

文章编号: 1000-0615(2010)05-0786-010

DOI:10.3724/SP.J.1231.2010.06769

缘管浒苔 *rbcL* 全长 cDNA 克隆与序列分析

应成琦^{1,2}, 尹顺吉¹, 林森杰³, 沈 轶¹, 何培民^{1*}

(1. 上海海洋大学农业部水产种质资源与养殖生态重点开放实验室, 上海 201306;

2. 中国水产科学研究院东海水产研究所, 中国水产科学研究院盐碱地渔业工程技术中心, 上海 200090;

3. Department of Marine Sciences, University of Connecticut, Groton, CT 06340, USA)

摘要: 对缘管浒苔光合作用第一关键酶 Rubisco 大亚基基因(*rbcL*)全长 cDNA 序列进行了克隆分离。首先应用 RT-PCR 方法获得了 *rbcL* 大片段 cDNA 序列, 序列长度为 1 101 bp, 并进行了测序分析。然后分别应用 5'RACE 方法和 3'RACE 方法获得 2 个克隆序列, 其序列长度分别为 371 bp 和 579 bp, 并进行了测序分析。在此基础上, 将 3 个片段序列进行拼接, 获得了 Rubisco 大亚基全长 cDNA 序列(1 472 bp), 其中包括 1 425 bp 的编码区序列及 47 bp 的前导序列(NCBI 登录号: DQ813496), 并推断出蛋白质序列(474 个氨基酸)。对序列进行了相关生物信息分析和同源性比对, 结果显示与已克隆出全长 *rbcL* 基因的 7 种绿藻比较具有较高同源性, 核苷酸序列和蛋白质序列同源性分别达到了 82.07% ~ 85.78% 和 88.00% ~ 94.11%, 其中与蛋白核小球藻蛋白质序列同源性最高, 为 94.11%。在此基础上, 使用生物信息学软件对缘管浒苔 *rbcL* 氨基酸序列与其它植物或藻类进行了多序列比对, 并且进一步应用在线软件程序进行了蛋白质的二级结构分析和三级结构预测。

关键词: 缘管浒苔; 核酮糖 1,5-二磷酸羧化酶/加氧酶大亚基基因(*rbcL* 基因); 前导序列; cDNA 末端快速扩增; cDNA

中图分类号: Q 75; S 917

文献标识码: A

核酮糖 1,5-二磷酸羧化酶/加氧酶(Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase, E. C. 4.1.1.39, 简称 Rubisco)是光合作用还原和光合氧化二个连锁循环的交叉点, 为调节光合作用和光呼吸、决定净光合速率的第一关键酶^[1-2]。在高等植物和绿藻中, Rubisco 由 8 个大亚基及 8 个小亚基组成, 大亚基由叶绿体基因组编码(简称为 *rbcL* 基因), 小亚基由核基因组编码(简称为 *rbcS* 基因)^[3]。

真核藻类 Rubisco 全酶的底物活化中心和催化中心都位于大亚基上, 因此研究 Rubisco 大亚基编码序列 *rbcL* 及其大亚基氨基酸序列对于深入研究其蛋白的特殊功能结构域和酶作用机理具有尤为重要的意义^[2,4]。此外, *rbcL* 基因序列已成为分子系统学研究中最为广泛的分子

指标之一, 用于探测植物间的系统关系和分子进化^[4]。因此, 近年来, *rbcL* 基因编码区序列资料迅速增长, 自 1980 年首次测定了玉米 *rbcL* 基因的序列以来, 迄今已有 1 500 余种植物的 *rbcL* 基因序列得以测定, 而且新的数据还在日益增多, 但完整 *rbcL* 基因序列仍然很有限^[4-5]。迄今为止, 在 Genbank 中, 已有许多关于绿藻的叶绿体 *rbcL* 基因序列的报道, 但绝大多数是使用通用引物, 得到的只是 *rbcL* 的一个大片段; 少数获得了 *rbcL* 基因全序列, 但全部为微型绿藻, 且通常只有一个 ORF, 缺少关于 3' 和 5' 非翻译区的信息; 未见到关于大型海藻 *rbcL* cDNA 完整序列的报道。

缘管浒苔 (*Ulva linza*) 隶属绿藻门 (Chlorophycophyta), 绿藻纲 (Chlorophyceae), 石

收稿日期: 2009-12-28 修回日期: 2010-03-09

资助项目: 国家海洋局绿潮灾害专项课题; 国家自然科学基金项目(30371101); 教育部博士点基金项目; 上海市科委优秀学科带头人计划项目(08XD14037); 上海市水生生物学重点学科资助项目(S30701)

通讯作者: 何培民, Tel: 15692165272, E-mail: pmhe@shou.edu.cn

莼目(Ulvales),石莼科(Ulvaceae)。作为一种大型经济海藻,其产品清香独特,具有抗癌及抗溃疡等药用价值^[6],在日本及韩国已大规模栽培生产,我国南方也已开始栽培生产。浒苔藻体生长很快,日相对生长率可高达 50%^[7],具有快速吸收氮磷改善水质等生态修复功能^[8]。但近年来由于海洋富营养化现象越来越严重,大型绿藻形成的绿潮发生的频率不断增加,这既是大自然对人类活动的一种警示,也是一种自我修复作用现象^[8]。为了揭开浒苔快速生长机制,选取缘管浒苔为材料,首次克隆了其光合作用关键酶 Rubisco 大亚基基因的完整 cDNA 序列,并应用生物信息学对序列进行了分析,推测了其氨基酸序列及蛋白质结构,为今后研究浒苔快速生长分子机制、海洋生态修复和绿潮控制等奠定理论基础。

1 材料与与方法

1.1 材料与藻体培养

实验所用藻类缘管浒苔采自江苏如东吕泗海区。经过清洗去除杂藻后,在 0.2% KI 溶液内处理 10 min,再用消毒过的自然海水将其冲洗干净,暂养于室内。从缘管浒苔藻体中各挑选 30 株进行单株分离和纯化培养。叶片洗净后,经过一定特殊培养使其放散孢子并形成小苗,然后在显微镜下分离单个小苗进行单独培养,并作为株系进行保存,编号为 YSJ01 ~ 30。选取株系 YSJ12 进行扩大培养,藻体培养条件为:光照强度为 35 ~ 45 $\mu\text{mol}/(\text{m}^2 \cdot \text{s})$,温度 18 $^{\circ}\text{C}$,光照周期为 12:12,充气培养,每隔 4 ~ 5 天更新藻体培养液^[1]。

1.2 缘管浒苔总 RNA 的提取

缘管浒苔总 RNA 的提取采用 Trizol RNA 提取试剂盒(Invitrogen 公司产品,美国),提取方法按照其说明书进行。收集一定量新鲜旺盛生长的藻体,用液氮快速充分研磨成粉状后迅速提取缘管浒苔总 RNA。得到的总 RNA 溶于经 DEPC 处理的水中,放入 -70°C 冰箱保存^[9]。

1.3 缘管浒苔 *rbcL* 基因 cDNA 部分序列克隆

cDNA 合成 cDNA 合成采用反转录试剂盒(FIRST STRAND cDNA SYNTHESIS KIT I, BBI),并按照反转录试剂盒说明书进行。以适量总 RNA 为模板,以 Oligo(dT) 为引物,在 AMV

逆转录酶的作用下,42 $^{\circ}\text{C}$ 温育 1 h,合成 cDNA 第一链。 -20°C 保存。

RT-PCR 利用 *rbcL* 基因在进化中高度保守的特点,根据已知的缘管浒苔近缘绿藻 *rbcL* 序列及叶绿体基因组全序列资料,设计一对引物(引物由上海生工生物工程技术有限公司合成)。

正向引物 *rbc* F1: 5'-GGTGTAAGATT-ACCGTTTAACTT-3';

反向引物 *rbc* R2: 5'-GTGAATACCACCT-GAAGCTACAGG-3'。

以 cDNA 为模板,进行 PCR 扩增。PCR 程序:样品 94 $^{\circ}\text{C}$ 预变性 3 min,94 $^{\circ}\text{C}$ 30 s,57 $^{\circ}\text{C}$ 30 s,72 $^{\circ}\text{C}$ 1 min,30 个循环,72 $^{\circ}\text{C}$ 延伸 10 min。得到的 PCR 产物送 TaKaRa 公司进行 TA 克隆及测序。TA 克隆使用的载体为 pMD[®] 19-T Vector,将连接 PCR 产物的载体转化大肠杆菌 JM109,采用蓝白斑筛选的方法挑选阳性克隆并测序。

1.4 缘管浒苔 *rbcL* 基因 5'端 cDNA 序列克隆

5'端 cDNA 序列克隆采用 5'-RACE 方法,试剂盒为 5'-Full RACE Core Set(Takara 公司产品),并按照 5'-RACE 试剂盒说明书进行。根据缘管浒苔 *rbcL* 大片段 cDNA 序列设计克隆 5'端 cDNA 序列所需的下列引物(引物由上海生工生物工程技术有限公司合成)。

反转录引物: 5'-P-GTCACGCCAACGCAT-3',

正向引物 F1: 5'-GAACCATTAGGAGA-AGACGACCAA-3',

F2: 5'-AGATTTACGTATTCCACCAGCT-TATGTT-3',

反向引物 R1: 5'-TACTTGATAATC-AGGCGTGTAAATAAGTT-3',

R2: 5'-TCAGGCGTGTAAATAAGTTAAAC-GGTAAT-3'。

克隆得到的 PCR 产物送 TaKaRa 公司进行 TA 克隆及测序。TA 克隆使用的载体为 pMD[®] 19-T Vector,将连接 PCR 产物的载体转化大肠杆菌 JM109,采用蓝白斑筛选的方法挑选阳性克隆并测序。

1.5 缘管浒苔 *rbcL* 基因 3'端 cDNA 序列克隆

3'端 cDNA 序列克隆采用 3'-RACE 方法,试剂盒为 3'-Full RACE Core Set(Takara 公司),步

骤按照其说明书进行。根据缘管浒苔 *rbcL* 大片段 cDNA 序列设计克隆 3' 端 cDNA 序列所需的下列引物(引物由上海生工生物工程技术有限公司合成)。

正向引物 *rbcL* F1: 5'-GCTCATTTCIGTC-GTGCTAGTG-3';

rbcL F2: 5'-TGTCGTGCTAGTGGATTAT-TATTAC-3'。

克隆得到的 PCR 产物送 TaKaRa 公司进行 TA 克隆及测序。TA 克隆使用的载体为 pMD[®]19-T Vector, 将连接 PCR 产物的载体转化大肠杆菌 JM109, 采用蓝白斑筛选的方法挑选阳性克隆并测序。

1.6 序列生物信息分析与蛋白质结构预测

对 3 个序列的测序结果进行拼接, 推断缘管浒苔 *rbcL* 基因全长 cDNA 序列。然后, 采用生物信息学软件 DNAMAN (LynnonBiosoft, 美国), 将全长 cDNA 序列翻译成蛋白质序列。在此基础上, 将推断的缘管浒苔 *rbcL* 氨基酸序列与菠菜 (*Spinacia oleralea*) 和玉米 (*Zea mays*) 及莱茵衣藻 (*Chlamydomonas reinhardtii*) 的 *rbcL* 氨基酸序列进行多序列比对, 以此来比较在氨基酸水平上缘管浒苔 Rubisco 酶的重要功能位点与高等植物和其他近缘藻类之间的差异。此外, 还在线 (http://npsa-pbil.ibcp.fr/cgi-bin/secpred_sopm.pl) 预测蛋白质二级结构, 以及使用计算机软件 Rasmol 对其三级结构进行推断, 并在线 (<http://swissmodel.expasy.org/repository/>) 对推断出的三级结构进行同源性比较。从 Genebank 中选取 7 种其他近缘绿藻的全长 *rbcL* 基因序列和蛋白质序列, 并与推断得到的缘管浒苔 *rbcL* 基因全长 cDNA 序列和蛋白质序列进行同源性比较。根据推断的缘管浒苔 *rbcL* 氨基酸序列, 使用 Neighbor-Joining 方法^[10] 构建进化树。

2 结果与分析

2.1 缘管浒苔总 RNA 提取与电泳检测

采用琼脂糖凝胶电泳检测缘管浒苔总 RNA 的完整性, 胶的浓度为 1%, 其电泳结果见图 1。从图中可以看见清晰的 28S、18S 及 5.8S RNA 条带, 所提取的缘管浒苔总 RNA 符合实验要求。

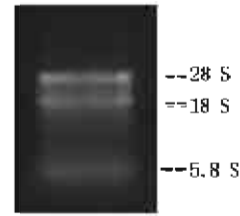


图1 缘管浒苔总 RNA 电泳图谱
Fig.1 Electrophoresis map of total RNA in *U. linza*

2.2 缘管浒苔 *rbcL* 基因 cDNA 部分序列克隆与分析

采用 RT-PCR 方法分离缘管浒苔 *rbcL* 大片段 cDNA 序列, 得到一条特异的扩增条带, 大小约为 1.2 kb (图 2)。PCR 产物经 TA 克隆并测序, 结果表明该序列大小为 1 101 bp (+93 ~ +1 193)。用该序列在 NCBI 上进行 BLAST 同源搜索, 结果表明: 所得相似基因均为 *rbcL* 基因且与 *Ulva linza* (同源性 99%)、*Ulva prolifera* (同源性 99%)、*Ulva procera* (同源性 99%)、*Ulva intestinalis* (同源性 98%)、*Ulva curvata* (同源性 98%)、*Ulva flexuosa* (同源性 98%)、*Ulva taeniata* (同源性 98%)、*Ulva ohnoi* (同源性 97%)、*Ulva armoricana* (同源性 97%) 以及 *Ulva gigantea* (同源性 97%) 等绿藻叶绿体 *rbcL* 基因序列高度同源。可以确定克隆得到的该序列为缘管浒苔 *rbcL* 基因 cDNA 序列。

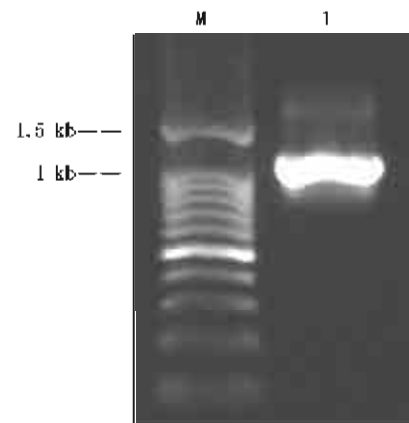


图2 缘管浒苔 *rbcL* 基因 cDNA 部分序列 PCR 产物电泳图谱
M: 100bp ladder Marker; 1: *rbcL* 大片段 cDNA 序列 PCR 产物。
Fig.2 Electrophoresis map of PCR amplification product of partial *rbcL* cDNA
M: 100 bp ladder Marker; 1: PCR amplification product of partial *rbcL* cDNA.

2.3 缘管浒苔 *rbcL* 基因 5'端 cDNA 序列克隆与分析

使用 Takara 5'-Full RACE Core Set 试剂盒克隆 *rbcL* 5'端 cDNA 序列,得到一个特异的扩增条带,大小约 400 bp(图 3-A)。PCR 产物经 TA 克隆并测序,获得一个 371 bp 的序列,其中包括完整的 5'端序列 130 bp。该 5'端序列中含有起始密码子 ATG 和一段 80 bp 的编码序列(+51 ~ +130),以及起始密码子 ATG 前面的一段 47 bp 的前导序列(+1 ~ +47)。

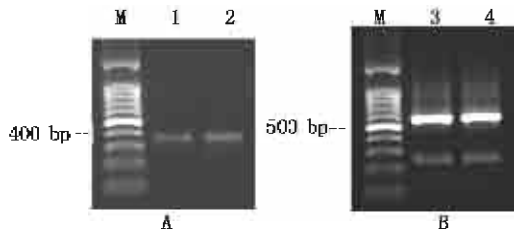


图 3 5'和 3'RACE, PCR 扩增产物电泳图谱

A:5'RACE; B:3'RACE; M:100 bp ladder Marker;1,2: *rbcL* 5'末端 cDNA 序列 PCR 产物;3,4: *rbcL* 3'末端 cDNA 序列 PCR 产物。

Fig. 3 Electrophoresis map of PCR amplification product of 5'end with 5'and 3'RACE method

A:5'RACE; B:3'RACE; M:100 bp ladder Marker;1,2: PCR amplification product of 5'end in *rbcL* cDNA; 3,4: PCR amplification product of 3'end in *rbcL* cDNA.

```
MAPQTEFKAGTGFKAGVKDYRLTYYPDYQVEDYDII.AAFRMPQPGVPAHEAGAAVAAS
STGTWTFVWTDGLTSLDRYKGRCYDIEPLGEDDQYIAIYA YPLDLEFEGISVTLNLFISVGNVFG
FKALRALRLEDLRIPPAYVKTFQGPPIIGIQRVERDKLNKYGRGLLGCTIKPKLGLSAKNYGRAV
YECLRGGLDFTKDDENVNSQPFMRWRDRFLFTA EAIYKSQSETGEVKGIYLNATAGTCEEM
MERGQFAKDLGVPIIMHDYITGGFTANTSLAHFCRASGLLLHHIRAMHAVIDRQRNHGIHFRV
LAKIIRMSGGDHLHSGTVVVGKLEGERHETLGFVDMRDDYIEKDRSRGIYFTQDWVSLPGTMP
VASGGIHVWHMPALVBIIFGLDACLQFGGGTLGHPWGNAPGAAANRVALEACTQARNEGRDL
ASEGGDVIRAACKWSPELAAACEVWKEIKFEFDAIDTL
```

2.6 同源性分析与进化树构建

目前在绿藻中, *Chlorella ellipsoidea* (EU038287), *Chlorella vulgaris* (EU038286), *Auxenochlorella protothecoides* (EU038285), *Tetraselmis suecica* (DQ173247), *Dunaliella tertiolecta* (AY882012), *Chloromonas* sp. ANT3 (CSU80809)和 *Chlorella pyrenoidosa* (AB240145)这 7 种绿藻的完整叶绿体 *rbcL* 基因已得到。在 *Chlorella pyrenoidosa* 中还得到了 *rbcL* 基因全长 cDNA 序列。将缘管浒苔 *rbcL* 基因全长 cDNA 序列与上述 7 种绿藻的 *rbcL* 基因的核苷酸序列和推测的蛋白质序列进行比较(表 1)。结果表明,在基因水平上缘管浒苔与 *Auxenochlorella*

2.4 缘管浒苔 *rbcL* 基因 3'端 cDNA 序列克隆与分析

使用 Takara 3'-Full RACE Core Set 试剂盒克隆 *rbcL* 3'端 cDNA 序列,获得一个约 600 bp 的扩增条带(图 3-B)。将得到的 PCR 产物经 TA 克隆并测序,获得一个长度为 579 bp 的序列,其中包含终止密码子 TAA(图 4 波浪下划线标记部分)。

2.5 缘管浒苔 *rbcL* 基因全长 cDNA 序列拼接与蛋白质序列推断

拼接上述得到的 3 个序列,推断出缘管浒苔 *rbcL* 全长 cDNA 序列,采用分子生物学软件 DNAMAN 将全长 cDNA 序列翻译成蛋白质序列。分析序列的 ORF,并与其他藻类 Rubisco 大亚基蛋白质序列对比,确定翻译起始密码子 ATG(加粗双下划线表示)及终止密码子 TAA(加粗波浪下划线表示)的位置,最终得到缘管浒苔 *rbcL* 全长 cDNA 序列,共 1 472 bp。在起始密码子 ATG 前有 47 bp 的前导序列。该序列已在 NCBI 上获得登录号:DQ813496(图 4)。

根据推断的缘管浒苔 *rbcL* 全长 cDNA 序列推算缘管浒苔 Rubisco 大亚基氨基酸顺序,得到含有 474 个氨基酸的蛋白质序列:

protothecoides 的 *rbcL* 同源性最高,达 85.78%;在氨基酸水平上,缘管浒苔与 *Chlorella pyrenoidosa* 的 *rbcL* 同源性最高,为 94.11%。

目前,在藻类和高等植物中,*rbcL* 基因已被用于探讨从科内属间、亚纲内甚至各主要类群之间的系统发育关系^[4]。从 genbank 中下载绿藻门中不同科属绿藻的 *rbcL* 氨基酸序列,并以红藻门海藻为外群(out group),根据推断的缘管浒苔 *rbcL* 氨基酸序列,使用 Neighbor-Joining 方法^[10]构建进化树。结果表明,缘管浒苔与其他浒苔属的藻类亲缘关系最近,其次是石莼科及绿藻门的藻类;与红藻的亲缘关系较远(图 5)。

```

1  GTTTGGGCAG ATCGAATGGT CAATTTATTA AATTTTAATT ATTAAAAAATG GCTCCAC
61  CTGAAACAAA AGCAGGTA CTGGTTAAAG CTGGTGTAAA AGATTACCGT TTAACCTT
121 ACACGCCTGA TTATCAAGTA GAAGATACTG ATATTTTAGC AGCGTTCCGT ATGACTC
181 ACCAGGAGT ACCGGCAGAA GAAGCAGGTG CAGCTGTTGC TGCTGAATCA TCAACAG
241 CTTGGACAA CTTATGGACT GATGGTTTAA CATCTTTAGA TCGTTATAAA GGTCGTT
301 ACCACATTGA ACCATTAGGA GAAGACCACC AATATATTGC TTATATTGCT TATCCTT
361 ACTATTTGA AGAAGGATCA GTTACAAACT TATTTACTTC AATTGTAGGT AACGTTT
421 GTTTTAAAGC TTTACGTGCT TTACGTTTAG AAGATTTACG TATCCACCA GCTTATG
481 AAACATTCCA AGGTCCACCG CATGGTATTC AGGTTGAACG TGATAAATA AACAAAT
541 GTCGTGGTTT ATTAGGTTGT ACGATTAAC CAAAATTAGG TCTTTCAGT AAAAACT
601 GACGTGCTGT TTATGAATGT TTACGAGGTG GTCTTGATTT TACAAAAGAC GATGAAA
661 TAAACTCACA ACCTTTCATG CGTTGGCGTG ACCGTTTCTT ATTTACTGCT GAAGCAA
721 ACAATCTCA ATCTGAAACT GGTGAGGTTA AAGGACATTA CTTAAATGCA ACTGCGG
781 CATGTGAAGA AATGATGGAA CGTGGTCAAT TTGCTAAAGA TTTAGGTGTT CCAATTA
841 TGCACTACTA CATTACTGGT GGTTTTACAG CTAACACTTC ATTAGCTCAT TTCTGTC
901 CTAGTGGATT ATTATTACAT ATTCACCGTG CTATGCACGC TGTTATTGAC CGTCAAC
961 ATCACGGTAT TCACTCCGA GTATTAGCGA AAATTTTACG TATGTCAGGT GGTGATC
1021 TACACTCAGG AACAGTAGTA GGTAAATTAG AAGGTGAACG TGAAATTACT TTAGGTT
1081 TTGACTTAAT GCGTGATGAT TACATTGAAA AAGATCGTAG TCGTGGTATT TACTTTA
1141 AAGATTGGGT TAGTTTACCA GGTACAATGC CTGTAGCTTC AGGTGGTATT CACGTTT
1201 ATATGCCGGC ATTAGTTGAA ATTTTTGGTG ACGACGCATG TTTACAATTC GGTGGTG
1261 CATTAGGACA CCCTTGGGGT AATGCTCCAG GAGCCGCTGC AAACCGTGTT GCTTTAG
1321 CTTGACACA AGCTCGAAAC GAGGGACGTG ATTTAGCATC TGAAGGTGGT GACGTAA
1381 GTGCTGCTG TAAATGGAGT CCTGAATTAG CTGCAGCTTG TGAAGTATGG AAAGAAA
1441 AATTTGAATT TGATGCTATT GATACATTAT AA
    
```

图 4 缘管浒苔 *rbcL* 基因全长 cDNA 序列

5'RACE 获得的序列用“_”标出,3'RACE 获得的序列用“~”标出;翻译起始密码子用加粗双下划线标出,终止密码子用加粗波浪下划线标出。

Fig. 4 Full-length cDNA sequence of *rbcL* in *U. linza*

Sequence of *rbcL* acquired by 5' RACE method was underlined with “_”, Sequence of *rbcL* acquired by 3' RACE method was underlined with “~”; Initiation codon was underlined with a bold “_”, stop codon was underlined with a bold “~”.

表 1 缘管浒苔 *rbcL* 的 cDNA 序列和氨基酸序列与其他物种的同源性比较

Tab. 1 Comparison identities of *rbcL* cDNA sequence and amino acid sequence

种名 species	cDNA 序列同源性 identities of cDNA	氨基酸序列同源性 identities of amino acid	NCBI 登录号 accession number
椭圆小球藻 <i>Chlorella ellipsoidea</i>	82.70%	92.63%	EU038287
小球藻 <i>Chlorella vulgaris</i>	84.10%	92.84%	EU038286
原壳小球藻 <i>Auxenochlorella protothecoides</i>	85.78%	93.89%	EU038285
亚心形扁藻 <i>Tetraselmis suecica</i>	84.45%	88.00%	DQ173247
杜氏盐藻 <i>Dunaliella tertiolecta</i>	84.87%	90.11%	AY882012
拟衣藻 <i>Chloromonas</i> sp. ANT3	82.07%	89.68%	CSU80809
蛋白核小球藻 <i>Chlorella pyrenoidosa</i>	85.64%	94.11%	AB240145

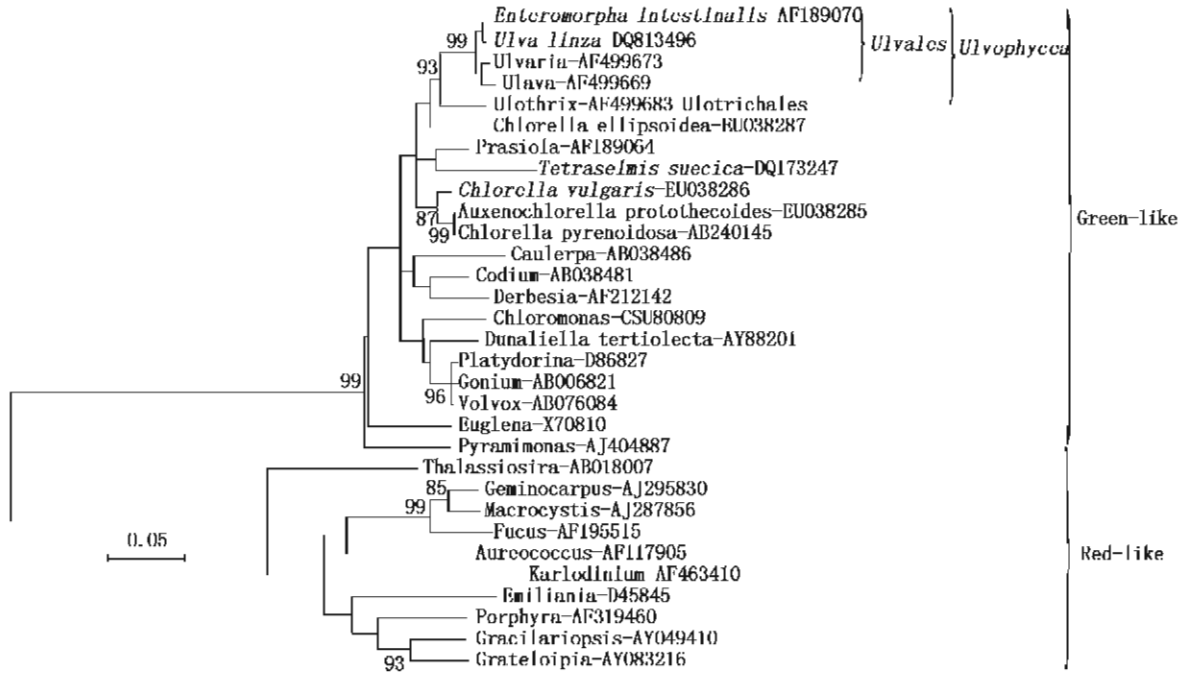


图5 根据 *rbcL* 氨基酸序列构建的进化树

Fig.5 Neighbor-Joining tree of *rbcL* based on amino acid sequences

2.7 氨基酸序列分析与蛋白质结构预测

为了比较在氨基酸水平上缘管浒苔 Rubisco 大亚基的重要功能位点与高等植物和其他近缘藻类之间的差异,选择菠菜和玉米这两种高等植物以及莱茵衣藻的 *rbcL* 氨基酸序列与缘管浒苔 *rbcL* 氨基酸序列进行了多序列比对,结果见图 6。与其他 3 个 *rbcL* 氨基酸序列相比,在缘管浒苔 *rbcL* 氨基酸序列中,在第 90 个氨基酸 L(亮氨酸)和第 91 个氨基酸 G(甘氨酸)之间有一个缺失突变位点。此外,在缘管浒苔 *rbcL* 氨基酸序列中,A,C,D,E,M 等用加粗标记的氨基酸(图 6)与莱茵衣藻中相应的氨基酸完全相同,但与菠菜和玉米这样的高等植物中相应的氨基酸并不完全相同。图 7 中的序列 I, II, III, IV 与莱茵衣藻中的相应的序列也完全相同,其中 I -A1(12 个氨基酸), III -A2(20 个氨基酸), IV -A3(13 个氨基

酸)在莱茵衣藻 *rbcL* 氨基酸序列中是与催化位点功能相关的序列而 II -C(17 个氨基酸)是与 CO₂ 活化肽相关的序列。

根据氨基酸序列在线预测蛋白二级结构 (http://npsa-pbil.ibcp.fr/cgi-bin/secpred_sopm.pl),得知缘管浒苔 Rubisco 大亚基蛋白二级结构有 α-螺旋、延伸直带和自由卷曲等三种形式,其中 α-螺旋有 199 处,延伸直带 52 处,自由卷曲 210 处。

据推测氨基酸序列在线预测蛋白三级结构 (<http://swissmodel.expasy.org/>),并利用计算机软件 Rasmol 进行分析,在线 (<http://swissmodel.expasy.org/repository/>)对根据缘管浒苔 *rbcL* 氨基酸序列推断出的三级结构进行同源性比较(图 7),结果发现:与莱茵衣藻的 Rubisco 蛋白(Parent PDB: 2v69)三级结构最为相似,同源性为 92%。

```

U. l  MAPQTETKACTGFKACVKDYRLTYTTPDYQVEDTDILAAFRMTPQPGVPAEEAGAAVAAESSTGTWTFVWT
C. r  MVPQTETKAGAGFKAGVKDYRLTYTTPDYVVRDTDILAAFRMTPQLGVPPEECGAAVAEAESSTGTWTFVWT
S. o  MSPQTETKASVGFKAQVKDYKLTYYTPEYETLDDTDILAAFRVSPQPGVPPPEEAGAAVAEAESSTGTWTFVWT
Z. m  MSPQTETKASVGFKAQVKDYKLTYYTPEYETPKDTDILAAFRVTPQLGVPPEEAGAAVAEAESSTGTWTFVWT

U. l  DGLTSLDRYKGRCYDIEPL-GEDDQYIAYIAYPLDLFEEGVSNTNLF1TSIVGNVFGFKALRALRLEDLRI2PP
C. r  DGLTSLDRYKGRCYDIEPV3GEDNQYIAYVAYPIDLFEEGVSNTNMTSIVGNVFGFKALRALRLEDLRI4PP
S. o  DGLTNLDRYKGRCYHIEPV5AGEENQYICYVAYPLDLFEEGVSNTNMTSIVGNVFGFKALRALRLEDLRI6PV
Z. m  DGLTSLDRYKGRCYHIEPV7GDPDQYICYVAYPLDLFEEGVSNTNMTSIVGNVFGFKALRALRLEDLRI8PP

I-A1                                II-C
U. l  AYVKTFFQPPHGIQVERDKLNKYGRLLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFMR
C. r  AYVKTFFQPPHGIQVERDKLNKYGRLLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFMR
S. o  AYVKTFFQPPHGIQVERDKLNKYGRLLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFMR
Z. m  AYSKTFQPPHGIQVERDKLNKYGRLLGCTIKPKLGLSAKNYGRACYECLRGGLDFTKDDENVNSQPFMR

U. l  WRDRFLFTAEEAIYKSQSETGEVKGHYLNATAGTCEEMMERGQFAKDLGVPIIMHDYITGGFTANTSLAHFC
C. r  WRDRFLFVAEAIYKAQ9AETGEVKGHYLNATAGTCEEMMKRAVCAKELGVPIIMHDYITGGFTANTSLAIYC
S. o  WRDRFLFCAEALYKAQ10AETGEIKGHYLNATAGTCEEMMKRAVFARELGVPIVMHDYITGGFTANTTSLHYC
Z. m  WRDRFVFCAEAIYKAQ11AETGEIKGHYLNATAGTCEEMIKRAVFARELGVPIVMHDYITGGFTANTTSLHYC

III-A2
U. l  RASGLLLHIHRAMHAVI DRQRNHGIFHRVLA12KILRMSGGDHLHSGTVVVGKLEGEREITLGFVDLMRDDYIE
C. r  RDNGLLLHIHRAMHAVI DRQRNHGIFHRVLA13KALRMSGGDHLHSGTVVVGKLEGEREITLGFVDLMRDDYVE
S. o  RDNGLLLHIHRAMHAVI DRQRNHGMHFRVLA14KALRLSGGDHHS15GTVVVGKLEGERDITLGFVDLLRDDYIE
Z. m  RDNGLLLHIHRAMHAVI DRQRNHGMHFRVLA16KALRMSGGDHIHSGTVVVGKLEGEREITLGFVDLLRDDFIE

U. l  KDRSRGIYFTQDWVSLP17GMTMPVASGGIHVWHMPALVEIFGDDACLQFGGGT18LGHPWGNAPGAAANRVALEA
C. r  KDRSRGIYFTQDWCSMPGVMPVASGGIHVWHMPALVEIFGDDACLQFGGGT19LGHPWGNAPGAAANRVALEA
S. o  KDRSRGIYFTQSWVSTPGVLPVASGGIHVWHMPALTEIFGDDSVLQFGGGT20LGHPWGNAPGAVANRVALEA
Z. m  KDRSRGIFF21TQDWVSM22PGVIPVASGGIHVWHMPALTEIFGDDSVLQFGGGT23LGHPWGNAPGAAANRVALEA

IV-A3
U. l  CTQARNEGRDLASEGGDVIRAACKWSP24ELAAACEVWKEIKFE-FDAIDTL
C. r  CTQARNEGRDLAREGGDVIRSACKWSP25ELAAACEVWKEIKFE-FDTIDKL
S. o  CVQARNEGRDLAREGNTIIREATKWSP26ELAAACEVWKEIKFE-FPAMD27TV
Z. m  CVQARNEGRDLAREGNEI28IKACKWSAELAAACEI29WKEIKFDGPKAMDTI

```

图6 缘管浒苔 *rbcL* 氨基酸序列与其他物种的比较

U. l: *U. linza*; C. r: *Chlamydomonas reinhardtii*; S. o: *Spinaci oleralea*; Z. m: *Zea mays*; 活化位点 A1, A2, A3 以及 CO₂ 活化区域用下划线标出。缘管浒苔 *rbcL* 氨基酸序列中与 *Chlamydomonas reinhardtii* 完全相同的氨基酸用加粗标出。位于 L (亮氨酸) 和 G (甘氨酸) 之间的缺失突变位点用方框标出。

Fig. 6 Comparison of the *rbcL* amino acid sequences from *U. linza* with those from other species

U. l: *U. linza*; C. r: *Chlamydomonas reinhardtii* with 475 amino acids; S. o: *Spinaci oleralea* with 475 amino acids; Z. m: *Zea mays* with 476 amino acids. The active sites A1, A2, A3, and the CO₂ activation region were underlined. The amino acids marked with bold were fully identified with those from *Chlamydomonas reinhardtii*. The deletion mutation between L (Leu) and G (Gly) were boxed.



图7 缘管浒苔 Rubisco 大亚基蛋白三级结构预测

Fig. 7 Predicted three dimension structure of Rubisco large subunit in *U. linza*

3 讨论

rbcL 基因是与光合作用调控有关的基因,因而备受人们的关注。近年来,关于 *rbcL* 基因的研究报道很多,本文所报道的缘管浒苔 *rbcL* 基因的全长 cDNA 序列,是石莼科大型经济海藻中所获得的第一个完整的 *rbcL* 基因序列。本文除了得到了缘管浒苔 *rbcL* 基因完整的 ORF 以外,还获得了位于 5' 端起始密码子 ATG 前的一段非翻译序列;而目前在绿藻中,关于完整的 *rbcL* 基因的报道还十分有限,关于 *rbcL* 基因非翻译区的信息则更少。

本文所报道的序列与已测定的绿藻和高等植物 *rbcL* 基因在结构上相似:编码区长度约为 1.4 kb 左右,编码 475 个氨基酸左右,在起始密

码子 ATG 前有一段前导序列。在绿藻 (*Euglena gracilis*) 中也有报道发现在该藻的 *rbcL* 基因中有一段 55 bp 的前导序列^[11-12]。本文所报道的缘管浒苔 *rbcL* 基因的前导序列为 47 bp, 与绿藻 (*Euglena gracilis*) 的 *rbcL* 基因前导序列的长度非常接近;然而龙眼、烟草、水稻高等植物 *rbcL* 基因的前导序列普遍都大于 150 bp, 水稻的甚至长达 313 bp^[13]。缘管浒苔等绿藻的 *rbcL* 基因前导序列远比高等植物的短, 这可能与两者的生长环境有关, 与系统进化地位也有一定的关系。此外, 在高等植物中, 在翻译起始密码子 ATG 上游 -6 ~ -10 bp 处, 有一段类似于原核生物中 SD 序列的片段 (GGAGG), 在高等植物中该片段的功能为核糖体结合位点^[13]; 有研究表明在莱茵衣藻中, *rbcL* 基因起始密码子上游 -20 ~ -23 bp 处有 SD 序列 GAAG, 该序列在 *rbcL* 基因的翻译过程中可以与莱茵衣藻叶绿体 16S rRNA 的 3' 端序列 CUUC 互补^[14]。而 Koo 等^[12] 在 1994 年报道了关于对绿藻 (*Euglena gracilis*) *rbcL* 基因 mRNA 翻译起始区域分析的研究, 文中指出绿藻 (*Euglena gracilis*) *rbcL* 基因 55 bp 的前导序列中没有 SD 序列, 并且翻译起始位点 AUG 前为一段富含 A/U 的序列。在本文获得的缘管浒苔 *rbcL* 基因前导序列中相应的位置上也没有发现这种在高等植物中普遍存在的 SD 序列, 且翻译起始位点 ATG 前是一段富含 A 的序列 (图 4)。这表明缘管浒苔与其他一些绿藻 (例如莱茵衣藻) 以及高等植物 *rbcL* 基因在前导序列上存在着一定的差异, 前导序列与基因的翻译有关, 所以推断两者的 *rbcL* 在基因表达的方式上可能有所不同。在高等植物和一些绿藻中, *rbcL* 基因的表达是通过 SD 序列与叶绿体 rRNA (16S rRNA) 结合, 形成 rRNA-mRNA 复合物, 然后由 30S 核糖体亚基来识别翻译起始位点 AUG; 而在缘管浒苔等一些绿藻中, 由于 *rbcL* 基因没有 SD 序列, 且翻译起始位点 AUG 前为一段富含 A/U 的序列, 因此 AUG 前的那段富含 A/U 的序列二级结构很不稳定, 30S 核糖体亚基可以直接识别翻译起始位点 AUG^[12]。

缘管浒苔与其他 7 种近缘绿藻 *rbcL* 基因的同源性比较结果发现, 蛋白质序列的同源性程度高于核苷酸序列同源性, 说明 Rubisco 大亚基在氨基酸水平上具有相当的保守性, 这可能与其自

身的功能密切相关。

Rubisco 是卡尔文循环的关键酶, 它既参与光合作用又参与光呼吸作用, 调节两者之间的关系, 是所有光合生物进行光合碳同化的关键性酶^[1-2], 在莱茵衣藻的 Rubisco 大亚基中有 3 个催化位点和 1 个 CO₂ 活化肽^[15]。在氨基酸水平上, 与莱茵衣藻 Rubisco 大亚基上的 3 个催化位点与 1 个 CO₂ 活化肽相应的氨基酸序列进行比较, 缘管浒苔 Rubisco 大亚基的相应氨基酸序列与之完全一致, 与菠菜和玉米这两种高等植物的相应序列也仅有个别氨基酸的差别, 这说明 Rubisco 大亚基在功能上具有很高的保守性。

对缘管浒苔 Rubisco 蛋白三级结构预测的结果显示, 缘管浒苔 Rubisco 蛋白三级结构与莱茵衣藻的 Rubisco 蛋白三级结构最为相似, 达到了 92%。这一结果也与 *rbcL* 基因的同源性比较结果以及 *rbcL* 氨基酸序列比对结果相吻合, 说明 Rubisco 在结构域上具有很高的保守性, 并且为进一步深入研究该酶的在光合作用中的功能奠定了基础。

参考文献:

- [1] 何培民, 吴庆磊, 吴维宁, 等. 条浒苔蛋白核超微结构和 Rubisco 及其活化酶分子定位 [J]. 水产学报, 2004, 28(3): 255-260.
- [2] 董晓丽, 周集体, 杜翠红, 等. Rubisco 的分子生物学研究 [J]. 高技术通讯, 2001, 12: 95-97.
- [3] Sugiura M. The chloroplast genome [J]. Plant Molecular Biology, 1992, 19: 149-168.
- [4] 崔杰, 杨谦徐, 徐德昌. 甜菜叶绿体 *rbcL* 基因的克隆与序列分析 [J]. 作物杂志, 2005, 5: 12-13.
- [5] Kono M, Satoh H, Okabe Y. Nucleotide sequence of the large subunit of the ribulose-1, 5-bisphosphate carboxylase/oxygenase from the green alga *Bryopsis maxima* [J]. Plant Mol Biol, 1991, 17: 505-508.
- [6] Ohma M. Recent developments in the seaweed cultivation and industry in Japan [C] // Phang S M, Critchley A T, Ang O P, eds. Advances in seaweed cultivation and utilization in Asia. University of Malaya Maritime Research Center, 2006: 1-20.
- [7] 张寒野, 吴望星, 宋丽珍, 等. 条浒苔海区试栽培及外界因子对藻体生长影响 [J]. 中国水产科学,

- 2006, 13(5): 781-786.
- [8] 徐姗楠, 何培民. 我国赤潮频发现象分析与海藻栽培生物修复作用[J]. 水产学报, 2006, 30(4): 1-9.
- [9] Sambrook J, Fritsch E F, Maniatis T. Molecular Cloning: A Laboratory Manual [M]. 2nd. Cold Spring Harbor Laboratory Press, New York. 1989: 208-212.
- [10] Saitou N, Nei M. The Neighbor-joining method: a new method for reconstructing phylogenetic trees [J]. Molecular Biology Evolution, 1987, 4: 406-425.
- [11] Ginrich J C, Hallick R B. The *Euglena gracilis* chloroplast ribulose-1, 5-bisphosphate carboxylase gene. II. The spliced mRNA and its product [J]. J Biol Chem, 1985, 260: 16162-16168.
- [12] Koo J S, Spremulli L L. Analysis of the translational initiation region on the *Euglena gracilis* chloroplast ribulose-bisphosphate carboxylase/oxygenase (*rbcL*) messenger RNA [J]. Journal Biological Chemistry, 1994, 269: 7494-7500.
- [13] 林同香, 陈振光. 龙眼 *rbcL* 基因结构分析[J]. 园艺学报, 1999, 26(5): 291-296.
- [14] Dron M, Rahire M, Rochaix J D. Sequence of the chloroplast DNA region of *Chlamydomonas reinhardtii* containing the gene of the large subunit of ribulose bisphosphate carboxylase and parts of its flanking genes [J]. Journal of Molecular Biology, 1982, 162: 775-793.
- [15] Chai Y R, Sun Y, Pan W D, *et al.* Isolation and characterization of *rbcL* gene from *Dunaliella salina* [J]. Life Science Journal, 2007, 4(1): 6-12.

Cloning and sequence analysis of the full length cDNA of *rbcL* from *Ulva linza* (Chlorophyceae, Chlorophycophyta)

YING Cheng-qi^{1,2}, YIN Shun-ji¹, LIN Sen-jie³, SHEN Yi¹, HE Pei-min^{1*}

(1. Key Lab of Aquatic Genetic Resources & Aquacultural Ecology Certificated by the Ministry of Agriculture, Shanghai Ocean University, Shanghai 201306, China;

2. East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Shanghai 200090, China;

3. Department of Marine Sciences, University of Connecticut, Groton, CT 06340, USA)

Abstract: Rubisco (Ribulose 1,5-bisphosphate Carboxylase / Oxygenase) is the first key enzyme in the photosynthesis pathway. In this study, the full length cDNA, encoding Rubisco large subunit (RbsL) was cloned from *Ulva linza* for the first time. Partial *rbcL* cDNA sequence of 1 101 bp was first cloned by RT-PCR from total RNA, then, both sequences at 3' and 5' ends were amplified using 3' -RACE and 5' -RACE technologies that yielded 579 bp and 371 bp DNA fragments respectively. A full length cDNA sequence of 1472 bp was obtained from these cloned sequences and was assigned to a new NCBI accession number DQ813496. It contained the 1 425 bp full coding sequence for Rubisco large subunit (RbsL) and a leader sequence of 47 bp, and it encoded a protein of 474 amino acids. Through bioinformatic analysis and BLAST search, the full cDNA sequence in *Ulva linza* showed high homologies to known full length cDNAs of the same gene from other green algae where *rbcL* cDNAs were previously cloned. The identity ranged from 82.07% to 85.78%, and 88.00% to 94.11%, respectively, at DNA and amino acid levels. The highest homology of the protein was obtained between *Ulva linza* and *Chlorella pyrenoidosa*. Based on the newly identified cDNA sequence, multiple alignment was performed by using bioinformatic software among amino acid sequences of *Ulva linza* and other two higher plants and one green alga. Then, the secondary and three-dimensional structures for Rubisco large subunit in *Ulva linza* were also predicted and analyzed with online programs.

Key words: *rbcL*; *Ulva linza*; leader sequence; rapid amplification of cDNA ends (RACE); cDNA

Corresponding author: HE Pei-min. E-mail: pmhe@shou.edu.cn