Journal of Experimental Botany, Vol. 56, No. 412, pp. 623–632, February 2005
doi:10.1093/jxb/eri065 Advance Access publication 10 January, 2005
This paper is available online free of all access charges (see http://jxb.oupjournals.org/open_access.html for further details)

Journal of Experimental Botany

RESEARCH PAPER

Evolution and expression analysis of starch synthase III and IV in rice

Weimin Dian*, Huawu Jiang* and Ping Wu†

State Key Laboratory of Plant Physiology and Biochemistry, College of Life Sciences, Zhejiang University, Hangzhou 310029, PR China

Received 23 August 2004; Accepted 25 October 2004

Abstract

Plants contain at least five subfamilies of starch synthases, granule bound starch synthase (GBSS) and starch synthases I, II, III, and IV (SSI, SSII, SSIII, SSIV). In this work, two members of SSIII and SSIV, respectively, were cloned and designated *OsSSIII-1/-2* and *OsSSIV-1/-2* in rice. Together with six other previously reported genes, the SS gene family in rice therefore is known to be duplicated and to comprise ten SS genes distributed among the five subfamilies. The starch synthase activity of each SS was confirmed by expression and enzyme activity assay in *E. coli*. Expression profile analysis with reverse transcription-PCR, western blotting and zymogram, indicates that OsSSIII-2 and OsSSIV-1 are mainly expressed in endosperm, while OsSSIII-1 and OsSSIV-2 are mainly expressed in the leaves. With a similar pattern of genes encoding other enzymes for starch synthesis, (such as GBSS, SSII, ADP-glucose pyrophosphorylases, and branching enzymes), it is suggested that two divergent groups of these genes should be classified in rice. Group I genes are preferentially expressed in the endosperm and function on storage starch synthesis. Group II genes are mainly expressed in leaves and some of them in the early developing endosperm, and function on transient starch synthesis in rice.

Key words: Gene duplication, gene expression, *Oryza sativa* L, starch biosynthesis, starch synthase.

Introduction

Plant starches are comprised of two classes of glucose homopolymers. Firstly, there is amylose, which is a lightly branched linear molecule with a degree of polymerization of 1000 to 5000 Glc units. Secondly, amylopectin, which has a much larger polymer unit (with a degree of polymerization of 10^5–10^6 Glc units) and contains frequent α-1,6 branch linkages. In higher plants, biosynthesis of starch occurs in plastids with the involvement of a series of biosynthetic enzymes, including ADP-Glc pyrophosphorylase (AGPase), starch synthase (SS), starch branching enzyme (BE), and debranching enzyme (DBE) (Smith *et al*., 1997; Myers *et al*., 2000; James *et al*., 2003).

The α-glucan chains of both types of polymers are elongated by starch synthase (ADP-Glc: α-1,4 glucan α-4-glucosyl-transferase; EC 2. 4. 1. 21). Starch synthase catalyses the transfer of α-D-glucose from ADP-Glc to the non-reducing end of the chain by an α-1,4-linkage. Five subfamilies of starch synthases have been identified in higher plants, including granule-bound starch synthase (GBSS), starch synthase I (SSI), starch synthase II (SSII), starch synthase III (SSIII), and starch synthase IV (SSIV). GBSS is essential for amylose synthesis and is exclusively bound to the starch granule. SSI, SSII, SSIII, and SSIV (named as SSV in dicots) are responsible for amylopectin chain elongation with their distribution between the granular and soluble fractions (Ball and Morell, 2003; Li *et al*., 2003). Each class of SS genes plays a distinct role in the synthesis of amylopectin. Analysis of a rice mutant with a retrotransposon inserted into a gene that encodes for SSI indicates that amylopectin is depleted in chains of DP8-12 and enriched in chains of DP6 and 7 in the mutant. This suggests a distinct capacity of SSI for the synthesis of chains with DP8-12 from DP6-7 (Nakamura, 2002). SSII has a specific role in the synthesis of intermediate-length chains (B2 and B3 chains) (Fontaine *et al*., 1993; Craig *et al*., 1998; Edwards *et al*., 1999; Morell *et al*., 2003). Mutations eliminating SSIII activity are known in maize as

* These authors contributed equally to this study.
† To whom correspondence should be addressed. Fax: +86 571 86971323. E-mail: clspwu@zju.edu.cn

du1 (Gao *et al*., 1998) and antisense suppression of *SSIII* has been reported in the potato (Edwards *et al*., 1999). The suppression of *SSIII* activity reduced amylopectin synthesis and therefore the length of development of the amylopectin long chains (DP25-35). This suggests that *SSIII* preferentially synthesizes long B1 and B2 chains.

The origin of these five starch synthase subfamilies is clearly ancient, given the conservation of the orthologous genes from *Chlamydomonas* through to the dicots and monocots (Ball and Morell, 2003). Evidence is increasing from a range of species that two members of GBSS (Fujita and Taira, 1998; Vrinten and Nakamura, 2000; Edwards *et al*., 2002; Dian *et al*., 2003), and two classes of SSII (Harn *et al*., 1998; Jiang *et al*., 2004) are present and functionally divergent with one member expressed in seeds and another member expressed in other tissues.

In rice, six distinct genes encoding starch synthases have been reported (Wang *et al*., 1990; Baba *et al*., 1993; Dian *et al*., 2003; Jiang *et al*., 2004). The identification of two genes encoding *SSIII* and two genes encoding *SSIV* in rice is reported here. The gene structure, expression patterns, genomic distribution, and evolution of the starch synthase genes have also been investigated and discussed. The expression pattern of the genes for starch synthesis and their role in starch accumulation in cereal seeds is also a key focus.

In the present paper, the two-letter prefix, *Os*, is used to indicate rice (*Oryza sativa* L.). The genes were designated *OsSSIIIa*, *OsSSIIIb*, *OsSSIVa*, and *OsSSIVb*, in chronological order of the submission to GenBank.

Materials and methods

Plant material

The indica rice (*Oryza sativa* L.) variety Zhe733 and japonica rice variety Zhenongda104 (for zymogram analysis) were grown in a greenhouse at 22/28 °C (night/day) and 80% relative humidity. The grains were harvested several times during maturation while the leaves and roots were taken from 4-leaf-stage seedlings.

For the sugar effect test, the 4-leaf-stage seedlings were cultured for 2 d under dark condition to deplete endogenous sugars before treatments with sugars. After 2 d, the second and third leaves were excised from the plants (now about 1 cm in length) and then transferred and incubated in rice nutrition solution (Yoshida *et al*., 1976) with different sugar or sugar analogues in the dark at 28 °C. Excised leaves were also incubated in a rice solution without sugar as a control. The leaf samples were harvested after 12 h of sugar treatments.

cDNA cloning of OsSSIII *and* OsSSIV

The TBLASTN searching of the GenBank (http://www.ncbi.nlm.nih.gov) databases with the conserved C-terminal amino acids of rice starch synthases yielded four highly homologous rice genomic clones: AL606645 (OSJN00079), AP005441, AP003292, and AC121365. The predicted polypeptides encoding by AL606645 and AP005441 bear an analogy to maize SSIII/DU1, named as *OsSSIII-1* and *OsSSIII-2*, while those of AP003292 and AC121365 bear an analogy to wheat SSIV, named as *OsSSIV-1* and *OsSSIV-2*, respectively. The *OsSSIII-1* cDNA clone containing the complete coding domain sequence (CDS) and an *OsSSIII-1* cDNA fragment (lack 560 bp at the 5′ region of the putative ORF) were isolated by screening the immature seed cDNA library (Jiang *et al*., 2004). For sequencing, the positive clones were converted to pTriplEx2. The plasmid was prepared using GFX™ Micro Plasmid Prep Kit (Amersham Pharmacia Biotech), and then the cDNA fragments were subcloned into the pBsSK vector and sequenced (MegaBACE™ 1000, Amersham Pharmacia Biotech). To clone the two putative *SSIV* cDNA, two primer pairs were designed for each gene. S411 [5′-AGT GCC CCT CCT CGC CCG CTT G (−24 to −3 bp, A –ATG is No. 1)] and S412 [5′-GTC GTA TTT GGG AAG AAC AAT CTC (+1540 to +1563 bp)], S413 [5′-AAA AAA GGT ACC GGA AAC AGT TCT GGC TTG CAC ATT GTC-3′ (+1414 to +1440 bp)] and S414 [5′-AAA AAA GAG CTC ATG TTT GAG CCG CTC CCC TTG CCT G-3′ (+2893 to +2917 bp)]; S421 [5′-TTC CCC AGC CTC CGC ATC CGA TTC (−26 to −3 bp)] and S422 [5′-TCA CAT CCG CCA AAC CAC CAA CC (+1308 to +1330 bp), S423 [5′-AAA AAA GGT ACC ACT AGT TCT GGC TTG CAC ATC ATC C (+1252 to +1276 bp)] and S424 [5′-AAA AAA GAG CTC TGC TGC TGC CCT CGC TCG AGC CAC TG (+2720 to +2745 bp)], respectively. The amplified fragments were cloned in the pUCm-T vector and sequenced (MegaBACE™ 1000, Amersham Pharmacia Biotech). GenBank accession numbers of these four cDNA clones were AF432915 (*OsSSIII-1*), AY100469 (*OsSSIII-2*), AY373257 (*OsSSIV-1*), and AY373258 (*OsSSIV-2*), respectively.

Expression of OsSSIII *and* OsSSIV *in* E. coli

The following primers were used for PCR modification of the N-termini of *OsSSIIIa*, *OsSSIIIb*, *OsSSIVa*, and *OsSSIVb*: pS311 [5′-AAA TTG CAT ATG ATG GAC TTT GTT TTC TCT-3′ (+2194 to +2214)] and pS312 [5′-AAA TTG GAT ATC TCA GTT CTT GCG AGC GGA ATG G-3′ (+3630 to +3651)]; pS321 [5′-AAA TTG CAT ATG ATG GAC TTT GTT TTC TCT-3′ (+3908 to +3930)] and pS322 [5′-AAA AAA GAG CTC TCA AAA TTT gGTG AGC TGA ATG-3′ (+5346 to +5367)]; S413 [5′-AAA AAA GGT ACC GGA AAC AGT TCT GGC TTG CAC ATT GTC-3′ (+1414 to +1440)] and S414 [5′-AAA AAA GAG CTC ATG TTT GAG CCG CTC CCC TTG CCT G-3′ (+2893 to +2917)]; S423 [5′-AAA AAA GGT ACC ACT AGT TCT GGC TTG CAC ATC ATC C (+1252 to +1276 bp)] and S424 [5′-AAA AAA GAG CTC TGC TGC TGC CCT CGC TCG AGC CAC TG (+2720 to +2745 bp)]. These primers were used to introduce *Nde*I/*Eco*RV, *Nde*I/*Sac*I, *Kpn*I/*Sac*I, and *Kpn*I/*Sac*I restriction sites at the C-termini of OsSSIII-1, OsSSIII-2, OsSSIV-1, and OsSSIV-2, respectively. The reconstructed sequences were subcloned into pET-29b (Novagen), named pSSIII-1, pSSIII-2, pSSIV-1, and pSSIV-2, respectively. The SDS-PAGE analysis of the expressed proteins and enzyme assays were carried out as described previously (Dian *et al*., 2003).

Semiquantitative RT-PCR analysis

Total RNA from leaves, roots, and grains of plants were isolated with Trizol (GIBICOL). PCR amplifications were performed on first cDNA strand corresponding to 2 μg of total RNA, using eight specific primer sets (Table 1). All specific primers were designed to span intron sequences, thus allowing the signals resulting from genomic DNA contamination to be accounted for. Primers (Table 1) that amplify *Actin* (Genbank accession number X16280) from rice were used as a control. The amplified fragments were cloned in the pUCm-T vector and sequenced (MegaBACE™ 1000, Amersham Pharmacia Biotech). Thermocycling time and temperature were as follows: 95 °C for 5 min, followed by 30 cycles (or as indicated) of 95 °C for 30 s, 58 °C for 30 s, 72 °C for 1 min, and a final extension period 72 °C for 7 min. PCR products were separated on 1% agarose gels, and stained with ethidium bromide.

Gel electrophoresis and immunoblot analysis

Native PAGE was carried out using 7.5% (resolving gel) and 2.5% (stacking gel) polyacrylamide and activity staining was performed as

Table 1. *Primer pairs for semiquantitative PCR amplification*

Gene	Primer pairs[a]		Fragment size[b]
OsSSI	F, 5′-GCGGATGGCAATTTCGACATACAG;	R, 5′-ATAAGAACAGCACAGGCGACAAAG	538
OsSSII-2	F, 5′-GGCGGCGGCGGATCCTATGTC;	R, 5′-TGTTTCTTGCCGGTGTCCAGCGTC	1487
OsSSIII-1	F, 5′-TTTATGCTGGTGCCGACTTC;	R, 5′-GAACTCACAACCGCAGGATAAC	577
OsSSIII-2	F, 5′-TCGGAAGAAGGTGGAATCTATG;	R, 5′-TTGTGGCTTTGTCGCAGTATG	618
OsSSIV-1	F, 5′-AGTGCCCCTCCTCGCCCGCTTG;	R, 5′-GTCGTATTTGGGAAGAACAATCTC	1587
OsSSIV-2	F, 5′-TTCCCCAGCCTCCGCATCCGATTC;	R, 5′-TCACATCCGCCAAACCACCAACC	1356
Actin	F, 5′-GGAACTGGTATGGTCAAGGC;	R, 5′-AGTCTCATGGATACCCGCAG	775

[a] F, forward primer; R, reverse primer.
[b] Numbers indicate the size (in bp) of amplified fragments.

previously described (Jiang *et al*., 2003). Each lane was loaded with 50 μg of the crude extract from leaves (L) or endosperms. Starch synthase activity was detected after incubation of the gel in 50 mM Tricine-NaOH (pH 8.5), 0.5 M sodium citrate, 25 mM potassium acetate, 2 mM EDTA, 2 mM DTT, 0.1% (w/v) potato amylopectin, and the indicated content of ADP-glucose at the indicated temperature for 10 h.

In immunoblot analysis, the following antisera were used: antisera (Anti-ZmSSIII), raised against the N-terminal region of the maize SSIII protein, were the generous gift of Dr Alan M Myers (Iowa State University, Ames) and are described by Cao *et al*. (1999). The rabbit polyclonal antisera (designated Anti-OsSSI and Anti-OsSSIII) were raised against the deduced mature OsSSI or the C-terminus of OsSSIII-1 expressed in *E. coli* (Jiang *et al*., 2004). The Anti-OsSSII-2 and Anti-OsSSII-3 are described by Jiang *et al*. (2004). Protein extracts were separated by 7.5% Native-PAGE or 7.5% SDS–PAGE, transferred to polyvinylidenedifluoride (PVDF) membranes (AMRESCO), and visualized using rabbit anti-SS antiserum followed by alkaline phosphatase–conjugated goat anti-rabbit serum (Santa Cruz, USA).

Results

Cloning and characterization of OsSSIII *and* OsSSIV *genes*

BLAST analysis of the rice genome (http://www.ncbi.nlm.nih.gov/BLAST/) using the conserved C-terminus of starch synthase amino acid sequences, revealed four new homologous genes. According to the putative amino acid sequences, two genes were homologous to maize *SSIII/DU1*, designated *OsSSIII-1* and *OsSSIII-2*, and the other two genes homologous to wheat *SSIV*, designated *OsSSIV-1* and *OsSSIV-2*. The cDNA of the four genes were cloned by screening a cDNA library or RT-PCR. According to the cDNA sequences, rice *SSIII* genes contain 16 exons separated by 15 introns as in wheat. The sequences of exon3 are the variable repeat region (Li *et al*., 2000), and vary between *OsSSIIIa* and *OsSSIIIb* in length. The structures of the *SSIV* genes contain 16 exons separated by 15 introns, the first three exons in the 5′ region of the genes are variable in length.

Duplication and distribution of SS *gene families in rice*

The four genes of *SSIII-1/-2* and *SSIV-1/-2* cloned in this case together with the other six genes cloned previously (Wang *et al*., 1990; Baba *et al*., 1993; Dian *et al*., 2003; Jiang *et al*., 2004), compose 10 members of the *SS* gene family in rice with five different subfamilies distributed over eight rice chromosomes. *SSIII-1, SSIII-2, SSIV-3*, and *SSIV-4* were located on chromosomes 4, 8, 1, and 5, respectively. However, many genes clustered around the loci of *SSIII-1* and *SSIII-2* and the loci of the *SSIV-1* and *SSIV-2* were identical (Fig. 1). This observation indicated that the duplications of *SSIII* and *SSIV* were associated with the duplication of a large chromosome segment during evolution. In addition to the 10 genes, a truncated SSI gene was identified on chromosome 10 and a truncated SSII-3 gene (Accession number BAC16084) on chromosome 7. The putative sequence of amino acids of the truncated SSI shared high identity with C-terminus of *OsSSI* (Fig. 2).

Domain organization of OsSSIII and OsSSIV and expression of OsSSIII and OsSSIV in E. coli

As is the case in the SSIII protein in maize and wheat (Gao *et al*., 1998; Li *et al*., 2000), rice SSIII proteins also contain four distinct regions. These are a putative transit peptide region (47/49 amino acids for OsSSIII-1/-2) (identified using the ChloroP neural network analysis of the 100 amino acids at the N-terminus of each sequence), a variable repeat region, a SSIII specific region, and a C-terminal region that contains the catalytic domain (Fig. 3A). The SSIII specific region in OsSSIII-1/-2 was composed by a set of three repeated units of amino acid similarity in each gene. Thirteen amino acids were conserved in these repeat units (Fig. 3B). Rice SSIV proteins contain three distinct regions: a putative transit peptide region (78/33 amino acids for SSIV-1/-2), a region homologous to Smc (COG1196)/myosin tail 1 (pfam01576) and a C-terminal catalytic domain region (Fig. 3A).

The deduced amino acid sequences of the C-terminal catalytic domain region from the ten OsSS proteins were aligned with one prokaryotic glucogen synthase (*E. coli* glucogen synthase, EcGS) using the Clustal W program. Nine conserved regions were identified. The invariant residues were listed (Fig. 3C). The first homologous region contains the consensus motif KXGGL, which is believed to be the ADP-Glc binding site of starch synthase (Furukawa *et al*. 1990).

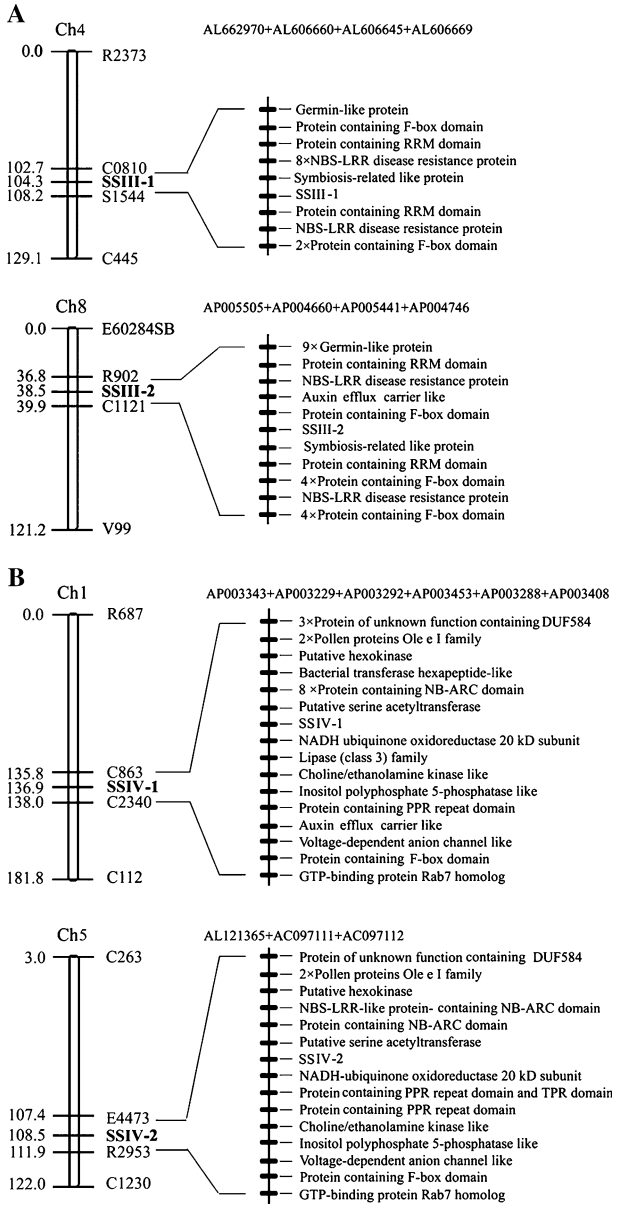

Fig. 1. Gene arrangement of the two *SSIII* gene loci (A) two *SSIV* gene loci (B) in rice. Location of the four genes, anchor RFLP markers and Kosambi values (cM) are indicated on the chromosomes (the High-Density Rice Genetic Map on http://rgp.dna.affrc.go.jp/ publicdata). The genomic sequences were analysed with GeneMark (http://www.ebi.ac.uk/genemark) and Genscan (http://genes.mit.edu/GENSCAN.html).The orthologous genes are listed.

To determine whether the *OsSSIII* and *OsSSIV* genes encode authentic starch synthases, expression (Fig. 4) of *OsSSIII-1*, *OsSSIII-2*, *OsSSIV-1* and *OsSSIV-2* genes in *E. coli* (BL21) were tested. The activities of starch synthase were increased by 4.1-fold, 4.8-fold, 2.7-fold, and 2.4-fold above baseline activity level of *E. coli* glycogen synthase, for each of the four genes, respectively. This indicated that the *OsSSIII* and *OsSSIV* genes encoded functional starch synthase enzymes in rice.

>AC092172 (OSJNBa0014J14) chr10 191397bp

```
OsSSI      244   QFRYTLLCYAACEAPLILELG 264
                 QFR TLLCYAACEAPLILELG
AC092172 86038 QFRDTLLCYAACEAPLILELG 86100

OsSSI      287   LLAAKYRPYGVY 298
                 LLAAKY PY VY
AC092172 86336 LLAAKYSPYDVY 86371

OsSSI      314   GVEPASTYPDLGLPPEWYGALEWVFPE 340
                 GVEPASTYP+LG+ PEWYG+LEWVFPE
AC092172 86434 GVEPASTYPNLGMSPEWYGSLEWVFPE 86514

OsSSI      486   GDPGFEGWMRSTESGYRDKFRGWVGFSVPVSHRITA 521
                 GDPGFEGWMRS ES Y+DKFRGWVGF+V VSHRITA
AC092172 86741 GDPGFEGWMRSIESEYKDKFRGWVGFTVLVSHRITA 86848

OsSSI      523   CDILLMPSRFEPCGLNQLYAMQYGTVPVVHGTGGLR 558
                 CDILLM SRFEPCGLNQLYAMQYG VPVVHGTGGLR
AC092172 86977 CDILLMSSRFEPCGLNQLYAMQYGIVPVVHGTGGLR 87084
```

Fig. 2. The amino acid sequence alignment of OsSSI and the other destroyed OsSSI orthologous gene in rice. The highly homologous rice genomic sequences were yielded by tBLASTn searching of the GenBank databases queried with OsSSI protein. The middle lane shows the conserved amino acids.

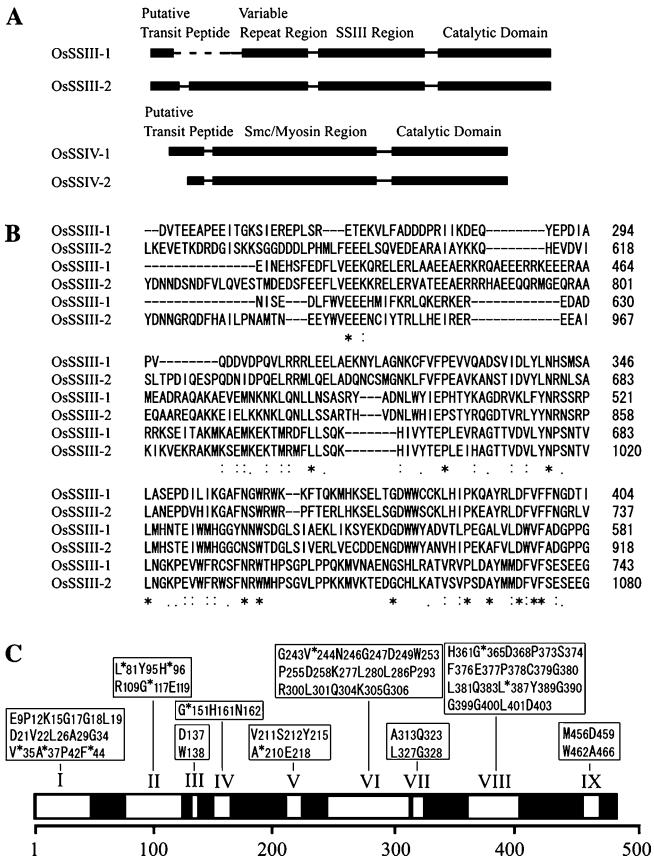

Fig. 3. (A) Domain structure of the rice SSIII and SSIV amino acid sequences showing the conserved motifs. Putative transit peptide cleavage sites were identified using the ChloroP neural network analysis of the 100 amino acids at the N terminus of each sequence. (B) SSIII-specific region. Alignment of the deduced amino acid sequence of the three repeated units of OsSSIII using Clustal W software. Asterisks indicate 100% conserved, double dots similar, single dots related amino acids; dashes depict gaps. Numbers refer to amino acids. (C) Identification of residues invariant to *E. coli* glycogen synthase (EcGS) and rice starch synthases. Invariant residues that are specific to rice starch synthases are marked with an asterisk. All amino acid numbers correspond to the sequence of EcGS (Accession number NC_000913).

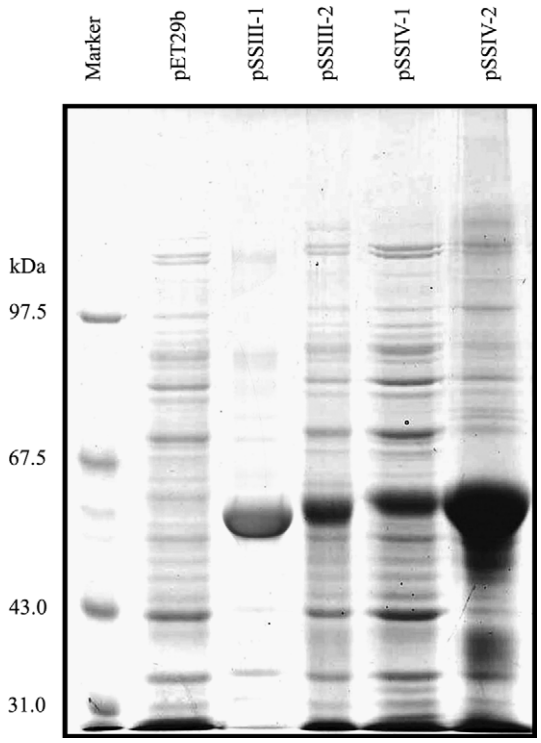

Fig. 4. SDS-PAGE of the bacterially expressed SS proteins. pET29b, cell lysate from *E. coli* BL21 transformed with native plasmid pET-29b; pSSIII-1, pSSIII-2, pSSIV-1, and pSSIV-2, cell lysate from *E. coli* expressing OsSSIII-1, OsSSIII-2, OsSSIV-1, and OsSSIV-2. Protein standard ladder (marker) was indicated on the left. The gel was stained with the dye of Coomassie Brilliant Blue R-250.

Organ expression profile of OsSSIII *and* OsSSI*V genes*

The spatial expression of *OsSSIII-1*, *OsSSIII-2*, *OsSSIV-1*, and *OsSSIV-2* was examined in storage and non-storage organs of rice by RT-PCR (Fig. 5A). The result shows that *OsSSIII-1* and *OsSSIV-2* were expressed mainly in leaves and weakly in endosperms, while *OsSSIII-2* and *OsSSIV-1* mainly in endosperms and weakly in leaves. Total proteins extracted from rice leaves and endosperms were electrophoresed in SDS–PAGE gels using antibodies to maize SSIII polypeptide (N-terminus) (anti-ZmSSIII) and anti OsSSIII-1 polypeptide (C-terminus) (anti-OsSSIII-1). Both antibodies were bound to 230 kDa proteins from rice endosperms and 170 kDa proteins from rice leaves and endosperms (Fig. 5B). The calculated molecular mass of OsSSIII-1 and OsSSIII-2 was 138 kDa and 201 kDa, respectively. The apparent molecular mass of SS as estimated on SDS-PAGE, was often noted to be larger than the calculated molecular mass (Knight *et al.*, 1998), suggesting that the 230 kDa protein should be OsSSIII-1 and the 170 kDa protein should be OsSSIII-1. This result is consistent with the expression patterns of *OsSSIII*-1 and *OsSSIII-2b* transcripts. *OsSSIII-1* was expressed at the early developing stage of endosperm, while *OsSSIII-2* reached maximum levels at the middle age. The transcription of

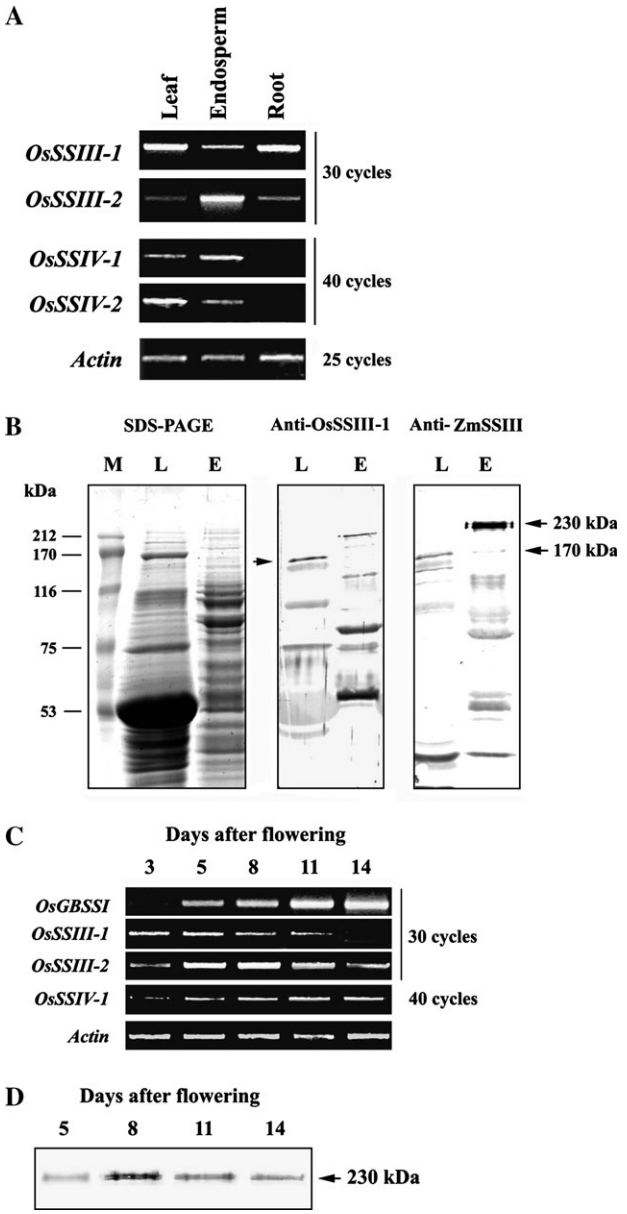

Fig. 5. Expression analysis of *OsSSIII* and *OsSSIV*. (A) Tissue-specific RT-PCR analysis. Total RNA samples from leaf, endosperm, and root were used for RT-PCR to measure RNA expression levels of the *OsSS* genes. Thermocycling time and temperature were as follows: 95 °C for 5 min, followed by indicated cycles of 95 °C for 30 s, 58 °C for 30 s, 72 °C for 1 min, and a final extension period 72 °C for 7 min. Lower panel shows the loading control of an *Actin* transcripts in each sample. (B) SDS-PAGE and immunoblot analysis of OsSSIII. Each lane was loaded with 50 μg protein. The proteins were extracted from 5 d after flowering endosperms (E) and 4-leaf stage seedlings leaves (L). The SDS–PAGE lanes indicate the Coomassie Brilliant Blue R stained proteins, while Anti-ZmSSIII and Anti-OsSSIII-1 indicate SDS–PAGE-separated proteins immunoblotted with polyclonal antibodies raised against the ZmSSIII proteins or OsSSIII-1 proteins. The protein markers (M) were indicated on the left. (C) RT-PCR analysis of endosperm samples. The RNA samples were taken at various days after flowering. The lower panel shows the loading control of the *Actin* transcripts in each sample. The PCR was performed as described above. (D) Immunoblot analysis of OsSSIII-2 proteins in endosperm samples. Each lane was loaded with total protein extracted from 1 mg (dry weight) of endosperms as the indicated days after flowering.

OsSSIV-1 reached maximum levels in endosperm at the late developmental stage (Fig. 5C). Western blotting also indicated that OsSSIII-2 protein maintains higher levels at the middle developmental stage of rice endosperms (Fig. 5D).

Zymogram analysis for SS activities in rice endosperm and leaf extracts

The zymogram experiment was employed to identify further the expression of different SS genes in rice endosperms and leaves. Figure 6A shows that at least four SS activity bands were detected in rice leaves and five SS activity bands in rice endosperms. Immunoblotting showed that the lowest mobility activity band in endosperms should be that of OsSSIII-2, and in leaves should be that of OsSSIII-1 (Fig. 6B). Rice SSIII-2 showed temperature-sensitive activity, with notably lower starch synthase activity at 37 °C than in 25 °C. Figure 6B also shows that the expression of OsSSI was not tissue-specific, whilst OsSSII-2 was leaf-specific and OsSSII-3 was endosperm-specific. The mobility of OsSSII-2 in leaves and OsSSII-3 in endosperms is seen as equal. SSI and SSII showed an identified band in many rice cultivars such as Zhenongda104, Nipponbare and Kasalath (the data of last two cultivars is not shown).

Influence of sugar level on SS gene expression in rice leaves

It has been demonstrated that the gene expression *of OsGBSSII* is regulated by sugar level in excised leaves (Dian *et al*., 2003). The effects of sugars on the other SS genes that expressed in rice leaves were tested in this case. Excised leaves were treated with 175 mM of mannitol (mal), sucrose (suc), 3-*O*-methyl-glucose (3-*O*MG), and 2-deoxyglucose (2-DOG), respectively. RT-PCR analysis indicated that accumulation of *OsSSI*, *OsSSII-2*, and *OsSSIII-1* transcripts were up-regulated by sucrose, but not for *OsSSIII-2* transcripts. Glucosamine eliminated the accumulation of these transcripts induced by sucrose (Fig. 7). Sequence analysis has revealed the sugar response elements in the promoter regions (up to 2000 bp from the initiation codon of ATG) of the three *SS* genes: *OsSSIIIa*, SP8BFIBSP8BIB (TACTATT) (Ishiguro and Nakamura, 1992), −1431, SURE2STPAT21 (AATACTAAT) (Grierson *et al*., 1994), −1758; *OsSS1*, SURE1STPAT21 (AATAGAAAA) (Grierson *et al*., 1994), −830; *OsSSII-2*, SURE1STPAT21, −362. This suggested that the sugar regulation of these SS genes in rice leaves is consistent and the sugar response elements are active in rice.

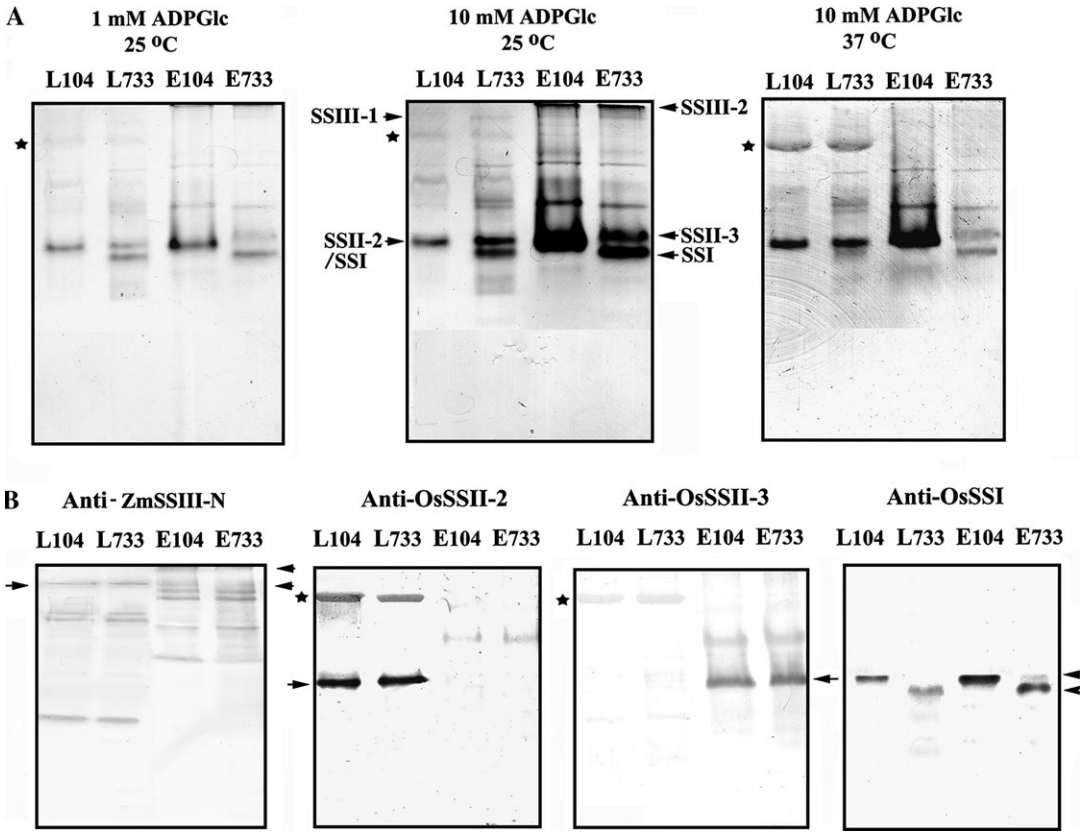

Fig. 6. Effects of various sugars and hexokinase inhibitor on partial *SS* gene expression in rice leaves. The excised leaves were incubated in nutrient solution with 175 mM of mannitol (Mal), sucrose (Suc), glucosamine (GlcN), sucrose and glucosamine (Suc/GlcN), 2-deoxyglucose (2-DOG), and 3-*O*-methyl-glucose (3-OMG), respectively, and maintained in the dark for 12 h at 28 °C. The excised leaves incubated in rice nutrient solution as a control (CK). The lower panel shows the normalization of total RNA levels in each sample used for reverse transcription.

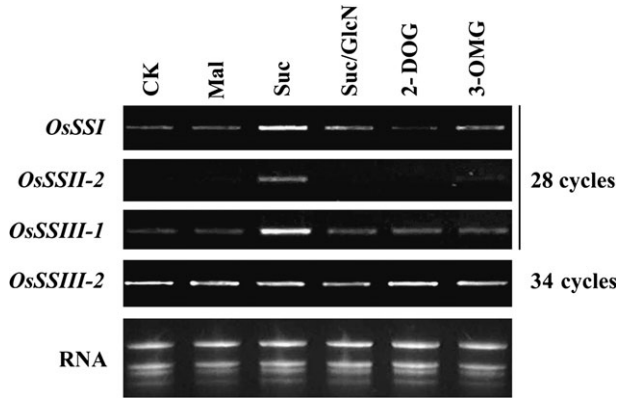

Fig. 7. Zymogram analysis and western blotting of *SS* of rice leaves and endosperms. (A) After separation of SS isoforms by native-polyacrylamide gel electrophoresis, the gels were incubated at 25 °C or 37 °C for 20 min, and then the enzyme reaction was performed with the addition of 0.1% (w/v) potato amylopectin, 0.5 mol l^{-1} citrate and indicated concentration of ADPGlc at 25 °C or 37 °C for 10 h. Each lane was loaded with 50 μg of the crude extract from leaves (L) or endosperms (E). 104, rice cultivar zhenongda 104 (japonica); 733 rice cultivar zhe733 (indica). (B) The proteins were separated by native-polyacrylamide gel electrophoresis, and immunoblotted by using the polyclonal antibodies raised against the ZmSSIII, OsSSII-2, OsSSII-3, and OsSSI proteins, respectively. Each lane was loaded with 30 μg soluble proteins of the crude extract from leaves or endosperms. Stars indicated the bands of Rubisco protein in leaves if it appeared.

Discussion

Characterization of SSIII *and* SSIV *genes in rice*

Five subfamilies of starch synthase genes have been reported in higher plants (Ball and Morell, 2003). Among them, three subfamilies have been characterized in rice, including *GBSS* (Wang *et al*., 1990; Dian *et al*., 2003), *SSI* (Baba *et al*., 1993; Tanaka *et al*., 1995), and *SSII* (Jiang *et al*., 2004). In this paper the genes of other two subfamilies, *SSIII* and *SSIV*, in rice were cloned and characterized. The gene structure analysis indicates that the arrangements of exon and intron of *SSIII* and *SSIV* genes are highly conserved in rice, wheat, and *Arabidopsis*, respectively, whereas they differ from those of *GBSS*, *SSI*, and *SSII* genes (Li *et al*., 2003; Dian *et al*., 2003; Jiang *et al*., 2004). SSIII proteins contain a SSIII-specific region with a set of three repeated units, while SSIV proteins contain a region homologous to the Smc motif. According to these observations, it is clear that the separation of GBSS, SSI, SSII, SSIII, and SSIV predates the separation of monocotyledonous and dicotyledonous plants. The proteins of OsSSIII-1/-2 and OsSSIV-1/-2 contain conserved domains characteristic of GlgA (glycogen synthase) at the C-termini. Each class of GlgA domain can catalyse starch synthesis when expressed in *E. coli*. This result indicates that the *SSIII* and *SSIV* genes encode functional starch synthase enzymes in rice.

Evolution of the SS *gene families in rice*

It has been reported that *GBSS* and *SSII* subfamilies have diverged through gene duplication events in cereals (Harn *et al*., 1998; Vrinten and Nakamura, 2000; Dian *et al*., 2003; Jiang *et al*., 2004). The present observation indicated that other two subfamilies, *SSIII* and *SSIV*, are also duplicated and diverged in rice. The homologous ESTs of *SSIII* and *SSIV* genes are detected in other cereals in GenBank and Tiger EST databases (data not shown), which suggests the duplication and divergence of *SSIII* and *SSIV* presented in other Gramineae. Although the duplicated *SSI* genes have been destroyed, gene duplication has also been identified (Fig. 2). This is the first report that the duplication event occurred in all the five subfamilies genes for starch synthase in Gramineae.

Subfunctionalization of the duplicated genes for starch synthesis in rice

The expression patterns of the five *SS* gene subfamilies were investigated as a first step toward understanding their respective functions. Previous observations indicated that the transcripts and proteins of the *OsSSII-2* and *OsGBSSII* genes were detected mainly in leaves, while *OsSSII-3* and *GBSSI* (*Wx*) primarily in endosperms (Harn *et al*., 1998; Vrinten and Nakamura, 2000; Dian *et al*., 2003; Jiang *et al*., 2004). The present observation with reverse transcription-PCR, western blotting or zymogram, indicate that *OsSSIII-2* and *OsSSIV-1* are mainly expressed in endosperm, while *OsSSIII-1* and *OsSSIV-2* are mainly expressed in leaves (Figs 5, 7). In rice endosperms, *OsSSIII-1* was expressed at an early developing stage, while *OsSSIII-2* reached maximum levels at the middle stage. In rice leaves, as for the *OsGBSSII*, the expression of *OsSSI*, *OsSSII-2*, and *OsSSIIIa* is regulated by sugars, suggesting that sugar regulation of these *SS* genes converged during rice evolution. The consistent regulation of these *OsSS* genes allows starch synthesis to be modulated in response to the accumulation of sugars in rice leaves under conditions such as high light and nitrogen or phosphate starvation.

In cereals, evidence is increasing that the different subfamilies of genes for starch synthesis may be partitioned into two gene groups, with one group expressed in the endosperm and a second mainly expressed in other tissues. For instance, *BEIIb (RBE3)* is specifically expressed in endosperms, while *BEIIa* (*RBE4*) is mainly expressed in vegetative tissues and in the early developing endosperms in cereals (Yamanouchi and Nakamura, 1992; Gao *et al*., 1996; Morell *et al*., 1997; Sun *et al*., 1998; Jiang *et al*., 2003; Mutisya *et al*., 2003). Transcripts of *AgpL1* and *AgpS1a* genes were specifically detected in endosperms, while *AgpL2*, *AgpS1b*, and *AgpS2* are mainly expressed in leaves of maize, wheat, barley, and rice (*AgpS1a* and *AgpS1b* are encoded by a single gene in wheat, barley, and rice). *AgpL2*, *AgpS1b*, and *AgpS2* are likewise expressed in the early developing endosperm in wheat, barley, and rice (Burton *et al*., 2002; Johnson *et al*., 2003). Thus, in cereals, the duplicated starch synthesis genes diverged into two

groups: Group I, preferentially expressed in endosperms, including *AgpL1*, *AgpS1a*, *GBSSI*, *SSII-3* (*SSIIa*), *SSIII-2* (*DU1*), *SSIV-1* (may be including *SSIV-2*), and *BEIIb (RBE3*), and Group II, mainly expressed in other tissues, and some in the early developing endosperms, including *ApgL2*, *AgpS1b*, *AgpS2*, *GBSSII*, *SSII-2 (SSIIb*), *SSIII-1*, *SSIV-2 (*may be including *SSIV-1*), and *BEIIa* (*RBE4*). The single member subfamily genes of *SSI* and *BEI* (*RBE1*) are expressed constitutively and function in both groups (Fig. 8). Thus the preservation of these duplicate genes is likely to be by subfunctionalization (Lynch and Force, 2000), with the primary function of the Group II genes on transient starch synthesis, and Group I genes on storage starch synthesis in rice.

Expression profiles of the corresponding genes suggest a two-step model for starch biosynthesis in rice

It is already known that angiosperm seeds accumulate storage protein, oil, and carbohydrate during seed filling and degrade them to support early seedling growth upon germination. The relative abundance of storage reserves varies among seeds of different species. Many oilseeds produce 50–70% oil, some legumes contain 40% protein, whereas most cereal seeds contain 70–85% of the seed dry weight as starch (Baud *et al*., 2002). In developing legume seeds, the model of detailed gene expression regarding the control of starch and protein synthesis has been established through the proteomics research on *Medicago truncatula* seed development (Gallardo *et al*., 2003). Contrapuntal networks of gene expression during the *Arabidopsis* seed-filling period reveal that the genes involved in carbohydrate metabolism peaks are expressed in early seed development, those required for oil synthesis are expressed later, and the genes for oil-body and storage-protein are expressed later still (Ruuska *et al*., 2002). Although the cytosolic localization of AGPase in cereal endosperms has been suggested to have functional significance for partitioning large amounts of carbon into starch when sucrose is plentiful (Beckles *et al*., 2001), the accumulation of storage components requires the co-ordination of many genes that encode the enzymes of the corresponding pathways (Ruuska *et al*., 2002). In rice, the expression of Group II starch synthesis genes peaks early in seed development, similar to that in the developing seeds of *Arabidopsis* (Fig. 8B), whilst the expressed pattern of Group I starch synthesis genes was consistent with the age of the rapid accumulation of starch in the endosperms. This suggests that this is one of the primary factors in the synthesis and storage of starch in rice endosperms (Fig. 8B).

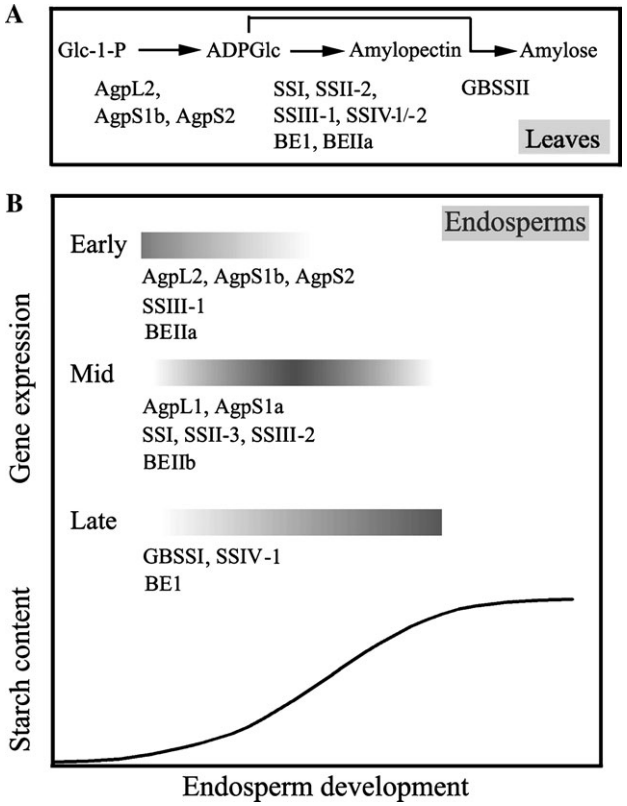

Fig. 8. Differential expression of the genes involved in starch biosynthesis in leaves and endosperms of cereals. (A) The expressed genes involved in starch biosynthesis in leaves. (B) Accumulation of starch in cereal seeds and expression patterns of genes that are involved in starch synthesis. The expression of BEI peaks at the late endosperm development in maize (Gao *et al*., 1996), wheat (Morell *et al*., 1997), and rice (Jiang *et al*., 2003).

The expression profiles of these genes are likely to suggest a two-step model for starch biosynthesis in rice endosperms. Group II genes synthesize the early transient starch mainly in leaves and Group I genes synthesize the storage starch in rice endosperms. This result is also in support of the hypothesis of subfunctionalization regarding the preservation of the duplicated genes in the genome (Lynch and Force, 2000). Therefore this can help in understanding the co-evolution of the relevant genes regarding the functional divergences in duplicated genes as revealed in gene expression changes in plants.

Acknowledgements

The authors thank Dr Alan M Myers for the generous donation of antibodies to the maize DU1 protein. This project was supported by the Natural Science Foundation of China, the Natural Science Foundation of Zhejiang Province (Grant No. ZA0106), and the China Postdoctoral Science Foundation (Grant No. 2003034068).

References

Baba T, Nishihara M, Mizuno K, Kawasaki T, Shimada H, Kobayashi E, Ohnishi S, Tanaka K, Arai Y. 1993. Identification, cDNA cloning, and gene expression of soluble starch synthase in rice (*Oryza sativa* L.) immature seeds. *Plant Physiology* **103,** 565–573.

Ball SG, Morell MK. 2003. From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. *Annual Review of Plant Biology* **54,** 207–233.

Baud S, Boutin J, Miquel M, Lepiniec L, Rochat C. 2002, An integrated overview of seed development in *Arabidopsis thaliana* ecotype WS. *Plant Physiology and Biochemistry* **40,** 151–160.

Beckles DM, Smith AM, Rees T. 2001. A cytosolic ADP-glucose pyrophosphorylase is a feature of graminaceous endosperms, but not of other starch-storing organs. *Plant Physiolgy* **125,** 818–827.

Burton RA, Jenner H, Carrangis L, *et al*. 2002. Starch granule initiation and growth are altered in barley mutants that lack isoamylase activity. *The Plant Journal* **31,** 97–112.

Cao H, Imparl-Radosevich J, Guan H, Keeling PL, James MG, Myers AM. 1999. Identification of the soluble starch synthase activities of maize endosperm. *Plant Physiology* **120,** 205–215.

Craig J, Lloyd JR, Tomlinson K, Barber L, Edwards A, Wang TL, Martin C, Hedley CL, Smith AM. 1998. Mutations in the gene encoding starch synthase II profoundly alter amylopectin structure in pea embryos. *The Plant Cell* **10,** 413–426.

Dian WM, Jiang HW, Chen QS, Liu FY, 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.

Edwards A, Borthakur A, Bornemann S, Venail J, Denyer K, Waite D, Fulton D, Smith A, Martin C. 1999. Specificity of starch synthase isoforms from potato. *European Journal of Biochemistry* **266,** 724–736.

Edwards A, Vincken JP, Suurs LC, Visser RG, Zeeman S, Smith A, Martin C. 2002. Discrete forms of amylose are synthesized by isoforms of GBSSI in pea. *The Plant Cell* **14,** 1767–1785.

Fontaine T, D'Hulst C, Maddelein ML, Routier F, Pepin TM, Decq A, Wieruszeski JM, Delrue B, Van den Koornhuyse N, Bossu JP. 1993. Toward an understanding of the biogenesis of the starch granule. Evidence that *Chlamydomonas* soluble starch synthase II controls the synthesis of intermediate size glucans of amylopectin. *Journal of Biological Chemistry* **268,** 16223–16230.

Fujita N, Taira T. 1998. A 56 kda protein is a novel granule-bound starch synthase existing in the pericarps, aleurone layers, and embryos of immature seed in diploid wheat (*Triticum monococcum* L.). *Planta* **207,** 125–132.

Furukawa K, Tagaya M, Inouye M, Preiss J, Fukui T. 1990. Identification of lysine 15 at the active site in *Escherichia coli* glycogen synthase. Conservation of a lys-X-gly-gly sequence in the bacterial and mammalian enzymes. *Journal of Biological Chemistry* **265,** 2086–2090.

Gallardo K, Le Signor C, Vandekerckhove J, Thompson RD, Burstin J. 2003. Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. *Plant Physiology* **133,** 664–682.

Gao M, Fisher DK, Kim KN, Shannon JC, Guiltinan MJ. 1996. Evolutionary conservation and expression patterns of maize starch branching enzyme I and IIb genes suggests isoform specialization. *Plant Molecular Biology* **30,** 1223–1232.

Gao M, Wanat J, Stinard PS, James MG, Myers AM. 1998. Characterization of *dull1*, a maize gene coding for a novel starch synthase. *The Plant Cell* **10,** 399–412.

Grierson C, Du JS, Zabala MT, Beggs K, Smith C, Holdsworth M, Bevan M. 1994. Separate *cis* sequences and *trans* factors direct metabolic and developmental regulation of a potato tuber storage protein gene. *The Plant Journal* **5,** 815–826.

Harn C, Knight M, Ramakrishnan A, Guan H, Keeling PL, Wasserman BP. 1998. Isolation and characterization of the zSSIIa and zSSIIb starch synthase cDNA clones from maize endosperm. *Plant Molecular Biology* **37,** 639–649.

Ishiguro S, Nakamura K. 1992. The nuclear factor SP8BF binds to the 5′-upstream regions of three different genes coding for major proteins of sweet potato tuberous roots. *Plant Molecular Biology* **18,** 97–108.

James MG, Denyer K, Myers AM. 2003. Starch synthesis in the cereal endosperm. *Current Opinion in Plant Biology* **6,** 215–222.

Jiang HW, Dian WM, Liu FY, Wu P. 2004. Molecular cloning and expression analysis of three genes encoding starch synthase II in rice. *Planta* **218,** 1062–1070.

Jiang HW, Dian WM, Wu P. 2003. Effect of high temperature on fine structure of amylopectin in rice endosperm by reducing the activity of the starch branching enzyme. *Phytochemistry* **63,** 53–59.

Johnson PE, Patron NJ, Bottrill AR, Dinges JR, Fahy BF, Parker ML, Waite DN, Denyer K. 2003. A low-starch barley mutant, riso 16, lacking the cytosolic small subunit of ADP-glucose pyrophosphorylase, reveals the importance of the cytosolic isoform and the identity of the plastidial small subunit. *Plant Physiology* **131,** 684–696.

Knight ME, Harn C, Lilley CE, Guan H, Singletary GW, MuForster C, Wasserman BP, Keeling PL. 1998. Molecular cloning of starch synthase I from maize (W64) endosperm and expression in *Escherichia coli*. *The Plant Journal* **14,** 613–622.

Li Z, Mouille G, Kosar-Hashemi B, Rahman S, Clarke B, Gale KR, Appels R, Morell MK. 2000. The structure and expression of the wheat starch synthase III gene. Motifs in the expressed gene define the lineage of the starch synthase III gene family. *Plant Physiology* **123,** 613–624.

Li Z, Sun F, Xu S, Chu X, *et al*. 2003. The structural organisation of the gene encoding class II starch synthase of wheat and barley and the evolution of the genes encoding starch synthases in plants. *Functional and Integrative Genomics* **3,** 76–85.

Lynch M, Force A. 2000. The probability of duplicate gene preservation by subfunctionalization. *Genetics* **154,** 459–473.

Morell MK, Blennow A, Kosar-Hashemi B, Samuel MS. 1997. Differential expression and properties of starch-branching enzyme isoforms in developing wheat endosperm. *Plant Physiology* **113,** 201–208.

Morell MK, Kosar-Hashemi B, Cmiel M, Samuel MS, Chandler P, Rahman S, Buleon A, Batey IL, Li Z. 2003. Barley *sex6* mutants lack starch synthase IIa activity and contain a starch with novel properties. *The Plant Journal* **34,** 173–185.

Mutisya J, Sathish P, Sun C, Andersson L, Ahlandsberg S, Baguma Y, Palmqvist S, Odhiambo B, Aman P, Jansson C. 2003. Starch-branching enzymes in sorghum (*Sorghum bicolor*) and barley (*Hordeum vulgare*): comparative analyses of enzyme structure and gene expression. *Journal of Plant Physiology* **160,** 921–930.

Myers AM, Morell MK, James MG, Ball SG. 2000. Recent progress toward understanding biosynthesis of the amylopectin crystal. *Plant Physiology* **122,** 989–997.

Nakamura Y. 2002. Towards a better understanding of the metabolic system for amylopectin biosynthesis in plants: rice endosperm as a model tissue. *Plant Cell Physiology* **43,** 718–725.

Ruuska SA, Girke T, Benning C, Ohlrogge JB. 2002. Contrapuntal networks of gene expression during *Arabidopsis* seed filling. *The Plant Cell* **14,** 1191–1206.

Smith AM, Denyer K, Martin C. 1997. The synthesis of the starch granule. *Annual Review of Plant Physiology and Plant Molecular Biology* **48,** 67–87.

Sun C, Sathish P, Ahlandsberg S, Jansson C. 1998. The two genes encoding starch-branching enzymes IIa and IIb are differentially expressed in barley. *Plant Physiology* **118,** 37–49.

Tanaka K, Ohnishi S, Kishimoto N, Kawasaki T, Baba T. 1995. Structure, organization, and chromosomal location of the gene encoding a form of rice soluble starch synthase. *Plant Physiology* **108,** 677–683.

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 Physiology* **122,** 255–264.

Wang ZY, Wu ZL, Xing YY, Zheng FG, Guo XL, Zhang WG, Hong MM. 1990. Nucleotide sequence of rice *waxy* gene. *Nucleic Acids Research* **18,** 5898.

Yamanouchi H, Nakamura Y. 1992. Organ specificity of isoforms of starch branching enzyme (Q-enzyme) in rice. *Plant Cell Physiology* **33,** 985–991.

Yoshida S, Forno DA, Cock JH, Gomez KA. 1976. *Laboratory manual for physiological studies of rice*, 3rd edn. Manila, The Philippines: International Rice Research Institute, 61–64.