Isolation and characterization of rbcL gene from *Dunaliella salina*☆

Yurong Chai[1], Yun Sun[1], Weidong Pan[2], Yanlong Jia[2], Lexun Xue[2,*]

[1]*Department of Histology and Embryology, Basic Medical College, Zhengzhou University, Zhengzhou, Henan* 450052, *China*;

[2]*Laboratory for Cell Biology, Zhengzhou University, Zhengzhou, Henan* 450052, *China*

Received September 12, 2006

Abstract

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is a key enzyme in photosynthesis, and its large subunit gene, from the chloroplast genome, encodes the enzyme's catalytic site. The present study was aimed to determine the nucleotide sequence of the large subunit of Rubisco (rbcL) from *Dunaliella salina* (*D. salina*). A pair of degenerate primer was designed according to the highly conserved amino acids of the known rbcL and was used to amplify the rbcL gene from *D. salina* by touch-down PCR (TD-PCR) and rapid amplification of cDNA ends (RACE). The resulting TD-PCR product was 1,347 bp in size, encoding a polypeptide of 449 amino acids. According to this result, 5′ RACE and 3′ RACE were performed to obtain the 5′-end and 3′-end of the rbcL gene. The complete open reading frame of *D. salina* rbcL was long 1,416 bp encoding a polypeptide of 472 amino acids, which shared high similarity of amino acid sequences with those of other species. It can be concluded that the full-length rbcL gene is potential isolated from the chloroplast genomic DNA of *D. salina*. [Life Science Journal. 2007;4(1):6-12] (ISSN: 1097-8135).

Keywords: *Dunaliella salina*; Rubisco; rbcL; degenerate primer; RACE

1 Introduction

Genetic engineering of microalgae has been greatly developed, including prokaryotic microalgae *Cyanobacterium*, and eukaryotic microalgae *Chlamydomonas reinhardtii*, *chlorella*, etc. *Dunaliella salina*, a unicellular green alga, which was one of the most halotolerant eukaryotes, originally described by Dunal in 1938[1], has a thin cellular membrane, and a single, large cup-shaped chloroplast with its photosynthetic thylakoid membranes, pyreniod, and abundance β-carotene globules, but without rigid cell wall. Studies on the genetics of *Dunaliella* have been conducted for decades at a few laboratories[2,3].

Ribulose-1, 5-bisphosphate carboxylase/oxygenase (Rubisco) is a bi-functional enzyme that catalyzes the initial step of photosynthetic carbonic reduction and photo-respiratory carbon oxidation cycles by combining CO_2 and O_2 respectively[4]. Rubisco occurs as a hecadecamer containing of eight large subunits (rbcL) and eight small subunits (rbcS) that assemble into a L8S8 holoenzyme in many eubacteria, cyanobacteria, algae, and higher plants. The rbcL gene encodes the enzyme's catalytic site, and it is conserved in evolution of photosynthetic plants[5]. The nucleotide sequence analysis of the rbcL can promote the studies of the function and structure, expression and regulation of the Rubisco as well as the phylogenetic relationship of *D. salina*. In the present report the gene coding for the large subunit of Rubisco was cloned from the chloroplast DNA (cpDNA) of *D. salina* and compared with other species.

2 Materials and Methods

2.1 Materials

Dunaliella salina strain: UTEX-LB-1644 was purchased from Marine Plant Collection, Texas University (Texas, USA). It was grown in batch cultures of liquid PKS medium at 25 ℃ under continuous irradiance of 150 μmol potons/m^2/s[6].

E. coli JM109 was from our laboratory. Plasmid pMD18-T vector, PCR purification kit, DNA gel extraction kit, Taq polymerase and restriction enzymes were purchased from the TaKaRa Co. Ltd (Dalian, China). Trizol was from Invitrogen (California, USA). AMV First strand cDNA synthesis kit was from Shanghai Sangon Co. Ltd(Shanghai, China), and FirstChoice RLM-RACE kit was from Ambion (Texas, USA).

☆Supported by National Natural Science Foundation of China (No. 30540067 and 30270031). The work was finished in Henan Key Laboratory for Molecular Medicine, Zhengzhou, China.

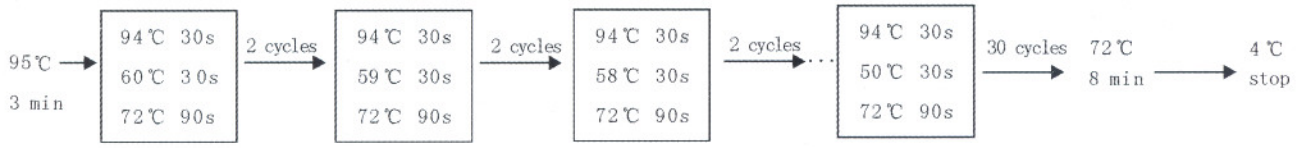
*Corresponding author. Tel: 86-371-6665-8332; Email: lxxue@public.zz.ha.cn

2.2 Methods

2.2.1 Primers. A pair of degenerate primers described below was synthesized based on the conservative regions: GFKAGV and ACEVWK of the rbcL amino acids sequence according to the GenBank data. The forward primer was 5′-GGN TTY AAR GCN GGN GT-3′ and the reverse primer was 5′-YTT CCA NAC YTC RCA NGC -3′ (where N was random base, Y stood for T or C and R was A or G).

2.2.2 Chloroplast DNA extraction and touch-down PCR (TD-PCR) amplification. The chloroplast DNA was isolated from *D. salina* according to Pan[7]. The cells of about 1×10^8 were harvested by centrifugation for 5 minutes at 4 ℃, 5,000 r/min, suspended by 350 μl NET solution (0.1 M NaCl, 50 mM EDTA, 20 mM Tris-HCl, pH 8.0), added 25 μl proteinase K (10 mg/ml) and 25 μl SDS (20 mg/ml), and incubated at 55 ℃ for 2 hours, cooled on ice. Then added 200 μl of 5 M KAc, on ice incubation for 30 minutes, centrifugated for 5 minutes at 4 ℃, 12,000 r/min. The aqueous was extracted three times with phenol and chloroform (1/1 volume), precipitated with ethanol, washed the pellet two times with 70% ethanol, and finally the pellet was resuspended in 30 μl of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

To get better amplification, a modified TD-PCR was performed using about 500 ng cpDNA of *D. salina* as template[8]. PCR amplification was performed using a Pekin Elmer DNA Thermal Cycler 480 (Pekin Elmer Biosystems, California, USA). PCR reaction system contained: template 2 μl, dNTP (each 2.5 mM) 2.5 μl, primers (20 μM)1μl, 10× buffer 5 μl, Taq polymerase (5 u/μl) 0.5 μl, ddH_2O 39 μl, with a final volume of 50 μl. PCR program was carried out as following:



2.2.3 Cloning and analyzing of TD-PCR product. The PCR product was fractioned in 1% agarose gel electrophoresis and then the investigated band was purified according to manufacturer's instruction of the DNA gel extraction kit and ligated to the pMD18-T vector. Competent *E. coli* JM109 cells were transformed with the ligation product, then grown on LB-agar plates containing 100 μg/ml ampicillin, 80 μg/ml X-gal and 80 μg/ml IPTG. White colonies were cultured in a 3 ml LB liquid medium containing 100 μg/ml ampicillin. Plasmid DNA mini-extraction, purification and enzyme digestion were referred to Sambrook[9].

The positive recombinant plasmid DNA was sequenced by the Shanghai Sangon Co. Ltd (Shanghai, China). The amino acid sequence was deduced from the above DNA data with DNAClub program. Sequence analysis was performed using GENTYX (Genetic Information Processing Software, Version5.0) program.

2.2.4 Amplification of rbcL 5′- and 3′-ends. To obtain the rbcL 5′- and 3′-ends of *D. salina*, the total RNA of *D. salina* was extracted with Trizol reagent and rapid amplification of cDNA ends (RACE) was performed by FirstChoice RLM-RACE kit as the protocol described. According to the sequencing result of TD-PCR, gene specific primers (GSP) were designed for 5′ RACE and 3′RACE. The 5′end of rbcL was amplified with GSP1 (5′-CTGGTGGTACACCAGGTTGTGG-3′) and the 3′end of rbcL was amplified with GSP2 (5′-GCACCAGGTGCCGTAGCTAAC-3′). The RACE products were separated by electrophoresis; the target bands were purified by PCR purification kit and ligated to the pMD18-T vector for sequence analysis.

3 Results

3.1 PCR and cloning of the rbcL gene

The agarose gel electrophoresis result of the TD-PCR from cpDNA of *D. salina* was shown in Figure 1. The fragment size of the resulting product was about 1,300 bp, then the PCR product was purified and subcloned into the pMD18-T vector. The recombinant plasmids were identified by *Eco*RI and *Hind* Ⅲ, which included in the multiple cloning sites (MCS) of pMD18-T vector and the correct one was shown in Figure 2. The candidate plasmid was sequenced and named pDS-rbcL.

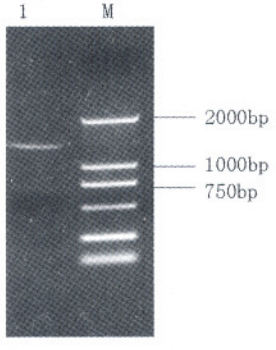


Figure 1. Agarose gel electrophoresis analysis of the PCR product
Lane M: marker, TaKaRa DL2000bp; Lane 1: 1.3 kb PCR product

To obtain the full length DNA sequence of rbcL, RACE was performed to isolate it's 5′end and 3′end. A 350 bp fragment corresponding to the 5′end of the gene was obtained using GSP1 and 5′RACE primer. A fragment about 400 bp was obtained by 3′RACE. The two fragments were then subcloned into the pMD18-T vector, respectively (Figure 3), named pDS-rbcL5, pDS-rbcL3, and sequenced.

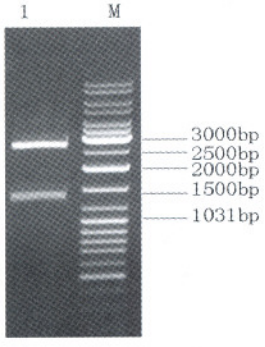


Figure 2. Identification of the recombinant plasmid containing the rbcL fragment
Lane M: marker, Sangon GeneRuler™ DNA Ladder mix; Lane 1: pDS-rbcL/*Eco*RI + *Hind* Ⅲ

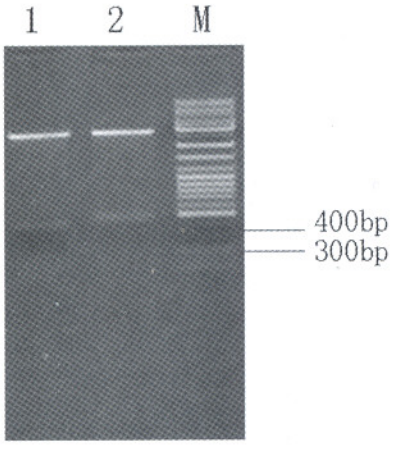


Figure 3. Identification of the recombinant plasmid containing the 5′end and 3′end of rbcL
Lane M: marker, Sangon GeneRuler™ DNA Ladder mix; Lane 1: pDS-rbcL5/*Eco*RI + *Hind* Ⅲ; Lane 2: pDS-rbcL3/*Eco*RI + *Hind* Ⅲ.

3.2 The nucleotide sequences and the deduced amino acid sequences of rbcL

Sequencing result showed that the TD-PCR product from cpDNA of *D. salina* was 1,347 bp long, encoding 449 amino acids residues. Putting the TD-PCR product and 5′end and 3′end fragments together, it was indicated that a complete open reading frame (ORF) of 1,416 bp was obtained. The full length of rbcL has an ORF encoding a putative polypeptide of 472 amino acids residues (Figure 4). The multiple alignments and BLAST analysis revealed that it shared high identities with rbcL gene of other species in nucleotide sequences, and the highest degree of identification was found with *Dunaliella parva* of 97%.

3.3 Comparison of the amino acid sequences

The amino acid sequences of the rbcL from *D. salina* were aligned in Figure 5. A high degree of similarity to other rbcL was obtained, 99% to *Dunaliella parva*, 93% to *Chlamydomonas reinhardtii*, 85% to *Spinacia oleralea* and 82% to *Zea mays*, respectively. In rbcL of *D. salina*, there are three amino acids deletion mutation between K-Lys and Y-Tyr. Besides, the amino acid sequences of rbcL from *D. salina* are totally identified with the published data of rbcL from *D. parva*, which contained 376 amino acids. Other amino acids like D, M, L, S, Q, I, C, etc., Marked by gray boxes in Figure 5 were identical with those in a green alga *Chlamydomonas reinhardtii* but not in higher plants such as spinach and maize, which was consistent to the phylogenetic relationship of the organisms. Sequences Ⅰ, Ⅱ, Ⅲ, Ⅳ were completely identical with those in *Chlamydomonas reinhardtii*, in which Ⅰ-A1 (12 amino acids), Ⅲ-A2 (20 amino acids), Ⅳ-A3 (13 amino acids), determined the sequences of the catalytic sites and Ⅱ-C (17 amino acids) determined the sequences of the CO_2 activator peptide in rbcL of *Chlamydomonas reinhardtii*[10].

Table 1 showed the codon usage in the rbcL gene of *D. salina*. It can be seen that T and A were used more frequently in codons ending than C and G. The preference for T and A as a third base had also been found previously in the rbcL genes of *Chlamydomonas reinhardtii* and *Euglena Gracilis*[10,11] and other chloroplast rbcL gene.

Based on the rbcL alignment from different species, a pair of degenerate PCR primes was designed to clone the gene of rbcL from *D. silina*. Because of the shorter degenerate primers of only 18 base pairs and the expected PCR product of about 1.3 kb in length, it is difficult to clone so long gene fragment of rbcL from the complex chloroplast genomic DNA, just as what we did in previous experiments. However, TD-PCR can avoid mispriming of the shorter primers and improve the specificity of the PCR[12]. In this study, the annealing temperature of first cycle was designed at 60 ℃ to increase the specificity of PCR reaction.

4 Conclusion

In the present study we determined the full length nucleotide sequences of rbcL using TD-PCR and RACE technique from the chloroplast DNA of *D. salina*, a unicellular green alga which can live under the environments with salinity ranging from 0.05 M to 5 M of NaCl. For its unique character,

D. salina would be a promising host of bioreactor and will potentially be a novel system to produce biologically active materials, for example oral vaccine, at lower cost. However, as a stress-tolerance living organism, relative slowly growth rate of *D. salina* made it non-convenient to be genetic modified and there is less known about the mechanism involved in the lower growth rate of it. One possible reason may lie in the photosynthesis. So it is critical to study the genes associated with the mechanism of photosynthesis of *D. salina* and improve the photosynthetic rate.

GCTGATGGCGATGAATGAACACTGCGTTTGCTGGCTTTGATGAAAAACTCCACCGACTACT
GCATTGAGGAGCTCTGTGGGCTACTAGTAGTAGTAAGAACTCTTGGACTTCCTTGAATACC
TAGGCATGCCAGCAACTCTGCTGAGGAAGTCATCTCTTTCGTCTGTTTCTCCTGCATGCTC
AACCATGAGAGTTGTCAACAACAAG

ATG GTA CCA CAA ACT GAA ACT AAA GCT GGT ACT GGA TTT AAG GCT GGT GTA AAAGAT
M V P Q T E T K A G T G F K A G V K D
TAC CGT TTA ACA TAT TAC ACT CCA GAC TAC GTA GTT AGC GAA ACT GAT ATT TTA GCA
Y R L T Y Y T P D Y V V S E T D I L A
GCT TTC CGT ATG ACT CCA CAA CCT GGT GTA CCA CCA GAA GAG TGT GGT GCA GCC GTA
A F R M T P Q P G V P P E E C G A A V
GCA GCT GAG TCA TCA ACA GGT ACA TGG ACT ACA GTA TGG ACT GAC GGT CTA ACA AGT
A A E S S T G T W T T V W T D G L T S
TTA GAC CGT TAC AAA GGT CGT TGT TAC GAT TTA GAA CCT GTA CCA GGG GAA GAA AAT
L D R Y K G R C Y D L E P V P G E E N
CAG TAC ATC GCT TAC GTT GCG TAC CCA ATC GAC CTT TTT GAA GAA GGT TCA GTA ACA
Q Y I A Y V A Y P I D L F E E G S V T
AAC TTA TTC ACT TCA ATT GTA GGT AAC GTA TTC GGT TTC AAA GCG TTA CGT GCA TTA
N L F T S I V G N V F G F K A L R A L
CGT TTA GAA GAC CTT CGT ATT TCA CCA GCT TAC GTT AAA ACA TTC GTT GGA CCA CCT
R L E D L R I S P A Y V K T F V G P P
CAC GGT ATC CAA GTT GAA CGT GAC AAA TAT GGT CGT GGT TTA TTA GGT TGT ACA ATT
H G I Q V E R D K Y G R G L L G C T I
AAA CCA AAA TTA GGT TTA TCA GCT AAA AAC TAC GGT CGT GCT GTT TAC GAA TGT TTA
K P K L G L S A K N Y G R A V Y E C L
CGT GGT GGT TTA GAC TTT ACG AAG GAT GAC GAA AAC GTA AAC TCT CAA CCA TTC ATG
R G G L D F T K D D E N V N S Q P F M
CGT TGG AGA GAC CGT TTC TTA TTC GTA GCT GAA GCT ATT TAC AAA GCA CAA ACA GAA
R W R D R F L F V A E A I Y K A Q T E
ACA GGT GAA ATT AAA GGT CAC TAC TTA AAC TGT ACA GCT GGT ACG TCT GAA GGT ATG
T G E I K G H Y L N C T A G T S E G M
CTT CAA CGT GCA CAA TGT GCT AAA GAA TTA GGT GTA CCA ATT GTA ATG CAT GAC TAC
L Q R A Q C A K E L G V P I V M H D Y
CTA ACT GGT GGT TTC ACA GCA AAC ACT TCA TTA GCA CAT TTC TGT CGT GAC CAC GGT
L T G G F T A N T S L A H F C R D H G
CTT TTA TTA CAC ATT CAC CGT GCG ATG CAC GCT GTA ATT GAC CGT CAA CGT AAC CAC
L L L H I H R A M H A V I D R Q R N H
GGT ATT CAC TTC CGT GTT TTA GCT AAA ACT TTA CGT ATG TCA GGT GGT GAC CAC CTT
G I H F R V L A K T L R M S G G D H L
CAC TCA GGT ACT GTA GTA GGT AAA CTA GAA GGT GAA CGT GAA GTA ACT TTA GGT TTC

H S G T V V G K L E G E R E V T L G F
GTA GAC TTA ATG CGT GAT AAC TTC GTA GAA AAA GAC CGT AGC CGT GGT ATC TAC TTC
V D L M R D N F V E K D R S R G I Y F
ACT CAA GAC TGG TGT TCA ATG CCA GGT GTA ATG CCA GTA GCT TCT GGT GGT ATT CAC
T Q D W C S M P G V M P V A S G G I H
GTA TGG CAC ATG CCA GCT CTA GTT GAA ATT GTC GGT GAT GAC GCT TGT TTA CAA TTC
V W H M P A L V E I V G D D A C L Q F
GGT GGT GGT ACT TTA GGT CAC CCT TGG GGT AAC GCA CCA GGT GCC GTA GCT AAC CGT
G G G T L G H P W G N A P G A V A N R
GTT GCT TTA GAA GCT TGT ACA CAA GCT CGT AAC GAA GGA CGT GAC CTT GCT CGT GAA
V A L E A C T Q A R N E G R D L A R E
GGT GGT AAC GTA ATT CGT TCA GCT TGT AAA TGG TCT CCT GAA TTA GCA GCT GCA TGC
G G N V I R S A C K W S P E L A A A C
GAA GTC TGG AAG GAA ATT AAA TTC GAA TTC GAT ACA GTT GAC AAA TTA TAA
E V W K E I K F E F D T V D K L stop
TATTTCTTTTTTCGTCGGAAAAAACATTTATCAAAGAAAATCTGGGATTTTCTTTGAGAAA
TCTCGAAGATTTCGAATATCCTTTAAGGATATTCGAAACTTCGACATTTAGTTTTTTCCCAA
GGGTCCCCTTCAGAGGCCCGCGTATATTAGAATTTCTTTTCTAATAGGTTAATCCCTACCT
AAAAAAAAAAACCTATAGTGAGTCGTATTAATTCGGATCCGCG

Figure 4. Nucleotide sequences and the deduced amino acids of rbcL from *D. salina* (GenBank Accession No: AY531529)
The primers of TD-PCR were wave-lined. Primers of GSP1and GSP2 were double-lined. The primers of 5′RACE and 3′RACE in the kit were underlined. Start codon and termination codon were boxed.

D.salina	MVPQTETKAGTGFKAGVKDYRLTYYTPDYVVSETDILAAFRMTPQPGVPPEECGAAVAAESSTGTWTTVWT
D.parva	GFKAGVKDYRLTYYTPDYVVSETDILAAFRMTPQPGVPPEECGAAVAAESSTGTWTTVWT
C.reinhardtii	MVPQTETKAGAGFKAGVKDYRLTYYTPDYVVRDTDILAAFRMTPQLGVPPEECGAAVAAESSTGTWTTVWT
S.oleralea	MSPQTETKASVGFKAGVKDYKLTYYTPEYETLDTDILAAFRVSPQPGVPPEEAGAAVAAESSTGTWTTVWT
Z.mays	MSPQTETKASVGFKAGVKDYKLTYYTPEYETKDTDILAAFRVTPQLGVPPEEAGAAVAAESSTGTWTTVWT
D.s	DGLTSLDRYKGRCYDLEPVPGEENQYIAYVAYPIDLFEEGSVTNLFTSIVGNVFGFKALRALRLEDLRISP
D.p	DGLTSLDRYKGRCYDLEPVPGEENQYIAYVAYPIDLFEEGSVTNLFTSIVGNVFGFKALRALRLEDLRISP
C.r	DGLTSLDRYKGRCYDIEPVPGEDNQYIAYVAYPIDLFEEGSVTNMFTSIVGNVFGFKALRALRLEDLRISC
S.o	DGLTNLDRYKGRCYHIEPVAGEENQYICYVAYPLDLFEEGSVTNMFTSIVGNVFGFKALRALRLEDLRIPV
Z.m	DGLTSLDRYKGRCYHIEPVPGDPDQYICYVAYPLDLFEEGSVTNMFTSIVGNVFGFKALRALRLEDLRIPP
	I-A1 II-C

```
D.s   AYVKTFVGPPHGIQVERDK-----YGRGLLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFMR
D.p   AYVKTFVGPPHGIQVERDKLNKYGRGLLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFMR
C.r   AYVKTFAGPPHGIQVERDKLNKYGRGLLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFMR
S.o   AYVKTFQGPPHGIQVERDKLNKYGRPLLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFMR
Z.m   AYSKTFQGPPHGIQVERDKLNKYGRPLLGCTIKPKLGLSAKNYGRACYECLRGGLDFTKDDENVNSQPFMR

D.s   WRDRFLFVAEAIYKAQTETGEIKGHYLNCTAGTSEGMLQRAQCAKELGVPIVMHDYLTGGFTANTSLAHFC
D.p   WRDRFLFVAEAIYKAQTETGEIKGHYLNCTAGTSEGMLQRAQCAKELGVPIVMHDYLTGGFTANTSLAHFC
C.r   WRDRFLFVAEAIYKAQAETGEVKGHYLNATAGTCEEMMKRAVCAKELGVPIIMHDYLTGGFTANTSLAIYC
S.o   WRDRFLFCAEALYKAQAETGEIKGHYLNATAGTCEDMMKRAVFARELGVPIVMHDYLTGGFTANTTLSHYC
Z.m   WRDRFVFCAEAIYKAQAETGEIKGHYLNATAGTCEEMIKRAVFARELGVPIVMHDYLTGGFTANTTLSHYC

                                  Ⅲ-A2

D.s   RDHGLLLHIHRAMHAVIDRQRNHGIHFRVLAKTLRMSGGDHLHSGTVVGKLEGEREVTLGFVDLMRDNFVE
D.p   RDHGLLLHIHRAMHAVIDRQRNHGIHFRVLAKTLRMSGGDHLHSGTVVGKLEGEREVTLGFVDLMRDNFVE
C.r   RDNGLLLHIHRAMHAVIDRQRNHGIHFRVLAKALRMSGGDHLHSGTVVGKLEGEREVTLGFVDLMRDDYVE
S.o   RDNGLLLHIHRAMHAVIDRQKNHGMHFRVLAKALRLSGGDHIHSGTVVGKLEGERDITLGFVDLLRDDYTE
Z.m   RDNGLLLHIHRAMHAVIDRQKNHGMHFRVLAKALRMSGGDHIHSGTVVGKLEGEREITLGFVDLLRDDFIE

D.s   KDRSRGIYFTQDWCSMPGVMPVASGGIHVWHMPALVEIVGDDACLQFGGGTLGHPWGNAPGAVANRVALEA
D.p   KDRSRGIYFTQDWCSMPGVMPVASGGIHVWH
C.r   KDRSRGIYFTQDWCSMPGVMPVASGGIHVWHMPALVEIFGDDACLQFGGGTLGHPWGNAPGAAANRVALEA
S.o   KDRSRGIYFTQSWVSTPGVLPVASGGIHVWHMPALTEIFGDDSVLQFGGGTLGHPWGNAPGAVANRVALEA
Z.m   KDRSRGIFFTQDWVSMPGVIPVASGGIHVWHMPALTEIFGDDSVLQFGGGTLGHPWGNAPGAAANRVALEA

                                  Ⅳ-A3

D.s   CTQARNEGRDLAREGGNVIRSACKWSPELAAACEVWKEIKFEFDTVDKL
D.p
C.r   CTQARNEGRDLAREGGDVIRSACKWSPELAAACEVWKEIKFEFDTIDKL
S.o   CVQARNEGRDLAREGNTIIREATKWSPELAAACEVWKEIKFEFPAMDTV
Z.m   CVQARNEGRDLAREGNEIIKAACKWSAELAAACEIWKEIKFDGFKAMDTI
```

Figure 5. Comparison of the rbcL amino acid sequences from *D. salina* with that from other species.
D.s: *D. salina*; D.p: *D. parva* with 376 amino acids; 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 CO_2 activation region were underlined. The nucleotides in gray boxes were fully identified with those from *Chlamydomonas reinhardtii*. The deletion mutation between K (Lys) and Y (Tyr) were boxed.

Table 1. Codon usage in the rbcL gene of *Dunaliella salina*

TTT-Phe	3 (0.6)	TCT-Ser	4 (0.8)	TAT-Tyr	2 (0.4)	TGT-Cys	11 (2.3)
TTC-Phe	15 (3.2)	TCC-Ser		TAC-Tyr	15 (3.2)	TGC-Cys	1 (0.2)
TTA-Leu	27 (5.7)	TCA-Ser	11 (2.3)	TAA		TGA	
TTG-Leu		TCG-Ser		TAG		TGG-Trp	8 (1.7)
CTT-Leu	7 (1.5)	CCT-Pro	5 (1.1)	CAT-His	2 (0.4)	CGT-Arg	29 (6.1)
CTC-Leu		CCC-Pro		CAC-His	13 (2.8)	CGC-Arg	
CTA-Leu	4 (0.8)	CCA-Pro	16 (3.4)	CAA-Gln	11(2.3)	CGA-Arg	
CTG-Leu		CCG-Pro		CAG-Gln	1(0.2)	CGG-Arg	
ATT-Ile	13 (2.8)	ACT-Thr	16 (3.4)	AAT-Asn	1 (0.2)	AGT-Ser	1 (0.2)
ATC-Ile	4 (0.8)	ACC-Thr		ACA-Thr	13 (2.8)	ACG-Thr	2 (0.4)
ATA-Ile		AAC-Asn	13 (2.8)	AAA-Lys	16 (3.4)	AGA-Arg	1 (0.2)
ATG-Met	10 (2.1)	AGC-Ser	2 (0.4)	AAG-Lys	3 (0.6)	AGG-Arg	
GTT-Val	9(1.9)	GCT-Ala	24(5.1)	GAT-Asp	6(1.3)	GGT-Gly	47(10.0)
GTC-Val	2(0.4)	GCC-Ala	2(0.4)	GAC-Asp	18(3.8)	GGC-Gly	
GTA-Val	25 (5.3)	GCA-Ala	11(2.3)	GAA-Glu	26(5.5)	GCG-Ala	3(0.6)
GTG-Val		GGA-Gly	3 (0.6)	GAG-Glu	2(0.4)	GGG-Gly	1 (0.2)

Numbers in parenthesis indicated per cent of polypeptide (472 amino acids) having this codon.

References

1. Avron M, Ben-Amotz A. *Dunaliella*: Physiology, Biochemistry and Biotechnogy. Florida, CRV Press, 1992.
2. Geng DG, Wang YQ, Wang P, Li WB, Sun YR. Stable expression of hepatitis B surface antigen gene in *Dunaliella salina* (Chlorophuta). J of Applied phycology 2003; 15(6): 451－6.
3. Li J, Ren HW, Huang JH, Yu MM, Ru BG. Transient expression of reporter gene EGFP in transformed *Dunaliella salina*. Chinese J of Biochem Mol Biol 2003; 19(3):338－42.
4. Andersson I, Taylor TC. Structural framework for catalysis and regulation in ribulose-1, 5-bisphosphate carboxylase/oxygenase. Arch Biochem Biophys 2003; 414(2):130－40.
5. McIvor L, Maggs CA, Stanhope MJ. RbcL sequences indicate a single evolutionary origin of multinucleate cells in the red algal tribe Callithamnieae. Mol Phylogenet Evol 2002; 23(3): 433－46.
6. Fisher M, Pick U, Zamir A. A salt-induced 60-kilodalton plasma membrane protein plays a potential role in the extreme halotolerance of the alga *Dunaliella*. Plant physiol 1994; 106: 1359－65.
7. Pan WD, Zhang GX, Yuan BM, Li J, Wang JM, Xue LX. Cloning of 16S RNA sequence of the *Dunaliella salina* chloroplast. Journal of Zhengzhou University (Medical Sciences) 2004; 39(1):9－12.
8. Zhang GX, Yuan BM, Xu PR, Wang JM, Xue LX. Modified touch down PCR efficiently improves specificity of PCR. Life Science Research 2002;6(4):314－7.
9. Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: A Laboratory Manual (2nd). New York: Cold Spring Harbor Laboratory Press, Cold Spring Harbor. 1989.
10. Dron M, Rahire M, Rochaix JD. Sequence of the chloroplast DNA region of *Chlamydomonas reinhardii* containing the gene of the large subunit of ribulose bisphosphate carboxylase and parts of its flanking genes. J Mol Biol 1982; 162: 775－93.
11. Gingrich JC, Hallick R. The *Euglena gracilis* chloroplast ribulose-1, 5-bisphosphate carboxylase gene. J of Biol Chem 1985; 260(30): 16162－8.
12. Fietto JL, DeMarco R, Verjovski-Almeida S. Use of degenerate primers and touch-down PCR for construction of cDNA libraries. Biotechniques 2002; 32:1404－8.