

1 **Supplementary Information**

2
3 **Soil classification predicts differences in prokaryotic communities across a range**
4 **of geographically distant soils once pH is accounted for**

5 Authors: Rachel Kaminsky, Blandine Trouche, and Sergio E. Morales

6 Department of Microbiology and Immunology, Otago School of Medical Sciences,
7 University of Otago, Dunedin, New Zealand

8
9
10 **Supplementary Figures**

11 **Figure S1:** Relationship between pH and diversity metrics.

12 **Figure S2:** ANOSIM test for correlations between land use and Bray-Curtis
13 distances.

14 **Figure S3:** Hierarchical clustering for all sample data.

15 **Figure S4:** ANOSIM test for correlations between soil classification and Bray-Curtis
16 distances.

17 **Figure S5:** Hierarchical clustering for high country sample data.

18 **Figure S6:** Hierarchical clustering for sheep and beef sample data.

19 **Figure S7:** Hierarchical clustering for dairy sample data.

20 **Figure S8:** Ordinations of regional subgroups (Land use).

21 **Figure S9:** Ordinations of regional subgroups (Soil classification).

22 **Figure S10:** ANOSIM test for correlations between soil classification and land use
23 and Bray-Curtis distances for South Canterbury soils.

24 **Figure S11:** ANOSIM test for correlations between soil classification and land use
25 and Bray-Curtis distances for Southland soils.

26 **Figure S12:** ANOSIM test for correlations between soil classification and land use
27 and Bray-Curtis distances for North Canterbury soils.

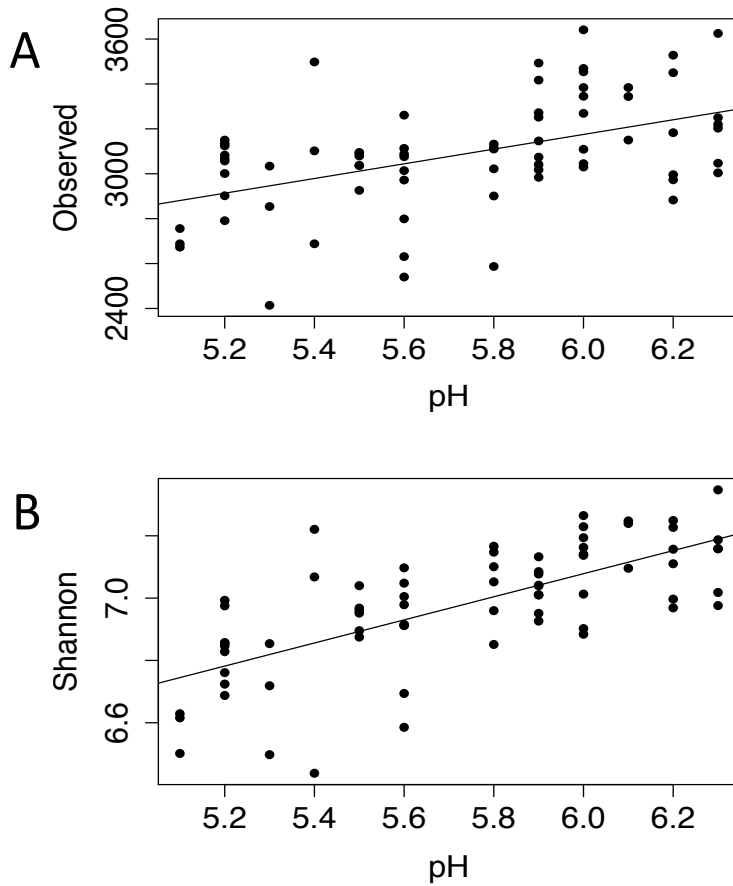
28 **Figure S13:** ANOSIM test for correlations between soil classification and land use
29 and Bray-Curtis distances for Otago soils.

30 **Supplementary Tables**

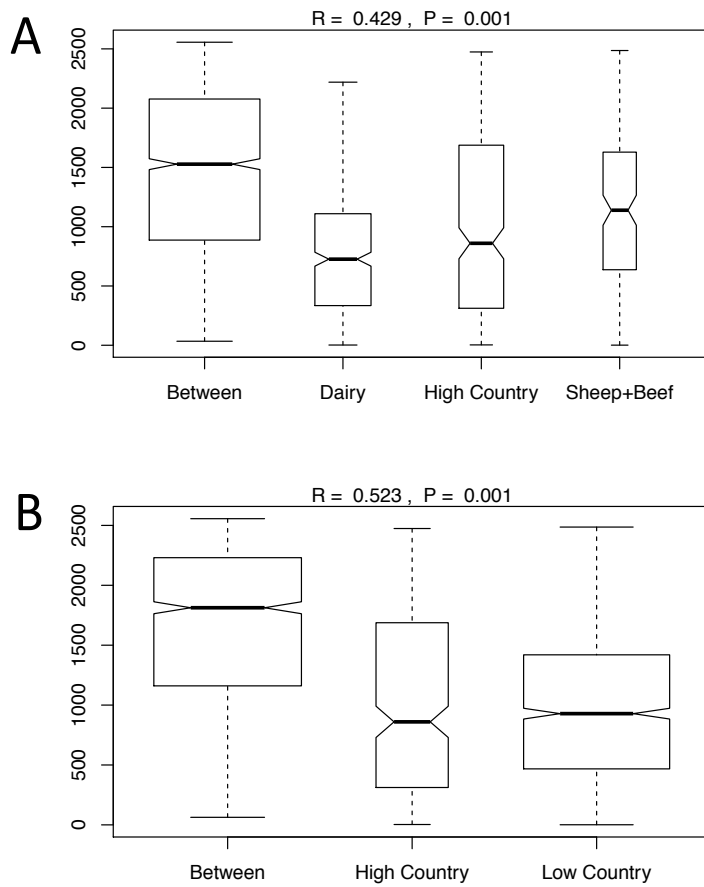
31 **Table S1:** Soil physicochemical properties measured in this study. Separate file.

32 **Table S2:** Mantel correlations for regions.

33 **Table S3:** Taxa that are significantly correlated to different physicochemical
34 parameters presented with statistical support. Separate file.
35
36

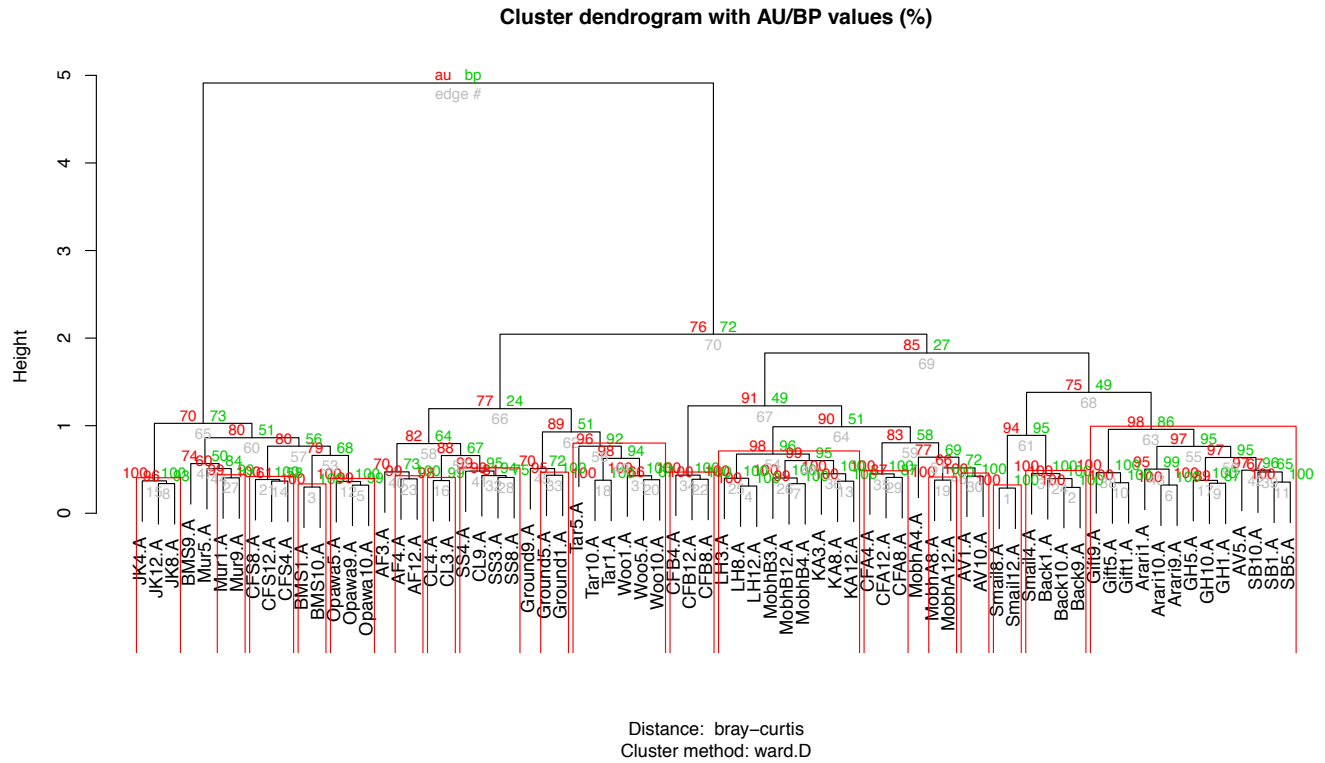


37
38 **Figure S1:** Relationship between pH and diversity metrics. (A) Observed species and
39 (B) Shannon diversity regressed with pH values. Line represents the best-fit linear
40 model.



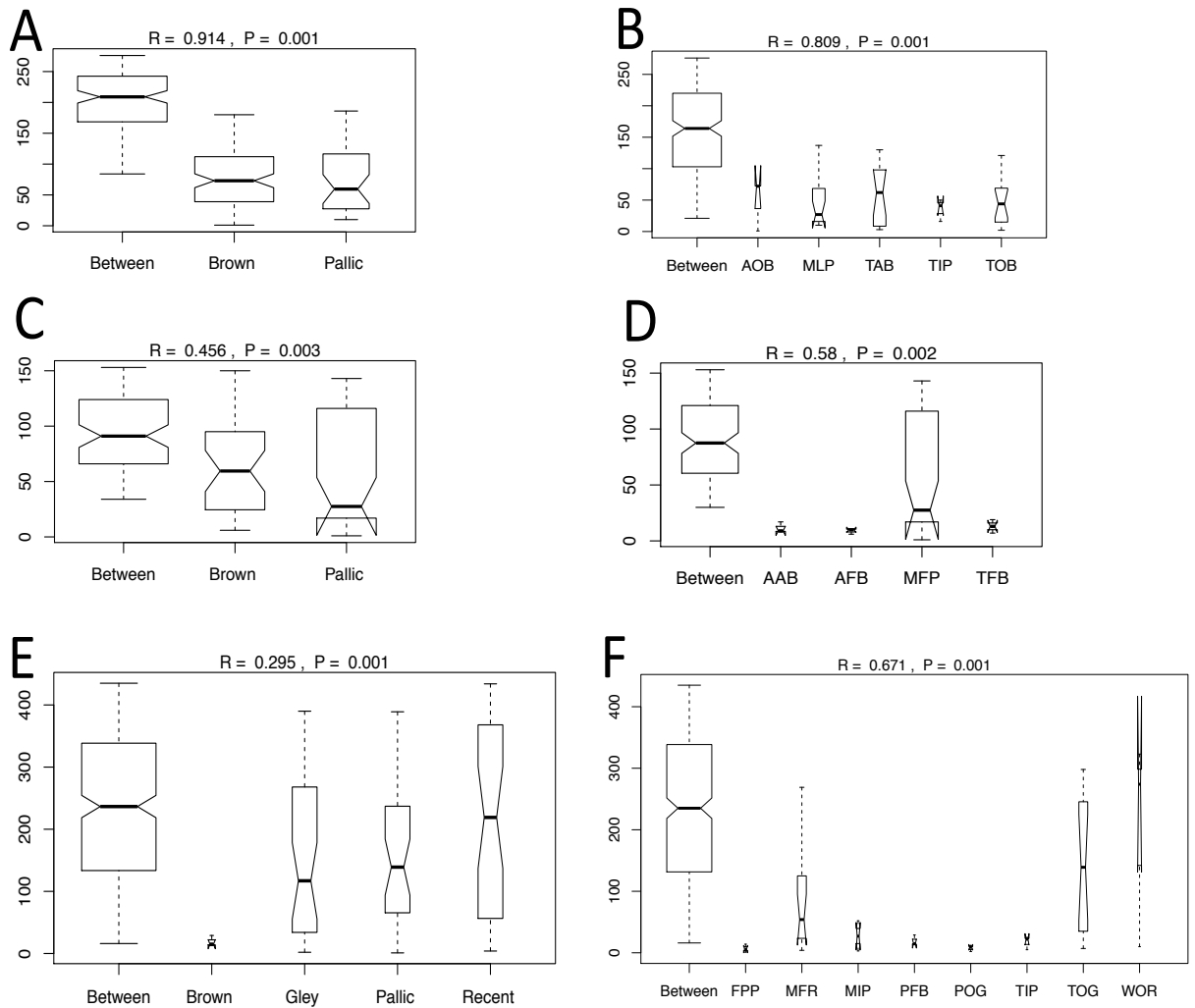
41

42 **Figure S2:** ANOSIM test for correlations between land use and Bray-Curtis
 43 distances. (A) Ranked dissimilarities plotted against the three land uses and (B) high
 44 country versus low country, where low country includes both sheep and beef and
 45 dairy samples. Boxplots are drawn with widths proportional to the square roots of the
 46 number of observations in each land use and with 95% confidence intervals.



47

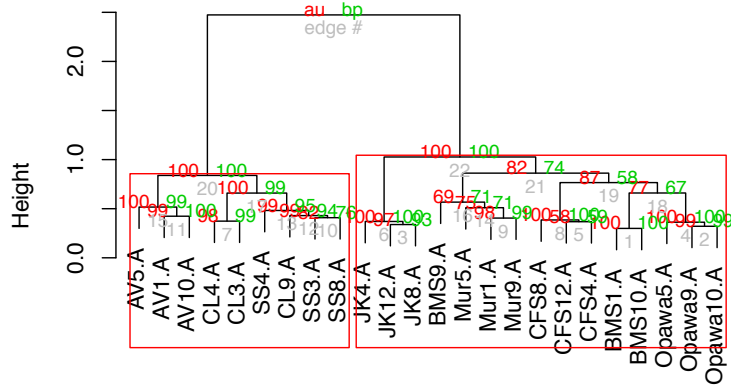
48 **Figure S3:** Hierarchical clustering for all sample data. Bray-Curtis distances were
 49 used to compare diversity between samples. This revealed two large, weakly
 50 supported clusters, with many well-supported sub-clusters. Branches are labeled with
 51 an approximately unbiased p-value (red), a bootstrap probability value (green).
 52 Bootstrapping was performed at nboot = 1000.



53

54 **Figure S4:** ANOSIM test for correlations between soil classification and Bray-Curtis
 55 distances. Ranked dissimilarities plotted against soil order (A, C, E) and soil
 56 subgroup (B, D, F) for high country soils (A, B) sheep and beef soils (C, D) and dairy
 57 soils (E, F). Boxplots are drawn with widths proportional to the square roots of the
 58 number of observations in each land use and with 95% confidence intervals. Soil
 59 subgroup abbreviations: AOB = acidic orthic brown; MLP = mottled laminar pallic;
 60 TAB = typic acid brown; TIP = typic immature pallic; TOB = typic orthic brown;
 61 AAB = acidic-pedal allophanic brown; AFB = acidic firm brown; MFP = mottled
 62 fragic pallic; TFB = typic firm brown; FPP = fragic perch-gley pallic; MFR =
 63 mottled-acidic fluvial recent; PFB = pallic firm brown; POG = peaty orthic gley; TIP
 64 = typic immature pallic; TOG = typic orthic gley; WOR = weathered orthic recent.

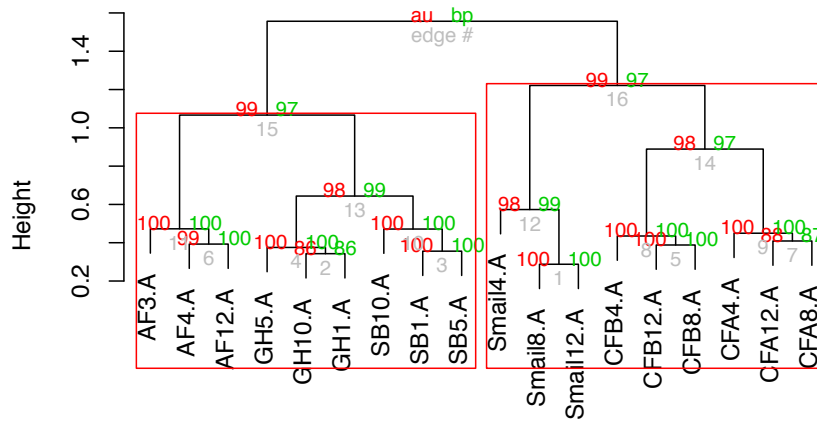
Cluster dendrogram with AU/BP values (%)



65

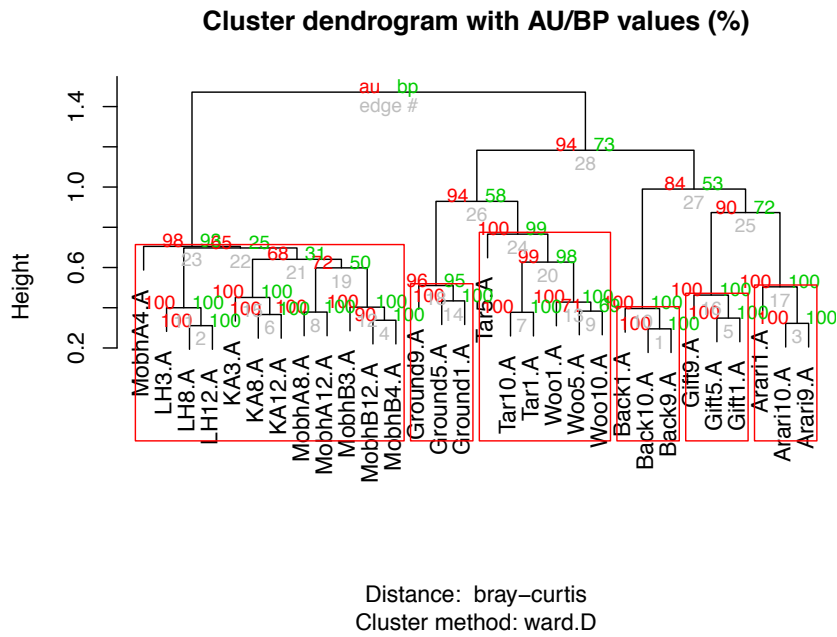
66 **Figure S5:** Hierarchical clustering for high country sample data. Bray-Curtis
67 distances were used to compare diversity between samples, revealing two large
68 clusters. Branches are labeled with an approximately unbiased p-value (red), a
69 bootstrap probability value (green). Bootstrapping was performed at nboot = 1000.

Cluster dendrogram with AU/BP values (%)



70

71 **Figure S6:** Hierarchical clustering for sheep and beef sample data. Bray-Curtis
72 distances were used to compare diversity between samples, revealing five clusters.
73 Branches are labeled with an approximately unbiased p-value (red), a bootstrap
74 probability value (green). Bootstrapping was performed at nboot = 1000.



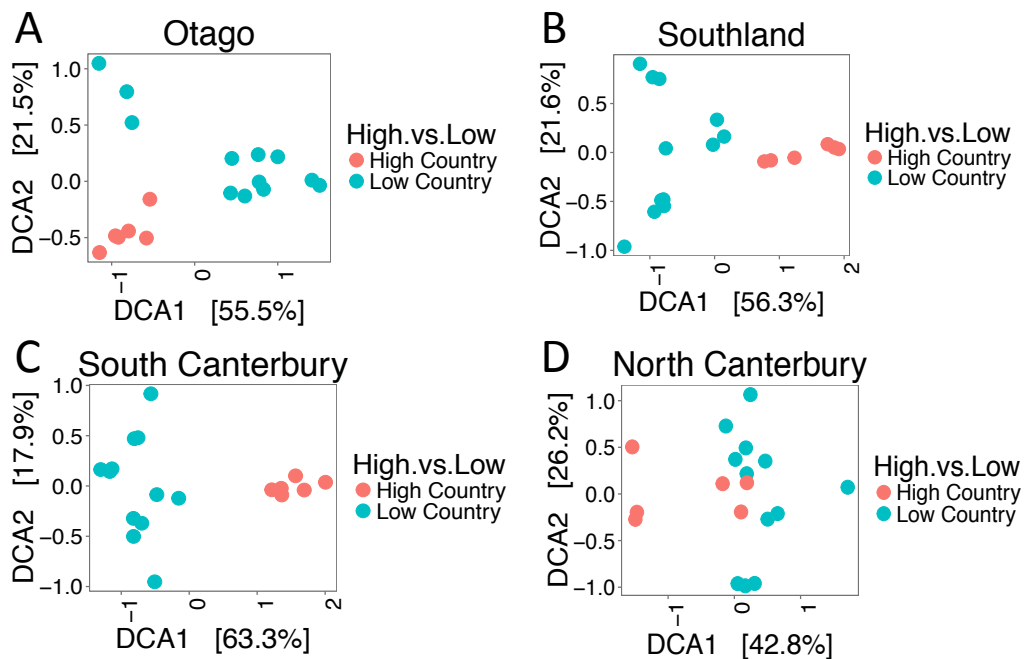
75

76 **Figure S7:** Hierarchical clustering for dairy sample data. Bray-Curtis distances were
 77 used to compare diversity between samples, revealing two large clusters. Branches are
 78 labeled with an approximately unbiased p-value (red), a bootstrap probability value
 79 (green). Bootstrapping was performed at nboot = 1000.

80

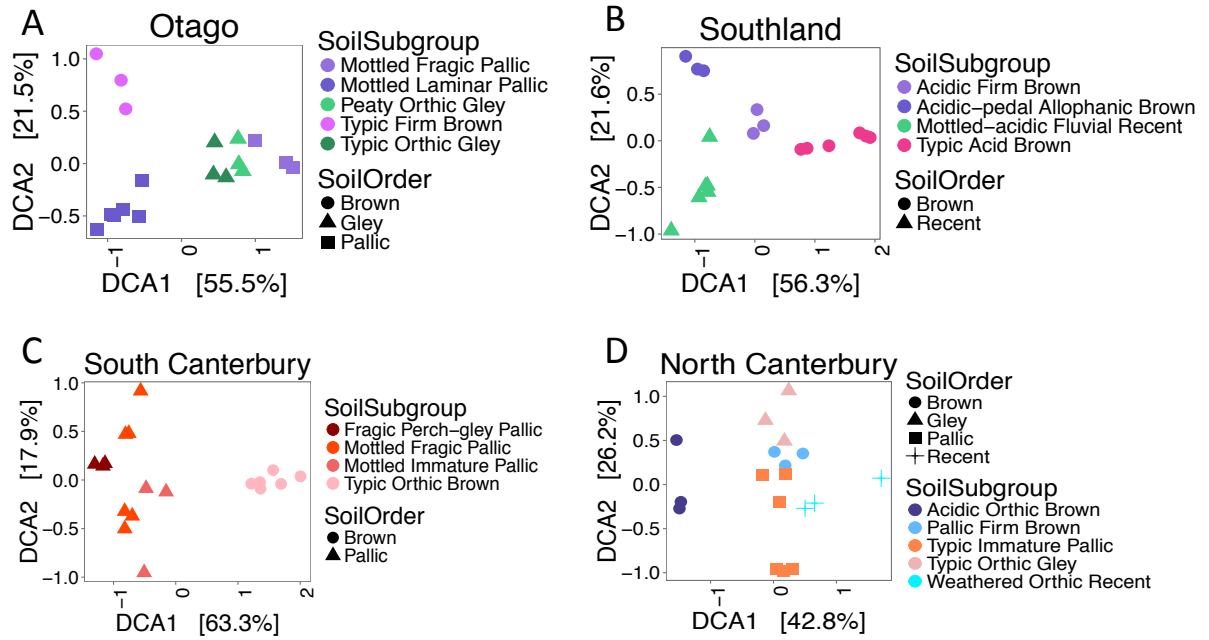
81

82



83

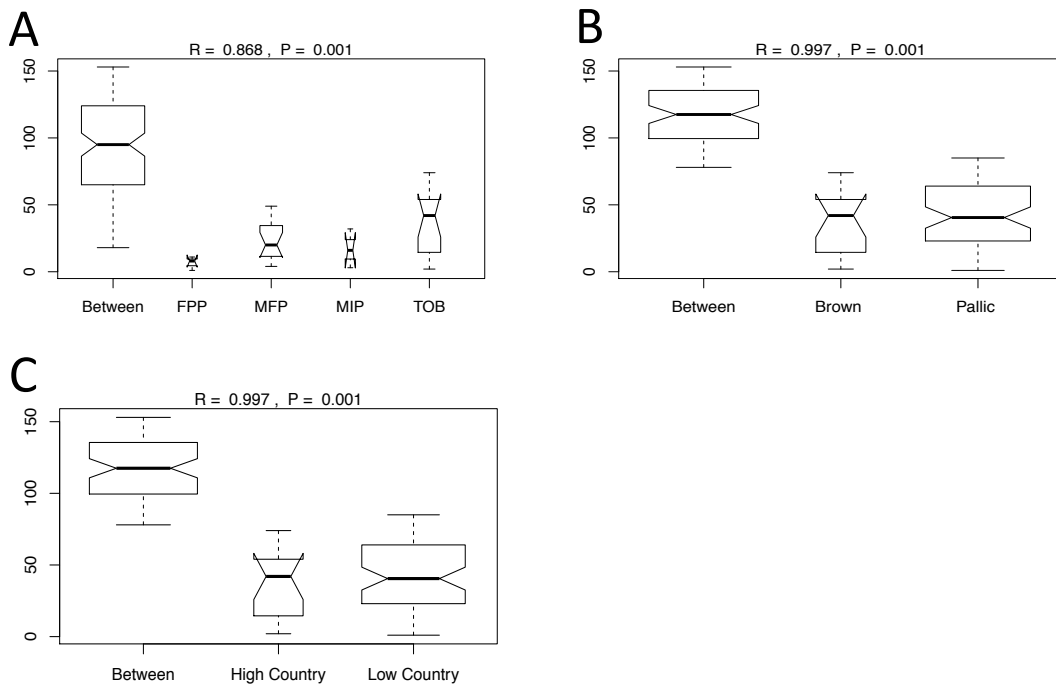
84 **Figure S8:** Ordinations of regional subgroups. DCA plots on a Bray-Curtis distance
 85 matrix showing the relationship between land use (simplified to low country versus
 86 high country) and and Otago soils (A), Southland soils (B), South Canterbury soils
 87 (C) and North Canterbury soils (D). Land use is indicated by color.



88

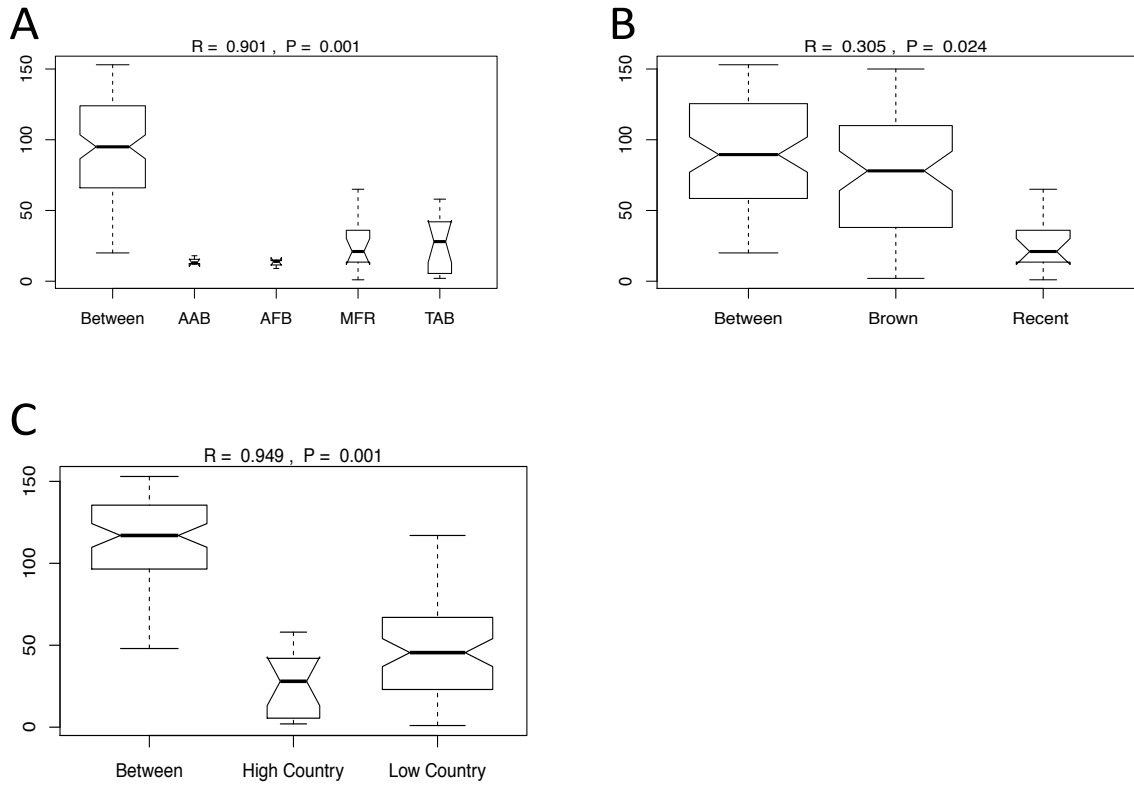
89 **Figure S9:** Ordinations of regional subgroups. DCA ordination based on Bray-Curtis
 90 showing relationships between soil classification and Otago soils (A), Southland soils
 91 (B), South Canterbury soils (C) and North Canterbury soils (D). Soil subgroup is
 92 indicated by color and soil order is indicated by shape.

93



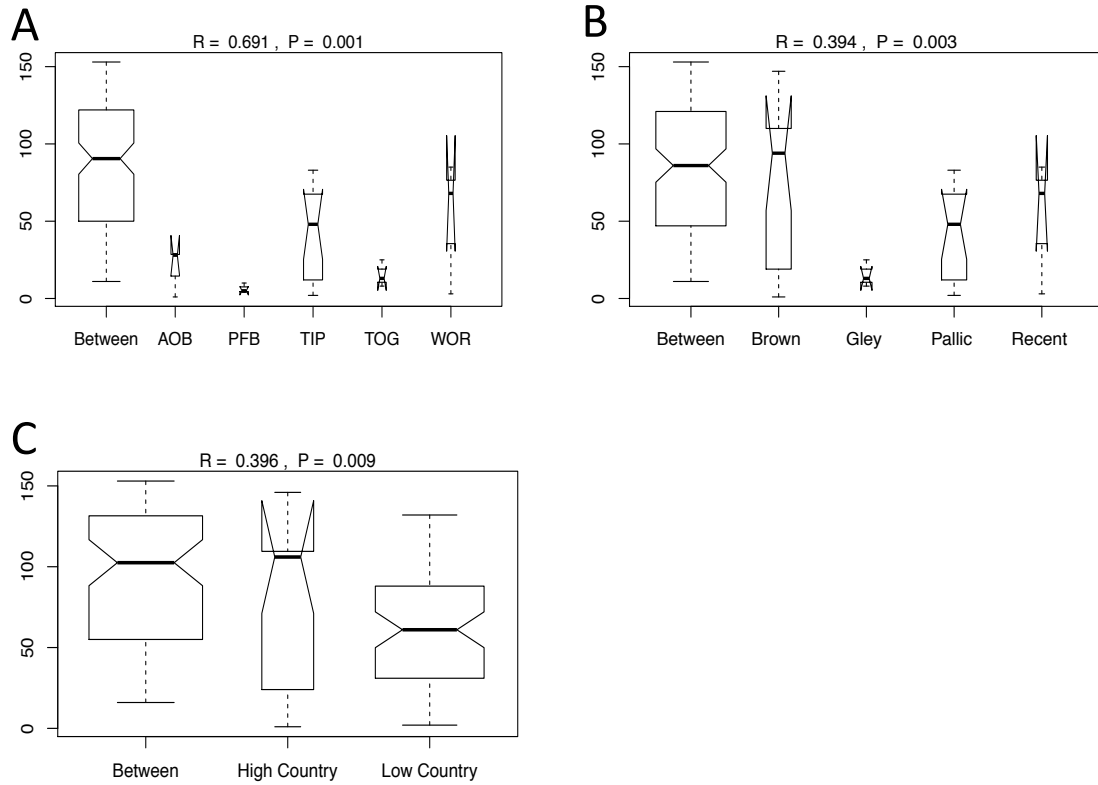
94

95 **Figure S10:** ANOSIM test for correlations between soil classification and land use
 96 and Bray-Curtis distances for South Canterbury soils. (A) Ranked dissimilarities
 97 plotted against soil subgroup, (B) soil order and (C) land use. Boxplots are drawn
 98 with widths proportional to the square roots of the number of observations in each
 99 land use and with 95% confidence intervals.



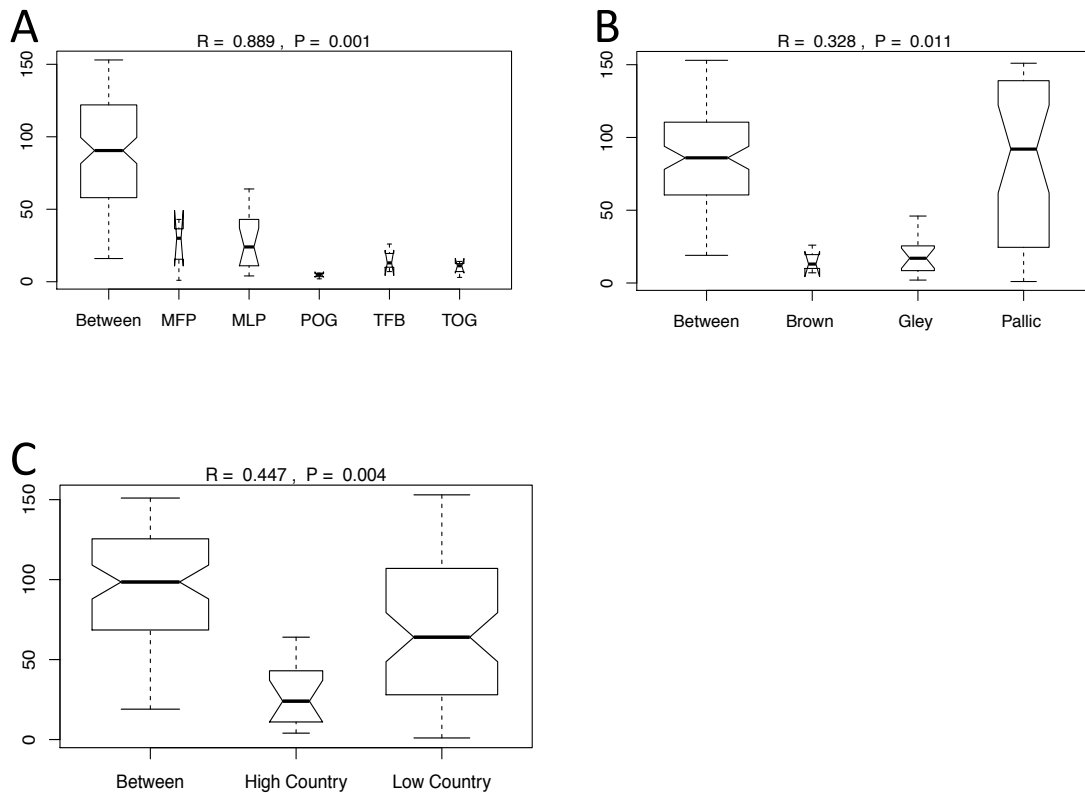
100

101 **Figure S11:** ANOSIM test for correlations between soil classification and land use
 102 and Bray-Curtis distances for Southland soils. (A) Ranked dissimilarities plotted
 103 against soil subgroup, (B) soil order and (C) land use. Boxplots are drawn with widths
 104 proportional to the square roots of the number of observations in each land use and
 105 with 95% confidence intervals. Soil subgroup abbreviations: TAB = typical acid brown;
 106 AAB = acidic-pedal allophanic brown; AFB = acidic firm brown; MFR = mottled-
 107 acidic fluvial recent
 108



109

110 **Figure S12:** ANOSIM test for correlations between soil taxonomy and land use and
 111 Bray-Curtis distances for North Canterbury soils. (A) Ranked dissimilarities plotted
 112 against soil subgroup, (B) soil order and (C) land use. Boxplots are drawn with widths
 113 proportional to the square roots of the number of observations in each land use and
 114 with 95% confidence intervals. Soil subgroup abbreviations: AOB = acidic orthic
 115 brown; PFB = pallic firm brown; TIP = typic immature pallic; TOG = typic orthic
 116 gley; WOR = weathered orthic recent



117

118 **Figure S13:** ANOSIM test for correlations between soil taxonomy and land use and
 119 Bray-Curtis distances for Otago soils. (A) Ranked dissimilarities plotted against soil
 120 subgroup, (B) soil order and (C) land use. Boxplots are drawn with widths
 121 proportional to the square roots of the number of observations in each land use and
 122 with 95% confidence intervals. Soil subgroup abbreviations: MFP = mottled fragic
 123 pallic; MLP = mottled laminar pallic; POG = peaty orthic gley; TFB = typic firm
 124 brown; TOG = typic orthic gley

125

126

127 **Table S2: Mantel Correlations between pH and Bray-Curtis distances for**

128 **within-region communities**

Region	Mantel R^2	p-value
North Canterbury	0.37	<0.002
South Canterbury	0.85	<0.001
Otago	0.48	<0.001
Southland	0.81	<0.001

129

130