Ecosphere

Artificial reforestation produces less diverse soil nitrogen-cycling genes than natural restoration

Yun Wang, Hua Zheng, Falin Chen, Yunfeng Yang, Jing Zeng, Joy D. Van Nostrand, Jizhong Zhou, and Zhiyun Ouyang

Appendix S1

Table S1. Comparisons of structure of tree layer of artificially reforested slash pine plantations (ARSP), artificially reforested Masson pine plantations (ARMP), and natural reforested secondary forests (NRSF)

Reforestation approach	Tree density ($1000/ha^{-1}$)	Tree biomass (Mg/ha^{-1})	Constructive species
ARSP	14.9 ± 5.5 b	89.3 ± 29.7 a	*P. elliottii* Engelm.
ARMP	19.1 ± 6.0 b	89.5 ± 18.0 a	*P. massoniana* Lamb.
NRSF	23.9 ± 9.3 a	95.2 (76.0, 151) a	*Liquidambar formosana* Hance, *Cinnamomum camphora* L., *P. massoniana* Lamb., and *Cunninghamia lanceolata* (Lamb.) Hook

Mean ± standard deviation are shown for normal and approximately normal data. Median (25% quartile, 75% quartile) is shown for non-normal data. Tree density ($F = 4.85$, $P = 0.014$). Tree biomass ($\chi 2 = 0.730$, $P = 0.694$). If the letters following the numerals differ, the values of the indicated parameter are significantly different among the three forest restoration approaches at $P < 0.05$.

Table S2. Adonis-based multiple variance analysis of N-cycling functional gene compositions in soils in reforested forests:14 stands of artificially reforested slash pine plantations (ARSP), 14 stands of artificially reforested Masson pine plantations (ARMP), and 11 stands of naturally reforested secondary forests (NRSF)

Comparisons	R^2	*F*	*P*
ARMP and NRSF	0.147	3.96	0.009
ARSP and ARMP	0.114	3.36	0.011
ARSP and NRSF	0.149	4.04	0.005

P, significant level.

Table S3. Comparisons of plant diversity of artificially reforested slash pine plantations (ARSP), artificially reforested Masson pine plantations (ARMP), and natural reforested secondary forests (NRSF)

Reforestation approach	Plant diversity	Tree diversity	Shrub diversity	Herb diversity
ARSP	1.82 (1.74, 1.90) b	0.077 (0.063, 0.261) b	1.17 ± 0.31 a	0.542 ± 0.337 b
ARMP	1.77 ± 0.12 b	0.139 (0.045, 0.373) b	1.41 ± 0.42 a	0.543 ± 0.310 b
NRSF	2.34 ± 0.11 a	1.44 (1.14, 1.53) a	1.33 ± 0.44 a	0.901 ± 0.368 a

Mean ± standard deviation are shown for normal and approximately normal data. Medians (25% quartile, 75% quartile) are shown for non-normal data. The diversity of plant, tree, shrub and herb was measured by Shannon-Wiener diversity index. Plant diversity (χ^2 = 12.9, P =0.002); Tree diversity (χ^2 = 19.9, P = 7.15×10^{-5}); Shrub diversity (F = 1.40, P = 0.026); Herb diversity (F = 4.48, P = 0.018). If the letters following the numerals differ, the values of the indicated parameter are significantly different among the three forest restoration approaches at $P < 0.05$.

Table S4. Spearman correlation coefficients between scores of the first axis of detrended corresponding analysis of N-cycling genes and plant diversity

Item	Index	Species richness			Shannon-Wiener diversity		
		Plant	Tree	Herb	Plant	Tree	Herb
ARMP–NRSF	*r*	0.257	0.544*	0.358◇	0.417◇	0.581*	0.372◇
	P	0.214	0.005	0.059	0.038	0.002	0.067
ARSP–NRSF	*r*	0. 603*	0.646*	0.445◇	0.600*	0.682*	0.471◇
	P	0. 001	4.90×10^{-4}	0.026	0.001	1.76×10^{-4}	0.018

ARMP–NRSF, the first axis of detrended corresponding analysis of N-cycling genes between artificially reforested Masson pine plantations (ARMP) and naturally reforested secondary forests (NRSF). ARSP–NRSF, for that between artificially reforested slash pine plantations (ARSP) and NRSF. * and ◇ are significant at $P < 0.05$ and $P < 0.10$, respectively.

Table S5. Correlation coefficients of N-cycling gene diversity vs. the plant diversity indices, species richness and Shannon–Wiener diversity index, and the soil factor of available N

Genes diversity index	Index	Plant species richness			Plant diversity Shannon–Wiener			Soil factor
		Tree	Shrub	Herb	Tree	Shrub	Herb	Available N
Number	*r*	0.320*	0.003	0.295◇	0.357*	–0.130	–0.203	0.285◇
	P	0.047	0.988	0.069	0.026	0.432	0.215	0.079
Abundance	*r*	0.296◇	0.014	0.245	0.326*	–0.065	0.166	0.280◇
	P	0.067	0.932	0.132	0.043	0.694	0.312	0.084
Diversity	*r*	0.342*	–0.033	0.283◇	0.377*	–0.178	0.183	0.311◇
	P	0.033	0.841	0.083	0.018	0.277	0.264	0.054

r, correlation coefficient; *P*, significant level; data tabulated are Pearson correlation coefficients except for the columns of tree Shannon–Wiener diversity index, plant species richness of tree and shrub, and available N that are Spearman correlation coefficients because of their non-normal distributions. * and ◇ following numerals indicate the correlations of the respective N-cycling functional gene diversity indices (rows) and plant diversity indices (columns) are significant at $P < 0.05$ and $P < 0.10$, respectively.

Table S6. Topological properties of molecular ecological networks of N-cycling genes in soils of forest restored with three approaches

Reforestation approach	Gene†	Similarity threshold‡	*r* of power law§	Average path length¶
ARSP	232	0.900	-0.696	2.39 ± 0.06 c
ARMP	271	0.880	-0.864	2.65 ± 0.08 b
NRSF	301	0.870	-0.936	2.97 ± 0.11 a

ARSP, artificially reforested slash pine plantations; ARMP, artificially reforested Masson pine plantations; and NRSF, naturally reforested secondary forests. †, Number of functional genes in molecular ecological network; ‡, Value used for transforming similarity matrix to neighboring matrix; §, r is the correlation coefficient between log(P(K)) and log(K), where K is connectivity, P(K) is the number of nodes with the degree of K. $P < 0.0001$ for all the three reforestation approaches. ¶, if the letters following the numerals differ, the values of the indicated network parameter are significantly different among the three forest soil types at $P < 0.05$. Average path length, $F = 154.8$, $P < 0.001$.

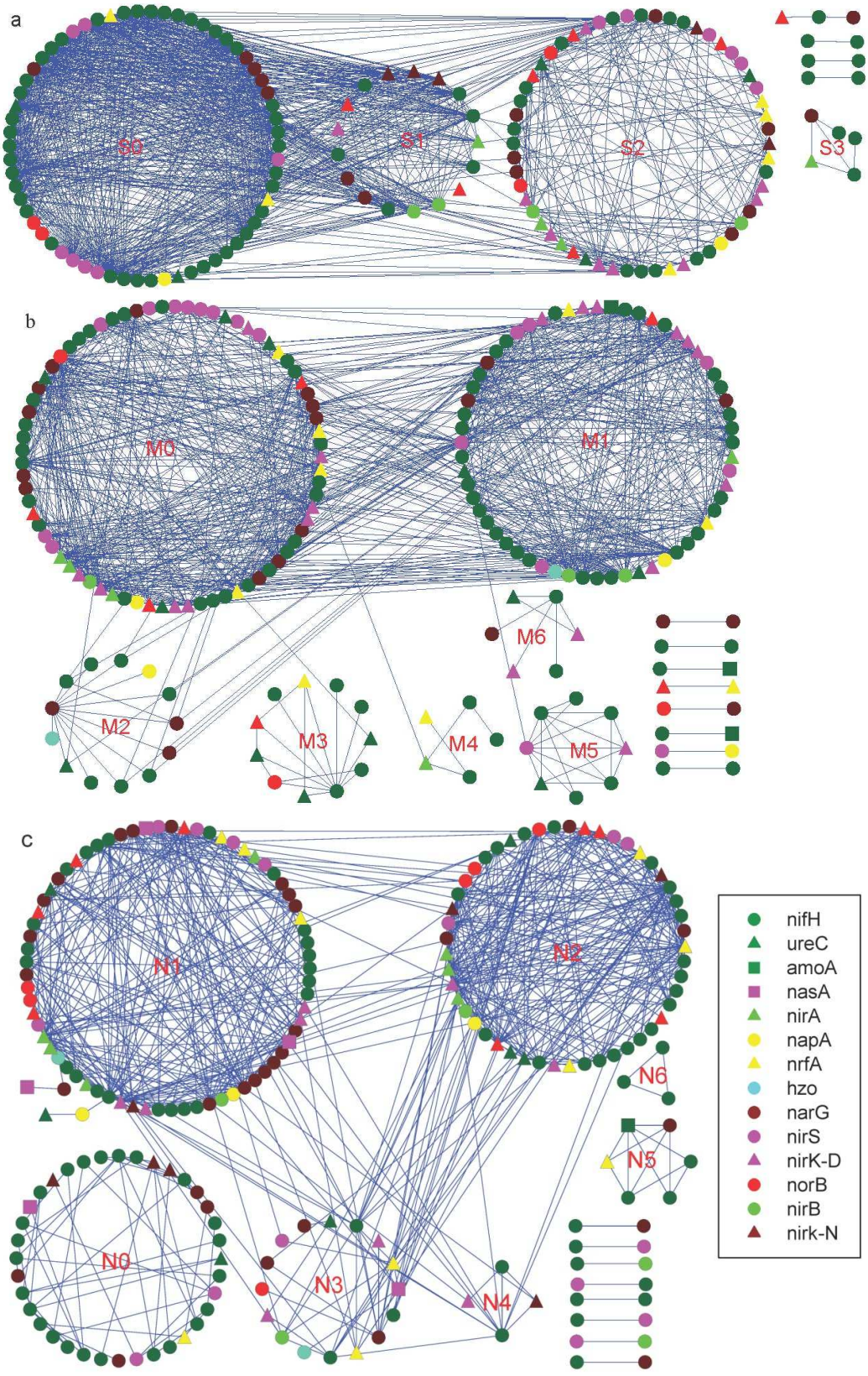
a
S0
S1
S2
S3
b
M0
M1
M2
M3
M4
M5
M6
c
N0
N1
N2
N3
N4
N5
N6
nifH
ureC
amoA
nasA
nirA
napA
nrfA
hzo
narG
nirS
nirK-D
norB
nirB
nirk-N

Fig. S1. Functional molecular ecological networks (fMEN) of N-cycling genes in the soils of artificially reforested slash pine plantations (ARSP, a), artificially reforested Masson pine plantations (ARMP, b), and naturally reforested secondary forests (NRSF, c). Each colored shape represents a different N-cycling gene detected by GeoChip 3.0 and are referred to as nodes. Lines connecting two nodes indicate interactions between those nodes, with blue lines indicating positive and red lines indicating negative interaction. Modules represent groups of genes with a high number of interactions among the node members. The four largest modules are labeled as reforestation approaches followed by module number 0–3; *i.e.* S0 to S3 denoted ARSP module 0 to ARSP module 3, M0 to M3 denoted ARMP module 0 to ARMP module 3, and N0 to N3 denoted NRSF module 0 to NRSF module 3 in ARSP, ARMP, NRSF, respectively.

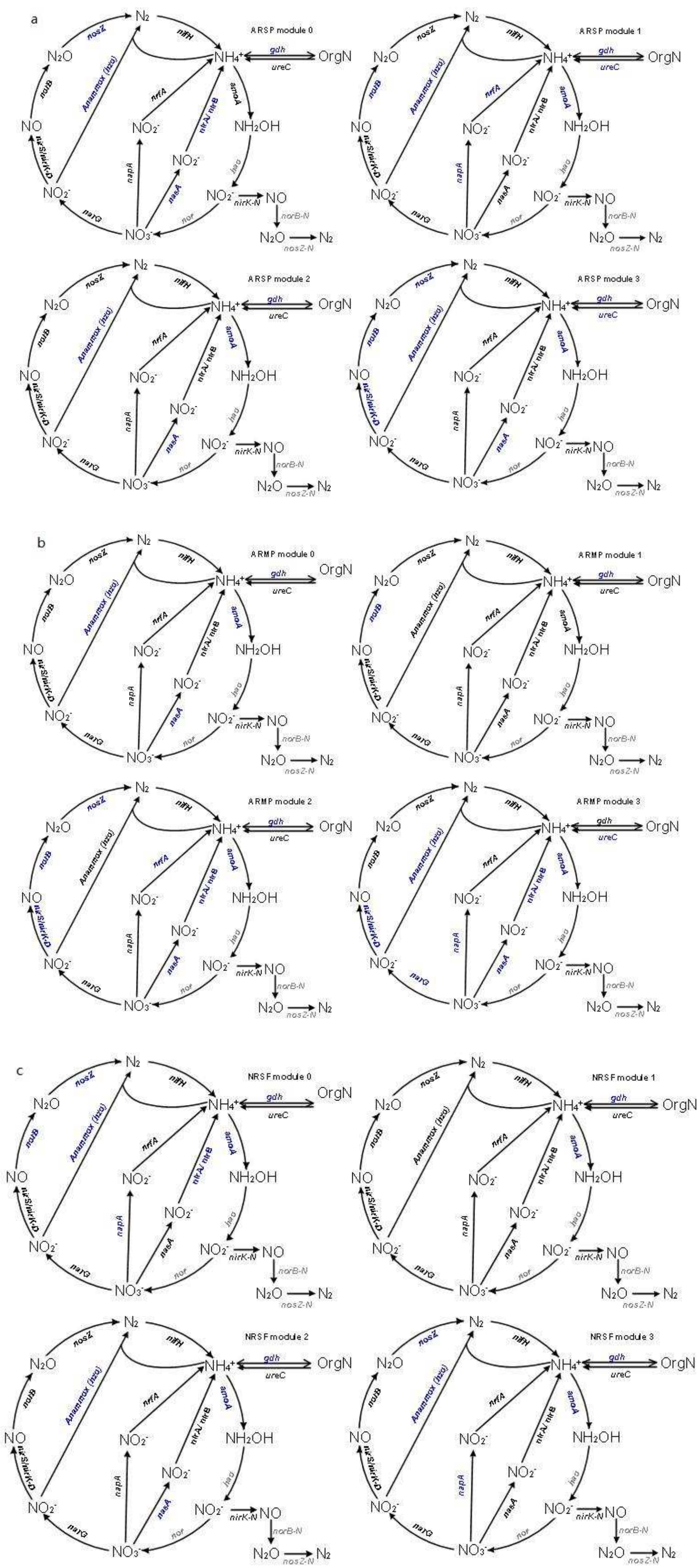
a
ARSP module 0
ARSP module 1
ARSP module 2
ARSP module 3
b
ARMP module 0
ARMP module 1
ARMP module 2
ARMP module 3
c
NRSF module 0
NRSF module 1
NRSF module 2
NRSF module 3
N2
N2O
NO
NO2-
NO3-
NH4+
NH2OH
OrgN
nosZ
nifH
norB
nirS/nirK-D
narG
napA
nasA
nrfA
nirA/ nirB
amoA
hao
nor
gdh
ureC
Anammox (hzo)
nirK-N
norB-N
nosZ-N

Fig. S2. Functional gene distributions in the four largest modules of N-cycling fMEN in soils of artificially reforested slash pine plantations (ARSP) (a), artificially reforested Masson pine plantations (ARMP) (b), and natural reforested secondary forests (NRSF) (c). Black and blue genes indicate that the functional genes of the processes are present and absent, respectively, and the arrows indicate reaction direction. The gray genes are not included in the functional microarray, GeoChip 3.0. As the probe number of *amoA* is far fewer than other genes in GeoChip 3.0, it is inferred that *amoA* is detected if genes involved in downstream processes such as assimilatory nitrate reduction, dissimilatory nitrate reduction, or denitrification were detected in the module. Similarly, if *nrfA* was detected, *napA* is expected to exist. If *nirA* or *nirB* was detected, *nasA* is expected to exist. *nosZ* is expected to exist if other genes involved in denitrification were all detected.

Ecosphere

Artificial reforestation produces less diverse soil nitrogen-cycling genes than natural restoration

Yun Wang, Hua Zheng, Falin Chen, Yunfeng Yang, Jing Zeng, Joy D. Van Nostrand, Jizhong Zhou, and Zhiyun Ouyang

Appendix S2

SUPPLEMENTARY METHODS

Plant survey and characteristics

In each of the 15 plots (500 m × 500 m), we surveyed the plant species composition and abundance at two to three tree subplots (10 m × 10 m), three to six shrub subplots (5 m × 5 m), and six to eight herb subplots (1 m × 1 m). In total, 40 tree, 76 shrub, and 112 herb subplots were surveyed for ARSP; correspondingly 42, 80, and 120 for ARMP; and 30, 57, and 87 for NRSF. Height and diameter at breast height (130 cm) were measured for each tree. Litter was collected from a 0.5 m × 0.5 m area within all tree subplots. Roots were collected using the same method as for soils. The understory plant composition and cover were described in previous study (Wang et al. 2013).

After washing with water, roots were scanned with an Epson Expression 836 XL scanning system (Epson Corp., Nagano, Japan), analyzed with WinRhizo V5.0 (Regent Instruments, Quebec, Canada), and dried at 80°C to constant weight. Then, parameters like root diameter and root surface area could be analyzed with WinRhizo V5.0. Thus the surface area of fine root (diameter < 2 mm) could be gotten. Litter was also dried at 80°C to constant weight. Tree biomass was calculated based on height and diameter at the breadth height of trees (Chen et al. 1993, Yao 2003, Wang et al. 2007, Zheng et al. 2008, Ouyang et al. 2010, Yang et al. 2010).

Soil characteristic analysis

Ammonium N, nitrite N, and nitrate N contents were measured as amounts in the extract of 2 mol l^{-1} KCl (analytical reagent, Sinopharm) aqueous solution by Continuous-Flow Analyzer (SKALAR SAN++, Breda, The Netherlands). Total carbon (C) and N were measured by Elementer Analyzer (VARIO, EL III, Hanau, Germany). Organic N was obtained from subtracting inorganic N from total N. Available phosphorus was measured by inductively coupled plasma

-mass spectrometry (ICP–MS; NexION 300, Perkin-Elmer, Norwalk, CT, USA) after extraction with HCl and H_2SO_4 (analytical, reagent, Sinopharm) (Bao 2000). Labile C was the C that hydrolyzed by sulfuric acid in the two-stage sulfuric acid hydrolysis method (Rovira and Vallejo 2002).

Soil particle size was measured with Mastersizer 2000 (Malvern Instruments, Malvern, UK) (Wang et al. 2008). Clay (particle size less than 2 μm) content was calculated as the ratio of volume of clay to volume of all soil particles.

Relationship of N cycling genes with environmental factors

Canonical correspondence analysis (CCA) was used to correlate N-cycling gene composition with environmental factors. The environmental factors included plant factors of Shannon–Wiener diversity index of tree layer, shrub layer, and herb layer; tree layer biomass; fine root surface area; litter stock; and root biomass; and soil factors of concentrations of nitrate N, nitrite N, available N, ammonium N, available P, labile C, clay, and sand. Except for the Shannon–Wiener diversity indices and soil pH, all other factors were $\log_{10}$ transformed, then normalized as Z scores (mean = 0, standard deviation = 1) to reduce the influence of units. Each factor group (i.e., plant and soil factors) was preselected separately using 999 Monte Carlo permutation tests. All significant factors were used in correlation analyses with abundance of N-cycling genes.

The relationship between abundance of N-cycling genes with environmental factors was measured by Mantel test, in which the Bray-Curtis distance was used for abundance of functional genes and the Euclidean distance for environmental factors.

The CCA was carried out in CANOCO for Windows 4.5. The Adonis and Mantel tests were performed in the R program. Figures were prepared using Sigmaplot (SPSS Inc.).

Construction of functional molecular ecological network

First, we constructed the Pearson correlation matrix based on the logarithm-transformed signal intensities. Then, a comparable matrix was built based on absolute values of the resulting Pearson correlation matrix. The adjacency matrix, which detected the connectivity between each pair of nodes, was converted based on the similarity threshold (s_t) in turn determined by random matrix theory (Luo et al. 2006). Greedy modularity optimization was used to check the module (Clauset et al. 2004). We defined a quality index (Q) when constructing modules. Q was the fraction of edges that fell within communities minus the expected value of the same quantity if edges fell at random without regard for community structure. The module was detected based on the highest Q (Newman 2004):

$$Q = \sum_i (e_{ii} - a_i^2) \quad (1)$$

$$a_i = \sum_j e_{ij} \quad (2)$$

where e_{ij} is the fraction of edges of group i to group j. The edge here is the connectivity between each pair of nodes.

Modularity (M) is calculated as follows:

$$M = \sum_{S=1}^{N_M} (I_S / I - (K_S / 2I)^2) \quad (3)$$

where N_M is the number of modules, I_s is the number of connections among nodes within module S, I is the number of connections within network, and Ks is the sum of connections of all nodes (Guimera and Amaral 2005).

The random networks were produced based on the Malsov–Sneppen procedure in which the number of nodes and connections remained the same while the position of connections changed so as to make the random network comparable to the original network (Zhou et al. 2010).

We plotted out the complete N cycle and determined which genes were present in the four

most prominent modules of each fMEN (fig. S2) – as ARSP had only four modules with more than five nodes, which may complete the entire N-cycling (from nitrogen gas cycled to either nitrogen gas or ammonium), only the four largest modules were chosen to compare among reforestation approaches. We defined a pathway as the presence of genes in a module with the potential to complete the entire N-cycling process. Absence of a gene indicates inability to complete a process.

Each of the following pathways can perform the entire N-cycling: (1) assimilatory nitrate reduction pathway included N fixation (*nifH*)/ammonification (*ureC*), nitrification (*amoA*), and assimilatory nitrate reduction (*nasA*, *nirA*/*nirB*); (2) dissimilatory nitrate reduction pathway included N fixation/ammonification, nitrification, and dissimilatory nitrate reduction (*nasA*, *nirA*/*nirB*)); (3) denitrification pathway included N fixation/ ammonification, nitrification, and denitrification (*narG*, *nirS*, *nirK*, *norB*, and *nosZ*); and (4) anammox pathway included N fixation/ammonification, nitrification, denitrification (*narG*), and anammox (*hzo*).

LITERATURE CITED

Bao, S. D. 2000. Soil and Agricultural Chemical Analysis. Third edition. China Agricultural Press, Beijing, China.

Chen Z. H., S. B. Wang, and Z.Q Zhang. 1993. Study on biomass and its allocation of the evergreen broadleaved forest in Heishiding, Guangdong. Acta Phytoecologica 17:289-298.

Clauset, A., M. E. Newman, and C. Moore. 2004. Finding community structure in very large networks. Physical Review E 70:1-6.

Guimera, R. and L. A. N. Amaral. 2005. Functional cartography of complex metabolic networks. Nature 433:895-900.

Luo, F., J. Zhong, Y. Yang, R. H. Scheuermann, and J. Zhou. 2006. Application of random matrix theory to biological networks. Physics Letters A 357:420-423.

Newman, M. E. J. 2004. Fast algorithm for detecting community structure in networks. Physical Review E 69:1-5.

Ouyang S. L., C. D. Dai, Y. N. Hou, Y. X. Xu, and J. Luo. 2010. Establishment of main constructive species biomass model for protection forest system around Dongting Lake. Science of Forestry in Hunan Province:22-24.

Rovira, P. and V. Vallejo. 2002. Labile and recalcitrant pools of carbon and nitrogen in organic matter decomposing at different depths in soil: an acid hydrolysis approach. Geoderma 107:109-141.

Wang, D., B. Fu, W. Zhao, H. Hu, and Y. Wang. 2008. Multifractal characteristics of soil particle size distribution under different land-use types on the Loess Plateau, China. Catena 72:29-36.

Wang X. Y, D. Hu, and J. S. He. 2007. Biomass research of *Fagus Engleriana* and *Quercus aliena var.acuteserrata* forest in Shennongjia Foest District. Journal of Capital Normal University (Natural Science Edition) 28: 62-67.

Wang, Y., Z. Ouyang, H. Zheng, F. Chen, S. Chen, J. Zeng. 2013. Effects of three forest restoration approaches on plant diversity in red soil region, southern China. Acta Ecologica Sinica. 33: 1202-1211.

Yang T. H., K. Song, L. J. Da , X. P. Li, and J. P. Wu. 2010. The biomass and aboveground net primary productivity of *Schima superba-Castanopsis carlesii* forests in east China Science in China, Series C Life Sciences:610-619.

Yao, Y. 2003. Biomass accumulation and nutrient distribution in *Cinnamommum camphora* plantation. Thesis. Central South University of Forestry and Technology. Changsha, Hunan, China.

Zheng, H., Z. Ouyang, W. Xu, X. Wang, H. Miao, X. Li, and Y. Tian. 2008. Variation of carbon storage by different reforestation types in the hilly red soil region of southern China. Forest Ecology and Management 255:1113-1121.

Zhou, J., Y. Deng, F. Luo, Z. He, Q. Tu, and X. Zhi. 2010. Functional molecular ecological networks. mBio 1:1-10.