Research note: Characterization of a cDNA encoding glutamine synthetase II from *Gelidium crinale* (Rhodophyta)

D. Wilson Freshwater,* David T. Thomas and J. Craig Bailey
Center for Marine Science, University of North Carolina at Wilmington, 5600 Marvin Moss Lane, Wilmington, NC 28409, USA

SUMMARY

A cDNA encoding glutamine synthetase (GS) was characterized from the red alga, *Gelidium crinale* (Turner) Gaillon, using reverse-transcriptase polymerase chain reaction and the 5′- and 3′-rapid amplification of cDNA ends. Sequence analysis of a 1231-bp GS cDNA transcript included both 5′ and 3′ untranslated regions and a 1056-bp open reading frame encoding a 352 amino acid polypeptide. Comparison with GS sequences from other organisms revealed that the *G. crinale* cDNA encodes a type-II GS, and the absence of a N-terminal plastid signal sequence suggests that it is a cytosolic isoenzyme. Phylogenetic analyses of GSII amino acid sequences supports the multiple origin of cytosolic and plastid isoenzymes during eukaryotic evolution.

Key words: cDNA, *Gelidium crinale*, glutamine synthetase, Rhodophyta, RT-PCR, 5′-RACE, 3′-RACE.

Glutamine synthetase (GS, EC 6.3.1.2) is an important primary biological catalyst that is responsible for the initial step at which nitrogen is brought into cellular metabolism. Glutamine synthetase uses ATP to condense ammonia (NH_4^+) with glutamate to form glutamine. Glutamine is subsequently the source of nitrogen in the biosynthesis of many nitrogen-containing metabolites. Three forms of GS have been characterized: (i) GSI, a holoenzyme of 12 polypeptide subunits, is found in prokaryotes (e.g. Eisenberg *et al*. 1987); (ii) GSII, a holoenzyme of eight polypeptide subunits, is found in both prokaryotes and eukaryotes (e.g. Eisenberg *et al*. 1987; Kumada *et al*. 1993); and (iii) GSIII, a holoenzyme of six polypeptide subunits, is found in prokaryotes (e.g. Goodman and Woods 1993). Recently, GSI-like genes have been described from higher plants, but it is unknown if the products of these genes act as glutamine synthetases (Mathis *et al*. 2000).

Characterization of GS genes from eukaryotes has been predominantly limited to those from animals and higher plants. Cytosolic and chloroplast GS genes have been isolated and characterized from the green algae *Chlamydomonas reinhardtii* (Chen and Silflow 1996) and *Dunaliella tertiolecta* (Robertson *et al*. 1999), and the gene for chloroplast GS has been described from the diatom *Skeletonema costatum* (Robertson *et al*. 1999).

Phylogenetic analyses of amino acid sequences for GS from these and other eukaryotic organisms have shown that whereas chloroplast and cytosolic forms of GS from *C. reinhardtii* represent an evolutionarily old divergence, chloroplast and cytosolic GS in higher plants have arisen from a more recent duplication event (Chen and Silflow 1996; Robertson *et al*. 1999). The non-monophyly of chloroplast or cytosolic forms within GS phylogenies indicates that GS isoenzymes have originated multiple times during the evolution of eukaryotes.

In this note, we present the sequence of a cDNA encoding type-II GS from the red alga *Gelidium crinale*. This is the first GS gene characterized from a red alga, and provides additional information towards our understanding of the evolution of nitrogen metabolism pathways in eukaryotes.

Specimens of *G. crinale* were collected from Fort Fisher, New Hanover County, North Carolina, USA and brought to the laboratory in seawater. Non-epiphytized growing tips were washed in a 1% betadine solution and used as starting material for total RNA extractions using the acid guanidinium thiocyanate/phenol/chloroform-based AMBION ToTALLY RNA Kit (Austin, TX, USA).

Reverse-transcriptase polymerase chain reaction (RT-PCR) using the QIAGEN OneStep RT-PCR Kit and protocol (Valencia, CA, USA) was used to amplify a *ca.* 650-bp section of GS. The forward RT-PCR primer was based on the GS01 primer of Robertson *et al*. (1999) and the reverse primer (GS03a) was designed based on an alignment of GS amino acid sequences available from GenBank (Table 1). The *ca.* 650-bp DNA band was isolated in a 1% low-melting-point agarose gel and cleaned with Wizard PCR Preps (Promega, Madison, WI, USA). The PCR product was used as template for sequencing in both directions with the Big Dye Sequencing Kit and primers GS01 and GS03a. The Big Dye Sequencing Kit protocol was followed to set up the

**To whom correspondence should be addressed.*
Email: freshwaterw@uncwil.edu
Communicating editor: G. H. Kim
Received 3 May 2001; accepted 14 August 2001.

Table 1. Primers used for RT-PCR, RACE and genomic DNA amplification and sequencing reactions. Primer position numbers are based on the location of their 5′ end on the sense strand.

Name	Primer sequence	Position
RT-PCR		
GO1	5′-CCY ITA RTG GAA YTA YGA YGG-3′	107
GO3a	5′-CCI GCI CCT TTC CAR TC-3′	775
RACE		
5RACE99	5′-CAC CAT TTG GAG TAT AAG TAT C-3′	333
5RACE78	5′-GTA TCA CAC ATA ACA AGA AGA TTT GG-3′	312
3RACE318	5′-GGT AGA GAT ATT GCT GAT GCA C-3′	552
3RACE348	5′-GCT TGT GTA TAT GCA GGT ATT-3′	582
Genomic		
F-36	5′-TCA GAT TCG TAC ATC CAC T-3′	–36 of 5′UTR
F28	5′-AGT CTC AAC GTA GAT GAA GGA-3′	28
F387	5′-GAA CGT ATT ATG AAT CAA GCA-3′	387
R778	5′-TGA CAA CCA GAA CCA TTC CA-3′	778
R + 43	5′-GGT AGA GAA GAA TCC TCA AGT-3′	+43 of 3′UTR

F, forward primer; R, reverse primer.

sequencing reactions except that 5-times the normal primer concentration was included in reactions. The sequenced region was then used to design nested primer sets for subsequent 5′- and 3′-rapid amplification of cDNA ends (RACE) reactions (Frohman *et al*. 1988).

RACE is a procedure for the amplification of DNA sequences from a messenger RNA (mRNA) template using primer sequences based on a defined internal site and either the natural mRNA poly A tail (3′-RACE) or homopolymeric tails added to the first strand cDNA (5′-RACE). RACE reactions were done using the Gibco-BRL 5′RACE System and 3′ RACE System kits and protocols (Life Technologies, Rockville, MD, USA). The 5′-RACE primer used in first-strand synthesis was 5RACE99, and the nested primer used in subsequent PCR reactions was 5RACE78 (Table 1). The initial and nested gene-specific PCR primers used for the 3′-RACE were 3RACE318 and 3RACE348, respectively. RACE reaction products were isolated and sequenced as previously described.

A GS cDNA sequence contig was generated with SEQUENCHER (version 3.1.1; Gene Codes Corp., Ann Arbor, MI, USA), and the translated amino acid sequence was included within a data set of GS amino acid sequences available from GenBank (Table 2). The amino acid sequences were initially aligned using CLUSTAL W (Thompson *et al*. 1994) and subsequently adjusted by eye in the program SEQ PUP (Gilbert 1992). This alignment is available as a nexus format file at http://snorkel.cmsr.uncwil.edu/pictures/people/wilson/gsalign. Characteristics of the *G. crinale* GS amino acid sequence were determined using MacDNasis (Hitachi Software Engineering Company Ltd, Yokohama, Japan), and phylogenetic analyses were performed using PAUP (version 40b8; Swofford 2001). Sites coded as gaps were treated as missing in parsimony analyses. Parsimony trees were generated from heuristic searches of 1000 random sequence additions, using the MULPARS and STEEPEST DESCENT options with the tree bisection reconnection (TBR) branch-swapping algorithm. Parsimony bootstrap analyses consisted of 1000 replications of simple sequence addition, MULPARS, STEEPEST DESCENT, and TBR branch-swapping. Distance trees were generated from a simple matching mean distance matrix with the neighbor-joining tree-building algorithm. Distance bootstrap values were calculated from 1000 replications of neighbor-joining tree-building.

Table 2. Species and glutamine synthetase (GS) protein sequence accession numbers for aligned GS sequences used for phylogenetic analyses and to develop RT-PCR primers. Where both forms were included, accession numbers for cytosolic and chloroplast isoenzymes are given, respectively.

Species	Accession number
Streptomyces hygroscopicus	AAA26749
Streptomyces viridochromogenes	CAA37028
Frankia sp.	AJKXQ
Bradyrhizobium japonicum	AJZJQ2
Rhizobium meliloti	P45626
Saccharomyces cerevisiae	S61058
Glomerella cingulata	AAB00322
Agaricus bisporus	CAA73235
Schizophyllum commune	AAF27660
Panulirus argus	JN0716
Paracentrotus lividus	JC4027
Opsanus beta	AAD34720
Cricetulus griseus	AAG43362
Pisum sativum	AAA33669, AAA33653
Arabidopsis thaliana	S18601, AAB20558
Oryza sativa	CAA32461, CAA32462
Pinus sylvestris	CAA52448
Chlamydomonas reinhardtii	AAB01817, AAB01818
Dunaliella tertiolecta	AAC77379
Skeletonema costatum	AAC77446
Gelidium crinale	AY034067

In order to determine if introns are present within the gene coding for *G. crinale* GSII, amplification and sequencing primers were designed based on the

previously determined cDNA sequence (Table 1). DNA fragment amplifications were performed using HotStar Taq DNA polymerase following the manufacturer's protocol and genomic DNA extracted as described in Freshwater and Rueness (1994). Amplification products were cleaned, sequenced and sequences assembled as previously described.

Despite the importance of GS as a primary biological catalyst, eukaryotic GS genes have been predominantly characterized from only animals and vascular plants. In this study, a cDNA coding for GSII was characterized from the red alga *G. crinale*.

Sequence analysis of the GSII cDNA indicated a total transcript length of 1231 bp (Fig. 1). The transcript included a 63-bp 5′ untranslated region (UTR), an open reading frame (ORF) of 1056 bp, and a 112-bp 3′ UTR. The normal polyadenylation signal, AATAAA (Hunt 1994), is located 20 bp upstream from the poly A tail. The G + C contents of the 5′ UTR, ORF and 3′ UTR were 30%, 34% and 27%, respectively. The G + C contents showed considerable variation between the first, second and third codon positions within the ORF, and were 52%, 39%, and 8%, respectively. The 1056-bp ORF encodes a 352 amino acid polypeptide with a predicted molecular weight of 39 kDa.

The PCR reactions utilizing *G. crinale* genomic DNA, and various combinations of forward and reverse primer pairs resulted in a single amplification product per reaction. The DNA sequence of the amplified gene was identical to that for the GSII cDNA and contained no introns. Studies of nuclear-encoded protein genes in red algae have generally revealed the presence of short introns near the 5′ ends (e.g. Zhou and Ragan 1994, 1995; Bouget *et al*. 1995), and the presence of these short 5′ splicesomal introns has been suggested as a characteristic feature of nuclear protein-coding genes (Liaud *et al*. 1995). Within eukaryotic GSII genes, intron presence or absence is variable. Although absent in *G. crinale* and the diatom *S. costatum* (Robertson *et al*. 1999), introns are common in mammalian GS genes (e.g. Gaunitz *et al*. 1997).

In order to determine the relationship of the *G. crinale* GSII with those from other organisms, an alignment of 25 amino acid sequences from 21 taxa was constructed. The GSII holoenzyme is composed of eight subunits and the active sites are formed at the subunit interfaces (Eisenberg *et al*. 1987). Four segments of GSII sequences are predicted to be involved in active site formation, based on the conservation of these residues with those for the GSI active site (Eisenberg *et al*. 1987). The predicted active site residues were conserved in the GSII alignment with a minimum pairwise similarity of 52%; the majority of pairwise similarities were greater than 70% for these regions.

Multiple GSII isoenzymes are present in eukaryotes, and those targeted to plastids or mitochondria can generally be identified by the presence of N-terminal

```
   1 tcgtctctatctctttatcacttttcttttcagattcgtacatccacttttgatctta

  59 tttcaatATGGAAAAATCTTTCGATCAACTTTTAAGTCTCAACGTAGATGAAGGATGT
        M  E  K  S  F  D  Q  L  L  S  L  N  V  D  E  G  C
 117 ATTGTTGAATACGTCTGGATTGATGGATATGATGCTGTTCGTTCTAAAGGTCGTTATC
     I  V  E  Y  V  W  I  D  G  Y  D  A  V  R  S  K  G  R  Y
 175 TTAAGAATATTCCATCAAAAGCTTCCGATTTACCTGATTGGAATTATGATGGTAGTTC
     L  K  N  I  P  S  K  A  S  D  L  P  D  W  N  Y [D  G  S  S]
 233 AACTAATCAAGCAAGTGGTGAAGATTCAGAAGTAATTATTAAACCTCAACGTATTTTT
      [T  N  Q  A  S  G  E  D  S  E  V  I  I] K  P  Q  R  I  F
 291 CCAGATCCATTTCGTGGAAAACCAAATCTTCTTGTTATGTGTGATACTTATACTCCAA
     P  D  P  F  R  G  K  P  N  L  L  V  M  C  D  T  Y  T  P
 349 ATGGTGAACCTTTAAAAACAAATACTCGTTTTGAATGTGAACGTATTATGAATCAAGC
     N  G  E  P  L  K  T  N  T  R  F  E  C  E  R  I  M  N  Q  A
 407 AAAAGATTTAAAACCTTGGTTTGGTTTAGAACAAGAATATTTTCTAATTAATCCTAAT
       K  D  L  K  P  W  F  G  L  E  Q  E  Y  F  L  I  N  P  N
 465 ACTGGTAAACCTTTAGGATTTCCTAATGATAAAGATCCTGAACCACAAGGTCCATATT
       T  G  K  P  L  G  F  P  N  D  K  D  P  E  P  Q  G  P  Y
 523 ATTGTGGTGTATCTGGATCTAAAATGTTTGGTAGAGATATTGCTGATGCACATTTTAA
     Y  C  G  V  S  G  S  K  M  F  G  R  D  I  A  D  A  H  F  K
 581 AGCTTGTGTATATGCAGGTATTAAAATTTCAGGTATTAATGCAGAAGTTGCACCTGGA
       A  C  V  Y  A  G  I  K  I  S  G  I  N  A [E  V  A  P  G]
 639 CAATGGGAATTTCAAGTAGGTCCATGTGTTGGTATTGAAGAAGGTGATGAATTATGGA
      [Q  W  E  F  Q] V  G  P  C  V  G  I  E  E  G  D  E  L  W
 697 TGGGAAGATATTTATTACAAAGAATTGCTGAAATGAAAGGTTTGGATGTTAATTATGA
     M  G  R  Y  L  L  Q  R  I  A  E  M  K  G  L  D  V [N  Y  D]
 755 TCCTAAACCTGTTAAAGGTGATTGGAATGGTTCTGGTTGTCATTCTAATTTTTCTACT
      [P  K  P  V  K  G  D  W  N  G  S  G  C  H] S  N  F  S  T
 813 AAACCTATGCGTGAAAAAGATGGTTATCAAAATGCTATTATACCTGCATTAGAAAAAT
       K  P  M  R  E  K  D  G  Y  Q  N  A  I  I  P  A  L  E  K
 871 TGAGTAAGAAACATAAAGAACATATTGCAGCTTATGGTATTGGTAATGAAGATCGACT
     L  S  K  K  H  K  E  H  I  A  A  Y  G  I  G  N  E  D  R  L
 929 TACTGGAAAACATGAAACTGCTAGTATTGAACAATTTAAATGGGGTGTAGCTGATCGT
       T  G  K  H  E  T  A  S  I  E  Q  F  K  W  G  V  A [D  R]
 987 GGTGCTTCATGTCGTGTTGGTCATGATGTAGAAAAGAAAGGTTATGGATATTTTGAAG
      [G  A  S  C  R  V] G  H  D  V  E  K  K  G  Y  G  Y  F  E
1045 ATAGACGTCCTGCTGGTAATTGTGATCCTTATGTTGTTACTATGATGCTTGTTAAGAC
     D  R  R  P  A  G  N  C  D  P  Y  V  V  T  M  M  L  V  K  T
1103 TTGTTGTCTTGAAGATTAGataagaaaatgctttcttgaaagaattgatagccttaag
       C  C  L  E  D  .
1161 aaaacttgaggattcttctctaccatatcttcctttgtaatacgttt<u>aataaa</u>caaat

1219 tgaatttattaccntaaaaaaaaaaaaaaaaaaaaaaaa
```

Fig. 1. DNA nucleotide and amino acid sequence of GSII from *Gelidium crinale*. Nucleotides of the open reading frame (ORF) are shown uppercase, and those of the 5′ and 3′ untranslated regions (UTR) are shown in lowercase. Amino acid residues predicted to form the active site are boxed. The polyadenylation signal within the 3′ UTR is underlined.

plastid or mitochondrial signal sequences (e.g. Robertson *et al*. 1999; Walsh *et al*. 1999). The absence of any signal sequence in the *G. crinale* GS characterized here suggests that it codes for a cytosolic isoenzyme.

The alignment of 25 GSII amino acid sequences from 21 taxa included 475 sites. The total length of the individual amino acid sequences ranged from 329 to 430 residues. The first 76 and last 30 sites of the alignment were excluded from phylogenetic analyses because of the difficulty in determining site homology between some sequences and the presence of chloroplast transit signal sequences in others. The resulting data matrix contained 243 parsimony informative sites. Parsimony analyses yielded a single tree of 1697 steps and a consistency index of 0.637 (Fig. 2). The distance analysis resulted in a tree that differed from the parsimony tree only in the position of branches that received little or no bootstrap support using either method of analysis.

The position of the *G. crinale* GSII sequence and that of the chlorophycean algae cytosolic isoenzymes is equivocal in the phylogenetic trees. Although their sister relationship in the parsimony tree supports the

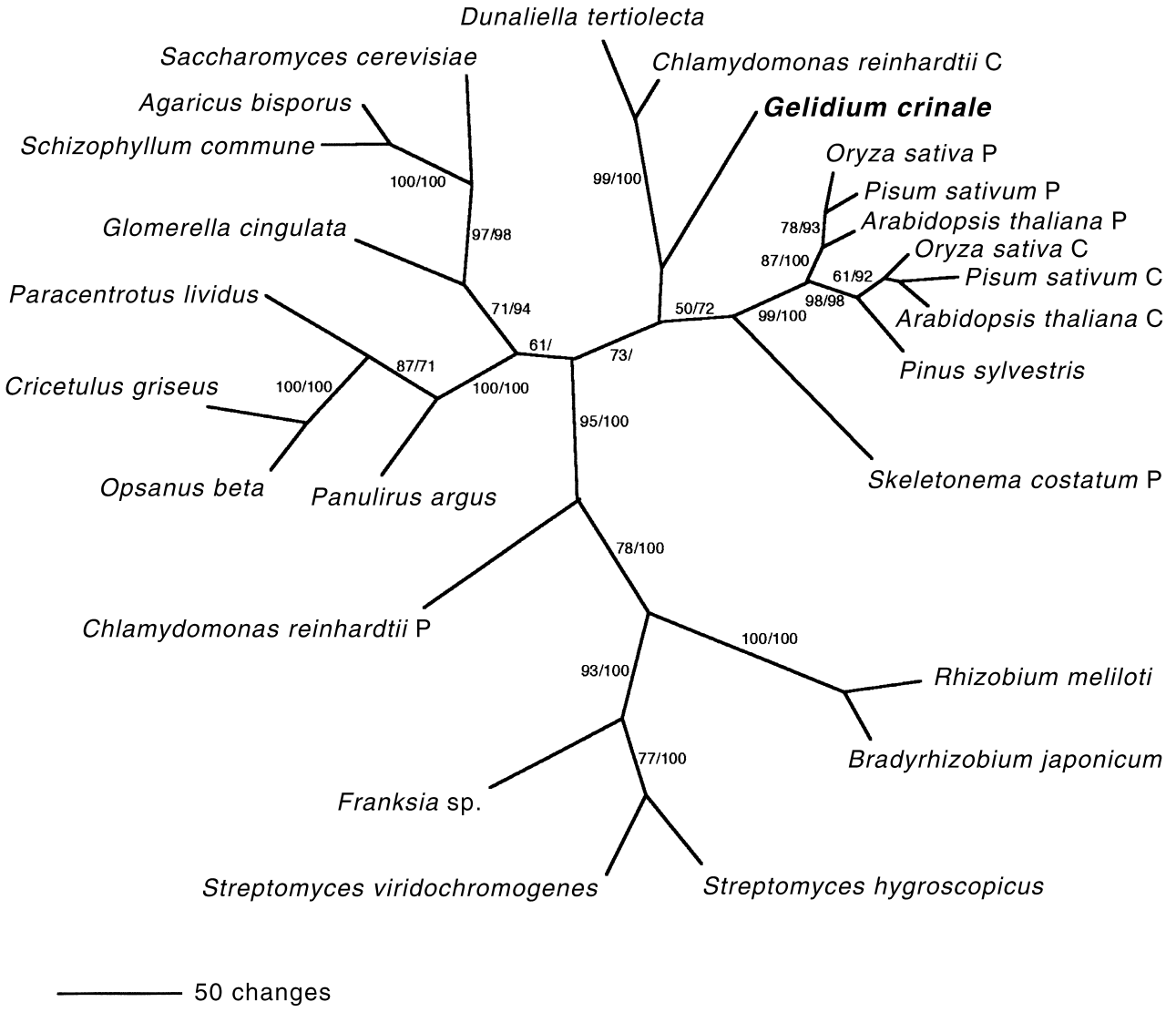

Fig. 2. Unrooted parsimony tree of 1697 steps resulting from analysis of 25 GSII amino acid sequences from 21 taxa. Where cytosolic and plastid isoenzymes are included for a taxon they are indicated by C and P, respectively. Bootstrap proportion values > 50 from parsimony and distance analyses are given for branches in the order parsimony/distance.

assertion of Ragan and Gutell (1995) that red and green algae are closely related, this relationship is not present in the distance tree and received no bootstrap support in either parsimony or distance analyses. The position of the chloroplast isoenzyme sequence for the diatom *S. costatum* receives very weak or no support. The chloroplast isoenzyme from *C. reinhardtii* is within a strongly supported monophyletic clade that contains the prokaryotic GSII sequences. Multiple origins for the chloroplast and cytosolic isoenzymes in eukaryotes are indicated by their varied positions in these taxon-limited trees. Within higher plants, an evolutionarily recent duplication event before the divergence of angiosperms and gymnosperms has occurred. In contrast, the chloroplast and cytosolic isoenzymes of GSII in *C. reinhardtii* diverged before the separation of the Chlorophyta from other eukaryotes. Additional GSII sequences from a variety of eukaryotic taxa are needed in order to better understand the evolutionary pathway of GSII.

ACKNOWLEDGEMENTS

This research was funded by a CMS pilot project grant to JCB and DWF, and NSF grant DEB-9726170 to DWF.

REFERENCES

Bouget, F.-Y., Kerbourc'h, C., Liaud, M.-F. *et al.* 1995. Structural features and phylogeny of the actin gene of *Chondrus crispus* (Gigartinales, Rhodophyta). *Curr. Genet.* **28**: 164–72.

Chen, Q. and Silflow, C. D. 1996. Isolation and characterization of glutamine synthetase genes in *Chlamydomonas reinhardtii*. *Plant Physiol.* **112**: 987–96.

Eisenberg, D., Almassy, R. J., Janson, C. A. *et al*. 1987. Some evolutionary relationships of the primary biological catalysts glutamine synthetase and RuBisCO. *Cold Spring Harb. Symp. Quant. Biol.* **52**: 483–90.

Freshwater, D. W. and Rueness, J. 1994. Phylogenetic relationships of some European *Gelidium* (Gelidiales, Rhodophyta) species, based on *rbc*L nucleotide sequence analysis. *Phycologia* **33**: 187–94.

Frohman, M. A., Dush, M. K. and Martin, G. R. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl Acad. Sci. USA* **85**: 8998–9002.

Gaunitz, F., Gaunitz, C., Papke, M. and Gebhardt, R. 1997. Cis-regulatory sequences from the first intron of the rat glutamine synthetase gene are involved in hepatocyte specific expression of the enzyme. *Biol. Chem.* **378**: 11–8.

Gilbert, D. 1992. SeqPup. Computer program provided by the author. Indiana University, Bloomington, IN, USA.

Goodman, H. J. K. and Woods, D. R. 1993. Cloning and nucleotide sequence of the *Butyrivibrio fibrisolvens* gene encoding a type III glutamine synthetase. *J. Gen. Microbiol.* **139**: 1487–93.

Hunt, A. G. 1994. Messenger RNA 3′ end formation in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **45**: 47–60.

Kumada, Y., Benson, D. R., Hillemann, D. *et al*. 1993. Evolution of the glutamine synthetase gene, one of the oldest existing and functioning genes. *Proc. Natl Acad. Sci. USA* **90**: 3009–13.

Liaud, M.-F., Brandt, U. and Cerff, R. 1995. The marine red alga *Chondrus crispus* has a highly divergent beta-tubulin gene with a characteristic 5′ intron: functional and evolutionary implications. *Plant Mol. Biol.* **28**: 313–25.

Mathis, R., Gamas, P., Meyer, Y. and Cullimore, J. V. 2000. The presence of GSI-like genes in higher plants: support for the paralogous evolution of GSI and GSII genes. *J. Mol. Evol.* **50**: 116–22.

Ragan, M. A. and Gutell, R. R. 1995. Are red algae plants? *Bot. J. Linn. Soc.* **118**: 81–105.

Robertson, D. L., Smith, G. J. and Alberte, R. S. 1999. Characterization of a cDNA encoding glutamine synthetase from the marine diatom *Skeletonema costatum* (Bacillariophyceae). *J. Phycol.* **35**: 786–97.

Swofford, D. L. 2001. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods), Version 4. Sinauer Associates, Sunderland, MA, USA.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–80.

Walsh, P. J., Handel-Fernandez, M. E. and Vincek, V. 1999. Characterization and sequencing of glutamine synthetase cDNA from liver of the ureotelic gulf toadfish (*Opsanus beta*). *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* **124**: 251–9.

Zhou, Y.-H. and Ragan, M. A. 1994. Cloning and characterization of the nuclear gene encoding plastid glyceraldehyde-3-phosphate dehydrogenase from the marine red alga *Gracilaria verrucosa*. *Curr. Genet.* **26**: 79–86.

Zhou, Y.-H. and Ragan, M. A. 1995. Characterization of the nuclear gene encoding mitochondrial aconitase in the marine red alga *Gracilaria verrucosa*. *Plant Mol. Biol.* **28**: 635–46.