Supporting Information

Zhou et al. 10.1073/pnas.1324044111

SI Materials and Methods

Site Description and Experimental Design and Sampling. Many groundwater ecosystems are contaminated with mixtures of metals, radionuclides, chlorinated solvents, and/or hydrocarbons (1). Subsurface amendment with slow release substrates such as emulsified vegetable soil (EVO) is believed to be one of the most effective approaches for sustained uranium (U) bioimmobilization in groundwater ecosystems (2). Thus, a long-term, integrated field-scale bioremediation experimental test was carried out at the US Department of Energy Oak Ridge Integrated Field Research Challenge (ORIFRC) study site, Oak Ridge, TN. The details of the experimental setup, manipulation, and monitoring of the groundwater system were described elsewhere (2). Briefly, in this experiment, there are three injection wells, one upgradient control well, and seven downgradient monitoring wells (2). These wells were very close together, only meters apart. The groundwater moves rapidly from upgradient to downgradient wells with the residence time of ~0.4 d based on hydraulic conductivity measurement. The EVO mixture (Terra Systems) contains 60% (wt/wt) soybean oil, 6% (wt/wt) food-grade surfactant, 0.3% yeast extract, and 0.05% $(NH_4)_3PO_4$ in water. Thus, EVO is a complex slow release carbon complex carbon substrate due to its poor water solubility, which is capable of yielding many intermediates, and supporting a variety of diverse groups of microorganisms. The EVO was mixed rigorously with groundwater pumped from the site to produce a 20% (vol/vol) EVO emulsion. A total of 3,400 L of such emulsion was equally injected into the three adjacent injection wells over 2 h on February 9, 2009. The majority of the injected EVO emulsion traveled downgradient, but a small amount of the amendment was forced to move against the prevailing hydrological gradient during the injection (2). Therefore, microbial changes were also observed in the upgradient wells at days 4 and 17. For this reason, the geochemistry data from the control well at these two time points were excluded from analysis.

Groundwater samples (2 L) were taken for geochemical and microbial analyses from the seven monitoring wells and the control well at different time points, 28 d before EVO injection, 4, 17, 31, 80, 140, and 269 d after injection (2). For convenience, 28 d before injection was considered as day 0 in this study. Although groundwater and sediment microbial populations may differ, groundwater samples were focused in this study because sampling sediment was not possible over the temporal and spatial scales of the experiment (2). Groundwater was filtered in situ with 8-μm filters to remove large particles, followed by filtering with 0.2-μm filters for sample collection (2). The filters were stored at −80 °C until use and distributed to different laboratories.

GeoChip Analysis. GeoChip-based metagenomic technology was used for dissecting microbial community functional structure as described elsewhere (3–5). Briefly, the frozen 0.2-μm filter (1/4 of the filter for day 269 and 1/2 of the filter for other samples) was grounded in liquid N_2 with pestle and mortar as described previously (6). The community DNA was extracted with the SDS-based chemical lysis method (6). The purified DNA was quantified with Quant-It PicoGreen kit (Invitrogen). A total of 50 ng DNA from each sample was amplified in triplicate using the Templiphi 500 Amplification kit (Amersham Biosciences) in a modified buffer DNA buffer (7) for 6 h. A total of 3 μg amplified DNA was labeled with Cy5 fluorescent dye and hybridized overnight with GeoChip 3.0 on a MAUI Hybridization system (BioMicro Systems) at 42 °C with 45% formamide (3). After hybridization, the slides were scanned with a ScanArray 5000 Microarray Analysis system (PerkinElmer) at a resolution of 10 μm, and the images were processed with ImaGene 6.0 (BioDiscovery).

Raw data from ImaGene were submitted to Microarray Data Manager on our website (http://ieg.ou.edu/microarray/default.cgi) 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 were removed as poor-quality spots. (*ii*) After removing the bad spots, the normalization was performed at three levels: individual subgrids on a single slide, technical replicates among samples, and across the whole data set. First, the mean Cy3 intensity of the universal standards in each subgrid was used to normalize the Cy5 intensity for probes in the same subgrid. Second, the Cy5 intensity after the first normalization was normalized again by the mean value of three technical replicates. In addition, the data were normalized by the mean intensity of universal standards (Cy3 channel) in all slides for Cy5 intensity of samples. (*iii*) If any replicates had (signal–mean) more than two times the SD, this replicate was removed as an outlier. This process continued until no such replicates were identified. (*iv*) If a probe appeared in only one sample among the total of seven wells for each time point, it was removed for all further analyses.

The signal intensities were normalized based on the mean signal intensity across all genes on the arrays. Because the same amounts of DNA from all samples were used for amplification, labeling, and hybridization, it is expected the average signal intensity across all of the genes should be approximately equal. A mean ratio was calculated for each positive spot by dividing the signal intensity of each spot with the mean signal intensity in each array. These normalized mean ratio data were then used for further analysis.

Statistical Analysis. The matrices of microarray data resulting from our pipeline were considered as "species" abundance in statistical analyses. Detrended correspondence analysis (DCA) was used to determine the overall functional changes in the microbial communities by R software version 2.9.1. 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 (8).

To understand whether EVO amendment impacts microbial community structure, the site-to-site variability in gene/population compositions, known as β-diversity, was measured with two commonly used dissimilarity indexes, Jaccard's incidence-based (D_J) and Bray–Curtis's abundance-based (D_{BC}) indexes, which both range from 0 to 1 (9, 10).

Different data sets of microbial communities generated by different analytical methods were used to examine whether elevated temperature has significant effects on soil microbial communities. Typically, it is difficult for all data sets to meet the assumptions (e.g., normality, equal variances, independence) of parametric statistics. Thus, in this study, three different complementary nonparametric analyses for multivariate data were used: analysis of similarity (ANOSIM) (11), nonparametric multivariate analysis of variance (Adonis) using distance matrices (12), and multiresponse permutation procedure (MRPP). We used the Bray–Curtis similarity index to calculate a distance matrix from GeoChip hybridization data for entire communities or individual functional gene categories for ANOSIM, Adonis, and

MRPP analyses. MRPP is a nonparametric procedure that does not depend on assumptions such as normally distributed data or homogeneous variances, but rather depends on the internal variability of the data (13). It is also interesting to notice that Adonis function in R program is also named as permutation multivariate analysis of variance (PERMANOVA) proposed by Anderson (12). 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 with the Vegan package (v.1.15–1) in R software version 2.9.1 (The R Foundation for Statistical Computing).

Null Model Analysis. β-Diversity represents the compositional variations among communities from site to site, which serves as a bridge to link local (α-diversity) and regional (γ-diversity) communities. β-Diversity is useful to understand patterns of species diversity across various spatial scales (14–17) and provides critical insights into the role of deterministic and stochastic processes in shaping community compositions and structure (14, 18–20). Low dissimilarity among communities that are otherwise identical in environmental conditions would imply a predominant role for deterministic assembly, whereas high dissimilarity would suggest a large role of stochastic assembly (19, 21). However, the measures of β-diversity are dependent on both α- and γ-diversity (22). It is not clear whether a change in β-diversity is due to the differences in the underlying assembly processes that generate β-diversity or to the differences in α- and γ-diversity (22). The apparent variations of the measured β-diversity could just simply be due to the differences in α- and γ-diversity. Thus, null model analysis, which assumes that an assemblage (a community) is not structured by species interactions (23), can provide a straightforward way to determine whether species/gene compositional differences among sites are caused by the changes in α- and γ-diversity or by the forces causing communities to be different from the expectations by random chance.

Two types of null model analyses were performed in this study. The first null model analysis is based on the method proposed by Chase et al. (22) with the following steps: (*i*) calculating the observed gene richness in each site (i.e., the control or monitoring well here) 1 (α_1) and 2 (α_2) at a particular time point and the number of shared species/genes (SS_{obs}) between these two sites; (*ii*) calculating the total number of species/genes detected in the "species/gene pool" (γ-diversity) from all sites at the particular time point, and the proportion of the sites occupied by each species/gene; and (*iii*) calculating the distribution of the expected shared species/genes from null model (SS_{exp}) by randomly drawing α_1 and α_2 species/genes from the species/gene pool. The probability of species/gene to be drawn is proportional to the among-site occupancy of this gene. Then SS_{exp} and the null expected Jaccard's similarity (J_{exp}) can be obtained for each drawing. This procedure is repeated 1,000 times. An average null expected Jaccard's similarity ($\bar{J}_{exp}$) and its SD can be estimated based on 1,000 drawings (σ_{exp}). The permutational analysis of multivariate dispersions (PERMDISP) was used to test the significance of the differences of the groundwater microbial communities at a particular time point from null model expectations (19).

Based on the observed Jaccard's similarity (J_{obs}) between two communities and the average null expected Jaccard's similarity and its SD, several similar metrics can be defined. One is standard effect size (SES), which can be calculated as: $SES = (J_{obs} - \bar{J}_{exp})/\sigma_{exp}$. SES is an index to measure the influence of deterministic factors on community composition and abundance (24). In this study, we introduced another similar quantitative metric, selection strength (SS), which is the proportion of the difference between the observed similarity and the null expected similarity divided by the observed similarity, $SS = (J_{obs} - \bar{J}_{exp})/J_{obs}$. Because it provides a quantitative estimation of the role of niche-based deterministic selection processes in shaping community composition and structure, such ratio is termed as selection strength. The complement of selection strength (1-SS) should provide the quantitative assessment of the importance of stochastic processes in regulating community composition and structure. Based on all pairwise comparisons, an average of each metric (SES, deviation, or SS) and its SD can be estimated. Hence, standard statistical significance tests such as *t* test and ANOVA can be used to test the significance of a metric across different experimental conditions (21). However, because such metrics are originally derived from every pairwise comparison, they could be not independent. Therefore, the nonparametric permutation test, permutational multivariate analysis of variance (PERMANOVA), was used to test whether these communities differed in their SES and SS.

Because the results from null model analyses are very sensitive to the models, approaches, and algorithms used (25), the second null model analysis (19) was also performed based on EcoSim (26), keeping the number of species per site and the number of sites occupied by each species constant. The null community is generated by randomly shuffling the original community with the independent swap algorithm by holding the number of genes/populations in each reactor and the number of reactors in which each gene/population appears constant. Because our microarray data set is too large to efficiently compute in EcoSim, the "randomizeMatrix function" of R program was carried out to generate 1,000 null communities. Then, PERMDISP was also used to assess the significance of the differences of the groundwater microbial communities at a particular time point from null model expectations (19).

A modified Raup–Crick metric was also used to assess whether the null-expected number of shared species between any two communities is different from the observed number of shared species (19). The proportion that the expected number of shared species between any two communities is greater than their observed number of shared species is calculated, which is termed the "modified Raup–Crick metric." This metric ranges from 0 to 1 (19). Low values indicate that the communities are highly deviant from the null expectation by random chance, whereas high values mean that the communities are more similar to those based on null model expectation. Similarly, PERMDISP was used to determine whether this metric is significantly different among the groundwater communities at various time points.

1. Wu WM, et al. (2006) Pilot-scale in situ bioremediation of uranium in a highly contaminated aquifer. 1. Conditioning of a treatment zone. *Environ Sci Technol* 40(12):3978–3985.
2. Gihring TM, et al. (2011) A limited microbial consortium is responsible for extended bioreduction of uranium in a contaminated aquifer. *Appl Environ Microbiol* 77(17): 5955–5965.
3. He Z, et al. (2010) GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. *ISME J* 4(9): 1167–1179.
4. He Z, et al. (2007) GeoChip: A comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *ISME J* 1(1):67–77.
5. Zhou JZ, et al. (2012) Microbial mediation of carbon-cycle feedbacks to climate warming. *Nat Clim Change* 2(2):106–110.
6. Zhou JZ, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol* 62(2):316–322.
7. Wu LY, Liu X, Schadt CW, Zhou JZ (2006) Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. *Appl Environ Microbiol* 72(7):4931–4941.
8. Hill MO, Gauch HG (1980) Detrended correspondence-analysis: An improved ordination technique. *Vegetatio* 42(1-3):47–58.
9. Bray JR, Curtis JT (1957) An ordination of the Upland Forest Communities of Southern Wisconsin. *Ecol Monogr* 27(4):326–349.

10. Faith DP, Minchin PR, Belbin L (1987) Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* 69(1-3):57–68.
11. Clarke KR (1993) Nonparametric multivariate analyses of changes in community structure. *Aust J Ecol* 18(1):117–143.
12. Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 26(1):32–46.
13. Mielke PW, Berry KJ, Brockwell PJ, Williams JS (1981) A class of nonparametric-tests based on multiresponse permutation Procedures. *Biometrika* 68(3):720–724.
14. Anderson MJ (2011) Navigating the multiple meanings of beta diversity: A roadmap for the practicing ecologist. *Ecol Lett* 14(2):19–28.
15. Jost L (2010) Independence of alpha and beta diversities. *Ecology* 91(7):1969–1974.
16. Tuomisto H (2010) A diversity of beta diversities: Straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography* 33(1):2–22.
17. Tuomisto H (2010) A diversity of beta diversities: Straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography* 33(1): 23–45.
18. Cadotte MW (2007) Concurrent niche and neutral processes in the competition-colonization model of species coexistence. *Proc Biol Sci* 274(1626):2739–2744.
19. Chase JM (2010) Stochastic community assembly causes higher biodiversity in more productive environments. *Science* 328(5984):1388–1391.
20. Chase JM, Myers JA (2011) Disentangling the importance of ecological niches from stochastic processes across scales. *Philos Trans R Soc Lond B Biol Sci* 366(1576): 2351–2363.
21. Chase JM (2007) Drought mediates the importance of stochastic community assembly. *Proc Natl Acad Sci USA* 104(44):17430–17434.
22. Chase JM, Kraft NJB, Smith KG, Vellend M, Inouye BD (2011) Using null models to disentangle variation in community dissimilarity from variation in α-diversity. *Ecosphere* 2(2):art24.
23. Ulrich W, Gotelli NJ (2010) Null model analysis of species associations using abundance data. *Ecology* 91(11):3384–3397.
24. Kraft NJB, et al. (2011) Disentangling the drivers of β diversity along latitudinal and elevational gradients. *Science* 333(6050):1755–1758.
25. Gotelli NJ (2000) Null model analysis of species co-occurrence patterns. *Ecology* 81(9): 2606–2621.
26. Gotelli NJ, Entsminger GL (2009) EcoSim: Null models software for ecology. Version 7.

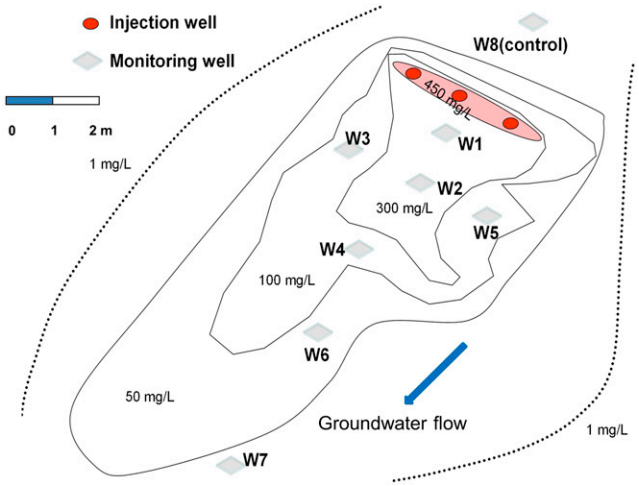

Fig. S1. Overview of the experimental system. The map of the wells for in situ U(VI) immobilization with injection of EVO. W1-7, downgradient monitoring wells; W8, upgradient control well.

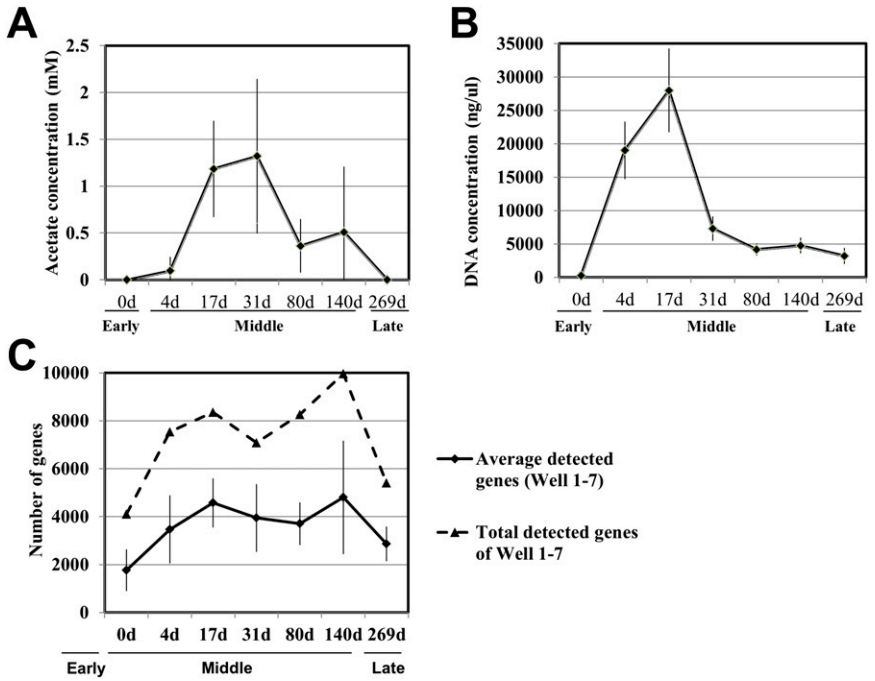

Fig. S2. Changes of acetate (*A*), DNA concentrations (*B*), and diversity (*C*) during the experiment. α-Diversity is represented by the average number of the genes detected across all monitoring wells at a particular time point. The γ-diversity is the total number of genes detected across all wells at a particular time point.

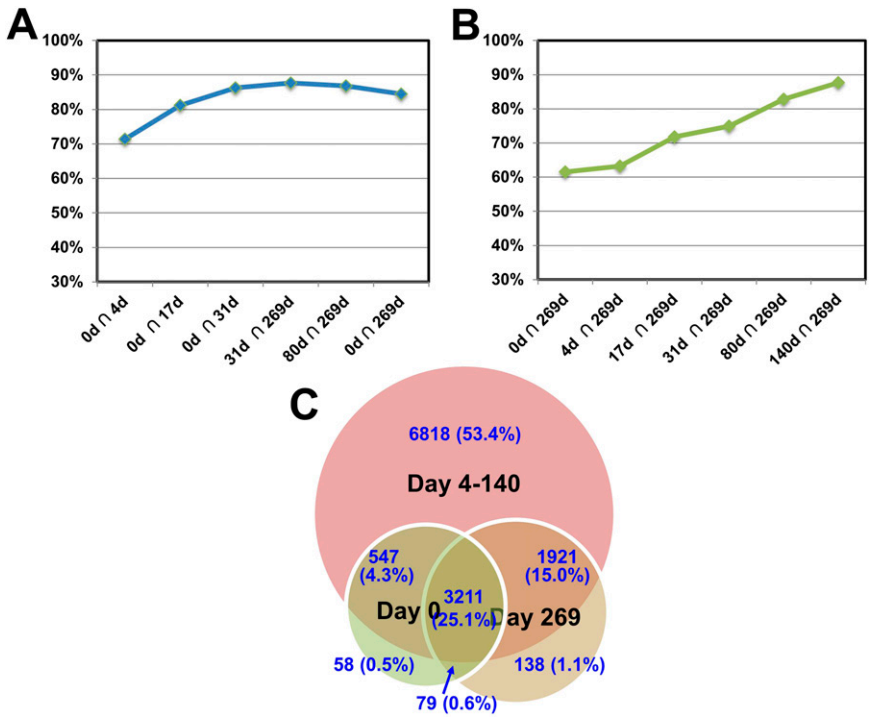

Fig. S3. Changes of microbial community structure across different time points. (*A*) Percentages of the overlap of genes detected at day 0 with the other time points. The percentage was calculated as the proportion of the overlapped gene number divided by the total number of genes detected at either day 0 or day 269. (*B*) Percentages of the overlap of genes detected at day 269 with the other time points. (*C*) The numbers and proportions of the unique functional genes detected in the early (day 0), middle (days 4–140), and final phase (day 269) of the succession.

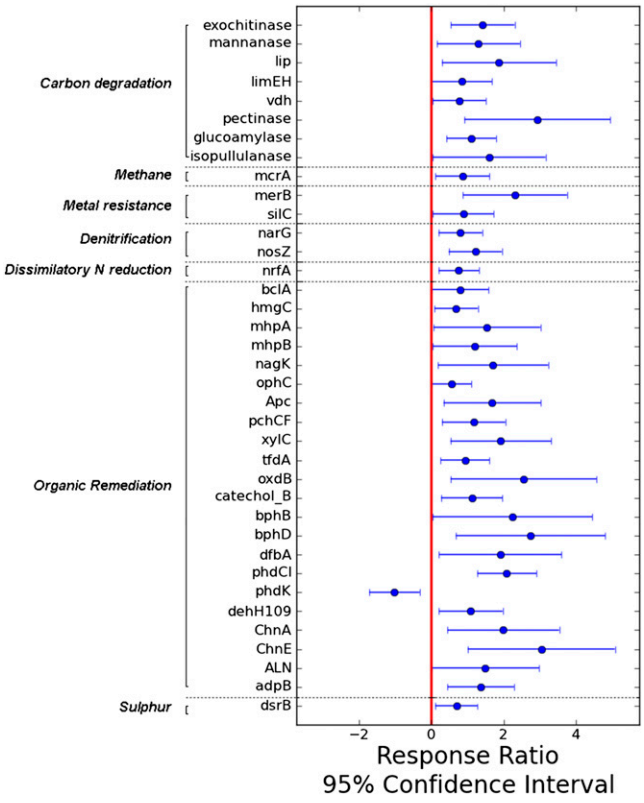

Fig. S4. Response ratios of major functional gene categories showing their significant changes between days 17 and 0 after EVO amendment. Based on DNA yields (Fig. S2*B*), day 17 represented the peak of the community shift so that it was selected for comparison, showing the maximum impacts of EVO on microbial community structure.

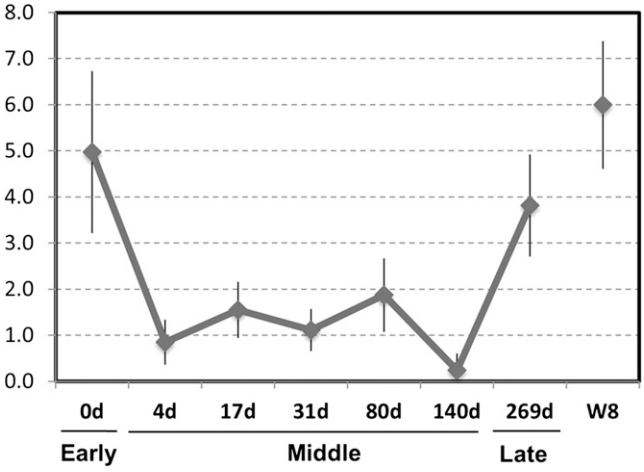

Fig. S5. Changes of standard effect size (SES) during the succession of the groundwater microbial communities. SES is the differences between the observed similarity and the null expected similarity divided by the SD of community similarity from null expectation, which is used to measure the influence of deterministic factors on community composition and abundance.

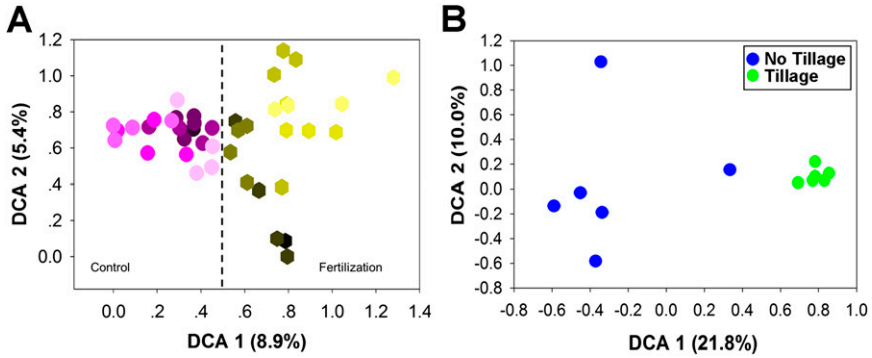

Fig. S6. Effects of long-term fertilization and tillage on microbial functional community structure. (*A*) Fertilization. Soil samples were from a long-term grassland fertilization experiment in Rothamstead, England. Diamond, samples from fertilization plots for 150 y; circle, samples from control plots without fertilization. (*B*) Tillage. Soil samples were collected from tillage and nontillage plots with monoculture of annually rotating corn, soybean, and wheat at the Kellogg Biological Station (KBS) Long-Term Ecological Research (LTER) experimental site at Hickory Corners, MI.

Table S1. Unique (bold) and overlapped detected probes in seven different wells (W1 to W7) at day 0

Sample name	W1_0 (%)	W3_0 (%)	W4_0 (%)	W5_0 (%)	W6_0 (%)	W7_0 (%)
W1_0	**39 (3.90)**	845 (37.27)	810 (50.63)	606 (45.33)	867 (43.92)	937 (28.10)
W3_0		**348 (16.49)**	1154 (48.77)	795 (35.21)	1282 (48.03)	1672 (45.08)
W4_0			**53 (3.76)**	694 (41.88)	1168 (56.13)	1307 (38.76)
W5_0				**51 (5.41)**	721 (34.98)	836 (24.76)
W6_0					**33 (1.79)**	1787 (53.78)
W7_0						**954 (29.17)**
Total detected probes	1,001	2,111	1,409	942	1,840	3,270

The data for W2 at day 0 were missing due to the low biomass concentration and an error during handling. The map of the well location is seen in Fig. S1.

Table S2. Significance tests of the differences of the microbial communities at days 269 and 0 and control wells

	Adonis*		ANOSIM†		MRPP‡	
Data sets	*F*	*P*	*R*	*P*	δ	*P*
Whole communities	4.870	0.004	0.512	0.001	0.408	0.007
Antibiotic resistance	7.921	0.004	0.713	0.003	0.429	0.004
Carbon cycling	5.265	0.002	0.557	0.005	0.408	0.002
Energy process	2.723	0.021	0.311	0.009	0.414	0.023
Metal resistance	4.032	0.008	0.456	0.004	0.400	0.007
Nitrogen	4.227	0.006	0.440	0.004	0.413	0.005
Organic remediation	4.293	0.005	0.429	0.003	0.411	0.004
Phylogenetic marker	3.435	0.028	0.356	0.012	0.351	0.019
Phosphorus	5.376	0.003	0.475	0.002	0.403	0.006
Sulfur	5.989	0.006	0.558	0.003	0.403	0.002

Different statistical approaches were used with Bray–Curtis distances, which were estimated based on GeoChip hybridization data. *P* values are of corresponding significance tests.

*Permutational multivariate analysis of variance using distance matrices. Significance tests were performed by *F* test based on sequential sums of squares from permutations of the raw data.

†Analysis of similarities. Statistic *R* is based on the difference of mean ranks between groups and within groups. The significance of observed *R* is assessed by permuting the grouping vector to obtain the empirical distribution of *R* under the null model.

‡Multiresponse permutation procedure. Statistic δ is the overall weighted mean of within-group means of the pairwise dissimilarities among sampling units. The significance test is the fraction of permuted δ that is less than the observed δ.

Table S3. List of the functional genes showing significant differences in abundance between days 269 and 0

Gene	Gene category	Subcategory	Abundance differences between days 269 and 0	*F* value	Pr (>*F*)
B_lactamase_A	Antibiotic resistance	Beta-lactamases	0.416	29.52	<0.001
B_lactamase_B	Antibiotic resistance	Beta-lactamases	0.634	4.27	0.046
B_lactamase_C	Antibiotic resistance	Beta-lactamases	0.115	8.74	0.003
Tet	Antibiotic resistance	Other	0.430	25.02	<0.001
ABC antibiotic transporter	Antibiotic resistance	Transporter	0.509	15.16	<0.001
MATE antibiotic	Antibiotic resistance	Transporter	0.368	25.94	<0.001
Mex	Antibiotic resistance	Transporter	0.644	17.23	<0.001
MFS antibiotic	Antibiotic resistance	Transporter	0.268	42.96	<0.001
SMR antibiotics	Antibiotic resistance	Transporter	0.502	292.86	<0.001
cellobiase	Carbon cycling	Carbon degradation	0.317	8.78	0.003
endoglucanase	Carbon cycling	Carbon degradation	0.141	4.99	0.026
exoglucanase	Carbon cycling	Carbon degradation	0.317	4.72	0.031
acetylglucosaminidase	Carbon cycling	Carbon degradation	0.193	9.86	0.002
endochitinase	Carbon cycling	Carbon degradation	0.154	11.47	0.001
ara_fungi	Carbon cycling	Carbon degradation	0.433	13.93	<0.001
xylA	Carbon cycling	Carbon degradation	0.127	7.72	0.006
phenol_oxidase	Carbon cycling	Carbon degradation	0.315	17.58	<0.001
AceA	Carbon cycling	Carbon degradation	0.150	22.35	<0.001
AceB	Carbon cycling	Carbon degradation	0.260	71.88	<0.001
AssA	Carbon cycling	Carbon degradation	1.478	177.03	<0.001
vdh	Carbon cycling	Carbon degradation	0.273	4.44	0.039
amyA	Carbon cycling	Carbon degradation	0.192	20.75	<0.001
isopullulanase	Carbon cycling	Carbon degradation	0.097	4.40	0.048
pulA	Carbon cycling	Carbon degradation	0.344	19.27	<0.001
CODH	Carbon cycling	Carbon fixation	0.121	35.27	<0.001
pmoA	Carbon cycling	Methane	0.244	15.37	<0.001
cytochrome	Energy process	Energy process	0.109	10.40	0.001
Al	Metal resistance	Aluminum	0.450	15.47	<0.001
arsB	Metal resistance	Arsenic	0.291	5.75	0.017
CadA	Metal resistance	Cadmium	0.200	12.40	<0.001
czcA	Metal resistance	Cadmium, cobalt, zinc	0.105	6.94	0.009
ChrA	Metal resistance	Chromium	0.227	43.45	<0.001
CopA	Metal resistance	Copper	0.149	29.12	<0.001
CueO	Metal resistance	Copper	0.456	5.92	0.018
CusF	Metal resistance	Copper	1.603	5.79	0.035
mer	Metal resistance	Mercury	0.217	7.26	0.007
merT	Metal resistance	Mercury	1.706	28.74	<0.001
silP	Metal resistance	Silver	0.392	19.74	<0.001
TerC	Metal resistance	Tellurium	0.262	8.56	0.004
TerZ	Metal resistance	Tellurium	0.306	25.58	<0.001
ZntA	Metal resistance	Zinc	0.108	11.20	0.001
ureC	Nitrogen	Ammonification	0.084	12.63	<0.001
hzo	Nitrogen	Anammox	0.294	10.40	0.002
nasA	Nitrogen	Assimilatory N reduction	0.259	4.57	0.033
nirA	Nitrogen	Assimilatory N reduction	0.349	34.69	<0.001
NirB	Nitrogen	Assimilatory N reduction	0.286	5.04	0.026
narG	Nitrogen	Denitrification	0.125	7.09	0.008
nirK	Nitrogen	Denitrification	0.067	4.06	0.044
nirS	Nitrogen	Denitrification	0.102	4.38	0.037
norB	Nitrogen	Denitrification	0.299	12.39	0.001
nosZ	Nitrogen	Denitrification	0.097	15.10	<0.001
nrfA	Nitrogen	Dissimilatory N reduction	0.208	12.22	0.001
amoA	Nitrogen	Nitrification	0.103	4.30	0.039
nifH	Nitrogen	Nitrogen fixation	0.247	71.28	<0.001
BpH	Organic remediation	Aromatics	0.370	4.70	0.033
GcdB	Organic remediation	Aromatics	0.516	16.44	<0.001
GCoADH	Organic remediation	Aromatics	0.172	15.72	<0.001
hmgA	Organic remediation	Aromatics	0.097	5.88	0.016
hmgB	Organic remediation	Aromatics	0.136	7.88	0.005
hmgC	Organic remediation	Aromatics	0.170	10.63	0.001
mdlA	Organic remediation	Aromatics	0.300	29.53	<0.001
mdlB	Organic remediation	Aromatics	0.817	10.47	0.002
mdlC	Organic remediation	Aromatics	0.182	11.49	0.001

Table S3. Cont.

Gene	Gene category	Subcategory	Abundance differences between days 269 and 0	*F* value	Pr (>*F*)
mhpA	Organic remediation	Aromatics	0.048	3.98	0.050
mhpB	Organic remediation	Aromatics	0.334	12.33	0.002
nagG	Organic remediation	Aromatics	0.247	16.14	<0.001
nagI	Organic remediation	Aromatics	0.172	5.22	0.023
nagK	Organic remediation	Aromatics	0.180	9.70	0.003
nagL	Organic remediation	Aromatics	0.427	10.89	0.002
ohbAB	Organic remediation	Aromatics	0.255	16.54	<0.001
phtA	Organic remediation	Aromatics	0.295	15.73	<0.001
pimF	Organic remediation	Aromatics	0.216	56.59	<0.001
PobA	Organic remediation	Aromatics	0.065	4.43	0.036
akbF	Organic remediation	Aromatics	0.167	6.03	0.015
catB	Organic remediation	Aromatics	0.159	12.53	<0.001
todC	Organic remediation	Aromatics	0.164	8.31	0.005
tutFDG	Organic remediation	Aromatics	0.365	7.41	0.007
fcbA	Organic remediation	Aromatics	0.844	16.36	0.001
tfdA	Organic remediation	Aromatics	0.196	16.87	<0.001
tfdB	Organic remediation	Aromatics	0.648	5.20	0.027
tftH	Organic remediation	Aromatics	0.149	10.00	0.002
mhqA	Organic remediation	Aromatics	0.283	5.87	0.034
nbzB	Organic remediation	Aromatics	0.164	10.11	0.003
nhh	Organic remediation	Aromatics	0.131	5.50	0.019
nphA	Organic remediation	Aromatics	1.521	4.39	0.047
AmiE	Organic remediation	Aromatics	0.215	6.31	0.014
Arylest	Organic remediation	Aromatics	0.322	5.94	0.016
BADH	Organic remediation	Aromatics	0.277	8.85	0.003
Catechol	Organic remediation	Aromatics	0.116	8.40	0.004
catechol_B	Organic remediation	Aromatics	0.122	5.67	0.019
nitA	Organic remediation	Aromatics	0.184	18.53	<0.001
pcaG	Organic remediation	Aromatics	0.163	13.99	<0.001
pheA	Organic remediation	Aromatics	0.175	7.29	0.007
bphC	Organic remediation	Aromatics	0.207	5.25	0.023
bphD	Organic remediation	Aromatics	0.199	6.40	0.016
nahF	Organic remediation	Aromatics	0.438	7.73	0.008
phdA	Organic remediation	Aromatics	0.618	20.83	<0.001
phdCI	Organic remediation	Aromatics	0.272	15.29	<0.001
phdJ	Organic remediation	Aromatics	0.760	31.76	<0.001
qorL	Organic remediation	Aromatics	0.434	14.75	0.001
cmuA	Organic remediation	Chlorinated solvents	0.621	4.88	0.034
dehH109	Organic remediation	Chlorinated solvents	0.099	6.97	0.009
exaA	Organic remediation	Chlorinated solvents	0.208	18.50	<0.001
rd	Organic remediation	Chlorinated solvents	0.274	5.94	0.015
mauAB	Organic remediation	Herbicides related compound	0.247	9.17	0.003
pcpB	Organic remediation	Herbicides related compound	0.535	11.79	0.001
pcpE	Organic remediation	Herbicides related compound	0.253	8.42	0.004
phn	Organic remediation	Herbicides related compound	0.129	7.39	0.007
trzA	Organic remediation	Herbicides related compound	0.153	6.23	0.014
trzN	Organic remediation	Herbicides related compound	0.272	10.03	0.002
alkB	Organic remediation	Other Hydrocarbons	0.314	20.22	<0.001
alkH	Organic remediation	Other Hydrocarbons	0.473	20.26	<0.001
chnB	Organic remediation	Other Hydrocarbons	0.154	7.11	0.008
alkK	Organic remediation	Others	0.219	10.77	0.001
dmsA	Organic remediation	Others	0.422	10.72	0.001
linB	Organic remediation	Pesticides related compound	0.194	11.19	0.001
linC	Organic remediation	Pesticides related compound	0.320	6.12	0.014
gyrB	Other category	Phylogenetic marker	0.113	8.84	0.003
ppk	Phosphorus	Phosphorus utilization	0.244	10.31	0.001
ppx	Phosphorus	Phosphorus utilization	0.247	30.84	<0.001
AprA	Sulfur	Other	0.285	12.47	<0.001
dsrA	Sulfur	Sulfite reductase	0.287	70.76	<0.001
dsrB	Sulfur	Sulfite reductase	0.340	48.99	<0.001
sox	Sulfur	Sulfur oxidation	0.231	14.47	<0.001

The abundance of a gene was the average of mean-signal-ratio across different probes within this gene. The mean-signal-ratio is the ratio of the signal intensity of a probe divided by the mean signal intensity across all probes detected.