Title:

Quantifying the relative roles of selective and neutral processes in defining eukaryotic microbial communities

Authors: Peter Morrison-Whittle & Matthew R Goddard

Affiliations: The School of Biological Sciences, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

Email addresses: pmor072@aucklanduni.ac.nz, m.goddard@auckland.ac.nz.

Running title: Selection regulates fungal community assembly

Subject category: Microbial population and community ecology

Type of article: Original article

Corresponding Author: Peter Morrison-Whittle. The School of Biological Sciences, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand. Phone: +64 (09) 9239537. pmor072@aucklanduni.ac.nz

Number of Words: 5 133 (Abstract – 179, main text – 5 162)

Number of references: 58.

Number of figures and tables: Two figures, two tables, five online supplementary tables and one online supplementary figure.

Abstract

We have a limited understanding of the relative contributions of different processes that regulate microbial communities, which are crucial components of both natural and agricultural ecosystems. The contributions of selective and neutral processes in defining community composition are often confounded in field studies because as one moves through space, environments also change. Managed ecosystems provide an excellent opportunity to control for this and evaluate the relative strength of these processes by minimizing differences between comparable niches separated at different geographic scales. We use next-generation sequencing to characterize the variance in fungal communities inhabiting adjacent fruit, soil, and bark in comparable vineyards across 1,000 kms in New Zealand. By compartmentalizing community variation, we reveal that niche explains at least four times more community variance than geographic location. We go beyond merely demonstrating that different communities are found in both different niches and locations by quantifying the forces that define these patterns. Overall, selection unsurprisingly predominantly shapes these microbial communities, but we show the balance of neutral processes also play a significant role in defining community assemblage in eukaryotic microbes.

Keywords: Biogeography/community compositions/deep community sequencing/ fungi/ microbes/selection

Introduction

Disentangling the processes regulating species' distributions makes a significant step toward achieving an integrated understanding of natural and managed ecosystems (Gaston 2000). A better understanding of the factors underpinning community structure also ensures greater accuracy for predictions concerning both species conservation and communities' fates under various environmental change scenarios (Fierer *et al.* 2003; Lennon & Jones 2011; Ferrenberg *et al.* 2013). Differences in the types and abundances of species in space may arise through both natural-selection-driven and non-selection driven processes, which are directly analogous to classic population genetic processes operating within species (Nemergut *et al.* 2013). Selective processes may drive differences between communities through species sorting in response to local conditions (Fierer & Jackson 2006; Hughes *et al.* 2008); non-selective processes may generate variation through a combination of other assembly processes including dispersal limitation, community drift, and speciation (Fierer *et al.* 2003; Vellend 2010; Hanson *et al.* 2012; Nemergut *et al.* 2013). These non-selective processes - which we define here as "neutral" processes - are considered in ecological neutral theories (Bell 2001; Hubbell 2005) and are predicted to produce variation in community structure through space without needing to invoke the actions of selection.

While we have a limited understanding of processes that drive broad community patterns for macrobes (Gaston 2000; Wiens 2011), we have far less data on how these translate to microbial communities, which are crucial components of natural, agricultural, and biotechnological ecosystems (Bardgett *et al.* 2008; Van Der Heijden *et al.* 2008; Fuhrman 2009; Hanson *et al.* 2012). Due to their cryptic nature, it is relatively more challenging to examine patterns in microbial than macrobial communities. Massively parallel next-generation sequencing technologies have allowed significant leaps forward in the power with which we may sample microbial communities. These methods do not have the constraint of having to culture species to analyse them, but instead

63 evaluate DNA directly isolated from the substrate of interest (Su *et al.* 2012; Kautz *et al.* 2013;
64 Bokulich *et al.* 2014; Taylor *et al.* 2014). So far, studies utilizing these techniques are providing
65 increasing evidence that many microbial communities are not homogenized through space, but
66 display significant structure just like many plant and animal communities do (Martiny *et al.* 2006;
67 Hanson *et al.* 2012; Nemergut *et al.* 2013). These findings have principally been based on studies
68 examining bacteria (Fierer & Jackson 2006; Bryant *et al.* 2008; Fulthorpe *et al.* 2008; Lauber *et al.*
69 2009; Knief *et al.* 2010; Ghiglione *et al.* 2012); the relatively fewer studies examining eukaryotic
70 microbes also show differential patterns of community composition in space (Dumbrell *et al.* 2010;
71 Scheckenbach *et al.* 2010; Bokulich *et al.* 2014; Taylor *et al.* 2014).
72
73 Early speculations concerning the forces that regulate microbial community composition are
74 represented by the classic Baas Becking hypothesis "everything is everywhere but the environment
75 selects". This suggests that, due to small individual sizes but large populations, dispersal is not
76 limiting and that only selection due to environmental gradients accounts for differences in
77 community structure (De Wit & Bouvier 2006). This idea directly informs a pervading view which
78 implicitly assumes that the actions of natural selection dominate, and many microbial community
79 ecology studies have correspondingly only attempted to evaluate the role of selection in community
80 assemblage. Overall there has been less focus on testing whether microbial communities may
81 become differentiated as a consequence of various neutral processes (Hanson *et al.* 2012; Nemergut
82 *et al.* 2013). While it is unrealistic to imagine that only one or the other of these 'selective' or
83 'neutral' processes are at play, it not clear is how important each process is in influencing any
84 structure we observe. Consequently, without accounting for relative contribution of neutral
85 processes we cannot assume the dominance of selection without empirically testing this first. In
86 some sense the field could be considered in an analogous position to evolutionary biology before
87 Gould and Lewontin's classic 'Spandrels' paper (Gould & Lewontin 1979) which suggested that
88 organisms' traits must be proved to have arisen via natural selection rather than assume they have.

89

90 The scant data attempting to disentangle these two drivers of microbial community composition are
91 conflicting, and primarily come from studies with bacteria (Hanson *et al.* 2012). Studies examining
92 gastrointestinal and soil subsurface bacterial communities show the primacy of selection, but also
93 indicate that neutral factors play a secondary role (Jeraldo *et al.* 2012; Stegen *et al.* 2012; Stegen *et*
94 *al.* 2013; Wang *et al.* 2013). In contrast bacterial communities in wastewater plants appear primarily
95 defined by neutral community assemblage (Ofiţeru *et al.* 2010), and neutral effects primarily
96 influence desert photosynthetic bacterial assemblages (Caruso *et al.* 2011). Dumbrell *et al.* (2010)
97 evaluated the factors regulating the formation of soil arbuscular mycorrhizal fungal communities
98 sampled less than 20 meters apart, and showed these communities are shaped primarily by
99 selection, but there was also a contribution from underlying neutral processes. However, we are
100 unaware of any study that has tested and quantified the relative role of selection in defining
101 eukaryotic microbial communities over larger scales, where processes like dispersal limitation and
102 community drift may become increasingly important.
103
104 The difficulty facing studies attempting to separate and quantify the effects of neutral and selective
105 processes as regulators of community structures is that as distance increases, environmental
106 similarity often decreases. Thus, studies that sample across reasonable distances tend also to sample
107 different environments, confounding one's ability to determine if neutral or selective processes
108 largely define any observed differences in communities (Bell 2001; Chase & Myers 2011; Ranjard *et*
109 *al.* 2013). So far statistical treatments have been developed to attempt to overcome this problem
110 separating these effects (Dumbrell *et al.* 2010; Chase & Myers 2011; Stegen *et al.* 2012; Stegen *et al.*
111 2013).
112

113 Here we employ a complimentary but alternative approach and attempt to minimize confounding
114 selective and neutral processes experimentally by evaluating community diversity in three adjacent
115 niches within artificially managed ecosystems (vineyards) replicated regions.
116 Vineyards naturally harbour a range of microbial taxa including diverse fungal communities (Bokulich
117 *et al.* 2014; Pinto *et al.* 2014; Taylor *et al.* 2014). The conserved design and management of vineyard
118 ecosystems (which comprise the same clone of plant, with comparable associated habitats and
119 niches) across large distances provide a powerful system to separately estimate the impact of
120 selective and neutral processes. Within any one vineyard, there are no apparent barriers to prevent
121 the mixing of species across physically adjacent niches through, for example, insect mediated
122 dispersal (Buser *et al.* 2014). At the same time these niches are separated ecologically by multiple
123 environmental gradients that provide an opportunity for selection to influence community
124 composition. Such habitat comparisons are rarely possible using unmanaged ecosystems, which vary
125 more greatly in space, and frequently compound extreme barriers to dispersal and environmental
126 gradients.
127
128 Here we use next generation sequencing to evaluate 106 contemporaneous fungal communities
129 inhabiting adjacent soil, bark, and fruit niches across six New Zealand regions spanning a thousand
130 kilometres. Using this design we empirically test and compare the effect of both selective and
131 neutral processes on community structure using permutational multivariate anova of community
132 dissimilarities. If selection dominates then communities will tend to differ more greatly between
133 niches regardless of distance, and if neutral processes dominate then communities will tend to differ
134 more greatly by distance regardless of niche. Here we attempt a major step forward by not just
135 robustly estimating whether any microbial community patterns exist, but going on to experimentally
136 quantify the relative contributions of selective and neutral processes in defining community
137 composition (Chase & Myers 2011; Hanson *et al.* 2012).
138

Materials and Methods

Sampling and extraction

Approximately 30g of vine bark, soil, and ripe fruit were aseptically sampled from *Vitis vinifera* var. Sauvignon blanc vineyards across six regions of New Zealand: Hawkes Bay, Martinborough, Nelson, Central Otago, and the Wairau and Awatere valleys in Marlborough. These regions span approximately 38° – 45° S and 168° – 177° E, around 1,000 km NE to SW gradient (Fig. 1 B). Six vineyards were selected in each region, and three sub-samples were taken evenly across each vineyard for each niche and pooled for a total of 108 samples (three niches sampled across 36 vineyards). All samples were taken at least five meters from row ends to avoid edge effects. Soil samples were taken from directly under vines 50 cm away from the central trunk. All samples were taken approximately two weeks before harvest in April 2011 and transported to the laboratory in sterile containers on ice. Fruit samples were washed with 300 mls of sterile water to remove epiphytes and then centrifuged; the resulting pellet was re-suspended in 500 µl of sterile water. Prior to DNA extraction, all samples were stored at -20°C. DNA was extracted from all samples using the Zymo Research Soil Microbe DNA MiniPrep™ kits. We empirically determined this procedure was sufficient to extract DNA from fruit and bark samples as well as soil (data not shown). The 600 bp D1/D2 region of the 26S ribosomal RNA locus was amplified using the NL1 and NL4 fungal specific primers (Kurtzman & Robnett 2003). This locus provides a readily aligned homogeneous PCR product and also provides good signal for community differentiation (Taylor *et al.* 2014). Thirty-six different multiplex identifiers were added to the primers to bioinformatically distinguish between samples. Two fruit samples failed to amplify from raw DNA extractions, reducing the total number of samples to 106. PCR products were cleaned with AmpureXP beads to remove all primer dimers, and their quality confirmed with Agilent DNA1000 chips. PCR products were uni-directionally sequenced on a full plate of a 454 Life Sciences GS FLX instrument by Macrogen (Korea).

Data Analysis

Sequence handling and processing was conducted with Mothur v. 1.30 (Schloss *et al.* 2009). Raw sequences for each sample are present in GenBank (accession number: SRP048520). Initially primers and reads <200bp were removed, and then low-quality reads and homopolymer errors were identified and removed using the pyronoise algorithm. PCR chimeras were identified and removed using the uchime algorithm. Individual sample identifiers were assigned to the resulting good quality reads, which were then merged and analysed together. Unique sequences were identified and compared to a fungal reference database, and those not assigned to Fungi (6.76% of all unique sequences, 5.75% of reads) were removed. While there will likely be differential genetic diversity within fungal species, empirical studies suggest that multiple species of Ascomycota and Basidiomycota differ by less than 2% at the 26s rDNA gene (Kurtzman & Robnett 2003; Romanelli *et al.* 2010). Thus the remaining 486 279 reads were aligned against a fungal reference database and clustered into groups that share >98% identity. These are considered the lowest level of molecular operational taxonomic units (MOTUs) and approximate species. We conservatively removed any 'singleton' MOTUs, those represented by just a single read, from all further analyses. An unequal number of reads were obtained across samples (ranging from 1 257 to 8 007), and so we re-sampled (rarefied) each community down to the sample with the lowest number of reads to produce a dataset with equivalent sampling effort for all communities, and this totalled 133 454 reads. MOTUs were compared to a fungal taxonomic database using a Bayesian approach, and each was classified to all levels above genus with the 'classify.seqs' command in Mothur. Consensus sequences that matched less than 70% at any one taxonomic level were listed as unclassified; the final data set is provided in Supplemental table S1.

Statistical analysis

Following Crist *et al.* (2003) we employed null models to test whether there is evidence to support the idea that any variance in observed species richness is no more than we would expect to see by chance given a random distribution of individuals across our samples. Species richness can be partitioned into α-diversity (the average number of species per sample unit), β-diversity (species richness difference between the average sample unit and the overall species pool), and γ-diversity (the total number of species observed in a defined area). α- and β-diversity was analysed for all three niches across all vineyards independently and compared with α- and β-diversity measures predicted by randomisation simulations. Community null-models are typically employed to analyse variation within niches and to our knowledge no analytical tools have been designed to allow inclusion of multiple niches. To verify whether within-niche regional patterns of richness held for overall vineyard richness we additionally tested whether overall vineyard richness was randomly assembled by pooling species presence from each of the three vineyard niches. Our experimental design, sampling replicate vineyards in replicate regions, allows comparisons of average species richness between vineyards and both their regional as well as national species richness.

To more comprehensively analyse the effects of niche and geographic location on both community composition (species richness) and community structure (species abundances), we conducted a two-way permutational multivariate ANOVA (Permanova) using a Jaccard (metric Bray-Curtis) community dissimilarity metric [PERMANOVA (Anderson 2001)]. These comparisons were then replicated using a number of other metrics (Bray-Curtis, Euclidean, and Manhattan). Permanova tests were also conducted separately with both niche and region as fixed effects, and the Benjamini-Hochberg multiple test correction (Benjamini & Hochberg 1995) was employed to evaluate the suite of results. Variance in community structure described by Jaccard dissimilarities was visualized using principal coordinate analysis in R (R Development Core Team 2013) and graphed using JMP (SAS Institute Inc 1989-2007). Mantel tests testing for correlations between community dissimilarity and geographic

spatial separation were conducted for each niche separately. In addition, since we have sequence data for all taxa, we evaluated communities using weighted-unifrac analysis (Lozupone *et al.* 2006). This analysis accounts for species abundances as well as phylogenetic relatedness and uses randomization to assess the degree to which observed community phylogenetic dissimilarity differs from those expected randomly. All statistical testing were carried out using the vegan package (Dixon 2003) in R with the exception of weighted-unifrac analysis that was performed in Mothur v. 1.30.

Results

Overall fungal diversity

Analyses of the sequences from the replicate soil, vine bark, and fruit samples revealed the presence of 3 583 fungal MOTUs (herein called species). Taxonomic assignment of representative sequences for each species, by comparison to a fungal reference database, reveal members from five fungal phyla, 24 classes, 59 orders, 104 families, and 166 genera (see Appendix S1). The Ascomycota dominated comprising 45.3% of all MOTUs, followed by the Basidiomycota with 9.6%. Table 1 shows species richness (i.e. presence/absence) broken down by both niche and region. Species richness significantly differs by niche but not region (two-way ANOVA, $F_{[2, 89]}$=662.471, P<0.0001 and $F_{[5, 89]}$=0.942, P=0.457 respectively). In contrast to evidence for latitudinal patterns in animal communities (Gaston 2000), there was no significant correlation between latitude and species richness for our sampled fungal communities in bark, fruit, or soil (Bark: P<0.117 and r<0.266, Fruit: P<0.095 and r<-0.291, Soil: P<0.954 and r<0.010). While the total number of species in each region was approximately similar (averaging around 1 100), there were 9-fold and 4-fold more species in soil and bark than fruit respectively. Just 2.1 % (76 species, Fig. 1A) of species were found in all three adjacent niches and accounted for 55.1%, 93.7%, and 37.1% of all reads in Bark, Fruit, and Soil communities respectively. If the number of reads assigned to each species is reasonably assumed to

represent taxa abundance, then species abundance approximately follows the typical power law decay distribution seen in communities generally (Bell 2001). A breakdown of the abundance of major phyla in all three niches is shown in Fig. 1C, and the breakdown over all taxonomic levels is shown in Appendix S1. Approximately 84, 72, and 87 per cent of species are unclassified at the genus level in bark, fruit, and soil respectively, indicating the relative extent of uncharacterised fungal diversity in these niches. The diversity of fungal species relative to sampling effort reported here is approximately in line with the only two other studies that employed next-generation sequencing approaches to analyse fungal communities associated with vineyards: 253 MOTUs in Taylor *et al.* (2014) and 158 MOTUs in Bokulich *et al.* (2014). However, fungal diversity in these 36 NZ vineyard soils appears much greater than estimates from just two Italian vineyards with around 300 MOTUs in Orgiazzi *et al.* (2012), but one might expect greater diversity with a greater sampling area. Previous reports of fungal diversity found on vine bark derived from live culture only and yielded just five species (Sabate *et al.* 2002); however, next-generation sequencing of fungal communities from vine leaves reveals around a thousand MOTUs and is thus in line with our inferences from bark here (Pinto *et al.* 2014).

Testing for patterns of community differences

Our aim was to quantify the relative roles of selective and neutral processes in defining differences in communities, but this aim critically requires that differences exist. We first tested whether any patterns in community composition were evident in these data. The results of our additive diversity partitioning hierarchical null model test revealed that both average vineyard and average regional diversity was significantly lower than expected under a random species distribution across these New Zealand sample sites ($P<0.0001$; See Appendix S2). Consequently regional and national β-diversity was both significantly different than predicted from random models ($P<0.0001$; See Appendix S2). Average national β-diversity (11.1) is approximately three times greater than average regional β-diversity (3.6), which shows individual communities on average tend to be more similar to

265 other communities within their region than to those of other regions. These patterns of α- and β-
266 diversity hold for each niche separately and for overall vineyard diversity. Together, these results
267 provide support to reject the concept that these communities are assembled randomly, at least in
268 terms of species presence and absence.
269
270 Permanova analyses also accounting for species abundances revealed that both geographic location
271 and niche significantly affect both community diversity and composition (See Table 2). The
272 significance of these effects, and their relative R^2 values, were stable when this analysis was
273 conducted using a range of different community dissimilarity measures (Bray-Curtis, Euclidean,
274 Manhattan; see Appendix S3). These effects can be visualized in the first two dimensions of classic
275 multidimensional scaling of Jaccard dissimilarities, and this is shown in Fig. 2. Reducing the data to
276 include only those species found in all regions (i.e. region-specific species removed) did not greatly
277 affect the P or R^2 values ($P<0.0001$, $R^2=0.070$), indicating that regional differentiation is not purely a
278 product of region-specific species, but differences in the proportion of shared species. The results of
279 permanova also revealed a significant interaction between both niche and region (See Table 2). To
280 dissect this interaction, we explored the 63 possible comparisons of communities between niches
281 within regions, and between regions for each niche. Once multiple-testing correction had been
282 applied, these unsurprisingly revealed differential patterns of regional delineations for niches (see
283 Appendix S4), with soil communities displaying the greatest number of significant regional
284 differences. The results of weighted-unifrac analysis also reveal a significant effect of both niche and
285 location on community composition (Average Niche comparison W-score= 1, $P<0.001$; Average
286 Regional comparison W-score=0.65, $P<0.001$; see Appendix S5). Together these various tests
287 converge to reveal a significant effect of both niche and geographic location on fungal community
288 composition.
289
290

The effect of geographic distance on ecological dissimilarity

Mantel correlations between Jaccard community dissimilarities and geographic distance matrices carried out for each niche independently revealed no effect of distance on bark communities, a borderline-significant correlation for fruit, and a significant correlation for soil communities (Bark: $P<0.111$, $r=0.103$; Fruit: $P=0.050$, $r= 0.148$; Soil: $P<0.0001$, $r=0.355$). Within-region mantel tests examining correlations between community and geographic distance for both soil and fruit niches show this relationship breaks down at smaller distances. The average distance between sites within regions is 12 kms, and the maximum is around 100 kms (see Appendix S6). After multiple pairwise test correction, only one of the 18 within region niche-specific community mantel tests reported a significant relationship. This result suggests that 100 kms is the cut-off for similarity of fungal communities inhabiting soil and possibly fruit niches.

Quantifying the effects of selection and neutral processes

Having demonstrated the significant effects of both niche and geographic location on fungal community composition, we went on to quantify the relative amount of observed variance in community diversity and composition explained by these factors. The R^2 values derived from permanova analyses allow us to assess this directly. Using community distance metrics based on only the presence/absence of species, geographic location explains an average of 6.1% of the variance in community diversity, but niche explains 27.1% (evaluated using both binary Bray-Curtis and Jaccard). When species abundances are taken into account, the average proportion of the total variation explained by niche was 31.9%, approximately four times the proportion of variation explained by region at just 7.2%. Thus, while both location and niche, and their interaction, appear to play a significant role in defining differences in eukaryotic microbial community diversity and composition, niche of isolation appears to explain at least four times more of the community variance than geographic location. Given this experimental design and analysis this shows that both selective and

316 neutral effects play a role in defining fungal communities, but that selective effects are
317 approximately four times stronger.
318
319 **Discussion**
320 We sampled fungal communities from three different adjacent niches that are relatively conserved
321 at scales of meters to thousands of kilometres with the aim of quantifying the relative degree to
322 which selective and neutral processes define community composition. Direct sequencing of DNA
323 from these niches reveals approximately the same extent of fungal diversity recovered from similar
324 niches in the handful of other studies that have used this approach (Bokulich *et al.* 2014; Pinto *et al.*
325 2014; Taylor *et al.* 2014). These communities are not homogenous; they significantly vary in terms of
326 species richness and community composition, and both niche and geographic location affect this
327 variance. In addition, there is a significant interaction between these main effects, and this is likely
328 driven by the fact that while communities in each niche demonstrate geographic differences overall,
329 the nature of these geographic patterns are slightly different within each of these three niches (see
330 Appendix S4). While the extent of latitude covered is relatively small, these data show no significant
331 correlation between species richness and latitude for communities in any niche. This finding aligns
332 with the lack of evidence for latitudinal patterns reported for microbes (Fierer & Jackson 2006), and
333 contrasts diversity patterns of macro-organisms (Gaston 2000; Hillebrand 2004). Further, we provide
334 evidence that there is significant distance decay between fungal communities, within at least soils,
335 over distances of about 100 kms. This overall result nicely recapitulates patterns revealed from the
336 population genetic analyses of just one species in this community sourced from these same samples:
337 *Saccharomyces cerevisiae* (Knight & Goddard 2014). The recovery of the same patterns at both the
338 species and community level lends confidence to the ability of these data to reveal underlying
339 biological signals.
340

By estimating the relative strength of selective and neutral processes in generating variation in community structure, these results go beyond merely demonstrating that different species are found in different places (either by niche or by geography). Given the individually conserved adjacent niches sampled over varying distances, these data show that communities tend to be more similar by niche than by distance. Selective forces will drive similarities between niches regardless of distance, and a range of neutral forces will drive similarities between geographic locations regardless of niche. Crucially, this means we show that niche effects may explain approximately one-quarter of the variance in community composition, but only one-twentieth of the variance, four-fold less, may be explained by geographic location. While regionally structured communities have been observed in vineyard communities before (Gayevskiy & Goddard 2011; Bokulich *et al.* 2014; Taylor *et al.* 2014), this is the first study to estimate of the relative strength of selective and neutral processes in defining microbial communities across regional scales for this managed ecosystem. Currently there are no studies examining whether this pattern holds more broadly across other ecosystems, or at larger scales. Thus, these data start to fill the gap in our understanding of the relative magnitude of the forces that regulate community composition for eukaryotic microbes, which are crucial components of both natural and agricultural ecosystems. While our data show selection is influencing community structure to a greater degree than the neutral processes are, we show the balance of neutral forces is significant. In principle any of the neutral forces of differential speciation, community drift, dispersal limitation or their interaction might underlie the neutral component defining community difference in this system. Our experiment was not designed to tease these apart, and our data do not allow us to do this. Instead here we evaluate the net effect of these processes (within niche variation) compared to selective processes (between niche variation).

Our experimental design aimed to capture and control for variation in community structure in response to two separate and frequently co-varying and potentially confounding niche and distance effects. In our attempt to disentangle these effects, we make an assumption that communities

residing in the same niche across different locations were exposed to approximately equivalent selection pressures. While vineyard ecosystems are managed and conserved by human design, these are ultimately open ecosystems. It is therefore unlikely that our assumption of ecological equivalence of the various niches across distances was fully satisfied. Of the three niches, soil likely differs most across NZ (Ranjard *et al.* 2013). Soil chemistry and physical properties were not controlled for in our sampling design, all of which have been implicated in microbial turnover (Dumbrell *et al.* 2010; Ranjard *et al.* 2013). As a consequence, it is reasonably likely that some of the differences between soil communities in different locations originate as a product of selection, not neutral effects. It seems reasonable that this effect will be lesser for the communities inhabiting bark and fruit as all samples originated from the same clonal variety of plant species. However, selective effects are unlikely to be entirely absent as different soils will likely translate into different plant phenotypes. We also acknowledge that region specific variation may also have clearly arisen in all niches from climatic differences between regions. A number of environmental gradients exist across New Zealand's regions, such as temperature (although this varies just 2-3 degrees), precipitation rates, and UV radiance. However, we feel these considerations do not prevent us from reaching meaningful inferences about the relative strength of the forces regulating these communities overall because if intra-niche community variation is jointly influenced by neutral processes and natural selection to undetected environmental variance, then intra-niche variation will simply represent the maximum limit of the influence of neutral processes. This means we can refine our summary by concluding that the influence of selection is at least four times stronger than neutral forces.

Very little work has been conducted to examine the relative balance of selection and neutral effects on regulating microbial community structure, and this has been largely focused on bacterial communities. However, our findings are in line with previous work by Stegen *et al*. (2012; 2013), who showed selection is the dominant effect shaping subsurface soil bacterial communities over small distances (<53 metres) while neutral forces also play a role; and Bell (2010) who concluded

393 that dispersal limitation played a minor role over distances of under one km. Our findings are nicely
394 in line with Dumbrell *et al.* (2010), who examined arbuscular mycorrhizal fungal communities less
395 than 25m apart, and concluded that niche partitioning was the primary mechanism regulating the
396 composition and diversity of communities, but these communities are also influenced by neutral
397 processes. Our results are also consistent with studies of larger plants and animals which indicate
398 the influence of both selection and neutral effects on community composition (Turnbull *et al.* 2005;
399 Farnon Ellwood *et al.* 2009). Broadly speaking then, along with other work, our results suggest the
400 primacy of selection as the consistently strongest effect on community regulation over large scales,
401 but show that balance of neutral effects are also important. How this balance changes across
402 different ecosystems for different microbial taxa remains to be seen. While we know that neutral
403 processes can generate variation, we do not know *a priori* how great this variation is in any
404 ecosystem, and must be measured, not assumed.
405
406 The primacy of selection in shaping communities generally gives rise to the prediction that
407 environmental change will have a greater effect on community composition than if neutral forces
408 were the dominant driver of community structure. Human-mediated habitat change and irregular
409 weather events with greater extremes of conditions predicted under climate change models will
410 likely manipulate the nature of selective forces and thus affect macrobial and microbial communities
411 in both natural and managed agricultural ecosystems. However, the absolute speed with which
412 species within these communities adapt to these changes will likely be greater for microbes due to
413 their faster generation times and larger population sizes (Goddard & Bradford 2003).
414
415

Acknowledgements

We thank Sarah Knight and Soon Lee who assisted in sample collection and processing of samples, Peter Tsai for bioinformatic assistance and Alexandria Leonard for her assistance in editing the manuscript. This work was funded by grants to MG from the New Zealand Ministry of Business, Innovation and Employment, Plant and Food Research Ltd and New Zealand Winegrowers. The completion of this research would not have been possible without the cooperation and assistance of the many collaborating companies who allowed access to their land: Amisfeild, Ata Rangi, Churton, Coal Pit, Constellation, Delegats, Domain Road, Frey Vineyard, Huia, Misha's Vineyard, Mt Difficulty, Mt Riley, Neudorf, Palliser, Pernod Ricard, Rippon, Seifried, Seresin, Te Kairanga, Te Whare Ra, Tohu, Trinity Hill, Villa Maria and Vita Brevis.

Conflict of Interest

We are not aware of any conflict of interest in carrying out this study.

Supplementary Information

Supplementary information is available at ISME's website.

References

1.

Anderson, M.J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, 26, 32-46.

2.

Bardgett, R.D., Freeman, C. & Ostle, N.J. (2008). Microbial contributions to climate change through carbon cycle feedbacks. *ISME Journal*, 2, 805-814.

3.

Bell, G. (2001). Neutral macroecology. *Science*, 293, 2413-2418.

4.

Bell, T. (2010). Experimental tests of the bacterial distance-decay relationship. *ISME Journal*, 4, 1357-1365.

5.

Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289-300.

6.

Bokulich, N.A., Thorngate, J.H., Richardson, P.M. & Mills, D.A. (2014). Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. *Proceedings of the National Academy of Sciences of the United States of America*, 111, E139-E148.

7.

Bryant, J.A., Lamanna, C., Morlon, H., Kerkhoff, A.J., Enquist, B.J. & Green, J.L. (2008). Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 11505-11511.

8.

Buser, C.C., Newcomb, R.D., Gaskett, A.C. & Goddard, M.R. (2014). Niche construction initiates the evolution of mutualistic interactions. *Ecology Letters*.

9.

Caruso, T., Chan, Y., Lacap, D.C., Lau, M.C.Y., McKay, C.P. & Pointing, S.B. (2011). Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. *ISME Journal*, 5, 1406-1413.

10.

Chase, J.M. & Myers, J.A. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366, 2351-2363.

11.

Crist, T.O., Veech, J.A., Gering, J.C. & Summerville, K.S. (2003). Partitioning Species Diversity across Landscapes and Regions: A Hierarchical Analysis of α, β, and γ Diversity. *American Naturalist*, 162, 734-743.

12.

De Wit, R. & Bouvier, T. (2006). 'Everything is everywhere, but, the environment selects'; what did Baas Becking and Beijerinck really say? *Environmental Microbiology*, 8, 755-758.

13.

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, 14, 927-930.

14.

Dumbrell, A.J., Nelson, M., Helgason, T., Dytham, C. & Fitter, A.H. (2010). Relative roles of niche and neutral processes in structuring a soil microbial community. *ISME Journal*, 4, 337-345.

15.

Farnon Ellwood, M.D., Manica, A. & Foster, W.A. (2009). Stochastic and deterministic processes jointly structure tropical arthropod communities. *Ecology Letters*, 12, 277-284.

16.

Ferrenberg, S., O'Neill, S.P., Knelman, J.E., Todd, B., Duggan, S., Bradley, D. *et al.* (2013). Changes in assembly processes in soil bacterial communities following a wildfire disturbance. *ISME Journal*, 7, 1102-1111.

17.

Fierer, N. & 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.

18.

Fierer, N., Schimel, J.P. & Holden, P.A. (2003). Influence of drying-rewetting frequency on soil bacterial community structure. *Microbial Ecology*, 45, 63-71.

19.

Fuhrman, J.A. (2009). Microbial community structure and its functional implications. *Nature*, 459, 193-199.

20.

Fulthorpe, R.R., Roesch, L.F.W., Riva, A. & Triplett, E.W. (2008). Distantly sampled soils carry few species in common. *ISME Journal*, 2, 901-910.

21.

Gaston, K.J. (2000). Global patterns in biodiversity. *Nature*, 405, 220-227.

22.

Gayevskiy, V. & Goddard, M.R. (2011). Geographic delineations of yeast communities and populations associated with vines and wines in New Zealand. *ISME Journal*.

23.

Ghiglione, J.F., Galand, P.E., Pommier, T., Pedrós-Alió, C., Maas, E.W., Bakker, K. *et al.* (2012). Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 17633-17638.

24.

Goddard, M.R. & Bradford, M.A. (2003). The adaptive response of a natural microbial population to carbon- and nitrogen-limitation. *Ecology Letters*, 6, 594-598.

25.

Gould, S.J. & Lewontin, R.C. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *PROC. R. SOC. LONDON B BIOL. SCI.*, 205, 581-598.

26.

Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C. & Martiny, J.B.H. (2012). Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nature Reviews Microbiology*, 10, 497-506.

545 27.

546 Hillebrand, H. (2004). On the generality of the latitudinal diversity gradient. *American Naturalist*,
547 163, 192-211.

548
549 28.

550 Hubbell, S.P. (2005). Neutral theory in community ecology and the hypothesis of functional
551 equivalence. *Functional Ecology*, 19, 166-172.

552
553 29.

554 Hughes, A.R., Inouye, B.D., Johnson, M.T.J., Underwood, N. & Vellend, M. (2008). Ecological
555 consequences of genetic diversity. *Ecology Letters*, 11, 609-623.

556
557 30.

558 Jeraldo, P., Sipos, M., Chia, N., Brulc, J.M., Dhillon, A.S., Konkel, M.E. *et al.* (2012). Quantification of
559 the relative roles of niche and neutral processes in structuring gastrointestinal microbiomes.
560 *Proceedings of the National Academy of Sciences of the United States of America*, 109, 9692-
561 9698.

562
563 31.

564 Kautz, S., Rubin, B.E.R., Russell, J.A. & Moreaua, C.S. (2013). Surveying the microbiome of ants:
565 Comparing 454 pyrosequencing with traditional methods to uncover bacterial diversity.
566 *Applied and Environmental Microbiology*, 79, 525-534.

567
568 32.

569 Knief, C., Ramette, A., Frances, L., Alonso-Blanco, C. & Vorholt, J.A. (2010). Site and plant species are
570 important determinants of the Methylobacterium community composition in the plant
571 phyllosphere. *ISME Journal*, 4, 719-728.

572
573 33.

574 Knight, S. & Goddard, M.R. (2014). Quantifying separation and similarity in a *Saccharomyces*
575 *cerevisiae* metapopulation. *ISME Journal*, in press (accepted).

576
577 34.

578 Kurtzman, C.P. & Robnett, C.J. (2003). Phylogenetic relationships among yeasts of the
579 'Saccharomyces complex' determined from multigene sequence analyses. *FEMS Yeast*
580 *Research*, 3, 417-432.

581
582 35.

583 Lauber, C.L., Hamady, M., Knight, R. & Fierer, N. (2009). Pyrosequencing-based assessment of soil pH
584 as a predictor of soil bacterial community structure at the continental scale. *Applied and*
585 *Environmental Microbiology*, 75, 5111-5120.

586
587 36.

588 Lennon, J.T. & Jones, S.E. (2011). Microbial seed banks: The ecological and evolutionary implications
589 of dormancy. *Nature Reviews Microbiology*, 9, 119-130.

590
591 37.

592 Lozupone, C., Hamady, M. & Knight, R. (2006). UniFrac - An online tool for comparing microbial
593 community diversity in a phylogenetic context. *BMC Bioinformatics*, 7.

594
595 38.

596 Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L. *et al.* (2006).
597 Microbial biogeography: Putting microorganisms on the map. *Nature Reviews Microbiology*,
598 4, 102-112.

599
600 39.

601 Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F. *et al.* (2013).
602 Patterns and processes of microbial community assembly. *Microbiology and Molecular*
603 *Biology Reviews*, 77, 342-356.

604
605 40.

606 Ofiţeru, I.D., Lunn, M., Curtis, T.P., Wells, G.F., Criddle, C.S., Francis, C.A. *et al.* (2010). Combined
607 niche and neutral effects in a microbial wastewater treatment community. *Proceedings of*
608 *the National Academy of Sciences of the United States of America*, 107, 15345-15350.

609
610 41.

611 Orgiazzi, A., Lumini, E., Nilsson, R.H., Girlanda, M., Vizzini, A., Bonfante, P. *et al.* (2012). Unravelling
612 soil fungal communities from different mediterranean land-use backgrounds. *PLoS ONE*, 7.

613
614 42.

615 Pinto, C., Pinho, D., Sousa, S., Pinheiro, M., Egas, C. & Gomes, A.C. (2014). Unravelling the diversity of
616 grapevine microbiome. *PLoS ONE*, 9.

617
618 43.

619 R Development Core Team (2013). R: A language and environment for statistical computing. R
620 Foundation for Statistical Computing,

621 Vienna, Austria.

44.

Ranjard, L., Dequiedt, S., Chemidlin Prévost-Bouré, N., Thioulouse, J., Saby, N.P.A., Lelievre, M. *et al.* (2013). Turnover of soil bacterial diversity driven by wide-scale environmental heterogeneity. *Nature Communications*, 4.

45.

Romanelli, A.M., Sutton, D.A., Thompson, E.H., Rinaldi, M.G. & Wickes, B.L. (2010). Sequence-based identification of filamentous basidiomycetous fungi from clinical specimens: A cautionary note. *Journal of Clinical Microbiology*, 48, 741-752.

46.

Sabate, J., Cano, J., Esteve-Zarzoso, B. & Guillamón, J.M. (2002). Isolation and identification of yeasts associated with vineyard and winery by RFLP analysis of ribosomal genes and mitochondrial DNA. *Microbiological Research*, 157, 267-274.

47.

SAS Institute Inc (1989-2007). JMP®. SAS Institute Inc. Cary, NC.

48.

Scheckenbach, F., Hausmann, K., Wylezich, C., Weitere, M. & Arndt, H. (2010). Large-scale patterns in biodiversity of microbial eukaryotes from the abyssal sea floor. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 115-120.

49.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B. *et al.* (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, 75, 7537-7541.

50.

Stegen, J.C., Lin, X., Fredrickson, J.K., Chen, X., Kennedy, D.W., Murray, C.J. *et al.* (2013). Quantifying community assembly processes and identifying features that impose them. *ISME Journal*, 7, 2069-2079.

51.

Stegen, J.C., Lin, X., Konopka, A.E. & Fredrickson, J.K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME Journal*.

52.

662 Su, C., Lei, L., Duan, Y., Zhang, K.Q. & Yang, J. (2012). Culture-independent methods for studying
663 environmental microorganisms: Methods, application, and perspective. *Applied*
664 *Microbiology and Biotechnology*, 93, 993-1003.

665
666 53.

667 Taylor, M.W., Tsai, P., Anfang, N., Ross, H.A. & Goddard, M.R. (2014). Pyrosequencing reveals
668 regional differences in fruit-associated fungal communities. *Environmental Microbiology*.

669
670 54.

671 Turnbull, L.A., Manley, L. & Rees, M. (2005). Niches, rather than neutrality, structure a grassland
672 pioneer guild. *Proceedings of the Royal Society B: Biological Sciences*, 272, 1357-1364.

673
674 55.

675 Van Der Heijden, M.G.A., Bardgett, R.D. & Van Straalen, N.M. (2008). The unseen majority: Soil
676 microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology*
677 *Letters*, 11, 296-310.

678
679 56.

680 Vellend, M. (2010). Conceptual synthesis in community ecology. *Quarterly Review of Biology*, 85,
681 183-206.

682
683 57.

684 Wang, J., Shen, J., Wu, Y., Tu, C., Soininen, J., Stegen, J.C. *et al.* (2013). Phylogenetic beta diversity in
685 bacterial assemblages across ecosystems: Deterministic versus stochastic processes. *ISME*
686 *Journal*, 7, 1310-1321.

687
688 58.

689 Wiens, J.J. (2011). The niche, biogeography and species interactions. *Philosophical Transactions of*
690 *the Royal Society B: Biological Sciences*, 366, 2336-2350.

691

692

693

694

695 **Figure Legends**

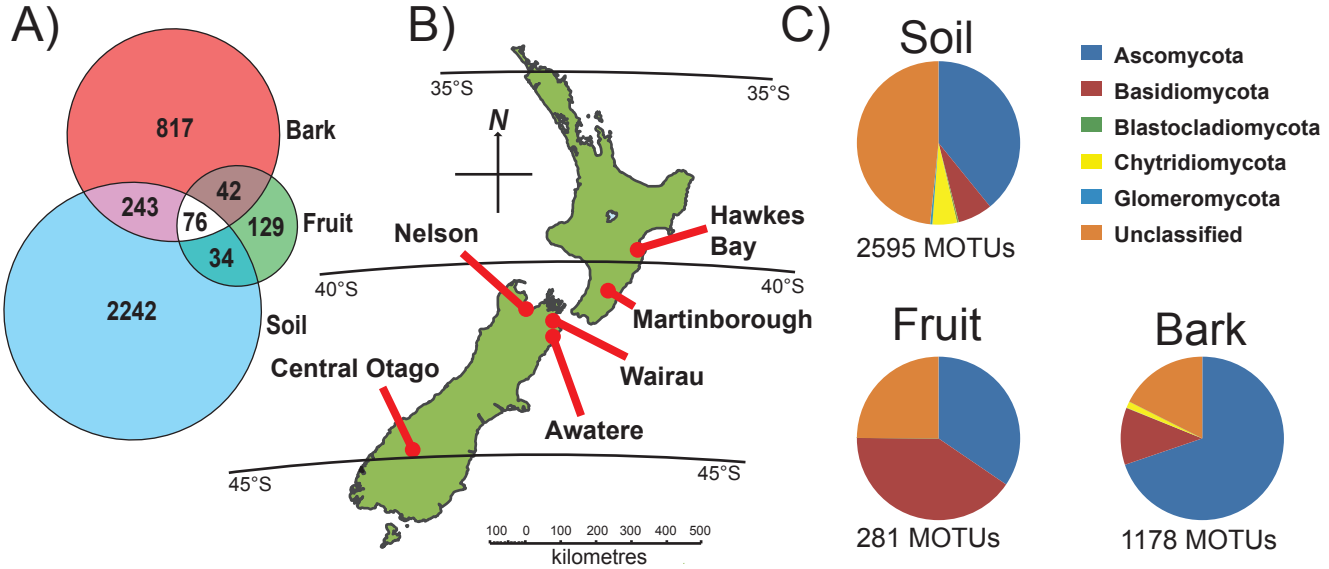
696 Fig. 1: A) Community diversity (number of species) among all three vineyard niches (bark, fruit, soil)
697 and their overlap. B) The six New Zealand wine-growing regions sampled. C) Relative proportions of
698 total MOTU diversity among niches for the five fungal phyla detected.

699

700

701 Fig. 2: First two dimensions of classic multidimensional scaling of Jaccard distances between samples
702 from different niches and regions.

703


A)
817
Bark
42
243
76
129
Fruit
34
2242
Soil
B)
35°S
35°S
N
Nelson
Hawkes Bay
40°S
40°S
Martinborough
Wairau
Central Otago
Awatere
45°S
45°S
100 0 100 200 300 400 500
kilometres
C)
Soil
2595 MOTUs
Ascomycota
Basidiomycota
Blastocladiomycota
Chytridiomycota
Glomeromycota
Unclassified
Fruit
281 MOTUs
Bark
1178 MOTUs

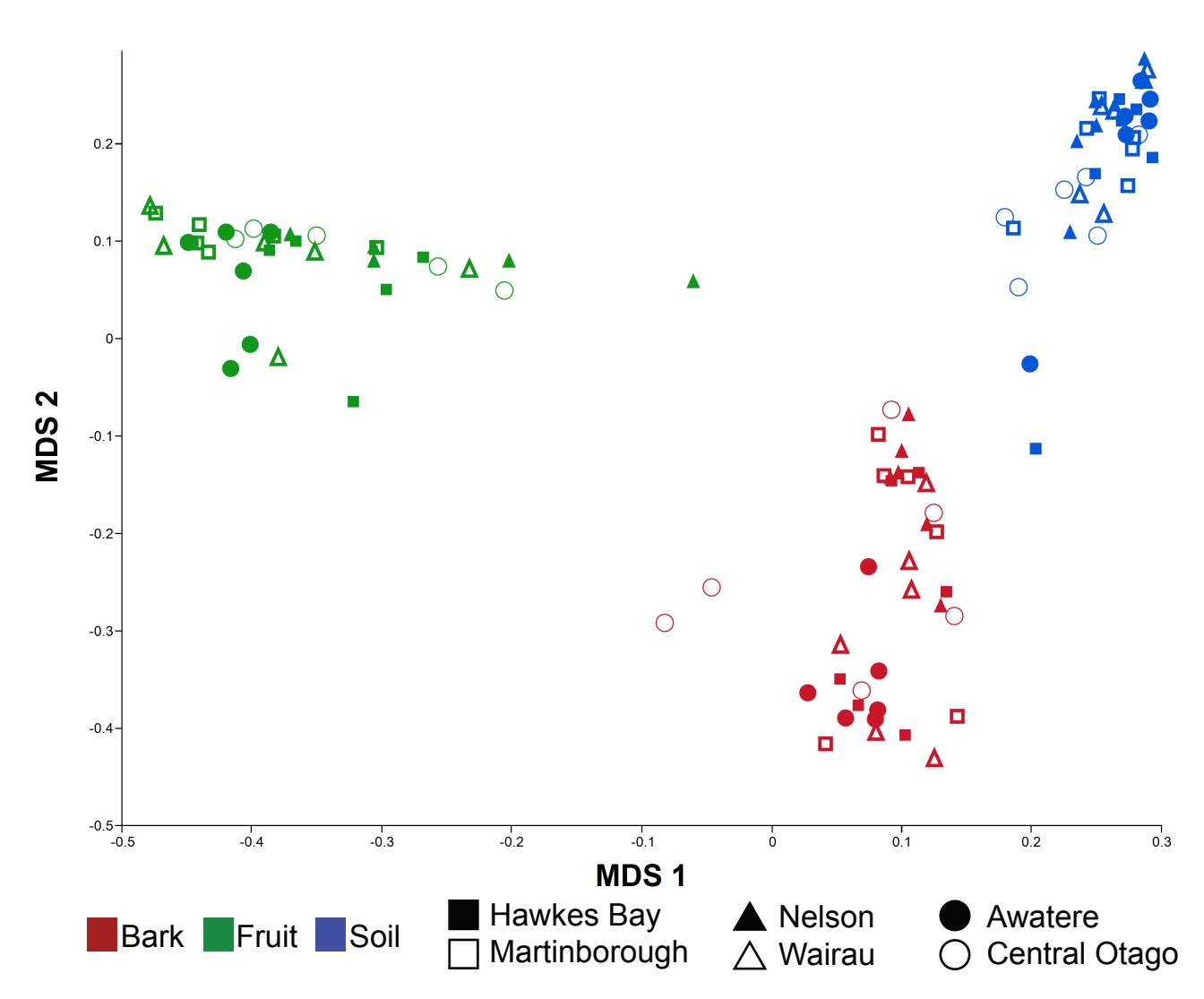

MDS 2
MDS 1
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
Bark
Fruit
Soil
Hawkes Bay
Martinborough
Nelson
Wairau
Awatere
Central Otago

Table 1: Species richness by region and niche. The numbers of species in all regions and niches are not the sum of the rows and columns as many species are present in more than one niche and location.

	Hawkes Bay	**Martinborough**	**Nelson**	**Wairau**	**Awatere**	**Central Otago**	**All Regions**
Bark	402	411	411	394	297	334	1178
Fruit	82	88	57	90	92	99	281
Soil	830	849	808	858	752	845	2595
All Niches	1202	1214	1171	1217	1037	1131	

Table 2: Results of Permutation ANOVA of Jaccard dissimilarities between fungal communities found in New Zealand vineyards (9999 permutations).

Effect	**df**	**SS**	**MS**	**Pseudo-F**	**R^2**	***P***
Region	5	3.017	0.603	2.211	0.07	<0.001
Niche	2	11.125	5.562	20.383	0.257	<0.001
Region x Niche	10	5.144	0.514	1.885	0.119	<0.001
Residuals	88	24.014	0.273	0.555		