

Microbial community dynamics and assembly follows trajectories of an early spring diatom bloom in a semi-enclosed bay

Huajun Zhang¹, Kai Wang^{1,2}, Lixin Shen¹, Heping Chen^{1,4}, Fanrong Hou¹, Xiaoyan Zhou³, Demin Zhang^{1,2*}, Xiangyu Zhu^{3*}

¹ School of Marine Sciences, Ningbo University, Ningbo, 315211, China

² Collaborative Innovation Center for Zhejiang Marine High-efficiency and Healthy Aquaculture, Ningbo, 315211, China

³ Ocean and Fishery Information Monitoring Center of Ningbo, Ningbo, 315010, China

⁴ Faculty of Architectural, Civil Engineering and Environment, Ningbo University, Ningbo, 315211, China

* For correspondence. E-mail: zhangdemin@nbu.edu.cn (Demin Zhang) or xiangshg@126.com (Xiangyu Zhu)

Summary of the numbers in supporting information:

The number of figures: 6

The number of tables: 3

Table S1. Spearman's correlation between microbial α -diversity indices and environmental variables.

	Richness	PD	Shannon	Evenness
Chl a	0.426*	0.494**	0.352	0.119
DO	0.386*	0.430*	0.070	-0.299
pH	0.289	0.368*	0.124	-0.125
Salinity	-0.491**	-0.531**	-0.331	-0.023
Ammonium	-0.467*	-0.475**	-0.090	0.337
Temperature	-0.423*	-0.465*	-0.195	0.131
Nitrite	-0.264	-0.318	-0.046	0.215
Nitrate	-0.024	-0.133	0.007	0.070
Phosphate	-0.245	-0.301	-0.025	0.237
Silicate	-0.116	-0.244	-0.072	0.046
COD	0.218	0.298	0.132	-0.035

** $P < 0.01$; * $P < 0.05$; PD: phylogenetic diversity; DO: dissolved oxygen; COD: chemical oxygen demand

Table S2. Random-forest analysis of the classification accuracy in each stage based on the 30 most abundant OTUs.

	ES	MS	LS	AB	Prediction accuracy (%)
ES	8	1	0	0	88.9
MS	0	10	0	0	100
LS	0	0	5	0	100
AB	0	0	0	5	100

Table S3. Simple Mantel tests for the correlations between environmental variables (Euclidean distance) and predicted gene table (Bray-Curtis distance) with 999 permutations.

Variables	ρ (Mantel test)	<i>P</i>
Nitrate	0.374	0.001
Phosphate	0.338	0.007
Nitrite	0.306	0.009
Silicate	0.187	0.004
Ammonium	0.134	0.092

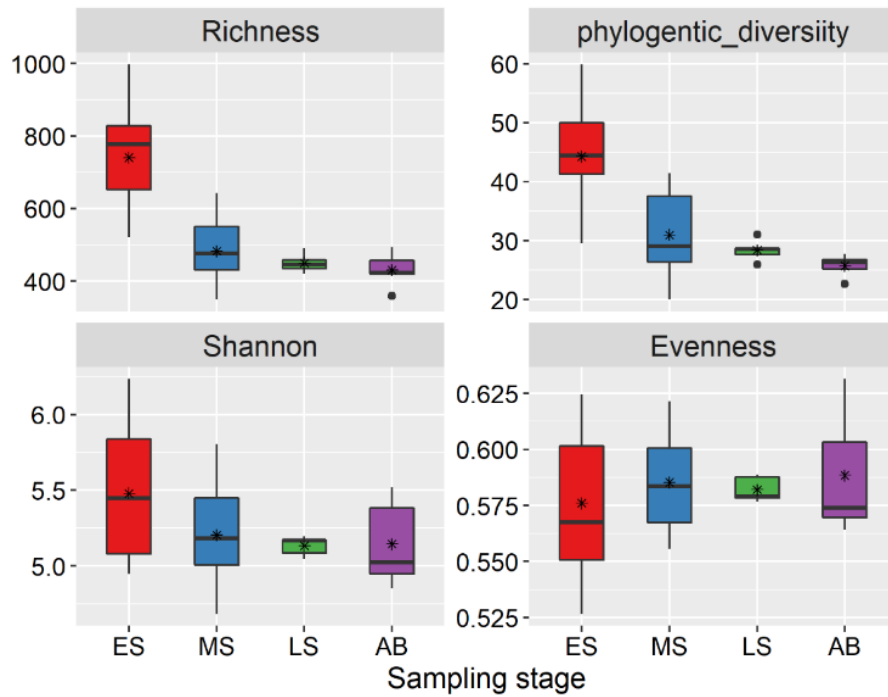


Figure. S1 Boxplots of α -diversity indices in each bloom stage, including richness, phylogenetic diversity, Shannon index, and evenness. The asterisk represents the average value.

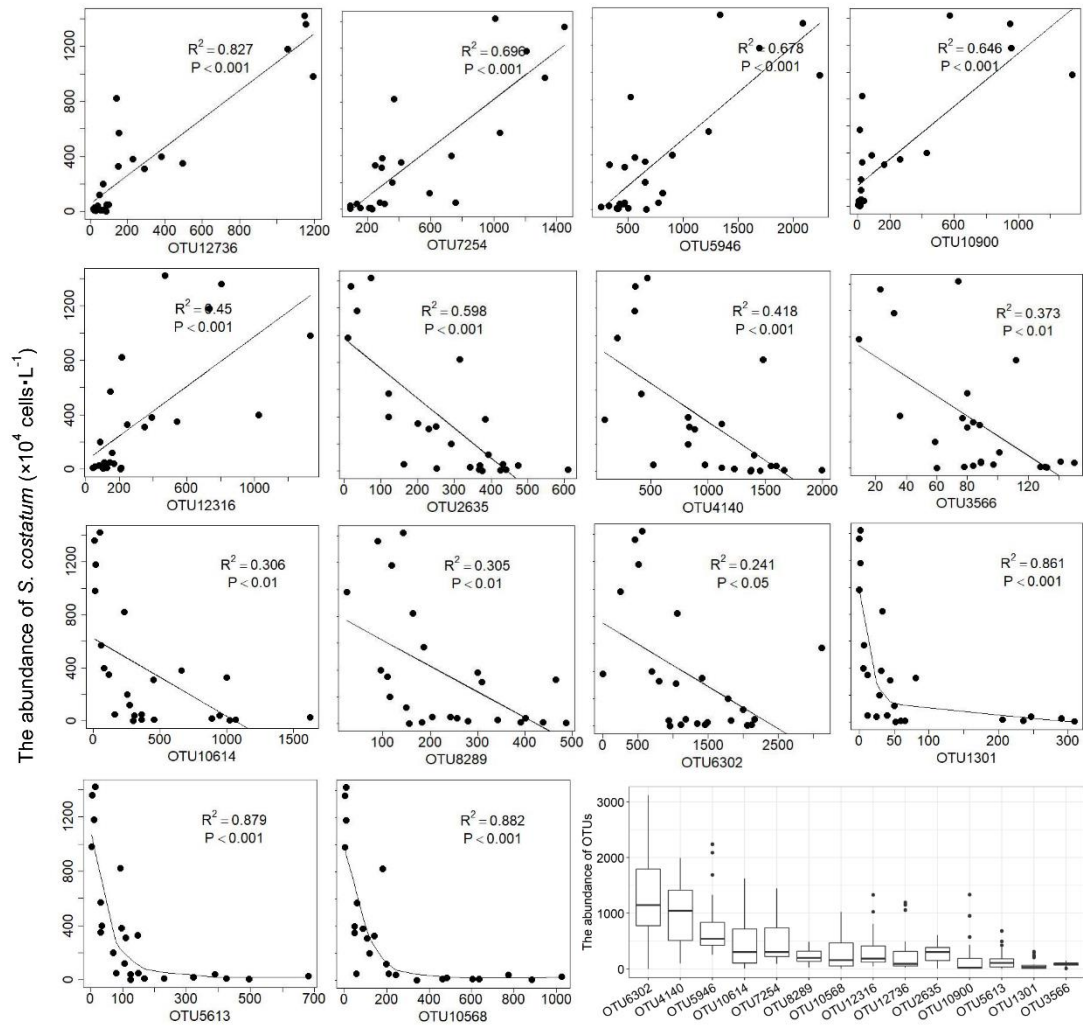


Figure. S2 Relationship between 14 of the 30 most abundant OTUs and the abundance of *S. costatum*. Linear model was done by “lm” in R, and non-linear model was calculated by “lowess” in R.

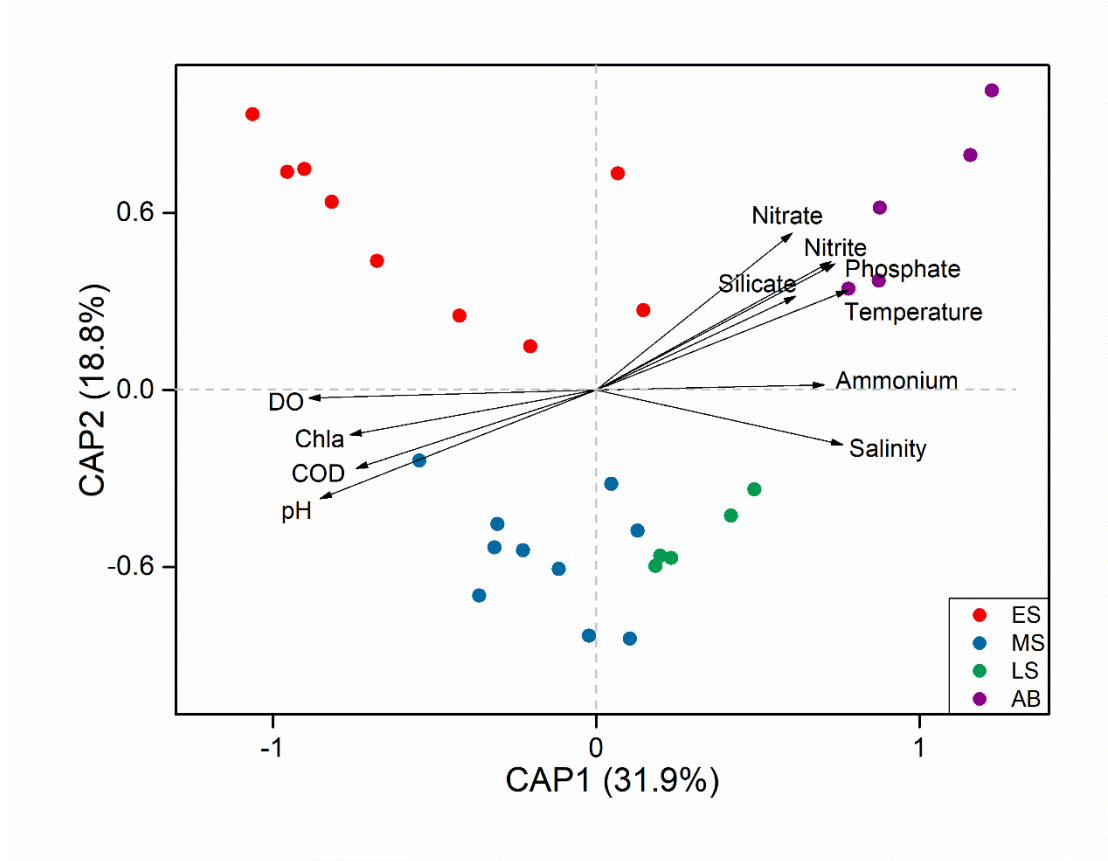


Figure. S3 Bray-Curtis distance-based redundancy analysis of the correlations between MCCs and environmental variables, according to bloom stages. DO, dissolved oxygen; COD, chemical oxygen demand.

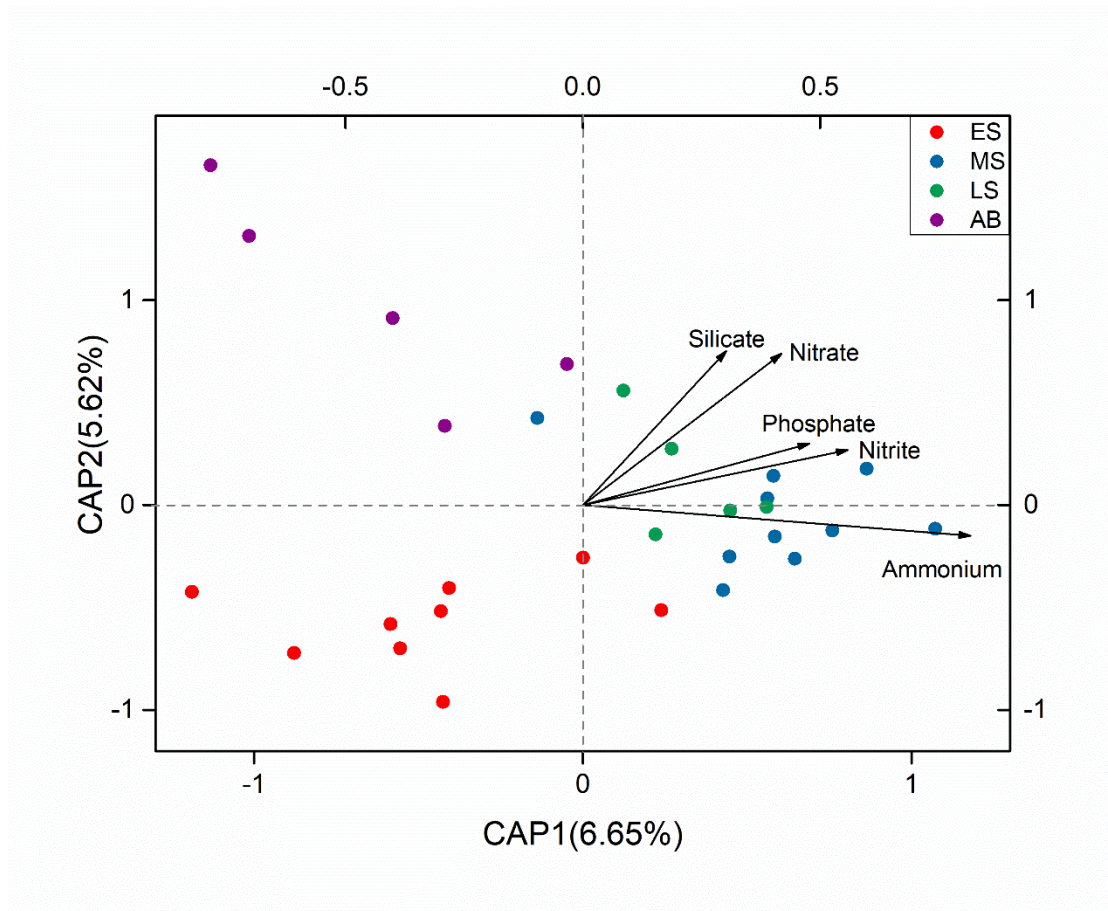


Figure. S4 Bray-Curtis distance-based redundancy analysis of the correlations between predicted gene ontologies and environmental variables, according to bloom stages.

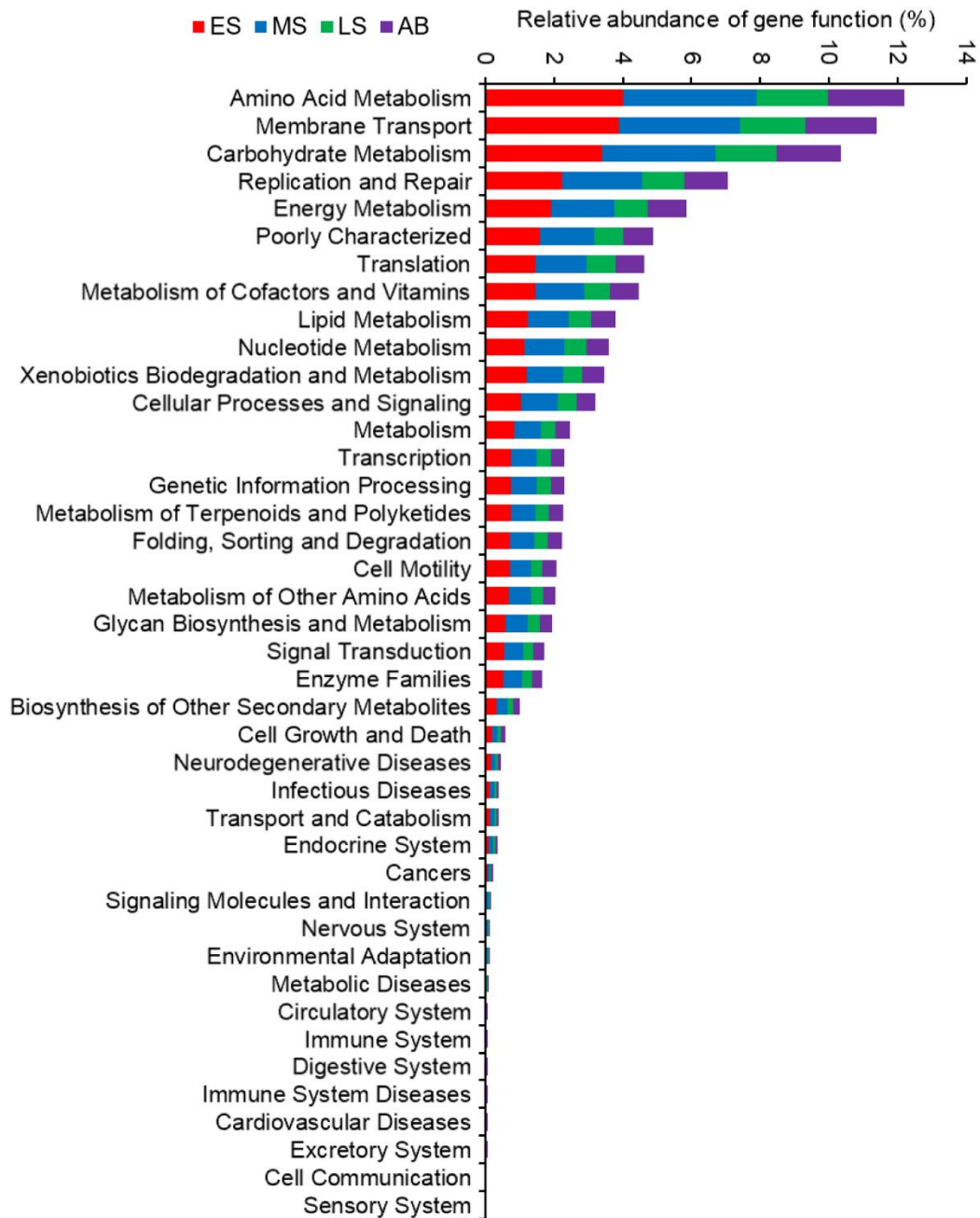


Figure. S5 Relative abundance of predicted functions of the microbial communities during the bloom process.

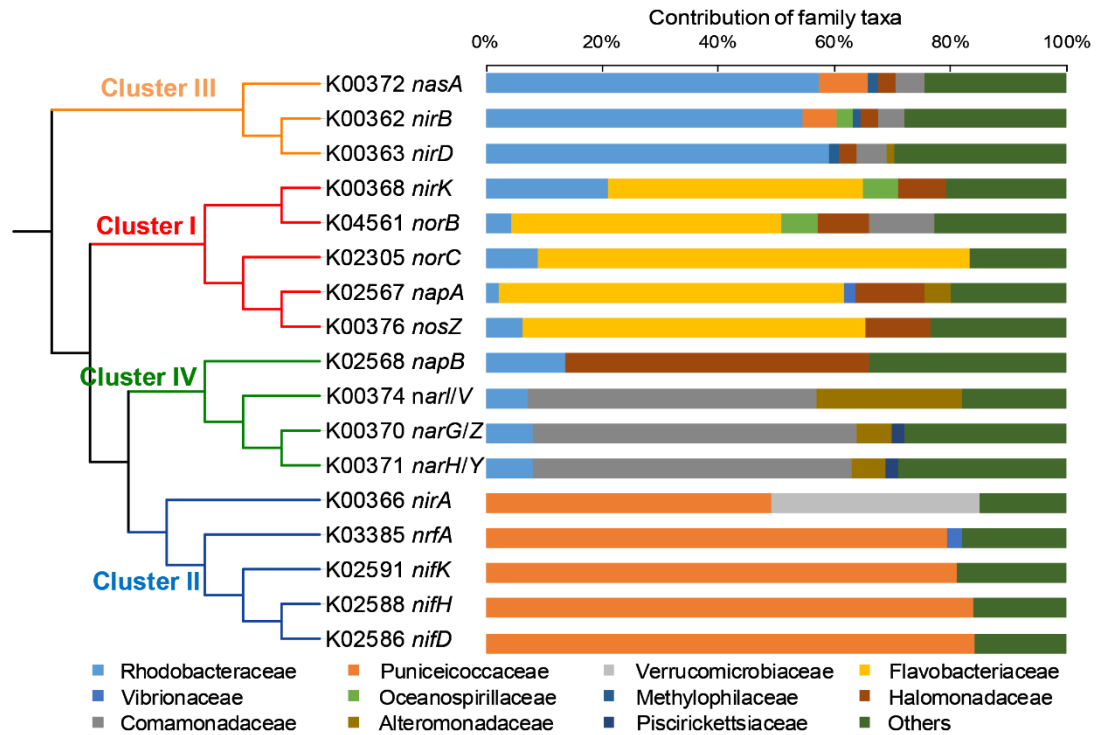


Figure. S6 Relative percentage contributions of bacterial families for nitrogen metabolism-related genes in imputed metagenome. Only contributions above 1% were selected. The predicted genes were clustered based on the Bray-Curtis distance.