Supporting Information

Peiffer et al. 10.1073/pnas.1302837110

SI Materials and Methods

Maize Germplasm, Microbiome Sample Collection, and Soil Sample Analysis. Twenty-seven diverse maize inbreds, all founder genotypes of the Nested Association Mapping panel, were selected to maximize genetic dissimilarity using previously established genotypic data (1). Seeds for each of the inbreds were obtained from a uniform stand grown at Muskgrave Research Station in Aurora, NY in 2009. In 2010, these lines were hand planted in a randomized complete block design in five field environments located in three states [University of Illinois, Crop Sciences Research and Education Center near Champaign-Urbana, IL (Well-Drained Drummer silty-clay loam soil); University of Missouri, South Farm near Columbia, MO (Well-Drained Mexico silt loam soil); Cornell University, Muskgrave Research Station near Aurora, NY (Well-Drained Honeoye silt loam soil); Cornell University, Ketola Organic Research Farm near Ithaca, NY (Well-Drained Erie Channery silt loam soil); and Willet Dairy near Lansing, NY (Well-Drained Lyons silt loam soil)]. Conventional culture practices were used, including ammonium nitrate-based fertilization, weed, and pest control in all locations except Ketola Research Farm, where an organic management regime was implemented, including manure-based fertilization and no pesticide or chemical weed control. The rhizosphere microbiota of all maize inbred plots, as well as bulk soil samples, were collected at their mean pollen shed, approximately 12 wk after planting. The last significant precipitation event occurring in all field environments was at least 3 d before the date of sample collection.

Within each field environment, plants were carefully removed from the soil using a drain spade. Roots of three random plants were sampled from the middle of each plot composed of between 12 and 25 plants (varying by environment) to avoid border effects potentially attributable to increased nutrient availability in the end plant of a plot. For each plant, a root segment of ∼5 cm in length and 0.5–3 mm in diameter was collected near the base of the plant, along with any adherent soil particles. All sample collection was performed in well-drained soils. However, no efforts were taken to collect or model data detailing covariation in soil moisture content within each of the microbiome samples. Variation in moisture content was assumed independent of the randomized and replicated plot design for genotypes within an environment. However, differences in soil moisture content, and thus adhesion, were likely a contributing factor to the variation in microbial diversity observed between environments. All root-adherent soil particles were less than 0.01 mm in diameter; most of these particles were also less than 0.002 mm. Bulk soil samples across each of the fields were also taken midrange between maize plots using a soil core sampler of 4-cm diameter and 20-cm depth. All samples were chilled on ice immediately following collection and stored at −80 °C before DNA extraction.

Soil analyses (Table S3) were performed by the Cornell University Nutrient Analysis Laboratory using standard operational procedures for measures of moisture content, extractable phosphorus and nitrate (using the Morgan test method), as well as potassium, calcium, magnesium, iron, manganese, zinc, and aluminum by an inductively coupled plasma atomic emission spectrometer. Buffer pH was measured using the Modified Mehlich buffer test and organic matter was discerned by loss on ignition.

DNA Extraction and Amplicon Generation. Total genomic DNA was isolated from the maize root tip and its associated soil (∼0.25 g) using the PowerSoil High-Throughput DNA Isolation Kit (Mo Bio Laboratories). The root and its loosely associated soil were placed into a 2-mL well of a 96-well plate for bead beating. Samples were homogenized using a bead beater (BioSpec; 2 min on high; note that this procedure gently scoured the root but did not pulverize it). It is noted that this protocol allows for the introduction of a small fraction of endophytic microbial communities scoured from the root epidermis; however, this fraction is reduced compared with the entire rhizosphere microbiome sample. Fifteen samples from Columbia, MO were used for the preliminary primer testing experiments. 16S rRNA genes were amplified using four different primer sets (27F-338R: AGAGTTTGATCCTGGCTCAG–TGCTGCCTCCCGTAGGAGT; 515F-806R: GTGCCAGCMGCCGCGGTAA–GGACTACHVGGGTWTCTAAT; 804F-1392R: AGATTAGATACCCDRGTAGTC–ACGGGCGGTGTGTRC; and 926F-1392R: AAACTYAAAKGAATTGACGG–ACGGGCGGTGTGTRC) (2), including barcodes and titanium adapters. For the full study, we used the 515F-806R primer pair. The PCR primers were constructed as follow: forward primer = 454 Titanium Lib-l Primer A/5-base barcode/forward 16S primer and reverse primer = 454 Titanium Lib-l Primer B/reverse 16S primer. All PCR reactions were carried out in triplicate 50-μL reactions with 1× of Easy-A buffer, 1.25 U Easy-A Taq, 0.2 μM of forward and reverse primers, 3.5 mM $MgCl_2$, 0.2 mM of dNTPs, and about 50 ng template DNA. Thermal cycling consisted of initial denaturation at 95 °C for 2 min, followed by 30 cycles of denaturation at 95 °C for 30 s, annealing at 53 °C for 20 s, and elongation at 72 °C for 60 s. Negative control samples were treated similarly with the exclusion of template DNA; these negative controls failed to produce visible PCR products. Following PCR, DNA amplicons were purified with Ampure magnetic purification beads (Agencourt) and quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen). Amplicons were then combined in equimolar ratios into a single tube with a final concentration of 12.5 ng/μL. Pyrosequencing was performed using Roche Titanium chemistry at the Department of Energy Joint Genome Institute.

Analysis of 16S rRNA Gene Sequence. Sequences were analyzed using the QIIME software package (Quantitative Insights into Microbial Ecology) using default parameters for each step (3). Sequences were removed if their lengths were shorter than 200 nt, their average quality score was <25, and they contained ambiguous bases, primer mismatches, homopolymer runs in excess of six bases or error in barcodes. Filtering of noisy sequences, chimera checking and operational taxonomic unit (OTU) picking was performed using the usearch series of scripts. De novo and reference-based chimera checking was performed and sequences that were characterized as chimeric by both methods were removed. More than 3.8-million quality-filtered reads were obtained for the samples, an average of 8,315 reads per sample (min = 2,225, max = 22,346). Sequences were chimera-checked and clustered into OTUs using Otupipe (4) and a minimum pair-wise identity of 97%. Each cluster was represented by its most abundant sequence. Representative OTUs sequences were then aligned to the Greengenes database (5) using the PyNAST algorithm (minimum percent identity was set at 80%) (6). A phylogenetic tree was built using FastTree (7). Taxonomy was subsequently assigned to each representative OTUs using the Greengenes database classifier with a minimum support threshold of 80% (5, 8).

Statistical Analyses. We used custom R scripts executed using R v2.13.2 (9) to calculate the percentage of classifiable reads. The

median proportion of Greengenes classifiable reads obtained from each primer set in the pilot experiment (Table S1) was calculated from 100 bootstrap samples of the surveyed microbiome extractions stratified by maize inbreds and bulk soil to maintain balance among these factors. Bootstrap sampling of microbiome extractions with replacement ensured equal representation of each inbred and bulk soil, and also provided a 95% confidence interval estimating the precision of estimates derived from the data. Given the lack of normality noted in the distributions of many populations tested, we used the function "aovp" from the R package lmPerm v1.1.2 (10) to discern variation in the proportion of classifiable reads between each primer set by permutation testing. Reported variances explained by each factor reflect the proportion of variance explained by that factor after accounting for the remaining factors and are calculated from the marginal sums of squares. The 95% confidence interval for variance explained was derived from the resulting distribution of variance estimates after fitting multiple regression models to each of the 100 bootstrap samplings of the data. A minimum of 5,000 permutations of the data were used to construct null distributions for each of the bootstrap samplings of the raw data in inferring significance. The reported significance values reflect the most conservative estimate obtained from the 100 bootstrap samplings. Significances for all pair-wise comparisons among the primer sets, soil, and maize inbreds in the pilot experiment were adjusted for multiple comparisons by Bonferroni correction.

Rarefaction was performed using QIIME to discern levels of OTU richness, Chao-1 diversity, and whole-tree phylogenetic diversity with respect to sequence depth (3). Following rarefaction, median abundances for each microbiome extraction were calculated at a level of 2,080 pyrosequence reads. Given an inability to accurately extrapolate OTU abundances beyond a microbiome extraction's maximum read depth, 2,080 reads was selected as a balance between removing microbiome extractions that did not possess this minimum and seeking to attain as many reads, and thus sensitivity, as possible in the included microbiome extractions. To address the unbalanced design resulting from removing extractions not possessing this minimum read depth, the microbiome extractions were bootstrapped for 100 samplings stratified by field environment, soil, and maize inbred. Permutation-based multiple regression analyses were performed in a manner similar to that implemented in discerning variation in the proportion of classifiable reads for partitioning variation in α-diversity, as measured by species richness among extractions. Reported variances in α-diversity explained by each biological factor, such as field environment, genotype, and genotype-within-field environment, reflect the proportion of variance explained after accounting for technical factors of amplification batch and pyrosequencing run and rarefying to the common depth of 2,080 reads. The reported 95% bootstrap confidence intervals were derived by reporting the distribution of variance explained upon sampling from the extractions.

To calculate β-diversity, unweighted and weighted UniFrac distance metrics were calculated and used to construct distance matrices using QIIME (3). Subsequently, the entries composing these matrices were bootstrapped for 100 samplings stratified by field environment, bulk soil, and maize inbred. The function "capscale" of the R package vegan v2.0.2 (11) was used in calculation of partial constrained principal coordinate analyses. The proportion of the total variance explained by each factor was calculated after conditioning on amplification batch, pyrosequencing run, and the remaining factors, and constraining variation to the factor of interest. The 95% confidence intervals for this variation explained were derived from the bootstrap samplings. Significances of factors within the model were calculated using vegan's permutation testing function "permutest" for constrained analysis of principal coordinates with 5,000 permutations (11). Comparisons of levels of within factor multivariate dispersion were performed using vegan's implementation of PERMDISP (12).

All comparisons of relative abundance of individual OTU as well as comparisons among soil characteristics were performed by permutation testing using the lmPerm package (10). Reported significance values are adjusted by Bonferroni correction. Normalization of the soil characteristics data and construction of the correlation matrix was performed using routines in the R base package (9). Estimations of the relatedness matrix among maize lines were performed using percent identity by state (12) as well as genotype data from the first-generation maize hapmap (13). Soil characteristic and maize kinship matrices were bootstrapped for 100 samplings stratified by field environment and maize inbred and performed using vegan's implementation of the Mantel test (11, 14).

1. Yu J, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. *Genetics* 178(1):539–551.
2. Kuczynski J, et al. (2012) Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet* 13(1):47–58.
3. Caporaso JG, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7(5):335–336.
4. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27(16):2194–2200.
5. McDonald D, et al. (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 6(3):610–618.
6. Caporaso JG, et al. (2010) PyNAST: A flexible tool for aligning sequences to a template alignment. *Bioinformatics* 26(2):266–267.
7. Price MN, Dehal PS, Arkin AP (2009) FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 26(7):1641–1650.
8. Werner JJ, et al. (2012) Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. *ISME J* 6(1):94–103.
9. R Development Core Team (2011) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria).
10. Wheeler RE (2010) lmPerm: Permutation tests for linear models. R package version 1.1-2. http://CRAN.R-project.org/package=lmPerm). Accessed August 5, 2012.
11. Oksanen J, et al. (2011) vegan: Community ecology package. R package version 2.0-2. http://CRAN.R-project.org/package=vegan). Accessed July 11, 2012.
12. Hardy OJ, Vekemans X (2002) spagedi: A versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* 2(4): 618–620.
13. Gore MA, et al. (2009) A first-generation haplotype map of maize. *Science* 326(5956): 1115–1117.
14. Legendre P, Legendre L (1998) *Numerical Ecology* (Elsevier, Amsterdam), pp 853.

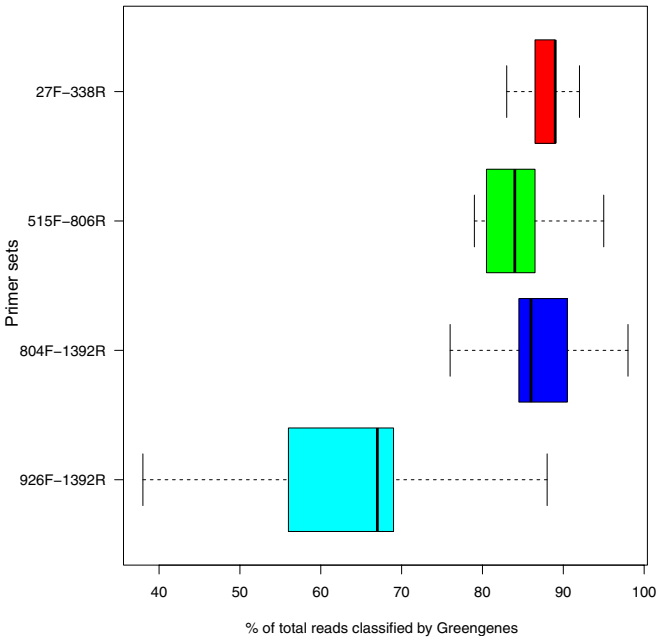

Fig. S1. Primer pair effects on proportion of classifiable sequence reads. The boxplots show the proportion of total sequence reads that matched the Greengenes database (*SI Materials and Methods*) and were obtained using the different primer sets (data are for samples shown in Fig. 1).

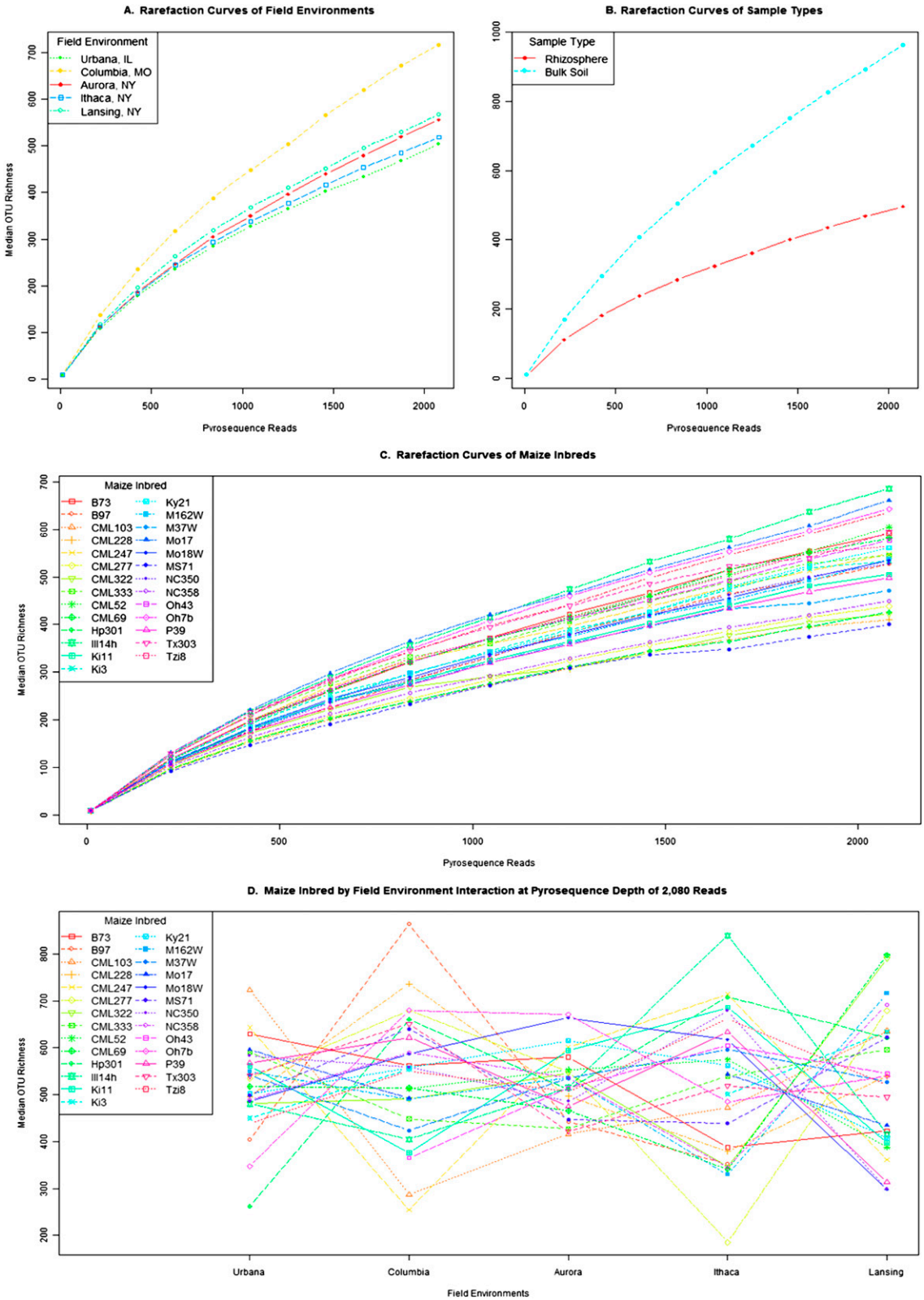

Fig. S2. Rarefaction curves for the full study. OTU counts as a function of sequence depth shown for (*A*) field environment, (*B*) sample type (bulk soil vs. the maize rhizosphere), (*C*) maize inbreds, and (*D*) maize inbreds by each soil.

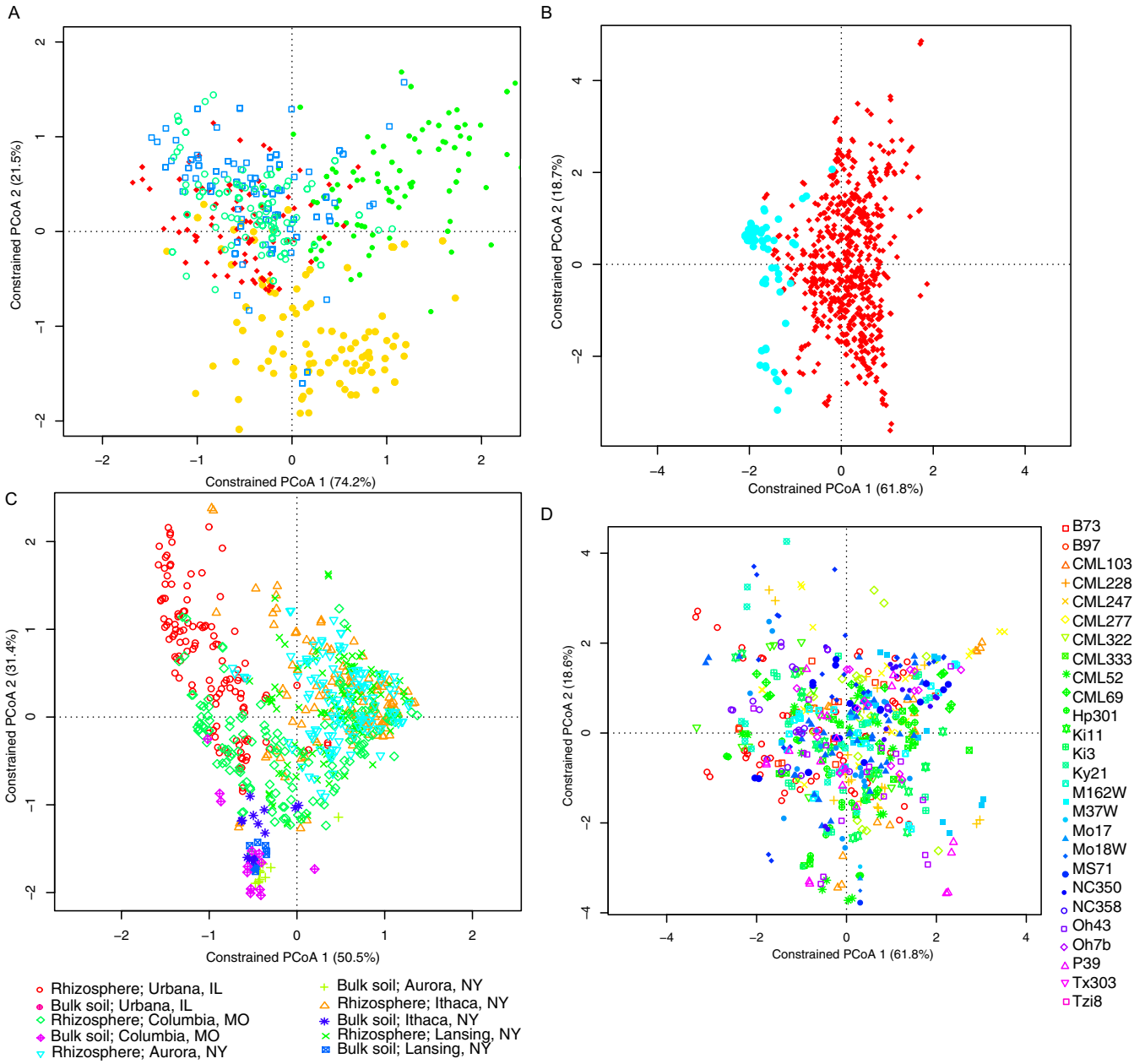

Fig. S3. Factors influencing rhizosphere and soil microbiome β-diversity. Variation in weighted UniFrac dispersion by (*A*) field environment ($P < 5.00E-03$); (*B*) sample type ($P < 5.00E-03$); (*C*) sample type within all field environments ($P < 5.00E-02$); (*D*) maize inbreds ($P < 5.00E-02$). Note that the percent variation explained by the principal coordinates (PCs) from this canonical analysis of principal coordinates (CAP) analysis is indicated on the axes and refers to the fraction of the total variance (indicated in the main text) explained by field (*A*), sample type (*B*), sample type by field (*C*), and maize inbred (*D*).

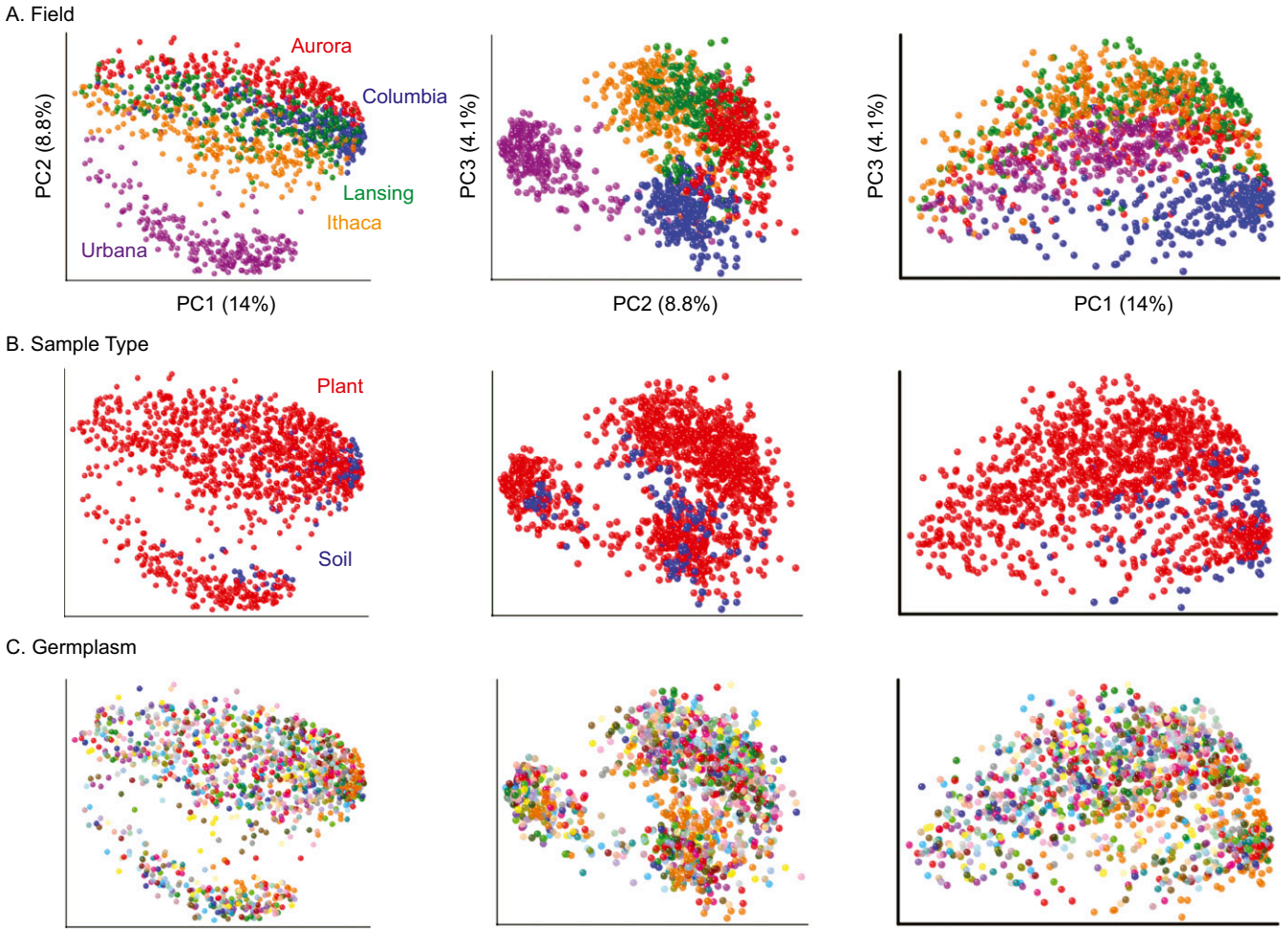

Fig. S4. Factors influencing rhizosphere and soil microbiome β-diversity. (*A*–*C*) Unconstrained principal coordinate analysis of weighted UniFrac distances. The percent variation explained by the PCs is indicated on the axes.

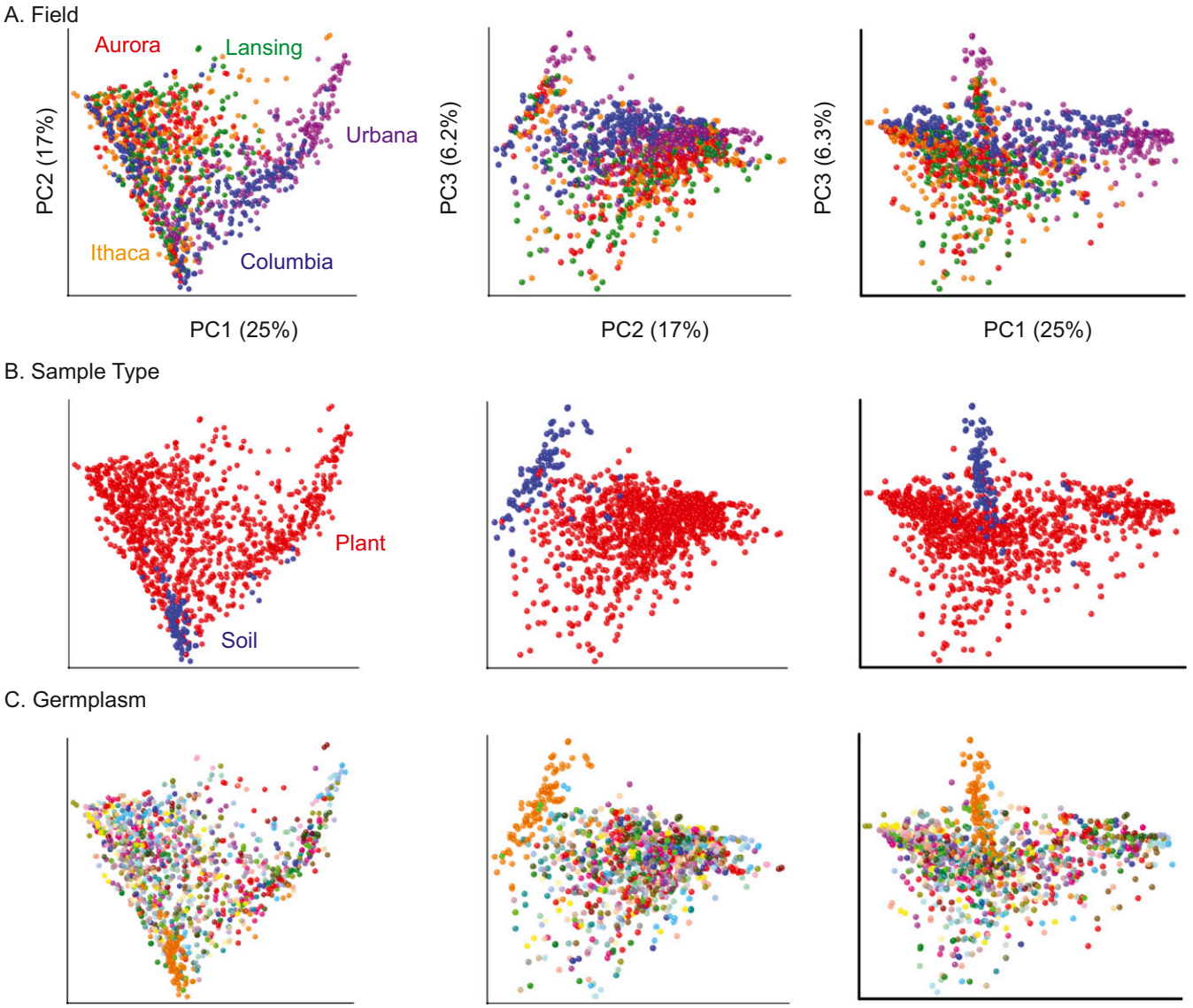

Fig. S5. Factors influencing rhizosphere and soil microbiome β-diversity. (*A*–*C*) Unconstrained principal coordinate analysis of unweighted UniFrac distances. The percent variation explained by the PCs is indicated on the axes.

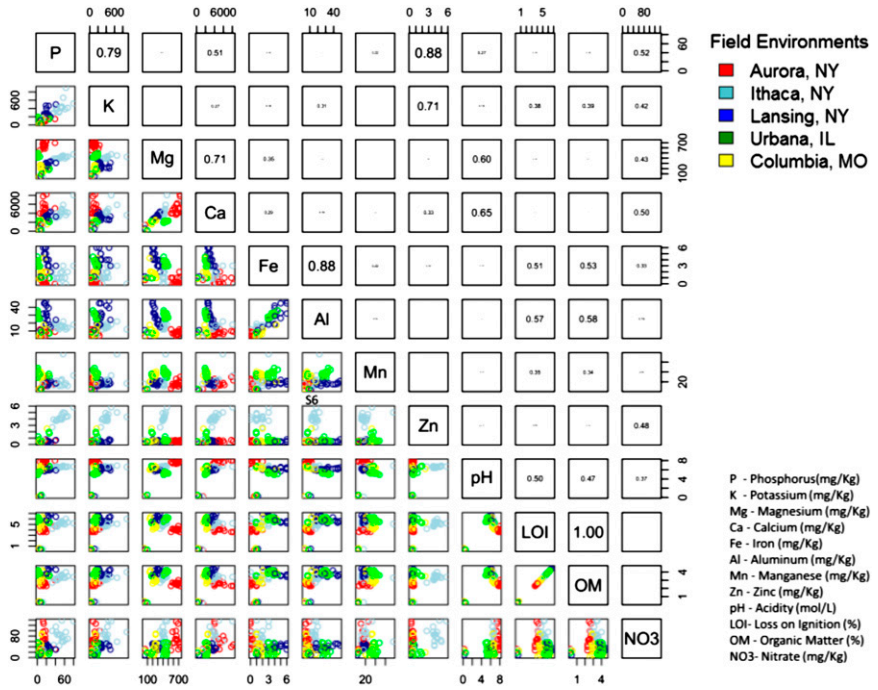

Fig. S6. Soil physiochemical properties. Scatterplots detailing relationships of physiochemical soil characteristics and field environment samples.

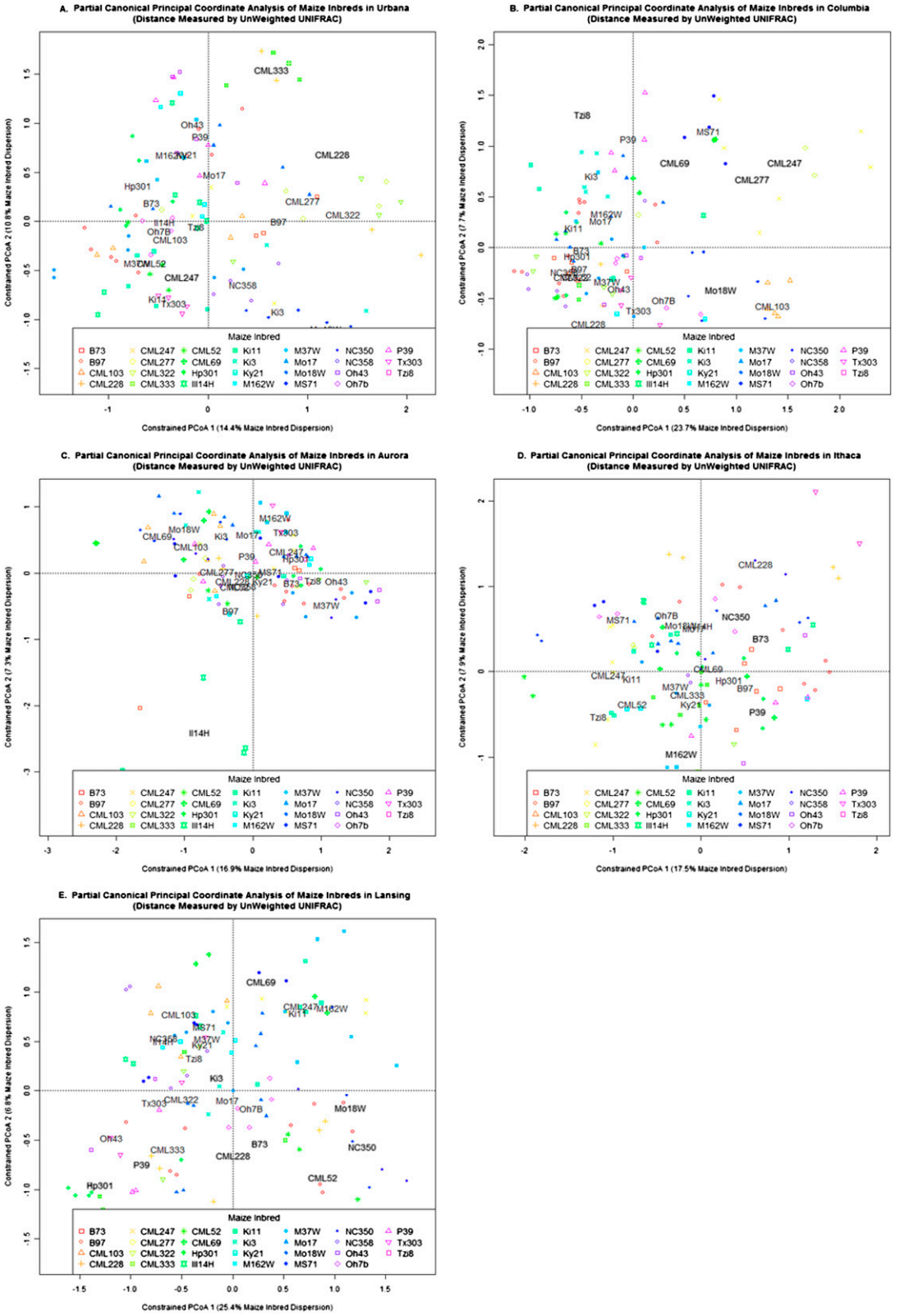

Fig. S7. Constrained principal coordinate analysis of unweighted UniFrac by field environment. Differences among maize inbreds within the different fields: (*A*) Urbana, IL; (*B*) Columbia, MO; (*C*) Aurora, NY; (*D*) Ithaca, NY; (*E*) Lansing, NY.

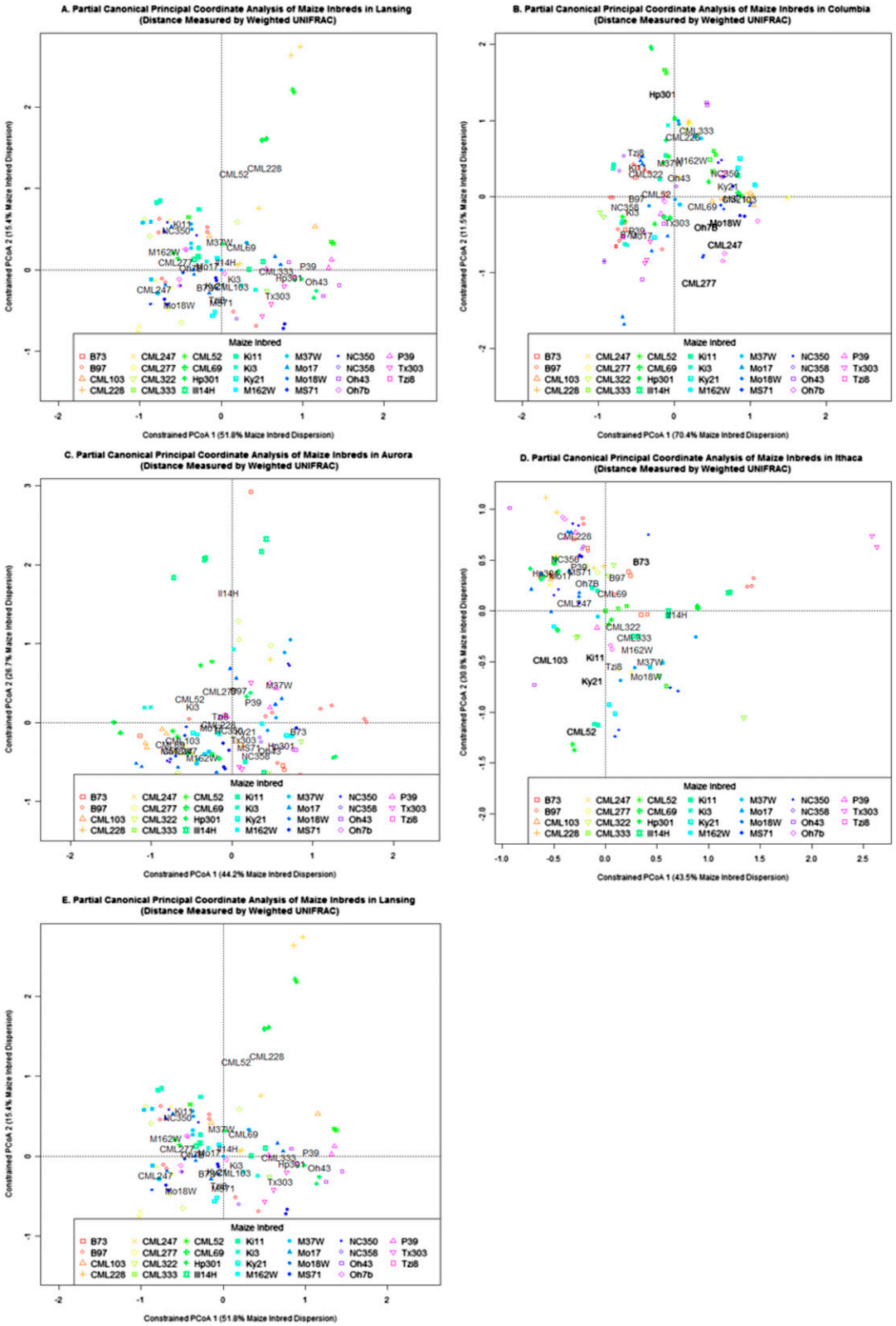

Fig. S8. Constrained principal coordinate analysis of weighted UniFrac by field environment. Differences among maize inbreds within the different fields: (*A*) Urbana, IL; (*B*) Columbia, MO; (*C*) Aurora, NY; (*D*) Ithaca, NY; (*E*) Lansing, NY.

Table S1. Summary of the 16S rRNA gene sequences obtained from the primer test experiment

Primer	Maize inbred	Sample type	Total pyrosequence reads	Pyrosequence reads (without singletons)	Greengenes classifiable reads
27F-338R (V1-V2)	B73	Rhizosphere	12,311	10,324	9,108
	B73	Rhizosphere	8,923	7,453	6,665
	B73	Rhizosphere	10,582	9,123	8,074
	B73	Rhizosphere	9,606	7,672	6,850
	Bulk soil	Bulk soil	13,859	9,804	8,183
	Bulk soil	Bulk soil	5,748	4,034	3,349
	Bulk soil	Bulk soil	12,957	10,143	9,039
	Bulk soil	Bulk soil	12,247	9,248	7,676
	Ill14h	Rhizosphere	10,501	9,035	8,532
	Ill14h	Rhizosphere	10,068	9,040	8,727
	Ill14h	Rhizosphere	12,845	10,401	9,608
	Mo17	Rhizosphere	10,411	8,154	7,077
	Mo17	Rhizosphere	11,704	8,819	7,549
	Mo17	Rhizosphere	17,217	13,941	12,419
	Mo17	Rhizosphere	9,436	7,340	6,380
515F-806R (V3-V4)	B73	Rhizosphere	17,113	15,227	12,586
	B73	Rhizosphere	37,551	34,684	29,009
	B73	Rhizosphere	31,550	29,991	25,256
	B73	Rhizosphere	29,148	24,828	21,489
	Bulk soil	Bulk soil	44,370	34,949	28,481
	Bulk soil	Bulk soil	13,809	10,880	8,693
	Bulk soil	Bulk soil	17,589	14,636	12,555
	Bulk soil	Bulk soil	16,239	13,637	10,774
	Ill14h	Rhizosphere	18,655	16,987	15,769
	Ill14h	Rhizosphere	13,499	12,648	12,012
	Ill14h	Rhizosphere	20,298	17,686	15,904
	Mo17	Rhizosphere	12,134	10,792	8,674
	Mo17	Rhizosphere	26,637	22,771	17,972
	Mo17	Rhizosphere	12,311	10,324	9,108
	Mo17	Rhizosphere	8,923	7,453	6,665
926F-1392R (V5-V8)	B73	Rhizosphere	4,791	3,640	2,034
	B73	Rhizosphere	5,220	4,148	1,944
	B73	Rhizosphere	4,552	3,856	1,451
	B73	Rhizosphere	5,693	4,471	3,085
	Bulk soil	Bulk soil	4,394	2,629	1,764
	Bulk soil	Bulk soil	4,492	3,081	2,092
	Bulk soil	Bulk soil	4,657	2,780	1,895
	Bulk soil	Bulk soil	4,493	2,956	2,124
	Ill14h	Rhizosphere	3,741	3,202	2,814
	Ill14h	Rhizosphere	4,489	4,121	3,900
	Mo17	Rhizosphere	5,386	3,883	2,520
	Mo17	Rhizosphere	4,780	3,337	1,952
	Mo17	Rhizosphere	6,732	5,403	2,861
804F-1392R (V6-V8)	B73	Rhizosphere	1,463	1,144	984
	B73	Rhizosphere	481	384	336
	B73	Rhizosphere	458	370	329
	B73	Rhizosphere	513	423	390
	Bulk soil	Bulk soil	625	429	305
	Bulk soil	Bulk soil	556	395	299
	Bulk soil	Bulk soil	1,378	1,031	879
	Bulk soil	Bulk soil	1,815	1,239	937
	Ill14h	Rhizosphere	664	583	556
	Ill14h	Rhizosphere	509	481	469
	Ill14h	Rhizosphere	405	342	320
	Mo17	Rhizosphere	426	324	277
	Mo17	Rhizosphere	572	428	358
	Mo17	Rhizosphere	558	440	380
	Mo17	Rhizosphere	455	339	290

Table S2. Summary of three measures of α-diversity

Factor	Level	Chao-1 (10 reads)	PD (10 reads)	Species richness (10 reads)	Chao-1 (838 reads)	PD (838 reads)	Species richness (838 reads)	Chao-1 (1,459 reads)	PD (1,459 reads)	Species richness (1,459 reads)	Chao-1 (2,080 reads)	PD (2080 reads)	Species richness (2,080 reads)
Type	Bulk Soil	48	3	9	1,310	47	483	1,687	62	726	1,973	73	929
	Rhizosphere	33	2	8	810	28	299	1,032	36	425	1,209	42	532
Field environment	Aurora	33	2	8	846	30	304	1,088	38	439	1,271	45	555
	Columbia	39	2	9	1,090	39	387	1,405	50	565	1,646	59	716
	Ithaca	33	2	8	716	28	294	920	35	416	1,075	41	517
	Lansing	35	2	8	887	29	318	1,110	37	451	1,297	44	567
	Urbana	32	2	8	757	25	285	958	32	402	1,134	38	502
Maize Inbred	B73	39	2	9	857	32	343	1,121	40	485	1,289	46	593
	B97	35	2	8	943	33	342	1,221	43	498	1,445	51	634
	CML103	25	2	8	643	22	232	850	29	336	937	32	399
	CML228	34	2	8	832	29	295	1,052	37	424	1,260	44	533
	CML247	27	2	8	633	22	238	816	29	343	917	33	409
	CML277	31	2	8	845	28	298	1,038	36	416	1,242	43	537
	CML322	30	2	8	781	27	274	998	35	398	1,179	41	498
	CML333	33	2	8	903	31	327	1,120	39	458	1,277	45	564
	CML52	36	2	9	848	30	320	1,062	38	460	1,317	47	605
	CML69	35	2	8	871	30	320	1,136	40	467	1,333	47	591
	Hp301	32	2	8	792	28	289	1,039	37	420	1,233	44	535
	Il14H	36	2	9	950	33	356	1,268	45	532	1,527	54	685
	Ki11	30	2	8	856	28	284	1,077	37	418	1,296	43	527
	Ki3	32	2	8	732	25	279	969	33	403	1,116	39	504
	Ky21	34	2	8	886	30	320	1,131	38	451	1,367	46	577
	M162W	32	2	8	669	25	268	771	29	341	922	34	423
	M37W	35	2	8	876	30	324	1,067	36	438	1,266	43	548
	Mo17	39	2	9	987	34	364	1,233	43	515	1,446	51	661
	Mo18W	32	2	8	618	22	238	812	29	344	955	34	426
	MS71	29	2	8	686	24	256	878	31	363	932	32	396
	NC350	34	2	8	747	26	275	966	34	396	1,104	38	470
	NC358	36	2	8	886	32	332	1,059	39	449	1,176	44	546
	Oh43	34	2	9	885	30	320	1,131	39	458	1,328	46	581
	Oh7B	37	2	8	908	32	348	1,178	43	509	1,392	50	642
	P39	27	2	8	682	24	245	897	32	357	1,023	36	437
	Tx303	31	2	8	798	29	297	1,045	37	427	1,274	46	560
	Tzi8	31	2	8	743	26	282	1,005	36	423	1,183	42	526

Chao-1 is an estimator of total species richness. It infers the abundance of unsampled diversity present within the community as a function of the abundance of singleton and doubleton species. This estimate is then added to the observed species richness. PD or phylogenetic diversity is a measure of biodiversity that incorporates phylogenetic differences between species. In this approach related individuals increase estimates of biodiversity less than unrelated individuals. Species richness is a measure of the observed number of unique OTU characterized at a given rarefaction level of reads.

Table S3. Physiochemical properties of the soils from the five field environments

Location	Sample no.	Moisture (%)	P (mg/kg)	K (mg/kg)	Mg (mg/kg)	Ca (mg/kg)	Fe (mg/kg)	Al (mg/kg)	Mn (mg/kg)	Zn (mg/kg)	pH (mg/kg)	LOI (%)	OM (%)	NO_3 (mg/kg)
Aurora, NY	1	2.21	18	95	715	5,230	2	12	18	0.3	7.8	4.0	2.6	56
	2	2.23	14	145	690	4,030	<det	5	11	0.3	7.4	4.2	2.7	126
	3	2.15	17	210	585	4,720	1	7	12	0.2	7.6	4.1	2.7	53
	4	2.12	18	235	650	3,670	<det	4	14	0.5	7.3	4.1	2.6	99
	5	2.17	17	155	660	5,940	<det	6	18	0.6	7.6	4.2	2.7	80
	6	1.86	13	80	730	7,770	1	10	30	0.7	8.0	3.6	2.3	40
	7	2.05	13	120	585	3,990	<det	7	16	0.6	7.4	4.2	2.7	96
	8	1.92	39	145	695	5,100	<det	9	15	0.6	8.0	4.7	3.0	31
	9	2.39	14	125	665	4,120	<det	6	14	0.3	7.4	4.1	2.6	68
	10	1.89	12	65	675	3,990	<det	5	16	0.5	7.5	3.9	2.5	66
	11	2.04	6	45	555	3,990	1	12	7	0.6	7.7	3.6	2.3	27
	12	2.14	12	95	675	6,230	1	9	14	0.3	7.8	3.8	2.4	65
	13	2.01	20	130	725	8,230	<det	9	23	0.5	7.9	3.6	2.3	48
	14	1.86	13	80	655	4,000	<det	9	26	0.3	7.4	3.9	2.5	134
Lansing, NY	1	2.27	36	295	390	3,410	1	19	13	3.9	6.6	4.9	3.2	47
	2	2.28	48	330	405	3,280	1	16	12	4.1	6.4	5.0	3.2	86
	3	2.02	41	430	395	2,930	3	20	16	3.7	6.4	4.7	3.1	115
	4	1.99	33	270	345	3,430	2	17	18	2.7	6.7	4.2	2.7	99
	5	2.26	54	410	430	4,560	2	17	24	4.2	6.7	5.5	3.6	75
	6	2.19	80	520	475	8,040	3	18	33	5.8	6.7	5.2	3.4	135
	7	2.19	48	305	390	3,590	<det	12	13	3.9	6.6	5.3	3.5	44
	8	2.31	54	450	440	4,980	1	15	25	4.7	6.6	5.3	3.5	138
	9	1.94	42	365	390	4,040	1	14	22	3.7	6.8	4.8	3.1	95
	10	2.30	54	420	415	4,140	2	17	12	3.9	6.6	5.6	3.7	61
	11	2.12	67	415	430	4,160	2	14	12	4.4	6.7	5.4	3.5	70
	12	2.24	63	900	425	3,600	<det	12	17	4.3	6.6	6.0	4.0	95
	13	2.12	57	660	450	5,870	3	22	76	5.0	6.9	5.1	3.3	59
	14	1.76	44	375	390	3,180	1	13	17	3.8	6.6	4.6	3.0	85
	15	2.04	61	385	395	4,340	1	15	14	4.9	6.7	5.0	3.3	65
Ithaca, NY	1	1.70	19	270	215	2,840	5	33	18	0.3	6.6	5.5	3.6	44
	2	1.75	15	270	195	2,460	5	45	16	0.5	6.4	5.3	3.5	51
	3	1.71	20	405	215	2,620	4	42	15	0.4	6.3	5.7	3.8	52
	4	1.76	25	465	240	2,660	4	39	15	0.5	6.2	6.3	4.2	43
	5	1.65	19	470	205	2,490	6	48	22	0.8	6.5	5.6	3.7	45
	6	1.68	17	230	200	2,680	6	45	19	0.4	6.1	5.5	3.6	52
	7	2.17	23	215	290	3,720	3	24	12	0.6	6.2	6.6	4.4	26
	8	1.84	21	295	245	3,000	4	31	15	0.3	6.1	5.7	3.8	36
	9	1.79	13	165	230	2,870	6	32	15	0.5	6.3	5.2	3.4	30
	10	1.76	16	265	245	3,080	5	26	24	0.4	5.7	5.5	3.6	41
	11	2.33	26	235	325	4,120	3	18	19	0.5	5.9	6.4	4.3	37
	12	2.24	41	500	310	3,270	3	24	17	0.7	5.9	6.4	4.3	30
	13	2.01	19	255	310	3,460	4	18	19	0.6	6.1	5.8	3.8	53
	14	2.05	19	170	290	3,460	4	19	15	0.4	5.8	5.6	3.7	28
	15	2.48	29	200	340	4,360	2	15	22	0.3	6.0	6.9	4.6	34
Columbia, MO	1	1.75	7	104	203	2,187	2	6	36	0.5	6.8	4.1	2.6	5
	2	1.73	9	108	217	2,322	2	5	38	0.6	6.5	4.4	2.9	37
	3	2.1	7	116	266	2,484	1	6	35	0.5	6.5	4.7	3.0	23
	4	1.96	9	104	233	2,490	1	5	27	0.5	6.9	4.5	2.9	10
	5	1.66	11	123	155	2,089	2	12	21	0.4	6.2	6.3	4.2	10
	6	2.21	7	122	270	2,462	1	6	41	0.5	6.4	4.5	2.9	39
	7	1.61	14	183	143	1,810	4	18	20	0.5	6.1	6.0	4.0	12
	8	2.07	10	131	256	2,482	1	5	32	0.5	6.5	4.5	2.9	85
	9	2.14	10	127	258	2,514	1	5	34	0.6	6.5	4.5	2.9	89
	10	2.03	8	131	247	2,482	1	6	38	0.5	6.6	4.6	3.0	53
	11	2.02	8	130	245	2,518	1	6	35	0.5	6.7	4.5	2.9	30
	12	2.13	7	115	267	2,407	1	6	35	0.4	6.6	4.5	2.9	52
	13	1.95	8	133	244	2,471	2	6	34	0.5	6.7	4.4	2.9	37
	14	2.07	9	154	254	2,468	1	6	36	0.5	6.7	4.3	2.8	49
	15	1.32	21	164	180	1,523	2	12	33	2.5	6.2	5.4	3.5	19

Table S3. Cont.

Location	Sample no.	Moisture (%)	P (mg/kg)	K (mg/kg)	Mg (mg/kg)	Ca (mg/kg)	Fe (mg/kg)	Al (mg/kg)	Mn (mg/kg)	Zn (mg/kg)	pH (mg/kg)	LOI (%)	OM (%)	NO_3 (mg/kg)
Urbana, IL	1	2.19	1	103	416	1,924	4	31	41	0.6	5.2	5.5	0.6	21
	2	2.17	1	115	414	1,892	4	33	48	0.6	4.9	5.5	0.6	55
	3	2.22	1	78	413	1,929	5	36	44	0.6	5.1	5.6	0.6	22
	4	2.39	1	167	455	2,247	4	29	30	0.7	5.2	6.3	0.7	27
	5	2.42	2	195	469	2,275	3	27	37	0.7	5.3	6.3	0.7	41
	6	2.44	1	195	500	2,332	3	27	36	0.7	5.2	6.2	0.7	54
	7	2.39	1	169	452	2,211	3	26	29	0.5	5.5	6.0	0.5	3
	8	2.3	2	170	449	2,176	4	29	39	0.7	5.2	6.0	0.7	39
	9	2.25	2	193	458	2,166	3	25	37	0.6	5.2	6.1	0.6	56
	10	2.37	2	188	439	2,117	3	28	34	1.3	5.4	6.1	1.3	13
	11	2.41	1	148	478	2,333	3	31	29	0.5	5.3	6.3	0.5	20
	12	2.37	1	119	466	2,146	4	33	41	0.6	5.3	5.9	0.6	9
	13	2.16	1	127	411	1,960	4	32	38	0.6	5.3	5.7	0.6	5
	14	1.35	27	155	196	2,092	1	7	31	1.8	6.7	5.4	1.8	21
	15	1.05	26	145	178	2,279	1	7	36	2.0	6.9	5.3	2.0	13

Al, aluminum measured in milligrams per kilogram; Ca, calcium measured in milligrams per kilogram; Fe, iron measured in milligrams per kilogram; K, potassium measured in milligrams per kilogram; LOI, percentage of mass lost on ignition; Mg, magnesium measured in milligrams per kilogram; Mn, manganese measured in milligrams per kilogram; Moisture, percentage of moisture present within sample; NO_3, nitrate measured in milligrams per kilogram; OM, percentage of organic matter; P, phosphorus measured in milligrams per kilogram; pH, acidity hydrogen ion concentration; Zn, zinc measured in milligrams per kilogram; <det, indicates quantity below detectable limit of instruments used.