Supporting Information

A. Supporting Materials and Methods

B. Supporting Tables

Table S1 Effects of block, eCO_2, eO_3, eCO_2+eO_3, depth and their combinations on soil microbial communities by Adonis analysis

Table S2. Significance tests of pairwise dissimilarities of overall microbial community structures between different treatments with three different statistical approaches

Table S3 Numbers of detected genes involved in carbon, nitrogen, sulfur and phosphorus cycling

Table S4 Correlations (*p* values) between soil properties or plant yield and signal intensities of functional genes by Mantel analysis in subsoil samples.

C. Supporting Figures

Fig. S1 DCA analysis of eCO_2, eO_3 and eCO_2+eO_3 effects on the microbial community structure of both surface soil and subsoil samples

Fig. S2 Maximum-likelihood phylogenetic tree of Rubisco gene sequences detected by GeoChip 3.0

Fig. S3 Effects of eCO_2, eO_3, and eCO_2+eO_3 on carbon degradation genes detected in the subsoil samples

Fig. S4 Simplified maximum-likelihood phylogenetic tree of *nifH* gene sequences detected by GeoChip 3.0

Fig. S5 Normalized signal intensities of key N cycling genes detected under eCO_2, eO_3, and eCO_2+eO_3 and control conditions in the subsoil soil

Fig. S6 The normalized signal intensities of key P cycling genes (A) and key S cycling genes (B) detected from surface soil (0-5 cm) and subsoil (5-15 cm) samples.

D. Supporting references

A. Supporting Materials and Methods

Site description and sampling

This study was conducted at the soybean free-air concentration enrichment (SoyFACE) facility located in Champaign, IL, USA (40°02′N, 88°14′W, 228 m above sea level) (http://www.soyface.uiuc.edu/index.htm) in 2008. The 32-ha SoyFACE experiment was established on a farmland that had been cultivated with an annual rotation of soybean, *Glycine max* (L.) Merr. and corn, *Zea mays* L. for more than 25 years, and the soil at the site is a Drummer fine-silty, mixed, mesic Typic Endoaquoll, typical of wet, dark-colored 'prairie soils' in northern and central Illinois (Pujol Pereira et al 2011). More soil background properties, including soil pH, moisture, Bray P, K, Ca, and Mg were previously documented (Peralta and Wander 2008). SoyFACE is designed to discover the effects of atmospheric change on the agronomy, productivity and ecology of Midwestern agro-ecosystems planted in a typical corn-soy rotation. The experiment was a randomized complete block design (n = 4) with each block containing four treatments: (i) ambient CO_2 (~400 ppm in 2008) and O_3 (~37.9 ppb in 2008), (ii) elevated CO_2 (~550 ppm), (iii) elevated O_3 (~ 61.3 ppb in 2008), and (iv) a combination of elevated CO_2 and O_3 (~550 ppm CO_2 and ~61.3 ppb O_3). Concentrations of CO_2 and O_3 were controlled by an adjustable segmented ring encircling each plot area that released high velocity gas just above the surface of the crop canopy. A total of 96 soil samples were collected in October 2008 from four soybean (*Glycine max* Merr.) grown plots under each of four treatments at both surface soil (0-5 cm) and subsoil (5-15 cm) layers with 48 samples for each soil layer, 12 samples for each treatment, and three subsamples for each plot. All soil samples were immediately transferred to the laboratory and stored at -80^{o}C until DNA extraction or soil property analyses.

Crop yield and soil property analysis

Annual crop yield data were collected for each plot as previously described (Morgan et al 2005) and previously reported (Twine et al 2013). To estimate the historical effects of elevated CO_2 on seed yield production prior to the time of our sampling, soybean yield data from 2004 and 2006 were averaged. Soil NO_3-N and NH_4-N were extracted with 1 M KCl solution and quantified by a Flow Injection Autoanalyzer (LACHAT 1994). Soil organic carbon and total N were determined using a LECO Truspec dry combustion carbon analyzer (Nelson and Sommers 1996).

Microbial community DNA extraction, purification and quantitation

Soil DNA was extracted by freeze-grinding mechanical lysis as described previously (Zhou et al 1996), and was purified using a low melting agarose gel followed by phenol extraction for 96 soil samples collected in October 2008. DNA quality was assessed by the ratios of 260 nm/280 nm, and 260/230 nm using a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE), and it should meet the following criteria: 260 nm/280 nm > 1.70, and 260 nm/230 nm > 1.80. The final soil DNA concentrations were quantified by the PicoGreen method (Ahn et al 1996) using a FLUOstar Optima (BMG Labtech, Jena, Germany).

Geochip analysis

A comprehensive functional gene array, GeoChip 3.0 was used to analyze the functional composition, structure and metabolic potential of all 96 microbial communities. GeoChip 3.0 contains > 28,000 probes covering approximately 57,000 gene variants from 292 functional gene families involved in carbon (C), nitrogen (N), phosphorus (P) and sulfur (S) cycling, energy metabolism, antibiotic resistance, metal resistance and organic contaminant degradation. It also has several other distinct features, such as a common oligonucleotide as the universal standard (CORS) for data normalization and comparison (Liang et al 2010), a software package for data management, and the *gyrB* gene for phylogenetic analysis (He et al 2010a).

a. Target amplification

In order to produce consistent hybridizations from all samples, a whole community genome amplification (WCGA) was used to generate approximately 3.0 µg of DNA with 50 ng purified

DNA as the template using the TempliPhi Kit (GE Healthcare, Piscataway, NJ) following the manufacturer's instructions (Wu et al 2006). Also, single-strand binding protein (267 ng μL^{-1}) and spermidine (0.1 mM) were added to the reaction mix to improve the amplification efficiency. The reactions were incubated at 30°C for 3 hours and stopped by heating the mixtures at 65°C for 10 min.

b. Target labeling

After amplification, all products were labeled with the fluorescent dye Cy-5 using random priming method as follows. First, the whole amplified products were mixed with 20 µL random primers, denatured at 99.9 °C for 5 min, and then immediately chilled on ice. Following denaturation, the labeling master mix containing 2.5 µL dNTP (5 mM dAGC-TP, 2.5 mM dTTP), 1 µL Cy-5 dUTP (Amersham, Piscataway, NJ), 80 U of the large Klenow fragment (Invitrogen, Carlsbad, CA), and 2.5 µL water were added, incubated at 37 °C for 3 hours, and heated at 95°C for 3 min. Labeled DNA was purified using the QIA quick purification kit (Qiagen, Valencia, CA) according to the manufacturer's instructions, measured on a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE), and then dried down in a SpeedVac (ThermoSavant, Milford, MA) at 45°C for 45 min.

c. GeoChip hybridization and imaging processing

The labeled target was resuspended in 120 µl hybridization solution containing 50% formamide, 3 x SSC, 10.0 µg of unlabeled herring sperm DNA (Promega, Madison, WI), and 0.1% SDS, and the mix was denatured at 95°C for 5 min and kept at 50°C until it was deposited directly onto a microarray. Hybridizations were performed with a TECAN Hybridization Station HS4800 Pro (TECAN, Durham, NC, US) according to the manufacturer's recommended method. After washing and drying, the microarray was scanned by ScanArray Express Microarray Scanner (Perkin Elmer, Boston, MA) at 633 nm using a laser power of 90% and a photomultiplier tube (PMT) gain of 75%. The ImaGene version 6.0 (Biodiscovery, El Segundo, CA) was then used to determine the intensity of each spot, and identify poor-quality spots.

d. Data pre-processing

Raw data from ImaGene were submitted to Microarray Data Manager in our website (http://ieg.ou.edu/microarray/) and analyzed using the data analysis pipeline with the following major steps: (i) The spots flagged as 1 or 3 by Imagene and with a signal to noise ratio (SNR) less than 2.0 (He and Zhou 2008) were removed as poor-quality spots; (ii) After removing the

bad spots, normalized intensity of each spot was calculated by dividing the signal intensity of each spot by the mean intensity of the microarray; (iii) If any of replicates had (signal–mean) more than two times the standard deviation, this replicate was moved as an outlier. This process continued until no such replicates were identified; (iv) At least 0.34 time of the final positive spots (probes), or a minimum of two spots was required for each gene; and (v) If a probe appeared in two or fewer samples among the total of 12 samples for each treatment condition, it was removed for data reliability (He et al 2010b).

Statistical analysis

Preprocessed GeoChip data were further analyzed using different statistical methods as previously described (He et al 2010b).

a. Analysis of Variance (ANOVA)

Analysis of variance (ANOVA) test was used to examine the differences among different treatments: ambient (the control), elevated CO_2 (eCO_2), elevated O_3 (eO_3), and their combination (CO_2+eO_3). The *F* value from *F*-test in ANOVA measures the ratio of the variance between-treatments and the variance within-treatments, and thus the relevant *p* value shows if the probability of the variance between-treatments equals the variance within-treatments. If *p* value is less than the significance level (e.g., 0.05), it means the difference among treatments significantly bigger the difference within treatments, indicating the treatments are significantly different. The ANOVA model in this study was set as: $Y \sim CO_2 + O_3 + CO_2{:}O_3$. For ANOVA test of functional genes, signal intensities of multiple probes from each functional gene in each sample were summed together, and then each functional gene was performed individually. The "aov" function in R software version 2.9.1 was used to implement the ANOVA test.

b. Multivariate and direct gradient analysis

Detrended correspondence analysis (DCA) was used to determine the overall functional changes in the microbial community, which was implemented by the Vegan package in R software. DCA is an ordination technique that uses detrending to remove the arch effect, where the data points are organized in a horseshoe-like shape, in correspondence analysis (Hill and Gauch 1980).

Different datasets of microbial communities generated with different analytical methods

were used to examine whether different treatments (e.g., ambient, eCO_2, eO_3, eCO_2+eO_3) have significant effects on soil microbial communities. Typically, there is some difficulty for all datasets to meet the assumptions (*e.g.*, normality, equal variances, independence) of parametric statistics. Thus, in this study, three different non-parametric analyses for multivariate data were used: analysis of similarities (ANOSIM) (Clarke 1993), non-parametric multivariate analysis of variance (permanova or adonis) using distance matrices (Anderson 2001), and multi-response permutation procedure (MRPP). We used Jaccard (non-quantitate) and Bray-Curtis (quantitate) similarity indexes to calculate distance matrix from GeoChip hybridization data for ANOSIM, adonis and MRPP analyses (McCune and Grace 2002). All three methods are based on dissimilarities among samples and their rank order in different ways to calculate test statistics, and the Monte Carlo permutation is used to test the significance of statistics. All three procedures (anosim, adonis and mrpp) were performed by the Vegan package in R (R Development Core Team 2011).

c. Mantel test

To elucidate the inter-relationships between soil geochemical variables and the abundance of functional genes of microbial community detected by GeoChip, the Mantel test was employed. Mantel test is an appropriate statistic method to measure the correlation between dissimilarity matrices and the significance of the statistic is evaluated by permuting the matrixes (Borcard et al 1992). The geochemical data were standardized to zero mean and unit deviation before calculation. The Bray-Curtis distance was used to construct the dissimilarity matrixes of communities and environmental variables respectively. All Mantel analyses were performed by functions in the Vegan package in R (R Development Core Team 2011).

B. Supporting Tables

Table S1 Effects of block, eCO_2, eO_3, eCO_2+eO_3, depth and their combinations on soil microbial communities by Adonis analysis. The statistical model of adonis is $Y = (CO_2 + O_3 + Block + Depth)^2$, where Y is the mean response of the community by CO_2, O_3, block, depth and their interactions.

	F-value	p-value	Contribution %
Block	2.36	0.001	5.2%
Depth	2.97	0.001	3.8%
CO_2	2.95	0.002	3.5%
O_3	2.57	0.001	2.5%
CO_2:O_3	2.83	0.001	3.1%
Block:CO_2	1.74	0.001	4.2%
Block:O_3	1.59	0.004	3.8%
Block:Depth	1.59	0.002	3.9%
CO_2:Depth	2.96	0.001	2.5%
O_3:Depth	2.56	0.001	2.2%
Block:CO_2:O_3	2.19	0.001	5.3%
Block:CO_2:Depth	1.61	0.001	3.9%
Block:O_3:Depth	1.57	0.002	3.8%
CO_2:O_3:Depth	1.95	0.006	1.6%
Block:CO_2:O_3:Depth	1.55	0.003	3.8%
Residuals			46.9%

Table S2. Significance tests of pairwise dissimilarities of overall microbial community structures among between different treatments with three different statistical approaches. **Bond** p values indicate no significant ($p > 0.05$) differences between those two treatments.

	ANOISM*		adonis**		MRPP***	
	R	p	F	p	δ	p
Surface soil (0-5 cm)						
eCO_2 vs ambient	0.451	0.007	0.313	0.002	0.554	0.003
eO_3 vs ambient	0.400	0.020	0.279	0.027	0.584	0.027
eCO_2+eO_3 vs ambient	0.310	0.028	0.251	**0.066**	0.576	**0.052**
eCO_2 vs eO_3	0.389	0.003	0.305	0.007	0.566	0.003
eCO_2 vs eCO_2+eO_3	0.491	0.004	0.342	0.001	0.573	0.001
eO_3 vs eCO_2+eO_3	0.366	0.015	0.275	0.036	0.595	0.022
Subsoil (5-15 cm)						
eCO_2 vs ambient	0.400	0.004	0.283	0.007	0.604	0.003
eO_3 vs ambient	0.249	**0.069**	0.219	**0.103**	0.591	**0.067**
eCO_2+eO_3 vs ambient	0.219	**0.125**	0.249	**0.065**	0.614	**0.067**
eCO_2 vs eO_3	0.390	0.005	0.279	0.002	0.576	0.005
eCO_2 vs eCO_2+eO_3	0.292	0.028	0.261	0.022	0.599	0.016
eO_3 vs eCO_2+eO_3	0.272	0.041	0.253	0.058	0.586	0.048

*Analysis of similarities (ANOSIM); **Non-parametric multivariate analysis of variance (MANOVA) with the adonis function; ***Multi-response permutation procedure (MRPP).

Table S3 Number of detected genes involved in carbon, nitrogen, sulfur and phosphorus cycling									
	Surface soil (0-5 cm)					Subsoil (5-15 cm)			
Gene/enzyme	Ambient	eCO2	eO3	eCO2+O3		Ambient	eCO2	eO3	eCO2+O3
CODH	8	9	9	13		13	14	11	12
pcc	63	73	64	66		75	77	63	60
rbcL	31	41	39	34		42	36	42	33
amyA	56	62	57	51		69	54	60	59
acetylglucosaminidase	8	11	6	13		12	10	11	10
ara	15	18	20	18		20	23	23	23
ara (fungi)	10	11	10	12		10	9	9	10
cellobiase	16	19	12	15		13	14	13	16
endochitinase	31	42	33	40		33	36	38	29
endoglucanase	4	6	4	7		8	7	3	5
exochitinase	2	3	3	3		3	2	3	3
exoglucanase	8	9	9	8		9	9	8	10
glucoamylase	8	8	8	9		7	8	9	10
pulA	12	28	14	21		26	18	20	25
vdh	4	4	1	4		5	4	4	5
vanA	28	30	27	30		34	29	33	28
limEH	6	6	6	8		9	6	6	8
xylanase	11	14	14	17		12	16	13	16
glx	6	5	5	6		5	6	6	6
lip	9	11	12	9		11	10	12	8
mnp	5	7	7	6		8	9	7	7
phenol_oxidase	24	33	31	30		29	35	32	28
Carbon cycling	365	450	391	420		453	432	426	411
amoA	7	12	6	7		12	5	9	4
gdh	4	4	4	2		6	3	4	5
napA	5	9	8	10		12	15	10	11
narG	58	73	63	58		78	70	68	74
nasA	21	26	20	21		27	26	31	24
nifH	123	136	116	131		146	121	147	124
nirK	57	77	53	57		70	52	59	63
nirS	40	54	35	50		48	55	43	48
norB	14	21	17	20		18	20	17	22
nosZ	41	54	44	45		43	46	48	42
nrfA	19	16	19	21		19	18	17	15
ureC	45	49	51	19		57	52	50	48
Nitrogen cycling	434	531	436	441		536	483	503	480
phytase	2	3	3	3		3	3	1	4
ppk	21	27	26	28		31	26	24	27
ppx	54	62	59	62		69	50	55	54
Phosphorus cycling	77	92	88	93		103	79	80	85
dsrA	76	104	82	82		98	99	82	85
dsrB	35	40	33	35		34	44	34	32
sox	47	51	48	47		49	49	47	48
Sulphur cycling	158	195	163	164		181	192	163	165
Total	**1034**	**1268**	**1078**	**1118**	**0**	**1273**	**1186**	**1172**	**1141**

Table S4 Correlations (*p* values) between soil properties or plant yield and signal intensities of functional genes by Mantel analysis of subsoil (5-15 cm) samples. Bold face indicates significantly changed *p* values ($p < 0.05$).

Gene/enzyme	Functional process	NO_3-N	NH_4-N	TN	TC	C/N	Yield
All detected		0.621	0.574	0.892	0.843	0.822	0.457
FTHFS	Acetogenesis	0.318	0.925	**0.035**	**0.024**	0.261	0.151
amyA	C degradation	0.509	0.860	0.184	0.206	0.503	**0.041**
pulA	C degradation	0.821	**0.041**	0.627	0.858	0.145	0.620
gdh	Ammonification	**0.016**	0.391	0.059	**0.010**	0.234	0.234
nrfA	N reduction	0.765	**0.004**	0.679	0.921	0.430	0.619
ureC	Ammonification	0.168	0.543	0.207	0.117	0.841	**0.023**
ppk	P utilization	**0.041**	0.299	0.761	0.267	0.775	0.066
sox	Sulphur oxidation	0.525	**0.001**	0.969	0.847	0.990	0.706
cnrA	Cobalt & Nickel	0.102	0.603	**0.006**	**0.033**	0.083	0.163
cusF	Copper	0.141	0.543	0.078	**0.030**	0.301	0.636
metC	Mercury	0.542	**0.005**	0.986	0.827	0.874	0.361
nreB	Nickel	0.496	**0.005**	0.878	0.594	0.653	0.526
terD	Tellurium	0.140	0.475	**0.028**	0.063	0.071	0.232
amiE	Aromatics	0.387	0.716	**0.029**	**0.034**	0.396	0.117
bphA	Aromatics	0.864	**0.010**	0.951	0.756	0.972	0.203
bphB	Aromatics	0.413	0.430	**0.043**	0.071	0.100	0.433
cumB	Aromatics	**0.049**	0.122	0.495	0.245	0.218	0.108
ebdABC	Aromatics	0.220	0.237	0.066	**0.044**	0.135	0.334
hcaB	Aromatics	0.220	0.751	0.066	**0.029**	0.173	0.385
hdnO	Aromatics	0.123	0.933	**0.004**	**0.030**	0.111	0.106
mdlB	Aromatics	0.830	**0.006**	0.889	0.850	0.855	0.559
nagI	Aromatics	0.300	**0.042**	0.063	**0.041**	0.346	0.708
pheA	Aromatics	0.105	0.563	0.135	**0.048**	0.365	0.190
phtA	Aromatics	0.111	0.906	0.286	0.411	0.610	**0.003**
tdnB	Aromatics	**0.048**	0.168	0.573	0.346	0.212	0.109
atzB	Herbicides	0.449	0.724	0.207	0.527	0.498	**0.003**
phn	Herbicides	0.295	0.488	0.511	0.161	0.354	**0.012**
trzE	Herbicides	0.210	0.765	0.061	**0.045**	0.320	**0.016**
trzN	Herbicides	0.725	0.329	0.752	0.835	0.502	**0.042**
alkJ	Hydrocarbons	0.978	**0.008**	0.799	0.741	0.866	0.896
linC	Pesticides	0.390	0.735	**0.001**	**0.012**	0.077	0.548

C. Supporting Figures

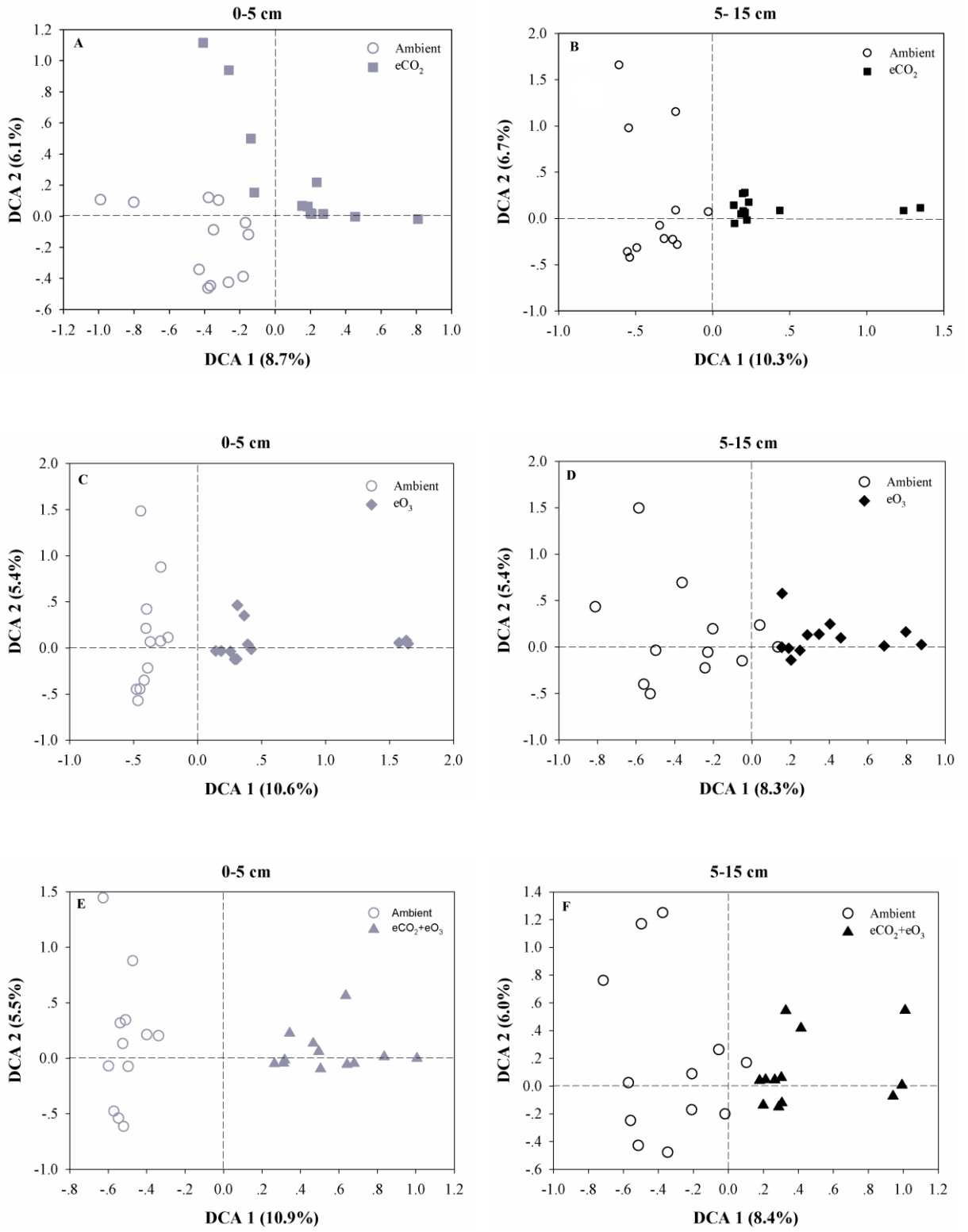
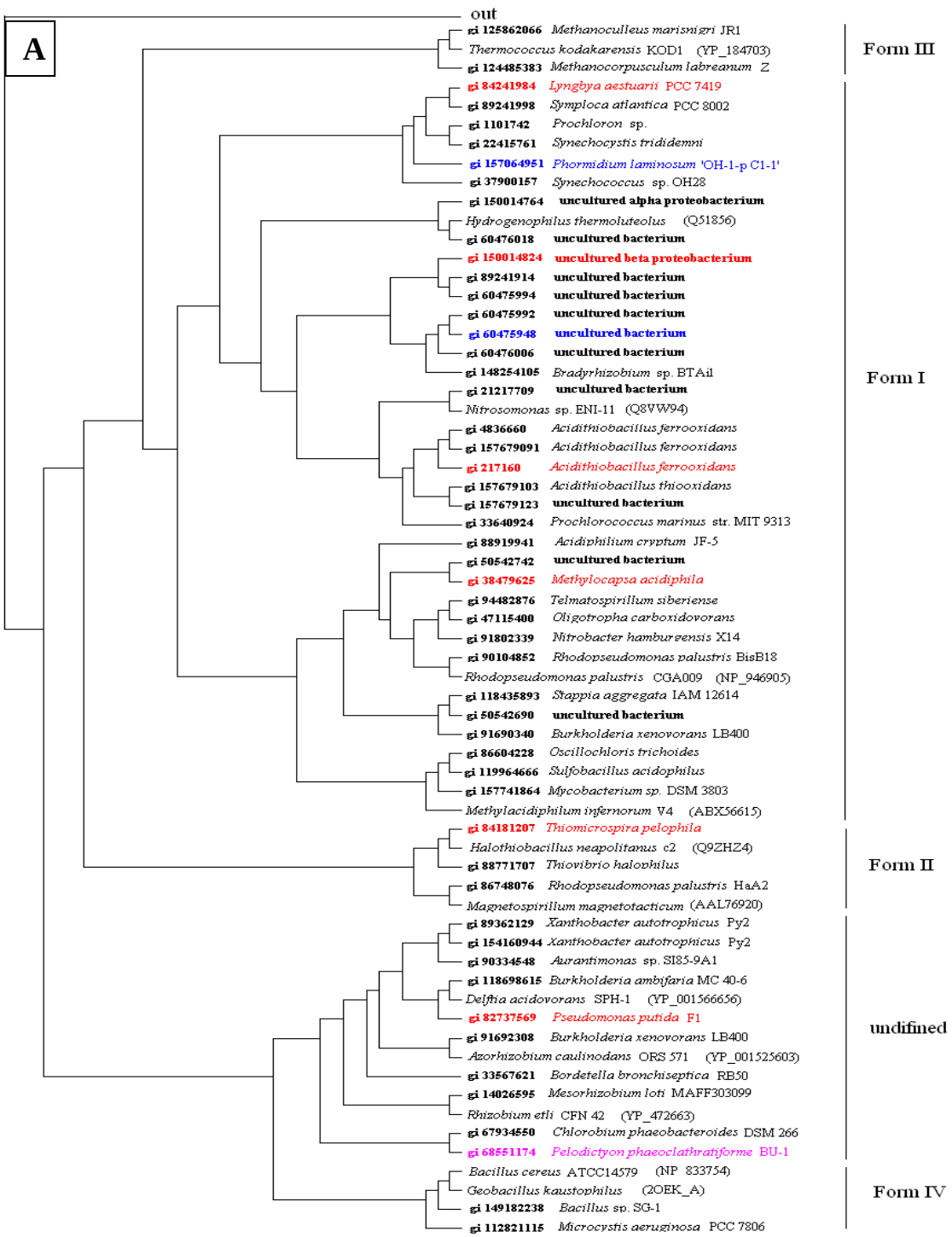


Fig. S1 Detrended corresponding analysis (DCA) of eCO_2 (A and B), eO_3 (C and D), and eCO_2+eO_3 (E and F) effects on the microbial community structure of both surface soil (A, C and E) and subsoil (B, D and F) samples.


A
out
gi 125862066 Methanoculleus marisnigri JR1
Thermococcus kodakarensis KOD1 (YP_184703)
gi 124485383 Methanocorpusculum labreanum Z
gi 84241984 Lyngbya aestuarii PCC 7419
gi 89241998 Symploca atlantica PCC 8002
gi 1101742 Prochloron sp.
gi 22415761 Synechocystis trididemni
gi 157064951 Phormidium laminosum 'OH-1-p C1-1'
gi 37900157 Synechococcus sp. OH28
gi 150014764 uncultured alpha proteobacterium
Hydrogenophilus thermoluteolus (Q51856)
gi 60476018 uncultured bacterium
gi 150014824 uncultured beta proteobacterium
gi 89241914 uncultured bacterium
gi 60475994 uncultured bacterium
gi 60475992 uncultured bacterium
gi 60475948 uncultured bacterium
gi 60476006 uncultured bacterium
gi 148254105 Bradyrhizobium sp. BTAi1
gi 21217709 uncultured bacterium
Nitrosomonas sp. ENI-11 (Q8VW94)
gi 4836660 Acidithiobacillus ferrooxidans
gi 157679091 Acidithiobacillus ferrooxidans
gi 217160 Acidithiobacillus ferrooxidans
gi 157679103 Acidithiobacillus thiooxidans
gi 157679123 uncultured bacterium
gi 33640924 Prochlorococcus marinus str. MIT 9313
gi 88919941 Acidiphilium cryptum JF-5
gi 50542742 uncultured bacterium
gi 38479625 Methylocapsa acidiphila
gi 94482876 Telmatospirillum siberiense
gi 47115400 Oligotropha carboxidovorans
gi 91802339 Nitrobacter hamburgensis X14
gi 90104852 Rhodopseudomonas palustris BisB18
Rhodopseudomonas palustris CGA009 (NP_946905)
gi 118435893 Stappia aggregata IAM 12614
gi 50542690 uncultured bacterium
gi 91690340 Burkholderia xenovorans LB400
gi 86604228 Oscillochloris trichoides
gi 119964666 Sulfobacillus acidophilus
gi 157741864 Mycobacterium sp. DSM 3803
Methylacidiphilum infernorum V4 (ABX56615)
gi 84181207 Thiomicrospira pelophila
Halothiobacillus neapolitanus c2 (Q9ZHZ4)
gi 88771707 Thiovibrio halophilus
gi 86748076 Rhodopseudomonas palustris HaA2
Magnetospirillum magnetotacticum (AAL76920)
gi 89362129 Xanthobacter autotrophicus Py2
gi 154160944 Xanthobacter autotrophicus Py2
gi 90334548 Aurantimonas sp. SI85-9A1
gi 118698615 Burkholderia ambifaria MC 40-6
Delftia acidovorans SPH-1 (YP_001566656)
gi 82737569 Pseudomonas putida F1
gi 91692308 Burkholderia xenovorans LB400
Azorhizobium caulinodans ORS 571 (YP_001525603)
gi 33567621 Bordetella bronchiseptica RB50
gi 14026595 Mesorhizobium loti MAFF303099
Rhizobium etli CFN 42 (YP_472663)
gi 67934550 Chlorobium phaeobacteroides DSM 266
gi 68551174 Pelodictyon phaeoclathratiforme BU-1
Bacillus cereus ATCC14579 (NP 833754)
Geobacillus kaustophilus (2OEK_A)
gi 149182238 Bacillus sp. SG-1
gi 112821115 Microcystis aeruginosa PCC 7806
Form III
Form I
Form II
undifined
Form IV

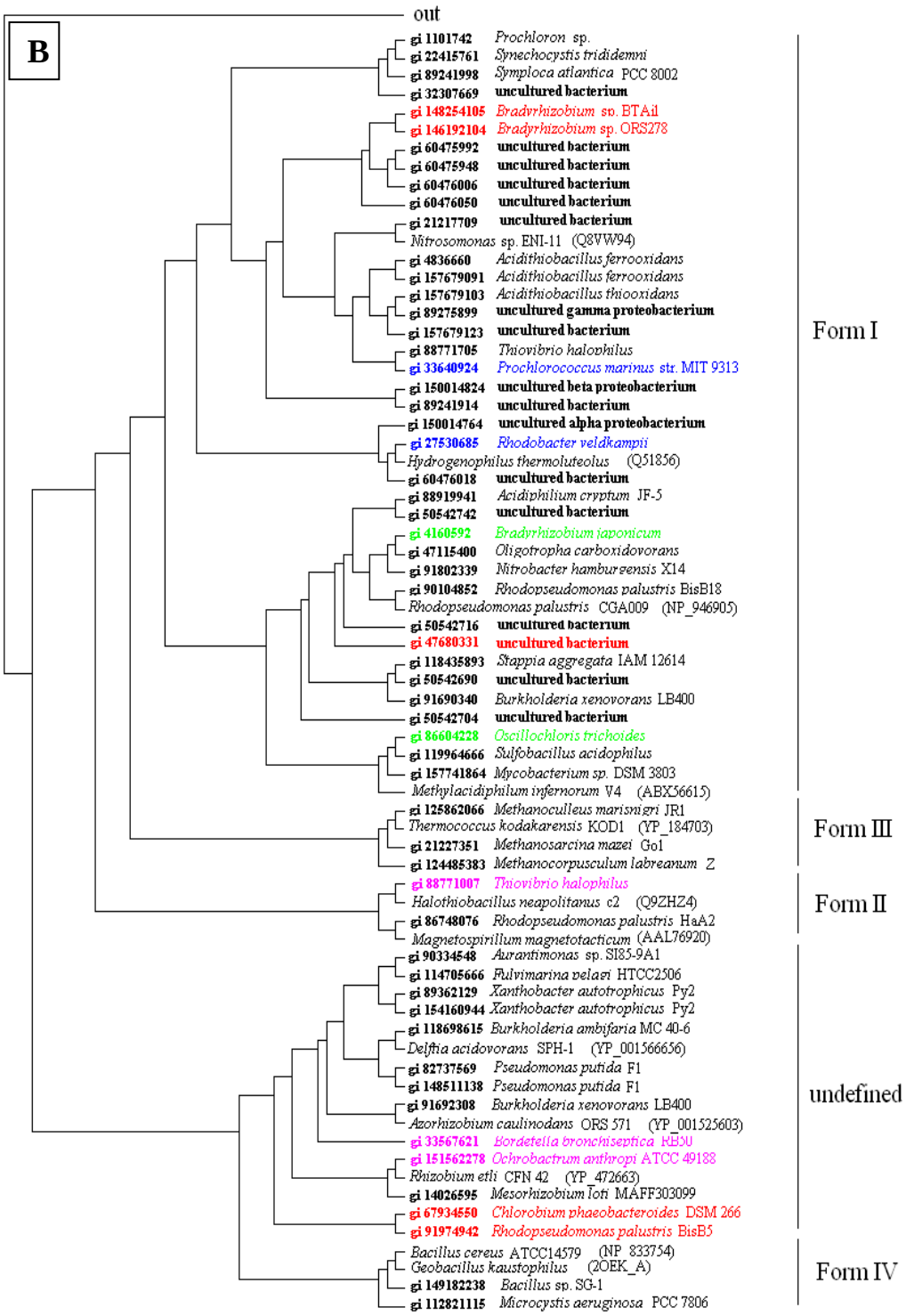


Fig. S2 Maximum-likelihood phylogenetic tree of Rubisco gene sequences detected by GeoChip 3.0. The phylogenetic relationship was shown among the five Rubisco forms; the reference sequences from GenBank are showed in regular with accession numbers, while the detected genes were showed in bold with the gene ID. **A.** 52 Rubisco gene sequences were detected in surface soil samples, and blue, red and pink colors are represented as unique genes at eCO_2, eO_3 and eCO_2+eO_3, respectively; **B.** 59 Rubisco gene sequences were detected in subsoil samples, and green, blue, red and purple colors represented unique genes in ambient, eCO_2, eO_3 and eCO_2+eO_3, respectively.

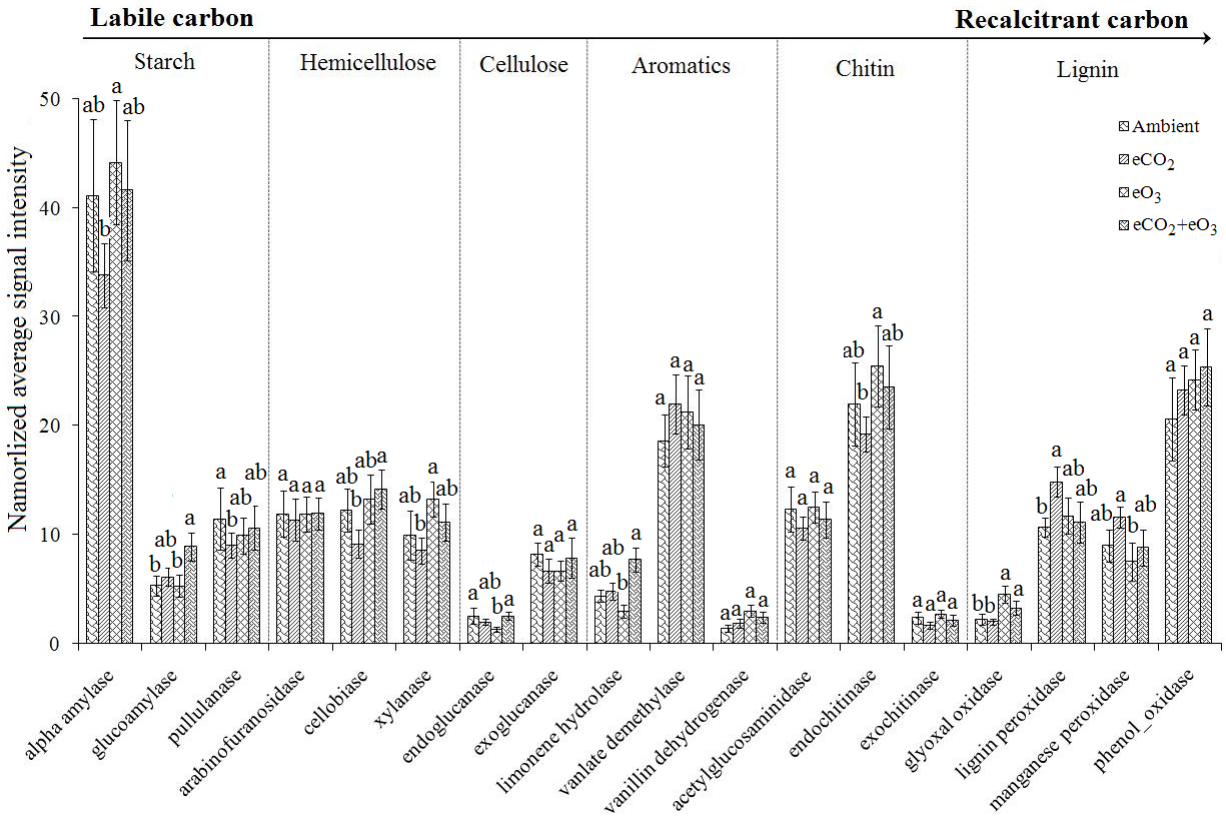
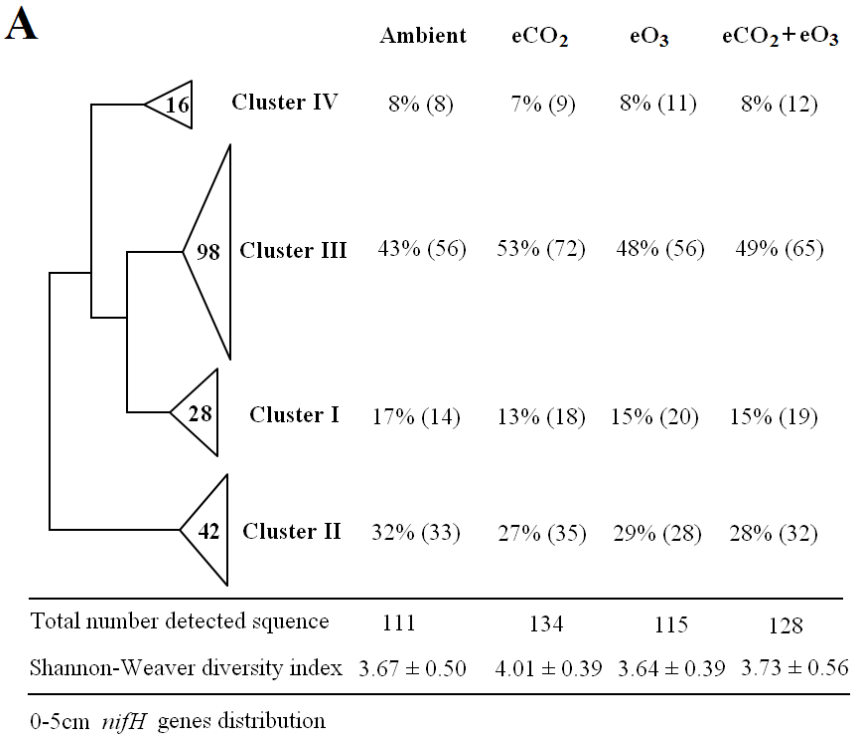


Fig. S3 Effects of eCO_2, eO_3, and eCO_2+eO_3 on carbon degradation genes detected in the subsoil samples. The complexity of carbon is presented in the order from labile to recalcitrant. All data are presented as mean ± SE. Significance among the treatments was analyzed by multi-way ANOVA.



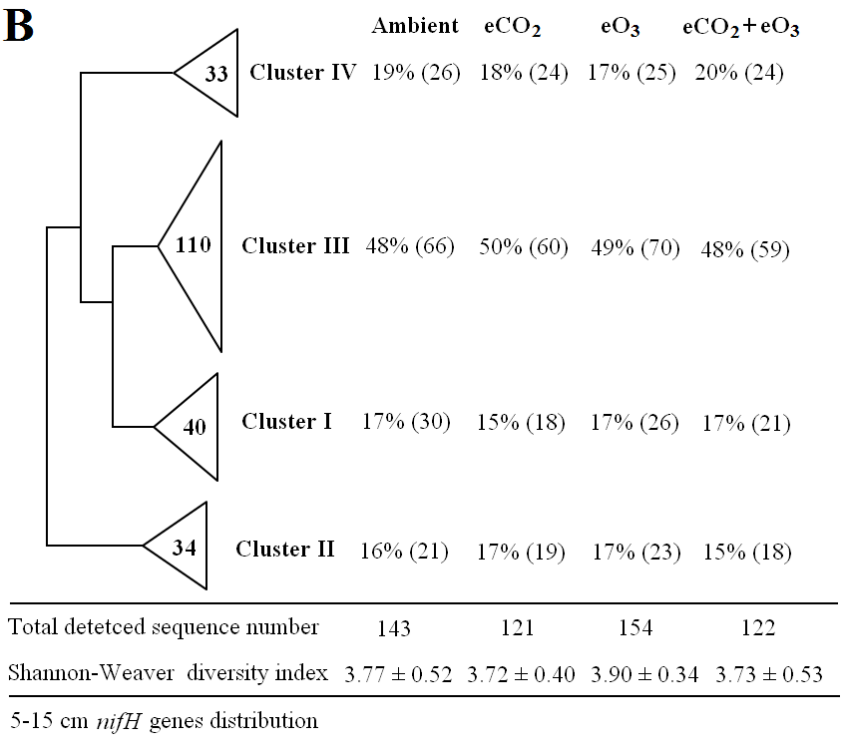


Fig. S4 Simplified maximum-likelihood phylogenetic tree of *nifH* gene sequences detected by GeoChip 3.0. **A.** 186 *nifH* gene sequences from surface samples. **B.** 217 *nifH* gene sequences from subsoil samples. The width of each wedge is the number of *nifH* sequences within each cluster. The percentages and numbers in bracket are the signal proportions and detected sequences in each cluster within each treatment, respectively. The significant differences of gene abundance were analyzed by ANOVA. The bottom of each tree shows the detected gene numbers and diversity indices.

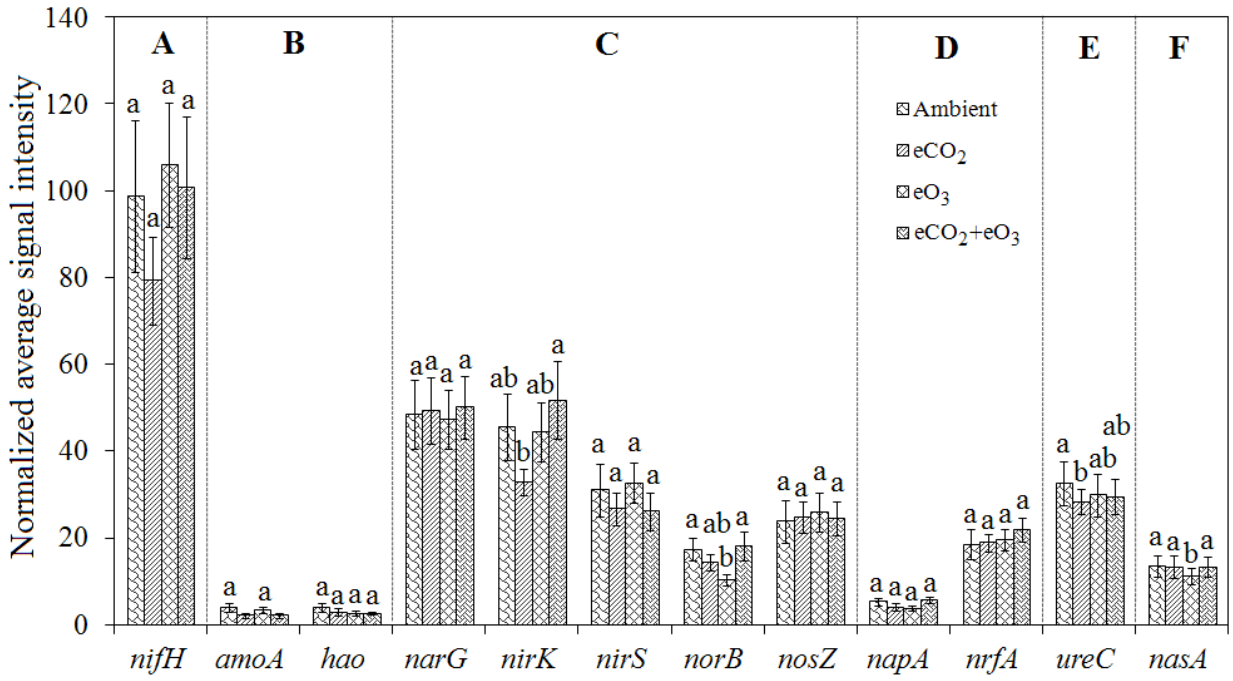


Fig. S5 Normalized signal intensities of key N cycling genes under eCO_2, eO_3, and eCO_2+eO_3 and control conditions in the subsoil (5-15 cm). (A) N_2 fixation, *nifH* encoding nitrogenase; (B) Nitrification, *amoA* encoding ammonia monooxygenase; (C) Denitrification, including *narG* for nitrate reductase, *nirS* and *nirK* for nitrite reductase, *norB* for nitric oxide reductase and *nosZ* for nitrous oxide reductase; (D) Dissimilatory N reduction to ammonium, including *napA* for nitrate reductase and *nrfA* for c-type cytochrome nitrite reductase; (E). Ammonification, including *gdh* for glutamate dehydrogenase and *ureC* encoding urease; (F) Assimilatory N reduction, *nasA* encoding nitrate reductase. All data are presented as mean ± SE. Significance among the treatments was analyzed by multi-way ANOVA.

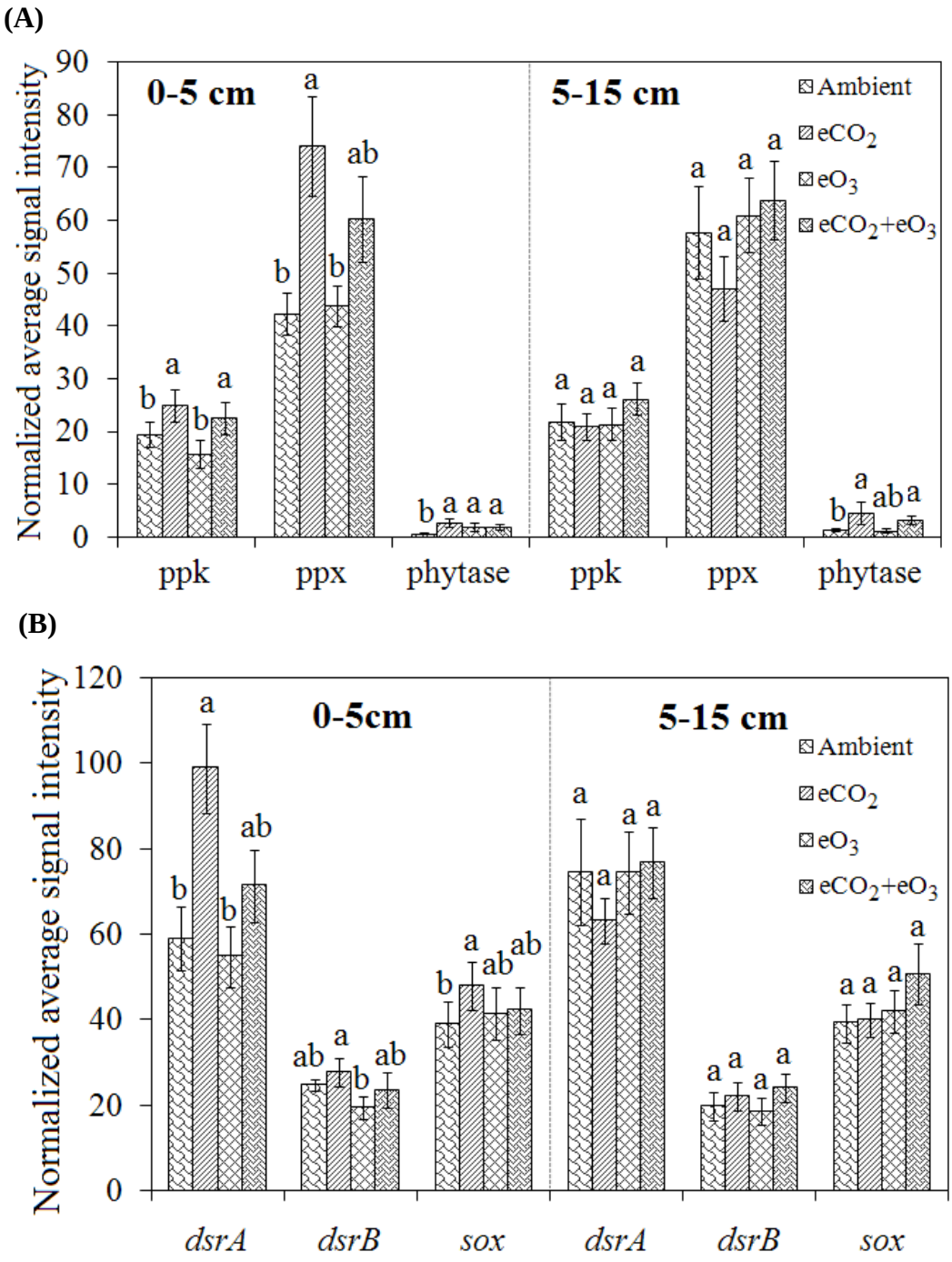


Fig. S6 The normalized signal intensities of key P cycling genes (A) and Key S cycling genes (B) detected from surface soil (0-5 cm) and subsoil (5-15 cm) samples. Ppx, exopolyphosphatase for inorganic polyphosphate degradation; Ppk, polyphosphate kinase for polyphosphate biosynthesis in prokaryotes, and phytase for phytate degradation. *dsrA* and *dsrB* are the genes for subunits of sulfite reductase for sulfur reduction, and *sox* is the genes for sulfur oxidation. All data are presented as mean ± SE. Significance among the treatments was analyzed by multi-way ANOVA.

D. Supporting reference cited

Ahn S, Costa J, Emanuel J (1996). PicoGreen quantitation of DNA: effective evaluation of samples pre- or post-PCR. *Nucl Acids Res* **24:** 2623-2625.

Anderson MJ (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecol* **26:** 32-46.

Borcard D, Legendre P, Drapeau P (1992). Partialling out the Spatial Component of Ecological Variation. *Ecology* **73:** 1045-1055.

Clarke KR (1993). Nonparametric Multivariate Analyses of Changes in Community Structure. *Aust J Ecol* **18:** 117-143.

He Z, Zhou J (2008). Empirical evaluation of a new method for calculating signal-to-noise ratio for microarray data analysis. *Appl Environ Microbiol* **74:** 2957-2966.

He Z, Deng Y, Van Nostrand JD, Tu Q, Xu M, Hemme CL *et al* (2010a). GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. *ISME J* **4:** 1167-1179.

He Z, Xu M, Deng Y, Kang S, Kellogg L, Wu L *et al* (2010b). Metagenomic analysis reveals a marked divergence in the structure of belowground microbial communities at elevated CO_2. *Ecol Lett* **13:** 564-575.

Hill MO, Gauch HG (1980). Detrended Correspondence-Analysis - an Improved Ordination Technique. *Vegetatio* **42:** 47-58.

LACHAT (1994). QuickChem Method 12-107-04-1-B. LACHAT Instrument: Milwaukee, WI.

Liang Y, He Z, Wu L, Deng Y, Li G, Zhou J (2010). Development of a common oligonucleotide reference standard for microarray data normalization and comparison across different microbial communities. *Appl Environ Microbiol* **76:** 1088-1094.

McCune B, Grace JB (2002). *Analysis of ecological communities*. MjM Software Design: Gleneden Beach, OR.

Morgan PB, Bollero GA, Nelson RL, Dohleman FG, Long SP (2005). Smaller than predicted increase in aboveground net primary production and yield of field-grown soybean under fully open-air [CO_2] elevation. *Glob Change Biol* **11:** 1856-1865.

Nelson DW, Sommers LE (1996). Total carbon, organic carbon, and organic matter. In: Sparks DL, Page AL, Helmke PA, Loeppert RH, Soltanpour PN, Tabatabai MA *et al* (eds). *Methods of Soil Analysis*. Soil Science Society of America: Madison, WI. pp 961-1010.

Peralta A, Wander M (2008). Soil organic matter dynamics under soybean exposed to elevated CO_2. *Plant Soil* **303:** 69-81.

Pujol Pereira EI, Chung H, Scow K, Sadowsky MJ, van Kessel C, Six J (2011). Soil nitrogen transformations under elevated atmospheric CO_2 and O_3 during the soybean growing season. *Environ Pollu* **159:** 401-407.

R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Twine TE, Bryant JJ, Richter K, Bernacchi CJ, McConnaughay KD, Morris SJ *et al* (2013). Impacts of elevated CO_2 concentration on the productivity and surface energy budget of the soybean and maize agroecosystem in the Midwest U.S. *Glob Change Biol* **(in press)**.

Wu L, Liu X, Schadt CW, Zhou J (2006). Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. *Appl Environ Microbiol* **72:** 4931-4941.

Zhou J, Bruns MA, Tiedje JM (1996). DNA recovery from soils of diverse composition. *Appl Environ Microbiol* **62:** 316-322.