

文章编号: 1674-5566(2012)05-0662-09

缺刻缘绿藻转录组测序及脂质代谢相关基因注释

陈思弘, 周志刚

(上海海洋大学 水产与生命学院, 上海 201306)

摘要: 为了能深入地了解缺刻缘绿藻花生四烯酸(ArA)和脂质的代谢过程, 利用 Roche 454 GS FLX 测序仪对该藻转录组进行高通量的焦磷酸测序。得到高质量读序(read)382 468 条, 占原始读序的 97.14%, 平均每条读序长 322 bp, 总大小达 123 Mb。经 CAP3 软件拼接得到 22 714 条重叠群、25 621 条 singleton。将这些序列与公共数据库进行同源性搜索、比较、基因功能注释和分类。基于转录组中所注释的基因构建缺刻缘绿藻脂质代谢途径: 脂肪酸是在叶绿体内从头合成, 然后游离脂肪酸进入胞质, 由内质网进行三酰甘油的合成, 最后可能在油体蛋白作用下形成油滴并储在于细胞中。ArA 自油酸是开始经过多个去饱和酶及延长酶的作用而产生的。油酸是由硬脂酰-ACP 去饱和酶作用而形成的, 而棕榈油酸是在叶绿体的糖脂上去饱和而产生的。三酰甘油最终被分解成自由脂肪酸和丙酮酸, 而长链脂肪酸是以酰基-辅酶 A 的形式逐级降解的。

研究亮点: 首次利用次世代大规模平行测序技术对缺刻缘绿藻的转录组进行测序, 得到大量的基因信息, 并利用生物信息学手段构建了缺刻缘绿藻脂质代谢途径, 为缺刻缘绿藻脂质代谢调控、基因改造等研究提供翔实的数据。

关键词: 缺刻缘绿藻; 转录组; 焦磷酸测序; 脂质代谢途径

中图分类号: S 917

文献标志码: A

花生四烯酸(arachidonic acid, ArA)是一种 ω -6系长链多不饱和脂肪酸, 是哺乳动物脑细胞中膜磷脂的主要组成成分之一^[1]。它广泛存在于动物组织中。研究表明, ArA 有助于神经系统信号的传递, 并能有效降低血糖、血脂以及胆固醇指标^[2]。鉴于 ArA 对人类健康的作用及其药理学价值, 加之其资源非常有限, 所以加强对 ArA 功效研究与开发富含 ArA 的生物资源, 具有十分重要的意义。

缺刻缘绿藻(*Myrmecea incisa*)是一种隶属于绿藻门(Chlorophyta)、共球藻纲(Trebouxiophyceae)的球状微藻^[3], 也是一种富含 ArA 的微藻, 其 ArA 的含量可达藻体干重的 7%^[4]。有关该藻 ArA 的合成代谢仅限于脂肪酸组分含量的变化^[4]、几个脂肪酸去饱和酶基因^[5]及延长酶基因^[6-7]的分析, 但对于该藻整个脂肪

酸及油脂的合成代谢途径了解甚少。因此, 为了阐明和完善缺刻缘绿藻 ArA 等脂质合成和分解途径及其调控机理等研究, 本研究借鉴新一代 454 焦磷酸测序系统近期进展与研究成果^[8-9], 利用它对缺刻缘绿藻转录组进行高通量测序, 在获得大量基因序列的基础上进行功能注释并构建脂质合成与分解代谢的完整途径, 为缺刻缘绿藻脂质代谢调控、基因改造等研究奠定扎实的基础。

1 材料与方法

1.1 藻种的培养

缺刻缘绿藻 H4301 购自布拉格查理大学藻类培养中心(CAUP)。将藻接种于 BG-11 培养基^[10], 采用内径为 3.5 cm 的圆柱体玻璃管(容量约 800 mL)作为光生物反应器, 培养温度为 25

收稿日期: 2012-04-16 修回日期: 2012-06-16

基金项目: 国家自然科学基金(30972243, 31172389); 国家海洋局海洋可再生能源专项基金项目(SHME2011SW02); 上海市教育委员会海洋生物学重点学科(J50701)

作者简介: 陈思弘(1986—), 男, 硕士研究生, 研究方向为藻类生物技术。E-mail: mcrommelc@163.com

通讯作者: 周志刚, E-mail: zgzhou@shou.edu.cn

℃。利用 Philips 40 W 冷荧光灯管提供光照强度为 $120 \mu\text{mol photons}/(\text{m}^2 \cdot \text{s})$ 的单侧光源^[11],光照和黑暗时间比为 12 h:12 h。同时向培养基中充入过滤空气。培养至对数期(约 14 d),5 500 r/min 离心 5 min 收集全部藻体,并用灭菌水清洗 3 次,然后全部转入新鲜培养基中再培养 2 d。同样方法离心、清洗和收集,并在液氮中冻存备用。

1.2 总 RNA 的提取及 mRNA 的纯化

冻存样本在研钵中加入液氮充分研磨,然后用 TRIzol 试剂(Invitrogen, USA)提取总 RNA^[12],提取后溶解在 200 μL 除 RNA 酶水中。用 NanoDrop(Thermo Scientific, USA)检测总 RNA 浓度,用 Agilent 2100 Bioanalyzer 生物分析仪(Agilent, USA)中的 RNA 6000 Pico LabChip 检测纯度。

总 RNA 用 10 单位的 DNase I (Ambion, USA)于 37 °C 处理 1 h,然后加入 250 μL 除核酸酶水。mRNA 使用 MicroPoly (A) 纯化试剂盒(Ambion, USA)进一步纯化。mRNA 溶解在 100 μL 的 RNA 保存液中,并用 NanoDrop 检测纯度。

1.3 互补 DNA 链的合成和文库制备

由 mRNA 合成双链 cDNA^[13]。用 10 μg mRNA 和 1 000 单位的 Superscript II 逆转录酶(Invitrogen),加入 *Gsu* I-oligo dT 引物先合成 cDNA 的第一条链。在 42 °C 下反应 1 h 后,mRNA 的 5'-CAP 结构被 NaIO_4 (Sigma, USA)氧化,即被生物素酰肼捆绑成完整的 mRNA/cDNA,后者与 Dynal M280 磁珠(Invitrogen)相结合以合成 cDNA 的第二链。当合成完毕后,用 *Gsu* I 去除 polyA 尾巴和 5' 接头。

利用 cDNA 分级柱(Agencourt, USA)按片段大小分离 cDNA。当 cDNA 片段大于 800 bp 时,用超声波将其打断至 300 ~ 800 bp,然后与其他的 300 ~ 800 bp 的 cDNA 样本混合。制备好的 cDNA 使用 GS DNA 文库制备试剂盒(Roche Applied Science, USA)转化为单链模板 DNA (single-stranded template DNA, sstDNA)文库。

1.4 454 焦磷酸测序

单链模板 DNA 文库使用 GS emPCR 试剂盒(Roche Applied Science, USA)以磁珠固定的方式扩增并用 454 GS FLX 系统(Roche Diagnostic, USA)测序(由国家人类基因组南方研究中心完成)。FLX 测序仪产生的读序由 SeqClean 软件

(<http://compbio.dfci.harvard.edu/tgi/>)去除接头序列、polyA 尾巴。并由一个内部开发的程序过滤掉低质量的读序,合格的读序交由 CAP3 软件^[14]以默认参数拼接。拼接后产生的序列称为重叠群,不能被拼接的读序称为 singleton。重叠群与 singleton 统称为单一序列。

1.5 功能注释

将上述重叠群和 singleton 用系统内部自带软件(基于 EMBOSS 软件包^[15]中的“GetORF”组件开发)预测开放阅读框(open reading frame, ORF)。以 $E\text{-value} \leq 1 \times 10^{-3}$ 为阈值,基于 BLASTp 算法与 Swiss-Prot 和 GenBank 数据库进行比对以标注 ORF,然后选择最佳结果作为基因注释。以 $E\text{-value} \leq 1 \times 10^{-3}$ 为阈值,利用 GoPipe 软件^[16]基于 BLASTp 算法与 Swiss-Prot 和 TrEMBL 数据库比对以进行基因本体论(gene ontology, GO)分析并选择最佳结果(<http://www.geneontology.org/>)。GO 注释信息分别按照分子功能、生物学途径和细胞组分进行分类。并加注酶学委员会(enzyme commission, EC)编号。

1.6 脂质合成途径的构建

利用 KEGG 数据库(<http://www.genome.jp/kegg/>)^[17],基于 BBH (bi-directional best hit) 算法^[18]将序列的注释信息映射到 KEGG 数据库。根据 KEGG pathway 注释系统输出的代谢途径信息,确定参与缺刻缘绿藻脂质代谢途径的酶基因信息以及它们在脂质代谢途径中的定位。利用绘图软件 SmartDraw (<http://www.smartdraw.com.cn/>)绘制脂质代谢图。

2 结果与讨论

2.1 454 测序和读序拼接

测序的原始数据为 393 722 条读序,平均长度为 333 bp。经过去除引物和 polyA 尾巴以及短的低质量序列等,得到 382 468 条高质量读序。它们占原始序列的 97.14%,平均长度为 322 bp,总大小达 123 Mb。利用序列拼装软件 CAP3^[14]对序列进行拼接,得到 22 714 条重叠群,平均长度为 639 bp,以及 25 621 条 singleton,平均长度为 277 bp(表 1)。77.7% 的重叠群是由 2 ~ 10 条读序拼接而成(表 2)。将本研究的高质量读序及平均长度与利用同样方法在杜氏藻(*Dunaliella tertiolecta*)^[19]和海洋弯曲甲藻(*Oxyrrhis*

marina)^[20]的测序结果作比较,发现它们均明显高于海洋弯曲甲藻(分别为79.66%和228 bp)但稍低于杜氏藻(分别为98.55%和400 bp)。由于插入片段大小不同会导致测序区间不同,同一基因可能被分在了不同的重叠群或 singleton 中,或者是由于前体 mRNA 转录后加工成不同的产物。因此,这些重叠群及 singleton 数并不代表这么多多个基因,为此将它们统称为单一序列。

表 1 重叠群和 singleton 的长度分布
Tab. 1 Length distribution of assembled contig and singletons

序列长度/bp	重叠群数	singleton 数
40 ~ 99	132	3 185
100 ~ 199	545	4 051
200 ~ 299	1 261	5 259
300 ~ 399	4 107	8 991
400 ~ 499	5 181	4 094
500 ~ 599	2 950	41
600 ~ 699	2 152	0
700 ~ 799	1 493	0
800 ~ 899	1 080	0
900 ~ 999	803	0
1 000 ~ 1 499	1 962	0
1 500 ~ 1 999	626	0
>2 000	422	0

表 2 拼接成每条重叠群的读序数量分布
Tab. 2 Summary of component reads per assembly

拼接每条重叠群的读序数	重叠群数
2 ~ 10	17 649
11 ~ 20	2 561
21 ~ 30	903
31 ~ 40	474
41 ~ 50	256
51 ~ 100	475
101 ~ 150	132
151 ~ 200	75
>200	189

2.2 功能注释

所有的单一序列用 BLASTp 算法与国家生物科技信息中心(national center for biotechnology information, NCBI)的非冗余蛋白质数据库进行比对。共有 7 633 条重叠群(占总数的33.76%)和 3 478 条 singleton(占总数的15.20%)有 BLAST 比对结果(表 3)。经 BLAST 比对分析,大约具有功能注释 32% 的单一序列比对结果与微藻有关(表 4),这些微藻主要是衣藻、团藻(*Volvox carteri*)、*Ostreococcus* 和微胞藻

(*Micromonas*),显示缺刻缘绿藻的转录组文库具有明显的微藻特性。

表 3 缺刻缘绿藻转录组注释概要
Tab. 3 *Myrmecea incisa* transcriptome annotation summary

	序列数		
	重叠群	singleton	单一序列
总计序列	22 612	22 878	45 490
BLAST 比对序列	7 633	3 478	11 111
基因本体论注释序列	2 297	1 156	3 453
加注酶学委员会编号序列	2 036	451	2 487

表 4 缺刻缘绿藻单一序列经 BLAST 比对有匹配的物种分布

Tab. 4 Distribution of top microalgal species against BLAST matches by *M. incisa* unisequences

匹配物种	单一序列数	占比对结果百分比/%
非藻类物种	7 454	67.09
衣藻	1 809	16.28
团藻	562	5.06
<i>Ostreococcus</i> spp.	359	3.23
微胞藻	315	2.84
其他微藻	490	4.40
大型藻类	122	1.10

GO 注释分类的结果(表 5)表明,从细胞构成的角度来看,约 80% 的单一序列是在细胞、胞内及胞质中起作用的;从分子功能的角度来分析,2 251 条单一序列具有催化活性,2 597 条起连接作用,它们占注释总量的 50%;从生物学过程来看,它们主要在代谢、细胞过程中起作用,占单一序列总量的 45%。这些转录组的数据反映出缺刻缘绿藻在细胞代谢、细胞周期等方面的旺盛程度。在 GO 注释的 3 453 条单一序列(2 297 条重叠群和 1 156 条 singleton)中,分别有 2 036 条重叠群和 451 条 singleton 标注了 EC 编号(表 3)。

由于缺刻缘绿藻并非模式生物,对其转录组的研究目前仍处于起步阶段,因此与其他藻类^[19-24]的研究相比,得到有注释信息(表 3)的序列占单一序列总量的比率偏低。随着微藻能源化工作的开展及微藻基因组信息的丰富,相信在不久的将来,这一比率会显著提高。

2.3 缺刻缘绿藻脂肪酸及脂类合成代谢途径的构建

在适宜条件下,微藻主要以酯化作用合成极性的甘油膜磷脂(主要包括糖基甘油酯和甘油磷

脂),而在环境胁迫下,会合成并积累中性的三酰

表 5 缺刻缘绿藻转录组 GO 注释结果分类
Tab.5 Distribution of most abundant Gene ontology (GO) terms assigned to the *M. incisa* transcriptome

单一序列数	
细胞	3 034
细胞内	2 786
细胞质	2 191
膜	889
细胞核	860
按细胞组 分分类	
细胞外区域	85
染色体	85
细胞外包被结构	58
细胞外空间	5
细胞表面	4
蛋白样细胞外基质	1
结合	2 775
催化活性	2 456
水解酶活性	993
蛋白质结合	818
核酸结合	721
转移酶活性	628
氧化还原酶活性	472
转运子活性	222
连接酶活性	219
裂解酶活性	200
结构分子活性	197
按分子功 能分类	
激酶活性	187
解旋酶活性	156
离子跨膜转运子活性	116
异构酶活性	107
肌动蛋白活性	89
转录调控子活性	76
电子载体活性	65
信号传导器活性	41
酶调控子活性	37
抗氧化剂活性	35
蛋白转运子活性	23
受体活性	22
通道活性	8
整合酶活性	1
细胞过程	2 819
代谢过程	2 708
大分子代谢过程	1 519
生物合成过程	1 295
核苷及核苷酸代谢过程	867
刺激应答过程	605
生物学过程调控	571
转运	569
细胞氨基酸及衍生物代谢过程	343
分解代谢过程	333
按生物学 过程分类	
多细胞体过程	269
多细胞体发育	235
多器官过程	63
细胞分化	54
细胞凋亡	19
细胞信息传递	15
分泌	15
细胞组分转移	14
行为	14
细胞膜融合	2
病理发生	2
细胞外功能结构	1

甘油^[23]。尽管全球范围内对真核生物脂肪酸合成途径的研究已经很详细^[25],但对微藻的脂肪酸的合成规律的研究仍很欠缺。由于与产生花生四烯酸的前体物质相关,需要对缺刻缘绿藻的脂质生物合成途径有一个完整的了解。基于转录组功能注释和在代谢途径中的定位,我们构建了缺刻缘绿藻的脂质合成途径。

缺刻缘绿藻脂肪酸的从头合成(图 1)与其他藻类和植物^[26]一样以酰基载体蛋白(acyl carrier protein,ACP)形式开始于叶绿体。乙酰-辅酶 A 羧化酶(acetyl-coenzyme A carboxylase, ACC, EC:6.4.1.2)催化乙酰-辅酶 A 和 HCO₃⁻生成丙二酰-辅酶 A(malonyl-coenzyme A, malonyl-CoA)。随后,丙二酰-CoA 在丙二酰-辅酶 A-ACP 酰基转移酶(malonyl-CoA ACP transacylase, MAT, EC:2.3.1.39)的催化下转化为丙二酰-ACP(malonyl-ACP),为随后的延长反应提供碳。在延长反应中,丙二酰-ACP 中的丙二酰基参与了以酰基-ACP 或乙酰-辅酶 A 为受体的一系列循环缩合反应,这一系列循环缩合反应主要由一个具有多种亚型的酶——酮酰基-ACP 合酶(ketoacyl-ACP synthase, KAS)催化。第一轮缩合反应是由 KAS III(EC:2.3.1.180)催化,产生一个四碳链(C4)产物。其余的缩合反应由 KASI(EC:2.3.1.41)催化,产生碳链长度不同的产物(C6 ~ C16)。

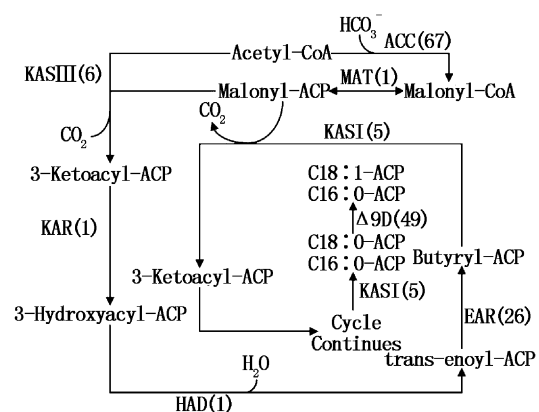


图 1 基于缺刻缘绿藻转录组注释结果重建的脂肪酸合成途径

Fig.1 Fatty acid biosynthesis pathway reconstructed based on the *de novo* assembly and annotation of *M. incisa* transcriptome

括号内的数字为该种酶基因的转录量数据。

每一轮缩合反应还有 3 个步骤:还原、脱水、

还原。为了形成饱和脂肪酸, β -酮酰基-ACP (β -ketoacyl-ACP) 被 β -酮酰基-ACP 还原酶 (β -ketoacyl-ACP reductase, KAR, EC:1.1.1.100) 还原, 再被 β -酮酰基-ACP 脱水酶 (β -ketoacyl-ACP dehydrase, HAD, EC:4.2.1.-) 脱水, 再被烯酰-ACP 还原酶 (enoyl-ACP reductase, EAR, EC:1.3.1.9) 还原。这四步反应成为一个完整的缩合延长反应, 每一轮反应使 ACP 上的碳链延长两个碳。经过六或七轮缩合反应, 最终产物为饱和的棕榈酰 (palmitoyl, C16:0)-ACP 和硬脂酰 (stearoyl, C18:0)-ACP。C18:0-ACP 被硬脂酰-ACP $\Delta 9$ 去饱和酶 ($\Delta 9$ desaturase, $\Delta 9D$, EC:1.14.19.2) 作用, 形成油酰 (oleoyl, C18:1)-ACP。

由于在测序结果中发现了叶绿体中糖脂合成相关酶的基因, 包括单半乳糖甘油二酯 (monogalactosyldiacylglycerol, MGDG) 合酶 (MGDG synthase, EC:2.4.1.46)、 β -1,3-半乳糖酰基转移酶 (β -1,3-galactosyltransferase, EC:2.4.1.179)、半乳糖脂半乳糖酰基转移酶 (galactolipid galactosyltransferase, EC:2.4.1.184)、棕榈酰单半乳糖甘油二酯 $\Delta 7$ 去饱和酶 (palmitoyl MGDG $\Delta 7$ desaturase), 我们推测 C16:0-ACP 在叶绿体中脱去 ACP 后形成 MGDG, 以棕榈酰-MGDG 的形式去饱和为棕榈油酰-MGDG, 后者在半乳糖脂半乳糖酰基转移酶的作用下形成棕榈油酰双半乳糖甘油二酯, 以糖脂的形式参与到叶绿体类囊体膜脂的建成^[27-28]。

缺刻缘绿藻中三酰甘油 (triacylglycerol, TAG) 的合成主要在内质网中 (图 2)。C16:0-ACP 及 C18:1-ACP 在酯酰基-ACP 硫酯酶 (fattyacyl-ACP thioesterase, FAT, EC:3.1.2.14) 的催化下脱去 ACP 成为游离脂肪酸后被转运出叶绿体, 在脂肪酸合酶 (fatty acid synthase, FAS, EC:2.3.1.85) 的催化下与酰基-辅酶 A 结合, 以酯酰-辅酶 A 的形式参与到三酰甘油的合成中。甘油在甘油激酶 (glycerol kinase, GK, EC:2.7.1.30) 的催化下形成 3-磷酸甘油 (glycerol-3-phosphate, G3P), 后者在 3-磷酸甘油酰基转移酶 (glycerol-3-phosphate acyltransferase, GPAT, EC:2.3.1.15) 的催化下与酯酰-辅酶 A 形成溶血磷脂酸 (lyso-phosphatidic acid, LPA), LPA 在溶血磷脂酸酰基转移酶 (lyso-phosphatidic acid acyltransferase, LPAT, EC:2.3.1.51) 的催化下与酯酰-辅酶 A

形成磷脂酸 (phosphatidic acid, PA)。PA 在磷脂酸磷酸酶 (phosphatidate phosphatase, PP, EC:3.1.3.4) 的催化下形成二酰甘油 (diacylglycerol, DAG), 遗憾的是在测序数据中没有发现编码 PP 的基因。随后 DAG 形成 TAG 的途径有两条: (1) 与酯酰-辅酶 A 在二酰甘油酰基转移酶 (diacylglycerol acyltransferase, DGAT, EC:2.3.1.20) 的催化下形成 TAG; (2) 由磷脂酰胆碱 (phosphatidylcholine, PC) 提供酰基, 在磷脂酰二酰甘油酰基转移酶 (phospholipid: diacylglycerol acyltransferase, PDAT, EC:2.3.1.158) 的催化下形成 TAG, PC 脱去一个酰基成为溶血磷脂酰胆碱 (lyso-phosphatidylcholine, lyso-PC)。所合成的 TAG 可能在一种叫做 caleosin 的油体蛋白作用下, 形成在显微镜下可见的油滴^[29]。

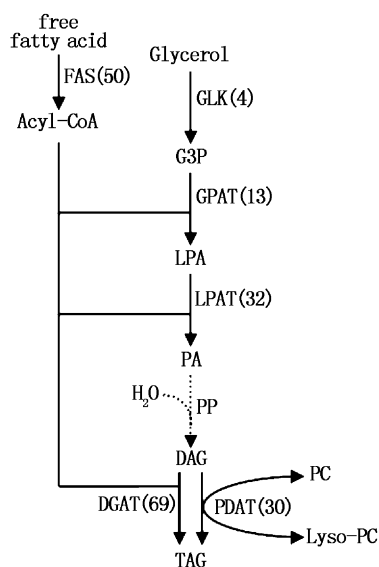


图 2 基于缺刻缘绿藻转录组注释结果重建的 TAG 合成途径

Fig. 2 TAG biosynthesis pathway reconstructed based on the *de novo* assembly and annotation of *M. incisa* transcriptome

括号内的数字为该种酶基因的转录量数据。

脂肪酸的去饱和和延长可能以磷脂酰胆碱 (phosphatidyl choline, PC) 等形式完成的。缺刻缘绿藻中 ArA 的合成途径是由油酸 (oleic acid, OA, C18:1) 开始, 经 $\Delta 12$ 脂肪酸去饱和酶 ($\Delta 12$ desaturase, $\Delta 12D$, EC:1.14.19.-) 去饱和为亚油酸 (linoleic acid, LA, C18:2), 再由 $\Delta 6$ 去饱和酶 ($\Delta 6$ desaturase, $\Delta 6D$, EC:1.14.19.3) 去饱和为 γ -亚麻酸 (γ -linolenic acid, GLA, C18:3), 再由

$\Delta 6$ 延长酶($\Delta 6$ elongase, $\Delta 6E$, EC: 2.3.1.-)延长为双高- γ -亚麻酸(dihomo- γ -linolenic acid, DGLA, C20:3),最后经 $\Delta 5$ 去饱和酶($\Delta 5$ desaturase, $\Delta 5D$, EC: 1.14.19.5)去饱和成为ArA(图3)。该途径与我们对该藻脂肪酸分析的基础上所建立的ArA合成代谢的 $\omega 6$ 途径完全一致^[3]。除此之外,我们在高通量的数据中还发现该藻具有编码 $\omega 3$ 去饱和酶的基因,该酶可能作用于亚油酸或ArA,使后者分别被去饱和为 α -亚麻酸和二十碳五烯酸(eicosapentenoic acid)。

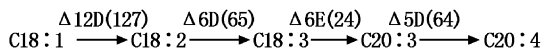


图3 基于缺刻缘绿藻转录组注释结果重建的脂肪酸去饱和与延长途径

Fig.3 Desaturation and elongation of fatty acid pathway reconstructed based on the *de novo* assembly and annotation of *M. incisa* transcriptome
括号内的数字为该种酶基因的转录量数据。

2.4 缺刻缘绿藻脂肪酸及脂类分解代谢途径的构建

图4为三酰甘油降解过程。三酰甘油在三酰甘油脂肪酶(triacylglycerol lipase, TAGL, EC: 3.1.1.3)的催化下分解为游离脂肪酸和甘油,甘

油在甘油激酶(glycero kinase, GLK, EC: 2.7.1.30)催化下形成3-磷酸甘油,3-磷酸甘油由3-磷酸甘油脱氢酶(glycerol-3-phosphate dehydrogenase, GPD, EC:1.1.1.8)脱氢形成磷酸二羟基丙酮,然后被磷酸丙糖异构酶(phosphotriose isomerase, PTI, EC: 5.3.1.1)催化形成3-磷酸甘油醛,之后被3-磷酸甘油醛脱氢酶(glyceraldehyde-3-phosphate dehydrogenase, GAPD, EC: 1.2.1.9)脱氢形成1,3-二磷酸甘油酸。1,3-二磷酸甘油酸会经两个途径形成3-磷酸甘油酸:途径(1)直接被3-磷酸甘油酸激酶(3-phosphoglycerate kinase, PGK, EC: 2.7.2.3)催化生成;途径(2)先被二磷酸甘油酸变位酶(biphosphoglycerate mutase, PM, EC: 5.4.2.1)催化成为2,3-二磷酸甘油酸,再由二磷酸甘油酸磷酸酶(biphosphoglycerate phosphatase, PGP, EC: 3.1.3.13)催化生成。但在重建这一步骤的过程中没有发现PM和PGP对应的单一序列,推测是由于转录量偏低而没有被检测到,或是在缺刻缘绿藻中只有途径(1)。之后3-磷酸甘油酸经磷酸甘油酸变位酶(phosphoglycerate mutase, PGM, EC:5.4.2.1)形成2-磷酸甘油酸,再由烯醇化酶(Enolase, ENO, EC: 4.2.1.11)催化生成磷酸烯醇式丙酮酸,再由丙酮酸激酶(Pyruvate kinase, PK, EC: 2.7.1.40)催化生成丙酮酸。

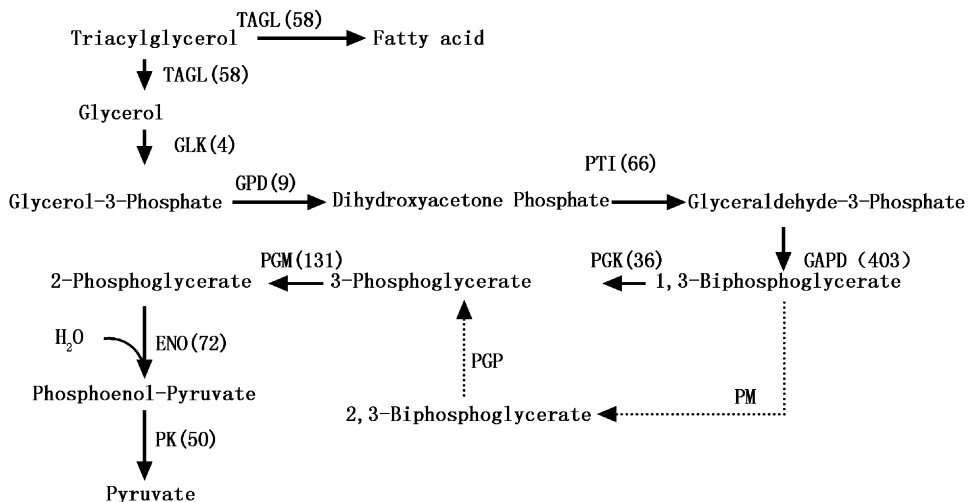


图4 基于缺刻缘绿藻转录组注释数据构建的三酰甘油降解途径

Fig.4 Degradation of TAG pathway reconstructed based on the *de novo* assembly and annotation of *M. incise* transcriptome
括号内的数字为该种酶基因的转录量数据。

脂肪酸的降解(图5)是以酰基-辅酶A的形式完成的,由酰基-辅酶A氧化酶(Acyl-CoA oxidase, ACOX, EC: 1.3.3.6)、烯酰-辅酶A水解酶(Enoyl-CoA hydratase, ECHS, EC: 4.2.1.17)、3-羟酰基-辅酶A脱氢酶(3-hydroxyacyl-CoA dehydrogenase, HCD, EC: 1.1.1.35, 1.1.1.211)和乙酰-辅酶A酰基转移酶(Acetyl-CoA acyltransferase, ACAA, EC: 2.3.1.16)4种酶共同催化完成一轮反应,自C16酰基-辅酶A开始,第一轮反应完成后脱下一个乙酰-辅酶A,生成C14酰基-辅酶A,这样经7轮相同的反应,C16酰基-辅酶A被分解成8个乙酰-辅酶A。

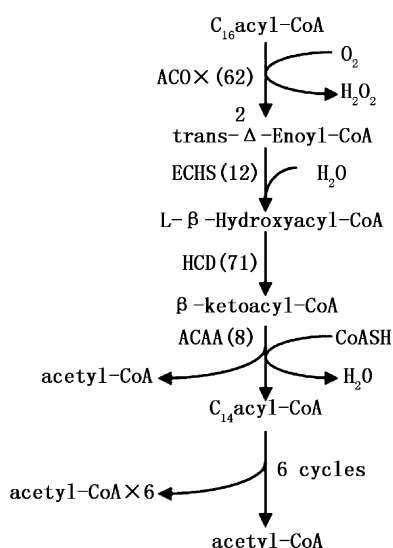


图5 基于缺刻缘绿藻转录组注释数据构建的脂肪酸降解途径

Fig. 5 Degradation pathway of fatty acids reconstructed based on the *de novo* assembly and annotation of *M. incise* transcriptome
括号内的数字为该种酶基因的转录量数据。

3 结论

通过对缺刻缘绿藻的转录组进行焦磷酸测序,并对产生的读序进行拼接。与同类研究相比,测序数据和拼接结果质量较高。之后对得到的单一序列进行功能注释和生物功能组分分类。基于注释信息构建了缺刻缘绿藻脂质代谢的完整途径:脂肪酸是在叶绿体内从头合成,然后游离脂肪酸进入胞质,由内质网进行三酰甘油的合成,最后可能在油体蛋白作用下形成油滴并储存于细胞中。ArA自油酸开始经过多个去饱和酶及

延长酶的作用而产生的。油酸是由硬脂酰-ACP去饱和酶作用而形成,而棕榈油酸是在叶绿体的糖脂上被去饱和而产生的。

本研究为缺刻缘绿藻脂质代谢途径的研究提供了一个新的思路和手段,为缺刻缘绿藻完整的脂质合成途径特别是多不饱和脂肪酸的合成研究提供了数据基础。由于缺刻缘绿藻在缺氮的条件下可以极大提高藻体内ArA的含量^[4],因此本研究还为深入探讨缺刻缘绿藻在氮胁迫条件下的脂质积累途径以及该藻对于氮胁迫信号的响应机理提供了理论基础。

参考文献:

- [1] 于长青,李丽娜.花生四烯酸研究进展[J].农产品加工,2007(4):10-12.
- [2] BELL J G, FARNDAL B M, BRUCE M P, et al. Effects of broodstock dietary lipid on fatty acid compositions of eggs from sea bass (*Dicentrarchus labrax*) [J]. Aquaculture, 1997, 149(1/2): 107-119.
- [3] REISIGL H. Systematik und ökologie alpiner Bodenalgae [J]. Österreichische Botanische Zeitschrift, 1964, 111(4): 402-499.
- [4] 童牧,于水燕,欧阳珑玲,等.氮饥饿与磷饥饿促使缺刻缘绿藻花生四烯酸含量增加的比较研究[J].水产学报,2011,35(5): 763-773.
- [5] 李春阳,杜道海,于水燕,等.缺刻缘绿藻 ω 3脂肪酸去饱和酶基因的特性及在氮饥饿过程中相对转录量的分析[J].水产学报,2010,34(9): 1343-1353.
- [6] YU S Y, LIU S C, LI C Y, et al. Submesoscale characteristics and transcription of a fatty acid elongase gene from a freshwater green microalga, *Myrmecea incisa* Reisingl [J]. Chinese Journal of Oceanology and Limnology, 2011, 29(1): 87-95.
- [7] YU S Y, LI H, TONG M, et al. Identification of a D6 fatty acid elongase gene for arachidonic acid biosynthesis localized to the endoplasmic reticulum in the green microalga *Myrmecea incisa* Reisingl [J]. Gene, 2012, 493(2): 219-227.
- [8] HUSE S M, HUBER J A, MORRISON H G, et al. Accuracy and quality of massively parallel DNA pyrosequencing [J]. Genome Biology, 2007, 8(7): R143.
- [9] GILLES A, MEGLÉCZ E, PECH N, et al. Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing [J]. BMC Genomics, 2011, 12: 245.
- [10] STANIER R Y, KUNISAWA M M, COHEN-BAZIR G. Purification and properties of unicellular blue-green algae (order Chlorococcales.) [J]. The Bacteriological Review, 1971, 35(2): 171-201.
- [11] 罗梦柳,桑敏,张成武,等.氮、磷对缺刻缘绿藻生长、总脂及花生四烯酸积累的影响[J].天然产物研究与开发,

- 2010,22(3): 378–382.
- [12] 杜道海,周志刚. 缺刻缘绿藻总 RNA 提取方法的比较研究[J]. 华北农学报,2008,23(s): 103–106.
- [13] NG P, WEI C L, SUNG W K, et al. Gene identification signature (GIS) analysis for transcriptome characterization and genome annotation [J]. *Nature Methods*, 2005, 2(2): 105–111.
- [14] HUANG X, MADAN A. CAP3: A DNA sequence assembly program [J]. *Genome Research*, 1999, 9(9): 868–877.
- [15] RICE P, LONGDEN I, BLEASBY A. EMBOSS: the European molecular biology open software suite [J]. *Trends in Genetics*, 2000, 16(6): 276–277.
- [16] CHEN Z Z, XUE C H, ZHU S, et al. GoPipe: streamlined gene ontology annotation for batch anonymous sequences with statistics [J]. *Progress in Biochemistry and Biophysics*, 2005,32(2): 187–190.
- [17] KANEHISA M, GOTO S, FURUMICHI M, et al. KEGG for representation and analysis of molecular networks involving diseases and drugs [J]. *Nucleic Acids Research*, 2010, 38: D355–D360.
- [18] SAPIEL G, QUINET M, HEIJDE M, et al. Genome-wide transcriptome analyses of silicon metabolism in *Phaeodactylum tricornutum* reveal the multilevel regulation of silicic acid transporters [J]. *PLoS ONE*, 2009, 4(10): e7458.
- [19] RISMANI-YAZDI H, HAZNEDAROGLU B Z, BIBBY K, et al. Transcriptome sequencing and annotation of the microalgae *Dunaliella tertiolecta*: Pathway description and gene discovery for production of next-generation biofuels [J]. *BMC Genomics*, 2011, 12: 148.
- [20] LOWE C D, MELLO L V, SAMATAR N, et al. The transcriptome of the novel dinoflagellate *Oxyrrhis marina* (Alveolata: Dinophyceae): response to salinity examined by 454 sequencing [J]. *BMC Genomics*,2011, 12: 519.
- [21] LIANG C, LIU Y, LIU L, et al. Expressed sequence tags with cDNA termini: previously overlooked resources for gene annotation and transcriptome exploration in *Chlamydomonas reinhardtii* [J]. *Genetics*, 2008, 179(1): 83–93.
- [22] MILLER R, WU G, DESHPANDE R R, et al. Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism [J]. *Plant Physiology*, 2010, 154(4): 1737–1752.
- [23] WISECAVER J H, HACKETT J D. Transcriptome analysis reveals nuclear-encoded proteins for the maintenance of temporary plastids in the dinoflagellate *Dinophysisacis minate* [J]. *BMC Genomics*, 2011, 11: 366.
- [24] YANG H, MAO Y X, KONG F N, et al. Profiling of the transcriptome of *Porphyrta yezeensis* with Solexa sequencing technology [J]. *Chinese Science Bulletin*, 2011, 56(20): 2119–2130.
- [25] GRIFFITHS G, STOBART A. Fatty acid biosynthesis and triacylglycerol assembly. *Handbook of Food Enzymology* [M]. CRC Press, USA,2002.
- [26] HARWOOD J L. Fatty acid metabolism [J]. *Annual Review of Plant physiology and Molecular Biology*, 1988, 39: 101–138.
- [27] ANDERSSON M X, STRIDH M H, LARSSON K E, et al. Phosphate-deficient oat replaces a major portion of the plasma membrane phospholipids with the galactolipid digalactosyldiacylglycerol [J]. *FEBS Letters*, 2003, 537(1/3): 128–132.
- [28] HÄRTEL H, DÖRMANN P, BENNING C. DGD1-independent biosynthesis of extraplastidic galactolipids following phosphate deprivation in *Arabidopsis* [J]. *Proceedings of the National Academy of Sciences USA*, 2000, 97(19): 10649–10654.
- [29] OUYANG L L, LI H, LIU F, et al. Accumulation of AA in a green microalga, *Myrmecea incisa*, enhanced by nitrogen starvation and its molecular regulation mechanisms [M]// DUMANCAS G G. *Arachidonic Acid: dietary sources and general functions*. New York: Nova science publishers, Inc, 2012.

The transcriptome pyrosequencing and gene function annotation of the green microalga *Myrmecea incisa*

CHEN Si-hong, ZHOU Zhi-gang

(College of Fisheries and Life Sciences, Shanghai Ocean University, Shanghai 201306, China)

Abstract: In order to understand the metabolic pathway of arachidonic acid and other lipids in *Myrmecea incisa*, the transcriptome pyrosequencing of this microalga was conducted by use of the sequencer Roche 454 GS FLX. Totally 393 722 reads (minimal size > 29 bp) averaging 333 bp were generated from one consecutive pyrosequencing run. Cleaning of the raw sequences resulted in a total of 382 468 high quality reads with an average length of 322 nucleotides totalling 123 Mb. After clustering and assembly, these reads were assembled into 22 714 contigs and 25 621 singletons. The average length for contigs and singletons were 639 bp and 277 bp, respectively. By annotating the unisequences, the metabolic pathways of lipids were constructed. Fatty acid was *de novo* synthesized in chloroplasts, and free fatty acids were transported into cytosol where triacylglycerol was synthesized by endoplasmic reticulum. Oil bodies were formed possibly with the help of caleosins. Arachidonic acid was synthesized by desaturation for several times and elongation from oleic acid. Oleic acid was formed by stearyl-ACP desaturase, whereas palmitic acid bound with glucolipid was generated by $\Delta 7$ desaturase. This research lays a foundation for systematic investigation into the manipulation of lipid metabolism and gene modification for higher production of ArA in *M. incisa*.

Key words: *Myrmecea incisa*; transcriptome; pyrosequencing; lipid metabolism