藻菌关系的生态网络研究方法: 回顾与展望

朱建明^{1,2}, 周进^{1*}, 王慧^{3*}, 陈国福², 蔡中华¹

1. 清华大学深圳国际研究生院, 海洋工程研究院, 深圳 518055;

2. 哈尔滨工业大学(威海)海洋科学与技术学院, 威海 264209;

3. 汕头大学理学院, 汕头 515063

* 联系人, E-mail: zhou.jin@sz.tsinghua.edu.cn; wanghui@stu.edu.cn

2021-03-30 收稿, 2021-05-13 修回, 2021-05-17 接受, 2021-05-18 网络版发表

国家自然科学基金(41976126)、中国大洋矿产资源研究开发协会课题(DY135-E2-5-4)和深圳市科技创新委员会计划(RCJC20200714114433069, JCYJ20200109142818589, WDZC20200817153116001, JCYJ20200109142822787)资助

摘要 藻类与细菌是海洋生态系统中重要的组成部分, 两者间复杂的互作行为在维持共生关系、物质代谢以及协同进化中具有重要的生态意义. 以往的研究证实了藻菌关系中多样化的表现, 包括互利、共生、拮抗以及互害等. 然而, 由于藻菌关系的复杂性, 传统工具无法充分挖掘暗藏在藻菌之间的深度信息. 鉴于全球变化对海洋生态系统的影响和微生物组学的发展, 藻菌关系的研究视角也逐渐上升到系统生态学层面. 得益于高通量测序技术的快速发展和大数据分析能力的建立, 藻菌关系的研究迎来了新机遇, 其中微生物生态网络就是核心的一环, 它影响着共生关系的建立、交互行为的发生以及生态事件的走向. 鉴于此, 本文以大数据背景下的信息为基础, 尝试回顾与总结藻菌生态网络的研究进展, 包括网络的属性、网络的构建方法、网络的功能, 以及藻菌关系中共生模式(分子网络、代谢网络)与生态意义. 最后, 提出了建立藻菌互作知识网络数据库、藻菌生物互作图形数据库的构想, 以期充分发挥网络的鲁棒性来挖掘藻菌关系背后的全局性信息, 为深入揭示藻际微生态的分子机制及其在海洋生态系统中的多维作用提供借鉴.

关键词 藻菌关系, 微生物组学, 大数据, 生态网络, 分子机制

浮游藻类是海洋生态系统中初级生产力的最主要贡献者之一, 它们通过光合作用合成有机物并在藻细胞周围形成一个相对稳定的生态位——藻际环境(phycosphere)^[1]. 以此生态位为基础, 细菌与微藻产生密切的相互作用, 形成“藻菌共生体(algal-bacterial symbiosis)”. 该共生体是海洋环境中物质循环和能量代谢的重要场所, 在微食物环乃至整个生物地球化学循环过程中具有重要的作用^[2~5]. 聚焦于藻际环境, 最典型的藻菌关系包括: 互利共生(mutualism)、偏利共生(commensalism)、竞争(competition)以及寄生(parasitism)^[6]. 互利共生是一种代表性的生物间互作关系, 它能使共

生体双方彼此受益, 比如盐单胞菌向宿主藻类提供维生素B₁₂(VB₁₂)(图1), 藻类则通过光合作用合成有机碳供给盐单胞菌生长^[7~11]. 偏利共生又称共栖, 共生体之间的受益天平会偏向一方, 例如莱茵衣藻利用细菌分泌的VB₁₂, 但莱茵衣藻却未给细菌提供有机碳^[12]. 竞争则是由于物种生态位重叠引起的一种对限制性营养物质争夺的生态行为(图1). 例如, 在低磷酸盐条件下, 硅藻的生长受到限制, 但藻际细菌对磷酸盐的同化能力却高于宿主藻类^[13]. 寄生指一方的增益是以牺牲另一方的利益为代价. 例如, 一些具有杀藻活性的溶藻细菌(algicidal bacteria)在藻华(algal blooms)消亡过程中可抑

引用格式: 朱建明, 周进, 王慧, 等. 藻菌关系的生态网络研究方法: 回顾与展望. 科学通报, 2021, 66: 4378–4394

Zhu J M, Zhou J, Wang H, et al. The ecological network approach to algal-bacterial relationships: Review and prospects (in Chinese). Chin Sci Bull, 2021, 66: 4378–4394, doi: [10.1360/TB-2021-0314](https://doi.org/10.1360/TB-2021-0314)

图1 (网络版彩色)藻际生命共同体

Figure 1 (Color online) Model diagram of phycobiont

制藻细胞的生长或裂解藻细胞(图1),从而利用藻细胞裂解物用于自身生长^[14-16]。然而,藻菌之间的相互作用模式没有严格的边界,也并非恒定不变,互动模式多受环境因素和多种交互关系的影响^[17-23],包括营养交换(nutrient exchange)、信号传递(signal transduction)和水平基因转移(horizontal gene transfer)等^[24-26]。营养交换是藻菌互作最常见的一种形式,藻类通过胞外分泌物选择性吸附细菌,藻菌之间可完成氮、碳、磷、硫及微量元素的交换过程以满足各自的生理需求(图1)^[12,27-29]。信号物质在群体感应(quorum sensing)、化感(allelopathy)等方面起重要作用,通过信号响应来影响藻或菌的生长^[30-32]。基因转移则是一种协同进化的过程,是适应环境变化的一种策略,比如硅藻中鸟氨酸——尿素循环相关的酶基因被认为是来源于其藻际细菌的水平基因转移^[33]。在藻际环境中微生物的种类、数量远远大于宿主藻类,且它们直接影响宿主藻类生态功能的发挥,被称为浮游藻类的第二基因组。类似于根际生命共同体^[34],由“藻类、微生物、藻际环境”形成的稳定循环系统,具有“藻际生命共同体”的特征(phycobiont),在物质循环、生态位营造以及生态生理

学过程中发挥着重要作用(图1)。

尽管藻菌关系的研究已经报道了很多,但基于藻际环境中复杂的物质流、信息流和能量流,一些隐含的信息尚未被揭示。例如,藻菌之间的跨界交流、藻菌关系的时序表现,以及多维视角下“物种组成-功能代谢-环境交互”的全局性特征等。这些空缺导致当前藻菌关系研究局限于特定的区域、特定的共生体系以及特定的环境场景,致使研究结论的分散化、碎片化,系统程度偏低,极大地限制了对藻菌关系全局性的鸟瞰和预测。在此情形下,寻求新的突破点成为当下亟需。大数据时代的到来为更全面更深入地认识藻菌关系提供了可能。

基于大数据的网络分析成为了解析藻菌关系的新工具。新近的研究发现,在群落组成水平藻菌的共发生模式受网络关系的紧密调节,物种间的正、负相关性决定了宿主的生理状态和藻际微生物的演替^[35,36]。除了物种水平,代谢水平上也发现,碳循环、氮循环等功能基因具有明显的模块化特征,并与主导物种有直接关联。这种功能上的共发生模式决定了藻菌共生体状态,并影响着藻际生态事件(如赤潮)的发展轨迹。随着

研究的深入,网络世界的生态关系越来越受到关注,重要性也日益凸显.为了更好地理解藻际微生态学机制,本文以藻际微生物为重点,总结了网络的形成背景、构建方法、属性与功能,以及两种水平(物种水平-代谢水平)的网络意义;并提出基于藻菌互作的知识网络数据库构想.旨在梳理最新进展和提出可能的研究途径,为深入认识藻菌关系及赤潮事件的形成机制提供新的理论思路.

1 藻菌关系的新机遇——大数据时代

大数据具有四“V”特征:数据量大(volume)、类型及来源广(variety)、处理时效快(velocity),以及数据蕴含价值高(value)^[37].高通量测序技术的发展和组学手段的应用将微生物生态学推入了大数据时代.基因芯片和DNA微阵列(DNA microarrays)技术为藻菌关系的研究带来了丰富的数据^[38].然而,受限于探针的检测范围,上述方法无法得到藻际环境中所有微生物的功能信息.随后宏基因组技术(metagenome)的出现弥补了这一不足^[39],它通过非培养的方法对环境样品中所有微生物进行研究,避开了传统培养方法的局限,获得了理论意义上的所有环境微生物信息^[40].但宏基因组学的研究对象是DNA,因而得到的数据只能证实某种基因的存在,并不能证明基因功能的切实体现.宏转录组学(metatranscriptome)的出现很好地解决了这一问题.它针对性地研究特定条件下的生物群落功能^[41],标志着基因组学的研究由结构水平进入了功能水平.与宏转录组类似,宏蛋白组学是在蛋白水平研究微生物的功能^[42].它通过揭示环境中蛋白质的组成和丰度、修饰程度与蛋白间的相互关系,更进一步体现微生物群落的发展、相互关系与代谢功能特征.若要在更为下游的阶段了解业已发生的生理水平,代谢组学才是真正意义上的表征手段.与上游的转录组学和蛋白质组学相比,代谢组学能更直接反映机体的生理功能.

上述几种组学方法应用于藻菌关系有过诸多探索,并获得了基于大数据特征的藻菌关系信息.有效解析这些数据成为当下的核心任务.考虑到微生物生态系统中各物种之间并不是单一存在,也并非简单的“加法”和“减法”,而是以综合的共生体形式存在,并通过物质交换、信号交流、能量传输等方式形成复杂的交互网络^[43].这要求我们有强大的生物信息学分析能力,整合生物分子之间的多维信息,并以可视化方式对系统功能作出解释,凸显出信息背后的生物学意义.为

此,这一背景为藻菌关系的网络研究铺设了道路.

2 生态网络的构建方法及其特征

系统生物学是研究复杂系统的一门新兴交叉学科,主要通过对多种来源的大数据进行分析,从整体水平解释系统内的各种机制^[44].实现这一目的最常见的工具是网络法,常见的网络参数及类型归纳于表1.随着研究的深入,系统生物学衍生出新的理论分支——复杂生物网络理论.它在传统网络分析的基础上,再辅以各种算法、软件等手段对藻菌大数据进行处理、分析和可视化,将微生物生态系统内多个物种的相互作用关系抽象成网络上的几何结构,通过网络的拓扑结构及鲁棒属性来映射现实的生物学意义^[48].对于许多数据实体间关系不明确或难以直接观测的数据集(如原位生态环境中无法直接观测到的彼此关系),分析数据间的相关性是建立实体间关联的一种有效方式.目前针对复杂生物网络的构建方法主要有两种:简单系统中是基于物种间的同现或者相关性(相似度、相关度、互信息)来预测成对关系;而对于复杂系统,则采用回归等方法来预测多物种之间的相互作用^[49],利用可视化软件(如Cytoscape^[50]、Gephi^[51]、R包igraph^[52]等)将预测关系进行数学转化,从而揭示微生物群落对生态系统功能构建、维持以及演替的影响.近年来,针对复杂生物网络的构建工具得到了长足发展,逐渐满足了各种特征数据的网络构建与分析(表2).

网络基本特征主要包括“点、连线”、模块化和关键节点等信息.最基本的单元是“点”和“连线”,“点”可代表现实生态系统中微生物、基因、蛋白质和环境因子等;而“连线”代表它们之间的相互作用,包括促进、抑制、拮抗或致死等.对模块化和关键节点特征分述如下.

(i) 网络的模块性(modular structure).基于网络的小世界性,具有相似属性或者功能节点的生物元件往往会聚集形成一个团体,即生物网络.生物网络分析从两个角度入手:自上而下(基于网络分析找出关键节点和功能模块)和自下而上(根据关注的节点分析与之相关的其他节点和功能模块).生物网络中的聚类或网络簇(network cluster)表示生态位的状态.当模块内的物种出现共发生或者正相关关系时,表示生态位的重叠;而当出现共排斥或者负相关关系时,则表示生态位的分化^[62].与网络簇相似,模块性也是生物网络中重要的表征指标.它可以指示环境因素和生物因素在微生物

表 1 常见参数及网络类型

Table 1 Common parameters and network types

网络参数	参数意义
度	在网络中节点的度等于连接节点的边个数. 网络节点的平均度(K)等于总边数的2倍除以总节点数, 它可以反映网络结构的稀疏程度. 网络的度分布定义为 k 值节点数目占总节点数的比例, 即 k 的概率 $p(k)$ ^[45]
最短路径	表示能联通网络中两个节点的最少边数. 把所有节点之间的路径长度最长的值称为网络的直径(diameter). 节点之间最短路径长度在网络中模块结构(modular structure)的形成过程中起重要作用
聚集系数	用来描述1个节点相邻节点之间边的连接情况. 如1个节点 r 有 n 个直接连接的相邻节点, 如果这 n 个节点之间都相互连接, 则应该有 $n(n-1)/2$ 条边, 但实际网络中只有 m 条边, 则该节点 r 的聚集系数为 $2m/(n(n-1))$. 聚集系数描述了节点的相邻节点之间连接的可能性. 整个网络的聚集系数(平均聚集系数, C)是网络中所有节点聚集系数的平均值, 反映了网络的模块性质, C 越大表明网络中的模块结构越多
中心度	中心度是衡量某个节点在整个网络中作用的重要定量指标. 中心度涉及节点之间的互作能力, 通过对节点中心度的计算, 可以找出网络的关键节点(key-node)或者枢纽节点(hub-node). 常用于衡量节点重要性的中心度指标有介数中心度、接近中心度、点度中心度等 ^[46] . 节点的介数是衡量节点中心度的重要指标, 节点的介数越大, 中心度越大, 枢纽性更强. 接近中心度大的节点可以快速影响网络的其他点, 在网络结构中也具有重要作用
网络的基本类型	
规则网络	指网络中各节点的连接关系服从一定的规则, 比如全局耦合网络(任意两节点之间都有连接边)具有最多的边数, 最大聚集系数 C 为1, 最小平均路径长度 L 为1
随机网络	与全局耦合网络对应, 随机网络的度分布服从泊松分布, 平均聚集系数小, 平均路径长度也小. 规则网络和随机网络代表了网络演化的两个极端情况, 是网络演化分析的两个基本参考模型
小世界网络	现实世界中, 许多网络都具有小世界网络的特性, 即网络的平均路径长度很小, 但是网络的平均聚集系数远远大于相同规模的随机网络, 但小世界网络的度分布也呈泊松分布
无标度网络	由于随机网络和小世界网络的度分布都是呈钟形曲线的泊松分布, 则其分布曲线具有明显的峰值——特征长度, 但许多实际网络如生物网络的度分布没有特征长度, 则成为一种新的网络模型——无标度网络. 无标度网络的度分布符合幂律分布, 即 $P(k) \sim k^{-r}$ (r 在2~3之间) ^[47] . 其显著的特点在于多数节点有少点的边连接, 而少数节点具有较多的连接边数, 即大多数节点的度很小, 但存在少数节点的度远远大于网络的平均度. 这些高连接点的节点可以被视为关键节点

群落构建中的相对重要性^[63]. 当系统发育水平上亲缘关系较近的物种表现为正相关, 亲缘关系较远的物种表现为负相关时, 意味着确定性过程在微生物群落构建中占主要地位; 反之则意味着竞争作用占主要地位^[64]. 对网络簇而言, 现有的主要分析方法有edge-betweenness、fastgreedy、label propagation、leading eigenvector、multilevel、spinglass、simulated annealing等^[65]. 而模块性分析可通过R软件包“igraph”和Cytoscape软件的功能插件“MCODE”^[66]等来实现.

(ii) 关键节点(key-node). 在微生物生态网络中关键物种和关键因子的识别非常重要, 这经常需要借助关键节点来判断. 网络中节点的属性可以映射为关键物种(或基因、蛋白等)、枢纽物种(connector)和连接器物种(hubs). 枢纽物种在网络中具有高度连接性, 即节点的“度”具有较高的赋值. 连接器物种虽然不具备较高的节点度, 但却是两个或者多个模块的连接节点, 在很大程度上影响着网络的连通性. 目前检测关键节点的方法主要有中心度指数. 该指数能够指出已知网络中哪些节点处于中心位置, 中心性更高的节点倾向于

拥有更多的链接, 能更快地与其他节点建立联系, 并控制它与其他节点之间的信息流^[67].

另外一种指示节点的方法是“模块内连通度(Z_i)和模块间连通度(P_i)”^[68]. 其中 Z_i 是衡量一个节点在模块内部中的作用, Z_i 值越高说明该节点在模块内部的作用越大; 同样 P_i 是衡量一个节点在其他模块的参与程度, P_i 值越高说明该节点参与其他模块的联系越紧密. 根据 Z_i 和 P_i 的范围可以将节点在网络中的拓扑角色分为4种: 外围节点(peripherals)($Z_i \leq 2.5$, $P_i \leq 0.62$), 这类节点只与模块内的某些节点连接且与其他模块的节点不连接; 模块核心(module hubs)($Z_i > 2.5$, $P_i \leq 0.62$), 这类节点与模块内的其他节点有多个连接; 连接器(connectors)($Z_i \leq 2.5$, $P_i > 0.62$), 该类节点不专属于某一个模块, 且与多个模块的点具有连接; 网络核心(network hubs)($Z_i > 2.5$, $P_i > 0.62$), 这里节点是整个网络的核心, 既是模块核心又是连接器^[69]. 由此看出, 网络核心节点最为重要, 它影响了整个网络的连通性和稳定性. 需要指出的是, 不是每一个生物网络都具有网络核心, 在缺失网络核心的情形下, 模块核心和连接器成为了替代重

表 2 用于复杂生物网络分析的常见工具

Table 2 Common tools for analysis of complex biological networks

工具名称	描述	相关信息	文献
CoNet	Co-occurrence network inference, 该方法引入一种新广义提升线性模型Generalized Boosted Linear Model(GBLM), 并整合了 Pearson、Spearman、Bray-Curtis等多种相关性、相似性和距离度量, 且可以整合环境因子数据	软件主页 http://psbweb05.psb.ugent.be/conet/	[53]
SPIEC-EASI	SParseInversECovariance Estimation for Ecological Association Inference, 估计偏相关关系性, 主要应用于解决数据的非独立性、数据的间接相关性和过度拟合等问题	软件代码 https://github.com/zdk123/SpiecEasi	[54]
SparCC	Sparse correlations for compositional data, 基于变量间(物种)相互作用关系稀疏的假设, 是一种比Pearson相关性更加稳健的相关性分析方法. 通过将数据归一化以消除成分效应, 以较高的准确度发现群落中的相互关联关系	软件代码 https://bitbucket.org/yonatanf/sparcc/src/default/	[55]
MENA	将物理学中的随机矩阵理论(random matrix theory, RMT)应用到共发生网络构建中, 来自动确定相关系数的阈值, 避免由人为设定相关系数阈值、显著性阈值等来构建共发生网络而存在的主观性	软件主页 http://ieg4.rccc.ou.edu/MENA/	[56]
Pearson, Spearman	基于R语言直接计算各变量(物种、环境因子等)的Pearson(线性相关)、Spearman(线性相关或者非线性相关)相关性系数, 经显著性检验后筛选P值小于0.05或0.01的关系对作为网络作图软件的输入数据作图	R语言 https://www.r-project.org/	[57]
eLSA	Extended local similarity analysis, 考虑个变量(物种)之间的相互作用关系存在时间上的延迟, 基于等距时间点的数据可以验证变量间的相互作用强度, 并基于变化的先后顺序在一定程度上可以解释关系对的因果关系	软件代码 https://bitbucket.org/charade/elsa/src/master/	[58]
Cross-correlation	交叉相关或者互相关, 计算出的两个时间序列数据之间的其中一个相对于另一个滞后的相关性. 主要应用于计算一个变量(物种)对另一个物种或环境参数的延迟响应. 所以要求数据是等距时间点的样本数据	主要通过R软件包stats中的ccf函数实现 https://cran.r-project.org/web/packages/statsr/index.html	[59]
DBN	Dynamic Bayesian network, 利用贝叶斯算法对大量基于时间序列数据的计算, 并对该复杂网络进行参数估计, 预测不同环境条件下群落结构的动态变化	CGBayesNets: Conditional Gaussian Bayesian Network and Continuous Dat. http://www.cgbayesnets.com	[60]
MDSINE	Microbial Dynamical Systems INference Engine for microbiome time-series analyses. 基于最大似然法或贝叶斯法拟合模型, 从而获得物种间的相互作用关系参数; 可以通过模拟不同外界环境条件下网络结构/群落结构的动态变化, 另一方面可以在此基础上进行关键物种的鉴别	软件代码 https://zenodo.org/record/50624#.XZmnwXaTfjx	[61]

点. 此外, 着眼网络内部节点时, 不能忽视外围节点的重要性, 如果外围节点的数目庞大, 它对生态系统的功能影响也具有不可忽略性, 因此内外节点需要并行考虑.

由此, 关键节点和模块属性在网络分析过程中具有重要地位. 前者在生物网络分析中可以找到影响系统的关键物种或因子, 后者可以找到相互作用的功能团体. 两个关键属性作为生物网络分析的重要基本单元, 为深入理解生态系统的演变机理提供了可能. 此外, 基于网络的模块性和关键节点的分析, 也可以找出生态系统中关键的功能群体, 简化复杂生物网络, 便于更针对性地理解生态系统中群落结构模型和生态过程.

从整体维度来看, 借助复杂生物网络对藻菌共生体进行研究将是藻菌共生关系研究的一个新视角. 基于组学大数据建立藻际微生物互作网络, 并将网络用生物信息学手段进行可视化, 可驱动藻菌体系的互作认识进入新层次.

3 藻菌关系的网络类型

3.1 分子生态网落(基于物种与功能水平)

在物种水平上, 基于组学数据或者16S rRNA/18S rRNA/ITS等扩增子测序数据, 以Pearson或Spearman等相关系数构建矩阵网络, 借助图形的可视化可以获得

基于丰度组成的藻菌关系,如正相关、负相关以及程度的强弱.在矩阵网络的构建中,阈值系数的设定非常关键,它影响着即时的相互关系及生态事件的后续走向.比如,在计算赤潮过程中藻菌相对丰度变化时,通常设置阈值的绝对值为 $r=0.6$,显著水平为 $P=0.01$,当两个分析单元的相关系数 r 的绝对值大于0.6且 P 小于0.01时,认为两者相关紧密;反之,则认为相关不紧密.阈值的设定除了依赖经验值,也需要遵循具体的场景.适当的阈值设定不会使阳性结果被掩盖,或阴性结果被夸大.在科学阈值探索中,周集中课题组^[70,71]提出了基于RMT的网络构建方法,该方法能自动获取阈值消除数据中的噪音干扰.在此理论基础上,Zhou等人^[72]开发了一种非参数工具(molecular ecological network analysis, MENA),对抵抗噪声具有很强的鲁棒性.依托该方法研究了长期升温对草地土壤微生物群落中分子生态网络复杂性和稳定性的影响,结果显示,升温显著增加了网络的复杂性,包括覆盖面大小、连接性、平均聚类系数以及相对模块性和关键种类的数量.且升温条件下的分子生态网络变得更加稳健,从而证明了气候变暖显著增强了网络结构功能潜力在整个生态系统中的地位^[73].MENA方法还适应于实验组别较少的环境.比如,针对藻菌共培养研究,如果只设计了对照组和实验组,此时只需将对照组和实验组的重复数设计到8个以上,便可利用MENA方法进行网络作图.通过对拓扑结构、节点属性和模块性进行对比,揭示藻菌之间的相互影响及其环境因子对它们的塑造作用.前期,我们实验室通过对深圳湾两个站点8个重复的采样工作,分析了16S和18S的物种共发生模式,证实变形杆菌和蓝细菌是甲藻中藻际细菌的关键物种,拟杆菌和厚壁菌是硅藻中藻际细菌的关键物种.这些相对致密的网络形态构成了整个原核生物和真核生物的稳定基石(宋晓等人,未发表数据).

上述方法是在不考虑时间序列下的简单模型,为了应对多维度的需要已开发出针对藻菌横截面和纵向面(如时间序列)的实验设计及数据集,以呈现藻际环境中宿主与微生物的网络互动.事实上,赤潮生消过程中藻类、细菌、环境因子、理化参数等随时间的变化趋势不是单纯的线性关系,而是非线性状态.因此,采样时间、频率和数量的设置将影响最终的分析结果.以赤潮事件为例,示例主导细菌物种(B1、B2、B3、B4)与宿主藻类在不同时间断面上的关系.图2展示了藻菌在时间序列下呈现非线性和周期性的变化规律,B1菌

图2 (网络版彩色)藻菌趋势时序变化关系示意图.实线表示藻(A)的相对丰度变化趋势,波形虚线表示不同种类的藻际细菌(B1、B2、B3、B4),纵轴表示相对丰度,横轴表示不同时间序列或者环境因子梯度

Figure 2 (Color online) Schematic diagram of time series change of algae-bacteria. The solid line indicates the trend of relative abundance of alga (A), while the wavy dotted line indicates different species of algal bacteria (B1, B2, B3, B4). The vertical axis represents relative abundance, and the horizontal axis represents gradients of different time series or environmental factors

与藻表现出延迟性正相关,且B1菌的变化先于藻,由此推断B1菌的改变影响其宿主的变化.B2菌与藻表现出正相关关系,B3菌与藻表现出延迟性正相关,B4菌与藻表现出显著负相关.如果考虑时间序列,以B3菌为例,当取样点聚焦于截面2~3,藻菌关系呈显著正相关;关注于截面2~4,则可能呈不显著的正或者负相关;当视角进入截面3~4,可能呈显著负相关.如果将截面1~6进行综合来看,则呈延迟性正相关,且藻的变化先于B3菌,即推断藻的变化对菌B3有明显影响.这些结果可以帮助我们判断藻菌关系中的主动因素和被动因素.而常规脱离时间维度的数据获取方式常常掩盖了这一信息.考虑到赤潮生消的时间尺度中藻菌变化趋势常常不同步,所以时序上的数据挖掘才能更准确地揭示藻菌间的真实规律.

为了挖掘延时性背后隐藏的信息,一种新的分析方法——局部相关性分析LSA(local similarity analysis)应运而生^[74].近年来,LSA及其延伸型eLSA被广泛应用于生物学研究中^[58,75~77].2013年,Paever等人^[78]通过LSA分析了不同时间点3个湖中的藻及细菌的群落组成,直观地展示了赤潮生消过程中与优势藻保持正相关和负相关的菌群.2014年,Liu等人^[79]通过eLSA分析了蓝藻到硅藻的演替过程中细菌的相对丰度和藻的生物量变化趋势,发现82.6%的OTUs(operational taxonomic units)与藻的演替有很强的相关性,且正相关大于负相关.这

些菌大多属于酸杆菌门(*Acidobacteria*)、放线菌门(*Actinobacteria*)和变形菌门(*Proteobacteria*)。2016年, Needham等人^[80]通过对圣佩德罗海春季赤潮的每天持续追踪采样, 基于eLSA算法构建了简单的(未涉及复杂网络拓扑结构分析)由细菌、古菌、浮游植物和环境因子组成的时滞性分子生态网络。研究发现, 浮游植物与环境因子之间的相关性较低(与营养物质之间的相关性为零), 浮游植物和原核生物群落之间的相互关系比与环境因子之间的相互关系更大。该结果突破了传统意义上赤潮生消过程由物理因子和营养盐主导的认知, 强调微生物间的相互作用在赤潮发生、维持和消亡过程同样具有重要作用。相比之前的二维关系, 增加的时间维度有利于将网络关系的相关性提升到因果性上来。此外, 我们也找到了影响赤潮生消进程的生物因子(细菌, 如海洋单胞菌、弧菌、海洋细杆菌等)和环境因子

(如硝酸盐、磷酸盐、铵盐等), 证实了藻和菌对环境因子的响应具有不同步性(图3)^[81].

鉴于单纯依赖物种丰度信息的局限性和不同生物类群可能执行相似生物功能的可能性, 基于功能网络的分析显得更为重要。组学数据下的功能网络为理解藻菌交互机制提供了新的切入点。宏基因组分子生态网络可以将各个特定物种间的关系联系起来, 也可以根据归类酶的功能和数量筛选出群落中特定代谢关键物种。宏转录组的分子生态网络可以了解群落的活性功能, 区分不同时间和不同条件下群落的活性基因, 展示全局性的调控因子和靶标基因之间的关系。而基于蛋白质层面的组学数据, 可以比较不同群落之间的蛋白质组成差异以及构建蛋白质互作网络, 将基因水平的可能性提升到蛋白水平的确定性。这些方法的实施多以模块为关注重点。根据模块的保守性和可变性可

图 3 (网络版彩色)时滞性藻菌分子生态网络。圆点表示藻际微生物, 菱形表示微藻, 三角表示环境因子, 实线表示两者正相关关系, 虚线表示两者负相关关系。连线顶端的图形表示指向变化滞后的一方(三角形表示两者互响应滞后一个时间点的延迟, 圆形表示两者互响应滞后两个时间点的延迟, 菱形表示两者互响应滞后三个时间点的延迟)^[81]

Figure 3 (Color online) Time-delayed molecular ecological network of algae-bacteria. The dots represent microorganisms, the diamonds represent microalgae, the triangles represent environmental factors, the solid lines indicate a positive correlation, and the dotted lines indicate a negative correlation. The graph at the top of the line points to creatures that respond with a lag time (triangle for one time lag, circle for two time lag, diamond for three time lag)^[81]

以在群落水平上将特定酶的信息联系到所属物种,一方面可以将各个特定物种间的关系联系起来,另一方面也可以根据归类酶的功能和数量筛选出群落中特定的关键物种^[82]。基于功能模块的分子生态网络,可以了解群落的活性功能;也可以发现调控因子和靶标基因之间的关系。例如,McClure等人^[83]基于基因共表达网络研究了土壤微生物组对干旱的响应机制,发现一些代谢途径(如糖异生、淀粉和蔗糖代谢)、呼吸、抗生素基因在网络中占据中心位置,并结合代谢组数据验证了次生代谢产物(例如安沙霉素和铁载体)的生物合成以及碳代谢(淀粉和蔗糖代谢、戊糖磷酸途径和糖酵解/糖异生)对土壤微生物群落应对干旱条件的作用。模块网络中除了基因元件,也可以在蛋白层面构建互作网络展现酶、底物、效应物之间的相关性,呈现代谢物种类、数量与生物群落功能、生化特征之间的直接联系。因此,这种基于蛋白和代谢水平上的藻菌互作网络分析将更有利于深入揭示藻菌之间的相互机制。

与物种水平的网络分析类似,功能水平上的分析也需要注意时间尺度问题。这是因为藻菌间的互作关系在不同时间段或者不同环境条件下有本质的差异。例如,赫氏颗石藻*Emiliania huxleyi*在生长初期为藻际细菌*Phaebacter inhibens*提供营养物质,该细菌为藻提供生长激素,此时为互利共生关系;但是当赫氏颗石藻衰亡时会释放p-香豆酸刺激细菌合成溶藻物质,两者即转化为敌对关系^[84]。动力学数据也表明,不同生物之间的相互关系随时间而不同,表明不同生物过程在不同时间尺度上运作^[85]。因而在藻菌关系的研究中,时间尺度的把握对研究结果的准确性至关重要。如果时间尺度过大,会造成较高的异质性,对网络的计算带来较大偏差。例如,在研究藻菌之间的相互作用关系时,如果脱离时间轴,那么相关性系数无法直观表示物种间复杂的演变过程,而是仅仅得到非正即负的片面结果。如果此时再辅以空间(不同水层)、季节(冬季和夏季)和环境因子(寡营养和富营养)的影响,那么获得的结构将会偏离实际情况更远。因此,时序性在藻菌关系的网络构建中是不可忽略的因素,是避免以偏概全的基本保障。

3.2 生态系统代谢网络

网络中的相关性数据只是表现展示了藻菌之间的相关关系,缺少对代谢能力的认识。将功能信息联合考虑才能形成实质性的网络系统,即生态系统代谢网络。区别于其他生态网络的特点在于它是基于代谢通路信

息的计算、归纳与演绎,而不是基于分子物质的相关统计和推断。作为复杂非线性动态调控网络,代谢网络是将一系列生化过程如合成代谢、分解代谢、运输系统整合成可调节的无尺度网络^[86,87]。基于它更强的系统性,代谢网络近年来越来越受到重视。

自Jeong等人^[88]首次提出代谢网络以来,该方法主要集中在代谢网络的重构和结构功能分析,包括酵母网络和其他生物体的代谢网络^[76]。但是这些代谢网络大多是基于模式物种或单个生物体。环境中功能微生物在生态系统中并不是独立的行使功能,而是通过网络结构彼此联系^[89]。这种网络结构称为微生物生态系统代谢网络(microbial eco-system metabolic network)。该网络以微生物为节点、以各个微生物之间的代谢关系为依托,从全局角度揭示特定环境条件下微生物之间的物质循环和能量流动规律,以及各个微生物之间的彼此作用。

除了功能上的聚焦,代谢网络也不再局限于单个物种,而是基于元素循环和信号传递构成整合型跨种甚至跨界间的物质代谢信息,因而具有更强的阐释系统应对外界环境变化的能力。功能水平上的生态系统代谢网络除了电子传递和物质代谢水平的互作关系,还包括由物种间营养缺陷所形成的相互依赖关系。相对于前者,后者更能体现出各物种之间的主动依赖性和专一性^[90]。研究发现,在双酚A的降解中,*Sphingomonas* spp.(鞘氨醇单孢菌)、*Pseudomonas* sp.(假单胞菌)和*Pusillimonas* sp.(极小单胞菌)具有网络协作功能。其中*Pseudomonas* sp.和*Pusillimonas* sp.的交互作用为*Sphingomonas* spp.提供了所需养分,*Pseudomonas* sp.对BPA的降解缓解了酚类物质对*Pseudomonas* sp.和*Pusillimonas* sp.两种菌的毒害。这种协作互惠的交流方式维系了共生体的稳态,也加速了对异质类物质的清除速度^[91]。这种复杂的多元互作关系得到了代谢网络的诠释,进而证实多物种共生关系维持了微生物群落结构的稳定性。此外,在群落功能水平,利用生态系统代谢网络探究了多种微生物互作对木质纤维素的消化机制,展现了生态系统代谢网络较好的应用前景。以白蚁肠道微生物为例,Kundu等人^[92]利用已公布的微生物代谢途径信息,开发了由205种微生物和265种代谢物组成的生态系统代谢网络。通过分析发现,37个关键微生物在维持微生物群落稳定和执行功能中具有重要的作用,进一步获得了15种关键微生物利用半纤维素的酶学机制。生态系统代谢网络不仅能展示特定生态系统

内物种间的相互作用,还可用于模拟和预测该系统内微生物对特定环境胁迫下的共同响应.基于流平衡分析(flux balance analysis, FBA)的计算方法通常用于模拟不同营养条件下微生物代谢网络的建模^[93].例如,基于功能基因的注释信息可以得到系统内微生物的代谢能力谱,通过基因组水平代谢建模工具(AutoKEE-GRec、AuReMe、CarveMe、ModelSEED、Kbase等^[94])和数据库(AGORA^[95]、BIGG^[96]等),将其转化为FBA模型,并重构成代谢模型(genomic-scale metabolic models, GEM)^[97],即生态系统代谢网络.利用该方法的分析过程可以将一些代谢物的流通量设置为零来模拟基因敲除或某些功能微生物缺失的情况.通过这种方式可以评估不同条件下微生物群落的生存力,以及向群落中添加新物种对宿主和生态系统的影响.FBA不仅可用于研究细菌之间的相互作用,还可以整合宿主的信息用于研究宿主与微生物的相互作用.例如,Heinken等人^[98]在代谢网络建模中整合了拟杆菌和小鼠细胞的代谢能力谱,通过研究宿主与其微生物之间的代谢依赖性证实,拟杆菌可能通过向宿主细胞提供必需和非必需氨基酸来影响小鼠细胞的生长.当前基于FBA的代谢网络建模方法还只局限于有限的物种互作关系,更大范围的模拟算法与数据运算能力仍是多物种生态系统代谢网络建模分析的瓶颈.

聚焦于藻际环境,生态系统代谢网络除了反映物质代谢和功能上的特征,也可反映藻际生态系统的形成和演化机理.这主要源自代谢网络的小世界性、无尺度性和模块化拓扑特征.代谢网络的小世界性特征意味着代谢物浓度的变化能迅速地传递到整个网络,使整个系统对外界环境变化作出及时响应,从而增强系统的稳定性^[99].代谢网络的无尺度特性表明,演化过程中新节点的出现将倾向于与高连接度的节点连接,

且服从幂函数分布,节点度越高意味着在网络中存在的时间越久^[100].代谢网络的模块化特征一般表示功能的聚类以及更高层次的嵌套特征^[101].基于生态系统代谢网络的拓扑结构特征对藻菌互作关系的深入研究具有重要的指导意义,如研究藻菌群落组成及生态功能的动力学过程、物种间的进化关系以及对环境的共适应机制等.此外,代谢网络还可以评估生物与环境之间的复杂关系.Dittami等人^[102]提出了藻菌之间生态系统代谢网络模型的构建方法,即首先基于泛基因组构建一个生态系统代谢网络框架,随后基于间隙填充算法校正代谢网络重建中缺失的环节;最后迭代进行网络的填补直到获得完整的生态系统代谢网络.由此产生的“泛网络”最大程度上表征了共生系统中所有的藻菌相互作用关系,为藻菌共生体响应外界环境变化研究提供了新的思路.综上所述,未来针对藻菌生态系统代谢网络,主要方法是通过组学技术(宏基因组、宏转录组、宏代谢组等)检测藻菌共生体系统中相关生物组分,获取生物标记物,再通过参考数据库或实验的方法建立表征藻菌间各种生物过程的代谢网络,借助不同条件下藻菌生态系统代谢网络的差异,找出关键的节点(基因、酶等),从而全面系统地揭示藻菌互作应对外界环境胁迫的机制(图4).

3.3 动态生物网络

分子生态网络的拓扑属性可以简化复杂的网络系统,这带来分析上的方便.然而,简化的信息也会带来失真的结论.例如,Connector或者Hubs等节点常被解读成关键种,它们的去留需要依据不同的情况加以分析.在某些情况下,中心节点的连接点数可能是由于生态位的重叠所致,为此中心节点的去除并不会引起微生物群落发生较大改变^[103].相反,如果连接数少的节点

图4 (网络版彩色)藻菌生态系统代谢网络分析

Figure 4 (Color online) Microbial eco-system metabolic network analysis of algal-bacterial ecosystem

是食物链顶端的捕食者, 它的去除会对微生物群落结构产生重大影响^[69,104]。为了规避分子生态网络的瑕疵, 动态模型网络应运而生。它更适合研究系统中某些物种的存在与否对系统造成的影响, 从而确定该物种在系统中作用的重要性。

生物系统本质上是随时间动态演化的, 但上述分子生态网络和基于FBA的代谢网络建模通过一次网络模拟只能捕捉到微生物群落静态特征, 无法捕获物种和代谢水平的动态变化。对动态生物网络中模块、关键节点的识别与演化分析, 可以有效地判断生物网络中的功能模块和关键物种(基因)。动态生物网络(dynamic biological network)是由一系列时间片段(或快照)组成, 每个时间片段表示在特定时间内所有变量的状态, 如动态流平衡分析(dynamic flux balance analysis, DFBA)。该方法是通过静态优化法(static optimization approach, SOA)等将FBA耦合到动力学模型, 并采用较小的时间步长制作一系列快照以准确地捕获生态系统的动态变化特征^[105,106]。藻际微生物的动态差异可以体现在多样性、物种丰度和功能等方面, 这些动态变量影响着赤潮的发生历程。藻菌动态网络分析框架主要分为三大层次(图5): 第一层是藻菌复杂网络建模, 该层次的数据来自动态复杂网络的快照; 第二层次是网络演化分析过程, 该阶段主要是对网络拓扑结构和数

据的分析处理; 第三层次是结果, 即通过第二层次的算法对网络的模块和边进行分析, 并归纳结果找出规律、相关性甚至因果关系, 对研究藻菌生态事件的发生、维持、演替过程提供指导。在研究动态网络的方法中, 最典型的是LV方程(generalized Lotka-Volterra, GLV), 它可以研究一个动态生态系统的稳定性和衡量某些微生物或者环境因子的重要性^[107]。此外, 还有Net-Shift方法, 该方法早期用于研究健康疾病状态之间的微生物关联网络的群落演替。移植于水体环境, 这些方法可用于量化赤潮不同发展阶段藻际微生物的群落演变, 从中筛选潜在的“赤潮推手”, 并将这些微生物描述为“驱动节点”^[108]。基于已有知识, 我们推测, 在藻际环境中, 通过动态网络分析有望解决3个关键问题: (1) 赤潮不同发展阶段或者藻类不同状态下微生物相互作用模式是否在整体上发生了重大改变; (2) 两个网络之间每个组成节点的关联是否有重大差异; (3) 这些节点是否代表了主要的网络组成, 它们在不同发展阶段或者藻细胞不同状态下网络中的重要性是否增加。

4 藻菌关系的网络延伸

4.1 藻菌互作关系知识网络

独立地分析单一组学并不能得出一个系统的整体

图5 (网络版彩色)藻菌动态网络分析框架

Figure 5 (Color online) A framework for dynamic network analysis of algae-bacteria

视图, 要获得全局性信息必须把各类组学数据有机融合并阐释其生物学功能. 为此, 更高水平的算法和数据处理能力得以开发, 知识网络也应运而生(图6). 这种网络展示的信息不是将某一生物层次的数据简单堆积, 而是由不同藻菌物种的基因组、转录组、蛋白质组、代谢组数据进行多层级结构的展示和知识梳理. 每一层级之间涉及多种因果、调控关系, 并且这些关系随着数据的上传、机器学习、训练而不断地更新和完善. 以机器学习为例, 它是一组从多维数据集中构建预测模型的数据分析技术, 基于预测任务来判断数据有无被标记(或显示特异性特征). 在这里, 微生物的大小、形状、分类、物质降解、有机质合成、光合作用、信号交流等都是分类的特征. 面对这些海量数据, 凭借鉴定算法很难进行提炼, 必须赋予机器学习的能力. 对于未标记的数据, 需要采用无监督的方法, 例如PCA或分层聚类; 而对于被标记的数据, 则可以应用监督方法(无监督方法生成带有标记的结果可以作为监督学习的输入数据), 将输入标签用于回归或数据分类. 经机器学习验证后可以生成新的数据并用于改进学习模型, 或衍生新的生物学假设^[109]. 在藻菌互作中, 知识网络所反映的核心特征是系统生物学中的多维数据整合,

包括个体水平、多物种水平以及跨界水平的整合. 以典型的藻菌共生体为例, 某些基因的存在可能对转录蛋白产生直接影响, 而蛋白联系着下游的代谢组信息, 代谢产物(如VB₁₂、跨界信号、抑菌素、杀藻物质等)又参与藻菌关系的反馈调解^[9]. 在这种环路背景下, 需要借助高度整合的知识网络进行解读, 才能挖掘出传统方法不能发现的互作机制.

4.2 生物互作图形数据库

基于藻菌数据库的更新和完善, 藻菌之间的生态网络也在数量和质量上得到提升. 图形数据不仅能以节点的形式表征生物(或者与环境因子)之间的关系, 还能以连接边属性的形式表示节点之间的关系类型和互作模式. 随着藻菌大数据时代的发展, 使得经典的关系数据库如Oracle、MySQL和MSSQL等在管理和存储高复杂、高依赖型的网络数据时面临着巨大挑战. 如传统的数据库无法对已经报道的生物互作关系进行高保真存储、精确查询和再度分析. 此外, 传统列表型数据库不便于比较不同生态系统、不同时间条件下的藻菌网络数据. 如果可以将生物互作信息以图形形式存储, 形成一种生物互作图形数据库(biological interaction graphic

图 6 (网络版彩色)藻菌互作关系知识网络

Figure 6 (Color online) Knowledge network database of algal-bacterial interactions

database, BIGDB), 将生物互作关系以网络图形形式储备,并赋予网络和节点不同属性,以标注不同外界条件下生物之间的互作模式,那么网络图形的拓扑属性和生物功能特性之间的关系将有一个更直观的展示。

图形数据库不仅仅是对表格数据的拓扑结构转化,还应可以进行搜索和计算分析,使得物种之间的关系及互作模式具有统一的存储、查询和分析标准。针对特定生境中生物互作关系的查询不再是仅仅对众多分散文献的逐个解读。Neo4j是一个有商业价值的开源图形数据库(<https://neo4j.com/>),它用高效的图数据结构代替传统的设计表,是一个高度可伸缩的图形数据库管理系统。Dai等人^[110]利用Neo4j整合了基因、蛋白质、化合物和小RNA之间的大量互作图形数据信息,并利用集成的图形搜索算法挖掘出了多种生物分子间的新关系。因此,对藻菌互作图形数据库的建立和完善,将会为藻菌跨界互作及生物地球化学循环中的机制研究提供重要的技术手段,为充分挖掘物种间新的级联关系模型提供技术支撑。

5 总结与展望

藻菌关系是海洋生态中的一对经典共生关系,目前已逐渐从表观的群落组成变化延伸到功能机制的解读。生态网络为藻菌关系的研究带来了新的契机,它能有效利用系统生物学思维从整体角度把握藻菌关系的全局性,不仅仅解释了藻菌彼此之间的相互作用、共生模式和互作机理,更为解读跨界交流、共适应特征甚至共进化机制提供了思路。然而,生态网络的研究也需要依赖一定的条件。首先,在藻菌大数据背景下,随着模型的不断完善和所要关注因子的逐渐增多,网络的构建和拓扑结构分析需要强大的计算能力,这对服务器的存储能力和运算能力提出了新要求。其次,网络整合的藻菌信息,在一定程度上为藻菌之间的互作机制提供了思路。但大多情况下,这种推断并不是一个坚实的结论,还需更多的野外实验或室内控制实验进行验证。基于此,未来的工作中我们还需要在如下方面进一步拓展和深入。

(i) 网络方法在藻菌关系研究中的应用需紧扣实践。网络方法可以帮助解释以下问题: (1) 藻菌共生体

的稳态维持及介导的物质循环偶联机制(如碳、氮、磷、硫等); (2) 赤潮事件中不同生物(细菌、古菌、真核藻类)基于时间延迟的网络时滞性特征; (3) 藻际微生物从个体水平到种群模式的代谢网络差异; (4) 藻际环境个体变异、关键物种变化对群落以及生态系统的影响程度,环境应激下的缓冲能力、恢复能力以及赤潮生消的维持成因; (5) 人类活动与全球变化(如海洋酸化)对藻类的影响和未来初级生产力的走向。这些方面都有不同程度的进展,也表现出非常大的潜力。但如何进一步将网络分析应用于自然生态系统,构建智能化模型并揭示其中的分子机制是我们需要努力探索的科学问题。

(ii) 新技术的应用。建立多组学、多数据之间的关联,即藻菌知识网络,将快速提升大数据背景下互作机制的效率。此外,积极探索人工智能AI技术、深度学习技术和神经网络模型工具,如卷积神经网络(convolutional neural networks, CNN)。这些工具可处理两个或更高维度的数据集,将数据信息转化为生物学信息并解析其背后的生态意义。

(iii) 模型与实测数据的结合。基于网络推断和预测的分析结果还需要更多的原位数据或者实验室结果来验证。高斯共培养试验是验证物种间相互作用关系的标准,共培养实验在受控的环境中创造了一个人工群落,从而为检验有关群落稳定性和动态性的生态学概念提供了理想的条件,但这些概念在宏观生态系统中难以测量。藻际环境中细菌群落与周边环境差异很大,发掘藻际环境新型微生物类群和未可培养细菌将有助于揭示藻菌之间复杂的作用和新颖的潜藏关系。单细胞测序技术将为藻菌群落在分类学和功能层面提供更高的分辨率,致力于藻际细菌单细胞测序和实验室内藻际细菌可培养化的研究将是藻菌关系研究的一个重要方向。此外,合成生物学手段和拉曼光谱技术也为研究藻菌关系提供了可选的途径。它可以根据实验需求设置不同的微生物群落,以研究物种间的相互关系和代谢产物的交互关系^[111,112]。类似于人类社会,如果我们能够基于网络挖掘藻际环境中的生态规律,厘清其中的物质流-信息流和能量流,将有望开启藻菌关系从生态学向社会学的跨越^[113]。

参考文献

- 1 Bell W, Mitchell R. Chemotactic and growth responses of marine bacteria to algal extracellular products. *Biol Bull*, 1972, 143: 265-277

- 2 Uribe P, Espejo R T. Effect of associated bacteria on the growth and toxicity of *Alexandrium catenella*. *Appl Environ Microbiol*, 2003, 69: 659–662
- 3 Arrigo K R. Marine microorganisms and global nutrient cycles. *Nature*, 2005, 437: 349–355
- 4 Cottrell M T, Kirchman D L. Natural assemblages of marine proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. *Appl Environ Microbiol*, 2000, 66: 1692–1697
- 5 Blackburn N, Fenchel T, Mitchell J. Microscale nutrient patches in planktonic habitats shown by chemotactic bacteria. *Science*, 1998, 282: 2254–2256
- 6 Fuentes J L, Garbayo I, Cuaresma M, et al. Impact of microalgae-bacteria interactions on the production of algal biomass and associated compounds. *Mar Drugs*, 2016, 14: 1–16
- 7 Seyedsayamdost M R, Case R J, Kolter R, et al. The jekyll-and-hyde chemistry of phaeobacter gallaeciensis. *Nat Chem*, 2011, 3: 331–335
- 8 Cooper M B, Smith A G. Exploring mutualistic interactions between microalgae and bacteria in the omics age. *Curr Opin Plant Biol*, 2015, 26: 147–153
- 9 Ramanan R, Kim B H, Cho D H, et al. Algae-bacteria interactions: Evolution, ecology and emerging applications. *Biotechnol Adv*, 2016, 34: 14–29
- 10 Croft M T, Lawrence A D, Raux-Deery E, et al. Algae acquire vitamin B₁₂ through a symbiotic relationship with bacteria. *Nature*, 2005, 438: 90–93
- 11 Helliwell K E, Wheeler G L, Leptos K C, et al. Insights into the evolution of vitamin B₁₂ auxotrophy from sequenced algal genomes. *Mol Biol Evol*, 2011, 28: 2921–2933
- 12 Kazamia E, Czesnick H, Nguyen T T V, et al. Mutualistic interactions between vitamin B₁₂-dependent algae and heterotrophic bacteria exhibit regulation. *Environ Microbiol*, 2012, 14: 1466–1476
- 13 Thingstad T R, Skjoldal E R, Böhne R A. Phosphorus cycling and algal-bacterial competition in sandsfjord, western norway. *Mar Ecol Prog Ser*, 1993, 99: 239–259
- 14 Kim M J, Jeong S Y, Lee S J. Isolation, identification, and algicidal activity of marine bacteria against cochlodinium polykrikoides. *J Appl Phycol*, 2008, 20: 1069–1078
- 15 Lee Y K, Ahn C Y, Kim H S, et al. Cyanobactericidal effect of *Rhodococcus* sp. isolated from eutrophic lake on *Microcystis* sp. *Biotechnol Lett*, 2010, 32: 1673–1678
- 16 Wang X, Li Z, Su J, et al. Lysis of a red-tide causing alga, *Alexandrium tamarense*, caused by bacteria from its phycosphere. *Biol Control*, 2010, 52: 123–130
- 17 Bruckner C G, Bahulikar R, Rahalkar M, et al. Bacteria associated with benthic diatoms from lake constance: Phylogeny and influences on diatom growth and secretion of extracellular polymeric substances. *Appl Environ Microbiol*, 2008, 74: 7740–7749
- 18 Carrillo P, Delgado-Molina J A, Bulles J F, et al. Complex interactions in microbial food webs: Stoichiometric and functional approaches. *Limnetica*, 2006, 25: 189–204
- 19 Mayali X, Doucette G J. Microbial community interactions and population dynamics of an algicidal bacterium active against *Karenia brevis* (dinophyceae). *Harmful Algae*, 2002, 1: 277–293
- 20 Sher D, Thompson J W, Kashtan N, et al. Response of *Prochlorococcus* ecotypes to co-culture with diverse marine bacteria. *ISME J*, 2011, 5: 1125–1132
- 21 Gurung T B, Urabe J, Nakanishi M. Regulation of the relationship between phytoplankton *Scenedesmus acutus* and heterotrophic bacteria by the balance of light and nutrients. *Aquat Microb Ecol*, 1999, 17: 27–35
- 22 Leung T L F, Poulin R. Parasitism, commensalism, and mutualism: Exploring the many shades of symbioses. *Vie Milieu*, 2008, 58: 107–115
- 23 Liu H, Zhou Y, Xiao W, et al. Shifting nutrient-mediated interactions between algae and bacteria in a microcosm: Evidence from alkaline phosphatase assay. *Microbiol Res*, 2012, 167: 292–298
- 24 Kouzuma A, Watanabe K. Exploring the potential of algae/bacteria interactions. *Curr Opin Biotechnol*, 2015, 33: 125–129
- 25 Krediet C J, Ritchie K B, Paul V J, et al. Coral-associated micro-organisms and their roles in promoting coral health and thwarting diseases. *Proc Biol Sci*, 2013, 280: 20122328
- 26 Amin S A, Hmelo L R, van Tol H M, et al. Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature*, 2015, 522: 98–101
- 27 Teplitski M, Rajamani S. Signal and nutrient exchange in the interactions between soil algae and bacteria. In: Witzany G, ed. *Biocommunication in Soil Microorganisms*. Berlin: Springer, 2011. 413–426
- 28 Bolch C J S, Subramanian T A, Green D H. The toxic dinoflagellate *gymnodinium catenatum* (dinophyceae) requires marine bacteria for growth. *J Phycol*, 2011, 47: 1009–1022
- 29 Takemura A F, Chien D M, Polz M F. Associations and dynamics of vibriaceae in the environment, from the genus to the population level.

- [Front Microbiol](#), 2014, 5: 38
- 30 Jha B, Kavita K, Westphal J, et al. Quorum sensing inhibition by *Asparagopsis taxiformis*, a marine macro alga: Separation of the compound that interrupts bacterial communication. [Mar Drugs](#), 2013, 11: 253–265
 - 31 Paul C, Pohnert G. Production and role of volatile halogenated compounds from marine algae. [Nat Prod Rep](#), 2011, 28: 186–195
 - 32 Twigg M S, Tait K, Williams P, et al. Interference with the germination and growth of *Ulva* zoospores by quorum-sensing molecules from *Ulva* - associated epiphytic bacteria. [Environ Microbiol](#), 2014, 16: 445–453
 - 33 Allen A E, Dupont C L, Oborník M, et al. Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. [Nature](#), 2011, 473: 203–207
 - 34 Shen J B, Bai Y, Wei Z, et al. Rhizobiont: An interdisciplinary innovation and perspective for harmonizing resources, environment and food security (in Chinese). [Acta Pedol Sin](#), 2021, 58: 805–813 [申建波, 白洋, 韦中, 等. 根际生命共同体: 协调资源、环境和粮食安全的学术思路与交叉创新. [土壤学报](#), 2021, 58: 805–813]
 - 35 Yang C, Wang Q, Simon P N, et al. Distinct network interactions in particle-associated and free-living bacterial communities during a microcystis aeruginosa bloom in a plateau lake. [Front Microbiol](#), 2017, 8: 1202
 - 36 Zhao D, Shen F, Zeng J, et al. Network analysis reveals seasonal variation of co-occurrence correlations between cyanobacteria and other bacterioplankton. [Sci Total Environ](#), 2016, 573: 817–825
 - 37 Gantz J, Reinsel D. The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east-western Europe. IDC, 2012, 2007: 1–16
 - 38 Green J L, Bohannan B J M, Whitaker R J. Microbial biogeography: From taxonomy to traits. [Science](#), 2008, 320: 1039–1043
 - 39 Handelsman J, Rondon M R, Brady S F, et al. Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. [Chem Biol](#), 1998, 5: 245–249
 - 40 Handelsman J. Metagenomics: Application of genomics to uncultured microorganisms. [Microbiol Mol Biol Rev](#), 2004, 68: 669–685
 - 41 Poretsky R S, Bano N, Buchan A, et al. Analysis of microbial gene transcripts in environmental samples. [Appl Environ Microbiol](#), 2005, 71: 4121–4126
 - 42 Wilmes P, Bond P L. The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. [Environ Microbiol](#), 2004, 6: 911–920
 - 43 Pande S, Kost C. Bacterial unculturability and the formation of intercellular metabolic networks. [Trends Microbiol](#), 2017, 25: 349–361
 - 44 Kitano H. Systems biology: A brief overview. [Science](#), 2002, 295: 1662–1664
 - 45 Newman M E J. The structure and function of complex networks. [SIAM Rev](#), 2003, 45: 167–256
 - 46 Opsahl T, Agneessens F, Skvoretz J. Node centrality in weighted networks: Generalizing degree and shortest paths. [Soc Networks](#), 2010, 32: 245–251
 - 47 Redner S. How popular is your paper? An empirical study of the citation distribution. [Eur Phys J B](#), 1998, 4: 131–134
 - 48 Alves R, Chaleil R A G, Sternberg M J E. Evolution of enzymes in metabolism: A network perspective. [J Mol Biol](#), 2002, 320: 751–770
 - 49 Jiang X P, Hu X H. A brief introduction to marine eco-systems biology (in Chinese). [Math Model Appl](#), 2013, 2: 12–21 [蒋兴鹏, 胡小华. 海洋生态系统生物学简介. [数学建模及其应用](#), 2013, 2: 12–21]
 - 50 Shannon P, Markiel A, Ozier O, et al. Cytoscape: A software environment for integrated models of biomolecular interaction networks. [Genome Res](#), 2003, 13: 2498–2504
 - 51 Bastian M, Heymann S, Jacomy M. Gephi: An open source software for exploring and manipulating networks. In: [Proceedings of the Third International AAAI Conference on Weblogs and Social Media](#), 2009, 3: 361–362
 - 52 Csardi G, Nepusz T. The igraph software package for complex network research. [Int J Complex Syst](#), 2006, 1695: 1–9
 - 53 Faust K, Lima-Mendez G, Lerat J S, et al. Cross-biome comparison of microbial association networks. [Front Microbiol](#), 2015, 6: 1200
 - 54 Kurtz Z D, Müller C L, Miraldi E R, et al. Sparse and compositionally robust inference of microbial ecological networks. [PLoS Comput Biol](#), 2015, 11: e1004226
 - 55 Friedman J, Alm E J. Inferring correlation networks from genomic survey data. [PLoS Comput Biol](#), 2012, 8: e1002687
 - 56 Deng Y, Jiang Y H, Yang Y, et al. Molecular ecological network analyses. [BMC Bioinf](#), 2012, 13: 1–20
 - 57 Ihaka R, Gentleman R. R: A language for data analysis and graphics. [J Comp Graph Stat](#), 1996, 5: 299–314
 - 58 Xia L C, Steele J A, Cram J A, et al. Extended local similarity analysis (eLSA) of microbial community and other time series data with replicates. [BMC Syst Biol](#), 2011, 5: 1–12
 - 59 David L A, Materna A C, Friedman J, et al. Host lifestyle affects human microbiota on daily timescales. [Genome Biol](#), 2014, 15: 1–15
 - 60 McGeachie M J, Chang H H, Weiss S T. Cgbytesnets: Conditional gaussian bayesian network learning and inference with mixed discrete and continuous data. [PLoS Comput Biol](#), 2014, 10: e1003676
 - 61 Bucci V, Tzen B, Li N, et al. Mdsine: Microbial dynamical systems inference engine for microbiome time-series analyses. [Genome Biol](#), 2016,

- 17: 1–7
- 62 Zhou J, Deng Y, Luo F, et al. Functional molecular ecological networks. *MBio*, 2010, 1: e00169
- 63 Barberán A, Bates S T, Casamayor E O, et al. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J*, 2012, 6: 343–351
- 64 Pérez-Valera E, Goberna M, Faust K, et al. Fire modifies the phylogenetic structure of soil bacterial co-occurrence networks. *Environ Microbiol*, 2017, 19: 317–327
- 65 Shen F, Huang R, Zeng J, et al. Progress of the molecular ecology network analysis (in Chinese). *Environ Sci Technol*, 2016, 39: 94–98 [沈烽, 黄睿, 曾巾, 等. 分子生态网络分析研究进展. *环境科学与技术*, 2016, 39: 94–98]
- 66 Bader G D, Hogue C W. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinf*, 2003, 4: 1–27
- 67 Freeman L C. Centrality in social networks conceptual clarification. *Soc Networks*, 1978, 1: 215–239
- 68 Guimerà R, Nunes Amaral L A. Functional cartography of complex metabolic networks. *Nature*, 2005, 433: 895–900
- 69 Olesen J M, Bascompte J, Dupont Y L, et al. The modularity of pollination networks. *Proc Natl Acad Sci USA*, 2007, 104: 19891–19896
- 70 Luo F, Zhong J, Yang Y, et al. Application of random matrix theory to biological networks. *Phys Lett A*, 2006, 357: 420–423
- 71 Luo F, Yang Y, Zhong J, et al. Constructing gene co-expression networks and predicting functions of unknown genes by random matrix theory. *BMC Bioinf*, 2007, 8: 1–17
- 72 Zhou J, Deng Y, Luo F, et al. Phylogenetic molecular ecological network of soil microbial communities in response to elevated CO₂. *MBio*, 2011, 2: e00122-11
- 73 Yuan M M, Guo X, Wu L, et al. Climate warming enhances microbial network complexity and stability. *Nat Clim Chang*, 2021, 11: 343–348
- 74 Ruan Q, Dutta D, Schwalbach M S, et al. Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. *Bioinformatics*, 2006, 22: 2532–2538
- 75 Durno W E, Hanson N W, Konwar K M, et al. Expanding the boundaries of local similarity analysis. *BMC Genomics*, 2013, 14: 1–14
- 76 Steele J A, Countway P D, Xia L, et al. Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *ISME J*, 2011, 5: 1414–1425
- 77 Xia L C, Ai D, Cram J, et al. Efficient statistical significance approximation for local similarity analysis of high-throughput time series data. *Bioinformatics*, 2012, 29: 230–237
- 78 Paver S F, Hayek K R, Gano K A, et al. Interactions between specific phytoplankton and bacteria affect lake bacterial community succession. *Environ Microbiol*, 2013, 15: 2489–2504
- 79 Liu L, Yang J, Lv H, et al. Synchronous dynamics and correlations between bacteria and phytoplankton in a subtropical drinking water reservoir. *FEMS Microbiol Ecol*, 2014, 90: 126–138
- 80 Needham D M, Fuhrman J A. Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nat Microbiol*, 2016, 1: 1–7
- 81 Zhou J, Lao Y M, Song J T, et al. Temporal heterogeneity of microbial communities and metabolic activities during a natural algal bloom. *Water Res*, 2020, 183: 116020
- 82 Chun S J, Cui Y, Lee J J, et al. Network analysis reveals succession of microcystis genotypes accompanying distinctive microbial modules with recurrent patterns. *Water Res*, 2020, 170: 115326
- 83 McClure R S, Lee J Y, Chowdhury T R, et al. Integrated network modeling approach defines key metabolic responses of soil microbiomes to perturbations. *Sci Rep*, 2020, 10: 1–9
- 84 Seyedsayamdost M R, Wang R, Kolter R, et al. Hybrid biosynthesis of roseobacticides from algal and bacterial precursor molecules. *J Am Chem Soc*, 2014, 136: 15150–15153
- 85 Fuhrman J A, Cram J A, Needham D M. Marine microbial community dynamics and their ecological interpretation. *Nat Rev Microbiol*, 2015, 13: 133–146
- 86 Bilke S, Peterson C. Topological properties of citation and metabolic networks. *Phys Rev E*, 2001, 64: 036106
- 87 Fiehn O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comp Funct Genomics*, 2001, 2: 155–168
- 88 Jeong H, Tombor B, Albert R, et al. The large-scale organization of metabolic networks. *Nature*, 2000, 407: 651–654
- 89 McGenity T, Folwell B, McKew B, et al. Marine crude-oil biodegradation: A central role for interspecies interactions. *Aquat Biosyst*, 2012, 8: 1–19
- 90 Zengler K, Zaramela L S. The social network of microorganisms—How auxotrophies shape complex communities. *Nat Rev Microbiol*, 2018, 16: 383–390
- 91 Yu K, Yi S, Li B, et al. An integrated meta-omics approach reveals substrates involved in synergistic interactions in a bisphenol a (bpa)-degrading microbial community. *Microbiome*, 2019, 7: 1–13

- 92 Kundu P, Manna B, Majumder S, et al. Species-wide metabolic interaction network for understanding natural lignocellulose digestion in termite gut microbiota. *Sci Rep*, 2019, 9: 1–13
- 93 Bauer E, Thiele I. From network analysis to functional metabolic modeling of the human gut microbiota. *mSystems*, 2018, 3: 209–217
- 94 Mendoza S N, Olivier B G, Molenaar D, et al. A systematic assessment of current genome-scale metabolic reconstruction tools. *Genome Biol*, 2019, 20: 1–20
- 95 Magnúsdóttir S, Heinken A, Kutt L, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat Biotechnol*, 2017, 35: 81–89
- 96 King Z A, Lu J, Dräger A, et al. Bigg models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res*, 2015, 44: 515–522
- 97 Zelezniak A, Andrejev S, Ponomarova O, et al. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc Natl Acad Sci USA*, 2015, 112: 6449–6454
- 98 Heinken A, Sahoo S, Fleming R M T, et al. Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes*, 2013, 4: 28–40
- 99 Wagner A, Fell D A. The small world inside large metabolic networks. *Proc R Soc Lond B*, 2001, 268: 1803–1810
- 100 Notebaart R A, van Enckevort F H, Francke C, et al. Accelerating the reconstruction of genome-scale metabolic networks. *BMC Bioinf*, 2006, 7: 1–10
- 101 Ravasz E, Somera A L, Mongru D A, et al. Hierarchical organization of modularity in metabolic networks. *Science*, 2002, 297: 1551–1555
- 102 Dittami S M, Eveillard D, Tonon T. A metabolic approach to study algal-bacterial interactions in changing environments. *Mol Ecol*, 2014, 23: 1656–1660
- 103 Lima-Mendez G, Faust K, Henry N, et al. Determinants of community structure in the global plankton interactome. *Science*, 2015, 348: 1262073
- 104 Astudillo-García C, Bell J J, Webster N S, et al. Evaluating the core microbiota in complex communities: A systematic investigation. *Environ Microbiol*, 2017, 19: 1450–1462
- 105 Höffner K, Harwood S M, Barton P I. A reliable simulator for dynamic flux balance analysis. *Biotechnol Bioeng*, 2013, 110: 792–802
- 106 Gomez J A, Höffner K, Barton P I. Dfbalab: A fast and reliable matlab code for dynamic flux balance analysis. *BMC Bioinf*, 2014, 15: 1–10
- 107 Hoffmann K H, Rodriguez-Brito B, Breitbart M, et al. Power law rank abundance models for marine phage communities. *FEMS Microbiol Lett*, 2007, 273: 224–228
- 108 Bhusan K, Chandrakar P, Sadhu S, et al. ‘Netshift’: A methodology for understanding ‘driver microbes’ from healthy and disease microbiome datasets. *ISME J*, 2018, 13: 442–454
- 109 Camacho D M, Collins K M, Powers R K, et al. Next-generation machine learning for biological networks. *Cell*, 2018, 173: 1581–1592
- 110 Dai X, Li J, Liu T, et al. Hrgn: A graph search-empowered integrative database of arabidopsis signaling transduction, metabolism and gene regulation networks. *Plant Cell Physiol*, 2016, 57: e12
- 111 Deng T, Guan X T, Wu B, et al. Applying mathematical models in the construction of synthetic microbial communities (in Chinese). *Chin J Appl Environ Biol*, 2020, 26: 73–83 [邓婷, 关晓彤, 吴波, 等. 数学模型在合成微生物群落构建中的应用. *应用与环境生物学报*, 2020, 26: 73–83]
- 112 Hong J K, Kim S B, Lyou E S, et al. Microbial phenomics linking the phenotype to function: The potential of Raman spectroscopy. *J Microbiol*, 2021, 59: 249–258
- 113 Zhou J, Cai Z. Microbial social interactions in biofilm. In: Pallaval V, ed. *Implication of Quorum Sensing System in Biofilm Formation and Virulence*. Singapore: Springer, 2018. 29–46

Summary for “藻菌关系的生态网络研究方法: 回顾与展望”

The ecological network approach to algal-bacterial relationships: Review and prospects

Jianming Zhu^{1,2}, Jin Zhou^{1*}, Hui Wang^{3*}, Guofu Chen² & Zhonghua Cai¹¹ Institute for Ocean Engineering, Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China;² School of Marine Science and Technology, Harbin Institute of Technology (Weihai), Weihai 264209, China;³ College of Science, Shantou University, Shantou 515063, China* Corresponding authors, E-mail: zhou.jin@sz.tsinghua.edu.cn; wanghui@stu.edu.cn

Algae and bacteria numerically dominate oceanic and freshwater planktonic communities. They play a central biogeochemical role in the marine ecosystem, including producing biomass to support aquatic food webs, cycling nutrients and trace elements, and decomposing organic matter. Algae and bacteria synergistically affect each other's physiology and metabolism and have coexisted ever since the early stages of evolution that revolutionized many aspects of life on Earth. The relationship between algae and bacteria influences ecosystems and represents all conceivable modes of interaction from mutualism to parasitism. The interactions between these two groups of plankton, and the influence of their interaction on each other and on a global scale have, therefore, become the focus of recent research interest. Previous studies have summarized a variety of algal-bacterial interactions and aimed to reveal the essence of those interactions. However, it is clear that their interactions are diverse and not static, and can be initiated and broken in response to environmental and developmental cues. Traditional research methods have been unable to reveal the complexity behind algal-bacterial symbiosis. As microbial omics technology has developed and the influence of global change on marine ecosystems has become apparent, the research perspective on algal-bacterial relationships has gradually risen to the level of systemic ecology. The rapid-blooming of high-throughput sequencing technology and the ability to analyze big data offer new opportunities for a thorough study of algal-bacterial interaction. Ecological networks are prominent among these new methods: They can integrate multiple types of information and might represent systems-level behavior. Network theory methods are powerful approaches for quantifying relationships between biological components and their relevance to phenotypes, environmental conditions or other external variables within a system. In this paper, we review the research advances in the algal-bacterial ecological network based on big data. We begin by introducing the network parameters and attributes commonly used in ecological network analysis and review the existing network modeling methods in microbial studies and algorithm features and application scenarios. By comparing some common algorithms or software, researchers can choose appropriate network construction methods when confronted with different data characteristics. Next, the ecological network analysis used in the study of algal-bacterial relationships is presented at multiple layers from taxonomy to function (e.g., molecular and metabolic networks). The problems and experiences encountered during research into algal-bacterial relationships are then discussed. For example, when focusing on the interaction between host and microorganism, the time/space shift correlations among the variables will lead to wrong conclusions, including false-positive or false-negative results. We also attempt to introduce metabolic and dynamic biological networks into the study of the algal-bacterial relationship to reveal further the mechanism of their interaction from the perspectives of function and causality. Finally, our novel proposal to establish a knowledge network database of algal-bacterial interactions and a graphic database of photosphere biological profiles is discussed. We highlight the need for microbial ecological network inference and suggest strategies to infer networks more reliably. Our study aimed to explore the holistic information behind algal-bacterial relationships and provide new cues for further revealing the molecular mechanisms of algal-bacterial interactions.

algal-bacterial interactions, microbial omics-tools, big data, ecological network, molecular mechanismsdoi: [10.1360/TB-2021-0314](https://doi.org/10.1360/TB-2021-0314)