基于玛氏骨条藻(*Skeletonema marinoi*)转录组的碳固定代谢途径分析

刘乾^{①②③}, 米铁柱^{②③④*}, 甄毓^{②③④}, 王华龙^{①②③}, 于志刚^{③⑤}

① 中国海洋大学海洋生命学院, 青岛 266003;

② 中国海洋大学海洋环境与生态教育部重点实验室, 青岛 266100;

③ 青岛海洋科学与技术国家实验室海洋生态与环境科学功能实验室, 青岛 266071;

④ 中国海洋大学环境科学与工程学院, 青岛 266100;

⑤ 中国海洋大学海洋化学理论与工程技术教育部重点实验室, 青岛 266100

* 联系人, E-mail: mitiezhu@ouc.edu.cn

2016-01-13 收稿, 2016-04-01 修回, 2016-04-05 接受, 2016-07-11 网络版发表

海洋公益性行业科研专项(201205031)、国家自然科学基金(41521064)和山东省自然科学基金(ZR2014DM007)资助

摘要 以玛氏骨条藻(*Skeletonema marinoi*)转录组信息为基础, 分析了其碳固定代谢途径, 共发现18种酶对应的34个编码基因, 构建了玛氏骨条藻进行碳固定代谢途径的通路图. 这些编码基因的序列比对结果表明, 其与假微型海链藻(*Thalassiosira pseudonana*)的同源基因具有较高的一致性. 对这些样品进行数字基因表达谱的差异基因分析, 获得了不同生长时期碳固定代谢途径酶编码基因的差异表达情况. 通过分析发现, C₃和C₄代谢途径中存在表达差异的基因分别有7和3个, 其中果糖-1,6-二磷酸酶和丙酮酸磷酸双激酶的编码基因表达在指数生长期之后出现显著上调. 这有助于对玛氏骨条藻碳固定代谢途径中关键编码基因调控过程的解析, 为进一步研究硅藻的固碳机制奠定了基础, 也为深入了解碳的生物地球化循环提供了新的研究方向.

关键词 玛氏骨条藻, 碳固定代谢途径, 转录组

光合作用可以将光能转化为分子形式的化学能, 并储存于植物体内, 之后再将能量传递给各级消费者, 这一过程在全球生态系统中起到了至关重要的作用. 根据植物碳同化途径所产生最初产物的不同, 可以将植物分为3类: C₃植物、C₄植物和景天酸代谢(crassulacean acid metabolism, CAM)植物. 在这3类植物中, C₃植物最为原始且常见, 其固定CO₂后的最初产物为3-磷酸甘油酸(3-phosphoglycerate, 3-PGA, 3C化合物), 这一过程由核酮糖-1,5-二磷酸羧化酶/加氧酶(ribulose-1,5-bisphosphate carboxylase/oxygenase, Rubisco)参与完成, 而拥有C₄和CAM途径的植物则由C₃植物进化而来^[1,2]. C₄途径过程极为复杂, 虽然消

耗能量较多, 但仍具有重要意义. 在低温、湿润和较高的大气CO₂浓度下适合C₃植物生长; C₄植物则由于自身较高的水分利用率和光合作用效率, 拥有较强的耐受能力, 因而在高温、干燥和较低CO₂浓度的环境下更具有生长优势^[3-5].

Rubisco作为固定CO₂的关键酶及限速酶备受关注^[6], 它只能利用CO₂而不能利用其他形式的无机碳^[7], 其对CO₂的平衡浓度(K_{mCO_2})为40~60 $\mu\text{mol L}^{-1}$ ^[8]. 由于现代海洋环境中无机碳浓度约为2 mmol L^{-1} , 其中的大部分还是以HCO₃⁻离子的形式存在, 同时光合作用的产物O₂也会同CO₂一起竞争Rubisco上的反应位点, 进行加氧反应(光呼吸作用), 因此, 水生植物

引用格式: 刘乾, 米铁柱, 甄毓, 等. 基于玛氏骨条藻(*Skeletonema marinoi*)转录组的碳固定代谢途径分析. 科学通报, 2016, 61: 2483-2493

Liu Q, Mi T Z, Zhen Y, et al. Description of carbon fixation pathway based on *Skeletonema marinoi* transcriptome (in Chinese). Chin Sci Bull, 2016, 61: 2483-2493, doi: 10.1360/N972016-00046

通常依赖CO₂浓缩机制(CO₂ concentrating mechanisms, CCM)来增加Rubisco周围的CO₂浓度,以此保证固碳反应的顺利进行,进而使Rubisco固定CO₂的能力显著提高^[9,10]. CCM可以分为生物物理机制以及生物化学机制. 生物物理机制主要通过碳酸酐酶(carbonic anhydrase, CA)^[11]在Rubisco周围催化HCO₃⁻脱水放出CO₂来实现,生物化学机制则是水生植物依赖类似于陆地C₄植物的浓缩机制^[12],通过特异性结构(Kranz结构^[13])以及磷酸烯醇式丙酮酸羧化酶(phosphoenolpyruvate carboxylase, PEPC)和磷酸烯醇式丙酮酸羧激酶(phosphoenolpyruvate carboxykinase, PEPCK)等一系列酶的作用生成4C酸,来完成对CO₂的浓缩. 而作为单细胞的硅藻,其虽然并不具备陆地C₄植物多细胞的复杂结构,但仍然可以通过细胞内的不同区域、各类膜结构以及C₄酶的共同作用完成浓缩反应,因此这种与C₄代谢途径十分相似的单细胞CCM机制也是研究者普遍关注的问题^[14]. 近年来,Reinfelder等人^[14,15]运用同位素标记方法和实时荧光定量PCR方法,证明在单细胞硅藻威氏海链藻(*Thalassiosira weissflogii*)中含有C₄途径的中间代谢产物及PEPC, PEPCK. 研究还发现,在限制CO₂浓度或抑制CA活性(对CCM的生物物理机制进行抑制)时,PEPC和PEPCK的转录本也出现了表达量增加的情况^[16,17],表明C₄活动在该条件下有所加强. 另外,通过基因组测序技术分析表明,在假微型海链藻(*T. pseudonana*)^[18]和三角褐指藻(*Phaeodactylum tri-cornutum*)^[19]中,同样存在C₄通路关键酶的编码基因.

海洋的初级生产力大约同陆地相当,其主要来源是藻类植物的光合作用^[20,21]. 藻类的初级生产除了维持海洋中的食物网和生态系统的正常运转外^[22,23],还在将CO₂从水体表面向深海迁移过程中发挥了重要作用^[24]. 硅藻是海洋初级生产的最大贡献者,通过光合作用,这一类群每年大约可以固定10¹⁶ g CO₂,相当于全球总固定量的20%^[25]. 因此,研究硅藻的碳固定代谢途径对于了解其在海洋生态系统中的重要作用,深入理解碳的生物地球化循环具有重要意义. 近年来,已有针对假微型海链藻和三角褐指藻^[18,19]的有关研究,而骨条藻(*Skeletonema* spp.)作为一类在全球近岸海域广泛分布的广温广盐型微型浮游硅藻,对其碳固定代谢途径的研究尚未见报道. 本研究以我国近岸海域优势种玛氏骨条藻(*Skeletonema marinoi*)为研究对象,首先确定了其碳固定代谢途径

相关酶的编码基因,并在此基础上研究其固碳作用的运行机制,从而阐明该藻在不同生长周期碳固定代谢途径对外界环境的分子响应.

1 材料与方法

(i) 玛氏骨条藻的培养. 自青岛近海海域分离得到玛氏骨条藻(*S. marinoi*),保存于中国海洋大学海洋环境与生态教育部重点实验室藻种室,在实验室条件下用常规f/2培养液对藻种进行培养,光照/黑暗周期设定为12 h/12 h,光照强度为80~100 μmol m⁻² s⁻¹,培养温度为20±1 °C. 实验所用海水取自青岛近海,盐度为33, pH 7.8~8.0.

(ii) 测序样品的准备. 将藻细胞按照生长时期分组的方式进行培养和收集:藻种接种后的指数生长期、稳定期和衰亡期分别取样. 每组设置3个平行样为生物学重复,离心收集藻细胞后,于液氮中保存. 样品由北京诺禾致源生物信息科技有限公司进行总RNA的提取及Illumina HiSeq2000(Illumina公司,美国)双端测序.

(iii) 数据预处理及从头测序(*de novo*)拼接. 高通量测序平台Illumina HiSeqTM2000所得到的原始数据文件经CASAVA碱基识别(base calling)分析转化为原始测序序列(sequenced reads),将其称之为raw data或raw reads,结果以FASTQ文件格式存储(其中包含测序序列的序列信息及其对应的测序质量信息)并对其进行质量评估,这主要包括测序错误率分布检查和A/T/G/C含量分布检查两部分. 随后,为了保证信息分析质量,必须对raw reads过滤,得到有效序列(clean reads). 数据过滤的过程由3个步骤组成:去除带接头的序列;去除无法确定碱基信息的比例大于10%的序列;去除低质量序列. 以Trinity^[26] (版本:v2012-10-05,参数设置: min_kmer_cov为2,其他参数为默认参数)拼接为一个转录组,以此作为后续分析的参考序列. 拼接得到的转录本序列信息以FASTA格式储存,文件中“>”后紧跟转录本的id号,每个转录本的id号构成都为comp seq,其中comp表示拼接过程形成的组成元件(component),可以看作为广泛意义上的基因,seq代表转录本编号. len=后面为转录本的长度,即该转录本的碱基数, path为拼接过程中经历的路径,其后为该转录的碱基序列.

(iv) 功能注释及代谢途径分析. 使用NCBI blast 2.2.28^[27]的算法将拼接成的转录组与NR, NT,

SwissProt, KOG数据库进行比对, NR, NT和SwissProt数据库比对的evalue设置阈值为 10^{-5} , KOG比对evalue阈值为 10^{-3} , 对转录本的功能预测并进行分类; 使用HMMER 3.0 package对PFAM蛋白结构域预测, 使用软件Blast2GO v2.5^[28]和自写脚本进行GO功能注释和KEGG(kyoto encyclopedia of genes and genomes)相关注释, 得到基于NR和PFAM两部分的蛋白注释结果; 使用KAAS服务器进行KEGG代谢途径分析^[29].

(V) 基因差异表达分析及差异基因KEGG富集分析. 基因差异表达的输入数据为基因表达水平分析中得到的read count数据. 采用DESeq^[30]以区分生物学重复之间以及不同生长时期分组之间的差异并对其进行筛选, 筛选阈值为 $\text{padj} < 0.05$.

由于在生物体内, 不同基因之间可以通过相互协调、相互作用共同行使其生物学功能, 因此通过代谢途径(pathway)显著性富集能确定差异表达基因参与的最主要生化代谢途径和信号转导途径. KEGG是有关pathway的主要公共数据库. Pathway显著性富集分析以KEGG pathway为单位, 应用超几何检验, 找出差异基因相对于所有注释的基因显著富集的pathway. 该分析的计算公式:

$$p = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}.$$

在这里 N 为所有基因中具有pathway注释的基因数目; n 为 N 中差异表达基因的数目; M 为所有基因中注释为某特定pathway的基因数目; m 为注释为某特定pathway的差异表达基因数目. FDR (false discovery rate) ≤ 0.05 的pathway定义为在差异表达基因中显著富集的pathway, 使用KOBAS(2.0), 设置参数——FDR为BH(即使用BH校正)进行pathway富集分析, 找出差异基因相对于所有注释的基因显著富集的代谢途径.

2 结果与讨论

2.1 转录组测序信息及拼接

对Illumina Hiseq2000双端测序获得的原始数据进行过滤, 去除低质量数据后共获得58592306 bp用于拼接的高质量有效序列, 占原始序列的96.15%,

将所有有效序列利用*de novo*进行拼接共获得39098个转录本, 其中有20319个非重复序列基因(unigene), 长度范围为200~20000 bp.

2.2 转录本的功能注释及其代谢途径分析

在20319个unigene中, 通过对前述7个数据库比对, 共有14693个unigene得到了注释, 占71.3%. 在NR, PFAM和GO数据库注释的成功率较高, 分别为65.2%, 51.44%和53.28%. 而在NT, KO, SwissProt和KOG数据库的注释成功率较低, 均未超过50%. 在对所有的转录本进行了GO注释以后, 将注释成功的转录本划分为46个类别, 分别属于生物过程、细胞组分和分子功能3大类. 同样与KOG数据库进行比对后, 可将转录本按照功能分成26个子家族. 进行KO注释后, 根据它们参与的KEGG代谢通路进行分类, 得到了31个类群, 其中有3816个转录本可以匹配到EC编号, 并注释为225条不同的代谢途径, 其中的碳固定代谢途径含有完整的 C_4 代谢途径, 并涉及二羧酸循环、糖酵解糖异生作用以及苯丙氨酸、酪氨酸和色氨酸的生物合成等3个间接反应(图1).

2.3 碳固定代谢途径分析

基于玛氏骨条藻转录本的功能注释及其代谢途径分析, 发现了碳固定代谢途径中的18个相关酶及34个编码基因的转录本, 每个酶有1~4个转录本并列出了每个转录本所对应的FPKM值^[31](每百万片段(fragments)中来自某一基因每千碱基长度的片段数目, 其同时考虑了测序深度和基因长度对片段计数的影响, 是目前最为常用的基因表达水平估算方法)(表1). 同时利用BLASTx比较玛氏骨条藻与已完成基因组测序的假微型海链藻和三角褐指藻的碳固定代谢途径同源基因的一致性, 结果显示, 玛氏骨条藻与同属中心纲的假微型海链藻在同源基因序列方面有着更高的一致性, 而与羽纹纲的三角褐指藻序列的一致性相对较低(表1).

通过对KEGG结果的分析, 构建了存在于玛氏骨条藻中的碳固定通路(图1). 其中含有的酶可以进行完整的 C_4 循环代谢途径, 这一机制的运行在碳固定代谢途径乃至光合作用中可以起到非常重要的作用. 碳固定代谢途径作为植物体内最基本的代谢途径, 是植物体进行物质合成和能量代谢的基础, 其中的Rubisco是碳固定代谢途径的起始酶, 广泛存在于光

图1 基于玛氏骨条藻转录组数据的 $de novo$ 拼接及注释结果构建的碳固定代谢途径。酶由长方形框出，代谢通路由椭圆框出

Figure 1 Carbon fixation pathway reconstructed based on the $de novo$ assembly and annotation of *S. marinoi* transcriptome. Enzymes are showed by rectangles, and metabolism pathways that are involved in the carbon fixation pathway are showed by ovals

合自养的植物体内，同时也是碳固定代谢途径中的关键性限速酶和调节酶^[32]。该酶存在 CO_2 和 O_2 共同的结合位点，两种类型酶活性反应的进行程度，完全取决于它们周围 CO_2 和 O_2 的比值，高比值有利于 CO_2 的羧化反应，反之则倾向于发生氧化反应，即光呼吸作用。光呼吸作用不同于普通的呼吸作用，不能提供能量，反而消耗巨大能量^[33]。因此， C_4 代谢途径对 CO_2 的浓缩情况不但引起光呼吸途径的强弱变化，还影响到光合作用的效率并最终作用于有机物的积累。通过 C_4 代谢途径的运行，Rubisco附近 CO_2 的浓度得以提高，减少Rubisco氧化反应的进行，光呼吸作用强度降低，达到减少有机物消耗的目的。所以相对于 C_3 植物来说， C_4 植物拥有更高的光合作用效率和生物量积累效率。这也很好地解释了硅藻在海洋初级生产中的巨大贡献。

C_4 途径在本质上是具有对 CO_2 气体的收集、浓缩和转运途径，它帮助植物完成了 CO_2 的重新分配，但 C_4 并不能代替 C_3 途径完成有机物的积累，所以无论是 C_4 植物还是 C_3 植物，以及为了适应干旱的沙漠气候所演变出来的含有CAM代谢途径的植物体内，由 C_3 途径所进行的碳固定都是必不可少的。虽然玛氏骨条藻是单细胞的海洋硅藻，但转录组的数据表明，其含有 C_4 代谢途径中全部酶类的编码基因，因此推测玛氏骨条藻可以像其他种类的硅藻一样^[14]，通过单细胞内不同细胞器和各类膜结构，以及胞内不同区域含有的相关酶类的共同作用，保证 C_4 代谢途径的进行。根据细胞内参与反应的酶的分布研究^[34]和KEGG的注释结果，得出了玛氏骨条藻进行碳固定代谢途径的示意图(图1)。其中，细胞在 C_4 代谢途径中最初参与固定 CO_2 的是磷酸烯醇式丙酮酸羧化酶

表 1 玛氏骨条藻转录组注释到的存在于碳固定途径的酶及转运蛋白的转录本^{a)}

Table 1 Enzymes and transporters transcripts which is associated with carbon fixation metabolism annotated in the *S. marinoi* transcriptome^{a)}

酶(蛋白)名称(英文名/简称)	EC编号	本转录组中 基因ID	氨基 酸数	***FPKM 值的表达情况	与其他硅藻序列比对一致性% (GenBank登录号)	
					假微型海链藻	三角褐指藻
磷酸烯醇式丙酮酸羧化酶 (PEPC)	4.1.1.31	Comp20135	1010	19.3±13.61 (10.02~52.98)	82 (XP_002290059)	65 (XP_002182829)
		Comp18857	1034	3.12±1.97 (0.97~1.97)	*NM	*NM
天冬氨酸转氨酶(aspartate aminotransferase, AAT)	2.6.1.1	Comp15766	447	29.13±31.31 (6.92~91.11)	85 (XP_002286190)	62 (XP_002183857)
		Comp20907	500	371.99±141.6 (87.77~542.36)	84 (XP_002296198)	69 (XP_002185116)
磷酸烯醇式丙酮酸羧激酶 (PEPCK)	4.1.1.49	Comp12883	632	2263.18±552.48 (1228.81~2742.45)	*NM	*NM
苹果酸脱氢酶(malate de- hydrogenase, MDH)	1.1.1.37	Comp15357	386	19.9±9.84 (4.99~37.42)	69 (XP_002296311)	25 (XP_002185521)
丙氨酸转氨酶(alanine transaminase, ALAT)	2.6.1.2	**Comp16171	608	52.96±33.37 (9.4~92.72)	81 (XP_002295840)	60 (XP_002295840)
		Comp20242	657	52.8±16.08 (16.65~66.01)	82 (XP_002289904)	52 (XP_002289904)
丙酮酸磷酸双激酶 (pyruvate orthophosphate dikinase, PPK)	2.7.9.1	**Comp19477	953	1315.45±588.67 (303.4~2318.67)	83 (XP_002290738)	64 (XP_002182372)
苹果酸脱氢酶(NADP ⁺ dependent malic enzyme, MAO)	1.1.1.39	**Comp20017	669	37.02±23.76 (5.76~74.79)	86 (XP_002290550)	64 (XP_002177890)
核酮糖-1,5-二磷酸羧化酶/加 氧酶大亚基 (ribulose-bisphosphate carboxylase large subu- nit, RBCL)	4.1.1.39	Comp60892	491	3.4±1.62 (1.82~5.66)	96 (XP_002297518)	92 (ABK20641.1)
磷酸甘油酸激酶 (phosphoglycerate kinase, PGK)	2.7.2.3	Comp16373	436	7.14±3.13 (3.03~13.34)	86(XP_002296689)	*NM
		Comp19107	549	11.1±5.72 (4.17~22.87)	68 (XP_002288341)	*NM
		Comp14200	435	3.98±2.06 (0.66~6.94)	92 (XP_002291583)	67 (XP_002183701)
		Comp15180	453	834.01±326.21 (199.55~1302.02)	84 (XP_002296699)	85 (XP_002182409)
NADP ⁺ -甘油醛-3-磷酸磷酸 脱氢酶 (glyceralde- hyde-3-phosphate dehy- drogenase, GAP)	1.2.1.13	Comp653909	90	1.32±3.37(0~11.36)	*NM	*NM
丙糖磷酸异构酶 (triosephosphate isomer- ase, TIM)	5.3.1.1	Comp18817	212	75.29±52.82 (26.83~170.8)	*NM	*NM
		Comp12808	305	23.28±7.09 (12.73~37.67)	*NM	*NM
		Comp12991	301	49.02±23.89 (26.15~92.46)	*NM	*NM
		Comp13024	296	206.83±153.42 (108.8~563.82)	83 (XP_002295570)	64 (XP_002181761)
果糖二磷酸醛缩酶 (fructose-bisphosphate aldolase, ALDO)	4.1.2.13	**Comp15487	317	1226.31±857.88 (9.06~2176.81)	81 (XP_002294849)	58 (XP_002176520)
		Comp15985	494	197.55±84.14 (97.65~376.61)	77 (XP_002288320)	77 (XP_002183771)
		**Comp13215	411	341.29±361.78 (76.05~1149.09)	83 (XP_002291506)	76 (XP_002185431)
		**Comp17762	405	1063.73±599.75 (175.48~1830.39)	*NM	75 (XP_002176182)

续表1

酶(蛋白)名称(英文名/简称)	EC编号	本转录组中 基因ID	氨基 酸数	***FPKM 值的表达情况	与其他硅藻序列比对一致性% (GenBank登录号)	
					假微型海链藻	三角褐脂藻
果糖-1,6-二磷酸酶 (fruc- tose-1,6-bisphosphatase, FBP)	3.1.1.11	**Comp19956	464	1.87±1.4 (0.67~4.41)	75 (XP_002295491)	56 (XP_002177290)
		Comp16315	585	31.42±8.62 (20.8~47.2)	63 (XP_002290348)	59 (XP_002178370)
		Comp15596	344	72.51±62.94 (7.69~198.22)	90 (XP_002288820)	78 (XP_002184240)
转羟乙醛酶(transketolase, TKT)	2.2.1.1	Comp16539	753	3.2±1.96 (1.14~7.18)	86 (XP_002286206)	61 (XP_002182454)
		**Comp13741	729	932.64±599.59 (101.42~1548.73)	90 (XP_002287200)	78 (XP_002178656)
景天庚酮糖-1,7-二磷酸酶 (sedoheptu- lose-1,7-bisphosphatase, SEBP)	3.1.3.37	Comp17773	405	3.92±1.75 (1.23~5.99)	80 (XP_002288369)	40 (XP_002177018)
核糖-5-磷酸异构酶(ribose 5-phosphate isomerase, RPI)	5.3.1.6	Comp14939	341	53.48±16.17 (27.8~85.35)	82 (XP_002287305)	70 (XP_002185690)
核酮糖-3-磷酸差向异构酶 (ribulose-3-phosphate epimerase, RPE)	5.1.3.1	**Comp15079	276	240.26±142.82 (27.77~375.97)	*NM	*NM
		Comp13123	247	3.69±3.18 (0.74~8.97)	*NM	38 (XP_002179664)
核酮糖磷酸激酶 (phosphoribulokinase, PRK)	2.7.1.19	**Comp13932	399	145.01±126.55 (2.7~344.03)	87 (XP_002289281)	74 (XP_0021864569)

a) *NM: 玛氏骨条藻中注释的转录本在另两种硅藻中未匹配到相应的序列; **Comp: 与指数期相比有显著差异表达的编码基因;
***FPKM: 该转录本全部样品的FPKM $\bar{x} \pm SE$ (极小值~极大值)

(phosphoenolpyruvate carboxylase, PEPC, EC 4.1.1.13), 该酶存在于细胞质基质中, 可以利用CO₂接触水面后所形成的HCO₃⁻与磷酸烯醇式丙酮酸发生羧化反应^[16], 产物则为草酰乙酸. 之后4C酸经过磷酸烯醇式丙酮酸羧激酶(phosphoenolpyruvate carboxykinase, PEPCK, EC 4.1.1.49)的脱羧反应, 重新释放出CO₂, 经C₃途径再次固定, 余下产物形成磷酸烯醇式丙酮酸, 从而完成C₄途径循环. 根据KEGG得到的代谢途径可以发现, 在C₄代谢途径中, 除了在PEPC和PEPCK的作用下发生的最基本的羧化和脱羧反应以外, 4C酸在从细胞质进入叶绿体的过程中会有多个选择. 一方面草酰乙酸可以直接运输进入叶绿体, 另一方面草酰乙酸也可以在天冬氨酸转移酶(aspartate aminotransferase, AAT, EC 2.6.1.1)和苹果酸脱氢酶(malate dehydrogenase, MDH, EC 1.1.1.82)的作用下, 生成天冬氨酸和苹果酸. 在转录组数据注释中, 只得到了定位在叶绿体内发生反应的苹果酸脱氢酶(EC 1.1.1.37)和依赖NADP⁺苹果酸脱氢酶(NADP⁺ dependent malic enzyme, MAO, EC 1.1.1.39)的编码基因, 所以未对其他类型的酶进行展示. 当3C酸重新

返回细胞质的过程中, 也同时存在多个选择. 磷酸烯醇式丙酮酸可以直接返回细胞质, 或者转化为丙酮酸或丙氨酸返回细胞质. 所以, 如果起运输作用的是丙氨酸转氨酶(alanine transaminase, ALAT, EC 2.6.1.2), 那么3C酸也会以丙氨酸的形式出现.

2.4 差异表达基因的KEGG富集分析以及碳固定途径中关键酶基因的差异分析

以P值的大小为排序基准(越接近于0, 表示富集越显著), 结合KEGG代谢途径, 对3个生长时期的差异表达情况进行富集分析. 在藻类生长进入稳定后, 一直到衰亡期, 差异表达基因所在的途径并不明显, 因此仍将两个时期分别同指数期进行比对. 在稳定期同指数期的对比中发现, 差异最为明显的代谢途径分别是核糖体、氨酰tRNA生物合成、脂肪酸生物合成、丙酮酸代谢、氮代谢和脂肪酸代谢; 而在衰亡期同指数期的比对中, 分别是核糖体、氨酰tRNA生物合成、脂肪酸生物合成、脂肪酸代谢、生物素代谢以及丙酮酸代谢途径中含有较显著的差异表达. 这一结果表明, 在藻类进入稳定期之后, 核糖体及

tRNA的调控活动加强,这也同蛋白质的翻译过程具有直接联系.此外,具有显著差异表达的途径还涉及丙酮酸代谢和脂肪酸代谢,这是植物细胞不断消耗储存在体内的能量以满足机体正常生长和种群延续的具体表现.同时,在稳定期差异十分显著的氮代谢途径,在衰亡期的比对中仅排第68位,这说明氮代谢途径在藻类稳定期的调控极为活跃,同时也是玛氏骨条藻对生活环境中氮元素变化响应机制的具体表现.而重点关注的碳固定代谢途径,在稳定期和衰亡期分别同指数期的比对结果中显示,其 P 值分别处于第27和11位,属于调控活动较为显著的代谢途径,因此,其固碳机制的分析对于在整体上研究玛氏骨条藻在不同生长时期转录水平上的调控也具有一定的意义.

同样以指数期的表达水平为基准,比较了玛氏骨条藻在 C_3 和 C_4 代谢途径中相关酶编码基因在不同生长时期的表达水平.玛氏骨条藻在稳定期和衰亡期,其碳固定代谢途径中的所有编码基因均无显著差异,而与指数期相比有着较显著的差异,在碳固定代谢途径中18个酶对应的34个编码基因中, C_3 和 C_4 代谢途径中存在差异表达的基因分别有7和3个.

如表1和图2所示,在玛氏骨条藻 C_3 代谢途径中,果糖二磷酸醛缩酶(fructose-bisphosphate aldolase, ALDO, EC 4.1.2.13)、果糖-1,6-二磷酸酶(fructose-1,6-bisphosphatase, FBP, EC 3.1.1.11)、转羟乙醛酶(transketolase, TKT, EC 2.2.1.1)、核酮糖-3-磷酸差向异构酶(ribulose-3-phosphate epimerase, RPE, EC 5.1.3.1)、核酮糖磷酸激酶(phosphoribulokinase, PRK, EC 2.7.1.19)的编码基因在稳定期和衰亡期出现显著差异情况. FBP的编码基因在稳定期与指数期表达无明显差异,但在衰亡期则表达上调.在玛氏骨条藻进入稳定期后,ALDO所对应的基因表达较为复杂,所对应的3个unigene中,Comp15487和Comp17762在稳定期和衰亡期都表现出了显著上调,Comp13215则在衰亡期出现显著下降,这与藻类在不同生长时期的生理特性有关. Comp13215的显著下调意味着其中一部分编码基因的表达量降低,并使得对应的代谢途径下调或停止,这样生物体可以节省一定的能量,用于其他必需的生理代谢途径.在 C_3 过程中核酮糖-1,5-二磷酸吸收 CO_2 生成果糖-6-磷酸需要经过6个酶的连续催化,分别是核酮糖二磷酸羧化酶/加氧酶(Rubisco)、磷酸甘油酸激酶(PGK)、甘油醛-3-磷酸脱

氢酶(GAPA)、丙糖磷酸异构酶(TIM)、果糖二磷酸醛缩酶(ALDO)、果糖-1,6-二磷酸酶(FBP).光合生物固定 CO_2 的限速步骤主要集中在 C_3 代谢途径,该途径又是所有光合生物碳素光合流动过程的中心环节^[35],而其中的ALDO是 CO_2 固定后第一个催化3C化合物生成6C化合物的酶,FBP则是控制果糖-6-磷酸合成的酶类,同时该酶催化的反应不可逆并处于代谢物离开 C_3 循环途径流入淀粉合成的分支点上.因此, C_3 代谢途径中的FBP在决定生物体碳素的最终分配上可以起到重要作用.2001年,Miyagawa等人^[36]将存在于藻类体内的FBP转入烟草(*Nicotiana tabacum*)的叶绿体中,在大气正常 CO_2 (360 ppm)浓度下,转基因烟草植株的光合作用效率和糖类的积累量有了显著提高,首次证明了通过改变一种酶的作用,就可以提高光合作用效率.所以在衰亡期表现出的FBP上调意味着糖类合成通路的加强,生物体内可以不断地生成蔗糖和淀粉并存在于体内.通过这样的生理过程,可以在体内最大程度地积累有机物,以此来抵御环境的胁迫作用,从而使种群最大限度地得以延续,这也是物种适应环境不断进化的重要表现.此外,合成核酮糖-1,5-二磷酸所需酶的编码基因表达量也存在明显变化. TKT, RPE和PRK编码基因的表达在稳定期和衰亡期均有所上调,且上调倍数在稳定期大于衰亡期,表明玛氏骨条藻在稳定期需要更多的核酮糖-1,5-二磷酸以固定更多的 CO_2 进入到新的 C_3 代谢途径中. TKT, RPE和PRK等基因表达的变化体现了藻类在不同生长环境下能量固定以及储能代谢途径的变化,这是碳固定代谢途径做出的分子响应.通过监测这些表达量的变化,有助于了解藻类发展、消亡的过程以及其内部有机物积累的变化情况.

作为在 C_4 代谢途径的限速酶,PPDK编码基因的上调表达可以转化更多丙酮酸以生成磷酸烯醇式丙酮酸而加速 C_4 代谢途径,为 C_3 代谢途径不断输送 CO_2 ,从而维持植物体对无机碳的高同化效率.在赤潮暴发和消亡过程中,由于初期种群数量的不断增加,可以利用的营养物质不断减少,同时光合作用所产生的 O_2 也在不断积累,造成周围环境中 O_2 的浓度增加,从而与 CO_2 共同竞争Rubisco上的活性位点.虽然在不同生长时期的比对中Rubisco大亚基编码基因的表达量并没有显著的变化,但光呼吸作用的消耗和应对逆境胁迫所需要的能量所占比例在不断上升,所以藻类为了保证足够的碳同化量,表现出PPDK编

图2 玛氏骨条藻碳固定代谢途径中差异表达的基因. N代表同指数期相比, 未出现差异表达. 数值表示 $\log_2$ (差异倍数) (指数期表达情况视为倍数1, 所以图中指数期数值为0)

Figure 2 Different gene expression involving in carbon fixation metabolism. The character N indicates that there are no significant differences in gene expressions in stationary phase. The numbers are the values of $\log_2$ (fold change). (The fold change in exponential phase is 1, so the value in exponential phase is 0)

码基因的显著上调. MAO编码基因在稳定期和衰亡期也有显著性上调表达, 这说明玛氏骨条藻在进入稳定期以后, 作为 C_4 代谢途径的中间产物, 磷酸烯醇式丙酮酸和丙酮酸通过不断积累促进了含碳有机物量的变化、重新分配以及输送转运. C_4 途径中可以催化羧化反应和脱羧反应的PEPC和PEPCK在3个时期中的编码基因表达方面均未出现显著性的变化, 这也可能是由于两种酶所催化的反应一直未达到饱和状态而造成的. 此外, 在这些显著差异表达的编码基因中, 还涉及氨基酸的转运过程. ALAT对应编码基因的显著上调, 表明植物体同时也可以通过丙氨酸和丙酮酸之间的转化或其他氨基酸的转氨反应生成丙氨酸, 从而满足 C_4 代谢途径中对于中间产物的需要, 并保证PEPC的高效运转. 由此可见, 开展这些编码基因

的差异性表达研究及分析, 有助于更加清晰地了解赤潮生物在不同生长阶段碳固定代谢过程中的中间产物积累和流向, 深入理解藻类碳固定代谢途径.

3 结论

通过对玛氏骨条藻转录组测序和分析, 发现存在18个与碳固定代谢途径相关的酶及34个对应的编码基因. 序列比对结果表明, 这些编码基因与假微型海链藻有较高的序列一致性. 基于KEGG注释结果, 分析了玛氏骨条藻的碳固定代谢途径, 发现玛氏骨条藻含有类似于陆地 C_4 植物 CO_2 浓缩机制所需的全部酶类, 这是该类群在 CO_2 浓度限制环境下仍然可以形成很高初级生产力的主要原因. 通过比较玛氏骨条藻不同生长时期的转录组以及KEGG差异分析的

表达结果, 发现在其碳固定代谢途径中一些关键酶的编码基因存在差异表达情况, 其中C₃途径中决定碳素最终分配去向的果糖二磷酸醛缩酶的编码基因在衰亡期有显著上调表达, 而C₄途径中的限速酶丙酮酸磷酸双激酶的编码基因在稳定期和衰亡期也表现为显著上调. 这些编码基因的表达变化可以影响磷酸烯醇式丙酮酸、果糖-6-磷酸以及核酮糖二磷酸

等产物的产量变化, 这是玛氏骨条藻在种群数量变化过程中能量固定和产物分配的适应性表现, 同时也是应对环境变化的分子响应. 这些能量固定和代谢途径方面的发现, 为今后从蛋白质组、代谢组角度探索碳固定代谢途径提供了新的思路, 也为深入了解藻类的碳固定代谢机制以及碳的生物地球化循环增添了新的认识.

参考文献

- 1 Cerling T E, Wang Y, Quade J. Expansion of C₄ ecosystems as an indicator of global ecological change in the late Miocene. *Nature*, 1993, 361: 344–345
- 2 Sage R F. The evolution of C₄ photosynthesis. *New Phytol*, 2004, 161: 341–370
- 3 Deines P. *Handbook of Environmental Isotope Geochemistry*. Amsterdam: Elsevier, 1980, 1: 329–406
- 4 Farquhar G D, Ehleringer J R E, Hubick K T. Carbon isotope discrimination and photosynthesis. *Annu Rev Plant Physiol Mol Biol*, 1989, 40: 503–537
- 5 Park R, Epstein S. Carbon isotope fractionation during photosynthesis. *Geochim Cosmochim Acta*, 1960, 21: 110–126
- 6 Haimovich-Dayana M, Garfinkel N, Ewe D, et al. The role of C₄ metabolism in the marine diatom *Phaeodactylum tricornutum*. *New Phytol*, 2013, 197: 177–185
- 7 Cooper T G, Filmer D. The active species of “CO₂” utilized by ribulose diphosphate carboxylase. *J Biol Chem*, 1969, 244: 1081–1083
- 8 Murray R B, John A T, Whitney S M, et al. The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO₂-concentrating mechanisms in algae. *Can J Bot*, 2011, 76: 1052–1071
- 9 Roberts K, Granum E, Leegood R C, et al. Carbon acquisition by diatoms. *Photosynth Res*, 2007, 93: 79–88
- 10 Giordano M, Beardall J R, John A. CO₂ concentrating mechanisms in algae: Mechanisms, environmental modulation, and evolution. *Annu Rev Plant Biol*, 2005, 56: 99–131
- 11 Tortell P D, Reinfelder J R, Morel F M M. Active uptake of bicarbonate by diatoms. *Nature*, 1997, 390: 243–244
- 12 Reinfelder J R. Carbon concentrating mechanisms in eukaryotic marine phytoplankton. *Ann Rev Mar Sci*, 2011, 3: 291–315
- 13 Bowes G. Single-cell C₄ photosynthesis in aquatic plants. In: *C₄ Photosynthesis Related CO₂ Concentrating Mechanisms*. Berlin: Springer Netherlands, 2010. 63–80
- 14 Reinfelder J R, Kraepiel A M, Morel F M. Unicellular C₄ photosynthesis in a marine diatom. *Nature*, 2000, 407: 996–999
- 15 Reinfelder J R, Milligan A J, Morel F M, et al. The role of the C₄ pathway in carbon accumulation and fixation in a marine diatom. *Plant Physiol*, 2004, 135: 2106–2111
- 16 Granum E, Raven J A, Leegood R C. How do marine diatoms fix 10 billion tonnes of inorganic carbon per year? *Can J Bot*, 2005, 83: 898–908
- 17 McGinn P J, Morel F M. Expression and inhibition of the carboxylating and decarboxylating enzymes in the photosynthetic C₄ pathway of marine diatoms. *Plant Physiol*, 2008, 146: 300–309
- 18 Armbrust E V, Berges J A, Bowler C, et al. The genome of the diatom *Thalassiosira pseudonana*: Ecology, evolution, and metabolism. *Science*, 2004, 306: 79–86
- 19 Kroth P G, Chiovitti A, Gruber A, et al. A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. *PLoS One*, 2008, 3: e1426
- 20 Field C B, Behrenfeld M J, Randerson J T, et al. Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science*, 1998, 281: 237–240
- 21 Keller A A, Riebesell U. Phytoplankton carbon dynamics during a winter-spring diatom bloom in an enclosed marine ecosystem: Primary production, biomass and loss rates. *Mar Biol*, 1989, 103: 131–142
- 22 Checkley D M, Entzeroth L C, Checkley D M, et al. Elemental and isotopic fractionation of carbon and nitrogen by marine, planktonic copepods and implications to the marine nitrogen cycle. *J Plankton Res*, 1985, 7: 553–568
- 23 Fry B, Wainright S C. Diatom sources of ¹³C-rich carbon in marine food webs. *Mar Ecol Prog*, 1991, 76: 149–157
- 24 Boyd P, Newton P. Evidence of the potential influence of planktonic community structure on the interannual variability of particulate organic carbon flux. *Deep-Sea Res Pt I*, 1995, 42: 619–639

- 25 Falkowski P G, Raven J A. Aquatic Photosynthesis. Princeton: Princeton University Press, 2007
- 26 Grabherr M G, Haas B J, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*, 2011, 29: 644–652
- 27 Camacho C, Coulouris G, Avagyan V, et al. BLAST⁺: Architecture and applications. *BMC Bioinformatics*, 2009, 10: 421
- 28 Götz S, García-Gómez J M, Terol J, et al. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res*, 2008, 36: 3420–3435
- 29 Moriya Y, Itoh M, Okuda S, et al. KAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res*, 2007, 35: W182–W185
- 30 Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*, 2010, 11: R106
- 31 Trapnell C, Williams B A, Pertea G, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*, 2010, 28: 511–515
- 32 Mei Y, Li H L, Xie J, et al. Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) (in Chinese). *Plant Physiol Commun*, 2007, 43: 363–368 [梅杨, 李海蓝, 谢晋, 等. 核酮糖-1,5-二磷酸羧化酶/加氧酶(Rubisco). *植物生理学报*, 2007, 43: 363–368]
- 33 Bauwe H, Hagemann M, Fernie A R. Photorespiration: Players, partners and origin. *Trends Plant Sci*, 2010, 15: 330–336
- 34 Hatch M D, Slack C R. Photosynthesis in sugar-cane leaves a new carboxylation reaction and the pathway of sugar formation. *Biochem J*, 1966, 101: 103–111
- 35 Iwaki T, Wadano A, Yokota A, et al. Aldolase—an important enzyme in controlling the ribulose 1,5-bisphosphate regeneration rate in photosynthesis. *Plant Cell Physiol*, 1991, 32: 1083–1091
- 36 Miyagawa Y, Tamoi M, Shigeoka S. Overexpression of a cyanobacterial fructose-1,6-/sedoheptulose-1,7-bisphosphatase in tobacco enhances photosynthesis and growth. *Nat Biotechnol*, 2001, 19: 965–969

Description of carbon fixation pathway based on *Skeletonema marinoi* transcriptome

LIU Qian^{1,2,3}, MI TieZhu^{2,3,4}, ZHEN Yu^{2,3,4}, WANG HuaLong^{1,2,3} & YU ZhiGang^{3,5}

¹ College of Marine Life Science, Ocean University of China, Qingdao 266003, China;

² Key Laboratory of Marine Environment and Ecology, Ministry of Education, Ocean University of China, Qingdao 266100, China;

³ Laboratory for Marine Ecology and Environmental Science, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266071, China;

⁴ College of Environmental Science and Engineering, Ocean University of China, Qingdao 266100, China;

⁵ Key Laboratory of Marine Chemical Theory and Technology, Ministry of Education, Ocean University of China, Qingdao 266100, China

Diatoms can fix approximately 10^{16} g CO₂ into organic carbon every year, equivalent to roughly 20% of global primary production. They are the basis for important food webs, having considerable impacts on carbon cycle and biological carbon pump that draws down CO₂ from the atmosphere to the ocean interior. Nevertheless, the $K_{1/2}$ of their ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is 40~60 $\mu\text{mol L}^{-1}$, higher than the apparent photosynthetic $K_{1/2}$ (CO₂) for ambient CO₂. Growth and kinetic studies have shown that diatoms can avoid CO₂ limitation through CO₂ concentrating mechanisms. Recently the carbon fixation metabolism pathway in diatoms is further elucidated by the presence of a complete set of genes essential for this metabolic route in *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* because of their entire available genome information. In this study, we used RNA-Seq technology to confirm the existence of necessary genes for plant-like C₄ activity in *Skeletonema marinoi*, which is a complex biological trait that enables *S. marinoi* to either has high water-use efficiency or accumulate biomass at a fast rate. The carbon fixation metabolism pathway of *S. marinoi* involves 20 enzymes and 34 coding genes, including the complete key enzymes especially for C₄ pathway, which are phosphoenolpyruvate carboxylase (PEPC), phosphoenolpyruvate carboxykinase (PEPCK) and pyruvate orthophosphate dikinase (PPDK). Comparing with the homologous genes of carbon fixation metabolism pathway in *T. pseudonana* and *P. tricornutum*, these genes showed little divergence. Besides, the gene differential expression of *S. marinoi* in carbon fixation metabolism pathway was identified at growth stages using digital gene expression profiling. The results indicated the number of differential expression genes was 7 and 3 in C₃ and C₄ pathway, respectively. As for the differential expression of coding genes in growth stages, there were 8 and 10 coding genes in the stationary and decline phase, respectively versus to those in the exponential phase. The gene expression of fructose-1,6-bisphosphatase (FBP), fructose-bisphosphate aldolase (ALDO) and PPDK was changed significantly after the exponential phase. The FBP and ALDO are not only responsible for the carbon fixation metabolism pathway, but also as the key enzymes in the glycolytic metabolism pathway, which links closely with the flow direction of carbon and the synthesis of carbohydrates. The PPDK, a cardinal enzyme of the C₄ pathway, catalyzing the regeneration of phosphoenolpyruvate (PEP), evolved as an adaptation to high light intensity, high temperature, low concentration of CO₂ and dryness. Therefore, the formation of PEP through PPDK is considered to be very important in the C₄ pathway. This study will contribute to analyze the gene regulation of key enzymes involved in the carbon metabolism of *S. marinoi*, explain the reason for huge amounts of primary production by diatom, and enhance our knowledge on the carbon fixation mechanism, which provides a new insight into understanding carbon biogeochemical cycle.

***Skeletonema marinoi*, carbon fixation metabolism pathway, transcriptome**

doi: 10.1360/N972016-00046