CrossMark

# Characterization of a new *granule-bound starch synthase* gene found in amaranth grains (*Amaranthus cruentus* L.)

**Young-Jun Park · Tomotaro Nishikawa ·
Kenichi Matsushima · Kazuhiro Nemoto**

**Abstract** To clarify the mechanism underlying amylose synthesis in the amaranth pericarp, we attempted to identify new *GBSS* isoforms. A new GBSS-encoding gene (i.e., *CrGBSSIb*) was isolated from amaranth leaf tissue. The *CrGBSSIb* gene consists of 4699-bp, including a 1938-bp open reading frame encoding 606 amino acids. A comparison of the cDNA and genomic sequences suggested that *CrGBSSIb* contains 12 introns and 13 exons. Interestingly, a phylogenetic analysis revealed that the amaranth *GBSSIb* gene evolved independently from the other *GBSSI* isoforms within this crop (i.e., intraspecies) and differed from the other plant *GBSSII* genes. The expression patterns of two *GBSS* isoforms revealed that *CrGBSSIb* and *CrGBSSI* are expressed during the early and late phases of seed development, respectively. Additionally, a high *CrGBSSIb* transcript level was detected in leaf tissue. This result indicates that *CrGBSSI* and *CrGBSSIb*, which affect amylose synthesis in amaranth grains, are active in the perisperm and pericarp, respectively. Therefore, *CrGBSSIb* encodes an enzyme associated with amylose synthesis. The enzyme may be primarily responsible for amylose metabolism in pericarp tissue.

**Keywords** Amaranth · Amylose · Granule-bound starch synthase · Pericarp

## Introduction

The genus *Amaranthus* includes approximately 60 species that grow in many areas of the world. All *Amaranthus* are drought-resistant $C_4$ photosynthetic plants that can grow well in saline, alkaline, acidic, or poor soil (Saunders and Becker 1984). This species originated in the New World are an ancient crop that was already under cultivation 5000 to 7000 years ago (Sauer 1967). Ancient amaranth grains still used to this day include the three species, *A. caudatus*, *A. cruentus*, and *A. hypochondriacus*. In recent years, amaranth is getting attention as a food and fodder crop with high seed protein content, a balanced amino acid composition, and high lysine content (Zheleznov et al. 1997). A seed of grain amaranth contains 50–60% starch and have two starch properties of non-waxy and waxy phenotypes (Okuno and Sakaguchi 1981; Saunders and Becker, 1984). This finding substantiated for the first time the presence of both non-waxy and waxy phenotypes in a dicot species.

Y.-J. Park
Rural Research Institute, Korea Rural Community Corporation, 870 Haen-ro Sannok-gu, Ansansi, Gyeonggi-do 15634, Republic of Korea

T. Nishikawa
Ministry of Agriculture, Forestry and Fisheries, 1-2-1 Kasumigaseki Chiyodaku, Tokyo 100-8950, Japan

K. Matsushima · K. Nemoto (✉)
Graduate School of Agriculture, Shinshu University, 8304 Minami-minowa, Kami-ina, Nagano 399-4598, Japan
e-mail: knemoto@shinshu-u.ac.jp

Springer

Generally, the amylose content of storage starch is considered as one of the most important traits in cereals (Morita et al. 2002). Granule bound starch synthase (GBSS), also known as the waxy protein, and it is responsible for the synthesis of amylose in amyloplasts in plant storage organs (Okuno and Sakaguchi 1981; Konishi et al. 1985; Park et al. 2009). Amylose content is a major factor affecting grain quality, and it is controlled almost exclusively by GBSS. There are two types of GBSS isoforms in higher plants based on localization. GBSSI is playing the major role in the storage-tissues such as endosperms and embryos of seeds, while the GBSSII (GBSSIb) producing transitory starch in non-storage tissues such as leaf, stem, root, and pericarp (Cheng et al. 2012). Multiple isoforms of *GBSS* have been reported in many plants (Denyer et al. 1997; Dian et al. 2003; Hirose and Terao 2004; Vrinten and Nakamura 2000). In amaranth, only one isoform of GBSSI has been identified by our previous study (Park et al. 2009). This study showed that the major protein which is responsible for the synthesis of amylose in amaranth perisperm is a GBSSI with the molecular mass of 68 kDa. Recently, we also found that the starch from *A. cruentus* pericarp tissue stained blueblack with I$_2$/KI and thus contained amylose in the initial stage of seed development (Park et al. 2011). That is considered to be an important temporary storage site in pericarp starch at initial developmental stage, as in rice and wheat, and this observation suggests the existence of another *GBSS* isoform such as *GBSSII* or *GBSSIb*.

In this study, to clarify the one responsible for the amylose synthesis in the amaranth pericarp, we will investigate the existence of a new *GBSS* isoform. For this goal, the pattern of starch accumulation was temporally and spatially re-observed during the development of seed by I$_2$/KI staining. Then, new gene encoding *GBSS* (designated *CrGBSSIb*) was isolated and characterized from amaranth leaf. Finally, the expression pattern of the *CrGBSSIb* gene was determined by analyzing the expression in storage- and non-storage tissues at different developmental stages.

The genus *Amaranthus* includes approximately 60 species with a global distribution. All *Amaranthus* species are drought-resistant C$_4$ photosynthetic plants that can grow well in saline, alkaline, acidic, or low-nutrient soil (Saunders and Becker 1984). These species originated as ancient crops that were being cultivated 5000–7000 years ago (Sauer 1967). Ancient amaranth grains still used today include those of *A. caudatus*, *A. cruentus*, and *A. hypochondriacus*. There has recently been increasing interest in amaranth plants as a food and fodder crop because of their high seed protein content, balanced amino acid composition, and high lysine content (Zheleznov et al. 1997). An amaranth seed contains 50–60% starch, which exhibits non-waxy and waxy phenotypes (Okuno and Sakaguchi 1981; Saunders and Becker 1984). This finding revealed for the first time the presence of non-*Cr* waxy and waxy phenotypes in a dicot species.

The amylose content of storage starch is generally considered one of the most important cereal traits (Morita et al. 2002). Granule-bound starch synthase (GBSS), which is also known as the waxy protein, is responsible for the synthesis of amylose in the amyloplasts of plant storage organs (Okuno and Sakaguchi 1981; Konishi et al. 1985; Park et al. 2009). Amylose content is a major factor affecting grain quality and is regulated almost exclusively by GBSS. The two types of GBSS isoforms in higher plants based on localization are GBSSI, which has an important function in storage tissues (e.g., seed endosperm or embryo), and GBSSII (GBSSIb), which produces transitory starch in non-storage tissues, including leaves, stems, roots, and pericarps (Cheng et al. 2012). Multiple GBSS isoforms have been identified in many plants (Denyer et al. 1997; Dian et al. 2003; Hirose and Terao 2004; Vrinten and Nakamura 2000). We detected the only known amaranth GBSSI isoform in a previous study (Park et al. 2009). This 68-kDa enzyme is responsible for the synthesis of amylose in the amaranth perisperm. We also recently observed that the starch from *A. cruentus* pericarp tissue during the initial stage of seed development appears blue-black after staining with I$_2$/KI (iodine-potassium iodide), indicating it contains amylose (Park et al. 2011). The amaranth pericarp is considered an important temporary starch storage site during the initial developmental stage, which is similar to what has been observed in rice and wheat. This suggests the existence of another GBSS isoform such as GBSSII or GBSSIb.

In this study, we investigated whether there is another amaranth GBSS isoform in an attempt to characterize the synthesis of amylose in the pericarp. We used I$_2$/KI staining to analyze the temporal and spatial starch accumulation patterns during seed development. A new *GBSS* gene (i.e., *CrGBSSIb*) was isolated and characterized from amaranth leaves. Finally, the *CrGBSSIb* expression patterns in storage and non-storage tissues at different developmental stages were determined.

## Materials and methods

### Plant materials

We used an *A. cruentus* line (accession number: Ames 22004) obtained from the USDA–ARS, National Genetic Resources Program. Plants were grown in a glasshouse for 90 days, and grains were collected at the following time points, which were selected based on external morphology, fresh weight, and size (Park and Nishikawa 2012): 1–3 days after pollination (DAP) (initial stage; approximately 0.03 mg), 3–5 DAP (early stage; approximately 0.12 mg), 4–8 DAP (early-late stage; approximately 0.26 mg), 8–12 DAP (middle stage; approximately 0.42 mg), 10–15 DAP (mid-late stage; approximately 0.54 mg), and 20 DAP (late stage; approximately 0.65 mg). The average fresh weight of ten seeds was recorded. Additionally, the leaves, petioles, stems, and roots were collected from seedlings at the six-leaf stage. Samples were immediately frozen in liquid nitrogen and stored at − 80 °C until used.

### Microscopic observations

The developing *A. cruentus* seeds were hand-sectioned (approximately 100 μm thick) using a DTK-1000 Microslicer (Dosaka, Japan). Seed cross-sections were stained with an iodine solution (0.1 g resublimated iodine and 0.2 g KI dissolved in 30 mL distilled water) and observed under an Olympus BX51 light microscope or an Olympus SZ dissection microscope.

### Isolation of total RNA and genomic DNA

Total RNA was isolated from plant tissue samples using the RNeasy Mini kit (Qiagen). Genomic DNA was extracted from young leaves using cetyl trimethyl ammonium bromide. The isolated RNA and DNA were quantified with a NanoDrop ND-1000 spectrophotometer (Nano-Drop Technologies, Wilmington, DE, USA). Finally, the RNA and DNA samples were stored at − 80 and 4 °C, respectively.

### Cloning of a new GBSS isoform from amaranth leaves

An alignment of the predicted cDNA sequences for rice *GBSSII* (accession number AY069940), cassava *GBSSII* (accession number AF173900), and wheat *GBSSII* (accession number AF109395) revealed several conserved domains. Two conserved regions were used to design sense (5′-TTCRTTGCAAATGAYTGGCAYAC-3′) and antisense (5′-AGRCCACATGGRTCAAARCT-3′) degenerate primers to amplify a new *GBSS* isoform from complementary DNA (cDNA) prepared using *A. cruentus* leaves. The first-strand *GBSS* cDNA was constructed using the SMARTer RACE cDNA Amplification Kit (Clontech) and then, using the degenerate primers, a 693-bp cDNA fragment was amplified and cloned. Based on the results of reverse transcriptase polymerase chain reaction (RT-PCR) analysis, the 693-bp cDNA fragment was used to design oligonucleotide primers. RT-PCR, 3′- and 5′-rapid amplification of cDNA end (RACE) were performed using the oligonucleotide primers (Electronic Supplementary Material [ESM]; Table 1). The PCR amplification was completed using the 50× Advantage 2 Polymerase Mix, 2.5 mM dNTPs, 10× Advantage 2 PCR Buffer (Clontech), and 10 pmol primers. The annealing temperatures used for each primer combination are listed in Table 1. Some amplified fragments were extracted from agarose gels using the QIAquick Gel Extraction Kit (Qiagen). All cDNA-amplified fragments were ligated into the pGEM-T Easy vector using the TA cloning kit (Promega), and then inserted into *Escherichia coli* strain JM109 cells.

### Genomic PCR amplification

The *GBSSIb* genomic DNA sequence was amplified with gene-specific primers that were designed based on a cDNA sequence analysis (accession number: LC15378) (ESM Table 1). The PCR was conducted using a thermal cycler and EX *Taq* polymerase (Takara). The 50-μL reaction solutions consisted of 2 μL total DNA, 5 μL 10× PCR buffer, 4 μL 2.5 mM dNTPs, 10 pmol primers, and 0.5 μL EX *Taq* polymerase. The PCR program was as follows: 35 cycles of 98 °C for 10 s, 62 °C for 30 s, and 72 °C for 1 min. The PCR products were purified using MultiScreen S384-SEQ Plates (Millipore). To confirm the amplification of the target sequence, 5-μL aliquots of the PCR products were analyzed by 1.5% agarose gel electrophoresis in 0.5× TBE buffer.

### Quantitative real-time PCR analysis

Amaranth grain cDNA was synthesized from 1 μL total RNA (50 ng) using the High Capacity RNA-to-cDNA

Kit (Applied Biosystems). A quantitative real-time PCR (qRT-PCR) analysis was completed using a 48-well plate, and the StepONE™ Real-Time PCR System. The data were analyzed with the StepONE™ (version 2.1) program. All qRT-PCR primers are listed in ESM Table 2. The gene-specific primers were designed based on the *A. cruentus GBSSI* (accession number AB506113) and *GBSSIb* (accession number LC153781) nucleotide sequences. *A. tricolor Actin* gene (AF024716) was used to normalize the amount of gene-specific RT-PCR products. This gene is typically a constitutive gene that is expressed at relatively constant levels in amaranth grain (Park et al. 2011, Park and Nishikawa 2012). All primers were designed using the Primer Express 3.0 software (Applied Biosystems). Transcript abundance was determined using the Power SYBR Green PCR Master Mix (Applied Biosystems). Each reaction consisted of 10 μL Power SYBR Green PCR Master Mix, 200 μM gene-specific primer pair, and 1 μL cDNA for a final volume of 20 μL. The PCR program was as follows: 95 °C for 10 min; 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Each sample was analyzed in triplicate to ensure the results were reproducible. The data were normalized using the values obtained for the *Actin* messenger RNA (mRNA).

Sequencing and phylogenetic analyses

The DNA sequences of the amplified products were determined in both directions with the BigDye Terminator Cycle Sequencing Kit (version 3.1; Applied Biosystems) and a 3100 Genetic Analyzer (Applied Biosystems). The sequencing primer (i.e., 3.2 pmol PCR primer; see Table 2) and dye terminator ready-reaction sequencing premix was added to each template. The sequencing program was as follows: 96 °C for 2 min and 25 cycles of 96 °C for 15 s, 50 °C for 1 s, and 60 °C for 4 min (i.e., dye terminator reactions). Sequences were assembled using the CodonCode Aligner program (CodonCode Co., Dedham, MA) to yield full-length genes. The nucleotide sequences were aligned using the ClustalX module of Geneious Pro 7.1.9 (Biomatters). Protein molecular weights were predicted using the Geneious Pro program. Additionally, BLAST analyses were completed using the default settings of an online tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The transit peptide and cleavage site were predicted using the ChloroP server (http://www.cbs.dtu.dk/services/ChloroP) (Emanuelsson et al. 1999).

The GBSS amino acid sequences from various plant species that were used for a phylogenetic analysis were obtained from the NCBI database. These sequences were aligned using the ClustalX program (Thompson et al. 1997). A phylogenetic tree was constructed using the neighbor-joining method (Saitou and Nei 1987) of the PAUP* 4.0 program (Swofford 1988). The bootstrap confidence values were obtained from 1000 replicates (Felsenstein 1985).

## Results and discussion

Spatiotemporal patterns of starch deposition

The spatiotemporal starch accumulation patterns in developing seeds were observed by staining with $I_2$/KI (Fig. 1). Starch granules were observed in the pericarp in the initial developmental stage (3 DAP). A few starch granules were detected in the perisperm in the early-late developmental stage (8 DAP), during which the pericarp starch contents rapidly decreased. Starch granules were distributed throughout the perisperm in the mid-late developmental stage (15 DAP). Similar results were reported for other cereal crops, including barley, rice, and sorghum (Yu et al. 2015). Starch granules in the pericarp are synthesized during the early seed developmental stages but are absent in mature seeds. We recently reported that starch deposits in the perisperm of developing amaranth seeds are detectable only after the initial developmental stage (Park et al. 2011). Prior to this stage, the pericarp is the major site of starch deposition. A recent study suggested that GBSSII isoforms are responsible for amylose synthesis in pericarps (Hirose and Terao 2004).

Amplification of a new *GBSS* isoform from amaranth leaves

To generate new amaranth *GBSS* isoforms, we used RT-PCR to amplify a 693-bp fragment spanning the N-terminal part of the enzyme. The RT-PCR degenerate primers (ESM Table 1) were based on the conserved sequence near the N-terminal of all known GBSSII enzymes. A BLAST search using the amplicon sequence as a query revealed that the closest matches were with available plant *GBSS* sequences. This result indicates that the amplified fragment was a partial sequence encompassing the conserved amaranth *GBSS* sequence.

**Fig. 1** Starch granules stained with I2/KI in a cross-section of a developing seed. *P* pericarp *Ps* perisperm

Innitial (3 DAP)

Earl-late (8 DAP)    Mid-late (15 DAP)
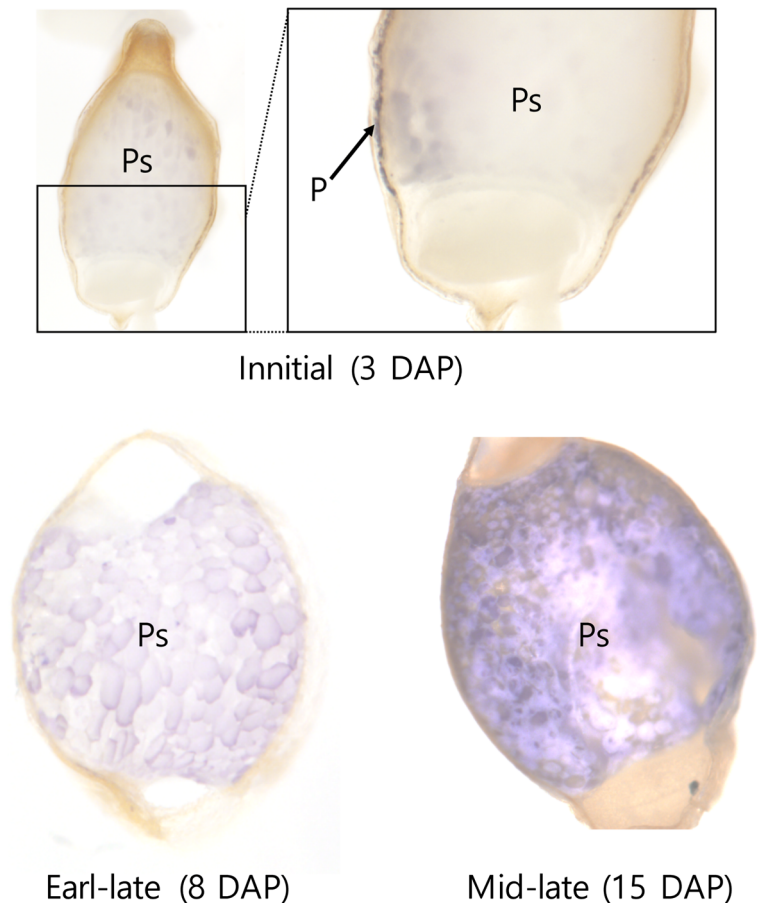
These fragments were then used to design gene-specific primers to amplify new *GBSS* isoforms from amaranth grains.

Isolation and characterization of *GBSSIb*

Further studies of amaranth starch will require the characterization of the *GBSS* gene primary structures. In this study, the full-length *CrGBSSIb* cDNA clones were identified from a cDNA library prepared from mRNA extracted from amaranth leaves. The clone sequences are provided in Fig. 2 and have been deposited in the DDBJ database (accession number: LC153781; http://www.ddbj.nig.ac.jp/). The *CrGBSSIb* cDNA sequence consists of 1938-bp, including a 58-bp 5′-untranslated region, a 1830-bp open reading frame, and a 50-bp 3′-untranslated region. This gene exhibits relatively low sequence identity (66%) with the amaranth *GBSSI* coding sequences (Park et al. 2009).

We analyzed the *GBSSIb* structure by comparing its genomic sequence with that of the *A. cruentus GBSSI* gene (Park et al. 2009). The *CrGBSSIb* gene contains 12 introns and 13 exons (ESM Fig. 1). The genomic DNA clones consist of 4699-bp, of which 2759-bp represent the introns. The intron and exon structures are similar to those of the amaranth *GBSSI* gene. All of the introns comply with the universal GT-AG rule (Breathnach and Chambon 1981) and contain AT-rich regions similar to those present in the genes of other dicot plants (Campbell and Gowri 1990; van der Leij et al. 1991).

Characterization of the GBSSIb enzyme

The structures of the enzyme encoded by *CrGBSSIb* were predicted based on the exon sequences (Fig. 2). The deduced CrGBSSIb enzyme consists of 609 amino acids, with an estimated molecular weight of approximately 66.5 kDa. A putative transit peptide, from amino acids 1–63, was identified using the

```
1          10          20          30          40          50          60          70          80          90          100
ATGGCATCTGTTGCCACTTCTCACTTTGTTTTCTAGAAGTGCACATTCTGCAGTTTCGGAAAGTAAAGCCGCTTTGGCTCAGATTGACCTTAAAACCCGACTC
M  A  S  V  A  T  S  H  F  V  S  R  S  A  H  S  A  V  S  E  S  K  A  A  L  A  Q  I  D  L  K  T  R  L

110         120         130         140         150         160         170         180         190         200
TGCACTCATAATGGGTTGAGGTCATTAAACAAAGCTGACATGCTACAGATGAAATCCAATGCAAAGGCATTGACTAAGAGTGTAAGAAGCAAAGTGAGCAAA
C  T  H  N  G  L  R  S  L  N  K  A  D  M  L  Q  M  K  S  N  A  K  A  L  T  K  S  V  R  S  K  V  S  K

210         220         230         240         250         260         270         280         290         300
ACCCAGAATGATCTGCCACACGGAGCTACTATTCGTGGAATGAACATTGCGTTCGTTGCAACTGAATGTGATAAGTTTTGCAAAACTGGTGGTCTTGGTGAT
T  Q  N  D  L  P  H  G  A  T  I  R  G  M  N  I  A  F  V  A  T  E  C  D  K  F  C  K  T  G  G  L  G  D

310         320         330         340         350         360         370         380         390         400
GTTGTTGGTGGACTCCCACCAGCTTTGGCTAAAAGAGGACATAGAGTAATGGCAATCATCCCTCGCTACGACCAATACAAAGATGCATGGGACACCAATGAA
V  V  G  G  L  P  P  A  L  A  K  R  G  H  R  V  M  A  I  I  P  R  Y  D  Q  Y  K  D  A  W  D  T  N  E

410         420         430         440         450         460         470         480         490         500         510
CTCGTTGAGTTGAAAGTTGGAGACTCACTTTTCACGGTCCGTCTTCTTCCATACTTACAAAGAGGTGTAGACATTGTTTTCGTGGACCATCCAATGTTCCTT
L  V  E  L  K  V  G  D  S  L  F  T  V  R  L  F  H  T  Y  K  R  G  V  D  I  V  F  V  D  H  P  M  F  L

520         530         540         550         560         570         580         590         600         610
GAGAAGGTATGGGGAAAAACGGGCTCAAAGCTCTATGGTCCTACAGCTGGATCAGATTTTGATGACGACCAAATTCGTTTCAGTTTGCTCTGCCAGGCAGCT
E  K  V  W  G  K  T  G  S  K  L  Y  G  P  T  A  G  S  D  F  D  D  N  Q  I  R  F  S  L  L  C  Q  A  A

620         630         640         650         660         670         680         690         700         710
TTGGAGGCTCCTATAGTTTTGAATTTGAACAACAATCCAAACTTCTCTGGCCCCTATGGAGAGGACATTGTGTTTGTTGCCAACGACTGGCACACTGGCCTG
L  E  A  P  I  V  L  N  L  N  N  N  P  N  F  S  G  P  Y  G  E  D  I  V  F  V  A  N  D  W  H  T  G  L

720         730         740         750         760         770         780         790         800         810
CTTCCATGCTACCTCAAGTCCATGTATCAATCTAAGGGTATCTACCAAAATGCCAAGGTTGTCTTCTGTATCCACAACATAGCTTACCAAGGGAGATTCGCC
L  P  C  Y  L  K  S  M  Y  Q  S  K  G  I  Y  Q  N  A  K  V  V  F  C  I  H  N  I  A  Y  Q  G  R  F  A

820         830         840         850         860         870         880         890         900         910
TTTGATGATTTCTCAAAACTCGGCCTTCCTGATGAATTCAAGGGTTCCTTTGACTTCATAGACGGGCACGTTAAGCCAGTTAAAGGAAGGAAATCAACTGG
F  D  D  F  S  K  L  G  L  P  D  E  F  K  G  S  F  D  F  I  D  G  H  V  K  P  V  K  G  R  K  I  N  W

920         930         940         950         960         970         980         990         1,000       1,010       1,020
CTGAAGGCCGGTATCTTGGAGTCCGACAGGGTAGTGACTGTCAGTCCATATTATGCTCAGGAACTCCTCTCAGGTCCTGCTAAAGGAGTTGAGTTGGACAGT
L  K  A  G  I  L  E  S  D  R  V  V  T  V  S  P  Y  Y  A  Q  E  L  L  S  G  P  A  K  G  V  E  L  D  S

1,030       1,040       1,050       1,060       1,070       1,080       1,090       1,100       1,110       1,120
GTATGTCGTGTGGCTACTCTTACAGGCATCACTAATGGCACCGACACCCAGATCTGGAACCCATCAACCGATGAATATATTTGTGTCAACTATGACAAAACT
V  C  R  V  A  T  L  T  G  I  T  N  G  T  D  T  Q  I  W  N  P  S  T  D  E  Y  I  C  V  N  Y  D  K  T

1,130       1,140       1,150       1,160       1,170       1,180       1,190       1,200       1,210       1,220
ACGGTAATGGATGCAAAACCACTTCTAAAAGAAGCCCTTCAAGCAGAAACTGGGTTAGAGATTGATCCAACAAAACCATTGATTGGCTTCATTGGTAGACTC
T  V  M  D  A  K  P  L  L  K  E  A  L  Q  A  E  T  G  L  E  I  D  P  T  K  P  L  I  G  F  I  G  R  L

1,230       1,240       1,250       1,260       1,270       1,280       1,290       1,300       1,310       1,320
GAGGAACAGAAAGGTTCCGATATTCTTACAGCAGCAATTCCAGAATTCATCGATGAGGATGTCCAAATAATTGTTTTGGGAACTGGTAAGAAAGCCATGGAG
E  E  Q  K  G  S  D  I  L  T  A  A  I  P  E  F  I  D  E  D  V  Q  I  I  V  L  G  T  G  K  K  A  M  E

1,330       1,340       1,350       1,360       1,370       1,380       1,390       1,400       1,410       1,420
AAGCAAATAGAAGAACTTGAGATCAAATATCCAGGCAAAGCTATCGGAGTGGCAAAGTTTAATAGCCCCTTGGCCCATATGATAATTGCTGGTGCTGATTTC
K  Q  I  E  E  L  E  I  K  Y  P  G  K  A  I  G  V  A  K  F  N  S  P  L  A  H  M  I  I  A  G  A  D  F

1,430       1,440       1,450       1,460       1,470       1,480       1,490       1,500       1,510       1,520       1,530
ATGCTAATTCCTAGCAGATTTGAGCCGTGCGGTCTTGTCCAGTTGCATGCCATGAGCTATGGAACTGTGCCAATTGTTGCCTCTACTGGTGGGTTAGTGGAC
M  L  I  P  S  R  F  E  P  C  G  L  V  Q  L  H  A  M  S  Y  G  T  V  P  I  V  A  S  T  G  G  L  V  D

1,540       1,550       1,560       1,570       1,580       1,590       1,600       1,610       1,620       1,630
ACAGTCAAGGAAGGCTATACTGGATTCCAGATGGGTGCATTTAGTGTTGAATGTGATGCTGTAGACCCACAAGATGTTACTGCTGTAGCCAACTGTCAAG
T  V  K  E  G  Y  T  G  F  Q  M  G  A  F  S  V  E  C  D  A  V  D  P  Q  D  V  T  A  V  A  T  T  V  K

1,640       1,650       1,660       1,670       1,680       1,690       1,700       1,710       1,720       1,730
AGAGCCCTCAAAACCTACGGTGCCCCTGCCATGAAAGAGATGATCCAGAACTGCATGTCTCAGGATGTTTCTTGGAAGGAACCGGCTAAACTGTGGGAAAG
R  A  L  K  T  Y  G  A  P  A  M  K  E  M  I  Q  N  C  M  S  Q  D  V  S  W  K  E  P  A  K  L  W  E  K

1,740       1,750       1,760       1,770       1,780       1,790       1,800       1,810       1,820       1,830
CTGTTGTTGAGCTTAGAAGTTCCTGGAAGTGAACCAGGTGTTGATGCAGTAGAAGAAATTACTCCCAAAACCGAGAAGGTTCCCGCTGCTGTATGA
L  L  L  S  L  E  V  P  G  S  E  P  G  V  D  A  V  E  E  I  T  P  K  T  E  K  V  P  A  A  V  *
```

**Fig. 2** Nucleotide and deduced amino acid sequences of the cDNA for *CrGBSSIb* gene

ChloroP program. This suggests that the mature enzyme contains 546 amino acids, with a molecular weight of 59.7 kDa. The 63-amino acid transit peptide likely has a molecular weight of 6.8 kDa. A BLAST search of each genus in the reduced dataset using the amino acid sequence revealed that the degree of sequence identity between the predicted amaranth GBSS enzyme and its homologs from various plants is 54.7–63.5%

Phylogenetic analysis

The relationships among the plant GBSS amino acid sequences were analyzed using the PAUP* 4.0 program, with the results presented in a neighbor-joining phylogenetic tree (Fig. 3). The *E. coli* glycogen synthase sequence was used as an outgroup. The plant GBSSs were separated into two groups, with monocot GBSSI forming one group and dicot GBSSs with monocot GBSSII forming the other. Overall, the monocot GBSSII appears to be closely related to GBSSIs from several dicot plants, which is consistent with the results of an earlier study (Cheng et al. 2012). The dicot GBSS could be grouped into three distinct clades, namely GBSSI, GBSSII, and amaranth GBSSs. Interestingly, the amaranth GBSSs form a sister group to the clade consisting of the rest of the dicot plants and are more closely related to
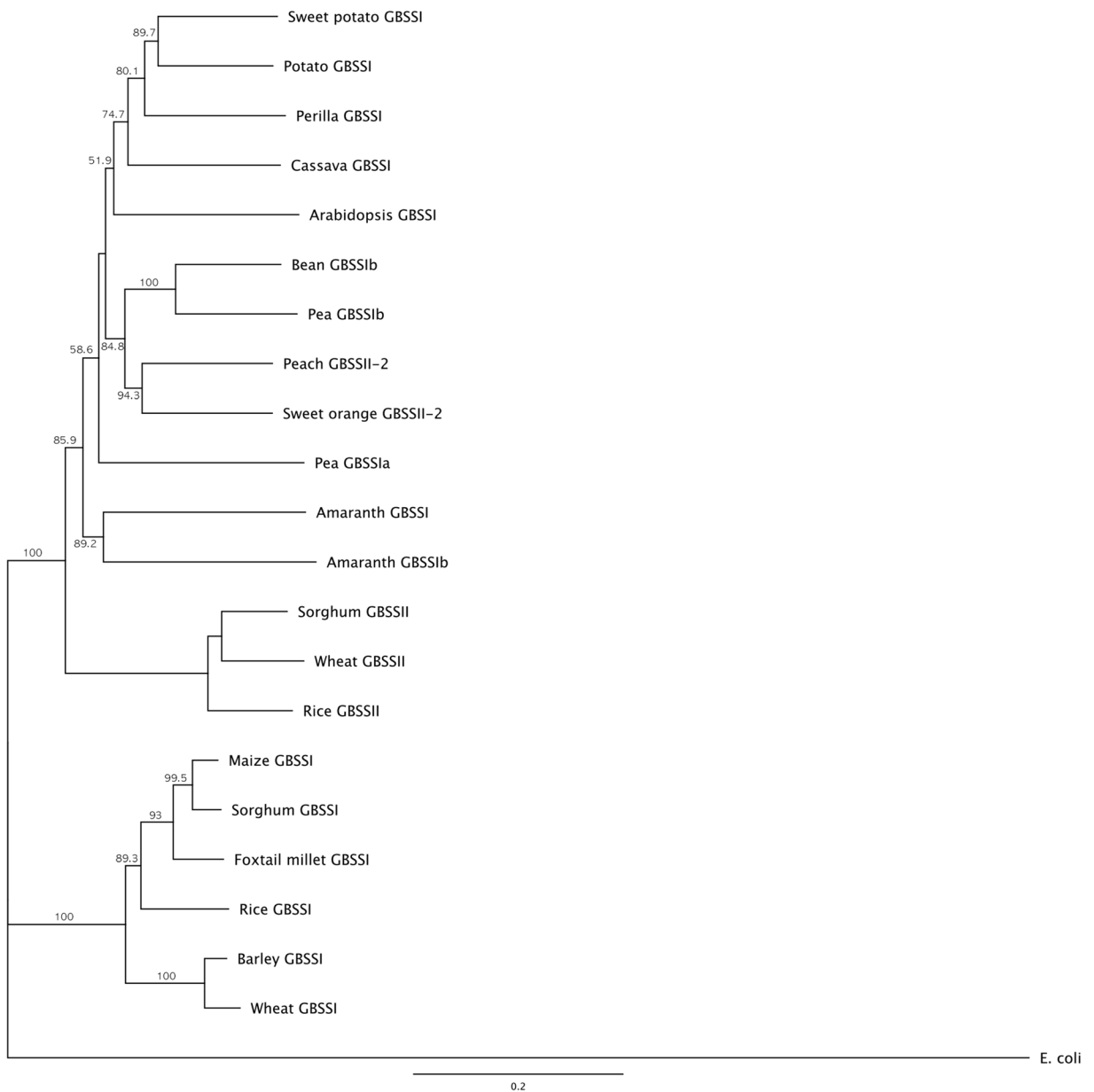
**Fig. 3** Phylogenetic relationships among the GBSS enzymes of various plant species and *E. coli* glycogen synthase. The tree was constructed using the neighbor-joining method. The scale bar represents the evolutionary distance. Bootstrap support values from 1000 replicates are provided at each node

GBSSIIs from monocot plants than to GBSSIIs from dicot plants. This tree implies that the amaranth GBSSIb pathway evolved differently than the pathways for the dicot GBSSIIs. Our results indicate that the amaranth *GBSSIb* isoform probably evolved independently from the *GBSSI* isoforms in this crop (i.e., intraspecies).

**Expression patterns of two *GBSS* isoforms in amaranth grains**

Similar to other cereals, the amylose content of amaranth grains has important consequences for grain cooking properties, yield, palatability, processing quality, and other agronomic traits. Therefore, clarifying the

**Fig. 4** Expression profiles of **a** *CrGBSSI* and **b** *CrGBSSIb* in developing *Amaranthus cruentus* seeds based on a qRT-PCR. **a** Total RNA was extracted from amaranth seeds at different developmental stages (i.e., st1, initial; st2, early; st3, early-late; st4, middle; st5, mid-late; and st6, late). The expression profiles of **a** *CrGBSSI* and **b** *CrGBSSIb* were analyzed by qRT-PCR. *Actin* mRNA was used as a reference to verify that equal amounts of RNA were present in all samples
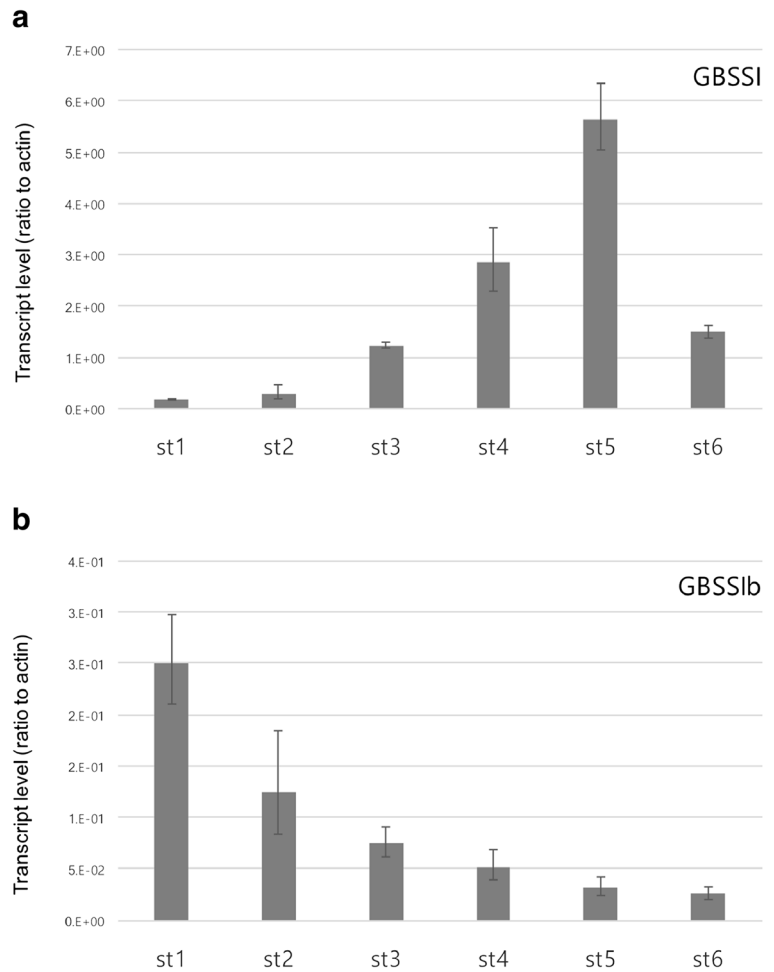
expression patterns of the *GBSS* isoforms in amaranth grains is important and possibly agriculturally relevant. In this study, the relative *CrGBSSI* and *CrGBSSIb* expression levels in seeds at different developmental stages were examined by qRT-PCR (Fig. 4). There were different expression patterns between the two *GBSS* isoforms. The *CrGBSSI* gene was expressed mainly during the mid-late developmental stage, with lower expression levels in the initial and/or early developmental stages (Fig. 4a). There have been several previous

**Fig. 5** *CrGBSSIb* expression profiles in various *Amaranthus cruentus* tissues based on a qRT-PCR. Total RNA was extracted from various tissues at the six-leaf stage (*L* leaf, *P* petiole, *S* stem, *R* root). *Actin* mRNA was used as a reference to verify that equal amounts of RNA were present in all samples
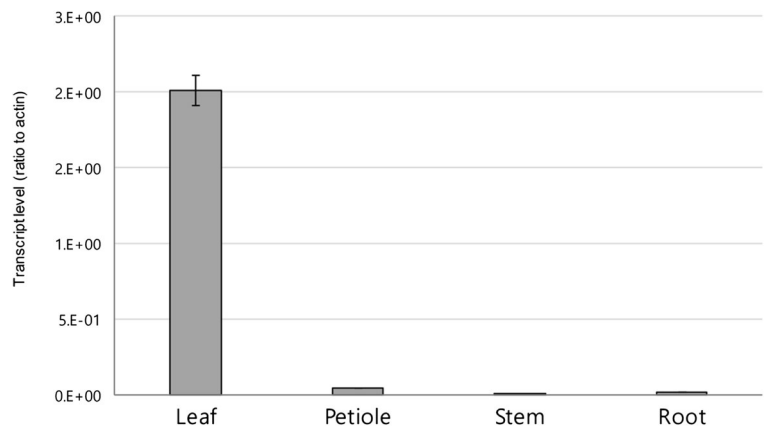
studies on *GBSSI* expression and the encoded enzyme in endosperm plants. This gene is designated as a "late expresser" because of the increased abundance of its transcripts in the later stages of starch granule formation during endosperm development (Hirose and Terao 2004; Ohdan et al. 2005; Park et al. 2011). Thus, the *CrGBSSI* gene is expressed in the later stages of perisperm development, similar to the expression patterns of its homologs in endosperm plants.

In developing seeds, *CrGBSSIb* was expressed during the initial developmental stage (st1) (Fig. 4b). However, the peak transcript levels occurred in the early-late developmental stage (st3), and then gradually decreased thereafter. Similar results were reported for wheat and rice (Nakamura et al. 1998, Hirose and Terao 2004). The *GBSSII* gene is generally expressed during the early stages of seed development, and this gene is considered responsible for the synthesis of amylose in pericarps. Pea plants carry a second *GBSSI* gene (i.e., *GBSSIb*) that is similar to the homologous wheat *GBSSII* gene and is expressed in leaves and pods rather than in embryos (Denyer et al. 1997). Our results also suggest that *CrGBSSIb* might affect the amylose content of pericarp starch. In other words, the encoded enzyme may be closely associated with the metabolism of amylose in pericarp starch. This hypothesis is supported by the fact it is consistent with the temporal pattern of starch deposition, in which intensely stained starch granules were observed in the pericarp during the initial seed developmental stage (Fig. 1). Additionally, the gene expression profiles in different tissues indicated that *CrGBSSIb* was expressed mainly in the leaves, and expression in the petiole, stem, and root tissues was at a very low level than that of the leaf tissues evaluated (Fig. 5). This observation is consistent with the results of a previous study. Vrinten and Nakamura (2000) reported that wheat *GBSSII* is specifically expressed in the pericarp and leaves. In rice, this gene is expressed in the pericarp tissues as well as in the leaves (Dian et al. 2003; Hirose and Terao 2004). The orange *GBSSII-2* and pea *GBSSIb* genes are predominantly expressed in leaves (Cheng et al. 2012). Thus, our results imply that *CrGBSSIb* encodes an enzyme that induces amylose synthesis and may be primarily responsible for amylose metabolism in pericarp tissues. We also conclude that different *GBSS* isoforms are present in the perisperm (*CrGBSSI*) and pericarp (*CrGBSSIb*) tissues of amaranth grains. Furthermore, we recently examined the *GBSSIb* expression patterns associated with the waxy phenotype of

*A. cruentus* (data not shown). The *wxCrGBSSIb* transcript levels exhibited the same patterns as the *CrGBSSIb* transcripts. This result is consistent with the amylose levels in the pericarp, including the levels in the pericarp of wheat plants exhibiting a waxy phenotype (Vrinten and Nakamura, 2000; Patron et al. 2002).

Finally, based on their expression patterns in maturing seeds reported herein and in a previous study (Park and Nishikawa 2012; Park et al. 2014a, b), the starch synthesis genes in amaranth grains can be divided into the following three groups: (i) *GBSSIIa* and *debranching enzyme III* (*DBEIII*) are early expresser genes, which are activated during the early stage of starch granule formation; (ii) *GBSSI* and *debranching enzyme I* (*DBEI*) are late expresser genes, which are induced during the middle or mid-late stages of starch granule formation; (iii) *soluble starch synthase I* (*SSSI*), *soluble starch synthase II* (*SSSII*), and *starch branching enzyme I* (*SBEI*) are steady expresser genes, which are constitutively expressed throughout the seed maturation process.

# References

Breathnach R, Chambon P (1981) Organization and expression of eucaryotic split genes coding for proteins. Annu Rev Biochem 50:349–383

Campbell WH, Gowri G (1990) Codon usage in higher plants, green algae, and cyanobacteria. Plant Physiol 92:1–11

Cheng J, Khan MA, Qiu WM, Li J, Zhou H, Zhang Q, Guo W, Zhu T, Peng J, Sun F, Li S, Korban SS, Han Y (2012) Diversification of genes encoding granule-bound starch synthase in monocots and dicots is marked by multiple genome-wide duplication events. PLoS One 7:e30088

Denyer K, Barber LM, Edward EA, Smith AM, Wang TL (1997) Two isoforms of the GBSSI class of granule-bound starch synthase are differentially expressed in the pea plant (*Pisum sativum* L.) Plant Cell Environ 20:1566–1572

Dian W, Jiang H, Chen Q, Liu F, Wu P (2003) Cloning and characterization of the granule-bound starch synthase II gene in rice: gene expression is regulated by the nitrogen level, sugar and circadian rhythm. Planta 218:261–268

Emanuelsson O, Nielsen H, von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci 8:978

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Hirose T, Terao T (2004) A comprehensive expression analysis of the starch synthase gene family in rice (*Oryza sativa* L.) Planta 220:9–16

Morita N, Maeda T, Miyazaki M, Mihura H, Ohtsuka I (2002) Dough and baking properties of high-amylose and waxy wheat flours. Cereal Chem 79:491–495

Nakamura T, Vrinten P, Hayakawa K, Ikeda J (1998) Characterization of a granule-bound starch synthase isoform found in the pericarp of wheat. Plant Physiol 118:451–459

Ohdan T, Francisco PB, Sawada T, Hirose T, Terao T, Satoh H, Nakamura Y (2005) Expression profiling of genes involved in starch synthesis in sink and source organs of rice. J Exp Bot 56:3229–3244

Okuno K, Sakaguchi S (1981) Glutinous and non-glutinous starches in perisperm of grain amaranths. Cereal Res Commun 9:305–310

Park YJ, Nemoto K, Nishikawa T, Matsushima K, Minami M, Kawase M (2009) Molecular cloning and characterization of granule bound starch synthase I cDNA from a grain amaranth (*Amaranthus cruentus* L.) Breed Sci 59:351–360

Park YJ, Nemoto K, Nishikawa T, Matsushima K, Minami M, Kawase M (2011) Genetic diversity and expression analysis of granule bound starch synthase I gene in the new world grain amaranth (*Amaranthus cruentus* L.) J Cereal Sci 53:298–305

Park YJ, Nishikawa T (2012) Characterization and expression analysis of the starch synthase gene family in grain amaranth (*Amaranthus cruentus* L.) Genes Genet Syst 87:281–289

Park YJ, Nishikawa T, Tomooka T, Nemoto K (2014a) Molecular characterization and expression analysis of a gene encoding an isoamylase-type starch debranching enzyme 3 (ISA3) in grain amaranths. Mol Breed 34:1975–1985

Park YJ, Nishikawa T, Tomooka T, Nemoto K (2014b) Molecular characterization of an isoamylase 1-type starch debranching enzyme (DBEI) in grain amaranth (*Amaranthus cruentus* L.) Mol Biol Rep 41:7857–7864

Patron N, Smith AM, Fahy BF, Hylton CM, Naldrett MJ, Rossnagel BC, Denyer K (2002) The altered pattern of amylose accumulation in the endosperm of low-amylose barley cultivars is attributable to a single mutant allele of granule-bound starch synthase I with a deletion in the 5′-non-coding region. Plant Physiol 130:190–198

Sauer JD (1967) The grain amaranths and their relatives: a revised taxonomic and geographic survey. Ann Missouri Bot Gard 54:103–137

Saunders R, Becker R (1984) *Amaranthus*: a potential food and feed resource. Adv Cereal Sci Technol 6:357–396

Swofford D (1988) PAUP*: phylogenetic analysis using parsimony and other methods, version 4.0 (test ver. 61–64). Sinauer associates publishers, Sunderland

van der Leij F, Visser R, Ponstein A, Jacobsen E, Feenstra W (1991) Sequence of the structural gene for granule-bound starch synthase of potato (*Solanum tuberosum* L.) and evidence for a single point deletion in the *amf* allele. Mol Gen Genet 228:240–248

Vrinten PL, Nakamura T (2000) Wheat granule-bound starch synthase I and II are encoded by separate genes that are expressed in different tissues. Plant Physiol 122:255–263

Yu XR, Li B, Wang LL, Chen XY, Wang WJ, Wang Z, Xiong F (2015) Systematic analysis of pericarp starch accumulation and degradation during wheat caryopsis development. PLoS One 10:e0138228

Zheleznov AV, Solonenko LP, Zheleznova NB (1997) Seed proteins of the wild and the cultivated *Amaranthus* species. Euphytica 97:177–182