Aquatic Botany 146 (2018) 23–30

Contents lists available at ScienceDirect

Aquatic Botany

journal homepage: www.elsevier.com/locate/aquabot

Identification and characterization of glutamine synthetase isozymes in *Gracilaria lemaneiformis*

Xiaojuan Liu[1], Qingfang Zhang[1], Zhongyan Huan, Mingqi Zhong, Weizhou Chen, Hong Du*

Guangdong Provincial Key Laboratory of Marine Biotechnology and STU-UNIVPM Joint Algal Research Center, College of Sciences, Shantou University, Shantou, 515063, Guangdong, China

ARTICLE INFO

Keywords:
Gracilaria lemaneiformis
Glutamine synthetase
Clone
Bioinformatics
Gene family

ABSTRACT

Glutamine synthetase (GS) isozymes are necessary during the conversion of inorganic to organic nitrogen (N) and they also play a crucial role in the growth of *Gracilaria lemaneiformis*. Six genes encoding GS-protein were found and identified from the transcriptome data of *G. lemaneiformis* genome. No intron was found in the six GS genes. Among these six GS genes, *Unigene0016999* and *Unigene0017002* genes encode two GSI subfamily proteins with 426 and 507 amino acid (AA), respectively. These two GSI proteins are localized to the mitochondria and contain only one conserved Gln-syn C domain. Three GSII subfamily proteins (Unigene0015608, Unigene0022741 and Unigene0002184 with 350 AA, 350 AA and 402 AA respectively) harbor a conserved Gln-synt_C and Gln-synt_N domains. While *Unigene0002194* gene encodes a GSIII protein (735 AA) and is comprised of a conserved Gln-synt_C and a specific GSIII_N domain. Phylogenetic analyses provided strong support for the classification of three GS subfamilies in *G. lemaneiformis* and insights into the evolutionary relationship of GS subfamilies between different organisms.

1. Introduction

Nitrogen (N) plays an important role in the process of plant growth and development. Nitrate and ammonium are important nitrogen forms obtained from environment, but these inorganic nitrogen cannot directly be used as substrate for protein synthesis and a range of secondary metabolites (Guan et al., 2015). The inorganic nitrogen must be processed and assimilated into organic nitrogen glutamine (Gln) through the consecutive action of enzymes glutamine synthetase (GS; EC: 6.3.1.2) and glutamine: 2-oxoglutarate amidotransferase (also termed glutamate synthase, GOGAT; EC: 1.4.7.1), called the GS/GOGAT cycle (Tobin and Yamaya, 2001; Hirel et al., 2005; Lea and Azevedo, 2007). As a key enzyme during the process of the nitrogen assimilation, GS gene superfamily has already been well classified and identified in higher plants, but very few information in algae. Thus, a better identification and characterization of GS family involved in the N assimilation will aid the understanding of their function in algae.

The GS gene family contains three different groups, GSI, GSII, GSIII, all of which varied on the molecular size and the number of subunits (van Rooyen et al., 2006; Ghoshroy et al., 2010). The three GS families are widely found in both prokaryotes and eukaryotes (van Rooyen et al., 2006). Genes encoding GSI mainly exist in Bacteria and Archaea (Robertson and Tartar, 2006a). Among these three groups GSII is divided into two distinct subfamilies based on their subcellular localization. Typically, GS1 is localized in the cytoplasm and encoded by a small multigene family. However, GS2 was found in the chloroplast and encoded by a single nuclear gene in most species (Tanaka et al., 2009; Masclaux-Daubresse et al., 2010; Swarbreck et al., 2011). GSII sequences were isolated from different algae, such as green algae (*Trebouxiophyceae*, *Chlorophyceae* and *Prasinophyta*) and diatom (*Skeletonema costatum* and *Thalassiosira psuedonana*) (Ghoshroy et al., 2010; Chen and Silflow, 1996; Robertson et al., 2001). And previous phylogenetic analyses suggested that the heterokont GSII gene arose via endosymbiotic gene transfer from the nuclear genome of the red – algal endosymbiont to the nuclear genome of the heterotrophic host (Robertson and Tartar, 2006b). Together with the GSII genes in algae, GSII were also found in different plants. For example, 6 genes coding a cytosolic GS isoenzyme (GS1) and 2 genes coding a plastidic GS isoenzyme (GS2) were identified in *Populus trichocarpa* (Castro-Rodríguez et al., 2011). The studies of transgenic poplars demonstrated that GS activity plays a fundamental role in the poplar growth and development, including enhance nitrogen assimilation, enhance vegetative growth and resistance to drought stress (Castro-Rodríguez et al., 2011; Suarez et al., 2002; Cánovas et al., 2007). In addition, three cytosolic

* Corresponding author.
E-mail address: hdu@stu.edu.cn (H. Du).
[1] These authors contributed equally to the article.

https://doi.org/10.1016/j.aquabot.2018.01.006
Received 27 October 2017; Received in revised form 20 January 2018; Accepted 23 January 2018

GS isoforms (*HvGS1_1*, *HvGS1_2* and *HvGS1_3*) were analyzed in barley and the results indicated that GS1 is important in N transport and remobilization, the primary assimilation of N and also in the protection against ammonium toxicity in roots (Goodall et al., 2013). The third type (GSIII) is a member of the last discovered family among the three GS types. Kinoshita S. et al., identified the full − length sequence of GSIII from the diatom *Chaetoceroscompressum* and proposed that eukaryotic GSIIIs are not within a monophyletic relationship with the possible occurrence of lateral gene transfer in GSIII during evolution (Robertson et al., 2001; Kinoshita et al., 2009). Therefore, till now many GS genes were reported in many organisms but not in *G. lemaneiformis*.

The photosynthetic microalgal *G. lemaneiformis* is an important primary producer in ecosystem. It's one of the most important red algae for agar production (Zhou et al., 2013). *G. lemaneiformis* are widely distributed around the world and may be significant for bioremediation (Freile-Pelegrín and Murano, 2005; Zhou et al., 2006). As mentioned above, the activity of GS isozymes was found to be very important in the plant growth and development. However, the lack of identification of the GS genes encoding different isozymes created limitations in the further research of GS isozymes in *G. lemaneiformis*. Therefore, in present study, six putative GS genes were isolated from *G. lemaneiformis* and analyzed the subcellular localization, functional motifs, as well as evolutionary relationship of these GS proteins employing a series of bioinformatic analysis. The results of this study will provide a useful resource for further research on the metabolic function of GS isoforms in *G. lemaneiformis* and closely related species.

2. Materials and methods

2.1. Material

G. lemaneiformis used in this study were collected from Nan'ao Island cultivation field (116.5°E, 23.3°N), Shantou, Guangdong, China. We domesticated the strain for one week in a light incubator. The temperature of light incubator was maintained at 20 ± 1 °C under 12:12 h LD photoperiod and the light intensity was maintained at 50 mol photons $m^{-2} s^{-1}$.

2.2. Design and synthesis of primers

Specific primers (Table 1) were designed according to the GS gene coding sequence annotated in the transcriptome database. All primers used in this work were supplied from Shanghai Sangon Biological Engineering Technology and Service Co., Ltd (Shanghai, People's Republic of China).

2.3. DNA extraction and gene amplification

Approximately 100 mg of algal material was dried using a vacuum freeze dryer, ground to powder with liquid nitrogen and total DNA was extracted using Plant Genomic DNA Kit (Tiangen Biotech Co., China). The total DNA of *G. lemaneiformis* was taken as template for PCR amplification. The elongation time was adjusted according to the size of PCR product (1 min per 1 kb).

Table 1
Oligonucleotides for PCR Cutting sites of restriction enzymes are marked in bold and underline. Protective bases are marked in bold before the restriction enzymes.

Name	Sequence (5′ → 3′)
Unigene0002184-F	**CCC**<u>**AAGCTT**</u>ATGAATGCCGCCTTTGGGTC
Unigene0002184-R	**CCG**<u>**CTCGAG**</u>TTCGGAAACTCCACACACGG
Unigene0002194-F	GATCTGTTTGGTGAGGATGTC
Unigene0002194-R	ATACAGCAGCTGCATGTATTTC
Unigene0015608-F	**CCG**<u>**GAATTC**</u>ATGCCCGTCGAAGCCGGC
Unigene0015608-R	**CCG**<u>**CTCGAG**</u>CTCATCCAAGCAGCAGGTTTG
Unigene0016999-F	**CGC**<u>**GGATCC**</u>ATGCGCCGATGCAAGCTCAT
Unigene0016999-R	**CCG**<u>**CTCGAG**</u>AATCTCATCACTCTTCGCGG
Unigene0022741-F	**CCG**<u>**GAATTC**</u>ATGAACACTTCTATCAAAGCCCT
Unigene0022741-R	**CCG**<u>**CTCGAG**</u>GTCAAGACAGCAAGTCTGCAC

2.4. Bioinformatic analysis of GS genes

Sequences of GS genes were aligned using the DNAMAN software to deduce open reading frames (ORFs) and amino acid residues, and to locate introns within the ORF. Theoretical isoelectric, mass values and physical properties such as stability and hydrophobicity for the protein were predicted and calculated using the ExPASy ProtParam Tool (http://web.expasy.org/protparam/).

Predictions of subcellular localization were achieved using iPSORT prediction (iPSORT Home Page), Target P prediction (http://www.cbs.dtu.dk/services/TargetP), SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP-3.0/), SignalP 4.1 Server (http://www.cbs.dtu.dk/services/SignalP/) and ChloroP 1.1 Server (http://www.cbs.dtu.dk/services/ChloroP/) as well as MitoProt (http://ihg2.helmholtz-muenchen.de/ihg/mitoprot.html). The structure and function of the protein were analyzed by the NCBI CDD database (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) and Expasy PROSITE database (http://prosite.expasy.org). Multiple protein sequences alignment was performed by ClustalX. Neighbor joining (NJ) and maximum parsimony (MP) methods in MEGA6 were used to construct the phylogenetic tree, in which the confidence level of each branch was determined by analyzing 1000 bootstrap replicates. Bootstrap values > 50% were generated.

3. Results

3.1. Gene cloning

The segments of six putative GS genes were generated by standard PCR amplification from genomic DNA of *G. lemaneiformis*. (Primers are shown in Table 1). The size of the amplified products is in accordance with the length of GS genes from transcriptome database (Fig. 1). Among these six genes *Unigene0002194* has the signal at the largest size of about 2000 bp (Fig. 1 lane 2). While the other genes contain a strong signal in the range of 1000 bp to 1200 bp. The sequences of *Unigene0016999* and *Unigene0017002* are conserved in the specific binding area of primers (Fig. S1a and b), therefore we could not distinguish the two products after PCR amplification (Fig. 1 lane 4). The results of DNA sequencing showed that the sequences of these six putative GS genes in

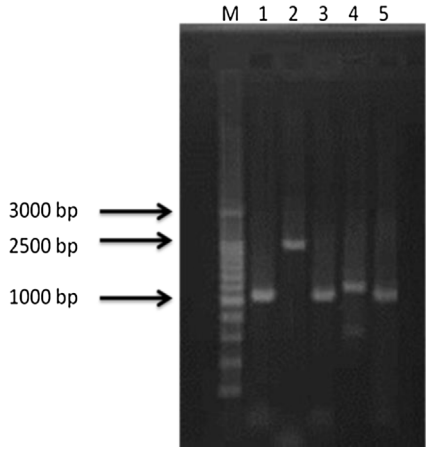

Fig. 1. PCR products of six GS genes.
M: DNA Marker (MD115, Tiangen). Lane 1–5: PCR products of *Unigene0002184*, *Unigene0002194*, *Unigene0015608*, *Unigene0016999/Unigene0017002* and *Unigene0022741*.