A. Supporting Tables

1. **Table S1.** The network topological indexes used in this study
2. **Table S2.** The topological comparisons of functional molecular ecological networks (fMENs) under aCO_2 and eCO_2
3. **Table S3.** Summary of the network complexity of individual functional genes involved in C, N, P and S cyclings
4. **Table S4.** The connectivity of all shared *nifH* genes detected under both aCO_2 and eCO_2

B. Supporting Figures

1. **Fig. S1.** The scatter plots showing the scale-free property of the MENs under both aCO_2 and eCO_2
2. **Fig. S2.** Modular organization of the fMENs with GeoChip-based metagenomics data
3. **Fig. S3.** Effects of eCO_2 on network interactions of several key functional gene categories under eCO_2 and aCO_2
4. **Fig. S4.** Impacts of eCO_2 on the functional network interactions of key microbial functional genes or populations

C. Supplemental References

A. Supporting Tables

Table S1. The network topological indexes used in this study.

Indexes	**Formula**	**Explanation**	**Note**	**References**
Part I: network indexes for individual nodes				
Connectivity	$k_i = \sum_{j \neq i} a_{ij}$	a_{ij} is the connection strength between nodes i and j.	It is also called node degree. It is the most commonly used concept for describing the topological property of a node in a network.	(1)
Stress centrality	$SC_i = \sum_{jk} \sigma(j,i,k)$	$\sigma(j,i,k)$ is the number of shortest paths between nodes *j* and *k* that pass through node *i*	It is used to describe the number of geodesic paths that pass through the i^{th} node. High Stress node can serve as a broker.	(2)
Betweenness	$B_i = \sum_{jk} \frac{\sigma(j,i,k)}{\sigma(j,k)}$	$\sigma(j,k)$ is the total number of shortest paths between *j* and *k*	It is used to describe the ratio of paths that pass through the i^{th} node. High Betweenness node can serve as a broker similar to stress centrality.	(2)
Eigenvector centrality	$EC_i = \frac{1}{\lambda} \sum_{j \in M(i)} EC_j$	*M(i)* is the set of nodes that are connected to the i^{th} node and λ is a constant eigenvalue.	It is used to describe the degree of a central node that it is connected to other central nodes.	(3)
Clustering coefficient	$CC_i = \frac{2l_i}{k_i'(k_i'-1)}$	l_i is the number of links between neighbors of node *i* and k_i' is the number of neighbors of node *i.*	It describes how well a node is connected with its neighbors. If it is fully connected to its neighbors, the clustering coefficient is 1. A value close to 0 means that there are hardly any connections with its neighbors. It was used to describe hierarchical properties of networks.	(4-5)
Vulnerability	$V_i = \frac{E - E_i}{E}$	*E* is the global efficiency and E_i is the global efficiency after the removal of the node *i* and its entire links.	It measures the decrease of node i on the system performance if node i and all associated links are removed.	(6)
Part II: The overall network topological indexes				
Average connectivity	$avgK = \frac{\sum_{i=1}^{n} k_i}{n}$	k_i is degree of node *i* and *n* is the number of nodes	Higher *avgK* means a more complex network.	(7)
Average geodesic distance	$GD = \frac{1}{n(n-1)} \sum_{i \neq j} d_{ij}$	d_{ij} is the shortest path between node *i* and *j.*	A smaller *GD* means all the nodes in the network are closer.	(7)

Geodesic efficiency	$E = \frac{1}{n(n-1)}\sum_{i \neq j}\frac{1}{d_{ij}}$	all parameters shown above	It is the opposite of *GD*. A higher *E* means that the nodes are closer.	(8)
Harmonic geodesic distance	$HD = \frac{1}{E}$	*E* is geodesic efficiency	The reciprocal of *E*, which is similar to *GD* but more appropriate for disjoint graph.	(8)
Centralization of degree	$CD = \sum_{i=1}^{n}(\max(k) - k_i)$	max(*k*) is the maximal value of all connectivity values and k_i represents the connectivity of i^{th} node. Finally this value is normalized by the theoretical maximum centralization score.	It is close to 1 for a network with star topology and in contrast close to 0 for a network where each node has the same connectivity.	(9)
Centralization of betweenness	$CB = \sum_{i=1}^{n}(\max(B) - B_i)$	max(*B*) is the maximal value of all betweenness values and B_i represents the betweenness of i^{th} node. Finally this value is normalized by the theoretical maximum centralization score.	It is close to 0 for a network where each node has the same betweenness, and the bigger the more difference among all betweenness values.	(9)
Centralization of stress centrality	$CS = \sum_{i=1}^{n}(\max(SC) - SC_i)$	max(*SC*) is the maximal value of all stress centrality values and SC_i represents the stress centrality of i^{th} node. Finally this value is normalized by the theoretical maximum centralization score.	It is close to 0 for a network where each node has the same stress centrality, and the bigger the more difference among all stress centrality values.	(9)
Centralization of eigenvector centrality	$CE = \sum_{i=1}^{n}(\max(EC) - EC_i)$	max(*EC*) is the maximal value of all eigenvector centrality values and EC_i represents the eigenvector centrality of i^{th} node. Finally this value is normalized by the theoretical maximum centralization score.	It is close to 0 for a network where each node has the same eigenvector centrality, and the bigger the more difference among all eigenvector centrality values.	(9)
Density	$D = \frac{l}{l_{\exp}} = \frac{2l}{n(n-1)}$	*l* is the sum of total links and l_{exp} is the number of possible links	It is closely related to the average connectivity.	(9)
Average clustering coefficient	$avgCC = \frac{\sum_{i=1}^{n} CC_i}{n}$	CC_i is the clustering coefficient of node *i*	It is used to measure the extent of module structure present in a network.	(4)

Transitivity	$Trans = \frac{\sum_{i=1}^{n}(2l_i)}{\sum_{i=1}^{n}\left[k_i'(k_i'-1)\right]}$	l_i is the number of links between neighbors of node *i* and k_i' is the number of neighbors of node *i*.	Sometimes it is also called the entire clustering coefficient. It has been shown to be a key structural property in social networks.	(9)
Connectedness	$Con = 1 - \left[\frac{W}{n(n-1)/2}\right]$	*W* is the number of pairs of nodes that are not reachable	It is one of the most important measurements for summarizing hierarchical structures. *Con* is 0 for graph without edges and is 1 for a connected graph.	(10)

Table S2. The topological comparisons of the fMENs under aCO_2 and eCO_2[a]

Index	**aCO_2**	**eCO_2**	**Significance**
Total nodes	184	245	NA
Total links	329	874	NA
Average connectivity (*avgK*)	3.58	7.13	NA
Average clustering coefficient (*avgCC*)	0.10	0.22	p<0.001
Average path (*GD*)	4.21	3.09	p<0.001
Geodesic efficiency (*E*)	0.20	0.27	p<0.001
Harmonic geodesic distance (*HD*)	5.08	3.73	p<0.001
The centralization of degree (*CD*)	0.06	0.14	p<0.001
The centralization of betweenness (*CB*)	0.18	0.09	p<0.001
The centralization of Stress Centrality Score (*CS*)	0.84	1.11	p<0.001
The centralization of Eigen vector centrality scores (*CE*)	0.28	0.24	p<0.001
Modularity	0.65	0.44	p<0.001
Density (*D*)	0.020	0.029	p<0.001
Transitivity (*Trans*)	0.18	0.29	p<0.001
Connectedness (*Con*)	0.80	0.82	p<0.001

[a] Detailed explanation of these indexes is provided in Table S1. For convenient comparison, several key indexes from Table 1 are also included in this Table.

Table S3. Summary of the network complexity of represented individual functional genes involved in C, N, P and S cyclings.

Gene Category	Functional group	Selected enzyme/gene names	Number of Shared nodes (%)	aCO_2			eCO_2		
				Number of all nodes	Average connectivity in shared nodes	Shannon index of connectivity	Number of all nodes	Average connectivity in shared nodes	Shannon index of connectivity
All enzymes/genes			129(43%)	184	3.75	4.879	245	9.11	5.027
Carbon cycling	Carbon degradation	Endochitinase (*chi*)	8 (57%)	10	3.25	1.885	12	8.25	1.918
		Endoglucanase (*bcsG*)	0 (0%)	0	0.00	0	2	0.00	1
		exochitinase (*chi36*)	1 (100%)	1	1.00	0	1	1.00	0
		Exoglucanase (*exg*)	1 (50%)	1	8.00	0	3	2.50	0.868
		Lignin peroxidase (*lip*)	0 (0%)	2	0.00	0.305	3	0.00	0.898
		Manganese peroxidase (*mnp*)	4 (100%)	4	4.25	1.202	4	11.00	1.121
		Pectinase (*pglA*)	2 (67%)	2	1.00	0.693	3	4.00	0.684
		Phenol oxidase (*phox*)	4 (36%)	7	2.75	1.681	8	2.00	1.955
		Xylanase (*xyn*)	5 (38%)	9	3.80	1.687	9	8.00	1.867
	Carbon fixation	Carbon monoxide dehydrogenase (*CODH*)	2 (40%)	4	4.50	1.137	3	4.50	1.028
		Tetrahydrofolate formylase (*FTHFS*)	0 (0%)	0	0.00	0	1	0.00	1
		Propionyl-CoA carboxylase (*pcc*)	9 (36%)	11	2.11	2.224	23	10.56	2.706
		Rubisco (*rbcL*)	3 (23%)	6	7.67	1.523	10	11.00	1.872
	Methane metabolism	Methyl coenzyme M reductase (*mcrA*)	0 (0%)	0	0.00	0	2	0.00	0.562
		Methane monooxygenase (*pmoA*)	1 (14%)	2	3.00	0.562	6	1.00	1.586
Nitrogen cycling	N fixation	Nitrogenase reductase (*nifH*)	27 (55%)	32	3.07	3.206	44	9.74	3.292
	Denitrification	Nitrite reductase (*nirK*)	6 (27%)	15	4.00	2.402	13	6.50	2.186
		Nitrite reductase (*nirS*)	8 (57%)	8	2.88	1.942	14	6.50	2.258
	Dissimilatory N reduction	c-type cytochrome nitrite reductase (*nrfA*)	4 (57%)	4	7.25	1.356	7	12.25	1.006
Phosphorus cycling		Polyphosphate kinase (*ppk*)	5 (50%)	7	4.20	1.769	8	11.80	1.671
		Exopolyphosphatase (*ppx*)	12 (48%)	19	4.83	2.621	18	11.25	2.363
Sulphur cycling	sulfite reductase	Dissimilatory sulfite reductase (*dsrA*)	13 (34%)	22	5.08	2.66	29	11.54	2.905
	Sulphur oxidation	Sulfite oxidase (*sox*)	14 (54%)	18	3.00	2.637	22	8.50	2.743

Table S4. The connectivity of all shared *nif*H genes detected under both aCO_2 and eCO_2 (sorted by connectivity under eCO_2)

GenBank ID	Organisms	Connectivity at eCO_2	Connectivity at aCO_2
110630622	Uncultured soil bacterium	42	7
76667345	Uncultured nitrogen-fixing bacterium	25	10
89512768	Uncultured nitrogen-fixing bacterium	23	2
10863129	Lactate SRB-Enrichment culture clone HBLac1	18	1
44829093	Uncultured bacterium	18	3
89512880	Uncultured nitrogen-fixing bacterium	18	3
37925835	Uncultured bacterium	17	7
61653195	Uncultured proteobacterium DelRiverFos13D03	16	6
116697525	*Syntrophobacter fumaroxidans* MPOB	15	2
3157524	Unidentified nitrogen-fixing bacteria	11	1
82698269	Uncultured methanogenic archaeon	10	2
158510468	*Candidatus Desulfococcus oleovorans* Hxd3	8	1
139004179	Uncultured nitrogen-fixing bacterium	6	1
99083393	Uncultured bacterium	5	3
89512536	Uncultured nitrogen-fixing bacterium	5	3
110631274	Uncultured soil bacterium	5	3
115519141	*Rhodopseudomonas palustris* BisA53	4	6
37925044	Uncultured nitrogen-fixing bacterium	4	1
46562231	*Desulfovibrio vulgaris* Hildenborough	2	7
73534357	Uncultured bacterium	2	1
139003137	Uncultured nitrogen-fixing bacterium	2	5
139004122	Uncultured nitrogen-fixing bacterium	2	2
138897063	*Geobacillus thermodenitrificans* NG80-2	1	2
73534215	Uncultured bacterium	1	1
70672905	Uncultured nitrogen-fixing bacterium	1	1
3157614	Unidentified nitrogen-fixing bacteria	1	1
3157662	Unidentified nitrogen-fixing bacteria	1	1

B. Supporting Figures

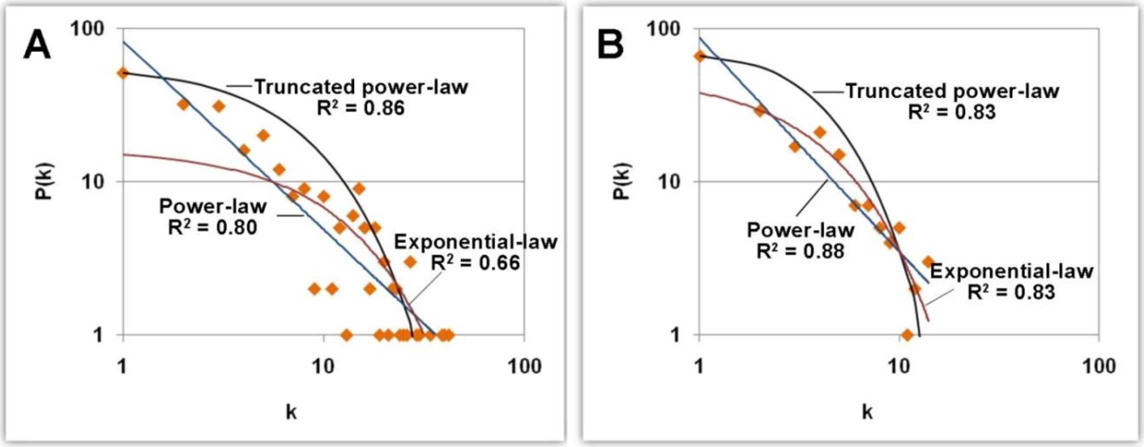


Fig. S1. The scatter plots showing the scale-free property of the fMENs under both aCO_2 and eCO_2. The x-axis is the node connectivity (*k*). The y-axis is the number of nodes under a given connectivity. The values in both axes were log-transformed. Lines and R^2 values are the best fit of the data to the model. (A). fMEN under eCO_2; (B) fMEN under aCO_2.

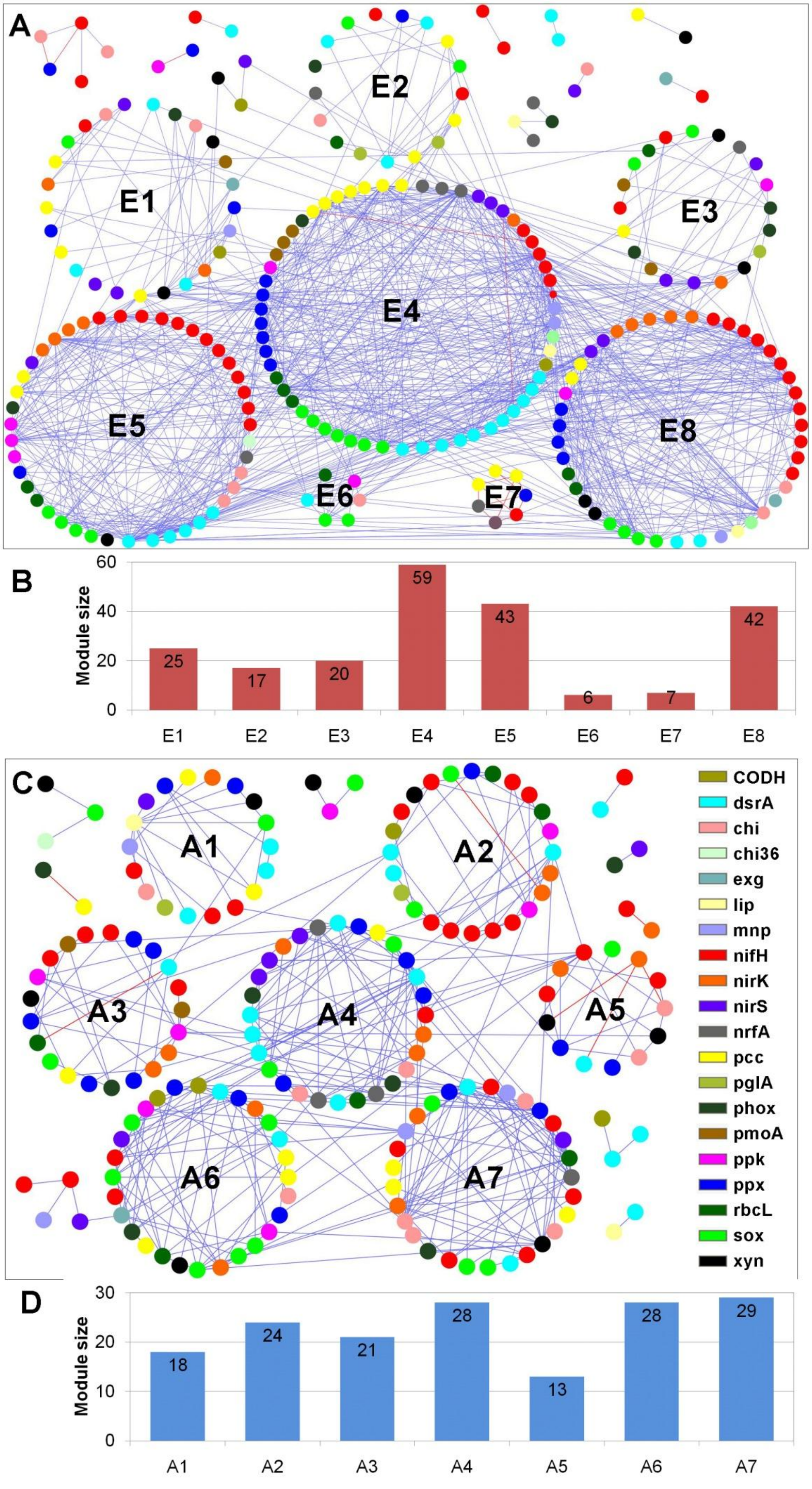

A
E1
E2
E3
E4
E5
E6
E7
E8
B
Module size
60
40
20
0
25
17
20
59
43
6
7
42
E1
E2
E3
E4
E5
E6
E7
E8
C
A1
A2
A3
A4
A5
A6
A7
CODH
dsrA
chi
chi36
exg
lip
mnp
nifH
nirK
nirS
nrfA
pcc
pglA
phox
pmoA
ppk
ppx
rbcL
sox
xyn
D
Module size
30
20
10
0
18
24
21
28
13
28
29
A1
A2
A3
A4
A5
A6
A7

Fig. S2. Modular organization of the fMENs with GeoChip-based metagenomics data. The networks were constructed with the RMT-based approach with the GeoChip data from (A) eCO_2 (12 samples) and (C) aCO_2 (12 samples). Clear modular architecture was observed in this fMEN. Each node signifies a gene, which could correspond to a microbial population. Colors of the nodes indicate different major functional genes. A blue line indicates a positive interaction between two individual nodes, while a red line indicates a negative interaction. The numbers indicate different modules or submodules determined by the fast greedy modularity optimization method. All data showed that the functional MENs have a modular architecture. Besides, the sizes for individual modules were plotted in B (eCO_2) and D (aCO_2).

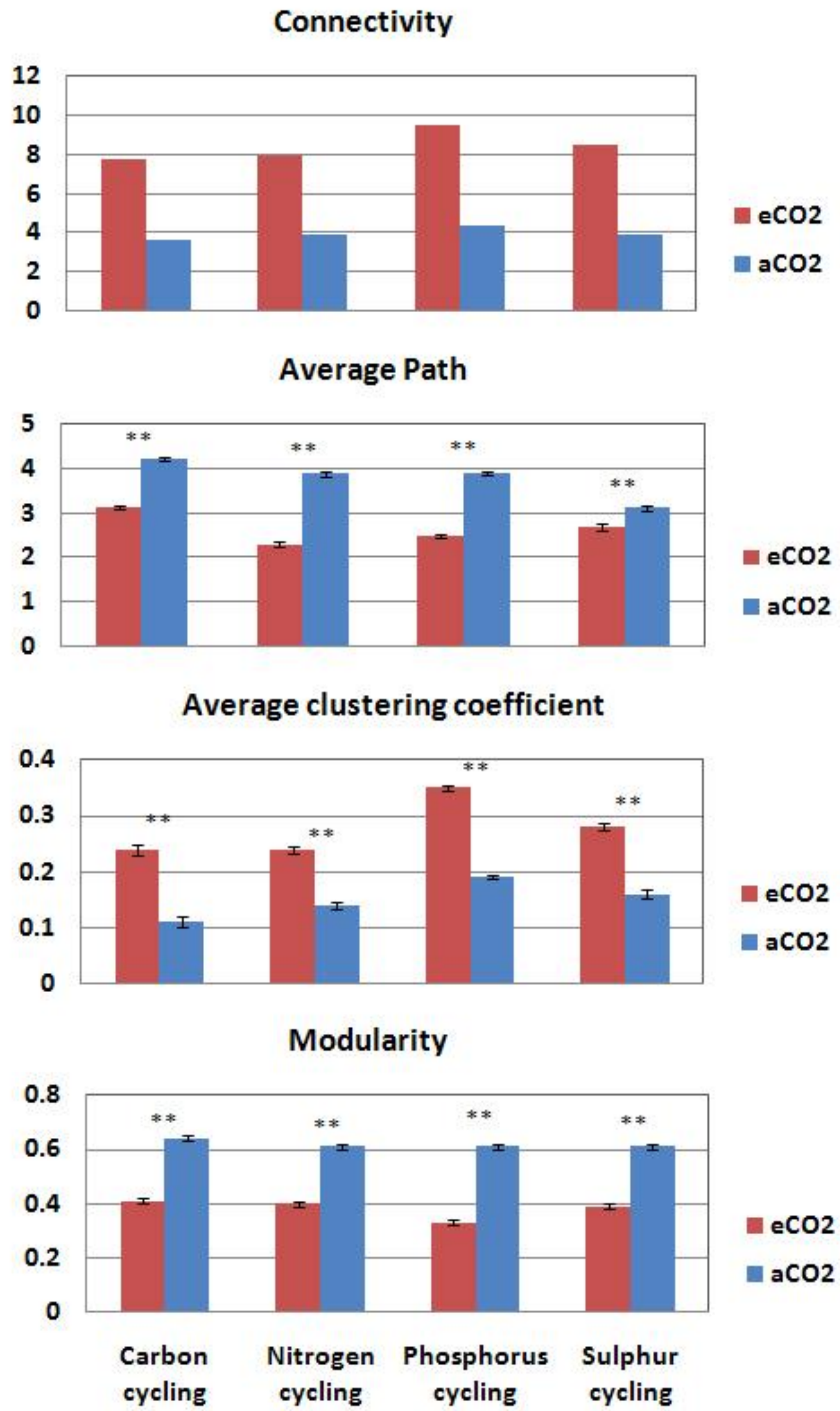
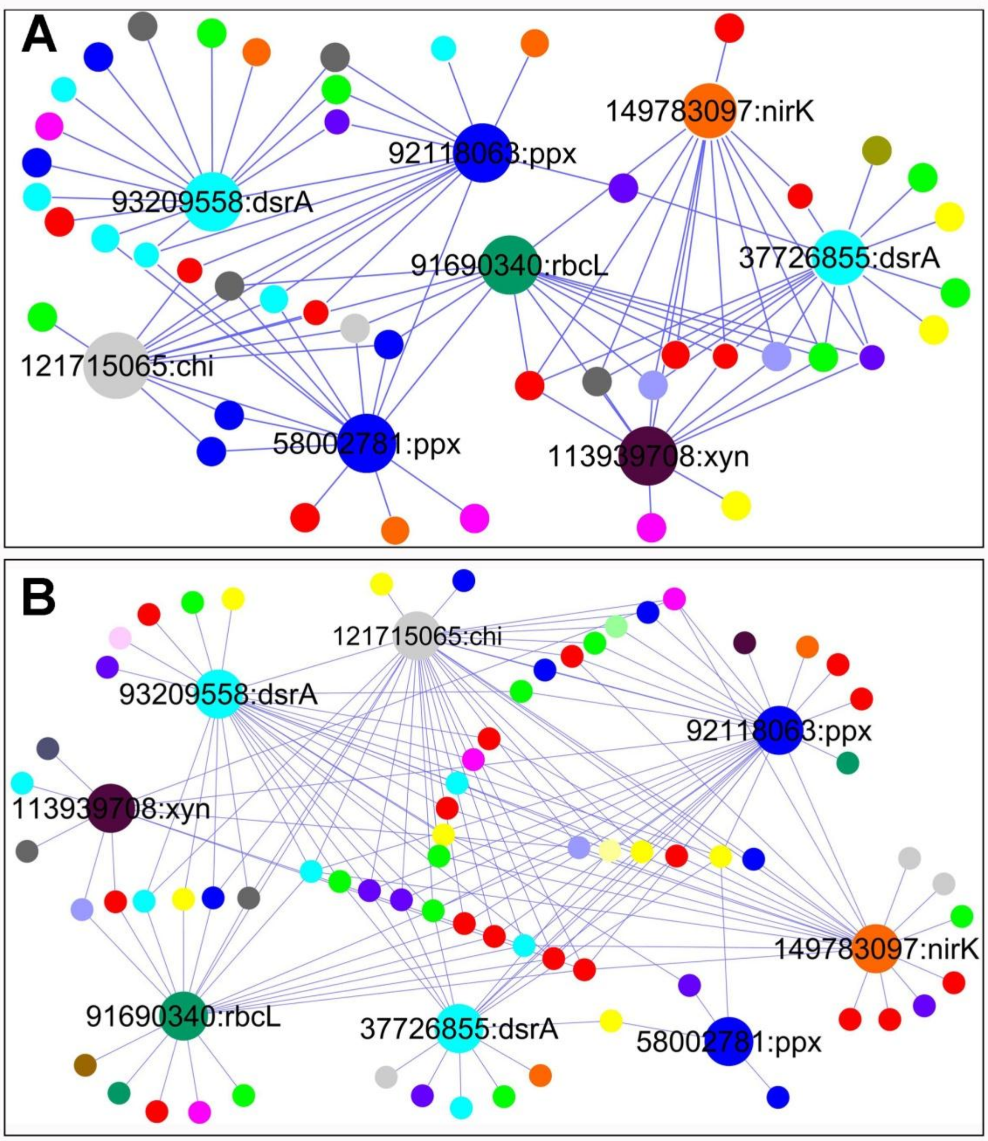


Fig. S3. Effects of eCO_2 on network interactions of several key functional gene categories under eCO_2 (red) and aCO_2 (blue). ** means statistical significance at $p=0.01$ based the standard deviation derived from random network simulation. No standard deviation can be estimated from random network for connectivity because the connectivities are identical between the empirical and random networks.



Communities	Size (n)	Links	Average connectivity (avgK)	Average path (GD)	Average clustering coefficient (avgCC)
aCO2	54	143	5.30	2.58	0.35
eCO2	78	436	11.18	2.22	0.48

CODH	🟫	nrfA	⬛	ppk	🟪	dsrA	🟦	nirK	🟧	sox	🟩
FTHFS	🟪	pcc	🟨	ppx	🟦	mnp	🟪	nirS	🟪	xylA	⬛
chi	⬜	pmoA	🟫	rbcL	🟩	nifH	🟥	xyn	🟪		

Fig. S4. Impacts of eCO$_2$ on the network interactions of key functional genes. (A) Network interactions of the top 8 functional genes with the highest connectivities under aCO$_2$; (B) Network interactions of the corresponding functional genes under eCO$_2$. The networks were constructed by the RMT-based approach with the GeoChip data. The meanings of some symbols were listed in Fig. 2.

C. Supplemental References

1. **Horvath, S., and J. Dong.** 2008. Geometric interpretation of gene coexpression network analysis. PLoS Comput Biol **4:**e1000117.
2. **Brandes, U., and T. Erlebach.** 2005. Network analysis: methodological foundations. Springer-Verlag, Berlin.
3. **Bonacich, P.** 1987. Power and Centrality - a Family of Measures. Am. J. Sociology **92:**1170-1182.
4. **Watts, D. J., and S. H. Strogatz.** 1998. Collective dynamics of 'small-world' networks. Nature **393:**440-442.
5. **Ravasz, E., A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A. L. Barabasi.** 2002. Hierarchical organization of modularity in metabolic networks. Science **297:**1551-1555.
6. **Costa, L. D., F. A. Rodrigues, G. Travieso, and P. R. V. Boas.** 2007. Characterization of complex networks: A survey of measurements. Adv. Phys. **56:**167-242.
7. **West, D. B.** 1996. Introduction to Graph Theory. Prentice Hall, Upper Saddle River, N.J.
8. **Latora, V., and M. Marchiori.** 2001. Efficient behavior of small-world networks. Phys Rev Lett **87:**198701.
9. **Wasserman, S., and K. Faust.** 1994. Social Network Analysis: Methods and applications. Cambridge Univerisity Press, Cambridge.
10. **Krackhardt, D.** 1994. Graph Theoretical Dimensions of Informal Organizations. Lawrence Erlbaum and Associates, Hillsdale, NJ.