

缺刻缘绿藻碳酸酐酶(CA)基因的序列克隆及特征分析

梅守华, 毕燕会, 周志刚*

(上海海洋大学水产与生命学院, 上海 201306)

摘要: 从缺刻缘绿藻的转录组数据库中搜索到 5 条编码该藻碳酸酐酶(carbonic anhydrase, CA)的 contig 序列, 据此序列设计基因特异性引物, 利用 cDNA 末端快速扩增(rapid amplification of cDNA ends, RACE)技术, 克隆得到 cDNA 全长序列, 分别命名为 Mi α CA1、Mi α CA2、Mi β CA1、Mi β CA2 和 Mi γ CA, 开放阅读框(ORF)长分别为 963、1 089、1 041、738 和 687 bp, 相应编码由 320、362、346、245 和 228 个氨基酸组成的蛋白。这些蛋白均富含疏水性氨基酸, 分别占氨基酸总量的 41.25%、45.31%、43.35%、42.45% 和 43.42%。基于缺刻缘绿藻和其他物种 CA 的蛋白序列所构建的 Neighbor-Joining 系统演化树显示, 这些 CA 很明显地被聚类成 α -、 β -和 γ -CA 等 3 支。缺刻缘绿藻的 2 个 α -CA 都存在与 Zn^{2+} 结合的 3 个 His 残基, 2 个 β -CA 也具有与 Zn^{2+} 结合的 2 个 Cys 残基和 1 个 His 残基, 但 Mi γ CA 中的 Zn^{2+} 结合位点分别为 Arg、His 和 Asn, 不同于报道中的 3 个 His。Mi α CA1 与莱茵衣藻的 CAH3 亲缘关系较近, 因具有 2 个信号肽, 它可能位于叶绿体的类囊体腔中并发挥作用。Mi α CA2 与莱茵衣藻的 CAH1 亲缘关系较近, 因具有 1 个信号肽, 可能在细胞的周质空间起着与 CAH1 类似的功能。Mi β CA1 和 Mi β CA2 与莱茵衣藻的 CAH7 和 CAH8 亲缘关系更近, 可能在细胞质中参与 CO_2 和 HCO_3^- 之间的转化。Mi γ CA 则与高等植物的 γ -CA 聚在一起, 可能位于线粒体内发挥作用。由此推测, 自缺刻缘绿藻所克隆的 5 个 CA 基因应在细胞的不同部位协同作用, 通过参与 CO_2 和 HCO_3^- 之间的转化以调节 pH 并实现 CO_2 在细胞内转运的目的。

关键词: 缺刻缘绿藻; 莱茵衣藻; 碳酸酐酶; 克隆; CO_2

中图分类号: Q 785; S 917.3

文献标志码: A

碳酸酐酶(carbonic anhydrase, CA, EC4.2.1.1)是一种含 Zn^{2+} (极少以钴离子或镉离子取代锌离子)的金属酶, 能够催化 CO_2 的可逆水合反应, 实现 CO_2 和 HCO_3^- 之间的快速转化^[1]。CA 最初是在哺乳动物的红细胞中被发现^[2], 随后在高等植物^[3]、藻类^[4]、蓝细菌^[5]和微生物^[6]中也相继发现 CA 的存在。已知的 CA 至少可分为 α -、 β -和 γ -CA^[3]等 3 个亚型。

CA 在藻类的光合作用中起着非常重要的作用^[7]。水域环境中, CO_2 的溶解度低, 且扩散速率慢^[8]。为了能在这样低 CO_2 浓度的水环境下生存并保持较高的光合效率, 藻类细胞发展了一

种 CO_2 浓缩机制(carbon concentrating mechanism, CCM), 从而提高核酮糖-1,5-二磷酸羧化酶/加氧酶(Rubisco)周围的 CO_2 浓度, 加强光合固碳的能力^[9-10]。在该机制中, CA 扮演着重要的角色^[8]。在莱茵衣藻(*Chlamydomonas reinhardtii*)这种模式生物中, 目前已发现 12 个编码 CA 的基因, 即 3 个 α -CA^[11-13]、6 个 β -CA^[14-16]和 3 个 γ -CA^[17]。这些 CA 分布在莱茵衣藻的细胞质、叶绿体、线粒体等处, 协同调节莱茵衣藻胞内及各相关细胞器的 pH 并向叶绿体基质中的 Rubisco 处输送 CO_2 ^[18], 增加 Rubisco 羧化活性位点处的 CO_2 浓度, 提高 Rubisco 的羧化能

收稿日期:2014-10-08 修回日期:2015-02-12

资助项目:国家海洋局可再生能源专项基金(SHME2011SW02);国家自然科学基金(31172389);上海高校海洋科学一流学科建设

项目
通信作者:周志刚, E-mail:zgzhou@shou.edu.cn

力,确保能固定更多的无机碳以满足细胞分裂、快速生长以及各种代谢活动等所需的物质与能量。

缺刻缘绿藻(*Myrmecea incisa* Reisi) 隶属绿藻门(Chlorophyta)、共球藻纲(Trebouxiophyceae),是一种淡水单细胞球形藻类,细胞常聚集在一起以形成非定型群体的细胞团^[19]。该藻富含长链多不饱和脂肪酸花生四烯酸(arachidonic acid, ArA),其含量在氮饥饿胁迫条件下可达藻体干重的7%^[20]。但在正常培养或自然生长条件下,因生长速度慢致使ArA的生产效率不高,因此,如何增加该藻的生长速率以提高ArA的生产效率是一个有待解决的问题。一般来说,藻类的生长速率与其碳同化效率成正相关,因为后者能够为前者提供所需的碳源与能量^[21]。藻类多生长在水域环境中,其同化的CO₂需在CA协助下才能有效地完成。因而,关于CA的探讨可为提高缺刻缘绿藻固定无机碳速率的研究奠定良好的理论基础。缺刻缘绿藻与莱茵衣藻同属绿藻门,亲缘关系较近,且莱茵衣藻的CA研究相对比较透彻。因此,本实验计划在缺刻缘绿藻高通量的转录组数据库^[22-23]中筛选到5个编码CA的contig的基础上,经cDNA末端快速扩增(RACE)技术克隆得到这5个CA基因的全长cDNA序列;然后对比莱茵衣藻的CA研究结果,通过生物信息学分析、蛋白序列比对、系统进化树的构建及同源建模等来分析缺刻缘绿藻CA的细胞定位与功能。

1 材料与方法

1.1 藻种与培养

缺刻缘绿藻 H4301 来源于捷克布拉格查理斯大学藻类收藏中心(Culture Collection of Algae of Charles University of Prague, CAUP),由暨南大学张成武教授馈赠。

将藻细胞接种于BG11液体培养基^[24]中,在25℃和115 μmol/(m²·s)的光照培养箱中培养,光周期为14 h:10 h(光照:黑暗)^[20],每天不定期地摇晃。待长至指数生长期,4℃下5 500 r/min离心10 min,收集藻细胞,用灭菌去离子水洗3次并离心收集藻体,液氮速冻后-80℃保存备用。

1.2 缺刻缘绿藻5个CA基因cDNA全长序列的克隆

采用TRIzol(Invitrogen公司)^[25]并按该产品说明取适量缺刻缘绿藻以提取其总RNA。使用

PrimeScript™ RT反转录试剂盒(TaKaRa公司),按其操作步骤合成cDNA第一链,作为后续PCR反应的模板。

在本实验室对缺刻缘绿藻进行高通量测序所获得的转录组数据^[22-23]中,搜索到5条编号分别为17538、2868、5885、374和1099的contig序列,用Primer Premier 5软件设计5对引物分别扩增这5个基因的cDNA片段(表1)。PCR反应体系为25 μL,包含10.5 μL灭菌双蒸水、12.5 μL的2×pfu PCR Master Mix(天根生化科技有限公司)、1 μL模板cDNA、上下游引物(10 μmol/L)各0.5 μL。反应程序为95℃预变性5 min;94℃变性45 s,在各自退火温度(表1)退火1 min,72℃延伸2 min,35个循环;最后72℃延伸10 min。

利用1%琼脂糖凝胶电泳检测PCR扩增反应的产物,并按照琼脂糖凝胶纯化试剂盒(北京艾德来生物科技公司)的说明回收PCR产物。将回收的产物与pMD19-T载体(TaKaRa公司)于16℃连接过夜,再转化到大肠杆菌(*Escherichia coli*)DH5α感受态细胞(天根生化科技有限公司)中,涂平板,将蓝白斑筛选后获得的阳性克隆送到上海生工生物工程有限公司进行测序。

利用RACE技术,根据上述测序结果设计这些CA基因的5'-RACE和3'-RACE特异性引物(GSP)(表1)。按照SMART™ RACE cDNA扩增试剂盒(Clontech公司)说明,首先进行反转录合成5'-RACE和3'-RACE的cDNA第一链,然后将获得的PCR产物作为反应模板,按照上述的PCR反应体系和程序进行CA基因的5'-和3'-末端序列扩增。然后按照上述方法进行克隆和测序,并将所得结果与已知contig序列比对,再利用DNAMAN软件(<http://www.lynnon.com/>)进行拼接得到cDNA全长序列。

1.3 缺刻缘绿藻CA基因序列的生物信息学及聚类分析

利用NCBI中的BLAST(<http://www.ncbi.nlm.nih.gov/BLAST/>)对获得的缺刻缘绿藻CA基因的cDNA全长序列进行核酸、蛋白质、氨基酸序列同源性搜索,并利用expasy中工具ProtParam(<http://www.expasy.org/tool/protparam.html>)在线分析蛋白质氨基酸组成、分子量、等电点。利用BioEdit软件^[26]分析其疏水性。利用软件TMHMM(<http://www.cbs.dtu.dk/services/>)

<http://www.sexuebao.cn>

TMHMM/)对蛋白质的跨膜区进行预测。使用 SignalP 4.1 Server (<http://www.cbs.dtu.dk/services/signalP/>)对其信号肽进行预测,通过 <http://www.sbg.bio.ic.ac.uk/phyre/>在线预测 CA 基因的蛋白质二级和三级结构。利用 MEGA 5.05 软件^[27]构建系统进化树。利用 TargetP 1.1

Server (<http://www.cbs.dtu.dk/services/Targetp/>)、Predotar (<http://urgi.versailles.inra.fr/prodotar/predotar.html>)和 Protein Prowler v. 1.2 (http://www.bioinf.scmb.uq.edu.au/pprowler_webapp_1-2/)对蛋白进行亚细胞定位预测。

表 1 实验所使用的引物及其序列
Tab.1 The sequence of primers used in the present research

引物 primer	序列(5'-3') sequence(5'-3')	退火温度/°C annealing temperature	产物大小/bp size of amplified product
cDNA 克隆			
CA1-U	TCCAGTGCAGTGCTCAGGG	64	1 182
CA1-D	AGCAGAGTGTCAGACTGTTAGA	58.2	
CA2-U	TGCAGCTGATTGCGCTATTTG	58	1 384
CA2-D	GCACTACTTGGCAGGGCTTGT	61.9	
CA3-U	AGGCCCGTTCTATCTACGTACC	61.9	1 420
CA3-D	AGGCAATGTGTGCTGTAAATGG	58.2	
CA4-U	CCTCTCTGCCCCACAAAGTCT	61.9	1 146
CA4-D	TGACACGCTCACAAACCCAG	61.9	
CA5-U	GTCCTGTGACCCCAGAGCA	61.9	1 020
CA5-D	ACCTTGCTCAGAGGCGTGTCC	63.9	
3'-RACE			
CA1-3-GSP1	GAAAGTATCCCCTCGTCTGTGGACATCA	64.9	-
CA1-3-GSP2	GAGCAGGTCGACTGGTTTGTCTTCACG	66.6	-
CA2-3-GSP1	TTGAGTTCGGGGCACATAGAGGAGAGC	66.6	-
CA2-3-GSP2	GATGACGTCCAGTTTGTGCTTCTTTAG	63.5	-
CA3-3-GSP1	GACCTGGACACCTTTGAAGCGACCGTT	66.6	-
CA3-3-GSP2	CTTCTCCAGAGCATGAGCACCCACCT	66.6	-
CA4-3-GSP1	TCAAGAAGCACACCAACTTCCAGGAGT	63.5	-
CA4-3-GSP2	TAGCACATAGCAAGTAGCGGCAGGATG	65.0	-
CA5-3-GSP1	AGAAGTCCTTCATGGTCACCTCGGCT	65.1	-
CA5-3-GSP2	AGTACCCCCACGTTGACAGGTCTGAC	66.7	-
cDNA 全长序列验证(下游)			
CA1-O-D	CGTGACCACAAACCAGCATTCT	60.1	1 624
CA2-O-D	CAGCACAACCACCACTGGCAAC	61.9	2 015
CA3-O-D	TTTTTGCGGTCAGGAATCAC	58.4	1 728
CA4-O-D	AGCAGTACCCAATCACGGCAT	60	1 427
CA5-O-D	GGCTACCTGTCATTGCCTTTC	60	1 468

2 结果

2.1 缺刻缘绿藻 5 个 CA 基因的 cDNA 全长序列

在缺刻缘绿藻高通量测序的转录组数据库中

搜索到 5 条长度分别为 1 269、1 906、1 702、1 557 和 1 202 bp 的 contig 序列(对应编号为 17538、2868、5885、374 和 1099),经 BLASTp 搜索,发现 1 269 与 1 906 bp 的 contig 序列与莱茵衣藻中已知的 α -CA 分别有 58% 和 35% 的相似性;1 702

<http://www.sexuebao.cn>

与1 557 bp的 contig 序列与莱茵衣藻 β -CA 分别有 60% 和 50% 的相似性;1 202 bp 的 contig 序列与拟南芥 γ -CA 有 58% 的相似性。因而,我们认为这 5 个 contig 可能都是编码缺刻缘绿藻不同 CA 的基因片段。根据这 5 条 contig 序列设计特异引物 GSP(表 1),应用 RACE 技术克隆缺刻缘绿藻 CA 基因的 3'-末端序列(因每个 contig 经比对都含有 5'-非翻译区序列)(图 1),经 TA 克隆、测序,将这些末端片段的序列与已知序列拼接得到这 5 个 CA 基因的 cDNA 全长。然后在每个基因 cDNA 序列的基础上,设计下游引物(表 1)并结合表 1 中不同 CA 基因 cDNA 克隆所使用的相对应上游引物进行序列验证。

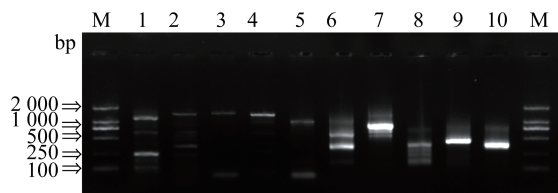


图1 缺刻缘绿藻 CA 基因全长 cDNA 序列克隆的电泳图

M. 2000 bp DNA 分子量标准;泳道 1、2、3、4、5:分别为从 cDNA 中扩增出的 $Mi\alpha CA1$ 、 $Mi\alpha CA2$ 、 $Mi\beta CA1$ 、 $Mi\beta CA2$ 、 $Mi\gamma CA$ 基因片段;泳道 6、7、8、9、10:分别对应着 $Mi\alpha CA1$ 、 $Mi\alpha CA2$ 、 $Mi\beta CA1$ 、 $Mi\beta CA2$ 、 $Mi\gamma CA$ 的 3'-RACE 末端扩增产物

Fig. 1 Electrophoresis patterns of full-length cDNA sequence of CA genes cloned from *M. incisa*

M. 2000 bp marker; Line 1 to 5: electrophoresis patterns of amplified product of $Mi\alpha CA1$, $Mi\alpha CA2$, $Mi\beta CA1$, $Mi\beta CA2$ and $Mi\gamma CA$. Line 6 to 10: amplified products of 3'-RACE of $Mi\alpha CA1$, $Mi\alpha CA2$, $Mi\beta CA1$, $Mi\beta CA2$ and $Mi\gamma CA$

根据 contig 17538 的序列克隆到编码缺刻缘绿藻第 1 个 CA 基因的 cDNA 长 1 624 bp,它的开放式阅读框(ORF)长 963 bp,5'-非翻译区(UTR)长 20 bp,3'-UTR 长 641 bp。经 BLASTp 搜索,发现该基因序列与莱茵衣藻的 CAH3(1 个 α -CA)有 80% 的同源性,因而,将其命名为 $Mi\alpha CA1$ (GenBank 登录号 KP036913)。

基于 contig 2868 的序列克隆到编码该藻第 2 个 CA 基因的 cDNA 长 2 015 bp,其中 ORF 长 1 089 bp,5'-UTR 长 23 bp,3'-UTR 长 903 bp。经 BLASTp 搜索,发现该基因序列与莱茵衣藻的 CAH1(1 个 α -CA)有 66% 的同源性,因而,将其命名为 $Mi\alpha CA2$ (GenBank 登录号 KP036914)。

根据 contig 5885 的序列克隆到编码该藻第 3 个 CA 基因的 cDNA 长 1 757 bp,它的 ORF 长 1 041 bp,5'-UTR 长 258 bp,3'-UTR 长 458 bp。经 BLASTp 搜索,发现该基因序列与莱茵衣藻的 CAH7(1 个 β -CA)有 69% 的同源性,据此,将其命名为 $Mi\beta CA1$ (GenBank 登录号 KP036915)。

基于 contig 374 的序列克隆到编码该藻第 4 个 CA 基因的 cDNA 长 1 427 bp,其中 ORF 长 738 bp,5'-UTR 长 99 bp,3'-UTR 长 590 bp。经 BLASTp 搜索,发现该基因序列与莱茵衣藻的 CAH8(1 个 β -CA)有 85% 的同源性,因而,将其命名为 $Mi\beta CA2$ (GenBank 登录号 KP036916)。

根据 contig 1099 克隆到编码该藻第 5 个 CA 基因的 cDNA 长 1 468 bp,它的 ORF 长 687 bp,5'-UTR 长 215 bp,3'-UTR 长 566 bp。经 BLASTp 搜索,发现此基因与拟南芥(*Arabidopsis thaliana*)的 γ -CA1 有 94% 的同源性,据此,将其命名为 $Mi\gamma CA$ (GenBank 登录号 KP036917)。

实验结果显示,自缺刻缘绿藻所克隆到的上述 5 个基因全长 cDNA 中,其 3'-UTR 均具有明显的 poly A 尾巴,说明它们均是由核基因组编码的。将这些序列进行比对,显示它们是不同的类型 CA 的基因序列。

2.2 缺刻缘绿藻 CA 基因的特性分析

预测缺刻缘绿藻 $Mi\alpha CA1$ 、 $Mi\alpha CA2$ 基因编码蛋白分别由 320、362 个氨基酸组成,其分子量分别为 34.21、38.10 ku,等电点分别为 9.18、5.18。将 6 个 α -CA 基因的编码蛋白进行比对发现(图 2-a),缺刻缘绿藻 $Mi\alpha CA1$ 、 $Mi\alpha CA2$ 与其他 α -CA 蛋白在 C 端具有 1 个共同的保守区,在该区域内都存在 1 个 Cys 残基($Mi\alpha CA1$ 中的 Cys266, $Mi\alpha CA2$ 中的 Cys262),经二级和三级结构预测,它与该蛋白质中的其他位置的 Cys 残基能形成二硫键(-S-S-),以稳固蛋白质的空间结构^[28-29]。另外, α -CA 中 3 个 His 残基($Mi\alpha CA1$ 中的 His172、His174 和 His191; $Mi\alpha CA2$ 中的 His157、His159 和 His176)也相当保守(图 2-a),推测它们与 Zn^{2+} 结合以构成该酶的催化活性中心^[30]。

预测缺刻缘绿藻 $Mi\beta CA1$ 、 $Mi\beta CA2$ 基因编码蛋白分别由 346、245 个氨基酸组成,其分子量分别为 37.26、27.34 ku,等电点分别为 5.64、

6.19. 将6个 β -CA基因的编码蛋白进行比对(图2-b),发现缺刻缘绿藻与其他 β -CA蛋白^[31]一样,其Mi β CA1、Mi β CA2也具有与Zn²⁺结合的2个保守Cys残基和1个His残基(在Mi β CA1上,它们位于Cys50、His106和Cys109处;而在Mi β CA2上,它们位于Cys49、His105和Cys108

处)。这2个Cys残基分别位于不同的保守区,其中后1个Cys残基和His残基位于同一保守区(图2-b)。经二级和三级结构预测显示,Mi β CA1和Mi β CA2的单体与已报道的 β -CA结构类似,它们的催化中心由4个 β 折叠片层构成,这些片层在空间结构上以2-1-3-4顺序平行排列^[32]。

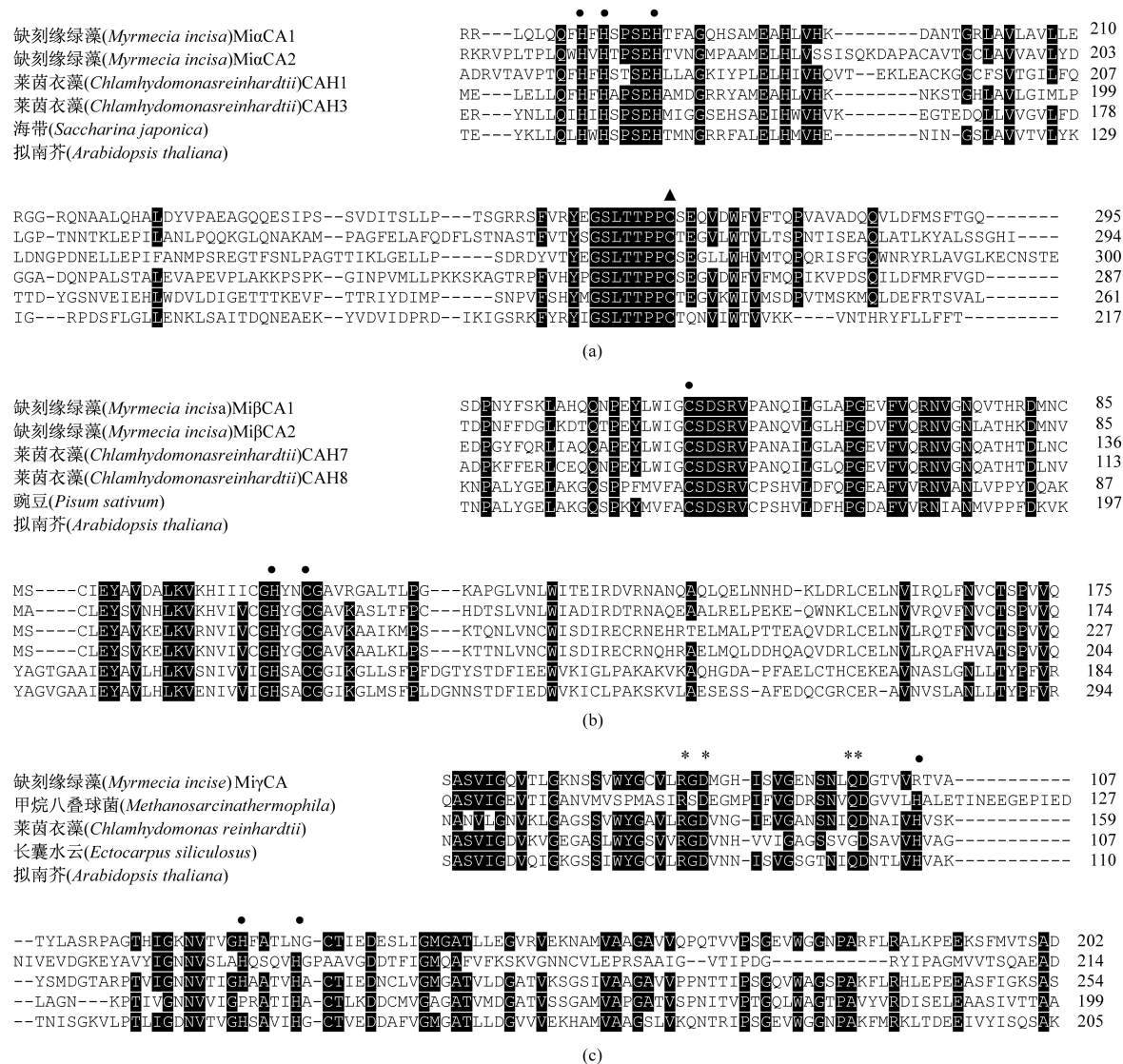


图2 缺刻缘绿藻CA基因编码蛋白的同源序列比对

(a) α -CA基因;(b) β -CA基因;(c) γ -CA基因。同源性70%以上的氨基酸用阴影表示,锌结合位点用“•”表示,“▲”表示 α -CA中保守的Cys,“*”表示 γ -CA中与催化活性有关的残基

Fig. 2 Alignment of deduced amino acid sequences of CA genes in *M. incisa*

(a) Alignment of deduced amino acid sequences of α -CA. (b) Alignment of deduced amino acid sequences of β -CA. (c) Alignment of deduced amino acid sequences of γ -CA. Amino acids of high identities more than 70% are indicated by shaded letters, the black dots show the Zn-binding sites, the triangle shows the conserved Cys residue of α -CA, the pentagrams show the important residues of γ -CA

缺刻缘绿藻 Mi γ CA 基因编码由 228 个氨基酸组成的蛋白,它的分子量为 24.48 ku,等电点为

8.18。将5个 γ -CA基因的编码蛋白进行比对(图2-c)发现,Mi γ CA与其他的 γ -CA具有一些

共同的保守序列。二级结构预测显示, $Mi\gamma CA$ 的 3 个 Zn^{2+} 结合位点分别为 Arg104、His127 和 Asn132(图 3)。这些位点与在甲烷八叠球菌 (*Methanosarcina thermophila*) 和拟南芥中已报道的 3 个 Zn^{2+} 结合位点均为 His 虽不同,但 $Mi\gamma CA$ 的单体折叠(图 3)与甲烷八叠球菌 γ -CA 的三级结构类似,都存在一个左手平行的 β -螺旋结构(left-handed parallel β -helix, $L\beta H$)^[33],且一些与催化活性有关的残基如 Arg83、Asp85、Glu98 和 Asp99 也很保守^[34]。

缺刻缘绿藻 CA 的氨基酸组成分析结果显示: $Mi\alpha CA1$ 、 $Mi\alpha CA2$ 、 $Mi\beta CA1$ 、 $Mi\beta CA2$ 和 $Mi\gamma CA$ 基因所编码的蛋白均富含疏水性氨基酸(Val、Leu、Ala、Ile、Phe、Trp、Pro、Met),分别占总氨基酸含量的 41.25%、45.31%、43.35%、42.45% 和 43.42%。表明这些蛋白具有较强的疏水性,有利于蛋白内部折叠形成二级结构和三级结构,保证蛋白空间结构的稳定性。

利用软件 TMHMM 对 CA 的跨膜区进行预测,结果显示除了 $Mi\alpha CA2$ 在信号肽区域有一个跨膜区,另外 4 个 CA 均无跨膜区,显示除 $Mi\alpha CA2$ 外其他 4 个 CA 蛋白均不是膜结合蛋白。

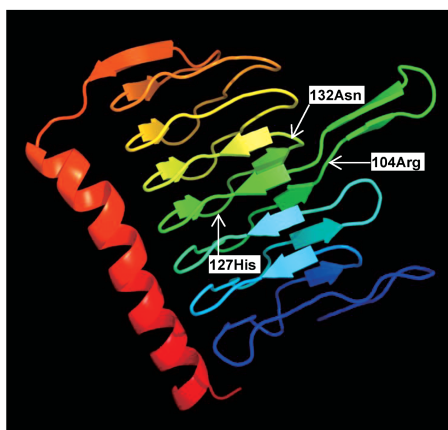


图 3 缺刻缘绿藻 $Mi\gamma CA$ 蛋白质的同源建模

104Arg、127His、132Asn 为锌结合位点

Fig. 3 Tertiary structure simulation of $Mi\gamma CA$ from *Myrmecia incisa*

The Zn-binding sites are 104 Arg, 127 His and 132 Asn

2.3 基于 CA 蛋白氨基酸序列的聚类分析

通过同源搜索,从 NCBI 中获得其他物种的 66 个 CA 基因,基于它们的蛋白序列,利用 MEGA 5.05 软件构建 Neighbor-Joining (NJ) 系统

演化树,结果显示,这些 CA 基因显著地(Bootstrap 值都为 100)被分成 α -CA、 β -CA、 γ -CA 等 3 支(图 4),这与 Hewett-Emmett 和 Tashian 通过比较 67 个 CA 序列获得的结果一致^[3]。表明它们的祖先虽起源不同,但在独立进化过程中功能趋于一致,即属于趋同进化;它们都在 CO_2 和 HCO_3^- 的可逆转化过程中起着催化作用^[35]。

从 NJ 系统演化树上,还可以了解到 $Mi\alpha CA1$ 、 $Mi\alpha CA2$ 与莱茵衣藻、长囊水云 (*Ectocarpus siliculosus*)、海带 (*Saccharina japonica*)、一些蓝藻以及动物的 CA 蛋白一起以组成 α -CA 支(图 4)。其中, $Mi\alpha CA1$ 与莱茵衣藻 CAH3 的亲缘关系最近;而 $Mi\alpha CA2$ 与多变小球藻 (*Chlorella variabilis*) CA 及莱茵衣藻 CAH1、CAH2 的亲缘关系更近。

缺刻缘绿藻的 $Mi\beta CA1$ 、 $Mi\beta CA2$ 与三角褐指藻 (*Phaeodactylum tricornutum*)、莱茵衣藻、一些蓝藻及一些高等植物聚在一起以组成 β -CA 支,与莱茵衣藻的 CAH7、CAH8 聚在一个分支上,说明它们的亲缘关系较近(图 4)。

$Mi\gamma CA$ 虽与一些细菌、藻类、高等植物等聚在一起以组成 γ -CA 支,但有趣的是 $Mi\gamma CA$ 与高等植物的亲缘关系更近(图 4)。

系统演化的结果也进一步清楚地表明,本研究自缺刻缘绿藻所克隆的 5 个 CA 基因中, $Mi\alpha CA1$ 、 $Mi\alpha CA2$ 属于 α -CA, $Mi\beta CA1$ 、 $Mi\beta CA2$ 属于 β -CA, 而 $Mi\gamma CA$ 则属于 γ -CA(图 4)。因此,这 5 个 CA 属于不同的基因家族。

2.4 CA 蛋白的亚细胞定位预测

应用 SignalP 4.1 预测 CA 蛋白的信号肽,结果显示, $Mi\alpha CA1$ 自 N 端开始到 Arg24 处为一个信号肽,从 Val25 到 Arg44 处为另一个信号肽; $Mi\alpha CA2$ 在 Ser20 之前有 1 个信号肽; $Mi\gamma CA5$ 从 N 端开始到 Ile19 处有 1 个信号肽;而 $Mi\beta CA1$ 和 $Mi\beta CA2$ 则无典型的信号肽酶切位点。

利用 TargetP 1.1 Server、Predotar、Protein Prowler v. 1.2 等 3 个软件对缺刻缘绿藻 CA 蛋白进行亚细胞定位进行在线预测。结果显示, $Mi\alpha CA1$ 应位于叶绿体中(表 2); $Mi\alpha CA2$ 因含有 1 个信号肽和 1 个跨膜区,能够借助分泌途径转运至细胞的相应部位,如细胞周质空间的膜上,以发挥其作用; $Mi\gamma CA$ 位于线粒体中的可能性较大,但得分值不高。而 $Mi\beta CA1$ 、 $Mi\beta CA2$ 的亚细

胞定位虽无法确定,但肯定不在线粒体或叶绿体中,极有可能在细胞质中发挥作用。显然,有关缺

刻缘绿藻 CA 的亚细胞精细定位还需要直接的细胞学证据。

表 2 缺刻缘绿藻 CA 蛋白亚细胞定位的预测结果

Tab.2 Subcellular localization prediction of various CA deduced proteins from *M. incisa* by using TargetP 1.1 Server

	叶绿体转运肽 cTP	线粒体转运肽 mTP	信号肽 SP	其他 other	定位 location
Mi α CA1	0.856	0.511	0.000	0.018	叶绿体 chloroplast
Mi α CA2	0.007	0.114	0.843	0.032	分泌途径 secretory pathway
Mi β CA1	0.152	0.211	0.045	0.761	-
Mi β CA2	0.198	0.071	0.140	0.860	-
Mi γ CA	0.155	0.493	0.005	0.274	线粒体 mitochondrion

3 讨论

真核生物的 CA 一般分为 3 个亚型: α -、 β -和 γ -CA^[3]。 α -CA 存在于所有哺乳动物、部分藻类以及一些细菌界的原核生物中,大部分 α -CA 是以单体的形式发挥作用,它们的 3 个保守 His 残基与 Zn^{2+} 结合(图 2-a)组成活性中心^[35]; β -CA 存在于植物、藻类、细菌及一些原始细菌中,与 α -CA 不同, β -CA 是以二聚体或多聚体的形式发挥作用,2 个保守的 Cys 和 1 个 His 残基能耦合 Zn^{2+} (图 2-b)以组成其活性中心^[31]; γ -CA 的研究较少,目前仅在一些植物、藻类、细菌和古细菌中有发现,已报道的莱茵衣藻含有 3 个 γ -CA 且全部定位于线粒体上^[18]。本实验通过 RACE 技术克隆得到了缺刻缘绿藻的 5 个 CA 的 cDNA 序列,再结合 CA 序列比对和 NJ 进化树,可以确定缺刻缘绿藻至少含有 3 类 5 个 CA 基因。

缺刻缘绿藻与莱茵衣藻这种模式生物同属绿藻门,因此可以借鉴后者来探讨缺刻缘绿藻 CA 的亚细胞定位及功能。在莱茵衣藻的 12 个 CA 基因中, α -CA 有 3 个,即 CAH1、CAH2 和 CAH3。CAH1 和 CAH2 在序列上仅有部分氨基酸的差异,但 CAH1 在低 CO_2 条件下表达,而 CAH2 则在高 CO_2 浓度的情况下表达^[11]。2007 年, Moroney 等^[36]通过免疫胶体金电镜技术证实了 CAH1 定位于细胞的周质空间,帮助莱茵衣藻在低 CO_2 环境下将细胞外的 HCO_3^- 转化为容易透过细胞膜的 CO_2 ,以满足细胞光合作用所需的 CO_2 。CAH3 的 N 端存在着两个信号肽序列,第

一个帮助进入叶绿体,第二个指引蛋白质到达类囊体腔中。在第二个信号肽后面,存在着一段由 Ala-X-Ala 组成的疏水区域,有助于 CAH3 到达类囊体腔^[37-38],与光系统 II 相结合以发挥其作用^[39]。在本研究中,缺刻缘绿藻的 Mi α CA1 与 CAH3 聚在一支上,信号肽预测结果显示, Mi α CA1 也含有 2 个信号肽序列,且在第二个信号肽序列后面找到了 Ala79-Lys80-Ala81 的疏水区域,鉴于此以及亚细胞定位预测结果,我们认为 Mi α CA1 可能定位于叶绿体的类囊体腔中发挥作用。Mi α CA2 与 CAH1 具有较高的相似性,且聚类在一支,结合 Mi α CA2 的信号肽序列预测结果,推测 Mi α CA2 与 CAH1 具有类似的功能,可能定位于周质空间中,帮助缺刻缘绿藻将海水中的 HCO_3^- 转化为 CO_2 ,使后者能容易透过细胞膜进入胞内,以满足藻体光合作用所需的 CO_2 。

Mi β CA1、Mi β CA2 与莱茵衣藻中的 CAH7、CAH8 聚在一起,且在序列上也分别有着 66% 和 85% 的相似性。据 Ynalvez 等^[16]报道,尽管 CAH7 的定位研究一直没有获得理想的结果,但 CAH8 不同于位于细胞壁周围的 CAH1,它更靠近细胞膜的周质空间。基于此,推测 Mi β CA1 和 Mi β CA2 定位于缺刻缘绿藻细胞的细胞膜内测,将转入到胞内的 CO_2 迅速转变成 HCO_3^- 以防止 CO_2 通过膜渗漏至体外;或者定位在叶绿体周围,将胞内的 HCO_3^- 转变成 CO_2 ,快速压入叶绿体。总之, Mi β CA1 和 Mi β CA2 可能通过参与胞质内 HCO_3^- 与 CO_2 的转化以调控其 pH 并达到 CO_2 在胞内转运的目的。

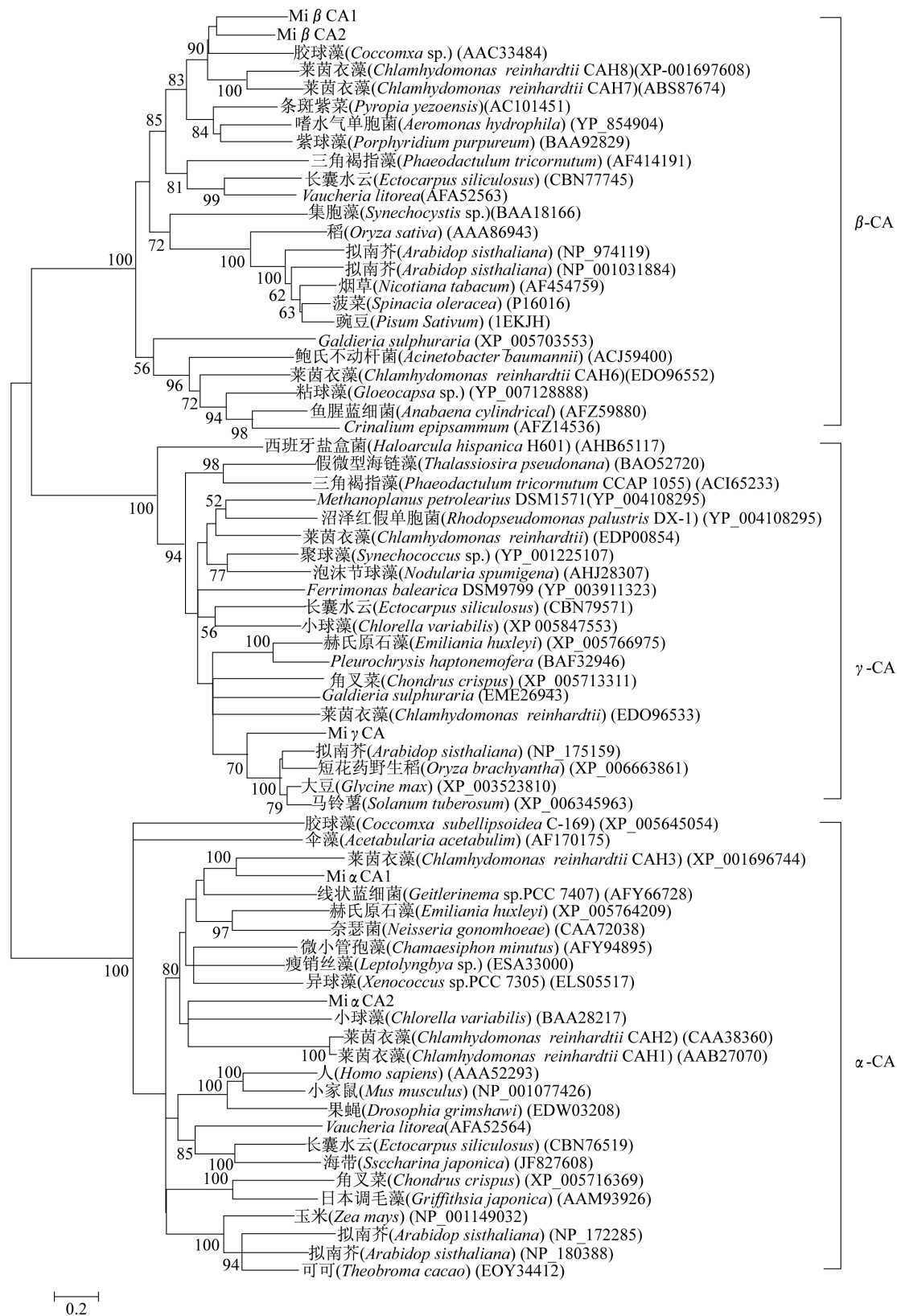


图4 基于CA蛋白氨基酸序列所构建的NJ聚类图

Fig. 4 Neighbor-Joining phylogenetic tree inferred from the deduced amino acid sequences of CA genes from several species

在 NJ 系统演化树中, $Mi\gamma CA$ 与高等植物的 CA 聚类在一起。在高等植物中, 拟南芥存在 5 个 γ -CA, 均定位于线粒体内膜的电子传递链复合体 I (complex I) 上^[40], 可能在转运由呼吸作用所产生的 CO_2 中发挥作用。结合亚细胞定位预测结果, 推测 $Mi\gamma CA$ 定位于线粒体上, 将呼吸作用产生的多余 CO_2 转化成 HCO_3^- , 调节线粒体内的 pH。

参考文献:

- [1] Khalifah R G. The carbon dioxide hydration activity of carbonic anhydrase I. Stop-flow kinetic studies on the native human isoenzymes B and C [J]. *Journal of Biological Chemistry*, 1971, 246(8): 2561 – 2573.
- [2] Meldrum N U, Roughton F J W. Carbonic anhydrase: Its preparation and properties [J]. *Journal of Physiology*, 1933, 80(2): 113 – 142.
- [3] Hewett-Emmett D, Tashian R E. Functional diversity, conservation, and convergence in the evolution of the α -, β -, and γ -carbonic anhydrase gene families [J]. *Molecular Phylogenetics and Evolution*, 1996, 5(1): 50 – 77.
- [4] Tsuzuki M, Miyachi S. The function of carbonic anhydrase in aquatic photosynthesis [J]. *Aquatic Botany*, 1989, 34(1): 85 – 104.
- [5] Fukuzawa H, Suzuki E, Komukai Y, *et al.* A gene homologous to chloroplast carbonic anhydrase (icfA) is essential to photosynthetic carbon dioxide fixation by *Synechococcus* PCC7942 [J]. *Proceedings of the National Academy of Sciences of the United State of America*, 1992, 89(10): 4437 – 4441.
- [6] Nimer N A, Iglesias-Rodriguez M D, Merrett M J. Bicarbonate utilization by marine phytoplankton species [J]. *Journal of Phycology*, 1997, 33(4): 625 – 631.
- [7] Park K Y, Lim B R, Lee K. Growth of microalgae in diluted process water of the animal wastewater treatment plant [J]. *Water Science and Technology*, 2009, 59(11): 2111 – 2116.
- [8] Badger M R, Price G D. The role of carbonic anhydrase in photosynthesis [J]. *Annual Review of Plant Biology*, 1994, 45: 369 – 392.
- [9] Aizawa K, Miyachi S. Carbonic anhydrase and CO_2 concentrating mechanisms in microalgae and cyanobacteria [J]. *FEMS Microbiology Letters*, 1986, 39(3): 215 – 223.
- [10] Graham L E, Wilcox L W. *Algae* [M]. USA: Prentice, 2000: 27 – 36.
- [11] Fujiwara S, Fukuzawa H, Tachiki A, *et al.* Structure and differential expression of two genes encoding carbonic-anhydrase in *Chlamydomonas reinhardtii* [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1990, 87(24): 9779 – 9783.
- [12] Rawat M, Moroney J V. Partial characterization of a new isoenzyme of carbonic-anhydrase isolated from *Chlamydomonas reinhardtii* [J]. *Journal of Biological Chemistry*, 1991, 266(15): 9719 – 9723.
- [13] Karlsson J, Hiltonen T, Husic H D, *et al.* Intracellular carbonic anhydrase of *Chlamydomonas reinhardtii* [J]. *Plant Physiology*, 1995, 109(2): 533 – 539.
- [14] Eriksson M, Karlsson J, Ramazanov Z, *et al.* Discovery of an algal mitochondrial carbonic anhydrase: Molecular cloning and characterization of a low- CO_2 -induced polypeptide in *Chlamydomonas reinhardtii* [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1996, 93(21): 12031 – 12034.
- [15] Mitra M, Lato S M, Ynalvez R A, *et al.* Identification of a new chloroplast carbonic anhydrase in *Chlamydomonas reinhardtii* [J]. *Plant Physiology*, 2004, 135(1): 173 – 182.
- [16] Ynalvez R A, Xiao Y, Ward A S, *et al.* Identification and characterization of two closely related β -carbonic anhydrases from *Chlamydomonas reinhardtii* [J]. *Physiologia Plantarum*, 2008, 133(1): 15 – 26.
- [17] Cardol P, Gonzalez-Halphen D, Reyes-Prieto A, *et al.* The mitochondrial oxidative phosphorylation proteome of *Chlamydomonas reinhardtii* deduced from the genome sequencing project [J]. *Plant Physiology*, 2005, 137(2): 447 – 459.
- [18] Moroney J V, Ma Y, Frey W D, *et al.* The carbonic anhydrase isoforms of *Chlamydomonas reinhardtii*: Intracellular location, expression, and physiological roles [J]. *Photosynthesis Research*, 2011, 109(1 – 3): 133 – 149.
- [19] Reisingl H. Zur Systematik und Okologie alpiner Bodenalgen [J]. *Osterreichische Botanische Zeitschrift*, 1964, 111(4): 402 – 499.
- [20] Tong M, Yu S Y, Ouyang L L, *et al.* Comparison of increased arachidonic acid content in *Mymecia incisa* cultured during the course of nitrogen or phosphorus starvation [J]. *Journal of Fisheries of China*, 2011, 35(5): 763 – 773. [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. 水产学报,2011,35(5):763-773.]
- [21] Teskey R O. A field study of the effects of elevated CO₂ on carbon assimilation, stomatal conductance and leaf and branch growth of *Pinus taeda* trees[J]. Plant, Cell & Environment, 1995, 18(5):565-573.
- [22] Ouyang L L, Li H, Liu F, et al. Accumulation of arachidonic acid in a green microalgae, *Myrmecea incisa* H4301, enhanced by nitrogen starvation and its molecular regulation mechanisms[M]//Dumancas G G, Murdianti B S, Lucas E A, eds. Arachidonic Acid: Dietary Sources and General Functions. New York: Nova Science Publishers, Inc. 2013:1-20.
- [23] Chen S H, Zhou Z G. The transcriptome pyrosequencing and gene function annotation of the green microalga *Myrmecea incisa* [J]. Journal of Shanghai Ocean University, 2012, 21(5):662-670. [陈思弘,周志刚. 缺刻缘绿藻转录组测序及脂质代谢相关基因注释. 上海海洋大学学报, 2012, 21(5):662-670.]
- [24] Stanier R Y, Kunisawa M M, Cohen-Bazir G. Purification and properties of unicellular blue-green algae (order Chroococcales) [J]. Bacteriological Reviews, 1971, 35(2):171-205.
- [25] Du D H, Zhou Z G. Comparison of total RNA extracted from *Myrmecea incisa* using different methods[J]. Acta Agriculturae Boreali-Sinica, 2008, 23(B10):103-106. [杜道海,周志刚. 缺刻缘绿藻总RNA提取方法的比较研究. 华北农学报, 2008, 23(B10):103-106.]
- [26] Hall T A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Widows 95/98/NT [J]. Nucleic Acids Symposium Series, 1999, 41:95-98.
- [27] Tamura K, Peterson D, Peterson N, et al. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods [J]. Molecular Biology and Evolution, 2011, 28(10):2731-2739.
- [28] Premkumar L, Greenblatt H M, Bageshwar U K, et al. Three-dimensional structure of a halotolerant algal carbonic anhydrase predicts halotolerance of a mammalian homolog[J]. Proceedings of the National Academy of Sciences of the United States of America, 2005, 102(21):7493-7498.
- [29] Thannhauser T W, Konishi Y, Scheraga H A. Sensitive quantitative analysis of disulfide bonds in polypeptides and proteins [J]. Analytical Biochemistry, 1984, 138(1):181-188.
- [30] Premkumar L, Bageshwar U K, Gokhman I, et al. An unusual halotolerant α -type carbonic anhydrase from the alga *Dunaliella salina* functionally expressed in *Escherichia coli* [J]. Protein Expression and Purification, 2003, 28(1):151-157.
- [31] Bracey M H, Christiansen J, Tovar P, et al. Spinach carbonic anhydrase: Investigation of the zinc-binding ligands by site-directed mutagenesis, elemental analysis, and EXAFS [J]. Biochemistry, 1994, 33(44):13126-13131.
- [32] Rowlett R S. Structure and catalytic mechanism of the β -carbonic anhydrases [J]. Biochimica et Biophysica Acta, 2010, 1804(2):362-373.
- [33] Parisi G, Fornasari M, Echave J. Evolutionary analysis of γ -carbonic anhydrase and structurally related proteins [J]. Molecular Phylogenetics and Evolution, 2000, 14(3):323-334.
- [34] Parisi G, Perales M, Fornasari M, et al. Gamma carbonic anhydrases in plant mitochondria [J]. Plant Molecular Biology, 2004, 55(2):193-207.
- [35] Moroney J V, Bartlett S G, Samuelsson G. Carbonic anhydrases in plants and algae [J]. Plant Cell and Environment, 2001, 24(2):141-153.
- [36] Moroney J V, Ynalvez R A. Proposed carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii* [J]. Eukaryotic Cell, 2007, 6(8):1251-1259.
- [37] Karlsson J, Clarke A K, Chen Z Y, et al. A novel α -type carbonic anhydrase associated with the thylakoid membrane in *Chlamydomonas reinhardtii* is required for growth at ambient CO₂ [J]. EMBO Journal, 1998, 17(5):1208-1216.
- [38] Sinetova M A, Kupriyanova E V, Markelova A G, et al. Identification and functional role of the carbonic anhydrase Cah3 in thylakoid membranes of pyrenoid of *Chlamydomonas reinhardtii* [J]. Biochimica et Biophysica Acta, 2012, 1817(8):1248-1255.
- [39] Park Y I, Karlsson J, Rojdestvenski I, et al. Role of a novel photosystem II-associated carbonic anhydrase in photosynthetic carbon assimilation in *Chlamydomonas reinhardtii* [J]. FEBS Letters, 1999, 444(1):102-105.
- [40] Sunderhaus S, Dudkina N V, Jansch L, et al. Carbonic anhydrase subunits form a matrix-exposed domain attached to the membrane arm of mitochondrial complex I in plants [J]. Journal of Biological Chemistry, 2006, 281(10):6482-6488.

Molecular cloning and characterization of carbonic anhydrase (CA) genes from *Myrmecia incisa* Reisi H4301

MEI Shouhua, BI Yanhui, ZHOU Zhigang*

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

Abstract: Based upon five contigs coding for carbonic anhydrase (CA) from a pyrosequencing transcriptome of *Myrmecia incisa*, gene-specific primers were designed for CA gene cloning from this green microalga. Five CA genes, named $Mi\alpha CA1$, $Mi\alpha CA2$, $Mi\beta CA1$, $Mi\beta CA2$ and $Mi\gamma CA$ were cloned by using the technique of rapid amplification of cDNA ends. The open reading frames of the cloned full-length cDNAs were 963 bp, 1 089 bp, 1 041 bp, 738 bp and 687 bp in length, encoding putative proteins composed of 320, 362, 346, 245 and 228 amino acids, respectively. The encoded CA proteins were rich in hydrophobic amino acids which accounted for 41.25%, 45.31%, 43.35%, 42.45% and 43.42% of total amino acids, respectively. Neighbor-joining (NJ) phylogenetic tree inferred from the putative proteins of CA genes indicated that these cloned CA genes were divided into three clades: $Mi\alpha CA1$ and $Mi\alpha CA2$ in α -CA clade, $Mi\beta CA1$ and $Mi\beta CA2$ in β -CA one and $Mi\gamma CA$ in γ -CA one. $Mi\alpha CA1$ and $Mi\alpha CA2$ are characterized by 3 conserved His residues, $Mi\beta CA1$ and $Mi\beta CA2$ are by 2 conserved Cys and 1 conserved His residues whereas $Mi\gamma CA$ is found to use Arg, His and Asn instead of the reported 3 residues of His to bind Zn^{2+} to form their catalysis centre. $Mi\alpha CA1$, clustered with CAH3 from *Chlamydomonas reinhardtii* in the NJ phylogenetic tree, had a two-part leader sequence, suggesting that it might be localized and associated with PSII in the thylakoid lumen like CAH3. $Mi\alpha CA2$, close to CAH1 after protein alignment, had a signal peptide which could lead the mature $Mi\alpha CA2$ to the periplasmic space resembling CAH1 in *C. reinhardtii*. It was predicted that this CA could function as CAH1 to transfer HCO_3^- into CO_2 for photosynthesis. Two β -CAs, $Mi\beta CA1$ and $Mi\beta CA2$, were clustered with CAH7 and CAH8 from *C. reinhardtii*, which might be involved in interconversion between HCO_3^- and CO_2 in the cytoplasm of *M. incisa*. $Mi\gamma CA$, grouped with γ -CA from higher plants, was expected to be associated with the complex I of mitochondrial electron transport chain to work. The 5 cloned CA genes from *M. incisa* should be localized at various regions of the algal cells to cooperate with each other for the reversible interconversion of CO_2 and HCO_3^- so as to modify pH and transfer CO_2 among the organelles.

Key words: *Myrmecia incisa*; *Chlamydomonas reinhardtii*; carbonic anhydrase; cloning; CO_2

Corresponding author: ZHOU Zhigang. E-mail: zgzhou@shou.edu.cn