

Rochester Institute of Technology

RIT Digital Institutional Repository

Theses

4-10-2018

Analysis of Grape Berry Epiphytic Microbiomes via QIIME

Isabelle O'Bryon
igo8214@rit.edu

Follow this and additional works at: <https://repository.rit.edu/theses>

Recommended Citation

O'Bryon, Isabelle, "Analysis of Grape Berry Epiphytic Microbiomes via QIIME" (2018). Thesis. Rochester Institute of Technology. Accessed from

This Thesis is brought to you for free and open access by the RIT Libraries. For more information, please contact repository@rit.edu.

Analysis of Grape Berry Epiphytic Microbiomes via QIIME

By Isabelle O'Bryon

A Thesis submitted in partial fulfillment of the requirements for the
degree of
Master of Science in Bioinformatics

Thomas H. Gosnell School of Life Sciences

College of Science

Rochester Institute of Technology

Rochester NY

April 10th, 2018

Thesis Committee Approval

Dr. Michael Osier, Thesis Advisor

Date

Dr. Lance Cadle-Davidson, Thesis Committee Member

Date

Dr. Jeffrey Lodge, Thesis Committee Member

Date

Acknowledgments

I'd like to give a huge thank you to Dr. Osier who helped me at every step and set me in the right direction whenever I didn't know where I was going. I must have sent you hundreds of emails and asked even more questions and you always had a helpful answer. Thank you for being a fantastic thesis advisor.

I would also like to thank my collaborator Megan Hall who collected and sequenced all of the samples, and reached out to me to join her project. A special thank you my thesis committee member Dr. Lance Cadle-Davidson for being a great teacher and a wonderful person to work with. I would also like to thank my thesis committee member Dr. Jeffery Lodge and Cornell's Dr. Wayne Wilcox for supporting me during this project.

Thank you so much to all of my friends in Bioinformatics and Biotechnology for listening to me talk about grapes and microbiomes constantly for the past year and a half. And a special thank you to Bethany Hartley and Gabe Fournier for being the most amazing and supportive people.

Table of Contents:

Abstract.....	1
Introduction.....	2
The Important Role of Grapes in the Economy.....	2
Current Problems the Wine Industry Faces.....	3
Pathogenic Fungi on Grapes.....	5
Issues with the Current Methods Used To Manage Grape Diseases.....	7
Developmental Ecology of the Grape Surface Microbiome & Grape Disease.....	9
Known Fungi Present on Grapes.....	10
Known Bacteria Present on Grapes.....	11
Potential Pitfalls and Complications when Analyzing Microbiomes.....	12
Methods.....	15
Data collection.....	15
Software.....	15
Preprocessing.....	16
Qiime Processing.....	17
OTU Filtering, Statistical Analysis, and Visualization.....	19
Results.....	20
Section 1: Data Filtering.....	20
Section 2: Unclassified Taxonomies.....	22
Section 3: Taxa Associated with Sour Rot Symptoms.....	23
Section 4: Drosophila spp. Microbiomes in the Vineyard.....	29

Section 5: Grape Epiphytic Microbiomes throughout Development.....	33
Section 6: Core Analysis across Sample Types.....	40
Discussion.....	45
Section 1: Data Filtering.....	45
Section 2: Unclassified Taxonomies.....	45
Section 3: Taxa Associated with Sour Rot Symptoms.....	46
Section 4: Microbiomes and Terroir.....	47
References.....	51

Table of Figures

Figure 1. 2000-2014 US Production of wine in tons per year	3
Figure 2. Comparison of grape disease phenotypes.....	4
Figure 3. Monetary effects of Powdery Mildew on grapes	7
Figure 4. Environmental effects of the top fungicides used	8
Figure 5. Variation between 16s RNA and WGS microbiome workflows.....	13
Figure 6. OTU assignment of de novo sequences.....	14
Figure 9. Per base quality scores without trimming	21
Figure 10. Per base quality scores using Qiime trimming	21
Figure 11. Per base quality scores using FastX trimming	21
Figure 7. Percentage of unidentified bacterial OTUs when identified to at least the kingdom level	22
Figure 8. Percentage of unidentified fungal OTUs when identified to at least the kingdom level	23
Figure 12. Comparison of fungi present in asymptomatic and sour rot symptomatic samples from New York in 2015.....	25
Figure 13. Comparison of fungi present in asymptomatic and sour rot symptomatic samples from Tasmania in 2016.....	25
Figure 14. Comparison of bacteria present in asymptomatic and sour rot symptomatic samples from New York in 2015	26
Figure 15. Percentages of Acetobacter in sour rot asymptomatic and symptomatic samples from New York in 2015.....	27

Figure 16. Comparison of bacteria present in asymptomatic and sour rot symptomatic samples from Tasmania in 2016	28
Figure 17. Heatmap comparing asymptomatic and symptomatic groups of NY 2015 and Tasmania 2016 at the genus level	29
Figure 18. Fungi found in <i>Drosophila</i> spp. from Tasmania 2016.....	30
Figure 19. Bacteria found in <i>Drosophila</i> spp. from Tasmania 2016.....	31
Figure 20. Fungi present on <i>Drosophila</i> spp. in NY 2014	32
Figure 21. Bacteria present on <i>Drosophila</i> spp. in NY 2014	33
Figure 22. Comparison of fungi present throughout the grape’s development in NY in 2015	35
Figure 23. Comparison of fungi present throughout the grape’s development in Tasmania in 2016	35
Figure 24. Comparison of bacteria present throughout the grape’s development in NY in 2014	37
Figure 25 Comparison of bacteria present throughout the grape’s development in NY in 2015 .	38
Figure 26. Comparison of bacteria present throughout the grape’s development in Tasmania in 2016.....	39
Figure 27. Developmental groups heatmap genus level	40
Figure 28. Heatmap of all Tasmanian groups	41
Figure 29. Heatmap of all New York 2015 groups.....	42
Figure 30. Heatmap of all New York 2014 groups.....	44

Table of Tables

Table 1. Number of samples in each group from Tasmania and NY for fungi and bacteria.....	24
Table 2. Sample sizes of <i>Drosophila</i> spp. in 2014 NY.....	32
Table 3. Number of samples in each group at the genus level for fungi	34
Table 4. Number of samples in each group at the genus level for bacteria	34

Abstract:

Research in the wine industry has previously characterized many aspects of the wine making process from soil fertility to the community of microbes in must and wine, but the epiphytic bacterial and fungal communities have not been studied throughout grape development. This project aims to investigate the epiphytic grape microbiome to better understand its role in disease and grape development, and the effects of terroir on the microbiome, by examining the epiphytic microbiome of developing and sour rot infected grapes in New York and Tasmania. Sour rot is characterized by a distinct vinegar smell that is caused by the combination of *Drosophila* spp., fermentative yeasts, the acetic acid producing bacteria *Acetobacter* and *Gluconobacter*. Total DNA was extracted from rinsate berry samples, and the bacterial 16S ribosomal RNA (rRNA) and fungal internal transcribed spacer (ITS) regions were amplified for Illumina sequencing. The tool Quantitative Insights into Microbial Ecology (Qiime) was used to quality filter the sequences and identify the operational taxonomic units (OTUs) present in each sample. The taxonomies of these OTUs were assigned using the Greengenes and Unite databases. Statistical analysis was done in Statistical Analysis of Metagenomic Profiles (STAMP), and hierarchical clustering and data visualizations were done in R. Sour rot infected grapes were found to have large communities of acetic acid producing bacteria in both terroirs, however the typical fermentative yeasts were not present in large quantities in New York. In general, when microbes known to cause sour rot were present on symptomatic grapes, they were also present on asymptomatic grapes, supporting an active role for *Drosophila* spp. in the disease complex, rather than simply acting as a vector. The core epiphytic microbiome remained relatively consistent throughout development but varied dramatically by terroir and year.

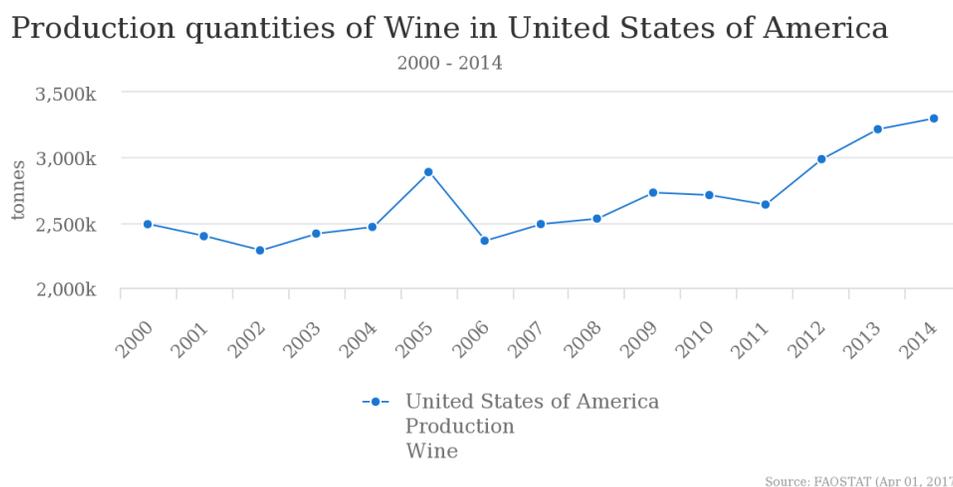
Introduction:

This project aims to investigate the epiphytic grape microbiome to better understand its role in disease and grape development and effects of terroir on the microbiome. Research in the wine industry had previously studied many aspects of the wine making process from the soil fertility to the community of microbes in the wine, but the bacterial and fungal community on the surface of the grapes as they develop has yet to be studied.

The Important Role of Grapes in the Economy

Grapes are an important part of the world's economy. The wine industry is prominent in 61 countries ("FAO Data - Dataset-Data-Filter - Crops Processed, National Production (FAOSTAT)"). In the United States alone wine sales totaled \$38 billion in 2015, and wine and grape production contributed \$162 billion to the United States' economy in 2007 (Gordon, 2016; "Economic Impact Study"; Insel, 2017). A study done in 2007 on the impact of wine in the US found that an estimated 1.1 million people in the US are employed in the wine industry and the wine industry is continuing to grow ("Economic Impact Study"). Since 2006 the US had seen a relatively consistent trend of the US producing more tons of wine each year, as seen in Figure 1 ("FAO Data - Dataset-Data-Filter - Crops Processed, National Production (FAOSTAT)"). With the wine industry playing such a significant role in the economy, research is needed to support the production of a consistent and viable crop of grapes.

Figure 1. 2000-2014 US Production of wine in tons per year (“FAO Data - Dataset-Data-Filter - Crops Processed, National Production (FAOSTAT)”).



Current Problems the Wine Industry Faces

Fungal pathogens are one of the major problems that grape growers face. Some of the more severe grape diseases are powdery mildew, Botrytis bunch rot, sour rot, and black rot (Fig. 2). These fungi have different environmental requirements and affect grapes via different mechanisms and therefore different treatments are needed to best manage each of them. For example, in California where low humidity and high heat is common, powdery mildew is able to thrive while other diseases such as sour rot are rarely an issue (Wunderlich et al., 2015). In New York, the higher humidity and moderate temperatures provide a suitable environment for many pathogens. Therefore vineyard managers cannot focus on managing a single pathogen but must consider the spectrum of diseases that can grow in their climate. Grapes are significantly more dependent on fungicide applications than other crops, with 95% of the grape crop yield being attributed to the use of fungicides compared to other crops which tend to range from 40% to 70% (Gianessi, 2006).

Figure 2. Comparison of grape disease phenotypes

a) *Black Rot*b) *Botrytis Bunch Rot*c) *Powdery Mildew*d) *Sour Rot*

a) ("Citrus Black Rot (*Phyllosticta Ampelicida*)", 2011)

b) (Taylor, 2017)

c) (Jones, 2014)

d) (Bordelon, 2016)

Pathogenic Fungi on Grapes

Each fungi can have different ideal climates, growing conditions, methods to infect the grape plant, means of propagating the infection, and ways to relocate to the next vine. This means that a treatment method that may be effective against powdery mildew may not affect Botrytis bunch rot.

Black Rot

A black rot infection is caused by the fungi *Guignardia bidwellii* and it typically begins infecting grapes when they are halfway through their development (Ries, 1999). Once a grape has been infected the fungi will harden the grape, turning it into what is known as a mummy that contains fungal spores (Fig. 2a). Fungicides and canopy management are used to reduce spread of black rot. Proper canopy management allows for air flow throughout the vineyard by manipulating how the vines are trimmed and grown. By keeping the vineyard more open it decreases the overall humidity, and thus makes the growing conditions less favorable for pathogenic bacteria and fungi. Canopy management and sanitation (removing diseased plant tissue) are particularly important to prevent a new crop of grapes from being infected by the previous year's grapes (Weigle, 2014).

Botrytis Bunch Rot

Botrytis bunch rot is caused by *Botrytis cinerea*. Infection can begin at flowering or when the grape is wounded by either a previous infection, pests, or an excess of water causing the grape to split open (Smith, 2016). The fungus breaks down 35%-45% of the sugars in the grape, increases the pH, and degrades the aromatic components on the grape causing the grapes to turn

brown (Fig. 2b) (Dharmadhikari, 2017). Overall, *B. cinerea* greatly changes the grape epiphytic environment and therefore may also change what other organisms are able to colonize the grape.

Powdery Mildew

Powdery mildew infections are caused by the fungus *Erysiphe necator* (Smith, 2016). This fungus infects young grapes and causes them to split open which then increases the grape's risk for other types of infections (Taylor, 2017). The grapes and leaves then become coated in the white, powdery fungi as seen in Figure 2c. At minimum, 3 to 15 applications of 50 to 125 gallons of fungicides are used per season to manage powdery mildew on grapes (Wunderlich, 2015).

Sour Rot

Unlike the other diseases there is no one specific fungus that causes sour rot which makes it more difficult to control. Sour rot is caused by a combination of the acetic acid producing bacteria *Acetobacter* and *Gluconobacter*, fermentative yeasts, and *Drosophila* spp. (Bordelon, 2016). It is characterized by the smell of vinegar from acetic acid, which makes the resulting wine acidic and unpalatable. For sour rot to start the grapes need to be wounded, yeasts then start fermenting the sugars to ethanol, which *Acetobacter* or *Gluconobacter* will then convert to acetic acid. Research has shown that *Drosophila* spp. lay their eggs in the grapes and create wounds which cannot heal quickly and therefore allow this process to start (Barata et al., 2012; Hall et al., 2018). After the grapes have been infected they become tan and soft and begin to disintegrate as seen in Figure 2d (Bordelon, 2016; Smith, 2014). The best way to manage sour rot is to prevent grapes from being damaged so as to prevent the infection from starting. This means that all aspects of the vineyard

from the fertilizer to irrigation to canopy management must be properly taken care of to avoid damaging the grapes and creating an environment for *Drosophila* spp. (Smith, 2014). Chemical management methods that use antimicrobials in conjunction with insecticides to control microbes and *Drosophila* spp. have shown to be effective in reducing sour rot.

Issues with the Current Methods Used To Manage Grape Diseases

The main expense that the wine industry faces is the cost of managing grape diseases. An estimated \$369 per acre is spent per growing season on managing just powdery mildew (Fig. 3) (Fuller). The cost of controlling fungal diseases includes chemical sprays, the means to distribute them, canopy management, and the labor to accomplish these jobs. In addition to the monetary costs there are also costs to the health of the workers. The chemical sprays that are used are toxic not only to pathogenic bacteria and fungi, but also to those who have to apply them onto the grapes. Those employed to maintain vineyards are at risk for developing allergic reactions, granulomatous fibrosin lung disease, skin and eye irritation, and asthma from being exposed to the pesticides and fungicides (Youakim, 2006).

Figure 3. Monetary effects of Powdery Mildew on grapes (Fuller)

Table 1. Powdery Mildew Costs				
	Annual PM Cost	Costs Attributed to PM as a Share of		
		Cultural Costs	Cash Costs	Total Costs
	<i>\$/acre</i>	-----percent-----		
Raisin Grapes				
Continuous Tray	222	8.7	4.6	3.4
Tray	222	12.4	6.9	4.5
DOV Open Gable	222	16.3	8.4	4.6
DOV Overhead Trellis	222	16.3	8.3	4.6
Wine Grapes				
Central Coast Chardonnay	369	19.6	12.4	7.7
Table Grapes				
Crimson Seedless	329	8.9	2.4	2.1

In addition to the health there are also significant environmental costs. Copper is one of the most used fungicides, being used 54% of the time for grape fungal infections (Wightwick, 2010). Figure 4 shows the severity of the ecotoxicological effects that the top four fungicides used in vineyards have on terrestrial and aquatic organisms (Wightwick, 2010).

Figure 4. Environmental effects of the top fungicides used (Wightwick, 2010)

	Copper-based	Dithiocarbamates	Triazoles	Strobilurins
Mobility in soil	Low	Medium to low	Medium to low	Low
Persistence in surface soils	High	Low	Moderate to high	Low to moderate
Field data on concentrations in soil	Good	None	Very limited	None
Field data on concentration in aquatic environments	Limited	None	Limited	Limited
Toxicity to earthworms	Moderate to low	Moderate to low	Moderate to low	Moderate to low
Toxicity to aquatic invertebrates	High	Low	Low	Moderate to low
Toxicity to fish	Moderate	Moderate to low	Low	Moderate to low
Toxicity to aquatic primary producers	Moderate	Low	Low	Low

Lastly, grape diseases negatively impact the wine produced with any diseased grapes. Wine made with 2%-3% of grapes being infected can taint the end product (Emmett, 2004). The resulting wine is poor quality and has a moldy and vinegary taste. Some of the fungi can survive fermentation and make the enzyme laccase which will degrade the pigmentation of the wine and spoil it (Steel, 2014).

Developmental Ecology of the Grape Surface Microbiome & Grape Disease

Microbiomes are the distinctive set of bacteria, fungi, and other microbes that have co-evolved to fill specific niches on plants, animals, or in the environment. Recently research done on grape microbiomes has primarily focused on the microbiome of the crushed mixture of the grapes that is fermented called the must. One study by Bokulich et al. (2013) investigated if the microbiome of the must differed due to the differences in terroir. Terroir is the term used among wine enthusiasts to describe the subtle effects on flavor and aroma from the environmental factors affecting wine grapes, such as the climate, soil fertility, elevation, and many other environmental factors. The Bokulich et al. study found that these factors influenced the microbiome of the must to some degree, however their work did not look into the original epiphytic grape communities. The present research aims to investigate this by examining the epiphytic grape microbiomes in New York and Tasmania which have very different terroirs. There is currently no way to measure terroir in quantitative terms so another goal of this project is to study the epiphytic microbes of grapes and the possible influence they may have on the terroir of the wine.

This study will also investigate if the healthy epiphytic grape microbiome changes as the grape develops. Currently, only the organisms present at harvest are known. The scientific community therefore does not know in what ways the epiphytic grape microbiome may be affected by the development of the grape. In this project, samples were taken throughout the growing season directly from the surface of the grapes and were sequenced and analyzed to investigate how the microbiome changes throughout the growing season. Samples were taken at the following developmental stages of grapes: pea sized berry, bunch closure, veraison, 15° Brix, and harvest. The first takes place when the grape is the size of a pea and the second is when the

grapes on the bunch have grown enough to close the gaps within the bunch. Veraison is defined as the point when the grapes begin change color, indicating that they have started to ripen. The Brix is the measurement used to assess the amount of sugar in the grape and therefore it is used to quantify ripeness. 15 ° Brix is slightly before harvest, when the grape are typically around 22° Brix.

This project also examines the effects of sour rot on the microbiome. By knowing if there are distinct changes in the microbiome that make the grapes more susceptible to sour rot then vineyard managers could test their grapes to detect when they would need to take preventative measures. This could prevent a widespread outbreak of the disease and therefore increase yield and prevent poor quality wine production. Ideally a test like this would replace the current methods used to determine if a plant is infected, which is typically done by “visual inspection [and] does not give an accurate estimate of levels of infection” (Emmett, 2004).

Known Fungi Present on Grapes

The scientific community has found that *Hanseniaspora*, *Metschnikowia*, *Candida*, and *Aureobasidium* are the most common yeasts found on the surface grapes (König et al., 2009). Previous work on the development of the epiphytic grape microbiome have found *Aureobasidium* to be one of the main genera on “immature, mature, and both damaged and undamaged grapes” (Prakitchaiwattana et al., 2004). This same study, done on grapes from New South Wales, Australia, found high abundancies of *Metschnikowia* and *Hanseniaspora* species on wounded grapes however they also state that their methods could not detect certain species even when they were known to be present in high quantities (Morgan, 2017). In addition to these benign fungi there are also many pathogenic ones that cause mildews and rots on the grape

berries. Some of these pathogens are “*Aspergillus* spp., *Botrytis cinerea*, *Cladosporium* spp., *Penicillium* spp. and *Rhizopus* spp.” (Hocking, 2007). There have been several fungi that have specifically been associated with sour rot which are: *Candida*, *Saccharomyces*, *Hanseniaspora*, *Pichia*, and *Zygosaccharomyces* (Barata et al., 2012). *Saccharomyces* is not common on healthy grapes but in rare cases has been found on damaged grapes (Barata et al., 2012).

Known Bacteria Present on Grapes

There have been fewer studies on the bacteria in the vineyard than on the fungi as most research projects tend to focus on the yeasts present to better understand how the organisms in the vineyard may be affecting fermentation. Recently however there have been more studies that are also examining the bacteria. The bacteria that are thought to be part of the natural grape microbiome include “*Enterobacter* spp., *Enterococcus* spp., *Bacillus* spp., *Burkholderia* spp., *Serratia* spp., *Staphylococcus* spp., among others, [that] have been isolated from grapes but do not have the ability to grow in wines” (Barata et al., 2012). A recent study by Mezzasalma et al. investigated the grape berry surface microbiomes at harvest in Italy and found that the grapes from the different locations had a “core composition characterized by Enterobacteriales, Pseudomonadales, Bacillales, and Rhodospirillales” (Mezzasalma et al., 2017). However, they also noted that the different areas each had additional specific microbial traits. Another study done in Italy in 2014 examined the bacterial communities on Merlot and Chardonnay grown using different agricultural practices and found high abundancies of *Burkholderia*, *Ralstonia*, *Staphylococcus*, and *Pseudomonas* by doing 16S rDNA pyrosequencing (Campisano et al., 2014).

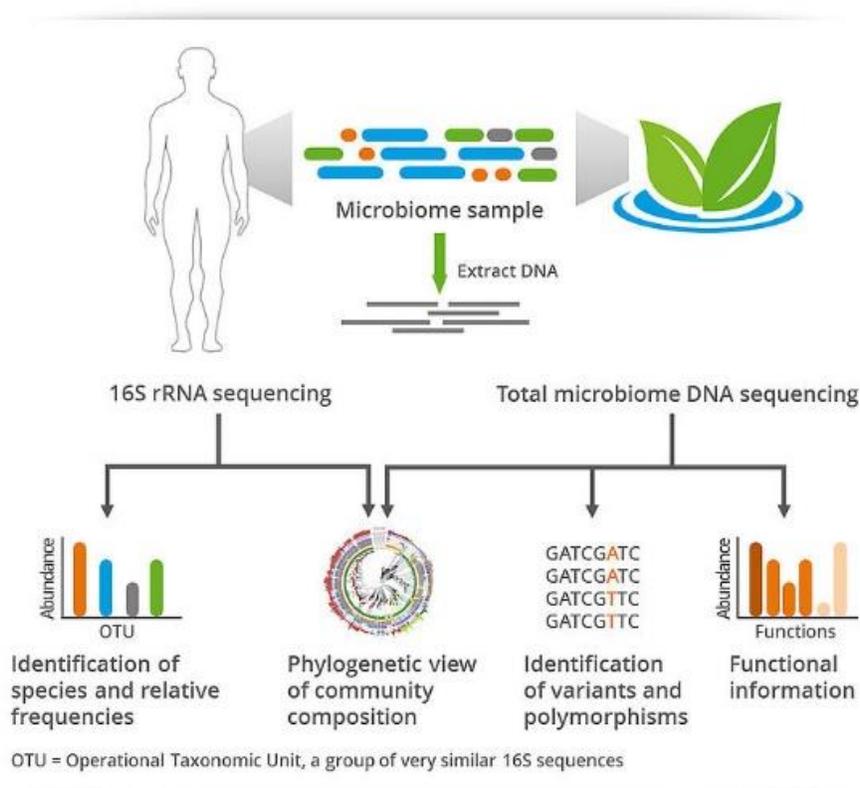
Potential Pitfalls and Complications when Analyzing Microbiomes

One of the first plant microbiomes to be studied was the microbiome surrounding roots, known as the rhizosphere, in 2010. The rhizosphere is a well-known and thoroughly studied area of plants that was first defined in 1904 and therefore was logically one of the first areas to be studied in the context of microbiomes (Burg, 2014). Just a few years later in 2012 the 28th New Phytologist Symposium: Functions and Ecology of the Plant Microbiome conference “brought together genetics/genomics, soil science, microbiology, computational biology, and plant and microbial physiology” to discuss the emerging field. This conference also started examining the “new challenges including analysis of huge amounts of data and the need to establish best practice standards for an emerging research field” (Lebeis, 2012). The speakers also acknowledged that this studying plant microbiomes would require a large interdisciplinary effort to correctly create the experiments and analyze the data in a consistent and replicable way.

The main goal when analyzing microbiome data is to determine what taxa are present and the abundances of those taxa. However, these data sets are large and complex and there are many ways that one analysis could differ from another. For example, when starting a microbiome experiment one of the first choices that needs to be made is what is going to be sequenced from the microbiome samples (Fig. 5). One option is to sequence a highly conserved region of the DNA that can be used to establish the taxa present. For bacteria the region used for this is the 16S rRNA gene and in fungi it is the internal transcribed spacer (ITS) region. Alternatively, all of the DNA present in the sample could be sequenced, which is called whole genome sequencing (WGS). While WGS is more expensive, however it gives a full picture of the genes and alleles are present in the microbiome. These two methods allow for different end results and therefore answer different questions, as seen in Figure 5. The WGS method returns the various genes and

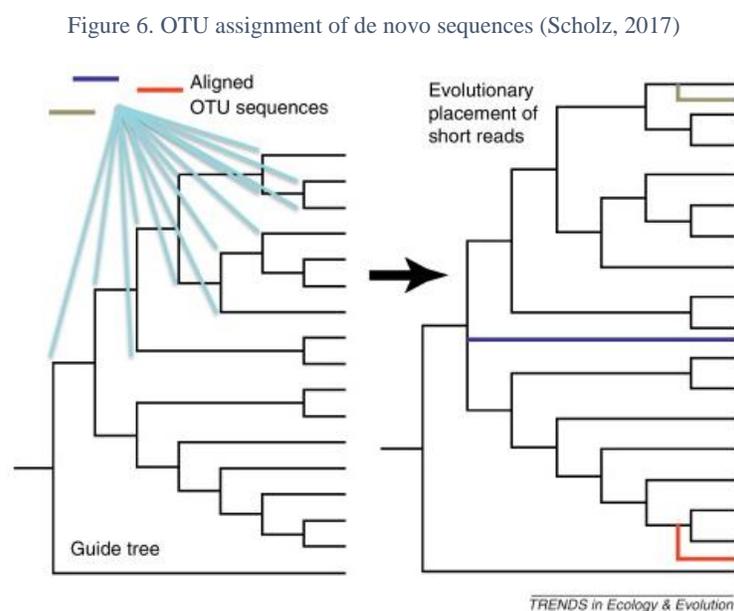
their functionalities, which is useful when investigating what the community as a whole is doing. The benefit of sequencing a highly conserved region is that it can be used to find the relative abundances of the taxa present. This is important to know when examining the composition of the community.

Figure 5. Variation between 16s RNA and WGS microbiome workflows (Microbiome Sequencing 16S rRNA Sequencing, 2017)



OTUs, operational taxonomic units, are identifiers that represent a group of similar sequences and are clustered by their similarity to other OTUs, in much the same way a species on a phylogenetic tree is clustered with other, closely related species. There are two methods that can be used to assign sequences to OTUs. One is the closed reference picking method, which

aligns the input sequences from the microbiome samples to those in the given database and groups them into OTUs if the sequences share a user-defined identity threshold. The standard threshold values are 95%, 97%, and 99% with 97% being used most commonly due to not being too lenient or too restrictive. In this method “[i]f the input sequence does not match any reference sequence at a user-defined percent identity threshold, that sequence is excluded” (Qiime). The taxonomic information for each OTU can then be obtained from the reference sequences. It is important to note that this is the only Qiime method that can be used for the 16S rRNA V2 and V4 regions because they are non-overlapping amplicons. This method works well for identifying known sequences but does not provide any taxonomic information for unknown, or *de novo*, sequences. The second method is open reference picking, also known as the *de novo* method, which assigns unmatched sequences to an OTU ID. This allows for some taxonomic information to be gained for these new sequences as it provides their relationship to known taxa (Fig. 6).



The database used as a reference also influences the end result. The different databases have different sequences in them which means that depending on what database is used there will be different reference sequences when building the OTUs. When analyzing 16S rRNA sequences in Qiime, either the Greengenes or SILVA database can be used. When analyzing the ITS region the UNITE database is the only given option (“Data Files and Other Resources”). It is important to note here that not all sequences will be able to be assigned to a full taxonomy and in those cases the taxa is assigned by using the closest taxonomic level that can be determined.

Methods:

Data collection

Data collection was done by Megan Hall in 2014, 2015, and 2016 using the methods described in Hall et al., manuscript in preparation. Briefly, 12 panels of grape vines were randomly selected in each vineyard from which to collect samples. Three grapes were collected from each cluster to make up one sample and the cluster was marked with flagging tape so they would not be selected at a later time point. For monitoring the epiphytic community during development, intact berries were rinsed with an extraction buffer, and the rinsate was used for DNA isolation. Primers were added to amplify the V2 and V4 16S rRNA and ITS regions. Samples were sequenced using an Illumina MiSeq.

Software

The raw compressed fastq files from Illumina were downloaded to the server for analysis. The following software were installed on this server to perform the analysis: Qiime 1.9.1,

BlastAll and its databases, FastX, FastQC, STAMP, and R v.3.3.2 (Caporaso, 2010; Andrews, 2010; FASTX-Toolkit; Babraham Bioinformatics; Parks, 2014; R Core Team (2013)). The software Qiime (Quantitative Insights into Microbial Ecology) is commonly used in the field of metagenomics as a robust tool for analyzing microbiomes (Gregory, 2010). The pipeline developed here used Qiime for the key components of analysis because of its capability to start with large datasets of raw sequences and produce an easy to work with file containing the number of sequences of each OTU in each sample and its taxonomy. Statistical analysis was done using STAMP (Statistical Analysis of Metagenomic and other Profiles) because easily uses Qiime's output file called a biom file along with a metadata file to group samples according to their metadata. STAMP then produces the relative mean frequencies (RMF) of taxonomies in these groups depending on the taxonomy level of interest. The RMF of a taxonomy in sample is calculated by dividing the number of sequences of that taxonomy by the number of all the sequences in that sample. In addition to STAMP, R was used to produce plots and heatmaps representing the populations of bacteria and fungi within the various experiments and treatments.

Preprocessing

In order to get the files ready to be processed, Qiime's `multiple_extract_barcodes` was executed on two folders, one containing all of the forward reads (R1) and the other containing all of the reverse reads (R2). The length of the barcodes was set to 17, location of the barcodes was specified to be in the label of the sequence, and “:0:” was set as the character delimiter. Fastq files of the barcodes were made using the `multiple_extract_barcodes` script in Qiime. Mapping files containing the metadata for each sample were created using the original Illumina demultiplexing file and were formatted according to Qiime's requirements. All file names were

changed to only contain alphanumeric and period characters, with the exception of an underscore before the specifier (ie `_barcode`, `_map`, `_R1`, `_R2`) to allow Qiime to correctly identify the specifiers.

Qiime Processing

Next, the fastq files were combined by executing `multiple_split_libraries_fastq` on a directory containing all R1 fastq files and their corresponding mapping and barcode files. The same was done for all R2 fastq files and their mapping and barcode files. The forward and reverse sequences could not be merged because the sequences did not have sufficient overlap. This was because the primer locations were greater than 150bp away. To still use both the forward and reverse sequences the pipeline was ran separately until the final biom files were made at which point they were joined together. `Multiple_split_libraries_fastq` was given the following parameters: mapping extension was set to `txt`, the demultiplexing method was `mapping_barcode_files`, and the read, barcode, and sample ID, and mapping indicators were `_R[1/2].fastq`, `_barcodes.fastq`, `'.'`, and `_map.txt` respectively. To determine the filtering parameters a comparison was done on a subset of the data between Qiime's filter method in `split_libraries_fastq` and FastX's `fastq_quality_filter`. The outputs from the trimming methods were then compared to the untrimmed file using FastQC. `Multiple_split_libraries_fastq` calls the function `split_libraries_fastq` multiple time and was given the following parameters: barcode length was 17, phred offset was 33, and phred quality threshold was 20.

Next, to assign the sequences to OTUs, the `seqs.fastq` file created from `multiple_split_libraries_fastq` was used as input for `pick_closed_reference_otus` and `pick_open_reference_otus`. Open and closed reference OTU-picking methods used `uclust` and

used a pairwise identity of 97% (Edgar, 2010). The bacterial database for 16S rRNA was Greengenes 13_5_97 as it has species level taxonomic identification, unlike the other database available Silva; the fungal database for ITS was UNITE 7_97 as it was the only option available in Qiime (DeSantis, 2006; Kõljalg, 2010).

Pick_closed_reference_otus was used to determine what bacterial taxonomies were present because pick_open_reference_otus could not be used. This was due to Qiime's specifications that states: "You must use closed-reference OTU picking if [y]ou are comparing non-overlapping amplicons, such as the V2 and the V4 regions of the 16S rRNA. Your reference sequences must span both of the regions being sequenced". The regions that had been sequenced for this study were the V2 and the V4 regions of the 16S rRNA. Pick_closed_reference_otus was executed with reverse strand match and assign taxonomy enabled to automatically assign taxonomies to OTUs after clustering. The Greengenes 13_5_97_otu_taxonomy.txt and 97_otus.fasta files were selected for alignment, which used the PyNAST alignment method and a 97% identity threshold (Caporaso, 2010).

For fungal taxonomies, pick_open_reference_otus was executed with the reference file path, the template file path, and reference sequence file path all set to the UNITE 97% file named sh_refs_qiime_ver7_97_28.06.2017.fasta. The ID to taxonomy file path was set to the UNITE file called sh_taxonomy_qiime_ver7_97_28.06.2017.txt. In addition to assigning OTU this function also does *de novo* OTU picking which creates new OTU IDs for sequences that do not belong to any pre-determined OTU. Reverse strand match and suppress lane mask filter were set to true and the assignment method was set to BLAST (Altschul, 1990). The biom files from the R1 and R2 reads containing the OTUs in each sample and their assigned taxa were then merged using Qiime's merge_otu_tables.

OTU Filtering, Statistical Analysis, and Visualization

Rare OTUs were defined as having less than 0.0001% of the total abundance from within that biom file and were removed from subsequent analyses. Biom files were converted into spf files using the biom_to_stamp script provided by STAMP. The mapping and spf files were read into STAMP and if an OTU was not identified to the genus level the lowest level of classification was used. If there was no classification the sequences were only used to calculate the total number of reads in the sample. ANOVA tests were done using the Tukey-Kramer method using a 95% CI, and a p-value filter of 0.05. The relative mean frequency (RMF) of each taxa in each sample was calculated. The RMFs of the taxa were plotted in R v.3.3.2 using ggplot2 with one standard error represented by errors bars (H. Wickham, 2009). The standard error was calculated by dividing the standard deviation by square root of the number of samples that contributed sequences to the group (*i.e.* asymptomatic and symptomatic).

Heatmaps were made in R v.3.3.2 using the pheatmap package (R Core Team, 2013; Kolde, 2012). The color of each cell represents the log of the RMF for each taxa. If a taxa was not present in a given group the value was assigned to the lowest value in the matrix. Hierarchical clustering was done using the complete method, the rows were clustered using the Euclidean method, and the columns were clustered using the Manhattan method. Euclidean was selected to cluster the taxonomies because of its ability to calculate distanced using continuous relative mean frequencies and weight strongly correlated variables. The Manhattan clustering method was chosen to cluster the groups because it accounts for multiple variables, in this case the many taxonomies, and reports the absolute distance between groups which was the desired result.

Results:

Section 1: Data Filtering

Two filtering methods were considered to improve the quality of the fastq files. Figure 9 shows the FastQC per base quality plots for one of the files before filtering, where the mean Phred scores are below 20 indicate trimming would improve the quality of the data. The same file was trimmed with Qiime (Fig. 10) and FastX (Fig. 11). The untrimmed and FastX-trimmed fastq files had large error bars frequently below a Phred score of 20 start at base pair 45 (Fig. 9 and 11) while the error bars in the Qiime-trimmed fastq file did not fall below a score of 20 until base pair 150 (Fig. 10). The Qiime trimming method was therefore selected as it provided a consistently better median quality score and smaller standard errors.

Figure 7. Per base quality scores without trimming

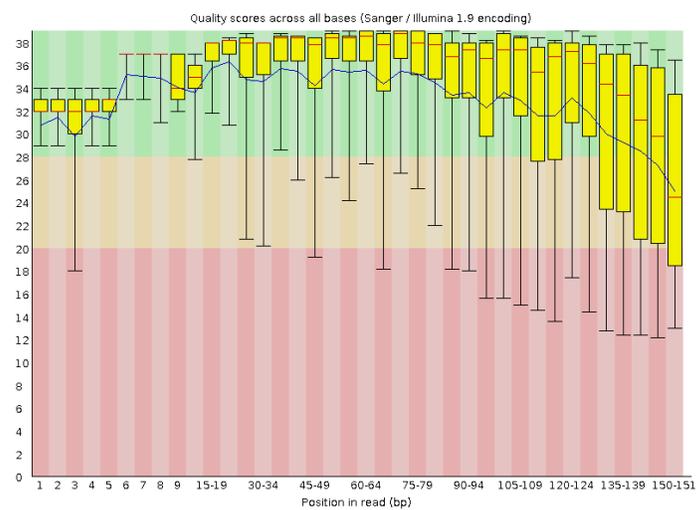


Figure 8. Per base quality scores using QIIME trimming

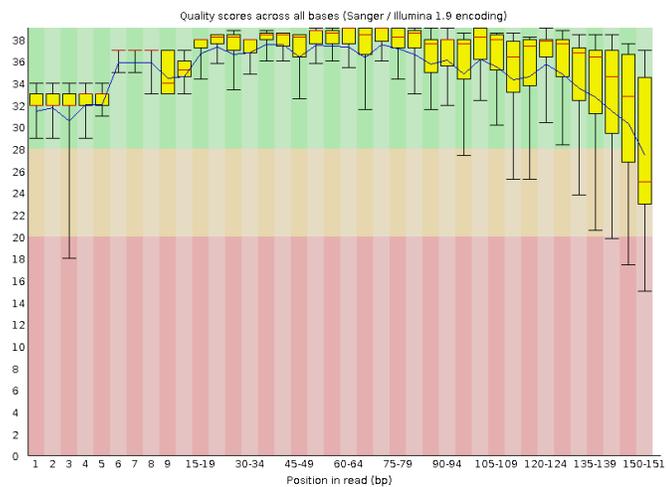
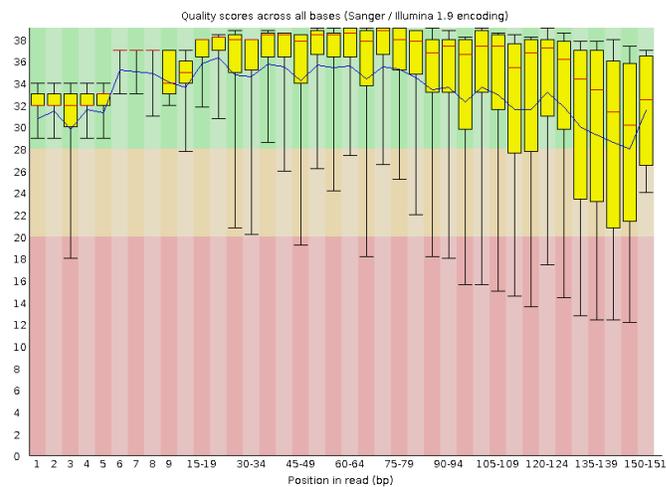


Figure 9. Per base quality scores using FastX trimming



Section 2: Unclassified Taxonomies

Of all OTUs that could be identified to at least the kingdom level, 99.25% were identified to the genus level. There was a higher percentage of unidentified orders in the bacteria (0.360%) compared to fungi (0.087%), but this difference was less distinct at the genus level (Figs. 7 and 8). The New York 2015 dataset overall all had more OTUs that were unidentified at the genus and order level compared to the other two data sets.

Figure 10. Percentage of unidentified bacterial OTUs when identified to at least the kingdom level

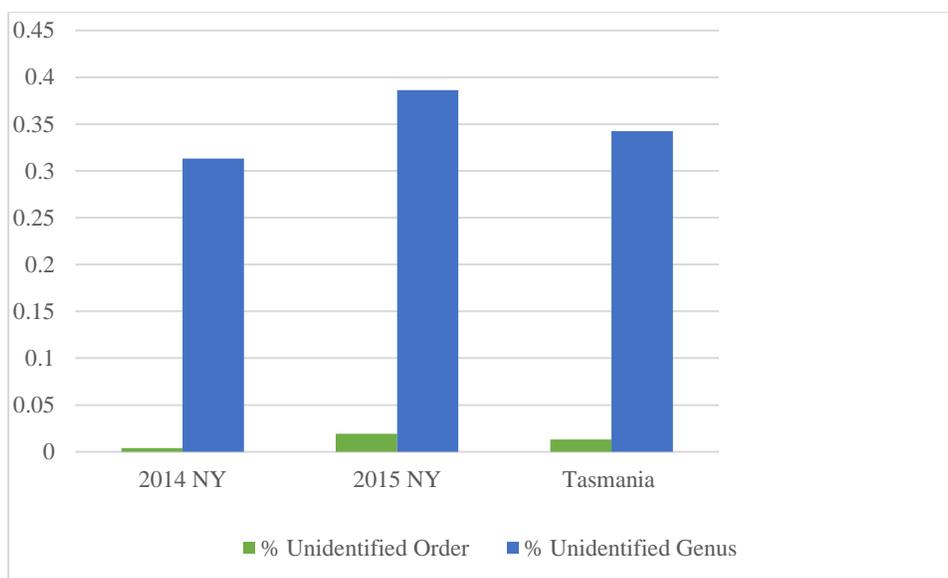
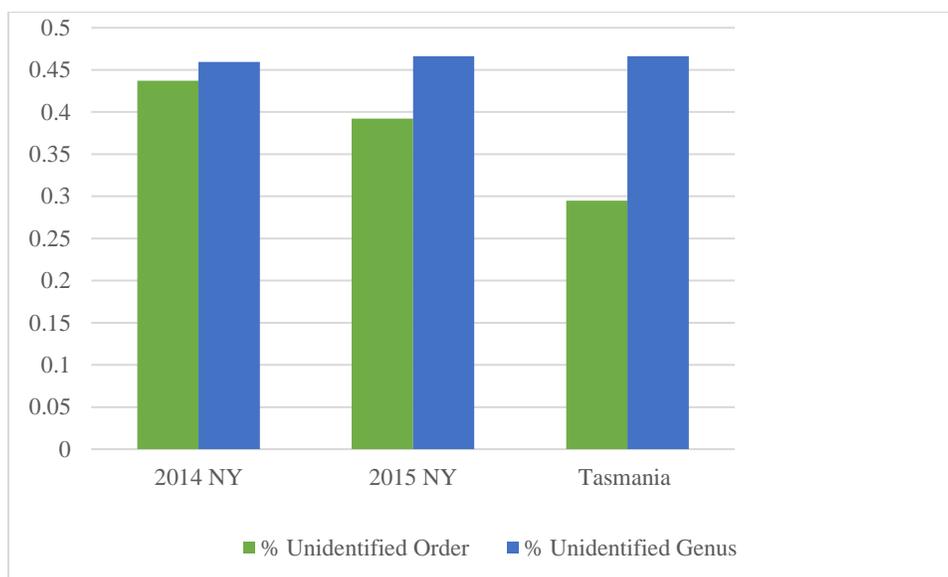


Figure 11. Percentage of unidentified fungal OTUs when identified to at least the kingdom level



Section 3: Taxa Associated with Sour Rot Symptoms

The average New York 2015 sample size was 22.5 (including fungi and bacteria) and the average Tasmania 2016 sample size was 39.5 (Table 1). Table 1 shows the number of samples with identified bacterial or fungal sequences in the asymptomatic and symptomatic groups after alignment and filtering. When interpreting the following bar plots, there are two reasons why each group may not add up to 100% RMF. First, of the sequences that remained after filtering and OTU assignment, those that did not have a taxonomic assignment are not shown but were used to calculate the RMF values. Second, only genera that had RMFs higher than a few percent and could be shown legibly on the plot were included. Any bacteria or fungi that were not able to be identified at to genus level were instead identified to the family or order level.

Table 1. Number of samples in each group from Tasmania and New York for fungi and bacteria

Fungi			Bacteria		
	Asymptomatic	Symptomatic		Asymptomatic	Symptomatic
NY 2015	22	29	NY 2015	18	21
Tas 2016	56	44	Tas 2016	24	34

In New York in 2015 and Tasmania in 2016 the most biologically relevant yeast genera to sour rot that were present were *Pichia*, *Candida*, and *Hanseniaspora* (Barata et al., 2012) (Figures 12 and 13). In addition to these the genera *Cladosporium* is also potentially interesting as this genus causes its own distinct rots. In New York 2015 *Talaromyces* had an increased RMF in the symptomatic group compared to the asymptomatic group (RMFs of 37.1+6.4% and 10.1+17% respectively) (Fig. 12).

In 2016 in Tasmania there was a higher RMFs of *Pichia* in the symptomatic group, with 15.5+19.4% in the asymptomatic group and 27.5+3.0% in the symptomatic (Fig. 13). *Cladosporium* was present only in the symptomatic group with a RMF of 2.54+2.2%. Both groups had similar abundancies of *Candida* and *Hanseniaspora* however the asymptomatic group again had larger standard errors (Fig. 13).

Figure 12. Comparison of fungi present in asymptomatic and sour rot symptomatic samples from New York in 2015

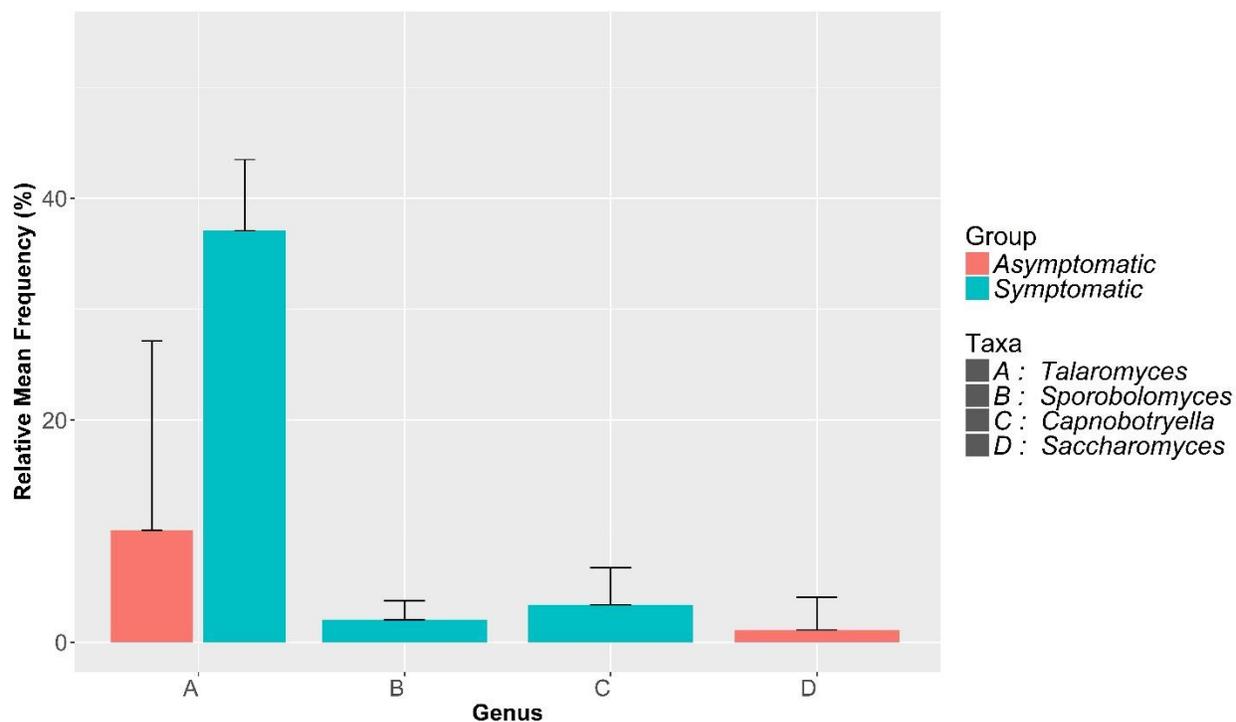
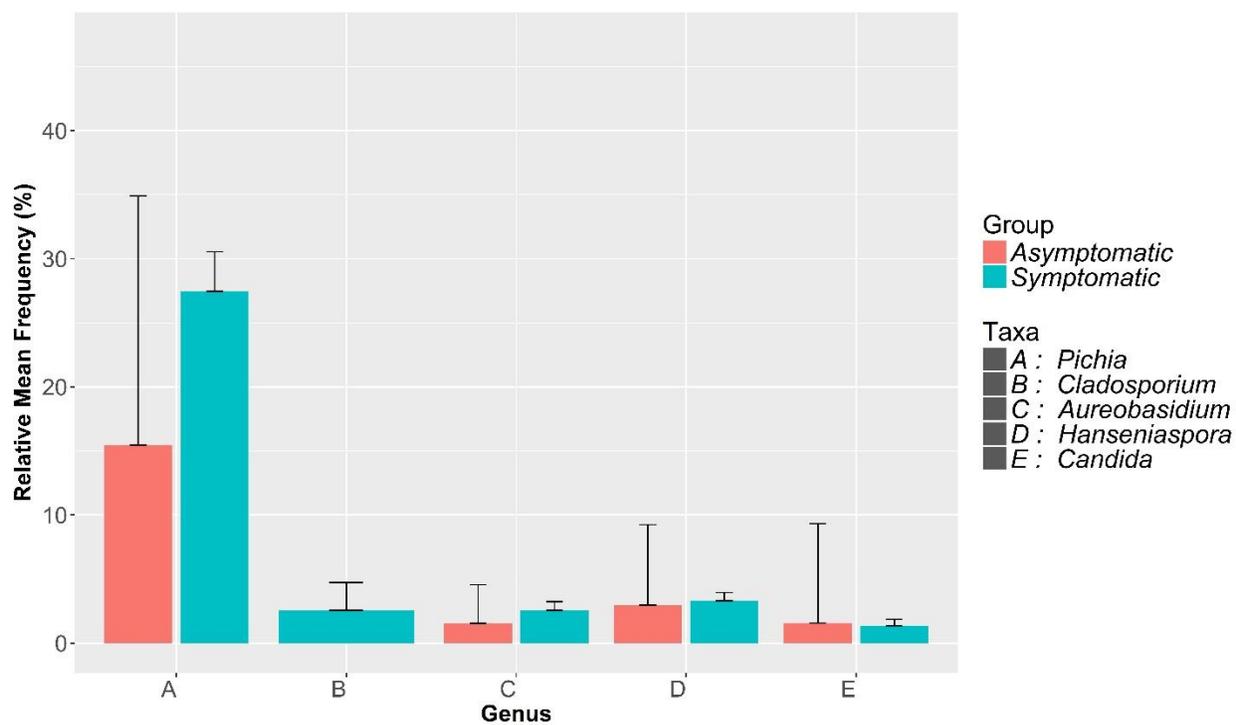


Figure 13. Comparison of fungi present in asymptomatic and sour rot symptomatic samples from Tasmania in 2016



The relevant bacteria to sour rot have been found to be *Acetobacter* and *Gluconobacter* as they are acetic acid producing bacteria. The 2015 New York data set had one of the most dramatic differences of bacteria between the two treatments, with nine species being present only in the asymptomatic or symptomatic samples (Fig. 14). *Acetobacter* was present in both treatments, but had a 25-fold greater RMF in symptomatic samples (Fig. 14). Further, 16 of the 21 (76.2%) symptomatic samples had *Acetobacter*, while only 3 of the 18 (16.7%) asymptomatic samples had *Acetobacter* present ($p=1.37 \times 10^{-5}$; Fig. 15). Note in red the *Acetobacter* positive controls both show 100% of their sequences being *Acetobacter*.

Figure 14. Comparison of bacteria present in asymptomatic and sour rot symptomatic samples from New York in 2015

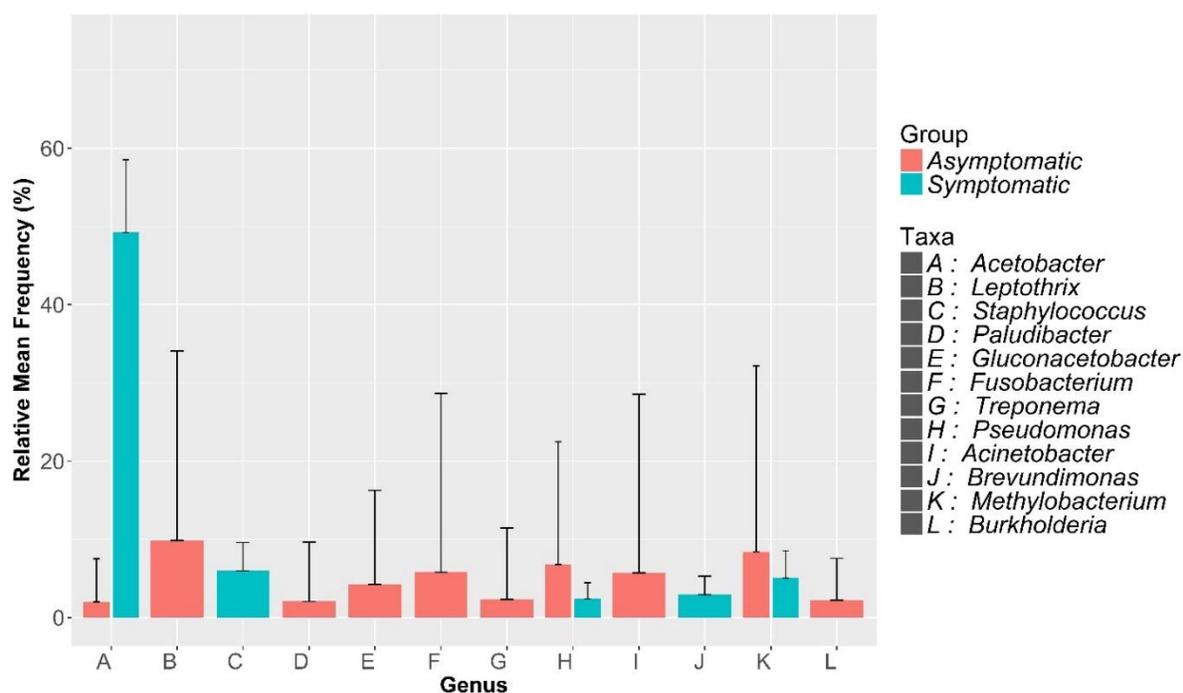
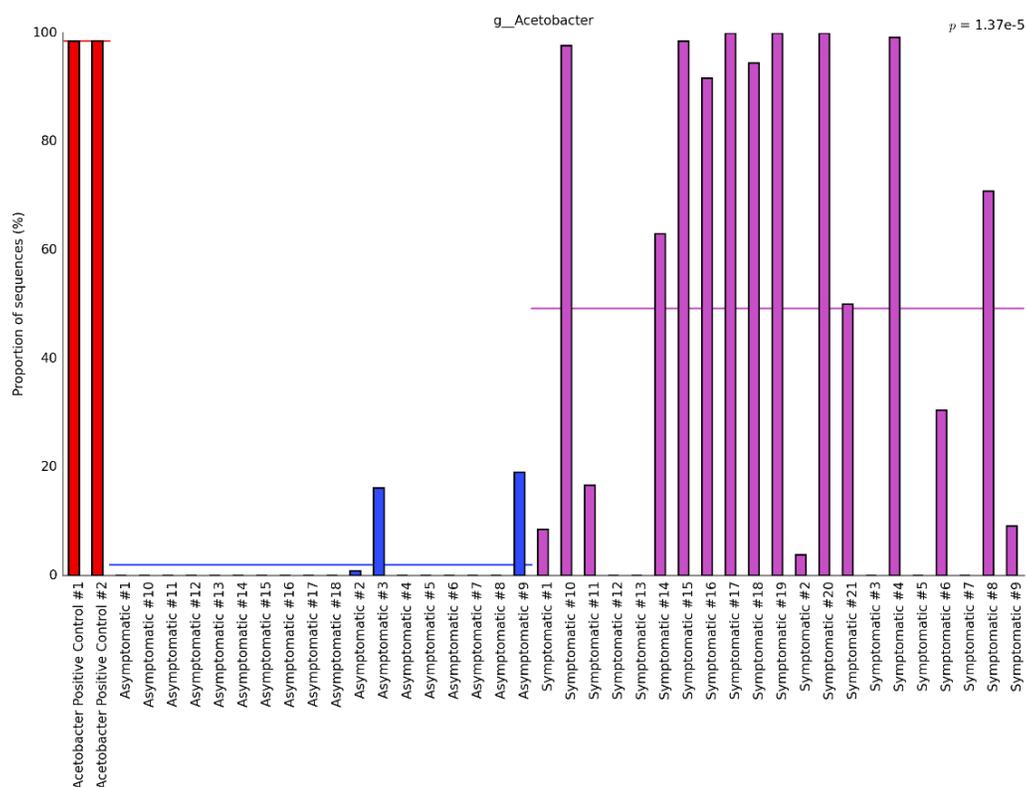
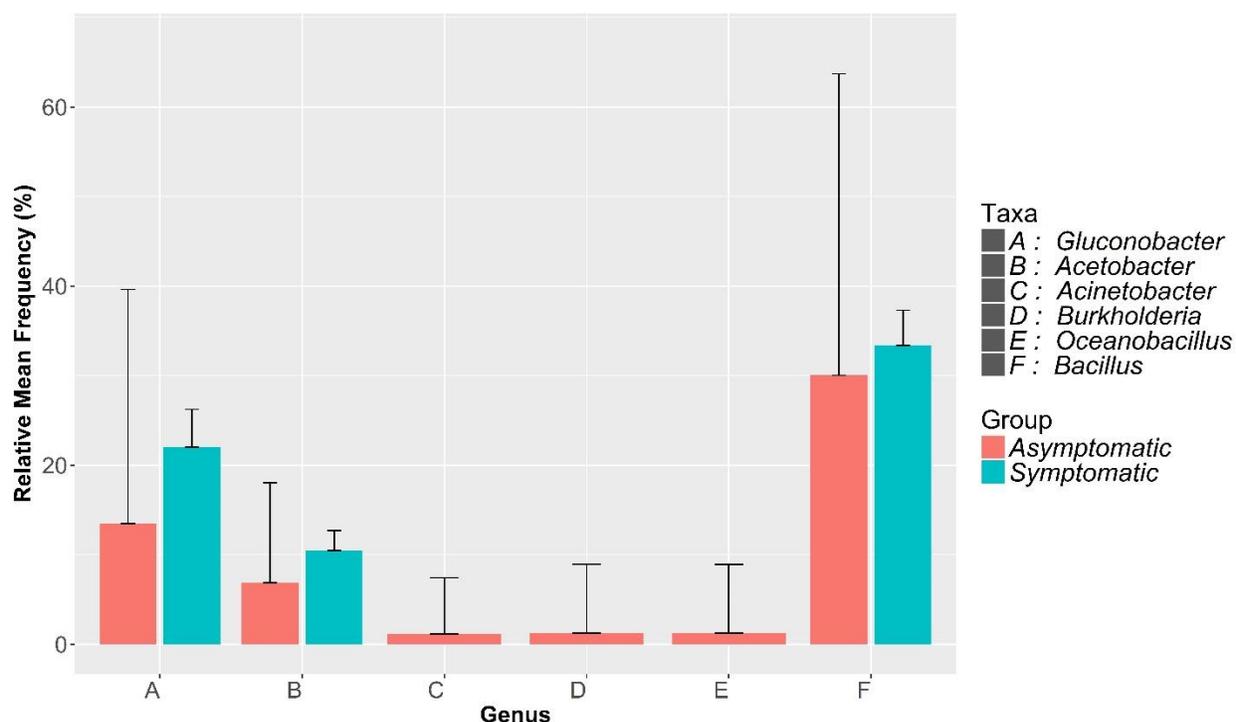


Figure 15. Percentages of *Acetobacter* in sour rot asymptomatic and symptomatic samples from New York in 2015

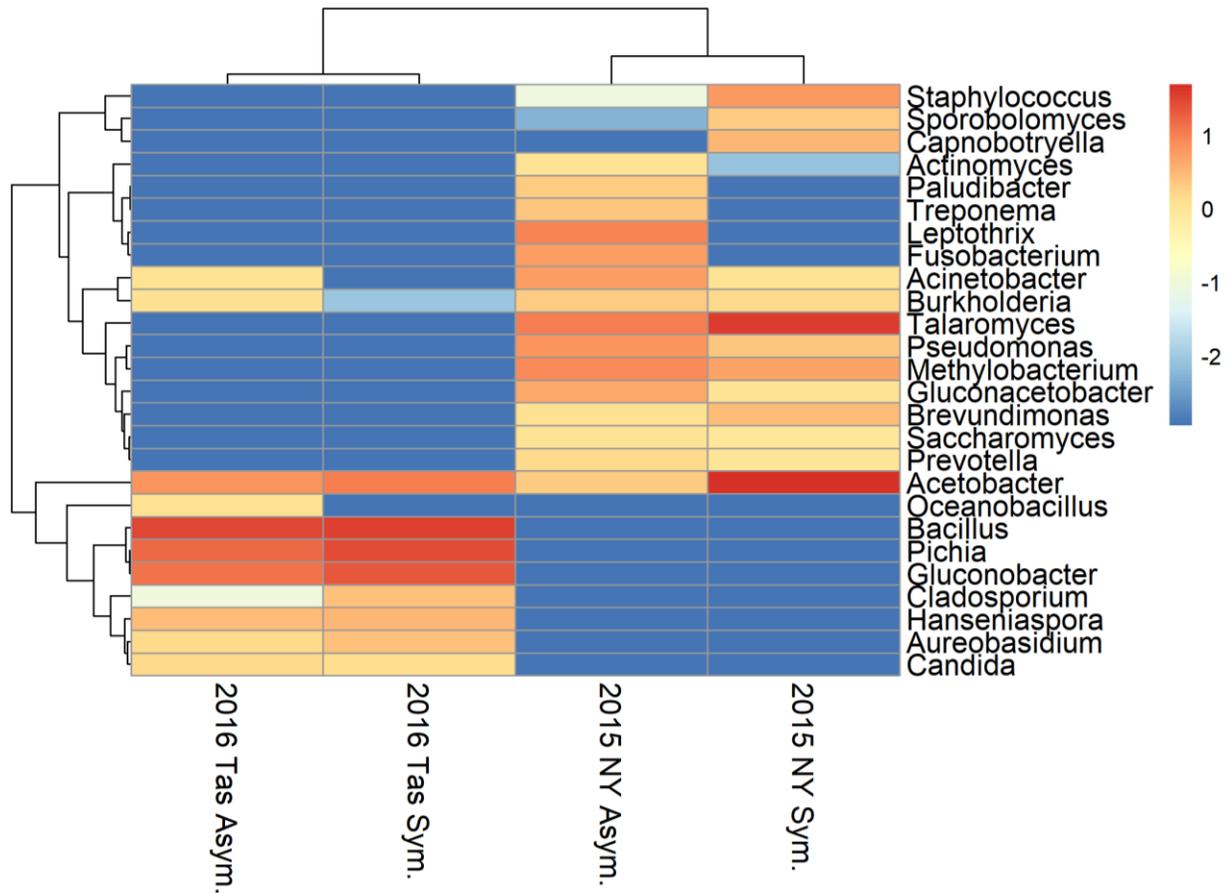
In the 2016 Tasmania data set both *Acetobacter* and *Gluconobacter* had slightly large RMFs in the symptomatic group compared to the asymptomatic group (Fig. 16), however those difference were non-significant ($p=0.215$ and 0.155 , respectively). *Gluconobacter* in the asymptomatic group had a RMF of $13.5+26.2\%$ and in the symptomatic group it was $22.04+4.1\%$. *Acetobacter* in asymptomatic was $6.9+11.1\%$ and in symptomatic it was $10.44+2.3\%$.

Figure 16. Comparison of bacteria present in asymptomatic and sour rot symptomatic samples from Tasmania in 2016



The samples from New York 2015 had vastly different microbiomes from those from Tasmania 2016 (Fig. 17). The heatmap shown in Figure 17 illustrates this, with the colors indicate the log of the RMFs of the bacterial and fungal in the asymptomatic and symptomatic groups. The groups clustered clearly by dataset, with asymptomatic and symptomatic samples from the same environment sharing many of the same taxa at similar abundances. The one taxa that was seen across all groups was *Acetobacter*, which had higher RMFs in the symptomatic groups compared to their respective asymptomatic groups. Other taxa that are associated with sour rot such as *Pichia*, *Gluconobacter*, *Hanseniaspora*, and *Candida* were clustered with or near to *Acetobacter*.

Figure 17. Heatmap comparing asymptomatic and symptomatic groups of NY 2015 and Tasmania 2016 at the genus level

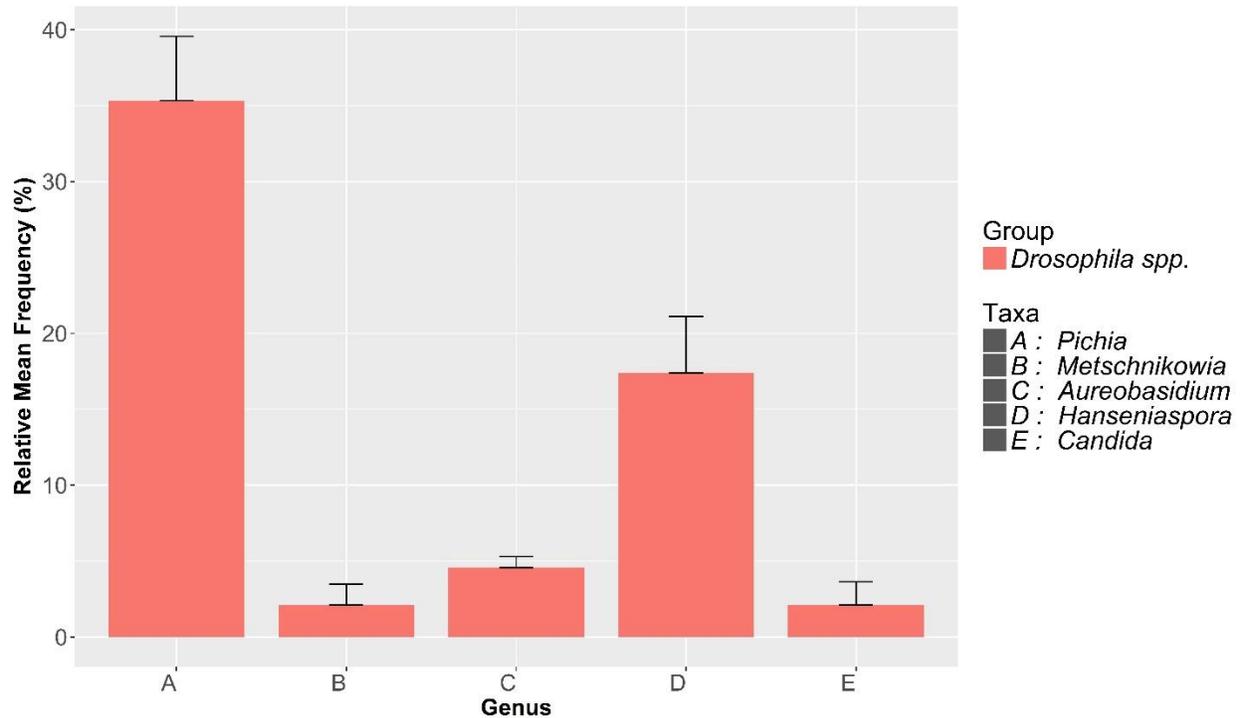


Section 4: *Drosophila* spp. Microbiomes in the Vineyard

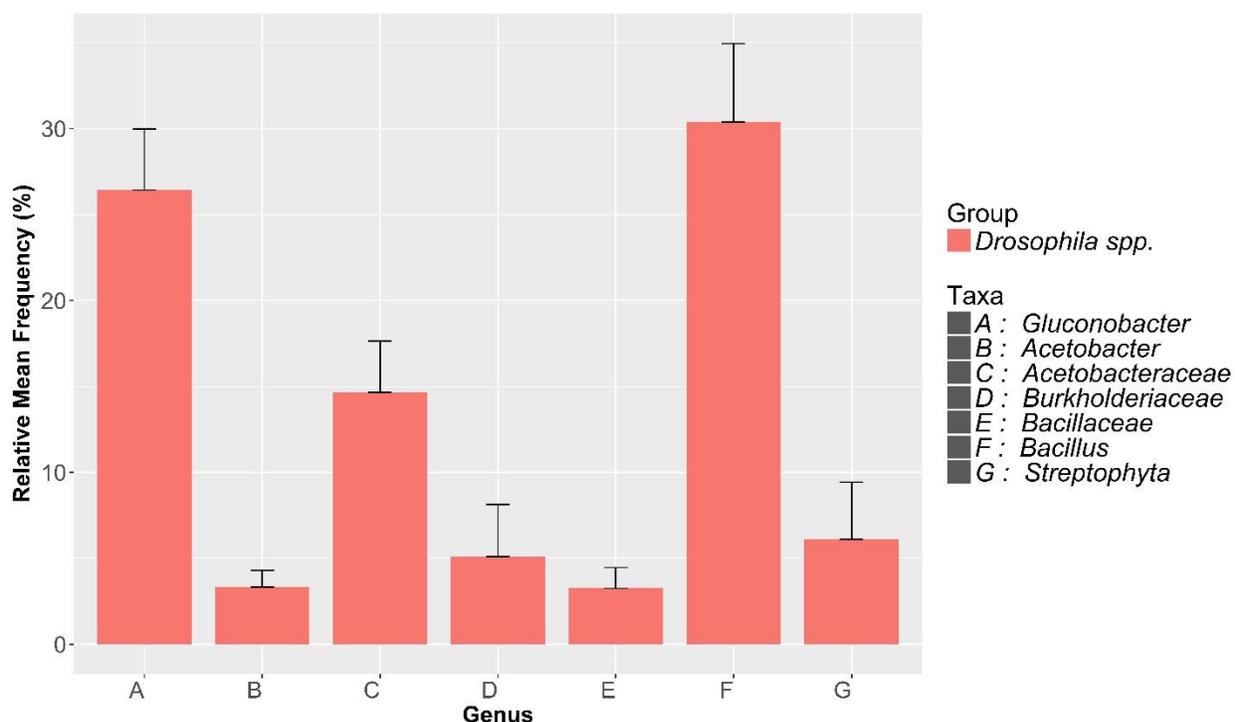
Bacterial and fungal DNA was amplified from 15 *Drosophila* spp. found in the vineyards sampled in Tasmania in 2016. The fungi present in these *Drosophila* spp. are commonly found on grapes in the later stages of development (Figure 18) (König et al., 2009). All of these fungi with the exception of *Metschnikowia* were also seen in the Tasmania symptomatic group with similar RMFs. *Pichia* contributed the most to this fungal community, having a RMF of 35.3±4.2% followed by *Hanseniaspora*, which had a RMF of 17.4± 3.7% (Figure 18). The

Pearson correlation between the *Drosophila* spp. microbiome and the other grape microbiomes sampled ranged from -0.05 to 0.33.

Figure 18. Fungi found in *Drosophila* spp. from Tasmania 2016



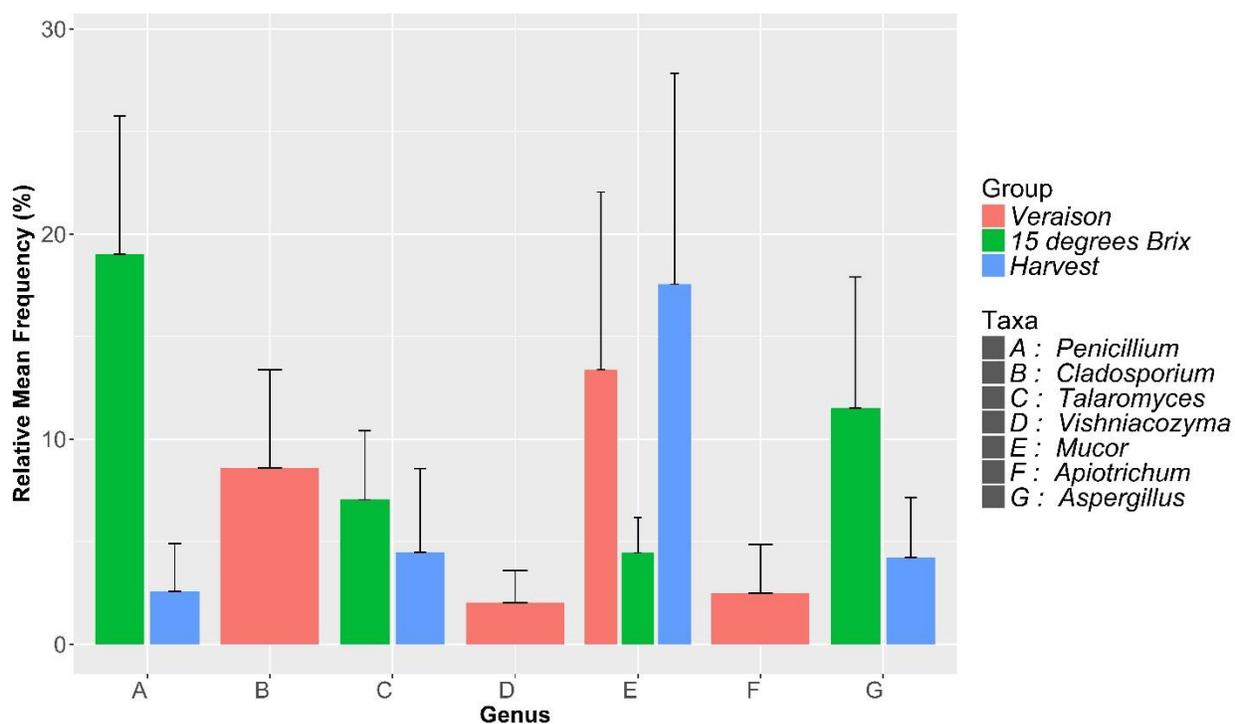
As with the fungi, that bacteria seen in the *Drosophila* spp. were mostly the same genera that were seen on the grapes in Tasmania (Figure 19). *Gluconobacter* and *Bacillus* had the highest RMFs on the *Drosophila* spp. with values of 26.4+3.6% and 30.4+4.6%, respectively. *Acetobacteraceae*, the family of acetic acid producing bacteria, had a RMF of 14.6+3.0%. Overall, the majority of the bacteria on the *Drosophila* spp. were acetic acid producing bacteria.

Figure 19. Bacteria found in *Drosophila* spp. from Tasmania 2016

Drosophila spp. were sampled at three time points in the later part of the growing season in New York in 2014. These time points were veraison, 15° Brix, and harvest and the sample sizes of bacteria and fungi at each time point is shown in Table 2. *Mucor* was present at veraison, 15° Brix, and harvest with RMFs of 4.5+1.7%, 13.4+8.7%, and 17.6+10.3% (Figure 20). *Mucor* causes a post-harvest rot on grapes called Mucor rot (König et al., 2009). The fungal community at 15° Brix was primarily composed of *Penicillium* (RMF of 19.0+6.7%) and *Aspergillus* (RMF of 11.5+6.4%). *Cladosporium*, *Vishniacozyma*, and *Apiotrichum* were only present in the veraison group. Overall, the 15° Brix and harvest communities were very similar to each other compared to veraison community.

Table 2. Sample sizes of *Drosophila* spp. in 2014 NY

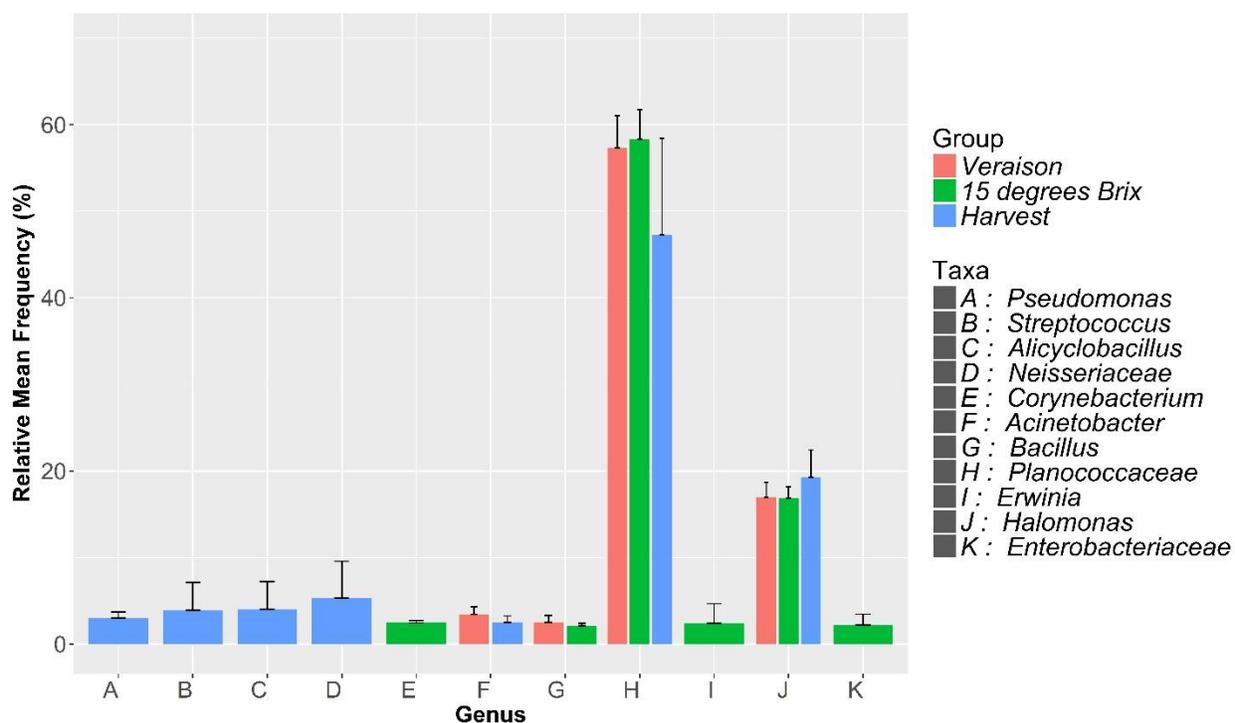
	Veraison	15° Brix	Harvest
NY 2014 fungi	10	16	6
NY 2014 bacteria	10	16	4

Figure 20. Fungi present on *Drosophila* spp. in NY 2014

The *Drosophila* spp. at all three time points in New York in 2014 had similar RMFs for the *Planococcaceae* family and *Halomonas* (Figure 21). Specifically, 15° Brix, veraison and harvest had RMFs of *Planococcaceae* of 58.3±3.4%, 57.3±3.7%, and 47.3± 11.1% and *Halomonas* RMFs of 16.8±1.3%, 17±1.7%, and 19.2±3.2%. In addition to this, the corresponding time points from the grapes in 2014 in New York, which are discussed in the next section, show the same abundancies of the *Planococcaceae* and *Halomonas*, around 50% and

10%-20%, respectively. The harvest community also had *Alicyclobacillus*, which a bacteria that is able to grow in acidic conditions.

Figure 21. Bacteria present on *Drosophila* spp. in NY 2014



Section 5: Grape Epiphytic Microbiomes throughout Development

Samples were taken at five time points of grape development in 2014 and 2015 in New York and 2016 in Tasmania. These time points in sequential order are: pea-sized berry, bunch closure, veraison, 15° Brix, and harvest. Note that there were not sufficient fungal samples after sequencing and filtering for a 2014 NY fungal data set (Table 3). The fungal sample sizes varied greatly, ranging from 10 to 100 (Table 3) and the same trend was seen in the bacterial samples (Table 4). There were no 2015 NY harvest samples that had sufficient bacterial or fungal DNA

to be included (Table 4). The Tasmania dataset overall had the largest sample sizes for both bacteria and fungi.

Table 3. Number of samples in each group at the genus level for fungi

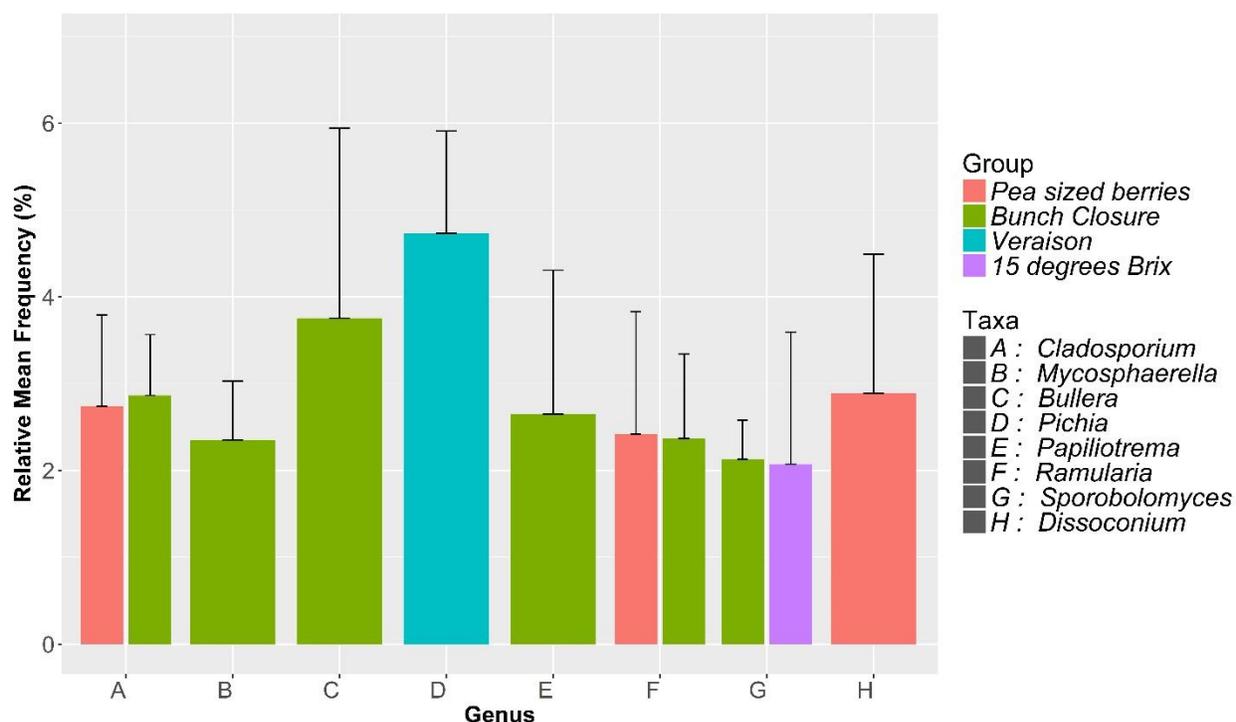
	Pea sized Berry	Bunch Closure	Veraison	15° Brix	Harvest
NY 2015	10	12	39	14	0
Tas 2016	40	40	86	100	40

Table 4. Number of samples in each group at the genus level for bacteria

	Pea sized Berry	Bunch Closure	Veraison	15° Brix	Harvest
NY 2014	4	7	13	7	4
NY 2015	10	12	39	14	0
Tas 2016	30	9	41	87	11

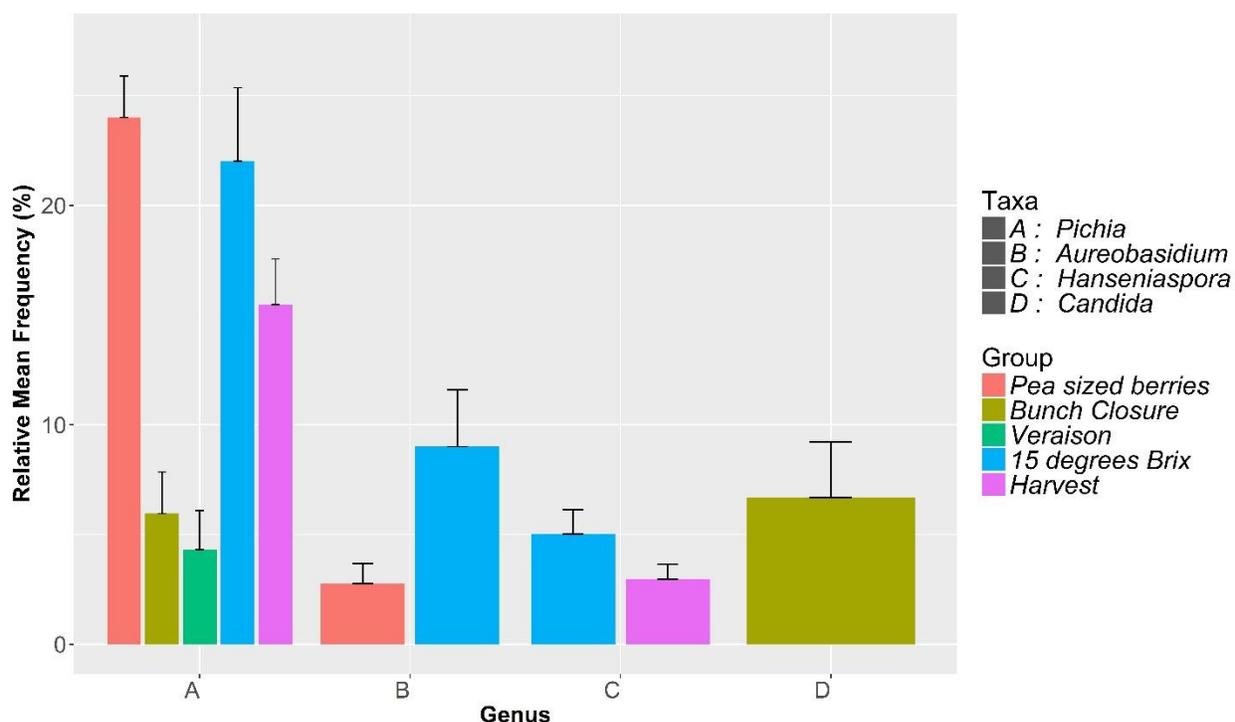
In the 2015 New York data set, only three taxa were identified in multiple developmental stages (Figure 22). Nearly all of the 39 veraison samples had either frequencies of approximately 8% for *Pichia* with a few having 0% (data not shown), resulting in a RMF of 4.7+1.2%. The other fungi seen in Figure 22 had on average RMFs of 2.4% and were either yeasts (*Bullera* and *Sporobolomyces*) or plant pathogens (*Cladosporium*, *Ramularia*, and *Dissoconium*).

Figure 22. Comparison of fungi present throughout the grape's development in New York in 2015



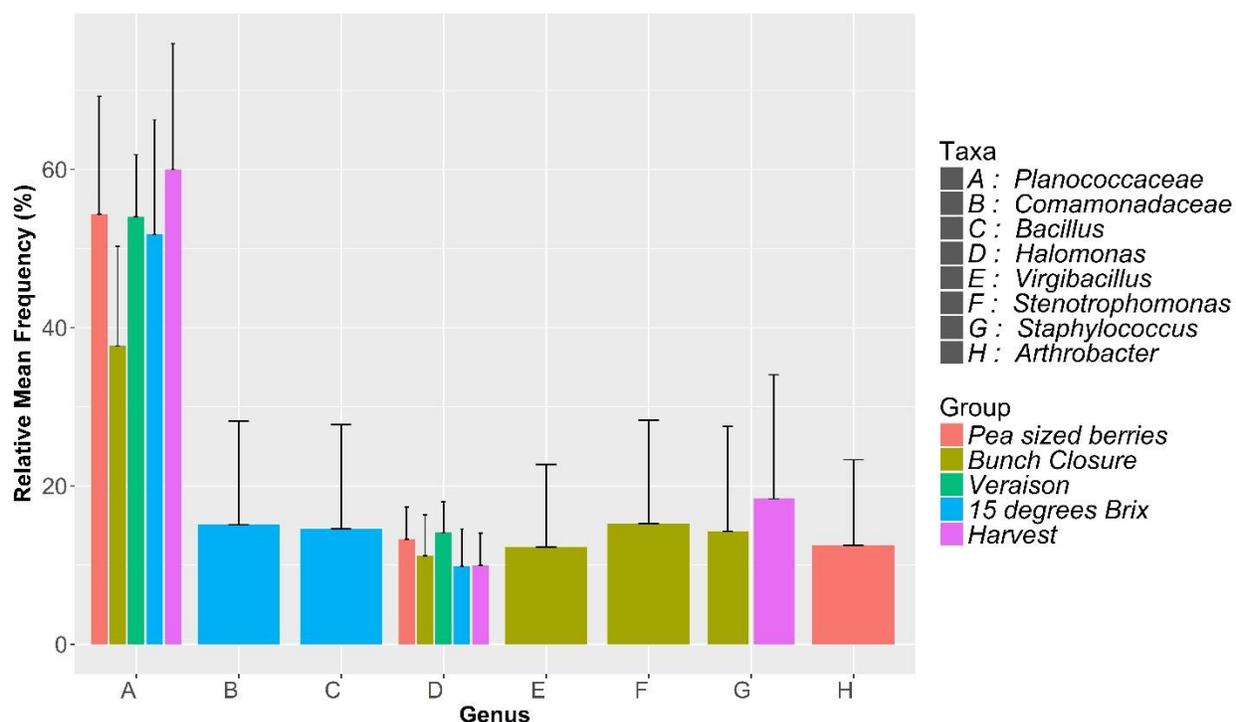
The Tasmania fungal data set had an average sample size of 61.2 (Table 3). *Pichia* was present in all groups and was a major contributor to the pea sized berry, 15° Brix, and harvest time points (24+1.8%, 22+3.4%, and 15.5+2.1% respectively) (Fig. 23). The pea-sized berry and 15° Brix time points also shared *Aureobasidium*, which as previously mentioned is likely to be a common contributor to the epiphytic grape microbiome as are the other fungi that were present, *Hanseniaspora* and *Candida*.

Figure 23. Comparison of fungi present throughout the grape's development in Tasmania in 2016



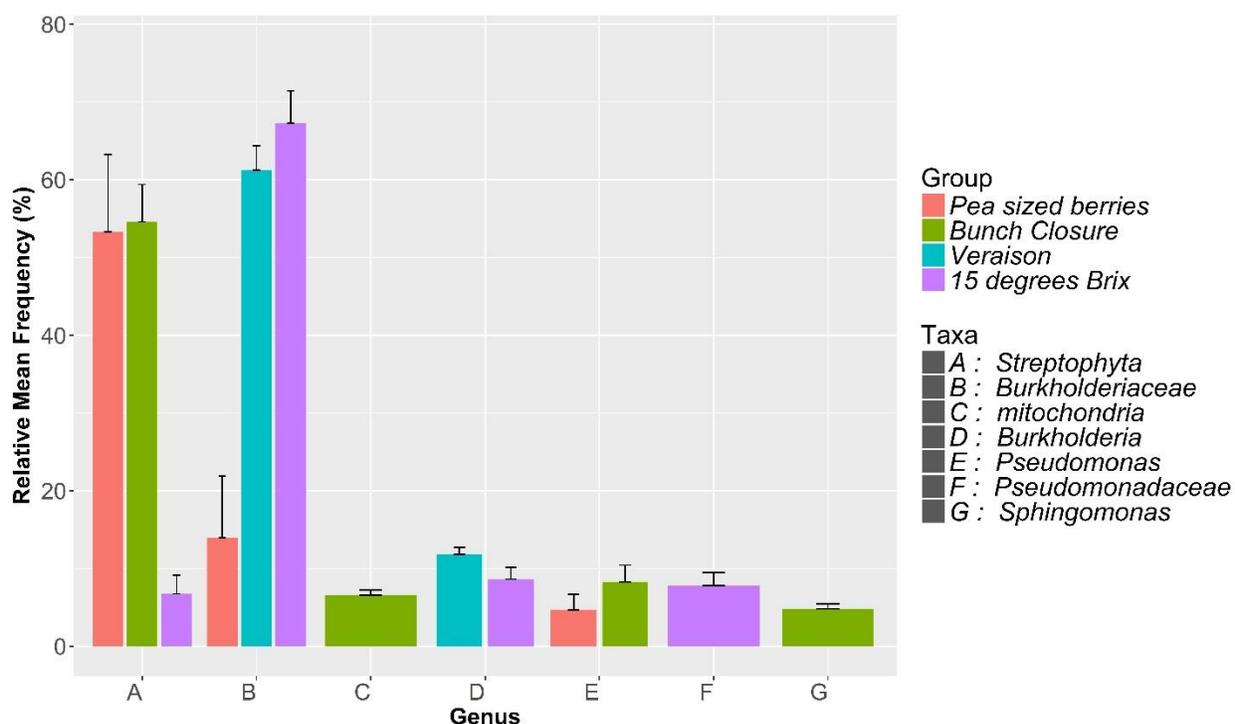
The sample sizes of the groups in New York 2014 for bacteria are relatively low, with the average sample size being 7 (Table 4). Across time points the family *Planococcaeae* had an average RMF of 51.6+13.2% and *Halomonas* had an average RMF of 14.5+4.4% (Figure 24). The former is in the Bacillales order and the latter is a halophilic bacteria that has previously been found on grape berries (Morgan, 2017). Other bacteria present were from orders and families that are also commonly found on grape such as Bacillales (*Bacillus*, *Virgibacillus*, and *Staphylococcus*), Burkholderiales (*Comamonadaceae*), and *Enterobacteriaceae*.

Figure 24. Comparison of bacteria present throughout the grape's development in New York in 2014



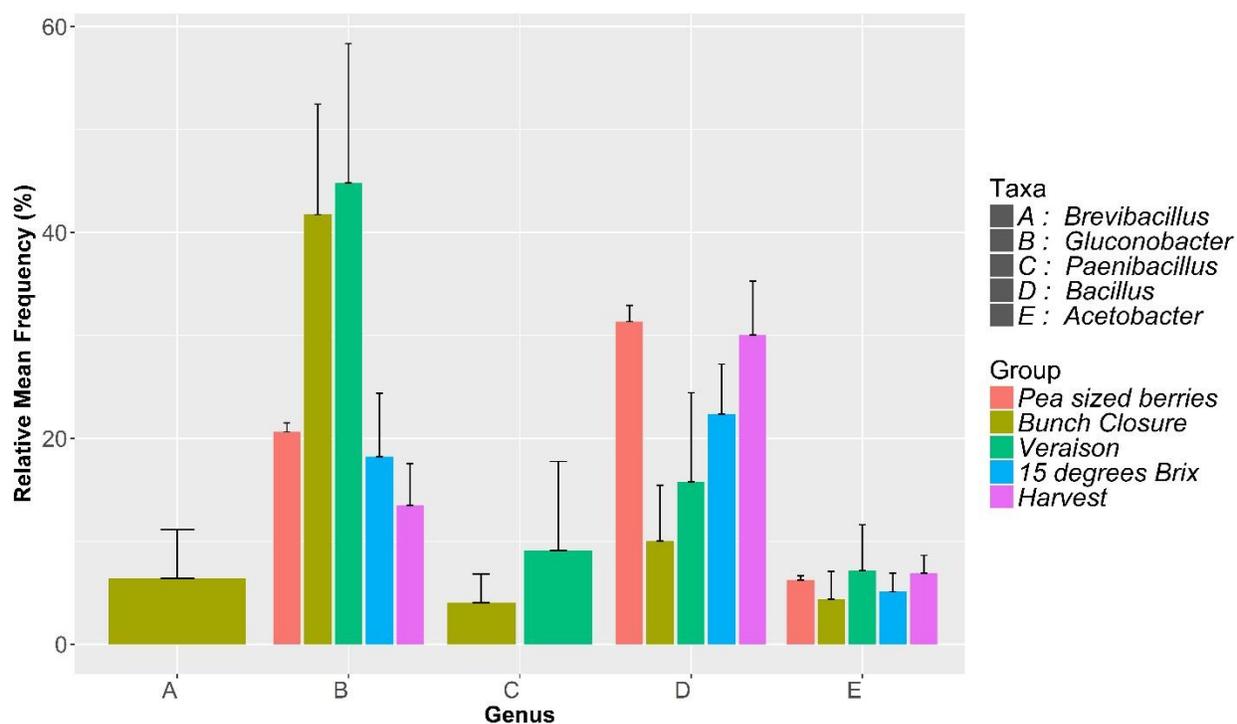
The family *Burkholderiaceae* was a major contributor to the veraison and 15 degree Brix communities in New York 2015, with each having 61.2+3.1% and 67.3+ 4.1% RMFs respectively (Fig. 25). The pea sized berry and bunch closure communities had RMFs of *Streptophyta* as 53.3+9.9% and 54.6+4.8%. The family *Pseudomonadaceae* was also present, along with the *Acetobacteraceae* family and *Sphingomonas*, which is found on roots and plants.

Figure 25 Comparison of bacteria present throughout the grape's development in New York in 2015



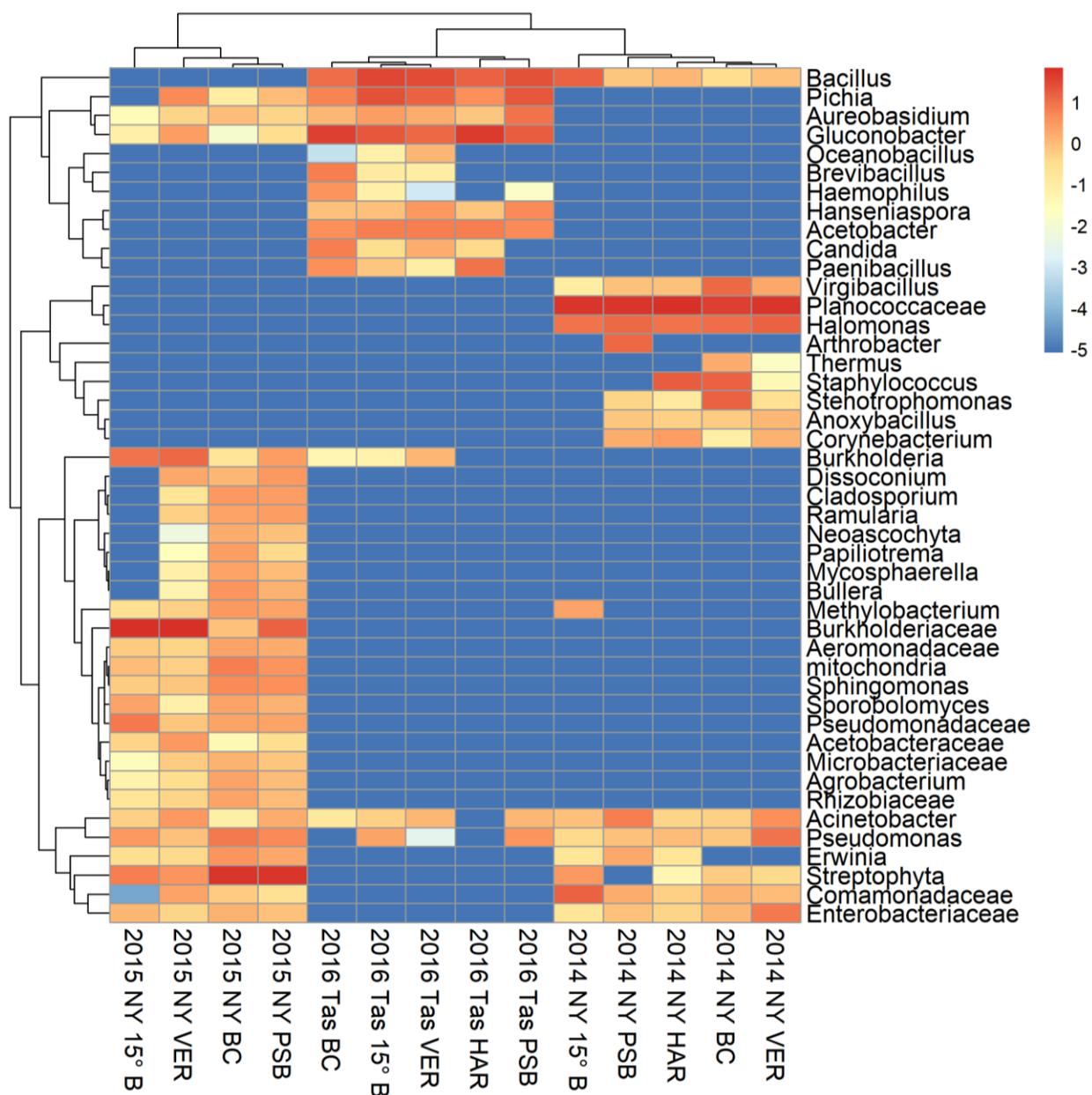
In the Tasmania 2016 data set *Gluconobacter*, *Bacillus*, and *Acetobacter* were seen in every time point and together composed the majority of the bacterial communities (Figure 26). The pea sized berry group had a RMF of 20.6+ 0.9% for *Gluconobacter* and 6.3+0.4% for *Acetobacter*. Samples in the other time points that had *Gluconobacter* tended to either have 100% of their RMF composed *Gluconobacter* or 0%. The pea sized berries, 15° Brix, and harvest time points had similar frequencies of *Gluconobacter*, and had an average RMF of 17.4+3.7%. The average RMF for the bunch closure and veraison groups was 43.5+12.1%. All of the time points had similar RMFs for *Acetobacter* and the average of their RMFs was 4.9+2.3%. *Bacillus* was found in 76.7% of all samples and had an average RMF of 21.8+5.2%.

Figure 26. Comparison of bacteria present throughout the grape's development in Tasmania in 2016



Each location and year had its own distinct set of bacteria and fungi regardless of developmental time points, as shown in Figure 27, and clustering of the groups was clearly determined by the data set (environment). *Acinetobacter* was the most common bacteria, followed by *Pseudomonas*, *Bacillales*, and *Gluconobacter*. Other than *Pichia* there were few fungi that were present in more than one data set. Each data set had as set bacteria and fungi that were seen in large abundancies throughout the growing season.

Figure 27. Developmental groups heatmap genus level



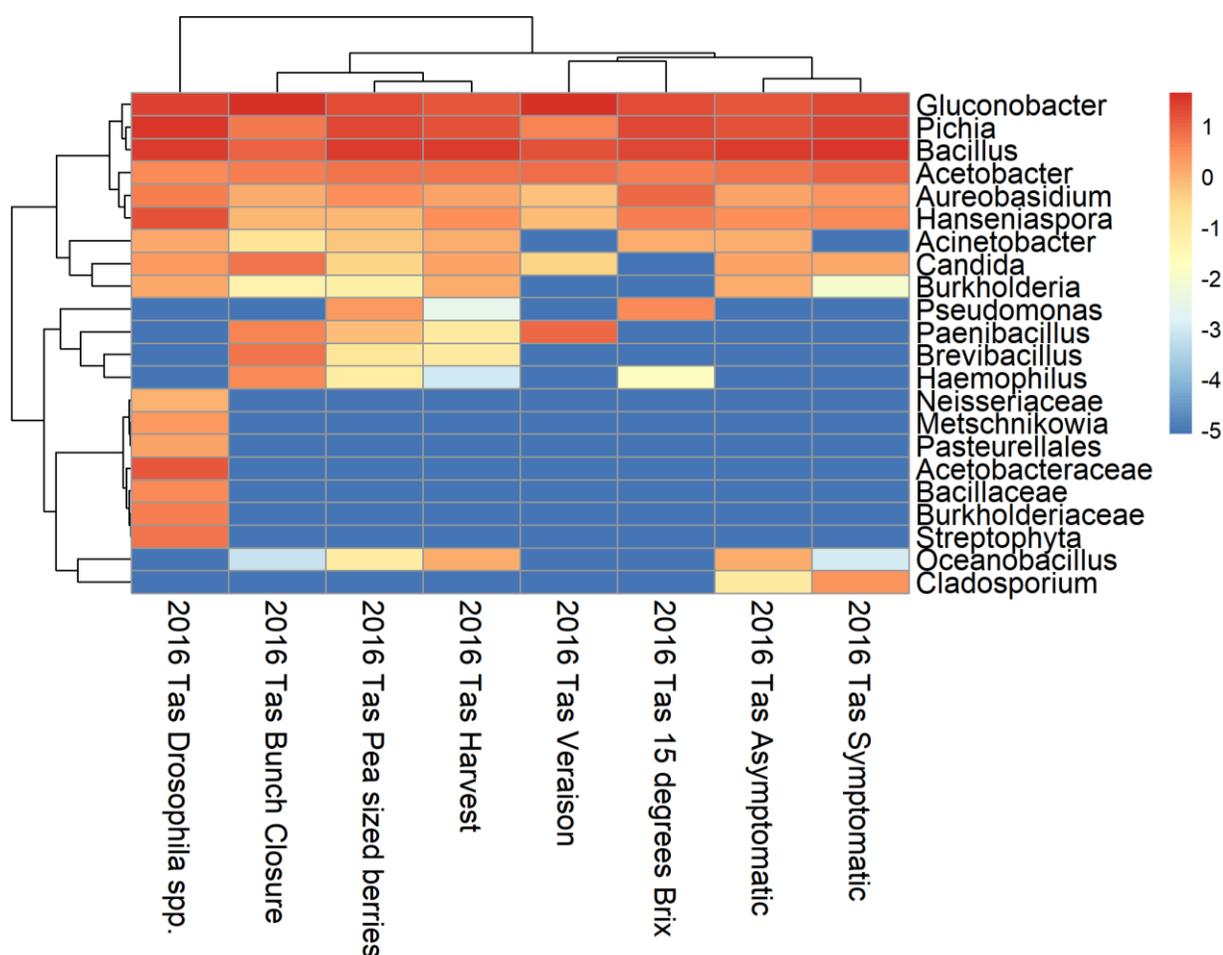
Section 6: Core Analysis across Sample Types

Above, data have been presented for three different sample types: 1) epiphytic washes from symptomatic versus asymptomatic whole berry samples; 2) epiphytic washes of whole berry samples over the course of development; and 3) macerated *Drosophila* fruit flies. In this

section, taxa frequencies were compared across these sample types within each environment to identify taxa that were shared and unique to each sample type.

The Tasmanian *Drosophila* spp., developmental stages, and sour rot asymptomatic and symptomatic grapes samples all shared a tightly clustered ‘core set’ of yeast and bacteria (Figure 28). *Drosophila* spp. had a set of bacteria and fungi that were unique to them.

Figure 28. Heatmap of all Tasmanian groups

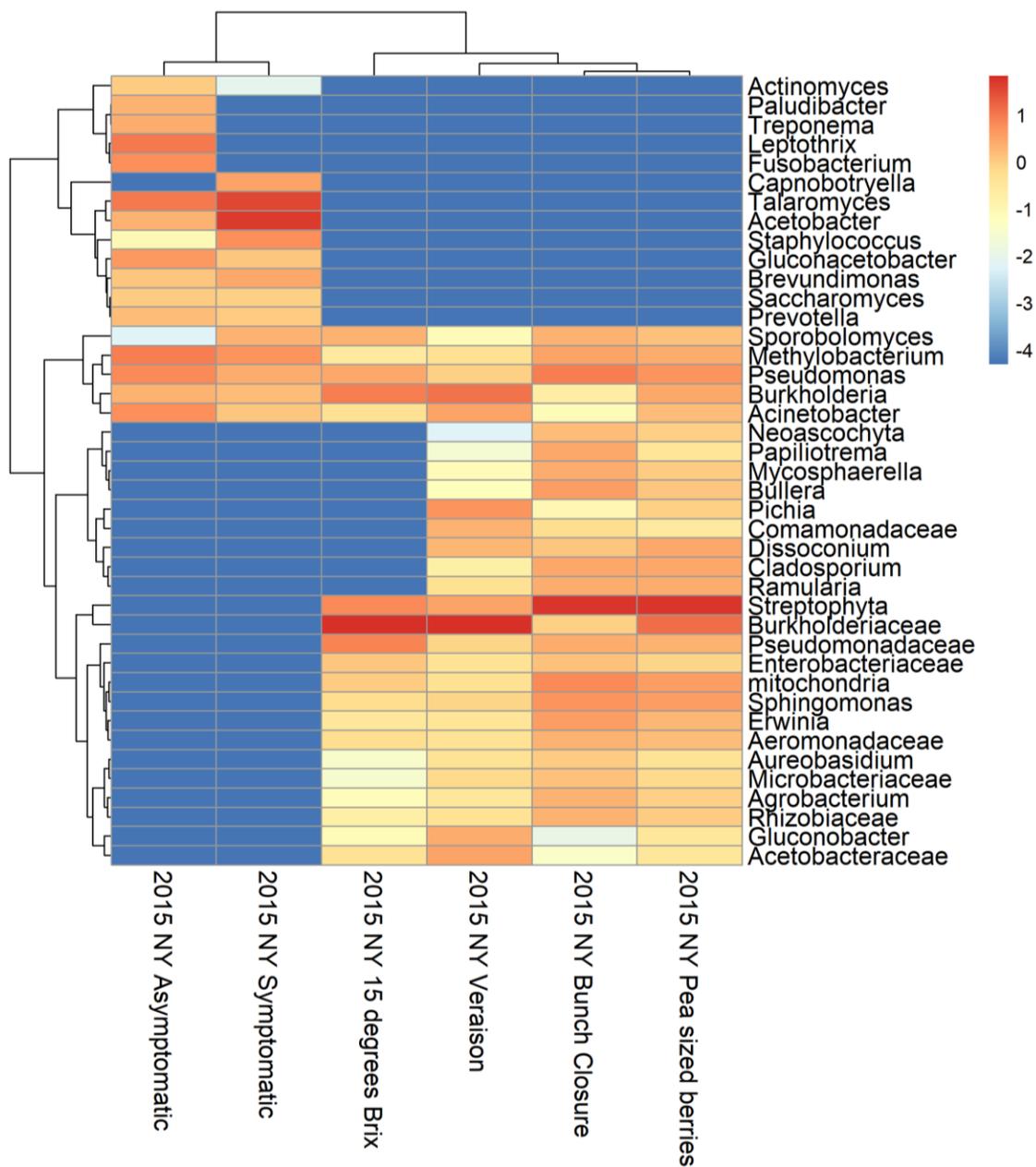


Across the New York 2015 developmental and disease data sets there were five bacteria and fungi shared among all of them (Figure 29), all of which are common grape epiphytes. This

is particularly interesting because the asymptomatic and symptomatic grapes were mancerated.

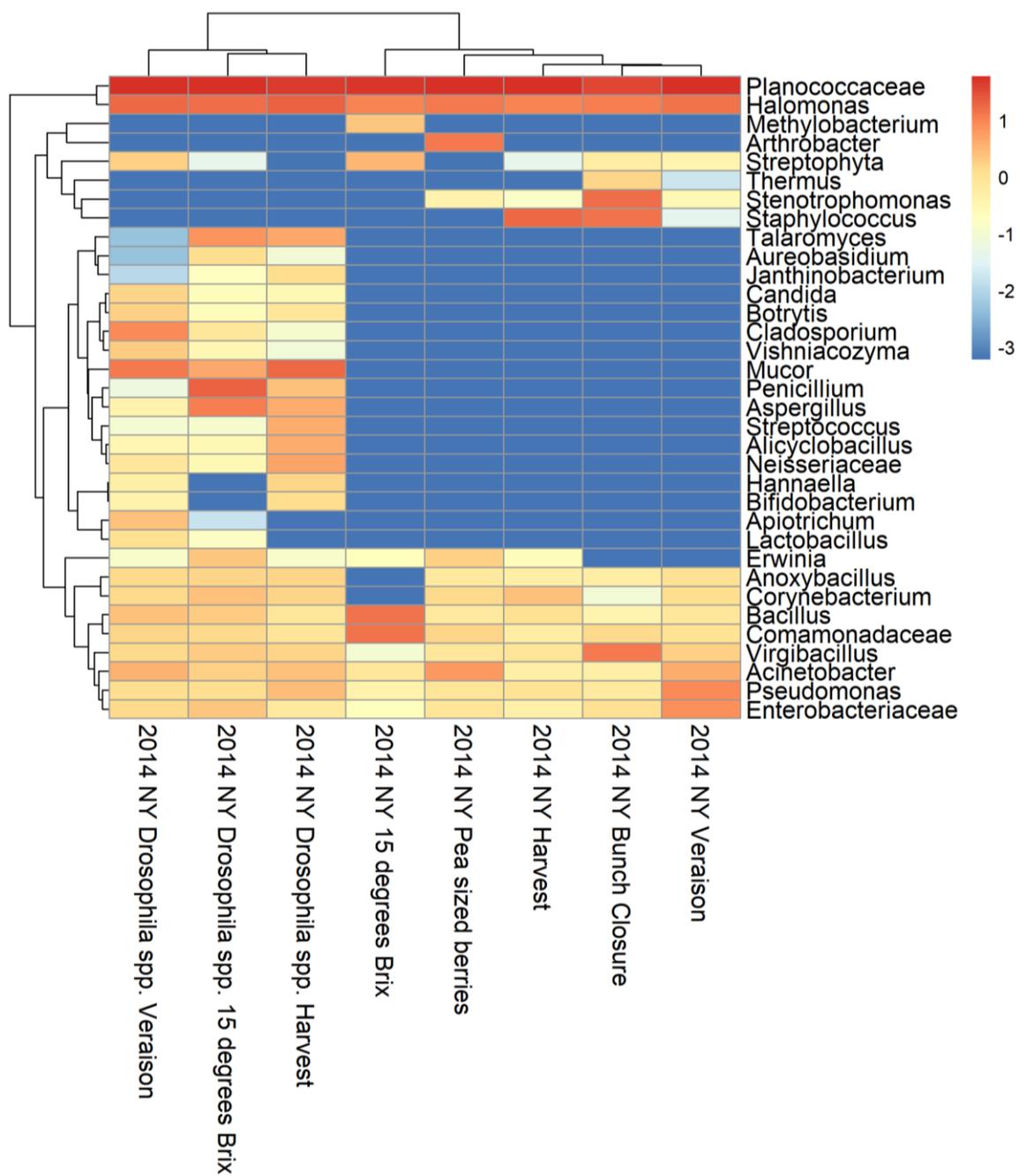
The 15° Brix group lacked ten microbes that were present in three earlier time points. There were many differences in bacterial and fungal populations between the two data sets.

Figure 29. Heatmap of all New York 2015 groups



Of all of the taxa present in the 2014 New York data sets, *Acinetobacter* was the only organism commonly present in all sample types in all environments (Figure 30). Many of the organisms only present in the *Drosophila* spp. are fungi as there was no grape fungal data. The bacteria present on the grapes and *Drosophila* spp. had relatively the similar abundances, especially *Planococcaceae* and *Halomonas*. The *Drosophila* spp. and grape communities each had several bacteria that were not present in the other.

Figure 30. Heatmap of all New York 2014 groups



Discussion:

Section 1: Data Filtering

FastQC was used to assess the quality of the sequences before and after quality filtering. Figure 7, Figure 8, and Figure 9 shows an example comparison between an untrimmed sequence and the same sequence trimmed using Qiime's and FastX's trimming methods. Qiime's trimming methods were used as they provided a better approach to trimming compared to FastX which only trims from the end until a base above the threshold is reached. In addition to this, Qiime's trimming method was incorporated into the function `split_libraries_fastq` that is used in this pipeline meaning it was easily integrated.

Section 2: Unclassified taxonomies

The difference in the number of OTUs that were unidentified at the genus and order levels is likely due to the well-known problem known as the '1000bp barrier of genus classification' (Brady, 2009). The '1000bp barrier' describes the problem of classifying bacteria to the genus level using less than 1000bp. Given this, achieving genus classification of 99.25% of classifiable OTUs is a very successful result. The percentage of unidentified bacterial genera was around 0.4% and percentage of unidentified bacterial orders was around 0.01% (Fig. 7). The fungal genera and orders had the same number of unidentified taxa, which could indicate that closed-reference OTU picking, in combination with the *de novo* method, contribute to a greater depth of identification (Fig. 8). The average percentage of unidentified fungal reads was 0.4%.

Section 3: Taxa Associated with Sour Rot Symptoms

The scientific community has found there to be three components of sour rot that must be present for the disease to start, and these are acetic acid producing bacteria, yeasts, and *Drosophila* spp. (Bisiach et al., 1986). As previously mentioned, these acetic acid producing bacteria are from the genera *Acetobacter* and *Gluconobacter* (Barata et al., 2012). In sour rot there are also large populations of the fermentative species *Pichia membranifaciens*, *Hanseniaspora uvarum*, *Metschnikowia pulcherrima*, and *Candida stellate* (Barata et al., 2012).

High RMFs of acetic acid producing bacteria were present the New York and Tasmania sour rot groups. The Tasmanian sour rot group also had *Pichia* compose the majority of its fungal community and had *Candida* and *Hanseniaspora* present as well. There was no significant differential abundance between the symptomatic group and asymptomatic group (Figure 13). In fact almost all groups from Tasmania had these fungi and bacteria as part of their core microbiome (Figure 28). For example, the *Drosophila* spp. sampled in Tasmania had *Pichia* and *Hanseniaspora* compose the majority of the *Drosophila* spp. fungal community and acetic acid producing bacteria were the most abundant bacteria (Figure 28). This supports previous research that have found *Drosophila* spp. are capable transporting the microorganisms needed to start sour rot (Rombaut et al., 2017).

The New York data set did not have any of the fungi that are typically associated with sour rot. This is likely be because *Pichia*, *Candida* and *Hanseniaspora* were not present during the later stages of development, as seen in Figure 22. This may have allowed for other fungi to colonize wounded grapes and start sour rot (Figure 29). There was a large population of the yeast *Talaromyces marneffe*, which until recently was named *Penicillium marneffe* (“Talaromyces

Marneffeii Infection”, 2018), that was present only in the sour rot symptomatic group. It therefore seems likely that *Talaromyces marneffeii* was able to take place of *Pichia* and the other fungi and start sour rot.

Section 4: Microbiomes and Terroir

The grape microbiome is known to influence “phenotypic characteristics, such as flavour, colour, and sugar content, thus influencing the winemaking process as well” (Mezzasalm et al. 2017). Papers such as Bokulich et al. have even speculated that the microbiome on grapes “could significantly affect grapevine and fruit health and development” (Bokulich et al. 2013). Because of this the relevant scientific community has researched how differences in location and terroir may be affected the microbiome and therefore the quality of the wine. So far the grape microbiome at harvest has been found to be effected by the cultivar, climate, terroir, and agricultural practices (Morgan, 2017). However, there has be little work done investigating the microbiome of the grape as it develops. The present research was therefore investigating if the microbiome changes as the grape develops or if the microbiome remains stable until harvest (Bokulich et al., 2013).

Figure 27 shows that the developmental stage of the grapes does not affect the core microbiome as much as the terroir, where the core microbiome is the set of bacteria and fungi that are present throughout the growing season. Each data set was clustered into its own clade and many organisms were only present in a specific data set. The common bacteria and fungi that were present in at least two data sets were: *Enterobacteriaceae*, *Comamonadaceae*, *Streptophyta*, *Erwinia*, *Pseudomonas*, *Acinetobacter*, *Burkholderia*, *Gluconobacter*, *Aureobasidium*, *Pichia*, and *Bacillus*.

The Tasmania 2016 data set had the largest sample sizes and had samples from sour rot infected grapes, *Drosophila* spp., and all five developmental time points. The most constant and abundant organisms present throughout all sample types were *Pichia*, *Gluconobacter*, *Bacillus*, and *Acetobacter* (Figure 28). Because of this it is likely that all that was needed for sour rot to develop was for *Drosophila* spp. to lay their eggs in the grapes to create the initial wounds.

As mentioned in the background, Prakitchaiwattana et al. found that *Metschnikowia* and *Hanseniaspora* were common on wounded grapes from New South Wales, Australia (Prakitchaiwattana et al., 2004). In the Tasmania data set, which is relatively close to New South Wales, *Hanseniaspora* was present not only on damaged sour rot symptomatic grapes but at all time points and on the *Drosophila* spp. However, *Metschnikowia* was not seen in any Tasmania group. A recent study investigating grape surface microbiomes at harvest in Italy found that the most ubiquitous and abundant fungal families that were *Dothioraceae*, *Pleosporaceae*, and *Saccharomycodaceae* (Mezzasalma et al., 2017). Unlike the Italian grape microbiomes, no fungi from either the *Pleosporaceae* or the *Dothioraceae* order was seen in Tasmania or New York. However, *Hanseniaspora* which is in the *Saccharomycodaceae* family was seen in all Tasmania groups.

In the New York 2015 growing season *Streptophyta* started as the predominate genus in the pea sized berry and bunch closure time points and then was replaced by bacteria from the *Burkholderia* family. All groups in this data set shared *Sporobolomyces* spp., a genus of non-fermenting yeast, *Burkholderia*, and several water and soil bacteria (*Methylobacteria*, *Pseudomonas*, and *Acinetobacter*). The three earlier time points clustered together and had more taxa in common than the 15° Brix group (Figure 29). This could be due to a natural progress in

the microbial community, a change in weather conditions, or a change in the vineyard such as pesticides or canopy management.

In 2014 in New York the pea sized berry group had small sample sizes of 4 and 6 for the bacteria and fungi respectively. The earlier developmental stages started with *Planococcaceae* and *Halomonas* and these genera remained a large contributors to the later communities. There were relatively few changes in taxa present on the *Drosophila* spp. at the three time points sampled. The *Drosophila* spp. groups had many bacteria that were not found in the grape samples, and they did not have several of the soil bacteria seen in the grapes. Because the 2014 New York data set lacked fungal information it is difficult to compare the differences and similarities between the 2014 and the 2015 New York data sets.

This project has shown that there are no distinctive changes in the epiphytic grape microbiome throughout its development from a pea sized berry to harvest. Instead it has clearly exemplified the unique microbiomes that are on the grape surface in a region for a given year. It is likely that the initial community at the beginning of the growing season along with factors such as climate, vineyard management practices, and other factors of terroir together affect the core epiphytic grape microbiome.

Further research should be done on possible yeasts such *Talaromyces marneffeii* that may be associated with sour rot. Another project should be done to determine if there are core microbiomes that contain sets of bacteria and fungi that are commonly found that could be classified into subgroups. Figure 27 shows the three distinct core microbiomes present in three different locations and years, however if there was a larger sample size of locations it could be possible to see further trends in what bacteria and fungi cohabitate grape surfaces together. Lastly, this field of study could be further enriched with a study that accesses the core

microbiome of the different terroirs and if there is a correlation between the core microbiome and traits in the resulting wine.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215(3):403-410.

Andrews, S (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>

Babraham Bioinformatics - FastQC A Quality Control Tool for High Throughput Sequence Data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed 20 Mar. 2018.

Barata, A., et al. "The Microbial Ecology of Wine Grape Berries." *International Journal of Food Microbiology*, vol. 153, no. 3, Feb. 2012, pp. 243–59. CrossRef, doi:10.1016/j.ijfoodmicro.2011.11.025.

Berg, Gabriele, et al. "Unraveling the Plant Microbiome: Looking Back and Future Perspectives." *Frontiers in Microbiology*, vol. 5, no. 148, 4 June 2014, pp. 1–7., doi:10.3389/fmicb.2014.00148.

Bisiach, M, et al. *Possible Integrated Control of Grapevine Sour Rot*. 1986, pp. 118–128.

Bokulich, N. A., J. H. Thorngate, P. M. Richardson, and D. A. Mills. "Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate." *Proceedings of the National Academy of Sciences* 111 (2013): 139-48. Web. 11 Apr. 2017.

Bordelon, Bruce. "Grapes: The Sour Rot Situation." Facts for Fancy Fruit, 25 Aug. 2016, <https://fff.hort.purdue.edu/article/grapes-the-sour-rot-situation/>.

Brady, Arthur, and Steven L. Salzberg. Phymm and PhymmBL: Metagenomic Phylogenetic Classification with Interpolated Markov Models. Vol. 6, no. 9, Sept. 2009, pp. 673–676, doi:10.1038/nmeth.1358.

Briceño, Erika X., and Bernardo A. Latorre. "Characterization of Cladosporium Rot in Grapevines, a Problem of Growing Importance in Chile." *Plant Disease*, vol. 92, no. 12, Nov. 2008, pp. 1635–42. *apsjournals.apsnet.org* (Atpyon), doi:[10.1094/PDIS-92-12-1635](https://doi.org/10.1094/PDIS-92-12-1635).

Campisano, Andrea, et al. "Bacterial Endophytic Communities in the Grapevine Depend on Pest Management." *PLoS ONE*, edited by Leonard Simon van Overbeek, vol. 9, no. 11, Nov. 2014, p. e112763. *CrossRef*, doi:[10.1371/journal.pone.0112763](https://doi.org/10.1371/journal.pone.0112763).

Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R. 2010. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 26:266-267.

“Citrus Black Rot (*Phyllosticta Ampellicida*).” Clemson University - USDA Cooperative Extension Slide Series, Bugwood.Org, 6 May 2011,
<http://www.forestryimages.org/browse/detail.cfm?imgnum=1234220>. Web. 19 Apr. 2017.

“Data Files and Other Resources.” *QIIME*, QIIME, qiime.org/home_static/dataFiles.html

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. 2006.
Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB.
Appl Environ Microb 72(7): 5069-5072.

Dharmadhikari, Murli. "Botrytis cinerea in Winemaking." Iowa State University Extension and Outreach, n.d. Web. 5 Apr. 2017.

"Economic Impact Study." *NGWI*. National Grape and Wine Initiative, n.d. Web. 01 Apr. 2017. < http://www.ngwi.org/economic-impact-study_226.html>

Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST.
Bioinformatics 26(19):2460-2461.

Emmett, Bob. *Fungal contaminants and their impact on wine quality*. Publication no. CRV 99/23. N.p.: Cooperative Research Centre for Viticulture, 2004. Print.

“Eurotiales”. New Brunswick Museum. <http://website.nbm-mnb.ca/mycologywebpages/NaturalHistoryOfFungi/Eurotiales.html>. Accessed 20 Apr. 2018.

“FAO Data - Dataset-Data-Filter - Crops Processed, National Production (FAOSTAT).” National Production (FAOSTAT), <http://ref.data.fao.org/dataset-data-filter?entryId=8dbed7b8-d99f-4933-b499-de864c476c39&tab=data>. n.d. Web. Accessed 01 Apr. 2017.

FASTX-Toolkit. Cold Spring Harbor Laboratories. Available at http://hannonlab.cshl.edu/fastx_toolkit/

Ficke, Andrea, et al. “Ontogenic Resistance and Plant Disease Management: A Case Study of Grape Powdery Mildew.” *Phytopathology*, vol. 92, no. 6, 2002, pp. 671–675.

Fuller, Katie. "The Value of Powdery Mildew RESistance in Grapes: Evidence from California." *Lodi Growers The Value of Powdery Mildew Resistance in Grapes Evidence from California Comments*. N.p., n.d. Web. 01 Apr. 2017.

Gianessi, Leonard, and Nathan Reigner. "The Importance of Fungicides in U.S. Crop Production." *Outlooks on Pest Management* 17.5 (2006): 209-13. Web. 2 Apr. 2017.

Golob, Jonathan L., et al. “Evaluating the Accuracy of Amplicon-Based Microbiome Computational Pipelines on Simulated Human Gut Microbial Communities.” *BMC Bioinformatics*, vol. 18, May 2017. *PubMed Central*, doi:[10.1186/s12859-017-1690-0](https://doi.org/10.1186/s12859-017-1690-0).

Gordon, Jim. "U.S. Wine Sales Total \$38 billion ." *Winesandvines.com*. N.p., 15 Jan. 2016. Web. 1 Apr. 2017.

H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2009.

Hall et al. "Temporal and Regional Shifts in the Surface Microbiota of Grapes within the Growing Season". 2018. Manuscript in preparation.

Hall et al. "Microbial Ecology of Sour Rot-Infected Grapes". 2018. Manuscript in preparation.

Hall, Megan, et al. "Causal Role of *Drosophila* Spp. in Sour Rot Development." American Society for Enology and Viticulture, <http://www.asev.org/abstract/causal-role-drosophila-spp-sour-rot-development>. Accessed 20 Apr. 2018.

Insel, Barbara, Bill Nelson, Nancy Light, Jim Bedient, Karen Ross, Tom Davenport, and Jim Trezise. *U.S. Wine, Grapes and Grape Products Contribute \$162 Billion to Economy*. Rep. WineAmerica, Wine Institute, Winegrape Growers of America, and the National Grape and Wine Initiative, 17 Jan. 2017. Web.

<<https://www.wineinstitute.org/resources/pressroom/01172007>>.

J Gregory Caporaso, Justin Kuczynski, Jesse Stombaugh, Kyle Bittinger, Frederic D Bushman, Elizabeth K Costello, Noah Fierer, Antonio Gonzalez Pena, Julia K Goodrich, Jeffrey I Gordon, Gavin A Huttley, Scott T Kelley, Dan Knights, Jeremy E Koenig, Ruth E Ley, Catherine A Lozupone, Daniel McDonald, Brian D Muegge, Meg Pirrung, Jens Reeder, Joel R Sevinsky, Peter J Turnbaugh, William A Walters, Jeremy Widmann, Tanya Yatsunenko, Jesse Zaneveld and Rob Knight; *Nature Methods*, 2010; doi:10.1038/nmeth.f.303

Jones, Laura, Summaira Riaz, Abraham Morales-Cruz, Katherine Ch Amrine, Brianna McGuire, W. Gubler, M. Walker, and Dario Cantu. "Adaptive genomic structural variation in the grape powdery mildew pathogen, *Erysiphe necator*." *BMC Genomics* 15.1 (2014): 1081. Web. 4 Apr. 2017.

Kolde, Raivo. "Pheatmap: pretty heatmaps." R package version 61 (2012).

König, Helmut, et al., editors. *Biology of Microorganisms on Grapes, in Must and in Wine*. Springer Berlin Heidelberg, 2009. CrossRef, doi:10.1007/978-3-540-85463-0.

Lebeis, Sarah L., et al. "Culturing a Plant Microbiome Community at the Cross-Rhodes." *New Phytologist*, vol. 196, no. 2, 17 Mar. 2012, pp. 341–344., doi:10.1111/j.1469-8137.2012.04336.x

Mezzasalma, Valerio, et al. “Grape Microbiome as a Reliable and Persistent Signature of Field Origin and Environmental Conditions in Cannonau Wine Production.” *PloS One*, vol. 12, no. 9, 2017, p. e0184615.

“Microbiome Sequencing (16s rRNA Sequencing).” *Gatc-Biotech.com*, GATC Biotech AG, 2017, www.gatc-biotech.com/en/expertise/targeted-sequencing/16s-rrna-analysis.html.

Morgan, Horatio H., et al. “The Grapevine and Wine Microbiome: Insights from High-Throughput Amplicon Sequencing.” *Frontiers in Microbiology*, vol. 8, May 2017. CrossRef, doi:10.3389/fmicb.2017.00820.

Parks DH, Tyson GW, Hugenholtz P, Beiko RG. (2014). STAMP: Statistical analysis of taxonomic and functional profiles. *Bioinformatics*, 30, 3123-3124.

“Pierce’s Disease.” The Wine Institute, <https://www.wineinstitute.org/initiatives/issuesandpolicy/piercesdisease>. n.d. Web. Accessed 02 Apr. 2017.

Prakitchaiwattana, Cheunjit J., et al. “Application and Evaluation of Denaturing Gradient Gel Electrophoresis to Analyse the Yeast Ecology of Wine Grapes.” *FEMS Yeast Research*, vol. 4, no. 8, Sept. 2004, pp. 865–77. *academic.oup.com*, doi:10.1016/j.femsyr.2004.05.004.

R Core Team (2013). R: A language and environment for statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

Ries, Stephen M. "Reports on Plant Diseases." *IPM : Reports on Plant Diseases : Black Rot of Grape*. University of Illinois at Urbana-Champaign, Dec. 1999. Web. 06 Apr. 2017.

Rombaut, A., et al. "Invasive *Drosophila Suzukii* Facilitates *Drosophila Melanogaster* Infestation and Sour Rot Outbreaks in the Vineyards." *Royal Society Open Science*, vol. 4, no. 3, Mar. 2017. PubMed Central, doi:10.1098/rsos.170117.

Scholz, Matthias. Operational Taxonomic Unit (OTU) - Metagenomics. <http://www.metagenomics.wiki/pdf/definition/operational-taxonomic-unit-otu>. Accessed 19 Oct. 2017.

Smith, Damon. "Botrytis Bunch Rot (Gray Mold) of Grapes." *EXtension*. N.p., 3 Oct. 2013. Web. 05 Apr. 2017.

Smith, Damon. "Ripe Rot of Grapes." *EXtension*. Oklahoma State University, Oct. 2013. Web. 05 Apr. 2017.

Smith, R. J., L. J. Bettiga, and W. D. Gubler. "UC Pest Management Guidelines: Grape Powdery Mildew." *UC IPM Online*. Statewide Integrated Pest Management Program, Feb. 2016. Web. 02 Apr. 2017

Smith, R. J., L. J. Bettiga, and W. D. Gubler. "UC Pest Management Guidelines: Grape Botrytis Bunch Rot." *UC IPM Online*. Statewide Integrated Pest Management Program, Jan. 2014. Web. 02 Apr. 2017

Smith, R. J., L. J. Bettiga, and W. D. Gubler. "UC Pest Management Guidelines: Grape Summer Bunch Rot (Sour Rot)." *UC IPM Online*. Statewide Integrated Pest Management Program, Jan. 2014. Web. 02 Apr. 2017

Steel, Christopher. "Secondary bunch rots of grapevine." *EPPO Bulletin* 35.3 (2014): 583-85. Web. 6 Apr. 2017.

"Talaromyces Marneffeii Infection." Leading International Fungal Education, <http://www.life-worldwide.org/fungal-diseases/talaromyces-marneffeii-infection>. Accessed 21 Mar. 2018.

Taylor, Andrew. "Know Your Grapevine Bunch Rots." Department of Agriculture and Food, 17 Jan. 2017, <https://www.agric.wa.gov.au/summer/know-your-grapevine-bunch-rots>. . Web. 04 Apr. 2017.

"VitisGen-Grape Industry Priorities." *VitisGen-Grape Industry Priorities*. Cornell University, 2015. Web. 01 Apr. 2017. <<http://www.vitisgen.org/industry.html>>

Weigle, Tim. "Managing Black Rot." *Managing Black Rot / Viticulture and Enology*. Cornell University, Mar. 2014. Web. 06 Apr. 2017.

Wightwick, Adam, Robert Walters, Graeme Allinson, Suzanne Reichman, and Neal Menzies. "Environmental Risks of Fungicides Used in Horticultural Production Systems." *Horticultural Production Systems, Fungicides* (2010): 273-304. Web.

Wunderlich, Lynn, Karen Klonsky, and Don Stewart. "2015 Sample Costs To Establish A Vineyard And Produce Wine Grapes." *University Of California Cooperative Extension* (2015): n. pag. Web.

Youakim, Sami. "Occupational health risks of wine industry workers." *BC Medical Journal* 48.8 (n.d.): 386-91. Web. Oct. 2006.