

## **Supporting Online Material for**

### **Human rigidly dominate urban metabolic network**

Yong Min<sup>1,2</sup>, Hong Li<sup>3,\*</sup>, Ying Ge<sup>4</sup>, Jie Chang<sup>4</sup>

<sup>1</sup>Center for Computational Communication Research, Beijing Normal University, Zhuhai, 519087.

<sup>2</sup>School of Journalism and Communication, Beijing Normal University, Beijing, 100875.

<sup>3</sup>Faculty of Architecture and Urban Planning, Chongqing University, Chongqing 400045, China.

<sup>4</sup>College of Life Sciences, Zhejiang University, Hangzhou 310058, China.

\*Corresponding Author: Hong Li, Ph.D.

Email: lihom@cqu.edu.cn, Phone: 86-13193093366

#### **This PDF file includes**

Section S1

Tables S1 to S4

Figures S1 to S4

References

## Supplementary Section S1 - Discussion

A macro-scale reconstruction of a biogeochemical network is a non-automated and iterative decision-making process (Fig.1 in the main text). Here, we discuss the details of decisions which should be made in the reconstruction process.

**(1) Fineness.** Unlike in cellular metabolic networks whose metabolites are biochemical reactants in molecular level (Francke et al., 2005), the fineness of metabolites in urban biogeochemical metabolic networks is not fixed. Here we can consider entire agricultural subsystem as a single metabolite, and we also can consider croplands, vegetables and fertilizers in agriculture as different metabolites, furthermore, we even can consider various N chemical forms in molecules level, such as  $\text{NH}_4^+$ ,  $\text{NO}_2^-$  and  $\text{N}_2\text{O}$ . The fineness is critical for reconstruction. A URC system will become to a single node in the coarsest situation, and lost all of structural properties. Contrarily, it will become to a very complex network which appears all micro-scale metabolic networks of various organisms, and it is impossible to obtain necessary data. Hence, we should depend on target problems and owned data to carefully decide fineness of networks.

**(2) Geography.** From geographic perspective, we can ignore geographic factors, such as the direction of stream and wind, the position of functional elements, such as inhabited area, cropland and forest. If we consider geographic differences, we should define the scale of an 'elementary block', in kilometer, meter or even centimeter, etc. In fact, geography is another type of fineness. The complex networks with geographical effects always have different properties (Warren et al., 2002; Huang et al., 2006). In the future, we want to import data from GIS to build more powerful network models with

geographic properties.

**(3) Composition.** In biogeochemistry, we usually study single chemical element independently, such as carbon, nitrogen, phosphorous and so on. In this situation, the quantity of a metabolite is the net quantity of the element. However, if there are enough data, we could also integrate multi-elements into one network by imitating biochemical reactants. For example, a metabolite, whose quantity of carbon and nitrogen is 2:1, can be written as  $C_2N$ . The study in this paper don't involve multi-elements, the operations of multi-elements would be studied in our future works.

**(4) Present.** After establishing all metabolites, we should know whether there exist fluxes between all pairs of metabolites.

**(5) Relationship.** If some fluxes have close relationship with each others, we should combine them and write in a conversion equation. For example, two fluxes, 'AgSl  $\rightarrow$  AgPr' and 'AgSl  $\rightarrow$  AgSt', could be written as: 'AgSl  $\rightarrow$  AgPr + AgSt'. Current, we lack of a concrete principle to judge the relationships between different fluxes. In the future, we need consider both ecological and structural characters to establish a reasonable principle and also need more cases to validate the principle. It is a difficult and time-consuming works.

**(6) Stoichiometry.** Stoichiometry rests upon the law of conservation of mass. Since biogeochemical conversions can neither create nor destroy matter, only change one metabolite into another, the amount of each element must be the same throughout the overall conversion. For single-element model, the stoichiometric coefficient in a basic conversion must be 1, such as '1 AgSl  $\rightarrow$  1 AgPr'. The stoichiometric coefficients in a model conversion should be determined by actual data, e.g. '1 AgSl  $\rightarrow$  0.2 AgPr + 0.8

AgSt', thus different systems usually have different stoichiometric coefficients for the same conversion. It is different with biochemical reactions: one reaction always has the same stoichiometry in different organisms. That is a reason why we insist on using basic conversion in our first model, which have unique forms in different URCs.

**(7) Inputs/Outputs.** For an open system, the inputs and outputs are very important, which represent the interaction between the system and outer environments. We can use external metabolites to summarize certain type of outputs, such as using ExPr to represent the output of various products.

**(8) Reversibility.** Two reversed conversions, such as biological fixation ' $1 \text{ N}_2 \rightarrow 1 \text{ Soil}$ ' and denitrification ' $1 \text{ Soil} \rightarrow 1 \text{ N}_2$ ', can also be represented in a reversible model conversion: ' $1 \text{ N}_2 \leftrightarrow 1 \text{ Soil}$ '. Reversibility, which is a special relationship between basic conversions, will affect the structural properties in network-based pathway analysis. The detail effects could be found at (Klamt et al., 2003; Papin et al., 2004).

In summery, it is the first attempt to reconstruct biogeochemical network models from a metabolic perspective. Just as reconstruction of cellular metabolic networks, the biogeochemical network reconstruction will be gradual and endless processes, which need both the accumulation of data and carefully studying of technical details. Hence, our work just is a start point.

**Table S1 Glossary**

Type	Cellular System	URC System	Explanation
<b>Model words</b>	Metabolite	Metabolite	Various N forms in different materials or places, such as N in soil and atmosphere, and food and fertilizer.
	Reaction, Transportation	Conversion	Describes both transformation and movement of metabolites.
		Convertor	The places, facilities or matters contain multi-form nitrogen, and translate them into different other convertors.
	Product	Product	A substance that forms as a result of a conversion, and locate in the right side of conversion equations.
	Substrate	Substrate	A substance that uptakes as a origin of a conversion, and locate in the left side of conversion equations.
<b>Mathematical words</b>	Convex solution cone		A mathematically defined space (whose axes correspond to reaction fluxes) that encompasses all valid steady-state flux distributions of a metabolic network.
	Exchange flux		A reaction flux that crosses the system boundary.
	Extreme pathways (EPs)		A minimal and unique set of pathways that defines the edges of the convex solution cone.
	Flux		The production or consumption of mass per unit area per unit time.
	Flux distribution		A particular set of reaction fluxes in a network.
	Internal flux		A reaction flux within defined system boundaries.
	Steady state		The network state in which all metabolite concentrations are constant, although there might be flux through the reactions.
	Stoichiometric matrix		A matrix with reaction stoichiometries as columns and metabolite participations as rows.
$\alpha$ -weighting	Flux of EPs		There is not a unique set of nonnegative weightings on the extreme pathways that produce a given steady state flux distribution but rather a range of possible values.

**Table S2 Subsystems for nitrogen cycles in the GHA system**

Subsystem Name	Functional Group	Abbreviation	Flux (Ggy <sup>-1</sup> )	Participation for Stem
Human	Consumer	Hm	83.18	0.564
Pets		Pt	4.16	0.037
Agriculture	Processor	Ag	126.72	0.806
Aquiculture		Aq	33.09	0.178
Livestock		Ls	39.64	0.436
Forest-Grassland		FG	38.09	0.359
Urban lawn		Lw	14.96	0.074
Waste water treatment	Remover	WTF	25.20	0.362
Surface water	Life-Supporter	SW	99.06	0.820
Near-Atmosphere		NA	75.24	0.422
Subsurface water		SsW	9.51	0.045
Solid waste		Swst	9.17	0.068

The GHA system was divided into four functional groups according to the role they play in N biogeochemical cycling: each functional group contains one or more subsystems. The processor group can process the input fixed N to food and other useful N containing products and then to support the nutrients and materials needed by consumers. The remover group mentions the artificial facilities to treat the waste N with processes converting active N ( $N_R$ ) into  $N_2$ , only including wastewater treatment subsystem here. The life-supporter group closely relate to almost all other subsystems. Subsystem participation is the percentage of extreme pathways that are related to a given subsystem, and subsystems that participate in a large number of extreme pathways means high influence and importance for deciding systemic state.

**Table S3 Metabolites in Nitrogen Cycles of GHA System**

Type	Component	Abbr. of Component	Subsystem	Name of Metabolite
<b>A</b>	Food / Feed	Fd	Human	HmFd
			Pets	PtFd
			Aquiculture	AqFd
			Livestock	LsFd
	Excretion	Ex	Human	HmEx
			Pet	PtEx
			Aquiculture	AqEx
	Fertilizer	Fr	Livestock	LsEx
			Agriculture	AgFr
	N-Chemicals	Nc	Urban lawn	LwFr
			Agriculture	AgFr
	Waste water	Ww	Human	HmNc
	Product	Pr	Human	HmWw
			Agriculture	AgPr
Aquiculture			AqPr	
Livestock			LsPr	
Soil	Sl	Agriculture	AgSl	
		Urban lawn	LwSl	
Straw	St	Agriculture	AgSt	
Pool water	Pw	Aquiculture	AqPw	
<b>B</b>	\		Forest-Grassland	FG
	\		Wastewater treatment	WTF
	\		Surface water	SW
	\		Near-surface Atmosphere	NA
	\		Subsurface water	SsW
	\		Solid waste	Swst
	\		Landfill (LF)	LF
<b>C</b>	\		Outer-Atmosphere (OA)	OA
	\		Accumulation (AC)	AC
	\		N <sub>2</sub> (N <sub>2</sub> )	N <sub>2</sub> In / N <sub>2</sub> Ot
<b>D</b>	Upstream	Uw		ExUw
	Downstream	Dw		ExDw
	Product	Pr	External metabolites (Ex)	ExPr
	Underground	Ug		ExUg
	Atmosphere	NA		ExNA

There are 4 kinds of metabolites: **(A)** for processors and consumers, except for forest-grassland subsystem, metabolites were inner components of subsystems. **(B)** all removers, life-supporters and forest-grassland subsystem were treated as a single metabolite. **(C)** some additional metabolites were considered to construct a complete

network, including outer-atmosphere (the source of wet deposition),  $N_2$  (we divide it to two metabolites:  $N_2In$  for the source of biological fixation and  $N_2Ot$  for the target of denitrification), and accumulation (an abstract metabolite for keeping mass balance). **(D)** five external metabolites are added in the system to simplify the expression of conversions and the representation of inputs/outputs. External metabolites don't change the structural properties of the network, and we don't consider them in calculating pathway length. The naming rule of a metabolite is: for type A and D, the name of a metabolite (4 letters) is the combination of its subsystem name (2 letters in abbreviation) and its component name (2 letters in abbreviation); for type B and C, the name of a metabolite just is the name of abbreviation of corresponding subsystem name (2-4 letters).

**Table S4 Conversions in nitrogen cycles of the GHA system**

Expression	Type	Flux (Ggy <sup>-1</sup> )	Participation of stem (%)	Relationship (%)
→ 1 AgFr	Inputs	75.12	0.088	0.538
→ 1 AqFd		33.09	0.164	0.641
→ 1 HmFd		17.23	0.091	0.513
→ 1 HmNc		42.10	0.096	0.551
→ 1 LsFd		34.26	0.165	0.551
→ 1 LwFr		9.80	0.032	0.487
→ 1 PtFd		4.16	0.032	0.500
→ 1 ExUw		20.90	0.036	0.487
→ 1 OA		29.76	0.156	0.577
→ 1 N2In		23.15	0.120	0.564
1 AC →	Outputs	47.18	0.277	0.731
1 ExPr →		27.01	0.097	0.667
1 ExNA →		67.44	0.122	0.667
1 ExDw →		50.58	0.092	0.654
1 ExUg →		7.58	0.058	0.628
1 N2Ot →		89.78	0.332	0.744
1 AgFr → 1 AgSl	Inner flux	75.12	0.088	0.538
1 OA → 1 AgSl		3.27	0.088	0.538
1 N2In → 1 AgSl		10.10	0.088	0.538
1 HmEx → 1 AgSl		12.10	0.110	0.513
1 LsEx → 1 AgSl		17.27	0.047	0.487
1 AgSt → 1 AgSl		4.80	0.002	0.026
1 SW → 1 AgSl		3.20	0.382	0.821
1 NA → 1 AgSl		0.86	0.000	0.000
1 AgSl → 1 AgPr		30.70	0.436	0.859
1 AgSl → 1 N2Ot		31.97	0.042	0.603
1 AgSl → 1 LF		0.71	0.000	0.000
1 AgSl → 1 AgSt		9.05	0.044	0.628
1 AgSl → 1 NA		22.95	0.139	0.692
1 AgSl → 1 SW		19.75	0.058	0.397
1 AgSl → 1 SsW		5.01	0.042	0.615
1 AgSl → 1 AC		6.58	0.042	0.603
1 AgPr → 1 HmFd		25.32	0.167	0.782
1 AgPr → 1 LsFd		5.38	0.269	0.833
1 AgSt → 1 NA		1.08	0.000	0.000
1 AgSt → 1 ExPr		3.17	0.042	0.615
1 AqFd → 1 AqPw		33.09	0.164	0.641
1 AqPw → 1 AqPr		6.31	0.091	0.551
1 AqPw → 1 N2Ot		2.36	0.002	0.051
1 AqPw → 1 AqEx		7.22	0.036	0.513
1 AqPw → 1 NA		1.59	0.000	0.000

1 AqPw → 1 SW	15.16	0.036	0.500
1 AqPw → 1 AC	0.45	0.000	0.000
1 AqEx → 1 SW	7.22	0.036	0.513
1 AqPr → 1 HmFd	6.31	0.091	0.551
1 LsFd → 1 LsPr	9.45	0.203	0.846
1 LsFd → 1 LsEx	28.71	0.232	0.833
1 LsFd → 1 AC	1.48	0.000	0.000
1 LsPr → 1 HmFd	9.45	0.203	0.846
1 LsEx → 1 SW	7.14	0.068	0.603
1 LsEx → 1 WTF	4.30	0.117	0.782
1 OA → 1 FG	20.24	0.032	0.487
1 N2In → 1 FG	12.56	0.032	0.487
1 NA → 1 FG	5.29	0.282	0.897
1 FG → 1 N2Ot	2.32	0.118	0.667
1 FG → 1 SW	4.57	0.110	0.679
1 FG → 1 AC	31.20	0.118	0.667
1 LwFr → 1 LwSl	9.80	0.032	0.487
1 OA → 1 LwSl	0.60	0.000	0.000
1 N2In → 1 LwSl	0.49	0.000	0.000
1 PtEx → 1 LwSl	3.92	0.032	0.500
1 NA → 1 LwSl	0.16	0.000	0.000
1 LwSl → 1 N2Ot	2.84	0.003	0.090
1 LwSl → 1 Swst	6.35	0.003	0.115
1 LwSl → 1 NA	1.67	0.000	0.000
1 LwSl → 1 WTF	2.59	0.058	0.500
1 LwSl → 1 SsW	1.52	0.000	0.000
1 PtFd → 1 PtEx	4.16	0.032	0.500
1 PtEx → 1 Swst	0.24	0.000	0.000
1 HmFd → 1 ExPr	23.84	0.055	0.590
1 HmFd → 1 Swst	2.58	0.055	0.615
1 HmFd → 1 HmEx	31.45	0.441	0.795
1 HmFd → 1 AC	0.44	0.000	0.000
1 HmEx → 1 SW	6.33	0.125	0.564
1 HmEx → 1 WTF	13.02	0.206	0.744
1 HmNc → 1 NA	32.81	0.031	0.474
1 HmNc → 1 HmWw	9.29	0.065	0.526
1 HmWw → 1 SW	5.38	0.036	0.500
1 HmWw → 1 WTF	3.91	0.029	0.462
1 Swst → 1 NA	0.92	0.000	0.000
1 Swst → 1 LF	8.25	0.058	0.628
1 LF → 1 WTF	1.38	0.000	0.000
1 LF → 1 ExUg	7.58	0.058	0.628
1 WTF → 1 N2Ot	17.46	0.075	0.628
1 WTF → 1 SW	5.47	0.261	0.744

1 WTF → 1 SsW	2.27	0.075	0.641
1 SsW → 1 N2Ot	1.77	0.000	0.000
1 SsW → 1 AC	7.03	0.117	0.667
1 ExUw → 1 SW	20.90	0.036	0.487
1 OA → 1 SW	5.65	0.036	0.487
1 NA → 1 SW	1.48	0.000	0.000
1 SW → 1 N2Ot	31.06	0.092	0.654
1 SW → 1 NA	14.21	0.233	0.692
1 SW → 1 ExDw	50.58	0.092	0.654
1 NA → 1 ExNA	67.44	0.122	0.667

The fluxes of all conversions list in the table, which compose the flux distribution of N cycle in the GHA system 2005. Additionally, two conversion properties, participation and correlation ratio, are also list in the table.

**Table S5 Additional restrictions for optimization calculation of flux capacity**

Previous conversion	Next conversion	Restriction
AgFr → AgSl	AgSl → AgPr	$v_{next} < 0.2 v_{pre}$
	AgSl → N <sub>2</sub>	$v_{next} < 0.35 v_{pre}$
	AgSl → AgSt	$v_{next} < 0.1 v_{pre}$
	AgSl → NA	$v_{next} < 0.3 v_{pre}$
	AgSl → SW	$v_{next} < 0.2 v_{pre}$
	AgSl → SsW	$v_{next} < 0.1 v_{pre}$
HmEx / LsEx → AgSl	AgSl → AgPr	$v_{next} < 0.3 v_{pre}$
	AgSl → N <sub>2</sub>	$v_{next} < 0.2 v_{pre}$
	AgSl → AgSt	$v_{next} < 0.1 v_{pre}$
	AgSl → NA	$v_{next} < 0.2 v_{pre}$
	AgSl → SW	$v_{next} < 0.2 v_{pre}$
	AgSl → SsW	$v_{next} < 0.1 v_{pre}$
OA / NA / SW / N <sub>2</sub> → AgSl	AgSl → AgPr	$v_{next} < 1 v_{pre}$
	AgSl → AgSt	$v_{next} < 1 v_{pre}$
	<i>others</i>	$v_{next} = 0$
<i>ALL</i>	LsFd → LsPr	$v_{next} < 0.3 v_{pre}$
	LsFd → LsEx	$v_{next} < 0.8 v_{pre}$
LsFd → LsEx	LsEx → AgSl	$v_{next} < 0.6 v_{pre}$
	LsEx → SW	$v_{next} < 0.3 v_{pre}$
	LsEx → WTF	$v_{next} < 0.2 v_{pre}$
LwFr → LwSl	LwSl → N <sub>2</sub>	$v_{next} < 0.25 v_{pre}$
	LwSl → Swst	$v_{next} < 0.5 v_{pre}$
	LwSl → NA	$v_{next} < 0.1 v_{pre}$
	LwSl → WTF	$v_{next} < 0.2 v_{pre}$
	LwSl → SsW	$v_{next} < 0.15 v_{pre}$
PtEx → LwSl	LwSl → N <sub>2</sub>	$v_{next} < 0.2 v_{pre}$
	LwSl → Swst	$v_{next} < 0.3 v_{pre}$
	LwSl → NA	$v_{next} < 0.2 v_{pre}$
	LwSl → WTF	$v_{next} < 0.2 v_{pre}$
	LwSl → SsW	$v_{next} < 0.2 v_{pre}$
OA / NA / N <sub>2</sub> → LwSl	LwSl → Swst	$v_{next} = v_{pre}$
	<i>others</i>	$v_{next} = 0$
<i>ALL</i>	HmFd → Swst	$v_{next} < 0.1 v_{pre}$
	HmFd → HmEx	$v_{next} < 0.95 v_{pre}$
HmFd → HmEx	HmEx → AgSl	$v_{next} < 0.4 v_{pre}$
	HmEx → Sw	$v_{next} < 0.2 v_{pre}$
	HmEx → WTF	$v_{next} < 0.4 v_{pre}$
<i>ALL</i>	Swst → LF	$v_{next} < 0.1 v_{pre}$
	Swst → NA	$v_{next} < 0.9 v_{pre}$
<i>ALL</i>	LF → WTF	$v_{next} < 0.2 v_{pre}$
	LF → EXUD	$v_{next} < 0.9 v_{pre}$

<i>ALL</i>	WTF → N <sub>2</sub>	$v_{next} < 0.7 v_{pre}$
	WTF → SW	$v_{next} < 0.25 v_{pre}$
	WTF → SsW	$v_{next} < 0.1 v_{pre}$
AgSI / LwSI → SsW	SsW → N <sub>2</sub>	$v_{next} = 0$
	SsW → AC	$v_{next} = v_{pre}$
WTF → SsW	SsW → N <sub>2</sub>	$v_{next} < 0.95 v_{pre}$
	SsW → AC	$v_{next} < 0.05 v_{pre}$
<i>ALL</i>	SW → AgSI	$v_{next} < 0.05 v_{pre}$
	SW → N <sub>2</sub>	$v_{next} < 0.35 v_{pre}$
	SW → NA	$v_{next} < 0.15 v_{pre}$
	SW → EXDW	$v_{next} < 0.6 v_{pre}$

## Supplementary Figures

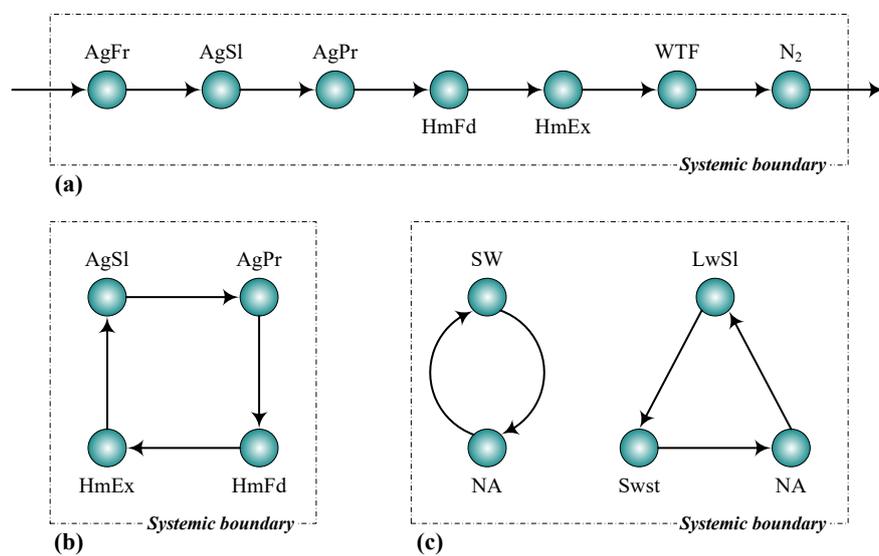
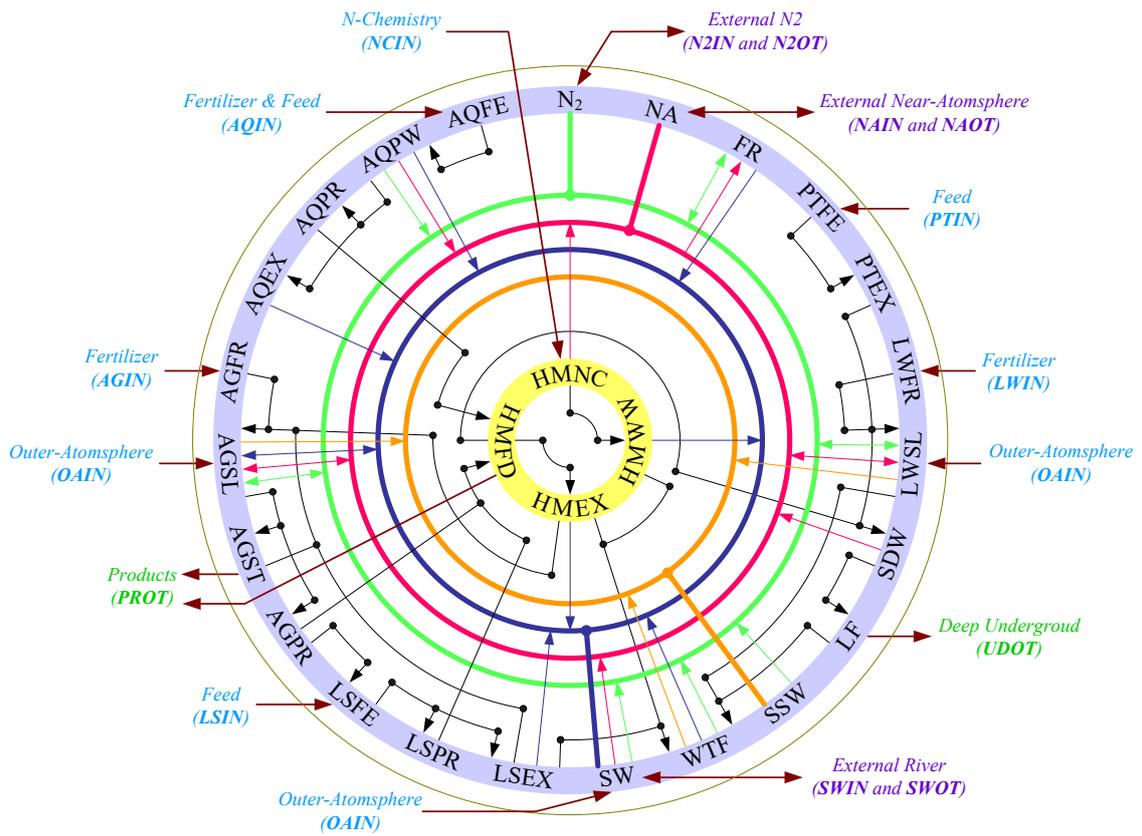
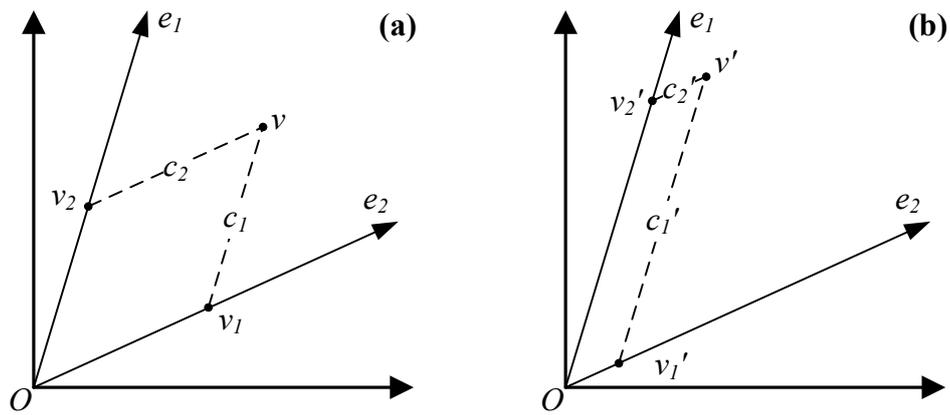


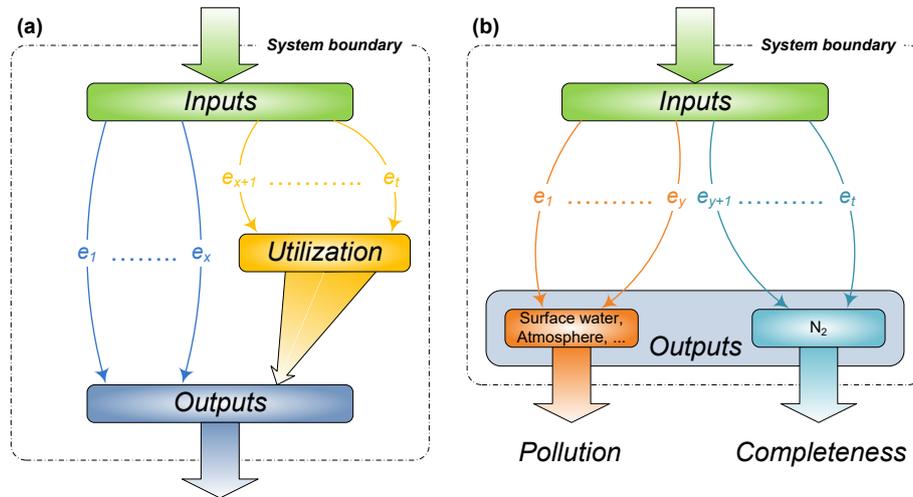
Figure S1 Three types of extreme pathways.



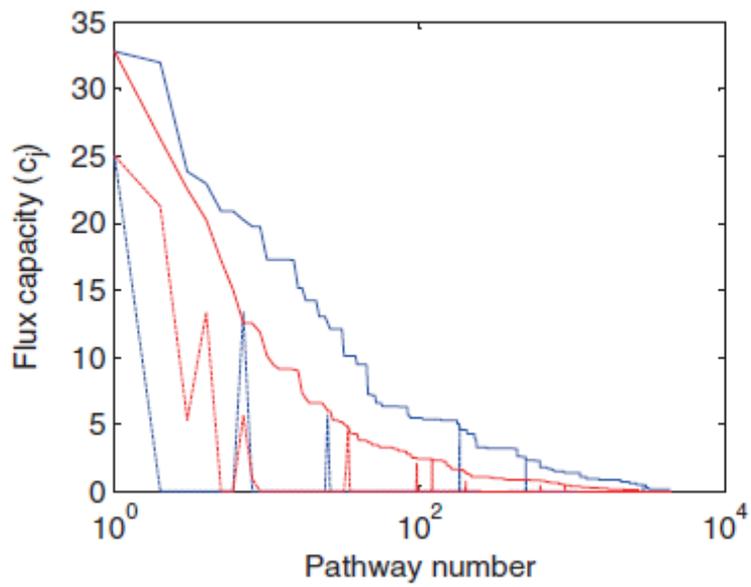
**Figure S2 The UBN of N cycles in GHA system.**



**Figure S3 The distribution of extreme pathway fluxes.** Different kinds of extreme pathway fluxes distribution will represent different systemic properties. We take a system with two extreme pathways  $e_1$  and  $e_2$  to illustrate such difference. **(a)** The flux of both extreme pathways,  $c_1$  and  $c_2$ , are basically the same, therefore, no matter which extreme pathway is removed, the new steady-state,  $v_1$  or  $v_2$ , will be significantly different with original steady-state  $v$ . **(b)** Contrarily, the flux  $c_1'$  is very larger than  $c_2'$ . By removing  $e_1$ , the new steady-state  $v_1'$  will deviate from original steady-state  $v'$ , however, the new steady-state  $v_2'$ , which is generated by removing  $e_2$ , is close to  $v'$ . It means  $e_1$  is more important than  $e_2$  to the given steady-state. For a system with large number of extreme pathways, the situations in (a) and (b) may approximately evolve to the Poisson and power-law distribution respectively.



**Figure S4 The calculation of NUE and PR.** (a) All EPs are divided into two categories by whether or not to be used by human: blue EPs don't pass through the human consumption, but N flows through yellow EPs are utilized by human. (b) All pathways are divided into two categories by different outputs: blue EPs completely translate inputting N into N<sub>2</sub>, but yellow EPs export active N through atmosphere, surface water and other life-supporters.



**Figure S5 The optimization of calculation of extreme pathway fluxes.** Blue lines represent the distribution of the range of flux without additional restrictions, and red lines are the result with restrictions in Tab.5. Solid lines indicate  $\max(c_i)$  and dash lines indicate  $\min(c_i)$ . The red range is obviously narrower than blue range, hence additional restrictions are effective.

## Supporting references and notes

- Francke, C., Siezen, R. J. & Teusink, B. Reconstructing the metabolic network of a bacterium from its genome. *Trends Microbiol* **13**, 550-558 (2005).
- Klamt, S. & Stelling, J. Two approaches for metabolic pathway analysis? *Trends Biotechnol* **21**, 64-69 (2003).
- Papin, J. A. *et al.* Comparison of network-based pathway analysis methods. *Trends Biotechnol* **22**, 400-405 (2004).
- Huang, L., Yang, L. & Yang, K. Q. Geographical effects on cascading breakdowns of scale-free networks. *Physical Review E* **73**, -, doi:Artn 036102 Doi 10.1103/Physreve.73.036102 (2006).
- Warren, C. P., Sander, L. M. & Sokolov, I. M. Geography in a scale-free network model. *Physical Review E* **66**, -, doi:Artn 056105 Doi 10.1103/Physreve.66.056105 (2002).