Network succession reveals the importance of competition in response to emulsified vegetable oil amendment for uranium bioremediation

Ye Deng,[1,2] Ping Zhang,[2] Yujia Qin,[2] Qichao Tu,[2] Yunfeng Yang,[3] Zhili He,[2] Christopher Warren Schadt[4] and Jizhong Zhou[2,3,5]*

[1]CAS Key Laboratory of Environmental Biotechnology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences (CAS), Beijing, China.
[2]Institute for Environmental Genomics and Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK, USA.
[3]State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing, China.
[4]BioSciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA.
[5]Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.

Summary

Discerning network interactions among different species/populations in microbial communities has evoked substantial interests in recent years, but little information is available about temporal dynamics of microbial network interactions in response to environmental perturbations. Here, we modified the random matrix theory-based network approach to discern network succession in groundwater microbial communities in response to emulsified vegetable oil (EVO) amendment for uranium bioremediation. Groundwater microbial communities from one control and seven monitor wells were analysed with a functional gene array (GeoChip 3.0), and functional molecular ecological networks (fMENs) at different time points were reconstructed. Our results showed that the network interactions were dramatically altered by EVO amendment. Dynamic and resilient succession was evident: fairly simple at the initial stage (Day 0), increasingly complex at the middle period (Days 4, 17, 31), most complex at Day 80, and then decreasingly complex at a later stage (140–269

days). Unlike previous studies in other habitats, negative interactions predominated in a time-series fMEN, suggesting strong competition among different microbial species in the groundwater systems after EVO injection. Particularly, several keystone sulfate-reducing bacteria showed strong negative interactions with their network neighbours. These results provide mechanistic understanding of the decreased phylogenetic diversity during environmental perturbations.

Introduction

Uranium (U) bearing waste is a critical pollutant at U mining and processing sites worldwide (Achtman and Wagner, 2008). Highly soluble U(VI) may migrate and spread into groundwater systems, imposing serious problems for human health and the environment. Previous studies showed that a variety of microorganisms, including metal-reducing bacteria, sulfate-reducing bacteria (SRB) and certain fermentative bacteria, have the ability to reduce soluble U(VI) to insoluble U(IV), resulting in immobilization of U *in situ* in the subsurface environment and consequently prevention of further U contamination in groundwater systems (Wu *et al.*, 2006a,b; Gihring *et al.*, 2011). To stimulate U immobilization, rapidly consumed substrates (e.g. acetate, ethanol) were injected as electron donors and carbon sources for enhancing microbial activity (Anderson *et al.*, 2003; Wu *et al.*, 2006b; Williams *et al.*, 2011). Since rapidly consumed substrates required multiple and periodic injections, more recently slow-releasing substrates (sparingly soluble and metabolized) like emulsified vegetable oil (EVO) (Gihring *et al.*, 2011; Tang *et al.*, 2013a,b) have also been tested for stimulating U(VI) reduction in a field experiment (Gihring *et al.*, 2011; Tang *et al.*, 2013a,b). The experimental results showed that one-time EVO injection significantly decreased U concentrations in groundwater for up to 4 months, suggesting that EVO injection can be a cost-effective, sustainable approach for bioremediation of U-contaminated sites (Tang *et al.*, 2013a).

Analysis of microbial diversity through pyrosequencing and quantitative polymerase chain reaction of 16S rRNA gene demonstrated that although total bacterial biomass dramatically increased after EVO injection, both bacterial

community richness and diversity rapidly declined (Gihring *et al.*, 2011). It also found that SRB (e.g. genus *Desulforegula et al.*) dominated groundwater microbial communities during EVO amendment. All these results suggested that only few special microbial populations (specialists) had been stimulated by EVO injection, but more generalists were inhibited. Meanwhile, functional gene microarray (GeoChip) analysis of the same groundwater microbial communities revealed that the community structure diverged substantially away from the initial community state after EVO injection but eventually converged on a state similar to pre-EVO injection, showing that the response of groundwater microbial communities to EVO amendment was resilient (Zhou *et al.*, 2014; Zhang *et al.*, 2015). Although such analyses have provided important insights into our understanding of the diversity, composition and structure of groundwater microbial communities in response to EVO amendment across different time points, little could be inferred quantitatively about the interactions among different microbial species/populations.

In natural habitats, microorganisms live together within complicated networks through various types of interactions, which could be either positive (e.g. mutualism) or negative (e.g. competition) (Faust and Raes, 2012). Such interactions can be depicted as a network model, in which each node represents a species, the edge linking two nodes represents the relationship between these two species, the edge weights represent the strength of the relationship, the arrow direction represents the temporal, regulatory or causal relationship from one species to another, and the size of nodes represents the abundance of species or the node properties. In the past two decades, ecological networks such as food webs and the mutualistic networks between plants and animal pollinators or seed dispersers have been intensively studied in macro-ecology because most of these interactions could be observed directly (Montoya *et al.*, 2006; Bascompte and Jordano, 2007). However, comparable network analyses in microbial ecology are only at infancy with the recent advances in metagenomic technologies (Zhou *et al.*, 2010; 2011).

The study of interactions among different microbial taxa is much more challenging than those in macro-ecology. Most of these interactions could not be directly observed largely due to their extremely high diversity and uncultivated status. Recently developed metagenomic technologies offer an unprecedented opportunity to examine the interactions among different microbial taxa. Based on metagenomics data (i.e. sequencing and microarrays) across many replicate samples, pairwise correlations between different taxa (genes, populations, Operational Taxa Units (OTUs)) can be obtained, which then were used to infer co-occurrence networks (Faust and Raes, 2012). The correlation-based asso-

ciation (or relevance) network method is one of the most commonly used network inference approaches in genomic biology. Many studies have shown that highly correlated proteins were often, but not always, physically interacting (Ge *et al.*, 2001; Dezso *et al.*, 2003; Barabasi and Oltvai, 2004). Among several correlation-based relevance network inference approaches, the random matrix theory (RMT)-based network method is one of the most commonly used approaches. It is advantageous in that the network inference is automatically implemented mathematically and robust to noise, which is an inherent problem associated with high-throughput metagenomic data (Luo *et al.*, 2007; Zhou *et al.*, 2010; 2011). The RMT has been used to delineate the effects of environmental variables on soil microbial community interactions. For instance, previous studies showed that an important environmental parameter of elevated carbon dioxide could dramatically alter the network interactions in terms of microbial functional or phylogenetic groups (Zhou *et al.*, 2010; 2011; Deng *et al.*, 2012a), and such shifts in network structure are also significantly correlated with soil properties.

Here, we described a time-series, RMT-based network inference approach to address the following questions: (i) How do different microbial populations interact with each other during the succession of the groundwater microbial communities? (ii) What are the changes of key populations over time in response to EVO amendment? and (iii) Are these changes of network interactions associated with important environmental chemistries? To this end, high-throughput, time-series GeoChip hybridization data from microbial community analysis during EVO amendment (Zhou *et al.*, 2014; Zhang *et al.*, 2015) were employed to construct these functional molecular ecological networks (fMENs). A modified RMT-based approach via incorporating time-series information was developed for determining microbial interaction networks. Our results suggested that microbial interactions were stimulated by EVO amendment, and the competition among microbial species resulted in the decline of richness and diversity of groundwater microbial communities in response to EVO amendment.

Results

Modified network construction method based on time-series data

To generate time-lagging networks, we modified the regular approach for reconstructing RMT-based networks. As described in *Experimental procedures*, we shifted the microbial abundance profile one time point forward and backward for calculating the Pearson correlation coefficient (r) between any two microorganisms (Fig. S3). Theoretically, if either r value from forward or backward abundance profiles is higher than the regular correlation

(Fig. S3B and C), the temporal order direction between these two microorganisms can be deduced based on the assumption that the abundance of the former appeared microorganism is changed earlier than the later appeared microorganism. This modified approach transformed the network to be directional.

The overall features of time-series network construction and topology

To discern the network interactions of the groundwater microbial communities, GeoChip-based time-series data were analysed by our modified approach. The time-lagging, directional fMENs had clear general network properties, such as scale-free, small-world and modularity (Table 1), consistent with our previous studies (Zhou *et al.*, 2010; 2011; Deng *et al.*, 2012a). The network topology fitted the power law distribution very well ($R^2 = 0.98$), meaning that fewer microorganisms in the fMEN had many more connections while most microorganisms had only few connections. The harmonic geodesic distance (HD) value of 5.06 was close to the logarithm of the total number of network nodes, suggesting that the fMEN had the typical property of small world. The modularity value ($M = 0.625$) was significantly higher than the M value from the corresponding randomized networks. Therefore, this fMEN appeared to be modular.

There were 348 microbial taxa with at least one connection in the fMEN. A total of 1435 edges were identified, including 948 (66%) positive and 487 (34%) negative interactions (Table S1). Notably, the percentage of negative correlations (34%) in this study was much higher comparing with our previous observations (9.2–14%) (Zhou *et al.*, 2010; 2011; Deng *et al.*, 2012a). In addition,

there were 458 edges with directions and 977 edges without directions. Among 458 directional edges, the majority (432 or 88.7%) were negative interactions.

Modularity and module eigengene analysis

The high modularity value (0.625) indicated that fMEN had an evident modular architecture, resulting in 29 modules by a simulated annealing approach (Guimera and Amaral, 2005). Among these, nine modules contained more than five nodes, with 97 nodes in the largest module (Fig. S4). To reveal higher order organization of fMENs, the eigengene analysis (Langfelder and Horvath, 2007; Horvath and Dong, 2008; Oldham *et al.*, 2008) was implemented. In this approach, each module had been decomposed into a single representative abundance profile, which is referred to as the module eigengene. The eigengene network analysis for Module 2 was illustrated in Fig. 1, which consists of a heat map showing abundance of all members within a module (Fig. 1A), abundance represented by singular value decomposition (SVD) analysis (Fig. 1B), module memberships (Fig. 1C) and network interaction map (Fig. 1D). The module eigengenes explained 37–89% variations of the abundance profiles of various modules across different wells in multiple time points (Fig. 1, Fig. S5I–VIII). Most eigengenes (6/9) explained more than 70% of the observed variations, which was considerably higher than our previous observations and the human eigengene network analysis (Dong and Horvath, 2007; Zhou *et al.*, 2011). In addition, the groundwater fMEN had two major eigengene clusters (Fig. S6): modules 1–4 formed a big cluster with almost 77% nodes in the whole network, while modules 5–9 formed the other cluster.

Table 1. Topological properties of the empirical functional molecular ecological networks (fMENs) during U bioremediation in comparison to the random networks.[a]

Datasets	Directional	Similarity threshold (s_t)	Network size (n)	R^2 of power law	Average connectivity (avgK)	Harmonic geodesic distance (HD)	Average clustering coefficient (avgCC)	Modularity (no. of modules)	Harmonic geodesic distance (HD)	Average clustering coefficient (avgCC)	Modularity (M)
									Random networks[b]		
Entire	Yes	0.870	348	0.98	8.247	5.060	0.394	0.625 (29)	2.809 ± 0.045	0.101 ± 0.008	0.355 ± 0.012
Day 0 (+W8)	No	0.974	206	0.98	2.175	30.016	0.179	0.854 (47)	7.142 ± 0.432	0.009 ± 0.005	0.421 ± 0.033
Day 4	No	0.974	401	0.97	5.292	11.998	0.333	0.768 (41)	3.520 ± 0.063	0.024 ± 0.004	0.401 ± 0.022
Day 17	No	0.974	559	0.99	6.537	13.756	0.271	0.553 (92)	3.406 ± 0.078	0.062 ± 0.006	0.378 ± 0.019
Day 31	No	0.974	565	0.98	9.526	9.310	0.355	0.550 (65)	3.005 ± 0.045	0.074 ± 0.006	0.396 ± 0.021
Day 80	No	0.974	490	0.97	21.069	5.435	0.406	0.367 (47)	2.512 ± 0.030	0.200 ± 0.007	0.159 ± 0.008
Day 140	No	0.974	533	0.94	5.373	8.020	0.338	0.738 (48)	3.583 ± 0.051	0.021 ± 0.003	0.345 ± 0.024
Day 269	No	0.974	459	0.97	4.135	8.580	0.272	0.710 (62)	4.067 ± 0.079	0.016 ± 0.003	0.412 ± 0.026

a. Various parameters of the empirical networks and generation of random networks are explained in Deng and colleagues (2012a).
b. The parameters of random networks were generated from 100 times of randomly rewired networks. The parameters presented here were the mean values and standard derivations from these 100 random networks.
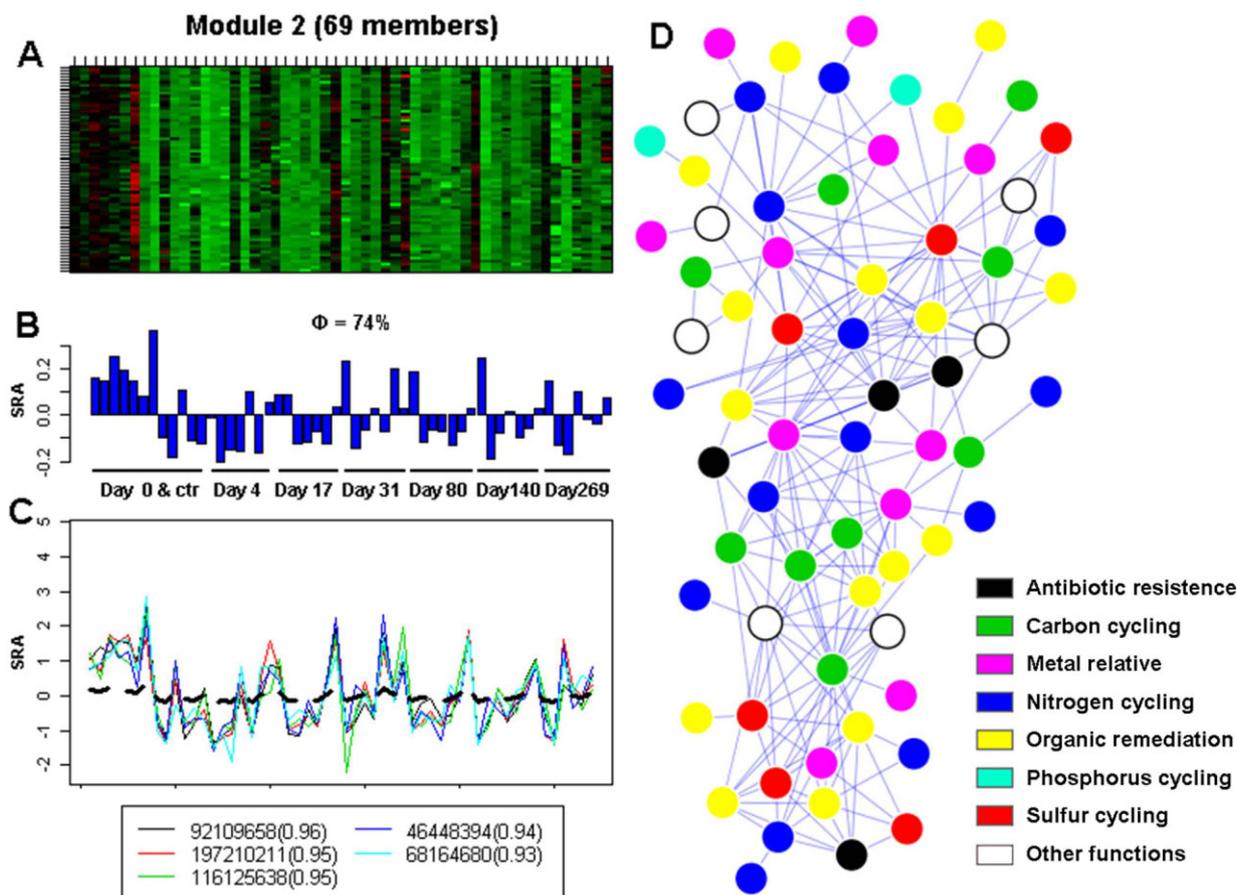
Fig. 1. A conceptual example of the eigengene network analysis on Module 2 of the entire fMEN of groundwater microbial communities.
A. Heat map of standardized relative abundance (SRA) of functional microorganisms across different samples. Rows correspond to individual organisms in the module, whereas columns are the samples. Red corresponds to the OTUs whose SRAs are more than 0, and green signifies whose SRAs are less than 0.
B. The corresponding eigengene (*y* axis) across the samples (*x* axis). The parameter indicates the percentage of the total variance explained by the eigengene.
C. Five nodes with the highest module membership values in this module. The values in parentheses are module memberships to identify nodes that could be well represented by corresponding module eigengene. The *y* axis is SRAs and the *x* axis is individual samples.
D. Module visualization showing the interactions among different nodes within the module. The different colours of nodes represent different functional groups.

Identification of keystone functional populations

Based on within-module connectivity (Z_i) and among-module connectivity (P_i) (Olesen *et al.*, 2007), seven key module hubs (Fig. 2A, Table 2) were observed, but no network hubs or connectors were identified (Fig. S7). Those seven module hubs were highly connected to many other nodes within their own modules; thus, they could be regarded as central nodes in the network. Three module hub microorganisms (46307850, 88062432 and 63029733) were all not-yet cultivated SRB with *dsrA* encoding sulfite reductase. The changes of these three sulfate-reducing microorganisms in abundance had strong correlations with the changes of acetate, nitrate, sulfate, iron and/or U(VI) (Table 2). Notably, the module

hub microorganism 92109658, *Nitrobacter hamburgensis*, contained multiple metal resistance genes *czcA*, *cadA* and *copA*, which were reported to be associated with heavy metal (e.g. cadmium, zinc, copper) resistance or transport (Tsai *et al.*, 1992; Solioz and Vulpe, 1996; Goldberg *et al.*, 1999). Although *N. hamburgensis* was mainly characterized in the nitrification process (Starkenburg *et al.*, 2008), no nitrification-related genes in this organism were detected. In contrast, metal resistance genes were frequently detected, suggesting it might participate in heavy metal recycling. In addition, three other module bacteria containing *ppx* encoding polyphosphatase, carbon monoxide dehydrogenase (CODH) or small multidrug resistance (SMR) genes were identified.
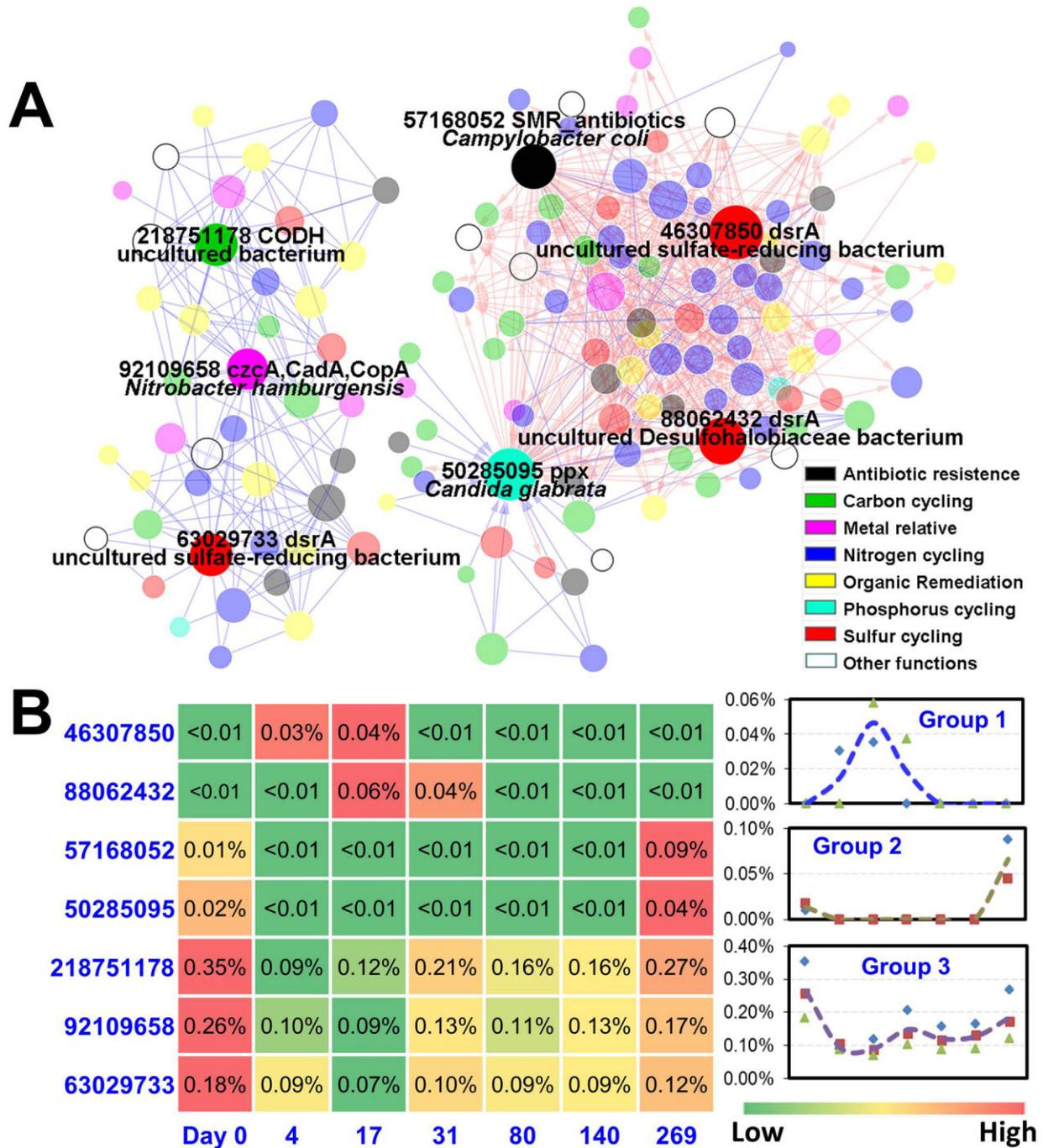
Fig. 2. Seven module hub microorganisms identified from network modular connectivity analysis of the entire fMEN.
A. The interactions of module hub organisms and their neighbours. The colours of nodes represent the ecological functions of microorganisms, and the node sizes represent their connectivity values. The blue lines represent the positive connections, while the red lines represent the negative connections. The arrow direction is assigned from ahead organism pointing to lagged one in time dynamics.
B. The abundance of the hub microorganisms in time points and three general groups can be classified according to their abundance profiles. The colour from red to green represents the abundance of this organism from high to low.

To further investigate the succession of keystone populations, these seven module hubs were classified into three groups based on their abundance profile (Fig. 2B): the first group included two SRB (46307850 and

88062432) carrying *dsrA* genes with high abundance at Day 17; the second group included two bacteria (57168052 and 50285095) carrying SMR and *ppx* genes, respectively, with high abundance detected at latest time

Table 2. The correlations between seven keystone species and important environmental parameters.

Node ID	Functional category	Microorganism	Correlations with environmental parameters				
			Acetate	NO_3^-	SO_4^{2+}	Fe^{2+}	U(VI)
46307850	Sulfur	Uncultured sulfate-reducing bacterium	−0.021	−0.439***	−0.167	0.565***	−0.160
88062432	Sulfur	Uncultured *Desulfohalobiaceae* bacterium	0.243	−0.584***	−0.246	0.497***	−0.071
63029733	Sulfur	Uncultured sulfate-reducing bacterium	−0.395**	0.306*	0.471***	−0.384**	0.360**
92109658	Metal resistance	*Nitrobacter hamburgensis*	−0.369**	0.449***	0.504***	−0.400**	0.330*
57168052	Antibiotic resistance	*Campylobacter coli*	0.000	−0.071	−0.427**	−0.937***	0.514***
50285095	Phosphorus	*Candida glabrata*	0.000	0.244	−0.258	0.217	−0.826***
218751178	Carbon cycling	Uncultured bacterium	−0.299*	0.453***	0.460***	−0.315*	0.247

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

point (Day 269); the third group included three hubs (218751178, 92109658 and 63029733) detected in all time points but with higher abundance before EVO injection and substantially decreased abundance after EVO injection. Interestingly, none of these hubs were from the most abundant microorganisms in the communities. Among about 1000 detected microorganisms, the most abundant microorganisms detected at each time point varied from 0.45% to 0.70%, but the highest relative abundance of hub microorganism 218751178 was 0.35%, which was ranked to the 32nd abundant microorganisms at Day 0, suggesting that keystone microorganisms in the network might not be the most abundant species in the community.

The positive and negative interactions and directions of these keystone microorganisms were further investigated. For example, the hub microorganism 92109658 (the member of Group 3) was positively connected with its 15 neighbours non-directionally (Fig. 3A), suggesting that this hub microorganism interacted with its neighbours synchronously during the U bioremediation process. On the contrary, another hub microorganism 88062432

(uncultured *Desulfohalobiaceae* bacterium, first group member) was negatively connected with its 54 neighbours. Based on the directions, this microorganism negatively influenced 34 neighbours and was negatively by 10 neighbours. The rest of 10 neighbours did not have defined directions, but were still found to be negatively associated with this hub organism. Interestingly, most neighbouring organisms that carried carbon cycling genes as detectable on the arrays were negatively affected by this module hub, implying that their growths may be restricted by this model hub bacterium.

Network succession of microbial communities after EVO amendment

In order to determine the network succession after EVO amendment, individual fMENs were constructed for each time point. An identical threshold was chosen for all of the fMENs to allow for direct comparisons (Table 1). In general, the constructed fMENs obeyed the principles of scale-free, small-world and modular, although their network sizes varied considerably (Table 1). The R^2 of
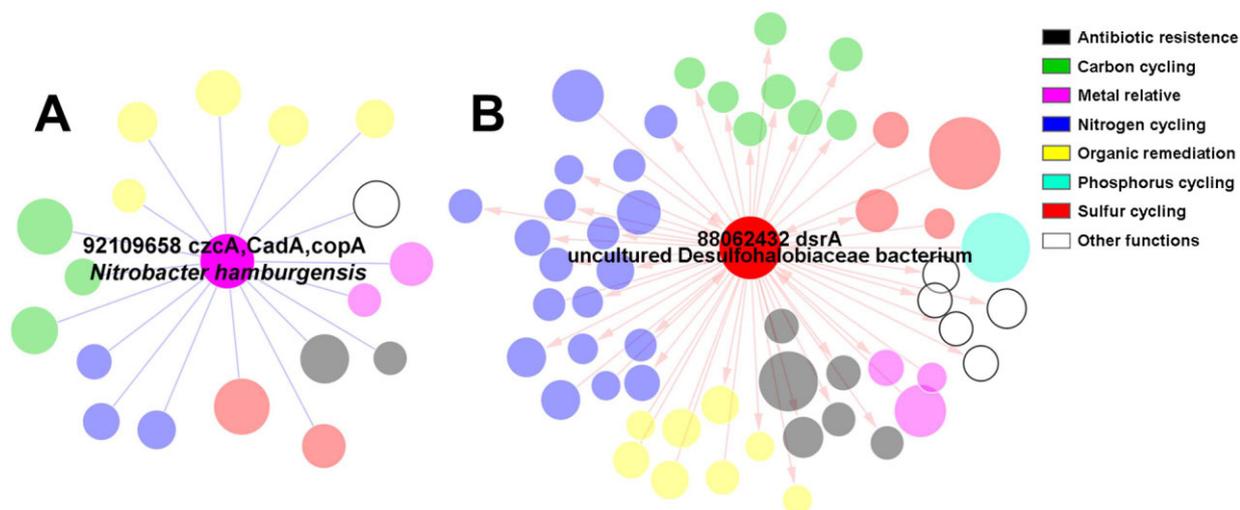

Fig. 3. The interactions between hub microorganism (A) *Nitrobacter hamburgensis* and (B) uncultured *Desulfohalobiaceae* bacterium and their neighbours in the entire fMEN.
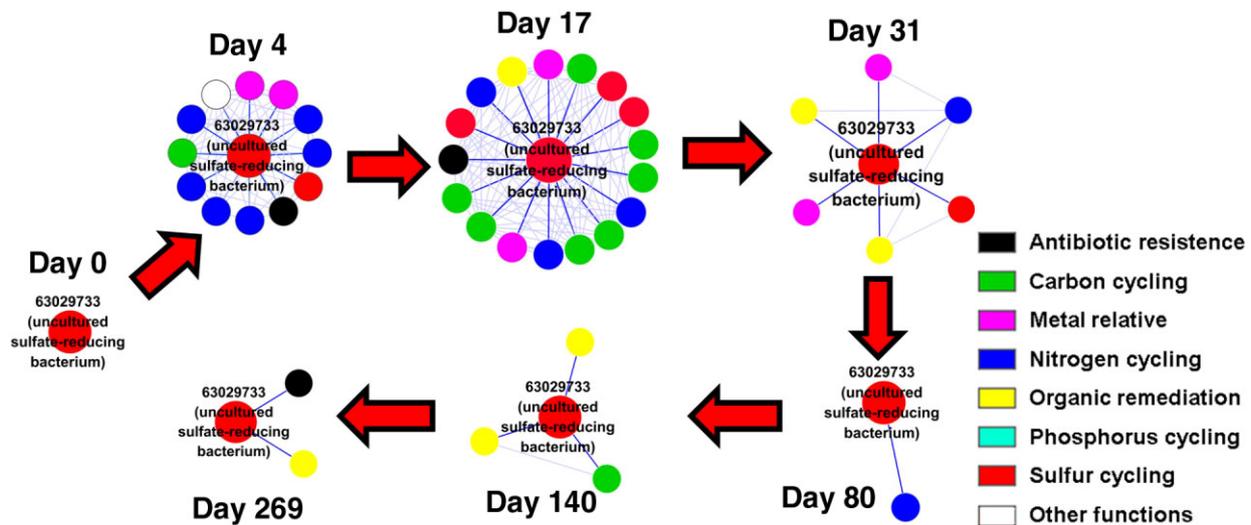
Fig. 4. The dynamic network connections of hub microorganism 63029733 (uncultured sulfate-reducing bacterium) in individual fMENs. The bright blue lines linked the hub organism with its neighbours and the opacified lines linked between neighbours.

power law was consistently high (0.94–0.99), indicating that all individual fMENs were scale-free. Their *HD* and *M* values were significantly higher than the random networks, indicating that small-world and modular properties were also prevalent. Average clustering coefficient (*avgCC*) was used to describe how well nodes were connected with the neighbours (Zhou *et al.*, 2011), which is another parameter to measure the small-world property (Steele *et al.*, 2011). It varied from 0.179 to 0.406, but all were greater than the random *avgCC*, showing the nodes in these networks were more connected than their corresponding random networks of same sizes.

The connections of the module hub microorganisms identified based on the entire network were examined for those time-series fMENs. Only three hub microorganisms (218751178 from uncultured bacterium for C cycling, 92109658 from *N. hamburgensis* for metal resistance and 63029733 for sulfate reduction) were frequently detected in most of fMENs. From their connections with the neighbours (Fig. 4, Fig. S8), the topologies of all three hubs in the networks were dramatically changed across all time points. Although the connections of three hubs were scarce in both the prior and last days, the most complex time points for hub microorganisms were slightly different. For instance, the uncultured SRB 63029733 had no neighbours at Day 0 but quickly gained 12 closest neighbours in Day 4 and reached a maximum of 17 connections in Day 17. After that, its connections gradually decreased (Fig. 4). Similar trends were observed in both 218751178 and 92109658, with a maximum of 55 and 16 connections at Day 80 and Day 4 respectively (Fig. S8). According to the functional genes they carried, three SRB hubs (46307850 uncultured sulfate-reducing bacterium, 88062432 uncultured *Desulfohalobiaceae* bacterium and 63029733 uncul-

tured sulfate-reducing bacterium) and one metal-related hub (92109658 *N. hamburgensis*) should play important roles in U bioremediations. Therefore, they either had higher abundance at the earlier stage after EVO injection (e.g. 46307850 and 88062432 had highest abundances at Day 17, Fig. 2B) or had higher connectivity/activity at the earlier stage after EVO injection (e.g. 63029733 in Fig. 4 and 92109658 in Fig. S8B). On the contrary, the other three hubs that carried functional genes less related to U bioremediation (57168052 *Campylobacter coli*, 50285095 *Candida glabrata* and 218751178 uncultured bacterium) either had higher abundance at the latest time point (e.g. 57168052 and 50285095 in Fig. 2B) or had higher connectivity/activity at the later stage (e.g. 218751178 in Fig. S8A). All these results implied that although the hub microorganisms were substantially affected by the EVO injection in succession, their activities might vary considerably. The U bioremediation-relevant microorganisms could be more abundant and active at the earlier stage after EVO injection.

The network properties of these fMENs showed some similar increased or decreased trends across different time points. The network of Day 0 appeared to have the sparsest connections among network members (Fig. S9A), which was reflected by its biggest value of harmonic geodesic distance (*HD*), the smallest values of average connectivity (*avgK*) and the *avgCC* (Table 1, Fig. S9B). Between Days 4 and 80 after EVO injection, the time-series fMENs became thicker and thicker with increasing network sizes and *avgK* but decreasing *HD* values (Fig. S9). Thereafter, the density of fMENs gradually decreased in terms of *avgK* and *avgCC*. However, the final fMEN (Day 269) still showed higher density than the initial state (Day 0) as indicated by both higher *avgCC* and

Table 3. The Mantel tests on connectivity versus the gene significance of important environmental variables[a] in overall fMEN.

Category	r_M	P
Whole fMEN	0.223	0.001*
Functional categories		
Nitrogen cycling	0.625	0.004*
nifH	−0.047	0.567
ureC	0.640	0.054
napA	−0.009	0.306
amoA	0.086	0.612
nirS	−0.208	0.547
Carbon cycling	0.063	0.637
CODH	0.458	0.121
amyA	−0.198	0.825
Rubisco	−0.172	0.644
AceB	0.544	0.170
Phenol oxidase	−0.297	0.724
Antibiotic resistance	0.113	0.245
Organic remediation	0.165	0.095
catB	−0.255	0.577
linB	−0.311	0.760
alkB	0.955	0.091
Metal resistance	0.580	0.022*
Phosphorus	0.944	0.047*
Sulfur	0.110	0.624
AprA	0.405	0.136
dsrA	0.144	0.089
dsrB	−0.205	0.755
Taxonomic categories		
Bacteria	0.205	0.013*
Proteobacteria	0.246	0.027*
α-	0.595	0.023*
β-	0.117	0.161
δ-	0.122	0.600
γ-	0.202	0.045*
ε-	0.211	0.213
Bacteroidetes	0.499	0.149
Cyanobacteria	0.221	0.967
Firmicutes	0.031	0.497
Actinobacteria	0.037	0.334
Fungi	0.944	0.045*
Archaea	0.100	0.505

a. The important environmental variables used for gene significance calculations included acetate, NO_3^-, SO_4^{2+}, Fe^{2+} and U(VI) concentrations.
*Significance of Mantel test $P < 0.05$.

much lower *HD*, suggesting that microbial interactions were greatly altered until the last time point (Day 269). By considering both *avgCC* and *HD*, which are commonly used to measure the small-world property (Watts and Strogatz, 1998; Watts, 1999; Chow *et al.*, 2014), the fMENs were dynamically changed from fairly simple at the initial stage (Day 0) to more complex in the middle period (Days 4, 17, 31), to the most complex at Day 80 and to less complex again at the late stage (140–269 days).

Relationships of network structure to environmental variables

First, a Pearson correlation was calculated between each environmental variable and each node in the entire fMEN. The *r* values ranged from −0.950 to 0.933 (Table S2) and

about 24.6% of them were significant ($P < 0.01$) by using the *t*-distribution tests without multiple *P* value corrections. However, only 13 of these *r* values were greater than 0.870, which was the minimal threshold for 1435 detected interactions among microorganisms in the fMEN. It suggested that association between available environmental variables and microorganisms was much weaker than the interactions among microorganisms themselves.

To reveal the relationship between network structure and environmental variables, several ecologically important environmental variables were selected. The associations between nodes' connectivity and environmental variables were examined by Mantel test as previously described (Deng *et al.*, 2012a). First, the trait-based gene significance (*GS*) was calculated as the square of Pearson correlation coefficient (r^2) of gene abundance profile with environmental variables. Then, the correlation between multiple *GS* of important environmental variables and nodes' connectivity degrees was analysed by Mantel tests. Our results showed that the nodes' connectivity was significantly correlated with the *GS* of acetate, NO_3^-, SO_4^{2+}, Fe^{2+} and U(VI) concentrations ($r_M = 0.223$, $P = 0.001$) (Table 3). By examining this association between the aforementioned variables and the nodes in each functional category, we found that microorganisms carrying C ($r_M = 0.625$, $P = 0.004$), P ($r_M = 0.944$, $P = 0.047$) and metal resistance ($r_M = 0.580$, $P = 0.022$) functions were significantly correlated. Very strong correlations were observed in both bacteria and fungi, and α- and γ-Proteobacteria within bacteria. Additionally, α- and γ-Proteobacteria and those microorganisms involved in N, P and metal cycling functions were most sensitive to the changes of important environmental variables.

Discussion

Microbes, including bacteria, archaea, fungi, protists and viruses, are the majority of the Earth's biosphere which play indispensable roles in various biogeochemical processes. With recent advances in metagenomic technologies, novel insights in terms of microbial diversity, distributions and responses to environmental changes have been obtained (e.g. Zhou *et al.*, 2008; 2012; 2014; He *et al.*, 2010a). However, the majority of current microbial ecology studies focused on the diversity, composition and structure of microbial communities across space and time, and/or experimental treatments, and hence little is known about the interactions among different microbial populations and with their environments (Zhou *et al.*, 2010; Faust and Raes, 2012; Bissett *et al.*, 2013). Surely, it is difficult to elucidate microbial interactions among various populations and link them to ecosystem processes and functions (Zhou *et al.*, 2010; 2011; Steele

et al., 2011) because the interactions among microbes, in contrast to the situations in plants and animals, have rarely been observed. In this study, we used RMT-based network approaches to examine the succession of co-occurrence networks of groundwater microbial communities in response to the EVO amendment. Our results indicated that EVO injection dramatically shifted the interactions of microbial populations. Particularly, it greatly enhanced competitive interactions. These results provided mechanistic insights on how the groundwater microbial communities respond to environmental perturbations.

One of the biggest advantages of network modelling is to simplify the most complicated relationships among species into one integrated net system (Bascompte, 2007). The interactions of microbial species could be mutualism, competition, predation, parasitism, commensalism and amensalism (Faust and Raes, 2012); however, all of these interactions are extremely hard to be observed in natural conditions. Therefore, multiple network approaches were developed to infer the interactions of species simply based on their co-occurrence (positive) or mutual-exclusion (negative) patterns. In the constructed network, a positive relationship is most likely due to the mutualism or commensalism, while a negative relationship may result from competition, predation, amensalism and so on (Faust and Raes, 2012). But it is also possible that multiple species might interact with each other indirectly via an external environmental driver. Here, we regarded most of negative relationships in our networks as competitions rather than predations because of two reasons. First, our current networks interpreted the relationships only among bacterial species. The typical predator–prey relationships in microbiology were ciliates–bacteria and bacteriophages–bacteria (Pernthaler, 2005), but very few of bacteria–bacteria predations were observed naturally and experimentally (Varon and Zeigler, 1978). Second, our experiment was implemented by providing additional nutrients to natural groundwater system, and thus we expected most negative relationships result from direct food competitions or indirect incompatible niche rather than their direct predations.

Various approaches have been developed and used to examine microbial co-occurrence networks by using both microarray and high-throughput sequencing (HTS) metagenomics technologies (Reshef *et al.*, 2011; Deng *et al.*, 2012a; Faust and Raes, 2012; Faust *et al.*, 2012; Friedman and Alm, 2012; Xia *et al.*, 2013; Kurtz *et al.*, 2015). Except technical differences (Zhou *et al.*, 2015), the data generated by these technologies are substantially different. The raw microarray data are obtained from signal intensities of scanned images, resulting in quantitative, continuous values. Usually, these raw signal intensities are also log-transformed and normalized by multiple steps, including both internal and external standards (Li *et al.*, 2014), and the final normalized signal intensities typically follow the normal distribution (Hoyle *et al.*, 2002). Therefore, Pearson correlations performed fairly well in network inferences of microarray data (Luo *et al.*, 2007; Zhou *et al.*, 2010; Deng *et al.*, 2012a), and it is appropriate to use it in our current study. However, HTS data, such as 16S rRNA gene amplicon sequences, are usually sparse and discrete with large numbers of zeros and positive integers. After transformed into relative fractions, these data are neither independent nor normally distributed. So the improved algorithms based on sparse data could be more appropriate than Pearson relevant approaches in network inferences for HTS datasets (Friedman and Alm, 2012; Kurtz *et al.*, 2015).

In this study, a modified RMT-based approach via incorporating time-lagging information was developed based on time-series data. Our results showed that the modified RMT-based network approach is effective in discerning the network interactions during the succession of the groundwater microbial communities in response to EVO amendment. Obtaining directional information for relevant networks is generally very difficult. Therefore, most of the current networks are non-directional. By considering time-lag information, we are able to assign directions to the constructed network (Fig. 2). Based on time-series datasets, not only ordinary associations but also other potentially time-delayed associations can be inferred since the response of one organism to the change in another organism may exist in time lags (Xia *et al.*, 2011). These time-lagging associations could be expressed as directional connections where the change of one organism leads to a change in another organism in a temporal order. As a result, the network can be transformed from undirected into directed network. The constructed directional networks provide more directions for further experimental testing and validation (Steele *et al.*, 2011; Chow *et al.*, 2014).

Interestingly, consistent with the increase of functional gene richness and diversity as defined by functional probes at the taxonomic resolution of strain and/or species (Zhou *et al.*, 2014; Zhang *et al.*, 2015), microbial network interactions were stimulated by EVO injection, but the overall phylogenetic diversity as defined by pyrosequencing of the V4 region at the taxonomic resolution of genus or higher decreased after EVO injection (Gihring *et al.*, 2011), which was different from previous observations that the overall phylogenetic diversity was stimulated by ethanol amendments (Cardenas *et al.*, 2008; Hwang *et al.*, 2009). Apparently, EVO injection dramatically stimulated a narrow group of bacterial taxa, especially SRB for sulfate reduction. Intuitively, additional carbon sources, especially compounded nutrients such as EVO which is a mixture of vegetable oil,

surfactants, yeast extract and $(NH_4)_3PO_4$, would stimulate the growth of broader groups of microorganisms. Thus, the overall phylogenetic richness and diversity of the microbial communities as a whole are expected to increase after EVO injection. However, it did not hold true in this study. One likely interpretation for the decrease of the overall phylogenetic diversity is that the rapid growth of certain directly stimulated microorganisms (e.g. SRB) would restrain the growth of other broad distantly related microorganisms due to competitive interactions (i.e. negative associations). This is supported by network results revealing a lot of negative interactions, which were much higher than our previous observations with soils, sediments and other groundwater samples (Zhou *et al.*, 2010; 2011; Deng *et al.*, 2012a). In particular, the two module hub microorganisms, SRB 46307850 and 88062432, showed higher abundance after EVO injection (Fig. 2B) with all negative interactions to its neighbours (Fig. 2A). In addition, hub organism 88062432 had 81% (44/54) of directional neighbours (Fig. 3B). Among them, the eight neighbour microorganisms with C cycling genes were all negatively linked. These results indicated that this network hub SRB restrained the growth of its neighbouring organisms. Collectively, all of the above network results suggested the competitions among microbial species were stimulated by EVO injection, and the stimulated microorganisms, especially SRB, restrained the growth of other broad distantly related microbial species, resulting in the decline of phylogenetic richness and diversity of the entire microbial community.

The dominant species were often regarded as functionally important species in a community (Walker *et al.*, 1999; Loreau *et al.*, 2001; Smith and Knapp, 2003). Especially in macro-ecology, dominant taxa were thought to be more functionally important than other taxa (Walker *et al.*, 1999). In HTS studies, abundant taxa and their distributions solicited even more attentions in both human microbiomes (Huttenhower *et al.*, 2012) and environmental microbiomes (Nealson and Venter, 2007; DeLong, 2009) since the dominant species in a habitat could contribute more to ecosystem function or performance for a long run (Walker *et al.*, 1999). However, when a sudden perturbation occurred, dominant and rare species might temporally 'switch' their positions to cope with the perturbation (Walker *et al.*, 1999). Here, seven hub microorganisms have been identified from the entire time-series fMEN, but none of them were the top 50 most abundant species. The most abundant hub microorganism, 218751178, was at the 55th position. Although they were not the most abundant species in this groundwater microbial community, they could play central roles in this short-term period in response to EVO amendment for U bioremediation.

Although the fMENs across different time points had the general topological features of scale-free, small-world and modular architecture as many microbial interaction networks, the properties were dynamically changed during the process of U bioremediation (Table 1 and Fig. S9), which could have important implications for the stability of microbial ecosystems (Barabasi and Oltvai, 2004; Zhou *et al.*, 2010). As indicated by the changes of *avgCC* and *HD* of the whole fMENs, the networks were greatly altered from fairly simple at the initial stage (Day 0) to more complex until Day 80 after EVO injection. When the perturbation gradually vanished, the network structure became simpler again. This result was consistent with our previous studies (Zhou *et al.*, 2010; 2011) showing that elevated atmospheric CO_2 would increase the interactions among microbial species, although the richness and diversity were slightly decreased (Deng *et al.*, 2012b). These experimental results were also in accordance with our general notion that nutrient input may stimulate microbial interactions as well (Zhou *et al.*, 2010; 2011). However, we also noticed the final fMEN on Day 269 was significantly different from that on Day 0 (Fig. S9) in terms of *avgCC* and *HD*, suggesting the EVO amendment fundamentally altered the stability of the groundwater microbial communities as our previous results showed (Zhou *et al.*, 2010; 2011).

In summary, a new RMT-based directional network approach was used to discern network association during the succession of the groundwater microbial communities in response to EVO amendment for uranium bioremediation. Our results showed that the most intense competition among microbial populations occurred immediately following EVO injection. This was reflected by EVO-stimulated microorganisms, especially SRB that restricted the growth of other species, resulting in the decline of richness and diversity of groundwater microbial communities. Our analyses also identified seven potential keystone functional microorganisms, and suggested they may play important roles in U bioremediation even if they were not the most abundant species in the community. Overall, this study offers novel insights and new quantifiable interaction evidence that increases our understanding of the dynamics of microorganisms during the succession of groundwater microbial communities in response to EVO injection.

Experimental procedures

The following is the summary of methods in this study. More details are provided in Appendix S1.

Site description and sampling

A long-term field-scale experiment for U bioremediation with EVO amendment was performed at the U.S. Department of

Energy's Oak Ridge Integrated Field Research Challenge study site, Oak Ridge, TN. A total of 3400 litters of EVO suspension containing 60% (w/w) vegetable oil were injected into three injection wells on 9 February 2009. Groundwater samples were collected from one upstream control well and seven downstream monitor wells at different time points: −28 (labelled as 0 for convenience), 4, 17, 31, 80, 140 and 269 days (Gihring *et al.*, 2011). According to previous groundwater flow experiment and measured environmental variables, the seven monitor wells were spatially adjacent, and their microbial communities were synchronously stimulated by the added vegetable oil but not identical (Zhou *et al.*, 2014; Zhang *et al.*, 2015). These wells were treated as biological replicates for network inferences.

GeoChip data preprocessing for network analysis

GeoChip-based metagenomic technology was used for dissecting microbial community functional structure as previously described (He *et al.*, 2007; 2010b; Zhou *et al.*, 2012). In the present study, except two missing samples, a total of 54 samples were analysed with GeoChip (Zhang *et al.*, 2015). In order to reconstruct the organism-oriented network with functional traits, each detected microorganism was explicitly linked to one of the ecological functions of carbon, nitrogen, phosphorus, sulfur cycling, metal resistance, organic remediation, antibiotic resistance and other categories. The assignment of ecological functions to microbial taxa was performed according to the function of detected genes and existing knowledge from the literature (Fig. S2). For the microorganism with only a single detected gene, its function was assigned by the category that the detected gene belonged to (Fig. S2, Step 1). If multiple genes were detected but all of them were derived from same functional category, its function was assigned by this category and the mean value of these genes was used (Fig. S2, Step 2). For the microorganism with multiple genes in various functional categories, the most relevant function was assigned according to its major function or original habitat based on literature information. If multiple functional genes were detected in a single functional category, the mean value was taken from the normalized signals of these genes (Fig. S2, Step 3). Finally, if there was no literature information showing the function of a microorganism, the micro-organism was assigned to the category with the maximum number of detected genes, and the mean value was calculated from this category (Fig. S2, Step 4). Through this procedure, each microorganism was assigned a representative ecological function, and only a single vector across a set of samples was retained.

The fMEN construction with regular RMT-based approach

The network generated from functional genes of microbial community was defined as functional molecular ecological network (fMEN) (Deng *et al.*, 2012a). In this study, fMEN for each time point was constructed by regular RMT-based network inference approach across seven monitor wells (Luo *et al.*, 2007; Zhou *et al.*, 2010; 2011; Deng *et al.*, 2012a). First, a symmetric correlation matrix was calculated at each sampling time point. The correlation between each of the two detected microorganisms was measured by Pearson correlation coefficient (r value). The correlation matrix was converted into the similarity matrix by taking the absolute values (Horvath and Dong, 2008). Thereafter, a series of thresholds from 0.30 to 1.00 with 0.01 intervals were applied to the matrix, and only the similarity values above certain threshold were kept for calculating matrix eigenvalues. The finest threshold was chosen when the nearest neighbour spacing distribution (NNSD) of eigenvalues followed Poisson distribution, which represents specific and non-random properties of complex systems (Luo *et al.*, 2007; Deng *et al.*, 2012a). In addition, in order to compare network topologies under the same condition, all fMENs of different time points were generated with a uniform threshold that was determined by two criteria: (i) all the NNSD of eigenvalues of the correlation matrices under this threshold still followed the Poisson distribution; and (ii) the threshold should be as low as possible.

The fMEN construction with a time-lagging RMT-based approach

For the entire network across all time points, we have made some modifications on correlation calculation in RMT-based approach. Because one microbial species may have a delayed response to the other species, we allowed one time point shifted when r value was calculated between two organisms. First, one r value was calculated across all 54 samples in an order without time lagging (Fig. S3A). Then two r values for time lagging were calculated by shifting the values to the previous and next time point within each well. If the absolute value of any time lagging r value was higher than the normal r value (Fig. S3B and C), this time-lagging r value was recorded as the similarity between these two populations, and meanwhile the direction was assigned according to the order of the maximum r value. If a maximum r value was derived from a regular correlation, there was no direction assigned (Fig. S3A). If the r value obtained from lagged correlation is greater than the regular one, the arrow of direction is assigned from the microorganism that appeared earlier to the one that appeared later (Fig. S3B and C). The new modified RMT-based network inference method for time-series experiments has been integrated into our network online analysis pipeline at http://ieg2.ou.edu/MENA, which is openly accessible worldwide. All other network topology characterization, module detection, module eigengene analysis, keystone node identification and the association of network topology with gene significance were described in Appendix S1.

References

Achtman, M., and Wagner, M. (2008) Microbial diversity and the genetic nature of microbial species. *Nat Rev Microbiol* **6:** 431–440.

Anderson, R.T., Vrionis, H.A., Ortiz-Bernad, I., Resch, C.T., Long, P.E., Dayvault, R., *et al.* (2003) Stimulating the in situ activity of Geobacter species to remove uranium from the groundwater of a uranium-contaminated aquifer. *Appl Environ Microbiol* **69:** 5884–5891.

Barabasi, A.L., and Oltvai, Z.N. (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet* **5:** 101–U115.

Bascompte, J. (2007) Networks in ecology. *Basic Appl Ecol* **8:** 485–490.

Bascompte, J., and Jordano, P. (2007) Plant-animal mutualistic networks: the architecture of biodiversity. *Annu Rev Ecol Evol Syst* **38:** 567–593.

Bissett, A., Brown, M.V., Siciliano, S.D., and Thrall, P.H. (2013) Microbial community responses to anthropogenically induced environmental change: towards a systems approach. *Ecol Lett* **16:** 128–139.

Cardenas, E., Wu, W.M., Leigh, M.B., Carley, J., Carroll, S., Gentry, T., *et al.* (2008) Microbial communities in contaminated sediments, associated with bioremediation of uranium to submicromolar levels. *Appl Environ Microbiol* **74:** 3718–3729.

Chow, C.E.T., Kim, D.Y., Sachdeva, R., Caron, D.A., and Fuhrman, J.A. (2014) Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *ISME J* **8:** 816–829.

DeLong, E.F. (2009) The microbial ocean from genomes to biomes. *Nature* **459:** 200–206.

Deng, Y., Jiang, Y.H., Yang, Y., He, Z., Luo, F., and Zhou, J. (2012a) Molecular ecological network analyses. *BMC Bioinformatics* **13:** 113.

Deng, Y., He, Z., Xu, M., Qin, Y., Van Nostrand, J.D., Wu, L., *et al.* (2012b) Elevated carbon dioxide alters the structure of soil microbial communities. *Appl Environ Microbiol* **78:** 2991–2995.

Dezso, Z., Oltvai, Z.N., and Barabasi, A.L. (2003) Bioinformatics analysis of experimentally determined protein complexes in the yeast *Saccharomyces cerevisiae*. *Genome Res* **13:** 2450–2454.

Dong, J., and Horvath, S. (2007) Understanding network concepts in modules. *BMC Syst Biol* **1:** 24.

Faust, K., and Raes, J. (2012) Microbial interactions: from networks to models. *Nat Rev Microbiol* **10:** 538–550.

Faust, K., Sathirapongsasuti, J.F., Izard, J., Segata, N., Gevers, D., Raes, J., and Huttenhower, C. (2012) Microbial co-occurrence relationships in the human microbiome. *PLoS Comput Biol* **8:** e1002606.

Friedman, J., and Alm, E.J. (2012) Inferring correlation networks from genomic survey data. *PLoS Comput Biol* **8:** e1002687.

Ge, H., Liu, Z.H., Church, G.M., and Vidal, M. (2001) Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. *Nat Genet* **29:** 482–486.

Gihring, T.M., Zhang, G.X., Brandt, C.C., Brooks, S.C., Campbell, J.H., Carroll, S., *et al.* (2011) A limited microbial consortium is responsible for extended bioreduction of uranium in a contaminated aquifer. *Appl Environ Microbiol* **77:** 5955–5965.

Goldberg, M., Pribyl, T., Juhnke, S., and Nies, D.H. (1999) Energetics and topology of CzcA, a cation/proton antiporter of the resistance-nodulation-cell division protein family. *J Biol Chem* **274:** 26065–26070.

Guimera, R., and Amaral, L.A.N. (2005) Cartography of complex networks: modules and universal roles. *J Stat Mech* **2005:** P02001.

He, Z., Gentry, T.J., Schadt, C.W., Wu, L., Liebich, J., Chong, S.C., *et al.* (2007) GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *ISME J* **1:** 67–77.

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

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

Horvath, S., and Dong, J. (2008) Geometric interpretation of gene coexpression network analysis. *PLoS Comput Biol* **4:** e1000117.

Hoyle, D.C., Rattray, M., Jupp, R., and Brass, A. (2002) Making sense of microarray data distributions. *Bioinformatics* **18:** 576–584.

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J.H., Chinwalla, A.T., *et al.* (2012) Structure, function and diversity of the healthy human microbiome. *Nature* **486:** 207–214.

Hwang, C.C., Wu, W.M., Gentry, T.J., Carley, J., Corbin, G.A., Carroll, S.L., *et al.* (2009) Bacterial community succession during in situ uranium bioremediation: spatial similarities along controlled flow paths. *ISME J* **3:** 47–64.

Kurtz, Z.D., Muller, C.L., Miraldi, E.R., Littman, D.R., Blaser, M.J., and Bonneau, R.A. (2015) Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput Biol* **11:** e1004226.

Langfelder, P., and Horvath, S. (2007) Eigengene networks for studying the relationships between co-expression modules. *BMC Syst Biol* **1:** 54.

Li, Y., He, J.Z., He, Z.L., Zhou, Y., Yuan, M.T., Xu, X., *et al.* (2014) Phylogenetic and functional gene structure shifts of the oral microbiomes in periodontitis patients. *ISME J* **8:** 1879–1891.

Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J.P., Hector, A., *et al.* (2001) Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* **294:** 804–808.

Luo, F., Yang, Y., Zhong, J., Gao, H., Khan, L., Thompson, D.K., and Zhou, J. (2007) Constructing gene co-expression networks and predicting functions of unknown genes by random matrix theory. *BMC Bioinformatics* **8:** 299.

Montoya, J.M., Pimm, S.L., and Sole, R.V. (2006) Ecological networks and their fragility. *Nature* **442:** 259–264.

Nealson, K.H., and Venter, J.C. (2007) Metagenomics and the global ocean survey: what's in it for us, and why should we care? *ISME J* **1:** 185–187.

Oldham, M.C., Konopka, G., Iwamoto, K., Langfelder, P., Kato, T., Horvath, S., and Geschwind, D.H. (2008) Functional organization of the transcriptome in human brain. *Nat Neurosci* **11:** 1271–1282.

Olesen, J.M., Bascompte, J., Dupont, Y.L., and Jordano, P. (2007) The modularity of pollination networks. *Proc Natl Acad Sci USA* **104:** 19891–19896.

Pernthaler, J. (2005) Predation on prokaryotes in the water column and its ecological implications. *Nat Rev Microbiol* **3:** 537–546.

Reshef, D.N., Reshef, Y.A., Finucane, H.K., Grossman, S.R., McVean, G., Turnbaugh, P.J., *et al.* (2011) Detecting novel associations in large data sets. *Science* **334:** 1518–1524.

Smith, M.D., and Knapp, A.K. (2003) Dominant species maintain ecosystem function with non-random species loss. *Ecol Lett* **6:** 509–517.

Solioz, M., and Vulpe, C. (1996) CPx-type ATPases: a class of p-type ATPases that pump heavy metals. *Trends Biochem Sci* **21:** 237–241.

Starkenburg, S.R., Larimer, F.W., Stein, L.Y., Klotz, M.G., Chain, P.S.G., Sayavedra-Soto, L.A., *et al.* (2008) Complete genome sequence of *Nitrobacter hamburgensis* X14 and comparative genomic analysis of species within the genus Nitrobacter. *Appl Environ Microbiol* **74:** 2852–2863.

Steele, J.A., Countway, P.D., Xia, L., Vigil, P.D., Beman, J.M., Kim, D.Y., *et al.* (2011) Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *ISME J* **5:** 1414–1425.

Tang, G.P., Watson, D.B., Wu, W.M., Schadt, C.W., Parker, J.C., and Brooks, S.C. (2013a) U(VI) bioreduction with emulsified vegetable oil as the electron donor – model application to a field test. *Environ Sci Technol* **47:** 3218–3225.

Tang, G.P., Wu, W.M., Watson, D.B., Parker, J.C., Schadt, C.W., Shi, X.Q., and Brooks, S.C. (2013b) U(VI) bioreduction with emulsified vegetable oil as the electron donor – microcosm tests and model development. *Environ Sci Technol* **47:** 3209–3217.

Tsai, K.J., Yoon, K.P., and Lynn, A.R. (1992) ATP-dependent cadmium transport by the cadA cadmium resistance determinant in everted membrane-vesicles of bacillus-subtilis. *J Bacteriol* **174:** 116–121.

Varon, M., and Zeigler, B.P. (1978) Bacterial predator-prey interaction at low prey density. *Appl Environ Microbiol* **36:** 11–17.

Walker, B., Kinzig, A., and Langridge, J. (1999) Plant attribute diversity, resilience, and ecosystem function: the nature and significance of dominant and minor species. *Ecosystems* **2:** 95–113.

Watts, D.J. (1999) *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton, NJ: Princeton University Press.

Watts, D.J., and Strogatz, S.H. (1998) Collective dynamics of 'small-world' networks. *Nature* **393:** 440–442.

Williams, K.H., Long, P.E., Davis, J.A., Wilkins, M.J., N'Guessan, A.L., Steefel, C.I., *et al.* (2011) Acetate availability and its influence on sustainable bioremediation of uranium-contaminated groundwater. *Geomicrobiol J* **28:** 519–539.

Wu, W.M., Carley, J., Fienen, M., Mehlhorn, T., Lowe, K., Nyman, J., *et al.* (2006a) Pilot-scale in situ bioremediation of uranium in a highly contaminated aquifer. 1. Conditioning of a treatment zone. *Environ Sci Technol* **40:** 3978–3985.

Wu, W.M., Carley, J., Gentry, T., Ginder-Vogel, M.A., Fienen, M., Mehlhorn, T., *et al.* (2006b) Pilot-scale in situ bioremediation of uranium in a highly contaminated aquifer. 2. Reduction of U(VI) and geochemical control of U(VI) bioavailability. *Environ Sci Technol* **40:** 3986–3995.

Xia, L.C., Steele, J.A., Cram, J.A., Cardon, Z.G., Simmons, S.L., Vallino, J.J., *et al.* (2011) Extended local similarity analysis (eLSA) of microbial community and other time series data with replicates. *BMC Syst Biol* **5:** 515.

Xia, L.C., Ai, D.M., Cram, J., Fuhrman, J.A., and Sun, F.Z. (2013) Efficient statistical significance approximation for local similarity analysis of high-throughput time series data. *Bioinformatics* **29:** 230–237.

Zhang, P., Wu, W.-M., Van Nostrand, J., Deng, Y., He, Z., Gihring, T., *et al.* (2015) Dynamic succession of groundwater functional microbial communities in response to emulsified vegetable oil amendment during sustained in situ U(VI) reduction. *Appl Environ Microbiol* **81:** 4164–4172. online.

Zhou, J., Kang, S., Schadt, C.W., and Garten, C.T., Jr (2008) Spatial scaling of functional gene diversity across various microbial taxa. *Proc Natl Acad Sci USA* **105:** 7768–7773.

Zhou, J., Deng, Y., Luo, F., He, Z., Tu, Q., and Zhi, X. (2010) Functional molecular ecological networks. *MBio* **1:** e169.

Zhou, J., Deng, Y., Luo, F., He, Z., and Yang, Y. (2011) Phylogenetic molecular ecological network of soil microbial communities in response to elevated CO_2. *MBio* **2:** e00122.

Zhou, J.Z., Xue, K., Xie, J.P., Deng, Y., Wu, L.Y., Cheng, X.H., *et al.* (2012) Microbial mediation of carbon-cycle feedbacks to climate warming. *Nat Clim Change* **2:** 106–110.

Zhou, J., Deng, Y., Zhang, P., Xue, K., Van Nostrand, J., Yang, Y., *et al.* (2014) Stochasticity, succession and environmental perturbations in a fluidic ecosystem. *Proc Natl Acad Sci USA* **111:** E836–E845.

Zhou, J.Z., He, Z.L., Yang, Y.F., Deng, Y., Tringe, S.G., and Alvarez-Cohen, L. (2015) High-throughput metagenomic technologies for complex microbial community analysis: open and closed formats. *MBio* **6:** e02288.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Fig. S1. The concentrations of uranium and sulfate in groundwater samples across different time points. The final sampling points have been marked in red. Both of U and sulfate were dramatically dropped within 31 days after EVO

injection, and then gradually resumed within 269 days. The final U concentrations were declined after bioremediation procedure.

Fig. S2. The steps for function assignment of each GeoChip detected microorganism. The percentage of microorganisms assigned in each step was signified in red.

Fig. S3. The modified step for Pearson correlation coefficient (*r* value) calculation for time-series dataset. The maximal *r* value from three situations was retained for following RMT-based modelling, and meanwhile the directions have been assigned from backward organism pointing to the forward one. (A) The maximal *r* value is derived from the normal unlagged order. (B) The maximal *r* value is derived from the forward lagging. (C) The maximal *r* value is derived from the backward lagging. In each panel, the above curve figure demonstrated the successional changes of two organisms; the middle colorful block figure showed the paired rule of all values in seven replicate wells across seven time points when Pearson correlation was calculated, and the bottom similarity values showed the way to assign direction.

Fig. S4. Modular organization of the fMEN in groundwater. The colours of the nodes indicate different functional categories. Clear modular architecture was observed in this fMEN. Each node signifies a microorganism that carried out certain ecological function detected by GeoChip. A blue line indicates a positive interaction between two individual nodes, while a red line indicates a negative interaction. The numbers indicate different modules determined by the simulated annealing approach. All data showed that the fMEN with time series has a clear modular architecture.

Fig. S5. Eigengene network analysis on all modules of entire fMEN except Module 2. For each figure from I to VIII: (A) Heat map of the standardized relative abundance (SRA) of functional microorganisms across different samples. Rows correspond to individual organisms in the module, whereas columns are the samples. Red corresponds to the OTUs whose SRAs are 0s, and green signifies whose SRAs are 0. (B) The corresponding eigengene (y axis) across the samples (x axis). The parameter indicates the percentage of the total variance explained by the eigengene. (C) Five nodes of highest module membership values in this module. The values in parentheses are module memberships. The module membership values were used to identify the nodes that could be well represented by corresponding module eigengene. The y axis is SRAs and the x axis is individual samples. (D) Module visualization showing the interactions among different nodes within the module. The different colours of nodes represent different functional groups.

Fig. S6. The correlations and heat map to show the relationships among module eigengenes. The right part is the hierarchical clustering based on the Pearson correlations among module eigengenes, and the below heat map shows the coefficient values (*r*). Red colour means higher correlation, whereas green colour signified lower correlation.

Fig. S7. ZP-plot showing distribution of microorganisms based on their module-based topological roles. Each dot represents a microorganism in the dataset. The topological role of each node was determined according to the scatter plot of within-module connectivity (Z) and among-module connectivity (P).

Fig. S8. The network connections of (A) hub microorganism 218751178 (uncultured bacterium) and (B) hub microorganism 92109658 (*Nitrobacter hamburgensis*) in individual fMENs for all of the time points.

Fig. S9. (A) Functional molecular ecological networks for individual time points from Day 0 to Day 269. (B) The biodiversity value (Shannon index) of the community in all time points and the corresponding average connectivity, harmonic geodesic distance and average clustering coefficient from network topologies. Since the scale of clustering coefficient is different with other variables, the values were marked in the dots directly.

Table S1. The identified edges in the entire network by using RMT network inference approach.

Table S2. The Pearson correlation between identified nodes with all available environmental variables. The sheet '*r* values' includes all Pearson correlation coefficient between nodes and environmental variables. The sheet '*P* values' includes all significance, and the *P* values less than 0.05 were marked in red.

Appendix S1. Supplemental materials and methods.