Cloning and characterization of ADP-glucose pyrophosphorylase small subunit gene in *Cyperus esculentus* (yellow nutsedge)

C. Cheng[1], J. Hu[1], Y. Zhi[1], J.J. Su[1], X.K. Zhang[2] and B.Q. Huang[1]

[1]Hubei Collaborative Innovation Center for Green Transformation of Bio-Resources,
College of Life Science, Hubei University, Wuhan, China
[2]Key Laboratory of Biology and Genetic Improvement of Oil Crops,
Ministry of Agriculture, Oil Crops Research Institute of the Chinese Academy of
Agricultural Sciences, Wuhan, China

Corresponding authors: X.K. Zhang / B.Q. Huang
E-mail: zhang.xk@139.com / huangbangquan@163.com

ABSTRACT. ADP-glucose pyrophosphorylase (ADPGlcPPase) controls the first committed step of starch synthesis by catalyzing the biosynthesis of ADP-glucose from glucose-phosphate and ATP. It is a tetrameric protein consisting of two small and two large subunits. The small subunits have a catalytic function, while the large subunits regulate the enzyme activity. *Cyperus esculentus* (yellow nutsedge) is a perennial C_4 plant grown from rhizomes and tubers. Previous studies on yellow nutsedge have mostly focused on the morphology and cultivation of tubers, their application in food, and biochemical analyses of the tubers. In this study, the gene encoding the ADPGlcPPase small subunit (*CeAGPS*) in yellow nutsedge was cloned and characterized. The full-length *CeAGPS* cDNA sequence contained an 81-bp 5'-untranslated region (UTR), a 188-bp 3'-UTR, and a 1539-bp open reading frame encoding 512-amino acid residues. The genomic sequence of *CeAGPS* comprises a nine exon-eight intron structure similar to the previously reported cotton and *Arabidopsis thaliana AGPS*

genes. The deduced translation product of the *CeAGPS* gene contained a well-conserved catalytic domain and regulatory elements typical of plant AGPS. Reverse transcriptase polymerase chain reaction amplification of the target gene in various plant parts using gene-specific primers indicated that the expression of *CeAGPS* was most abundant in the tuber, and relatively lower in nutsedge roots.

Key words: *Cyperus esculentus*; ADP-glucose pyrophosphorylase; Small subunit; Gene cloning

INTRODUCTION

ADP-glucose pyrophosphorylase (ADPGlcPPase; EC 2.7.7.27) catalyzes the synthesis of starch in plants and glycogen in bacteria, respectively; in these reactions, ADP-glucose (ADPGlc) is synthesized from glucose-phosphate (Glc-1-P) and ATP, which is then utilized as the glycosyl donor for elongation of α-1,4-glucosidic chains (Ballicora et al., 2004; Preiss, 2009). This reaction is the first committed step of starch synthesis (Preiss, 1982); hence, ADPGlcPPase plays a pivotal role in the carbon metabolism in plants. The physiological importance of ADPGlcPPase in plants has been illustrated by numerous studies (Preiss et al., 1991; Smith et al., 1995; Kavakli et al., 2000; Slattery et al., 2000). Seed-specific inhibition of ADPGlcPPase activity in *Brassica napus* resulted in the inhibition of starch synthesis (Vigeolas et al., 2004). On the other hand, the overexpression of *TaLSU*I-encoding ADPGlcPPase cytosolic large subunit in wheat remarkably enhanced the ADPGlcPPase activity, endosperm starch weight, grain number per spike, and single grain weight (Kang et al., 2013). *Arabidopsis thaliana* mutant TL25 accumulated only 2% starch compared to the wild type because of the lack of both subunits of ADPGlcPPase (Lin et al., 1988). These studies provide biochemical evidence of the strong correlation between starch synthesis in plants and the activity of ADPGlcPPase.

So far, ADPGlcPPases identified from higher plants are all heterotetramers consisting of two α- and two β-subunits (Thorbjornsen et al., 1996; Burgess et al., 1997; Chen et al., 1998; Beckles et al., 2001; Burton et al., 2002). The α-subunit, commonly known as the small subunit, is the catalytic center that interacts with the substrates (Glc-1-P and ATP). The β-subunit, or the large subunit, plays a regulatory role (Ballicora et al., 2004). However Cross et al. (2004) suggested that both subunits of ADPGlcPPase played an equally important regulatory role in potato and maize. The ADPGlcPPase small subunit is highly conserved in the small subunit, and less conserved in the large subunit (Ballicora et al., 2004). Although the integral enzyme activity is a function of both subunits, previous studies have shown that the small subunit alone could form a homotetramer (α_4) and be active under saturating concentrations of substrates and 3-PGA (Ballicora et al., 1995). The ADPGlcPPase small subunit (*AGPS*) gene has been identified in various higher plants (Thorbjornsen et al., 1996; Okita et al., 2001; Rosti et al., 2006; Rösti et al., 2007; Taliercio, 2010; Huang et al., 2013). In barley (Thorbjornsen et al., 1996) and rice (Okita et al., 2001), a single gene can encode two different transcripts for *AGPS*, whereas in maize, two paralogous genes encode two different *AGPS*s (Rösti et al., 2007).

Cyperus esculentus (yellow nutsedge) of the sedge family Cyperaceae is a perennial C_4 plant grown from rhizomes and tubers (Defelice, 2002). The tubers of yellow nutsedge contain large amounts of storage nutrients, including starch (42%), oil (30%), sugar (20%), and protein (9%) (Linssen et al., 1989); therefore, yellow nutsedge could potentially become an ideal model

to describe the carbon partition in non-seed tissues (Turesson et al., 2010). Previous studies on yellow nutsedge have mostly focused on the morphology (Wills et al., 1980) and cultivation (Ayeh-Kumi et al., 2014) of tubers, their application in food (Defelice, 2002), and biochemical analyses of the tubers (Turesson et al., 2010; Manek et al., 2012). In this study, we have reported the cloning and characterization of full-length cDNA and genomic DNA sequences of *CeAGPS* gene from yellow nutsedge. The expression profiles of *CeAGPS* in different tissues were also examined.

MATERIAL AND METHODS

Plants and tissue collection

Tubers of yellow nutsedge were provided by the laboratory of oil crops, College of Life Sciences, Hubei University, China. The tubers were half dipped in double distilled water in a growth chamber (70.80 $\mu m \cdot m^2 \cdot s^{-1}$), which was set at 22°C for germination; the chamber was subjected to a 16 h light/8 h dark cycle. A week later, the germinated leaves and roots were collected, immediately placed in liquid nitrogen, and subsequently stored at -80°C for total RNA extraction. Fresh and tender tubers were collected from plants growing in the field, immediately placed in liquid nitrogen, and then stored at -80°C for total RNA extraction. Plantlets regenerated from calli derived from yellow nutsedge tubers were collected for genomic DNA extraction.

Cloning of *CeAGPS* gene

Total RNA was isolated from different tissues (stored at -80°C) using the TRIzol total RNA extraction reagent (TaKaRa Bio Inc., Otsu, Japan) according to the manufacturer protocols. First strand cDNA was synthesized using 1 µg total RNA and 1 µL ReverTra Ace (100 U; Toyobo, Tokyo, Japan) according to the procedure detailed by the manufacturer. AGPS protein sequences from different higher plants (*Solanum tuberosum*, accession No.: AAA66057; *Nicotiana tabacum*, accession No.: ABD60582; *Fragaria x ananassa*, accession No.: AAS00541; *Gladiolus hybrid cultivar*, accession No.: AHN15416; *Triticum aestivum*, accession No.: AAF61173; and *Hordeum vulgare*, accession No.: AAU06191) were aligned by ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/), and highly conserved sequences were retrieved. Based on these highly conserved sequences, a pair of degenerated primers (*AGPS* conserved primer, Table 1) was designed to amplify the conserved region of the *CeAGPS* gene using cDNA from the leaf, root, and tuber as templates. The PCR reaction was performed in a 25 µL reaction volume, containing 2.5 µL 10X PCR Buffer (Mg^{2+} Plus), 0.5 µL dNTP (2.5 mM each), 0.5 µL of each primer (10 mM), 19.8 µL ddH$_2$O, 0.2 µL 5 U/mL Taq DNA polymerase (Tiangen Biotech, Shanghai, China), and 1 µL cDNA template. The PCR conditions were set as follows: pre-denaturation at 94°C for 4 min, denaturation for 35 cycles at 94°C for 30 s, annealing at 58°C for 30 s, extension at 72°C for 30 s, and a final extension at 72°C for 8 min. 5'-RACE and 3'-RACE were performed using the SMART RACE Amplification kit (Clontech Laboratories, Mountain View, CA) with gene-specific 5'-RACE (GSP1) and 3'-RACE primers (GSP2), respectively, in order to clone the upstream and downstream sequences of *CeAGPS* (Table 1). The PCR system (50 µL) was designed according to the recommendations provided in the kit, and the reaction was set as follows: five cycles at 94°C for 30 s and 72°C for 3 min; five pre-denaturation cycles at 94°C for 30 s, annealing at 68°C for 30 s, and extension at 72°C for 3 min; 27 cycles at 94°C for 30 s and 67°C for 30 s; and a final 72°C extension for 8 min.
Two primers pairs (CDP and GDP, Table 1) were designed based on the 5'- and 3'-RACE

sequences to amplify the full-length *CeAGPS* gene from cDNA and genomic DNA. Reverse transcriptase PCR (for cDNA) conditions included a pre-denaturation step at 94°C for 4 min, denaturation for 35 cycles at 94°C for 30 s, annealing at 56°C for 30 s, and extension at 72°C for 1 min 30 s, and a final extension at 72°C for 8 min. The genomic DNA was amplified under the following reaction conditions: pre-denaturation at 94°C for 4 min, denaturation for 35 cycles at 94°C for 30 s, annealing at 63°C for 30 s, and extension at 72°C for 4 min, and a final extension at 72°C for 8 min. All PCR products were cloned into a pMD-18T vector (TaKaRa Bio Inc.), and sequenced. The sequencing accuracy was ensured by sequencing at least five-independent clones for each PCR product.

Table 1. List of primer sequences.

PCR system	Primer name	Primer sequence (5'-3')	Amplicon size/bp
Core fragment PCR	AGPS conserved FP	ACATGGGCGGCTACAARAAYGARGG	213
	AGPS conserved RP	ATGGAYTAYGAGAAGTTCATCCAGGCC	
5'RACE	GSP1	TTGGGGTTCTCCGGGCTCTGCTGTG	1000
3'RACE	GSP2	CGGCGGCTTCTTCTTCTGTTGCGTATGTTC	850
cDNA full length PCR	CDP-FP	GCACACTCTCCCCTCCTCTGGT	1694
	CDP-RP	TGGCACAGATATATATTGATGAAACAGACA	
Genomic DNA PCR	GDP-FP	GCACACTCTCCCCTCCTCTGGT	4913
	GDP-RP	TGCACATGCTCGCTTATTTCAGTC	
qRT PCR	AGPS qRT PCR FP	GATCGCGGGAGGGCAGAGGAG	400
	AGPS qRT PCR RP	GTTGGGGTTCTCCGGGCTCTGC	
	actin qRT PCR FP	CGCCGCTTAACCCTAAGGCTAACAG	322
	actin qRT PCR RP	TTCTCTTTAATGTCACGGACGATTT	

Footnote: AGPS: ADP-glucose pyrophosphorylase small subunit; RACE: rapid-amplification of cDNA ends; GSP: gene specific primer; CDP: cDNA full length PCR; GDP: genomic DNA PCR.

Sequence analysis

Open reading frame (ORF) Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) was used to perform analyze the ORFs, and to deduce the amino acid sequence. The genomic organization of *CeAGPS* was determined by the Splign software (http://www.ncbi.nlm.nih.gov/sutils/splign/splign.cgi). The theoretical isoelectric point (pI) and molecular weight (Mw) was computed using the Compute pI/Mw tool (http://web.expasy.org/compute_pi/). The nucleotide and derived amino acid sequences of the *CeAGPS* gene were BLAST-searched against the reported *AGPS* gene sequences at the National Center for Biotechnology Information (NCBI) server (http://www.ncbi.nlm.nih.gov/BLAST). Conserved domains of the deduced CeAGPS protein were analyzed using the NCBI CD-Search tool (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). The protein secondary structures were predicted using the SOPMA software (http://npsa-pbil.ibcp.fr/). The phylogenetic tree was constructed using DNAMAN (Lynnon Biosoft, USA)

Analysis of gene expression by quantitative real-time PCR (qRT-PCR)

The mRNA expression profile of the *CeAGPS* gene was evaluated by qRT-PCR, using total RNA extracted from the leaf, root, and tuber of the plants. The constitutively expressed actin gene was used as the internal control to normalize transcript levels. Spanned intron primers (*AGPS* qRT-PCR primer) were designed (Table 1) to detect the levels of expression of *CeAGPS*. qRT-PCR was conducted using the comparative Ct method, according to the manufacturer (Bioer, Hangzhou Bioer Technology Co., Ltd.) instructions: the PCR amplification was

performed in a reaction mix with a final 15 µL volume, comprising 1.2 µL cDNA, 0.2 µL each gene-specific primer, 7.5 µL 19 PCR mix, and 5.9 µL purified water. The program was run using the following optimal thermal cycling conditions: one cycle at 95°C for 10 min, 40 cycles at 95°C for 15 s, 60°C for 1 min, 95°C for 15 s, 60°C for 30 s, and 95°C for 15 s. The experiments were performed in triplicate, and the results were represented as the mean ± standard deviations (SD).

Statistical analysis

Statistical analyses were carried out using the SPSS software (v.18.0; SPSS Inc., Chicago, IL, USA). The relationship between the levels of expression of *AGPS* in different tissues of yellow nutsedge was assessed using Spearman's correlation. Gene transcripts that showed significant ($p < 0.05$) changes in the relative abundance of the *AGPS* gene in different tissues were identified by analysis of variance (ANOVA).

RESULTS AND DISCUSSION

Cloning and sequence analysis of yellow nutsedge *AGPS* gene

A conserved fragment (213-bp) was amplified using mRNA extracted from the leaf, root, and tuber of yellow nutsedge and degenerated primers designed based on the highly conserved sequences in the *AGPS* gene of different species. By aligning and splicing the sequences of the conserved *CeAGPS* region, 3'-, and 5'-RACE PCR product, a 1,808-bp putative full-length cDNA of the *CeAGPS* gene was predicted (Figure 1). A 1694-bp cDNA and a 4913-bp genomic DNA sequence were obtained using full length cDNA primers and genomic DNA primers (CDP and GDP, Table 1). NCBI-ORF Finder analyses indicated that the *CeAGPS* cDNA sequence contained an 81-bp 5'-UTR, a 188-bp 3'-UTR, and a 1539-bp ORF encoding 512 amino acid residues (Figure 1). BLASTX search showed that the 1539-bp ORF sequence shared high identity (E value = 0.0) with the *AGPS* of lotus (GenBank accession No.: AHZ08828.1). Splign analysis revealed that the *CeAGPS* gene has a nine exon-eight intron structure (Figure 2). Yellow nutsedge is a monocot that is more closely related to cereals than to dicots. However, previous studies have indicated that the *AGPS* gene in cereals is generally comprised of ten exons and nine introns (Huang et al., 2013); alternately, *AGPS* in cotton (Taliercio, 2010) and *Arabidopsis* (Hadrich et al., 2012) expressed a nine exon-eight intron structure, similar to that seen in this study.

Two types of *AGPS* have been identified in cereals and other grasses: type 1 contains two different transcripts encoded by a single gene via a classic alternative splicing scheme that uses two diverse first exons, and shares the remaining nine exons, as seen in barley (Thorbjornsen et al., 1996); the second type occurs in maize (Rösti et al., 2007), wherein two paralogous genes encode two homologous transcripts. The alignment of diverse *AGPS*s indicated that the difference mainly occurred at the 5'-end (Rösti et al., 2007); moreover, the analysis of *CeAGPS* identified the same pattern of 5'-end divergence (data not shown). In this study, although RNA from different tissues was used as templates to amplify *AGPS*, only one copy of *CeAGPS* gene was obtained.

```
1      CATGGGGATTTCTATATCACTCTCTCTTTCTCTTTTTCTCGAGCACACTCTCCCCTCCTC
61     TGGTTGCTCTCGATTTTTGCGATGGCGGCAATGGCGGTGACCATGGGGACGTCCGGATAC
                           M  A  A  M  A  V  T  M  G  T  S  G  Y
121    CTCTCGTACAGCACGAACAGGGCGTCACGGTCTGCTCCGGGTGGTAGCATGCCGTCCTTC
       L  S  Y  S  T  N  R  A  S  R  S  A  P  G  G  S  M  P  S  F
181    CTGACGACCGGCCGGACCCCGATCGCGCCGATCGCGGGAGGGCAGAGGAGGAGGCAGTGT
       L  T  T  G  R  T  P  I  A  P  I  A  G  G  Q  R  R  R  Q  C
241    ACCCGTGGTGCCCCCCTCGTTGTCCGGGCCGTCTCGGATTCCCGCGGCTCCCAGACCTGC
       T  R  G  A  P  L  V  V  R  A  V  S  D  S  R  G  S  Q  T  C
301    CTCGATCCCGATGCCAGTAGGAGTGTATTGGGTATAATTCTTGGTGGTGGGGCCGGTACC
       L  D  P  D  A  S  R  S  V  L  G  I  I  L  G  G  G  A  G  T
361    CGTCTGTACCCGCTAACCAAGAAGCGCGCCAAGCCGGCTGTGCCGCTTGGGGCCAACTAC
       R  L  Y  P  L  T  K  K  R  A  K  P  A  V  P  L  G  A  N  Y
421    CGGCTGATTGACATTCCTGTAAGCAACTGCCTGAACAGTAACGTCTCCAAGATATACGTG
       R  L  I  D  I  P  V  S  N  C  L  N  S  N  V  S  K  I  Y  V
481    TTGACACAGTTTAACTCCGCTTCTCTCAACCGTCACCTCTCGTGTGCGTACTCGAGCAAC
       L  T  Q  F  N  S  A  S  L  N  R  H  L  S  C  A  Y  S  S  N
541    ATGGGTGGCTACAAGGACGAAGGGTTCGTGGAGGTGCTGGCCGCACAGCAGAGCCCGGAG
       M  G  G  Y  K  D  E  G  F  V  E  V  L  A  A  Q  Q  S  P  E
601    AACCCCAACTGGTTTCAGGGTACAGCTGATGCAGTTAGGCAGTACCTTTGGCTTTTTGAA
       N  P  N  W  F  Q  G  T  A  D  A  V  R  Q  Y  L  W  L  F  E
661    GAGCACAACATAATGGAGTTCCTGGTGCTAGCAGGTGACCACCTTTACCGTATGGACTAT
       E  H  N  I  M  E  F  L  V  L  A  G  D  H  L  Y  R  M  D  Y
721    GAGAAGTTCATCCAAGCGCACCGTGAGTCTAATGCAGATATCACTGTCGCTGCCCTTCCC
       E  K  F  I  Q  A  H  R  E  S  N  A  D  I  T  V  A  A  L  P
781    ATGGATGAGAAGCGTGCCACTGCATTTGGTCTCATGAAGATCGATGATGAAGGCAAGATC
       M  D  E  K  R  A  T  A  F  G  L  M  K  I  D  D  E  G  K  I
841    ATTGAATTTTCCGAGAAACCCAAAGGGGATGCACTCAAAGCTATGAAGGTCGACACCACA
       I  E  F  S  E  K  P  K  G  D  A  L  K  A  M  K  V  D  T  T
901    ATTCTGGGCCTTAATGAGGAGAGAGCTCAAGAGATGCCCTACATTGCAAGTATGGGCATC
       I  L  G  L  N  E  E  R  A  Q  E  M  P  Y  I  A  S  M  G  I
961    TATGTGTTCAGCAAACACGCAATGCTTCAACTCCTTCGAGACAAATTCCCAGATGCCAAT
       Y  V  F  S  K  H  A  M  L  Q  L  L  R  D  K  F  P  D  A  N
1021   GATTTTGGTAGTGAGGTCATTCCTGGGGCCACCAATGTTGGCATGAGGGTCCAGGCATAC
       D  F  G  S  E  V  I  P  G  A  T  N  V  G  M  R  V  Q  A  Y
1081   CTATTTGACGGCTATTGGGAAGACATTGGGACCATAGAGGCGTTCTATAATGCCAATCTA
       L  F  D  G  Y  W  E  D  I  G  T  I  E  A  F  Y  N  A  N  L
1141   GGCATTACCAAGAAACCTTTCCCTGATTTCAGCTTCTATGACCGATCAGCCCCCATATAT
       G  I  T  K  K  P  F  P  D  F  S  F  Y  D  R  S  A  P  I  Y

1201   ACCCAGCCTCGGTATTTACCCCCATCCAAGATAAAAAGTGCTGATATAACAGAGAGTGTC
       T  Q  P  R  Y  L  P  P  S  K  I  K  S  A  D  I  T  E  S  V
1261   ATCGGTGAGGGCTGCATTATCAAAAATTGCAGGATACACCACTCTGTTGTTGGTCTTCGC
       I  G  E  G  C  I  I  K  N  C  R  I  H  H  S  V  V  G  L  R
1321   TCGTGTGTATCGGAAGGTGCAGTCATTGAGGACTCCCTTCTTATGGGTGCTGATTATTAT
       S  C  V  S  E  G  A  V  I  E  D  S  L  L  M  G  A  D  Y  Y
1381   GAGACTGATAACGACAAGAAGTATCTGGCGGGTGTAGGCGGCATCCCCATCGGGATCGGG
       E  T  D  N  D  K  K  Y  L  A  G  V  G  G  I  P  I  G  I  G
1441   AGGAATGCTCACATTAAGCGGGCTATAATTGATAAAAATGCCCGCATAGGGGATGATGTT
       R  N  A  H  I  K  R  A  I  I  D  K  N  A  R  I  G  D  D  V
1501   AAAATCATTAACAAAGAGGGTGTCCAAGAGGCAGCCAGAGAAACAGACGGATACTTCATC
       K  I  I  N  K  E  G  V  Q  E  A  A  R  E  T  D  G  Y  F  I
1561   AAGAGCGGCATCGTGACTGTCATAAAGGATGCAGTCATACCCAATGGCTGGGTTATATAG
       K  S  G  I  V  T  V  I  K  D  A  V  I  P  N  G  W  V  I  *
1621   TTTTGTACCTTTAGACACTAAATAATTGTGTAGTAATCAATCAAATAAACTCTTATTAAT
1681   ATTAGCTAATATGACAGAGAGAGCATATTGCTGGACTGAAATAAGCGAGCATGTGCAGCT
1741   TGTTAAAAAATTTTGTCTGTTTCATCAATATATATCTGTGCCAAAAAAAAAAAAAAAAAA
1801   AAAAAAAA
```

Figure 1. Full-length yellow nutsedge AGPS cDNA sequence and the deduced amino acid sequence. The initiator codon and stop codon are shown in black boxes; sequences before the initiator codon and after the stop codon are the 5'-UTR and 3'-UTR, respectively.

Figure 2. Schematic presentation of the nutsedge AGPS gene. The exons are shown as solid gray boxes and labeled by exon numbers. Lines between the exons represent the introns. The start codons and stop codons are indicated by bars. The length of the gene is labeled by numbers under the structure. Distances indicated are drawn to scale.

Analysis of the deduced amino acid sequence

The Compute pI/Mw Tool was used to determine the theoretical isoelectric point (pI; 7.08) and molecular weight (Mw; 56,062.00 Da) of the CeAGPS protein. The putative secondary structure of the deduced amino acid sequence is presented in Figure 3. BLASTP analysis revealed that the deduced amino acid sequence of CeAGPS shares the highest identity (87%) with *Nelumbo nucifera* AGPS (accession No.: AHZ08828.1). The amino acid sequences and biochemical properties of AGPS from different plant species are summarized in Table 2. Multiple alignments of the yellow nutsedge CeAGPS protein with those of other species is presented in Figure 4. Analyses with the NCBI CD-Search tool predicted two conserved domains in the Glyco-transf-GTA-type and LbetaH superfamilies (data not shown). Some functional motifs are shown in Figure 4.

Previous studies have indicated that the amino acid sequences encoded by exon 1 of *AGPS* show greater variability (Taliercio, 2010). This phenomenon is also evident in CeAGPS (Figure 4). Cys 81 controls the dimerization of APS1 in *Arabidopsis thaliana* via an intermolecular cysteine bridge between the two small subunits, which may play modulate starch metabolism (Hadrich et al., 2012). Two amino acid residues, glutamine and tryptophan, are believed to be critical for the 3-PGA activation of ADP-glucose pyrophosphorylase in potato tubers. Mutagenesis of either of these two residues dramatically disrupts the allosteric activation of AGPS (Figueroa et al., 2013). These regulatory sites are conserved in yellow nutsedge AGPS (Figure 4), suggesting that this enzyme could potentially be constructed into dimers and activated by 3-PGA, although the actual functions of these specific sites remain to be experimentally verified. ADPGlcPPase is present in the amyloplasts of most dicots, with a dominant portion of ADPGlcPPase activity occurring in the amyloplasts in pea seeds (Smith et al., 1990). Immunogold-labeling of ADPGlcPPase in potatoes revealed that this enzyme was exclusively located in the amyloplast (Kim et al., 1989). Different isoforms of AGPS have been isolated from cereals; the AGPS enzyme from cereal endosperm is strictly cytosolic (Denyer et al., 1996; Rosti et al., 2006). The characteristic plastid-targeting leader sequence motif VSDSQNSQ (Morell et al., 1987) was found to be nearly identical to a sequence of VSDSR(Q)G(N)SQ in CeAGPS (Figure 4), indicating that the CeAGPS could also be located in the plastid.
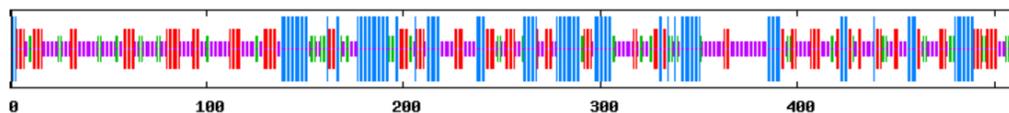

Figure 3. Secondary structure of nutsedge AGPS predicted by SOPMA software. The longest, second longest, third longest, and the shortest lines stand for alpha helices, extended strands, beta turns, and random coils, respectively.

Table 2. Comparison of amino acid sequence homology and biochemical properties of AGPS from different plant species.

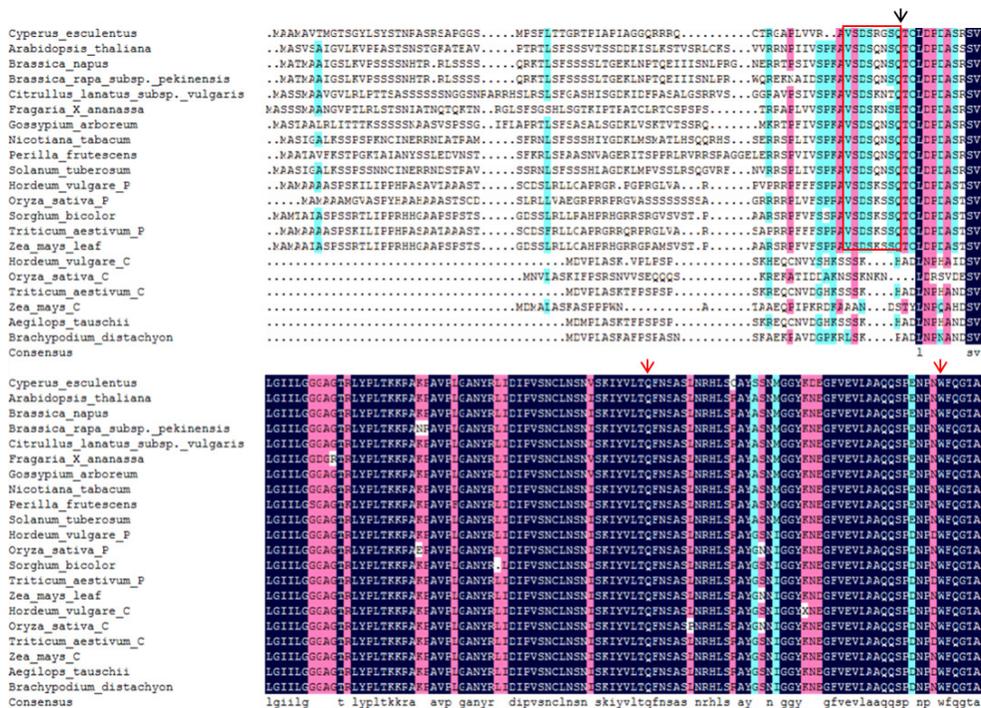| AGPS from different plant species	Accession No.	Number of amino acid	Amino acid		Molecular weight	Isoelectric point
			Similarity (%)	Identity(%)		
Cyperus esculentus	-	512	-	-	56,062.00	7.08
Arabidopsis thaliana	AAB09585.1	520	85.4	76.3	56646.28	6.14
Aegilops tauschii	EMT26197.1	473	80.7	74.4	52089.37	5.54
Brachypodium distachyon	XP_003573790.1	474	81.3	74.8	51997.33	5.51
Brassica napus	CDY00843.1	520	85.2	77.9	57040.69	5.87
Brassica rapa	AAK27684.1	519	84.4	76.7	57083.68	5.87
Citrullus lanatus	AAB91466.1	526	85.2	76.2	57252.80	7.74
Fragaria X ananassa	AAS00541.1	521	84.3	77.5	57106.99	7.19
Gossypium arboreum	KHG13573.1	520	84.2	77.1	56759.41	7.18
Hordeum vulgare	AAU06191.1	513	84.0	77.2	56044.96	6.12
Nicotiana tabacum	ABD60582.1	520	86.5	78.1	57380.14	6.61
Oryza sativa	ACJ86315.1	498	82.1	75.1	54478.80	6.42
Perilla frutescens	AAF66434.1	523	84.5	75.9	57470.24	6.76
Solanum tuberosum	AAF66434.1	521	84.6	77.4	57259.09	6.79
Sorghum bicolor	ABK97531.1	516	84.3	77.0	56335.97	6.94
Triticum aestivum	ACN32622.1	514	84.8	77.2	56285.11	6.42
Zea mays	XP_008664734.1	517	84.9	77.4	56478.21	6.50

Figure 4. Multiple alignments of N-termini of the AGPS protein from various species. The black arrow, red arrow, red box, and red line indicate the dimerization site, 3-PGA activation sites, plastid leading sequence, and Glyco-transf-GTA-type motif, respectively.

Phylogenetic relationships among different plant AGPS proteins

A phylogenetic tree (Figure 5) was constructed based on the multiple sequence alignment of AGPS protein sequences from various species, including rapeseed (*Brassica napus*, accession No.: CDY00843.1), turnip (*Brassica rapa* subsp *pekinensis*, accession No.: AAK27684.1), watermelon (*Citrullus lanatus* subsp *vulgaris*, accession No.: AAB91466.1), yellow nutsedge (*Cyperus esculentus*), strawberry (*Fragaria x ananassa*, accession No.: AAS00541.1), cotton (*Gossypium arboretum*, accession No.: KHG13573.1), barley (*Hordeum vulgare* P, accession No.: AAU06191.1; *Hordeum vulgare* C, accession No.: AAU06190.1), tobacco (*Nicotiana tabacum*, accession No.: ABD60582.1), rice (*Oryza sativa* P, accession No.: ACJ86315.1; *Oryza sativa* C, accession No.: AAA33890.1), purple perilla (*Perilla frutescens*, accession No.: AAF66434.1), potato (*Solanum tuberosum*, accession No.: AAF66434.1), wheat (*Triticum aestivum* P, accession No.: ACN32622.1; *Triticum aestivum* C, accession No.: AAF61173.1), maize (*Zea mays* P, accession v: XP_008664734.1; *Zea mays* C, accession No.: ADD17357.1), sorghum (*Sorghum bicolor*, accession No.: ABK97531.1), brachypodium (*Brachypodium distachyon*, accession No.: XP_003573790.1), goat grass (*Aegilops tauschii*, accession No.: EMT26197.1), and *Arabidopsis* (*Arabidopsis thaliana*, accession No.: AAB09585.1). The phylogenetic tree analysis indicated that all species were clustered in two distinct groups: monocot and dicot (Figure 5). Our result is consistent with those of previous studies (Georgelis et al., 2008; Li et al., 2012); the only difference is that yellow nutsedge AGPS occupies a single branch in the monocot group (Figure 5), which implied that yellow nutsedge may be distinct from other monocots. All monocot species analyzed in this study belonged to the Gramineae family; AGPS was not identified in other Cyperaceae species. The specific differentiation point of AGPS between the grass and sedge families remains to be identified. Li et al. (2012) proposed that a higher evolutionary rate and a greater number of gene duplications were associated with higher starch content in grass seed . The evolution of the starch biosynthetic pathway in cereals and other grasses indicate that the genes encoding the starch synthesis pathway originated from a whole genome-duplication event in an early ancestor of grass (Comparot-Moss et al., 2009); this may partly explain why yellow nutsedge AGPS is clustered out of the grass group.

Tissue-specific expression of *CeAGPS*

Expression of *AGPS* in different tissues, including the leaf, root, and tuber, of yellow nutsedge was examined using qRT-PCR. The highest expression of *CeAGPS* was observed in the tuber; relatively lower *CeAGPS* expression was observed in the leaf and root of yellow nutsedge, in order. The *CeAGPS* expression level in tuber is nearly 17 orders of magnitude higher than that in the root (Figure 6). The expression of *AGPS* was also high in starch-rich potato tubers (Okita and Nakata, 1995). In yellow nutsedge, starch is accumulated as granules in distinct organs for diverse functions. Fixed carbon occurs as transitory starch and is then utilized for sucrose synthesis in the photosynthetic organ, the leaf. Tubers accumulate starch as a storage nutrient (Ballicora et al., 2004), and amyloplasts within the root cell may be related to its gravitropism (Schwuchow et al., 1995). Starch occupies approximately 40% of the dry weight among the carbohydrate reserves in yellow nutsedge tubers (Linssen et al., 1989). *AGPS* is responsible for the rate-limiting step of starch biosynthesis. Therefore, the abundant expression of *CeAGPS* in the tuber is consistent with the need for synthesis of storage starch.
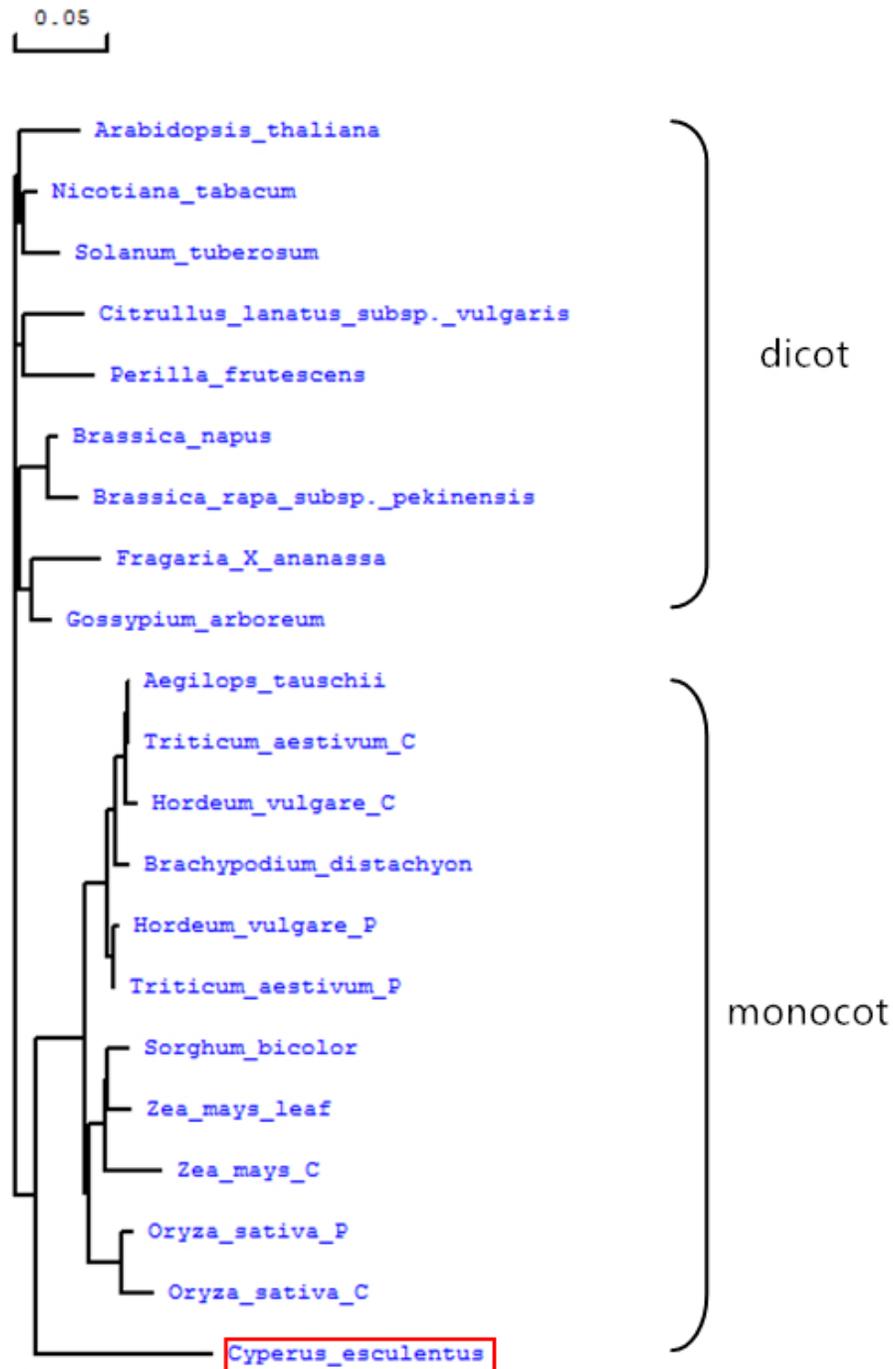
Figure 5. Phylogenetic tree of AGPS proteins from different species. Branches of the phylogenetic tree were labeled with the species name; yellow nutsedge (*Cyperus esculentus*) is indicated by a red box.
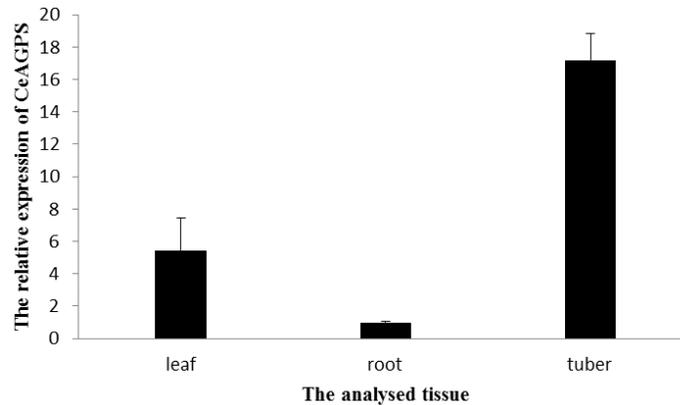
Figure 6. Relative mRNA expression of AGPS yellow nutsedge in different tissues. Total RNA from 3 tissues of yellow nutsedge were used to perform the qPCR, run in triplicates. The expression of yellow nutsedge AGPS was normalized to the amount of actin mRNA. Data are presented as means ± standard deviations (SD).

CONCLUSIONS

In this paper, the full-length cDNA and genomic DNA sequences of the *AGPS* gene in yellow nutsedge were cloned, and the exon-intron structure was analyzed. Subsequently, we characterized the structure and conservative sites in the deduced amino acid sequence; this was used to construct a phylogenetic tree of AGPS among different species. The mRNA expression profile of *CeAGPS* in nutsedge tubers, leaves, and roots was also presented. This study provides a theoretical basis for further research into the possible roles of the *AGPS* gene in yellow nutsedge.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

REFERENCES

Ayeh-Kumi PF, Tetteh-Quarcoo PB, Duedu KO, Obeng AS, et al. (2014). A survey of pathogens associated with *Cyperus esculentus* L (tiger nuts) tubers sold in a Ghanaian city. *BMC Res. Notes.* 7: 1-9.

Ballicora MA, Laughlin MJ, Fu Y, Okita TW, et al. (1995). Adenosine 5'-diphosphate-glucose pyrophosphorylase from potato tuber. Significance of the N terminus of the small subunit for catalytic properties and heat stability. *Plant Physiol.* 109: 245-251.

Ballicora MA, Iglesias AA and Preiss J (2004). AGP;ADP-glucose pyrophosphorylase; a regulatory enzyme for plant starch synthesis. *Photosynth. Res.* 79: 1-24.

Beckles DM, Smith AM and ap Rees T (2001). A cytosolic ADP-glucose pyrophosphorylase is a feature of graminaceous

endosperms, but not of other starch-storing organs. *Plant Physiol.* 125: 818-827.

Burgess D, Penton A, Dunsmuir P and Dooner H (1997). Molecular cloning and characterization of ADP-glucose pyrophosphorylase cDNA clone isolated from pea cotyledon. *Plant Mol. Biol.* 33: 431-444.

Burton RA, Johnson PE, Beckles DM, Fincher GB, et al. (2002). Characterization of the genes encoding the cytosolic and plastidial forms of ADP-glucose pyrophosphorylase in wheat endosperm. *Plant Physiol.* 130: 1464-1475.

Chen B-Y, Janes HW and Gianfagna T (1998). PCR cloning and characterization of multiple ADP-glucose pyrophosphorylase cDNAs from tomato. *Plant Sci.* 136: 59-67.

Comparot-Moss S and Denyer K (2009). The evolution of the starch biosynthetic pathway in cereals and other grasses. *J. Exp. Bot.* 60: 2481-2492.

Defelice MS (2002). Yellow nutsedge *Cyperus esculentus* L.-Snack food of the gods. *Weed Technol.* 16: 901-907.

Denyer K, Dunlap F, Thorbjornsen T, Keeling P, et al. (1996). The major form of ADP-glucose pyrophosphorylase in maize endosperm is extra-plastidial. *Plant Physiol.* 112: 779-785.

Figueroa CM, Kuhn ML, Falaschetti CA, Solamen L, et al. (2013). Unraveling the activation mechanism of the potato tuber ADP-glucose pyrophosphorylase. *PLoS One* 8: e66824.

Georgelis N, Braun EL and Hannah LC (2008). Duplications and functional divergence of ADP-glucose pyrophosphorylase genes in plants. *BMC Evol. Biol.* 8: 232.

Hadrich N, Hendriks JHM, Kotting O, Arrivault S, et al. (2012). Mutagenesis of cysteine 81 prevents dimerization of the APS1 subunit of ADP-glucose pyrophosphorylase and alters diurnal starch turnover in *Arabidopsis thaliana* leaves. *Plant J.* 70: 231-242

Huang B, Hennen-Bierwagen TA and Myers AM (2013). Functions of multiple genes encoding ADP-glucose pyrophosphorylase subunits in maize endosperm, embryo, and leaf. *Plant Physiol.* 164: 596-611.

Kang G, Liu G, Peng X, Wei L, et al. (2013). Increasing the starch content and grain weight of common wheat by overexpression of the cytosolic AGPase large subunit gene. *Plant Physiol. Biochem.* 73: 93-98.

Kavakli IH, Slattery CJ, Ito H and Okita TW (2000). The conversion of carbon and nitrogen into starch and storage proteins in developing storage organs: an overview. *Aust. J. Plant Physiol.* 27: 561-570.

Kim WT, Franceschi VR, Okita TW, Robinson NL, et al. (1989). Immunocytochemical localization of ADPglucose pyrophosphorylase in developing potato tuber cells. *Plant Physiol.* 91: 217-220.

Li C, Li QG, Dunwell JM and Zhang YM (2012). Divergent evolutionary pattern of starch biosynthetic pathway genes in grasses and dicots. *Mol. Biol. Evol.* 29: 3227-3236.

Lin TP, Caspar T, Somerville C and Preiss J (1988). Isolation and characterization of a starchless mutant of *Arabidopsis thaliana* (L.) *Heynh* lacking ADPglucose pyrophosphorylase activity. *Plant Physiol.* 86: 1131-1135.

Linssen JPH, Cozijnsen JL and Pilnik W (1989). Chufa (*Cyperus esculentus*): A new source of dietary fibre. *J. Sci. Food Agri.* 49: 291-296.

Manek RV, Builders PF, Kolling WM, Emeje M, et al. (2012). Physicochemical and binder properties of starch obtained from *Cyperus esculentus*. *AAPS Pharm. Sci. Tech.* 13: 379-388.

Morell MK, Bloom M, Knowles V and Preiss J (1987). Subunit structure of spinach leaf ADPglucose pyrophosphorylase. *Plant Physiol.* 85: 182-187.

Okita TW and Nakata PA (1995). Differential regulation of ADP-glucose pyrophosphorylase in the sink and source tissues of potato. *Plant Physiol.* 108: 361-369.

Okita TW, Sikka VK, Choi S-B, Kavakli IH, et al. (2001). Subcellular compartmentation and allosteric regulation of the rice endosperm ADP glucose pyrophosphorylase. *Plant Sci.* 161: 461-468.

Preiss J (1982). Regulation of the biosynthesis and degradation of starch. *Annu. Rev. Plant Physiol.* 54: 431-454.

Preiss J, Ball K, Hutney J, Smith-White B, et al. (1991). Regulatory mechanisms involved in the biosynthesis of starch. *Pure Appl. Chem.* 63: 535-544.

Rösti S and Denyer K (2007). Two paralogous genes encoding small subunits of ADP-glucose pyrophosphorylase in maize, Bt2 and L2, replace the single alternatively spliced gene found in other cereal species. *J. Mol. Evol.* 65: 316-327.

Rosti S, Rudi H, Rudi K, Opsahl-Sorteberg HG, et al. (2006). The gene encoding the cytosolic small subunit of ADP-glucose pyrophosphorylase in barley endosperm also encodes the major plastidial small subunit in the leaves. *J. Exp. Bot.* 57: 3619-3626.

Schwuchow JM, Kim D and Sack FD (1995). Caulonemal gravitropism and amyloplast sedimentation in the moss *Funaria*. *Can. J. Bot.* 73: 1029-1035.

Slattery CJ, Kavakli IH and Okita TW (2000). Engineering starch for increased quantity and quality. *Trends Plant Sci.* 5: 291-298.

Smith AM, Denyer K and Martin CR (1995). What controls the amount and structure of starch in storage organs? *Plant Physiol.* 107: 673-677.

Smith AM, Quinton-Tulloch J and Denyer K (1990). Characteristics of plastids responsible for starch synthesis in developing pea embryos. *Planta.* 180: 517-523.

Taliercio E (2010). Characterization of an ADP-glucose pyrophosphorylase small subunit gene expressed in developing cotton

(*Gossypium hirsutium*) fibers. *Mol. Biol. Rep.* 38: 2967-2973.

Thorbjornsen T, Villand P, Kleczkowski LA and Olsen OA (1996). A single gene encodes two different transcripts for the ADP-glucose pyrophosphorylase small subunit from barley (*Hordeum vulgare*). *Biochem. J.* 313: 149-154.

Turesson H, Marttila S, Gustavsson KE, Hofvander P, et al. (2010). Characterization of oil and starch accumulation in tubers of *Cyperus esculentus var. sativus* (Cyperaceae): A novel model system to study oil reserves in nonseed tissues. *Am. J. Bot.* 97: 1884-1893.

Vigeolas H, Mohlmann T, Martini N, Neuhaus HE, et al. (2004). Embryo-specific reduction of ADP-Glc pyrophosphorylase leads to an inhibition of starch synthesis and a delay in oil accumulation in developing seeds of oilseed rape. *Plant Physiol.* 136: 2676-2686.

Wills GD, Hoagland RE and Paul RN (1980). Anatomy of yellow nutsedge (*Cyperus esculentus*). *Weed Sci.* 28: 432-437.