

Molecular Cloning and Characterization of Starch Branching Enzyme in Sorghum Grain

Yue-ie C. Hsing¹, Hsing-fang Lee^{1,2}, Bai-chan Lin², Jaw-shu Hsieh^{2*}

1. Institute of Botany, Academia Sinica, Taipei, Taiwan

2. Department of Agronomy, National Taiwan University, Taipei, Taiwan

*The contribution of starch-branching enzyme (EC 2.4.1.18) to starch synthesis in developing sorghum (*Sorghum bicolor*) grains is very important. In the present work, the cDNA encoding sorghum starch branching enzyme was cloned, sequenced and characterized. This clone was termed SorBE and its accession number in public database is AF169833. SorBE consisted of an open reading frame of 2499 bp encoding a protein of 832 amino acids, with the putative molecular weight of 94.1 kDa. The upstream region of the SorBE mRNA was highly GC rich which caused the difficulty for the 5'-RACE cloning. This SorBE hybridized to mRNA of approximately 2,700 nucleotides whose accumulation was detected in developing sorghum seeds. Southern blot analysis of the sorghum leaf genomic DNA revealed SorBE was encoded by a single gene. Phylogenetic study indicated that this is a seed-specific type I starch branching enzyme.*

Keywords: DNA sequencing, grain development, sorghum grain, starch branching enzyme, type I, phylogenetic study

Introduction

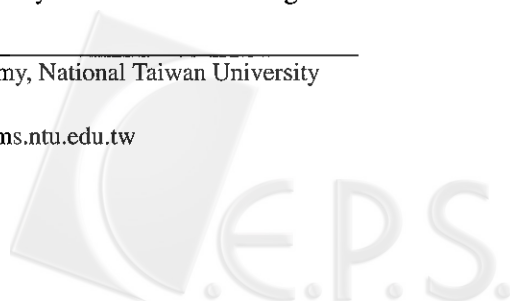
Starch is plant reserve polysaccharide, an end product of carbon fixation by photosynthesis. Most green leaves contain some starch, termed transitory starch, while some organs such as cereal grains, potato tubers, sweet potato roots accumulate large amounts of it, also known as reserve starch. Starch is a polymer of α -D-glucose. At least two polymers can be distinguished: amylose, which is essentially linear, and amylopectin, which is highly branched. Amylose is mainly found as linear chains of about 840 to 22,000

units of α -D-glucopyranosyl residues linked by α -(1 \rightarrow 4) bonds. Its branching percentage is about 0.2 to 1.0 % and its D.p.n. (number-average molecular weight) is about 1,000. In contrast, amylopectin, which usually constitutes about 70% of the starch granule, is more highly branched, with about 4 to 5% of the glucosidic linkages being α -1 \rightarrow 6. Its D.p.n. value is 100,000 to 1,000,000. Both of them are insoluble in water and the absorption maximum of the iodine coloration for amylase and amylopectin are 660 nm and 540 nm, respectively [1].

Starch biosynthesis occurs through the

*Corresponding author: Prof Jaw-shu Hsieh, Department of Agronomy, National Taiwan University
Taipei, Taiwan

TEL: 886-2-2363-5348, FAX: 886-2-2362-0879, eMail: jawar@ccms.ntu.edu.tw



action of four enzymes: ADP-Glc pyrophosphorylase (EC 2.7.7.27), starch synthase (EC 2.4.1.21), starch-branching enzyme (EC 2.4.1.18), and debranching enzyme (EC 2.4.1.41) [2]. The starch branching enzyme catalyzes the formation of these branch points by breaking α -1,4 linkages and re-attaching the reducing ends of the glucan chains by α -1,6 bonds. Introduction of branches into glucan chains increases the number of non-reducing ends, thereby facilitating starch synthesis. In this way, BEs can affect both the structure and quantity of the starch produced.

The importance of starch branching enzyme had been demonstrated by many genetic studies on branching enzymes-deficient mutants. The first instance was even demonstrated by Mendel in 1865. Wrinkled peas, one of the seven traits he used, were designated as *rr* and the regular seeds as *RR*. These *rr* seeds had a reduced starch level, about 66-75% of that in the round (*RR*) seeds. Besides, the amylose content was about 33% in the round form while 60-70% in the wrinkled form. Smith discovered that the genotype *rr* was associated with the absence of one isoform of branching enzyme [3]. This locus was cloned later and found out that the branching enzyme gene contained a 0.8 kb insertion in the *rr* seeds which causing it to express an inactive branching enzyme [4]. Carbon that would have been stored as starch accumulated as sucrose that sweetened the flavor of the pea seeds. The excess sucrose also lowered the osmotic potential of the developing seeds, causing them to swell with fluid. Upon seed maturation, the loss of water resulted in the seeds a wrinkled appearance. Another example was maize amylose extender, *ae*, mutant line that contained high amylose content in the kernels. Detail studies using this line indicated that *Ae* is the structural gene for branching enzyme II and the mutation on the branching enzyme would cause increment of

amylose content [5].

The branching structure of amylopectin appeared to result from the actions of the different isoforms of starch branching enzyme. It had been shown that these branching enzyme isoforms could be divided into two major classes with distinctly differing properties [6]. They had been shown to have different specificities for chain length and also different substrate affinities - with branching enzyme I branching amylose more effectively while branching enzyme II was more active against amylopectin. For instances, in maize endosperm, there were three branching enzyme isoforms [7, 8, 9]. They were BEI, BEIIa, and BEIIb. Of them, BEIIa and BEIIb have been shown to share several biochemical properties [9, 10].

Several branching enzyme genes had been cloned and sequenced. For instances, maize cDNA sequences for BEIIa and BEIIb shared high homology, however, they differed in terms of their expression pattern [11]. *Sbe2a* and *Sbe2b*, corresponding respectively to maize BEIIa and BEIIb cDNA, had a 2.1-kb region showing 78% identity, flanked by divergent 5' and 3' regions. *Sbe2a* had high expression levels in vegetative tissues and moderate expression in developing kernels, whereas the expression of *Sbe2b* was restrictedly to very high levels in kernels during development. These expression patterns suggested that the combinations of isoforms, which interacted to form the transitory starch formed in leaf chloroplasts, were different than the combinations responsible for the reserve starch formed in amyloplasts of kernel endosperm. There was no report regarding to the starch branching enzyme of sorghum yet. The goal of this work was to identify and characterize the sorghum grain starch branching enzyme to gain an understanding of the role of BE in sorghum grain starch synthesis, including quality and quantity.



Materials and Methods

Plant materials

Sorghum (*Sorghum bicolor* L.) plants were grown to maturity in a field environment. Seeds were harvested at 5, 10, 15, 20, 25, 30 days after pollination (DAP). After harvesting or treatment, plant samples were immediately frozen in liquid N₂ and stored at -70 °C prior to extraction.

Construction and screening of the cDNA library

cDNA was synthesized from poly(A)⁺RNA derived from 10 to 12 DAP sorghum seeds. The first strand of the cDNA was synthesized using reverse transcriptase primed with oligo(dT) and random primer (dN₆), and the second strand using RNase H and DNA polymerase (Klenow fragment). The cDNA was ligated into the *Eco*RI linkers and the whole cassette was ligated into the *Eco*RI site of LambdaZapII (Stratagene).

RT-PCR amplification of the sorghum BE fragment

Degenerate oligonucleotide primers for RT-PCR were designed based on consensus amino acid sequence conserved in the BE type I sequences of maize, rice, and wheat (with the accession number U17897, AF36268 and AF076679, respectively). Seven degenerated primers were used, i.e. BEa (5'-GTTTGATGGCTCCGATTTGATGGAG-3'), BEb (5'-GGGTGACCAAACCTCATTTCCCATAA-3'), BEc (5'-GAGGGAAGTCTTGAATCTTTTTTC-3'), BEd (5'-ACCCACATTGATACCATGGTGATGA-3'), BEe (5'-CACCTACAATTGATCGAGGGATTGCA-3'), BEf (5'-CUACUACUAGGCCACGCGTCTGACTGTAC-3'), and BEg (5'-CCATTCGCGATATACAGTTCC-3'). Five µg of total RNA extracted from 10 DAP sorghum grain were used in a standard RT-PCR reaction to generate the PCR template first strand

cDNA. An 1 µl aliquot of the RT-PCR reaction mix was used in a PCR reaction with annealing temperatures. PCR cycle conditions consisted of 3 min at 94 °C, then 30 cycles of 94 °C for 30 sec, 55 °C for 30 sec, and 72 °C for 1 min, followed by 10 min at 72 °C.

cDNA sequencing and molecular analysis

The sequences of pSorBE1 were determined with the dideoxy chain termination method [12] using the Prism Ready Reaction Dye Deoxy Terminator Cycle Sequencing Kit and ABI377 automated DNA-sequencing apparatus (Applied Biosystems) following subcloning into pBluescript. Both strands were sequenced and the sequence data were analyzed using the Genetics Computer Group Sequences Analysis software package Version 9.0 [13]. The nucleotide sequence data appeared in the database under the accession number AF169833.

Preparation of RNA and Northern blot analysis

Sorghum seeds with specific developing time were pooled and RNA was extracted according to McCarty [14]. Total RNA from developing seeds was separated in 0.8% agarose gels containing formaldehyde and transferred to nitrocellulose membrane [15]. The full-length cDNA clones of SorBE1 were used for the hybridization. Blots were hybridized overnight with radioactive DNA fragments and washed twice for 20 min at room temperature in 2X SSPE (1X SSPE is 0.18 M NaCl, 10 mM NaHPO₄, pH 7.5, 1 mM EDTA), 0.1% (w/v) SDS, and twice for 20 min at 50 °C in 0.1X SSPE, 0.1% SDS.

Genomic Southern analysis

Total genomic DNA was extracted from young sorghum leaves as described by Hsing *et al.* [16]. Five restriction enzymes were

used - *EcoRV*, *HindIII*, *ScaI*, *SpeI* and *PvuII*. DNA prepared were digested with each of the five enzymes and hybridized with pSorBE1 probe. The digested DNA sample were resolved on 0.8% agarose gels and transferred to nylon membranes. DNA labeling was performed with α - ^{32}P dCTP using a modified oligolabeling method [17]. Hybridization was performed at 65 °C for 16 hr. The membrane was then washed in a series consisting of SSC buffer (1X SSC is 0.15M NaCl and 0.015M sodium citrate, pH 7.0) - 5X SSC buffer containing 0.5% SDS at room temperature for 20 min, 1X SSC containing 0.5% SDS at 37 °C for 20 min, 1X SSC containing 0.1% SDS at 65 °C for 20 min, and then 0.1X SSC containing 0.1% SDS at 65 °C for 20 min.

Phylogenetic analysis

The DNA sequences of all available plant starch branching enzyme cDNA were pulled out from database and aligned using the CLUSTALV program [18] by setting the fixed and floating gap penalties of 5 each and weighted toggle transitions in the multiple

alignment parameter option. Phylogenetic reconstruction was performed by the maximum likelihood method based on the procedure by Felsenstein [19]. The distance trees were constructed by the neighbor-joining method [20] using the Phylogenetic Inference package (PHYLIP version 3.5) [21].

Results and Discussion

Isolation of SorBE1 and sequence analysis

To isolate the sorghum grain BE cDNA, degenerate primers were prepared based on amino acids sequences from highly conserved regions of the rice, wheat and maize BE sequences. Totally 7 primers were used and each of them overlapped at least 100 bp. Both strands of the products were sequenced and BLAST searched right away to confirm if the products were BE fragments. The detail cloning strategy is illustrated in Fig 1. From the 5'-most fragment, the primer BEg and 5'RACE were

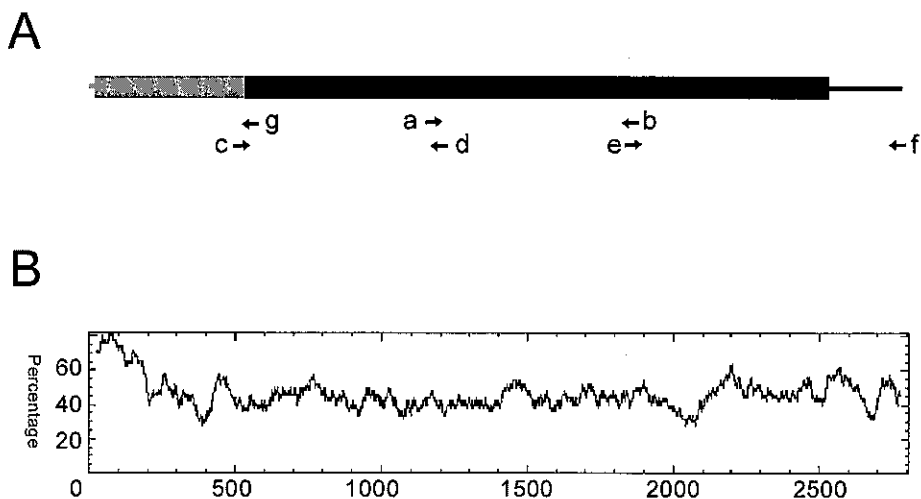


Fig. 1. Cloning strategy for the sorghum branching enzyme cDNA. Panel A, the protein-coding region is boxed. The black bar indicates the region pulled out by PCR cloning and the grey bar indicates the region pulled out by cDNA library method. Arrows indicate the location and direction of primers used for PCR cloning, 5' RACE and 3' RACE. Panel B. GC content of the whole cDNA region, with the window size of 50 bp.

used to try to amplify the upstream fragment but no PCR product was generated. In order to find the whole mRNA sequence, we then constructed a 10 DAP sorghum grain cDNA library with dN₆ as primer instead of the regular oligo(dT) primer, as indicated in Materials and Methods. Fragments cd, ab, and ef were pulled out by PCR cloning. Since primer SBg could not be used successfully to perform the 5'-RACE, the most 5' clone was pulled out from the phage lifts using fragment BEbc as the probe. The whole sequence designated pSorBE and its accession number in database was AF169833.

Panel B in Fig 1 illustrated the GC content of the whole SorBE cDNA sequence. The GC content in the 5' end was very high, compared with the remaining region. This high-GC region consisted about 200 bp nucleotides containing the start codon, and corresponding to the region obtained through cDNA cloning strategy instead of PCR cloning. This coincided with the observation that many *Gramineae* genes contained high GC-content region at the 5' end [22].

The nucleotide sequence of pSorBE contained 2805 bp with 13 bp and 293 bp at 5'- and 3'- untranslated regions, respectively. The putative SorBE protein comprised 832 amino acids with predicted molecular mass of 94.1 kDa. The cDNA sequence and its corresponding amino acid sequences, along with the position of start codon, stop codon and much other information are indicated in Fig 2. This protein contained a transit peptide targeting to chloroplast or amyloplast and were presumably post-translationally modified. Accordingly, BE amino terminus should contain some features in common with chloroplast transit peptides, i.e. a high content of Ser and Thr residues and a central, positively charged domain. SorBE protein did consist this property. Moreover, with comparison to the branching enzyme of other species, the transit peptide length should be 64 residues with the molecular mass of 6.6

Table 1. The amino acid number and its percentage of the mature sorghum branching enzyme.

a.a.	Number	Percentage
Ala	54	7.031
Cys	5	0.651
Asp	62	8.073
Glu	45	5.859
Phe	43	5.599
Gly	57	7.422
His	30	3.906
Ile	29	3.776
Lys	53	6.901
Leu	49	6.380
Met	25	3.255
Asn	37	4.818
Pro	33	4.297
Gln	15	1.953
Arg	38	4.948
Ser	51	6.641
Thr	37	4.818
Val	53	6.901
Trp	15	1.953
Tyr	37	4.818

kDa. The mature peptide comprised 769 amino acids with predicted molecular mass of 87.5 kDa. The pI values of the mature peptide are 6.5 with the net charge of -16. Table 1 lists the amino acid residue numbers and composition of the sorghum BE mature protein and Fig 3 illustrates the hydrophobicity of the protein. Only three amino acid residues consisted less than 2%, they were Cys, Gln and Trp. Hydropathy plot indicated that there were several hydrophilic and hydrophobic domains in this amyloplast-targeted enzyme.

SorBE belonged to type I BE

There were two types of BE, as indicated in the introduction section. For instances, they could be separated into BEI and BEII from maize endosperm using ion exchange chromatography [7]. There were also differences in substrates preferences, with

```

1  CGAAGCGCGCGCAATGCTGTGCTCTGTGTCGGCTCTTCCTCGCGACTCGCTCCGCGCGCGCGGCGCTCTGTTG3CCTGCTGATCGGCGCGCGCGCCAGGATCG3GGTGGCAG
    M L C L V S P S S S P T P L P P P R R S R S P A D R A A P P G I A G G S 36
121 CAATGTGGCGCTGAGTGTGCTGTCTGTCCAGTGGAGGCTCGCGGTGAGGCGCCCGGAAAGGTCAGAGCAAAATGGCCACTGCGACTTATGTCAGAAAGATAAAACTATGGCAACCGC
    N V R L S V L S V Q C E A R R S G P R K V K S K F A T A A I V Q E D K T M A T A 76
241 CAAAGCGCAAGCTGACCATCTTCCCATATAGACCTGACCCCAAGTGTGTAAATTCAGAGACCATTTGACGTACCGGATGAAAAATTCCTGGACAGAAAGGCTCCATTGAGGAAAA
    K G N V D H L P I Y D L D P K L V K F K D H F S Y R M K K F L D Q K G S I E E N 116
361 TGAGGGAAGTCTTGAAGAAATTTCTAAAGGCTATTGAAATTTGGGATCAATACAAGTGAAGATGGAAGTGTATATCGGGAATGGGACCTGCTGCAAGGAGGCGAGACTTATGGTGA
    E G S L E E F S K G Y L K F G I N T S E D G T V Y R E W A P A A Q E A E L I G D 156
481 CTTCATGAGTGAAGTGGTGCAAAOCCACAAGATGGAGAAGGATAAAATTTGGTGTGTTGGTGCATCAAAATTGACCATGTCAAAGGAAACCTGCTATCCCTCAATAATCCAAAGGTTAAAT
    F N E W N G A N H K M E K D K F G V W S I K I D H V K G K P A I P H N S K V K F 196
601 TCGCTTCTACATGGTGGAGTATGGGTGTAGTGTATTCCTGCATGGATTCGTTATCGGACTGCTGATGCCCTCAAATTTGGAGCTCCCTATGATGGTGTTCATTGGGATCTCCCTGCTTC
    R F L H G G V W V D R I P A W I R Y A T A D A S K F G A P Y D G V H W D P P A S 236
721 TGAAAGGTACACTTTTAAAGCATCTCGCGCTTCAAAGCCTGCTGCTCCACGTTACTATGAGGCCATGTTGGGTATGAGTGGTGAAGGCAAGCAGTAAAGCACATATAGGGAATTTGCGA
    E R Y T F K H P R P S K P A A P R I Y E A H V G M S G E K P A V S T Y R E F A D 276
841 CAATGTGTGCCACGCATACGAGCAAAATCAACACCGTTCAGTGTGAGGAGTATGGAACATTTCTACTATGCTTCTTTGGGTACCATGTCAGAAATTTCTGCGGTAGCAG
    N V L P R I R A N N Y T V Q L M A V M E H S Y Y A S F G Y H N P F A V S S 316
961 CAGATCAGGCAACCAGAGGACCTAAAATATCTGTGTGATAAGGCAACAGTTTGGSTTTGGGAGTCTGATGGATGTTGTTACAGCAGCATCAAGTAATAATGTCACAGATGTTTAAA
    R S G T P E D L K Y L V D K A H S L G L R V L M D V V H S H A S N N V T D G L N 356
1081 TGGTATGATGTGGCAAAAGCACCAGAGTCTTATTCATATGGAGATAGAGGTATATCAAACTTTGGGACAGTCCGCTGTTCACACTGCTAACGGAGGATGAGGTTTCT
    G Y D V G Q S T Q E S Y F H M G D R G Y H K L W D S R L F N Y A N W E V L R F L 396
1201 TCTTTCACACCTGAGATATGGCTAGATGAATTCATGTTGATGGCTTCGCGTGTGATGGAGTACATCAATGCTGTATCATCACCATGGTATCAATGGTGGGTTTACTGGTAATACCA
    L S N L R Y W L D E F M F D G F R F D G V T S M L Y H H G I N V G T G F T G N Y Q 436
1321 GGAATATTCAGTGTGGACACAGATGGGATGCACTGTTTACATGATGCTTGCACCAACATTTAATGCACAAACTCTTGGCAGAACTGTTGTTGCTGAAGATGTTTCAGGATGCC
    E Y F S L D T D V D A V Y M M L A N H L M H K L L P E A T V V A E D V S G M P 476
1441 AGTCCCTTCCCGTTCAGTGTGATGAAGTGGAGTGGGTTTGCATATCCCTGCGCAATGGCTATCCCTGATAGATGGATGACTACCTGAAGAATAAAGATGACTCTGAGTGGTGGATGGG
    V L C R S V D E G G V G F D Y R L A M A I P D R W I D Y L K N K D D S E W S M G 516
1561 TGAAATAGGCACTACTTGAACAACAGGAGATATCTGAAAATGTCATGCGATATGCTGAGAGCAGTATGATGCTACTGTTGGCGACAAAACATAGCATTTCTCTGATGAGCAAGGA
    E I A H T L T N R R Y T E K C I A Y A E S H D Q S I V G D K T I A F L L M D K E 556
1681 AATGTACATGGCATGTCAGACTGCGCTGCTTCACTTACAATTTGATGGAGGATTTGCACCTCCAAAAGATGATTCACITCAACAAATGGCCCTGGAGGATGATGGCTACTTGAATTT
    M Y T G M S D L Q P A S P T I D R G I A L Q K M I H F I T M A L G G D G Y L N F 596
1801 TATGGGAATGAGTGTGGTCCACGAGAATGATGATTTCCAAGAGAAGGAACAACCTGGAGCTATGATAAAATGAGAGCTGAGTGGAGCTTGTGGACACTGACTTGGGTACA
    M G N E F G H P E W I D F P R E G N N W S Y D K C R R Q W S L V D T D H L R Y K 636
1921 GTACATGAATGGTGTGATCAAGCAATGAATGGCTGCGATGAGAGATTTCCCTCCCTTTTGTCAATCCAAGCAGATTTGTAAGCGACATGAATGATGAGAAAAAGGTTATTTGCTTTGAAGC
    Y M N A F D Q A M N A L D E R F S F L L S S K Q I V S D M N D E K K V I V F E R 676
2041 TGGAGATTTGGTGTGTTTCAATTTTCAATCCCAAGAAAACCTATGACGTTACAAAAGTGGATGATGTTGGCTGGGAAATACAGAGTAGCAGTGGACTGATGCTTTCGCTTTGG
    G D L V F V F N F H P K K T Y D G Y K V G C D L P G K Y R V A L D S D A F V F G 716
2161 TGGACATGGAAGTCTGGCGAGCTGAGTACTTCACTGCTGCGCTGAGGGATGACAGGGGTGCTGAAAACAACTCAACACACCGGCAAACTCATTCAAAGTCCCTTCCGCCACCGCG
    G H G R V G H D V D H F T S P E G I P G V P E T N F N N R P N S F K V L S P P R 756
2281 CACTTGTGCTGCTTATATGCTGTAGATGAAGAGCTGAAAGCTTCAAGCGAAAGCAGAGACAAGTGGTCTAGAAAAGACTTCTCCAGACATCATGATGATGATGCTACCCCGTCAA
    T C V A Y Y R V D E E A E R L Q A K A E T T G S R K T S P D I I D V D A T P V K 796
2401 AACTTCTACAGCTACTAAAGAGACAAGGAGGCAACTCTGCTGGCAAGGAGATGCAAGCAAGGAGGATGGAAGTTTGCACCGCAGTCACTCAAATAAAACACCAAAATGAAGCCAGAA
    T S T A T K E D K E A T S G G K E D A S K K G W K F A R Q S S N K N T K * 832
2521 CTCCTTGATCAGGACTGGACTGGCTGCCAGCGCCCTCTTAATAGTCTGCTCTACTGAGACTAGCGCTGGTGCCTTGTAAACGTCCTTCTGCTGACTCTTAGATGACTGGTGCATCT
    CCGAGCAAGCACTGCTTGTATAGTTTTCTAGATAGTAGACAAGAATAATTCAGGGATGGATGTTGTTGATGATGATCAGGAGGAGCCATGGCTGGGTCACTGATTTGACTGGGTACTCT
2641
2761 GGCTAAATGTGCCCTGCCAGAAATTAACACTTGGGCTTCTGGA
    
```

Fig. 2. Nucleotide and deduced amino acid sequences of the cDNA encoding sorghum branching enzyme. The deduced amino acid sequence is shown below the nucleotide sequence numbered in the 5' to 3' direction. The possible polyadenylation signal is underlined.

branching enzyme I branching amylase more effectively while branching enzyme II more active against amylopectin. The BEII isoform differed from SBI by having an

acidic amino-terminal extension and a shorter carboxyl-terminus. Moreover, transcript of BEII was highly expressed in leaf tissues, whereas transcript of BEI was

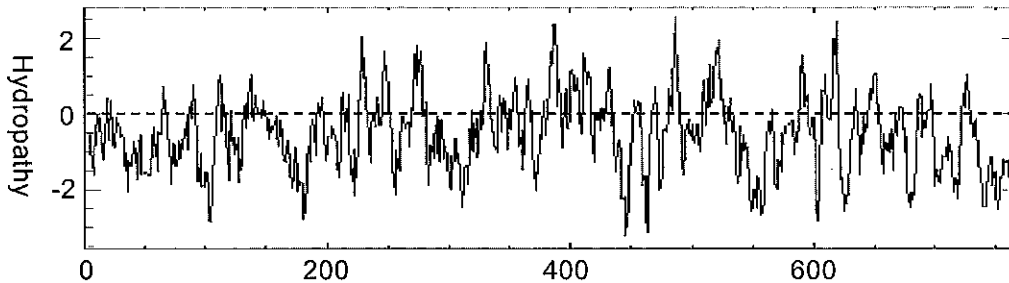


Fig. 3. Hydropathy plots of the deduced SorBE protein. Hydropathy plots were predicted according to the method of Kyte and Doolittle [32] using a window of six amino acids. The abscissa represents the position of the amino acid residues. Positive values indicate hydrophobicity.

mainly detected in storage organs such as rice seeds or potato tubers [23]. The core region of the SorBE protein, i.e. excluding the amino-terminal and carboxy-terminal extension, showed more than 50% identical on the amino acid level to any available BEI or BEII sequences from plant species and contained the presumed active site of the starch synthesizing enzymes. Because of the sequence similarity of SorBE with the known

BE, along with the longer carboxy terminus and shorter amino terminus, it was confirmed that SorBE belonged to type I BE.

SorBE was highly expressed in developing sorghum grains

To investigate the expression of the SorBE genes, Northern blot analysis was performed on total RNA extracted from developing sorghum grains at 5, 10, 15, 20,

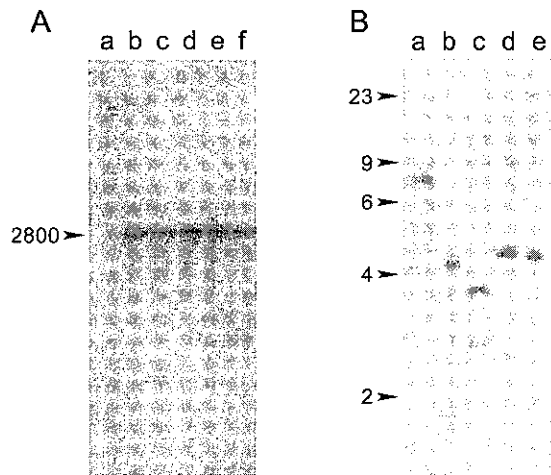
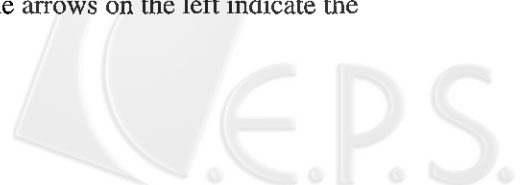


Fig. 4. Southern and Northern blot analysis of sorghum starch branching enzyme gene. Panel A. Northern blot analysis. Twenty μg total RNA were subjected to northern blot analysis. The probe was probed by ^{32}P -labeled SorBE1 encoding sorghum branching enzyme. Lanes a to f represent RNA from 5 DAP, 10 DAP, 15 DAP, 20 DAP, 25 DAP and 30 DAP seeds, respectively. Panel B. Southern blot analysis. Ten μg genomic DNA was digested with restriction enzyme and subjected to Southern blot analysis. Hybridization was carried out using the ^{32}P -labeled SorBE. Lanes a to e, digested with *EcoRV*, *HindIII*, *ScaI*, *SpeI* and *PvuII*, respectively. The arrows on the left indicate the molecular size marker (kb).



25 or 30 DAP (Fig. 4A). An RNA of about 3000 nucleotides was detectable in the developing sorghum seeds. High level of SorBE transcripts were detected at 10 days after pollination and kept at high level till the end of maturation process. This expression pattern was thus very similar to the pattern observed for many other genes involved in starch synthesis, such as granule bound starch synthase or soluble starch synthase in sorghum gram [24, 25]. Southern blot analysis (Fig. 4B) was also performed to reveal the copy number for the genes encoding SorBE. Totally 5 restriction enzymes were used, i.e. *EcoRV*, *HindIII*, *ScaI*, *SpeI* and *PvuII*. The results indicated that SorBE was encoded by a single-copy gene.

Phylogenetic analysis

Database searching using SorBE as the query sequence gave many protein sequences with similar BE core region. The aligned enzymes included type I or type II BE from Arabidopsis, barley, cassava, common bean, maize, pea, potato, rice, sweet potato, sorghum, and wheat. The information about the species, its corresponding accession number and references are indicated in Table 2. Phylogenetic analysis was performed according to Materials and Methods and the results are displayed in Fig. 5. This figure indicates that BE may be divided into two distinct clades, and SorBE belongs to the type I BE. Like most of the other cereal plants, sorghum genome should contain the type II BE. The sequence homology between type I and type II BE should be relatively low so that the Southern blot analysis shown in Fig 4 only gives the signal of the gene encoding

Table 2. The information of sequences used in Fig 6.

	Species	Accession	Reference
po_ I	Potato	<u>Y 08786</u>	Khoshnoodi et al., 1996 [26]
po_ II		<u>AJ 011885</u>	Unpublished
bar_ IIa	Barley	<u>AF 064560</u>	Sun et al., 1998 [27]
bar_ IIb		<u>AF 064561</u>	Unpublished
mz_ I	Maize	<u>U 17897</u>	Fisher et al., 1995 [28]
mz_ II		<u>L 08065</u>	Fisher et al., 1993 [29]
ri_ I	Rice	<u>AF 136268</u>	Unpublished
ri_ 3		<u>D 16201</u>	Mizuno et al., 1993 [30]
ri_ 4		<u>AB 023498</u>	Unpublished
wh_ I	Wheat	<u>AF 076679</u>	Unpublished
wh_ Ia		<u>AF 286318</u>	Unpublished
wh_ Id		<u>AF 286317</u>	Unpublished
wh_ II		<u>AF 286319</u>	Unpublished
wh_ IIa		<u>AF 338432</u>	Rahman et al., 2001 [31]
ara_ IIa	Arabidopsis	<u>NM 112935</u>	Unpublished
ara_ IIb		<u>NM 120446</u>	Unpublished
sp_ II	Sweet potato	<u>AB 071286</u>	Unpublished
ca	Cassava	<u>X 77012</u>	Unpublished
pea_ I	Pea	<u>X 80009</u>	Burton et al., 1995 [6]
pea_ II		<u>X 80010</u>	Burton et al., 1995 [6]
be_ I	Bean	<u>AB 029548</u>	Unpublished
be_ 3		<u>AB 029549</u>	Unpublished
sor	Sorghum	<u>AF 169833</u>	this study

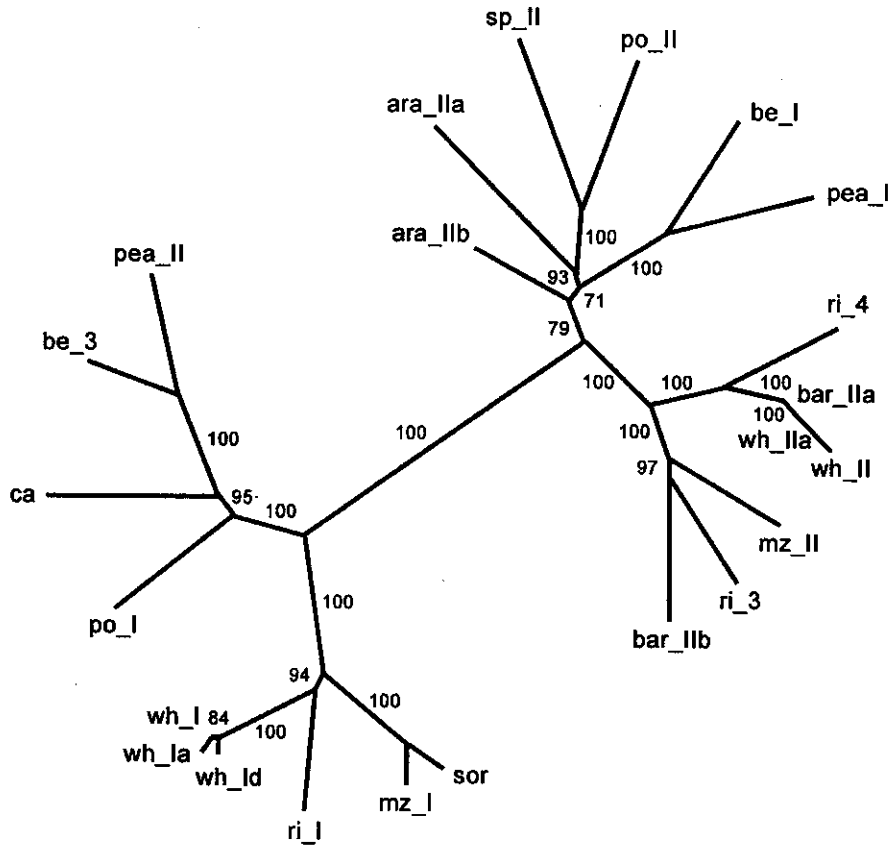


Fig. 5. Phylogenetic tree of higher plants starch branching enzymes. The aligned enzymes are from Arabidopsis, barley, cassava, common bean, maize, pea, potato, rice, sweet potato, sorghum, and wheat. Abbreviations are as in Table 2. The numbers beside branches indicate the bootstrap values (%) derived from 1000 replicates.

type I BE since high stringency was used in our study.

Acknowledgements

This work was supported by grants from the National Science Council, ROC and Academia Sinica, ROC to Hsing YIC and Hsieh JS.

References

1. Nakamura Y: Some properties of starch debranching enzymes and their possible role in amylopectin biosynthesis. *Plant Sci*, 1996; 121:1-18.
2. Myers AM, Morell MK, James MG, Ball SG: Recent progress toward understanding biosynthesis of the amylopectin crystal. *Plant Physiol*, 2000; 122: 989-997.
3. Smith AM: Major differences in isoforms of starch-branching enzyme between developing embryos of round and wrinkled-seeds peas (*Pisum sativum* L). *Planta*, 1988; 175, 270-279.
4. Bhattacharyya MK, Smith AM, Ellis THN, Hedley C, Martin, C: The wrinkled-seed character of pea described by Mendel is caused by a transposon-like insertion in a gene encoding starch branching enzyme. *Cell*, 1990; 60: 115-121.
5. Fisher DK, Gao M, Kim KN, Boyer CD, Guiltinan MJ: Allelic analysis of the maize amylase-extender locus suggests that independent genes encode starch-branching enzymes IIa and IIb. *Plant Physiol*, 1996;

- 110: 611-619.
6. Burton RA, Bewley JD, Smith AM, Bhattacharyya MK, Tatge H, Ring S, Bull V, Hamilton WDO, Martin C: Starch branching enzymes belong to distinct enzyme families are differentially expressed during pea embryo development. *Plant Journal*, 1995; 7: 3-15.
 7. Boyer CD, Preiss J: Multiple forms of (1-4)- α -D-glucan, (1-4)- α -D-glucan-6-glucosyl transferase from developing *Zea mays* L. kernels. *Carbohydr Res*, 1978; 61: 321-334.
 8. Singh BK, Preiss J: Starch branching enzymes from maize. Immunological characterization using polyclonal and monoclonal antibodies. *Plant Physiol*, 1985; 79: 34-40.
 9. Guan HP, Preiss J: Differentiation of the properties of the branching isozymes from maize (*Zea mays*). *Plant Physiol*, 1993; 102: 1269-1273.
 10. Takeda Y, Guan HP, Preiss J: Branching of amylase by the branching isoenzymes of maize endosperm. *Carbohydr Res*, 1993; 240: 253-263.
 11. Gao M, Fisher DK, Kim KN, Shannon JC, Guiltinan MJ: Independent genetic control of maize starch branching enzymes α and β . Isolation and characterization of a *sbe2a* cDNA. *Plant Physiol*, 1997; 114: 69-78.
 12. Sanger RF, Nicklen S, Coulson AR: DNA sequencing with chain-termination inhibitors. *Proc Natl Acad Sci USA*, 1977; 74: 5463-5476.
 13. Deverux J, Haerberli P, Smithies O: A comprehensive set of sequence analysis programs for VAX. *Nucl Acids Res*, 1984; 12: 387-395.
 14. McCarty DR: A simple method for extraction of RNA from maize tissue. *Maize Genet Coop Newsl*, 1986; 60: 61.
 15. Hsieh JS, Liu C, Hsing YC: Molecular cloning of a sorghum cDNA (Accession No. U23945) encoding the seed waxy protein. *Plant Physiol*, 1996; 112: 1735.
 16. Hsing YIC, Hsieh JS, Peng CI, Chou CH, Chiang TY: Systematic status of the *Glycine tomentella* and *G. tabacina* species complexes based on ITS sequences of nuclear ribosomal DNA. *J Plant Res*, 2001; 114: 435-442.
 17. Hodgson CP, Fisk RZ: Hybridization probe size control: optimized 'oligolabelling'. *Nucl Acids Res*, 1987; 15: 62-95.
 18. Higgins DG, Bleasby AJ, Fuchs R: CLUSTAL V: improved software for multiple sequence alignment. *Comput Appl Biosci*, 1992; 8: 189-191.
 19. Felsenstein J: Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 1985; 39: 783-791.
 20. Saitou N, Nei M: The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evo*, 1987; 4: 406-425.
 21. Felsenstein J: PHYLIP (Phylogeny inference package) version 3.5s. Department of Genetics, University of Washington, Seattle, Wash, 1993.
 22. Wong GK, Wang J, Tao L, Tan J, Zhang J, Passey DA, Yu J: Compositional gradients in Gramineae genes. *Genome Res*, 2002; 12:851-856.
 23. Larsson C, Khoshnoodi J, Ek B, Rask L, Larsson H: Molecular cloning and characterization of starch-branching enzyme II from potato. *Plant Mol Biol*, 1998; 37: 505-551.
 24. Lin CL: Studies on the waxy gene expression in sorghum endosperm at protein and RNA levels. M. S. Thesis, Research Institute of Agronomy, National Taiwan University. 1991.
 25. Lee HF, Chen MR, Hsing YIC, Hsieh JS: Characterization of a sorghum cDNA encoding soluble starch synthase in endosperm. *Chinese Agronomy*, in press.
 26. Khoshnoodi J, Blennow A, Ek B, Rask L, Larsson H: The multiple forms of starch-branching enzyme I in *Solanum tuberosum*. *Eur J Biochem*, 1996; 242: 148-155.
 27. Sun C, Sathish P, Ahlandsberg S, Jansson C: The two genes encoding starch-branching enzymes IIa and IIb are differentially expressed in barley. *Plant Physiol*, 1998; 118: 37-49.
 28. Fisher DK, Kim KN, Gao M, Boyer CD, Guiltinan MJ: A cDNA encoding starch branching enzyme I from maize endosperm. *Plant Physiol*, 1995; 108: 1313-1314.
 29. Fisher DK, Boyer CD, Hannah LC: Starch branching enzyme II from maize endosperm. *Plant Physiol*, 1993; 102: 1045-1046.
 30. Mizuno K, Kawasaki T, Shimada H, Satoh H, Kobayashi E, Okumura S, Arai Y, Baba T: Alteration of the structural properties of starch components by the lack of an isoform of starch branching enzyme in rice seeds. *J Biol Chem*, 1993; 268:19084-19091.
 31. Rahman S, Regina A, Li Z, Mukai Y, Yamamoto M, Kosar-Hashemi B, Abrahams S, Morell MK: Comparison of starch-

branching enzyme genes reveals evolutionary relationships among isoforms. Characterization of a gene for starch-branching enzyme IIa from the wheat genome donor *Aegilops tauschii*. *Plant Physiol*, 2001; 125: 1314-1324.

32. Kyte J, Doolittle RF: A simple method for displaying the hydropathic character of a protein. *J Mol Biol*, 1982; 157: 105-132.

