

The amylose content in rice endosperm is related to the post-transcriptional regulation of the *waxy* gene

Zong-Yang Wang¹, Fei-Qin Zheng¹, Ge-Zhi Shen¹,
Ji-Ping Gao¹, D. Peter Snustad², Min-Gang Li²,
Jing-Liu Zhang¹ and Meng-Min Hong^{1,*}

¹Institute of Plant Physiology, Academia Sinica,
300 Fenglin Road, Shanghai 200032, People's Republic of
China, and

²Department of Genetics and Cell Biology, and Plant
Molecular Genetics Institute, University of Minnesota, St
Paul, MN 55108, USA

Summary

The *waxy* (*Wx*) gene of rice encodes a granule-bound starch synthase (GBSS = *waxy* protein) required for the synthesis of amylose in endosperm. An analysis of *Wx* transcripts, *Wx* protein, and amylose content of 31 rice cultivars revealed that endosperm amylose and *Wx* protein contents are correlated with the ability of the cultivar to excise intron I from the leader sequence of the *Wx* transcript. Cultivars with high endosperm amylose content (group I) contain high levels of amylose, *Wx* protein, and the mature 2.3 kb *Wx* mRNA. Cultivars with intermediate amylose content (group II) produce substantial amounts of a large 3.3 kb *Wx* pre-mRNA, with intron I still present, in addition to the mature *Wx* mRNA, and intermediate levels of *Wx* protein. Glutinous rice (group III cultivars) contains no amylose, no *Wx* protein, and no mature *Wx* mRNA; only the incompletely spliced 3.3 kb *Wx* pre-mRNA is present in group III cultivars. Based on these results, it is hypothesized that the amylose content of rice endosperm is regulated at the level of *Wx* transcript processing, and, more specifically, at the stage of intron I excision from the *Wx* pre-mRNA.

Introduction

Starch is the major energy reserve in many of the world's most important foods such as the seeds of cereal grains, the potato tuber, and the root of yam. The storage organs of these plants contain two types of starch: amylose, composed of largely linear chains of $\alpha(1\text{--}4)$ -linked glucose residues, and amylopectin, a branched starch with $\alpha(1\text{--}6)$ linkages at the branch points. The starch reserves of plants

usually consist of about 20–30% amylose and 70–80% amylopectin (Preiss, 1990).

Starch biosynthesis takes place within the chloroplasts and amyloplasts of plants. Synthesis of the glucose chains in starch is catalyzed by the enzyme starch synthase (ADP-glucose glucosyl transferase; EC 2.4.1.2.1), which occurs in two major forms, one soluble and the other bound to starch granules. Maize and pea both contain two forms of granule-bound starch synthase (GBSS), a 59 kDa protein (GBSS I) and a 77 kDa protein (GBSS II) (Dry *et al.*, 1992; MacDonald and Preiss, 1983, 1985; Smith, 1990). However, these two GBSS isoforms are synthesized in different tissues and probably play distinct roles in starch biosynthesis during plant development. The *waxy* (*Wx*) gene of plants encodes the 59 kDa GBSS, which is more commonly called the '*Wx* protein' (Ainsworth *et al.*, 1983; Echt and Schwartz, 1981; Hovenkamp-Hermelink *et al.*, 1987; Hsieh, 1988; Konishi *et al.*, 1985; Okuno and Sakaguchi, 1982; Sano, 1984; Tsai, 1974).

Maize *Wx* mutants were first shown to lack amylose in kernels (Sprague *et al.*, 1943) and GBSS I in endosperm (Nelson and Rhines, 1962). Similar results were subsequently obtained with *Wx* mutants of rice, barley, wheat, potato, sorghum, and Amaranths (Ainsworth *et al.*, 1983; Hovenkamp-Hermelink *et al.*, 1987; Hsieh, 1988; Konishi *et al.*, 1985; Okuno and Sakaguchi, 1982; Rohde *et al.*, 1988; Sano, 1984). In potato, the expression of a *Wx* antisense RNA chimeric gene was shown to inhibit GBSS I synthesis, providing direct evidence that GBSS I is the *Wx* gene product (Visser *et al.*, 1991). We previously cloned and sequenced *Wx* genes from three rice cultivars: *Oryza sativa* subspecies *japonica* Hanfeng cultivar (Wang *et al.*, 1990), *O. sativa* subspecies *indica* 232 cultivar, and *O. sativa* L. *f. spontanea* Dongxiang (wild rice) (Wang *et al.*, 1994). The coding region of the *Wx* gene from the Hanfeng cultivar consists of 13 exons interrupted by 12 introns (Wang *et al.*, 1990). Subsequently, a large (1124 bp) intron was identified in the leader region of the rice *Wx* gene (Figure 1; Wang and Hong, unpublished results). A comparison of the nucleotide sequences of the *Wx* genes of domesticated rice cultivars (Hanfeng and 232) and wild rice (Dongxiang) revealed the presence of two types of transposon-like elements, designated RTL-1 and RTL-2, located in introns 1, 10, and 13 of the rice *Wx* gene (Wang *et al.*, 1994).

As in other cereals, the amylose content of rice endosperm is thought to play important roles in grain yield, palatability, and processing quality. Therefore, an understanding of the genetic control of starch composition in rice seeds will be important in designing breeding and/or

Received 4 July 1994; revised 23 November 1994; accepted 23 December 1994.

*For correspondence (fax + 86 21 4042385).

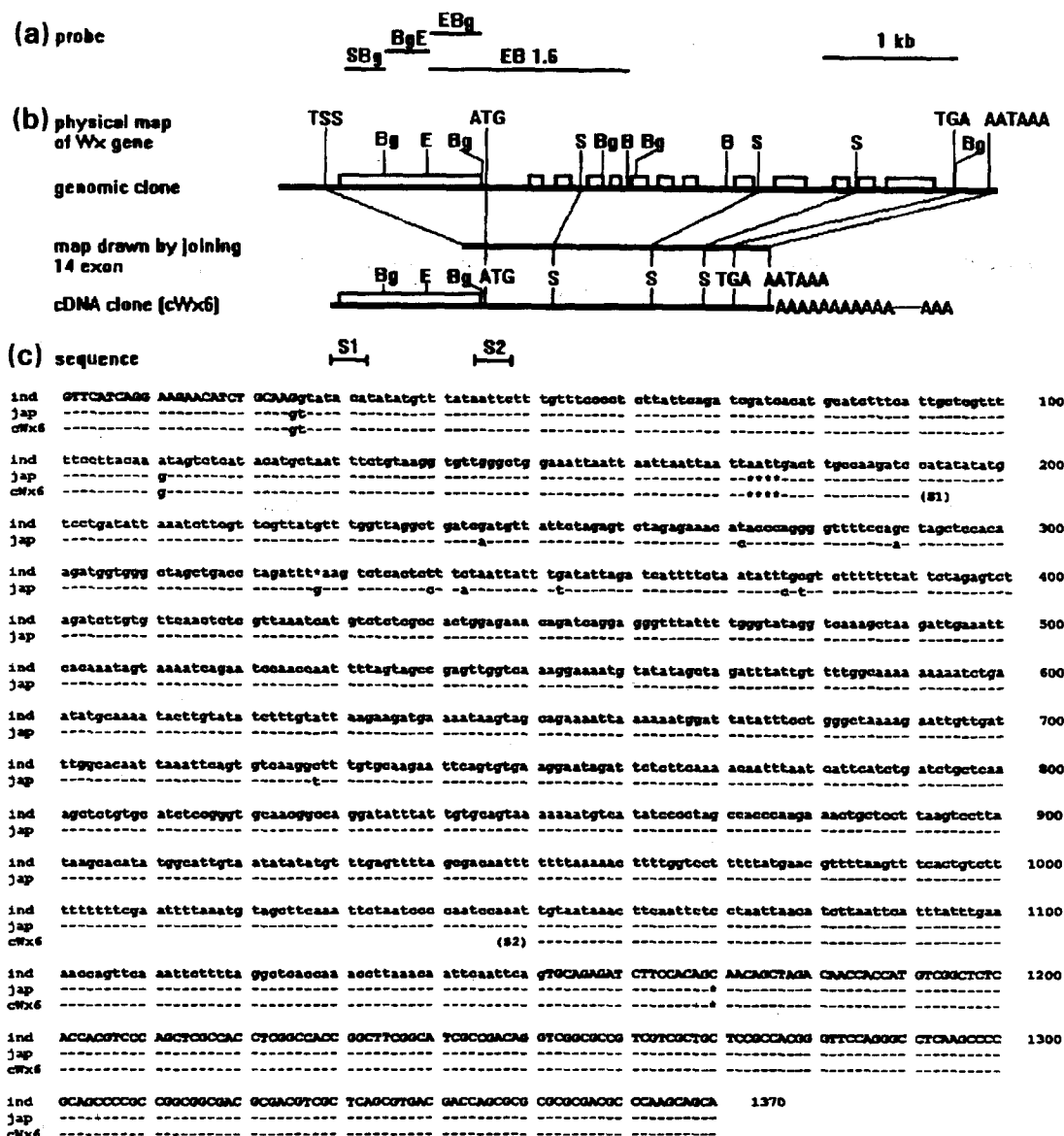


Figure 1. Comparison of the physical maps of the rice *Wx* gene and *Wx* cDNA cWx6.

(a) Diagram showing the size and locations of the *Wx* hybridization probes used as the experiments described in this paper.

(b) Comparisons of the restriction maps of cDNA cWx6 and the 14 exons of the *Wx* gene. Introns are represented by boxes; exons are shown as lines.

Abbreviations for specific sequences are as follows: transcription start site, TSS; translation start codon, ATG; transcription termination codon, TGA; polyadenylation site, AATAAA; restriction sites, S, *Sal*I; E, *Eco*RI; Bg, *Bgl*I; B, *Bam*HI.

(c) Comparisons of the nucleotide sequences of intron 1 and adjacent 5' and 3' regions of the *Wx* gene of the *indica* 232 cultivar (group I), the *japonica* Hanfeng cultivar (group II) and intron 1 junction regions of cDNA cWx6 from the Hanfeng cultivar. Genomic sequences of rice *Wx* gene from both 232 cultivar and Hanfeng cultivar were previously analyzed (Wang *et al.*, 1990, 1994). Sequences of the 232 cultivar including the entire intron 1 and the coding region are available from EMBL accession no. X64108. Sequences of the Hanfeng cultivar that contain the complete coding region and a 3' portion of intron 1 can also be found in EMBL accession no. X53694. Sequences extending 5' portion of intron 1 of the Hanfeng cultivar are shown in this study.

Exon sequences are shown in capital letters. Intron sequences are presented using lower case letters. The complete sequences are given only for the 232 cultivar (ind). For the Hanfeng cultivar (jap) and two fragments (S1 and S2) of the cDNA cWx6, only those nucleotides that differ from the corresponding nucleotides are shown. The dashes (-) indicate identity of nucleotides shown on the top line for the indica 232 cultivar. The asterisk (*) denotes those nucleotides absent in one of the cultivars, but present in the other.

genetic engineering strategies aimed at enhancing both yield and the nutritional quality of this important food source.

In this report, we show that the amylose contents of 31

different rice cultivars are related to their ability to process properly the *Wx* gene transcript. Rice cultivars with high amylose content produce large amounts of mature *Wx* mRNA and *Wx* protein. Glutinous rice produces no mature

Wx mRNA, *Wx* protein, or amylose; instead, a large *Wx* transcript containing intron 1 accumulates in glutinous rice. Cultivars with intermediate levels of amylose and *Wx* protein produce both the mature *Wx* mRNA and the incompletely spliced *Wx* pre-mRNA with intron 1 still present. These results indicate that the efficiency of excision of intron 1 from the *Wx* transcript is a major factor in the control of the amylose content of rice endosperm. Studies to identify the elements that control the efficiency of intron 1 excision from the rice *Wx* gene transcript are in progress.

Results

Three distinct patterns of *Wx* transcripts in different cultivars

Given the role of the *Wx* gene in controlling amylose synthesis in rice endosperm (Okagaki and Wessler, 1988) and the variation in amylose content in rice cultivars (Sano *et al.*, 1986), we thought that the *Wx* gene might be differentially expressed in various rice cultivars. To investigate this possibility, we used RNA gel blots to compare the relative levels of steady-state *Wx* transcripts in seeds of cultivars with different amylose content. *O. sativa* subsp. *indica* 232 cultivar (amylose content 20%), *O. sativa* subsp. *japonica* Hanfeng cultivar (amylose content 16%) and *O. sativa* subsp. *japonica* Fengnuo cultivar (amylose content 0%; glutinous rice) were used in the initial experiments.

Total RNAs were isolated from developing seeds harvested at 12 days after pollination (dap). RNA gel blots were prepared, and the RNAs on the blots were hybridized with a ³²P-labeled restriction fragment that contains both 5' non-coding and coding sequences of the rice *Wx* gene (probe EB 1.6; see Figure 1a). (Figure 1 presents a physical map of the rice *Wx* gene showing the sizes and locations of all restriction fragments that were used as hybridization probes in this study.) The RNA gel blots from the three cultivars exhibit distinct patterns of probe hybridization (Figure 2). The RNA sample from the 232 cultivar contained a single RNA species that hybridized to the *Wx* probe. This RNA was approximately 2.3 kb in size, the predicted size of mature *Wx* mRNA. In the case of the Hanfeng cultivar, the mature *Wx* mRNA and a larger RNA species (3.3 kb in length) were detected. In the Fengnuo cultivar (no amylose), only the 3.3 kb RNA species was present. In addition to the two major hybridizing RNAs described above, a minor hybridizing RNA was present in all three cultivars. The significance of this minor hybridizing RNA is unknown.

Changes in *Wx* transcripts during seed development

To investigate whether the pattern of *Wx* transcript accumulation might change during seed development, total RNAs

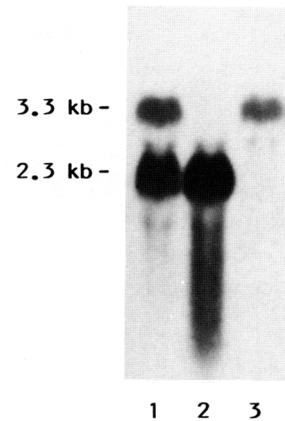


Figure 2. RNA gel blot comparison of the *Wx* gene transcripts in three rice cultivars.

Comparison of the patterns of *Wx* RNAs from three rice cultivars. Lane 1, Hanfeng cultivar; lane 2, 232 cultivar; lane 3, Fengnuo cultivar. Total RNAs (20 µg per sample) were separated in agarose gels containing formaldehyde, and RNAs on the membrane blots were hybridized to ³²P-labeled probe EB 1.6 (see Figure 1a).

were extracted from developing seeds of rice cultivar IR36 (non-glutinous indica rice; amylose content 24.5%), Hanfeng and Fengnuo and were analyzed for *Wx* transcripts by RNA gel blot experiments carried out with restriction fragment EB 1.6 (see Figure 1a) as the hybridization probe. The results of these experiments are shown in Figure 3. Although the levels of *Wx* transcripts were found to vary during seed development in a cultivar-specific manner, the pattern of transcripts present in a given cultivar did not change during seed development in any of the cultivars examined. The *Wx* transcripts in the IR36 and Hanfeng cultivars were present at the highest levels in developing seeds harvested at 7 and 17 dap, respectively. On the other hand, the 3.3 kb *Wx* transcript was present at about the same level from 8 and 23 dap in the Fengnuo cultivar. The levels of *Wx* transcripts began to decrease at the onset about 20 and 26 dap of seed maturation and desiccation in IR36 and Hanfeng cultivars.

Characterization of the 3.3 kb *Wx* RNA

Hybridization with intron 1 sequences. The 3.3 kb *Wx* RNA species present in group II and III rice cultivars is about equal in size to the mature *Wx* mRNA (2.3 kb) plus intron 1 (1.1 kb). To test for the presence of the intron 1 sequences in the 3.3 kb RNA, RNA gel blots of total RNA from developing seeds of the Hanfeng cultivar were probed with *Wx* DNA restriction fragments SBg, BgE and EBg (see Figure 1a) derived from the left, internal, and right portions of intron 1, respectively. These three intron 1 probes hybridized with the large 3.3 kb RNA species, but not with the mature *Wx* mRNA (Figure 4). These results demonstrate that the 3.3 kb *Wx* RNA species contains major portions of

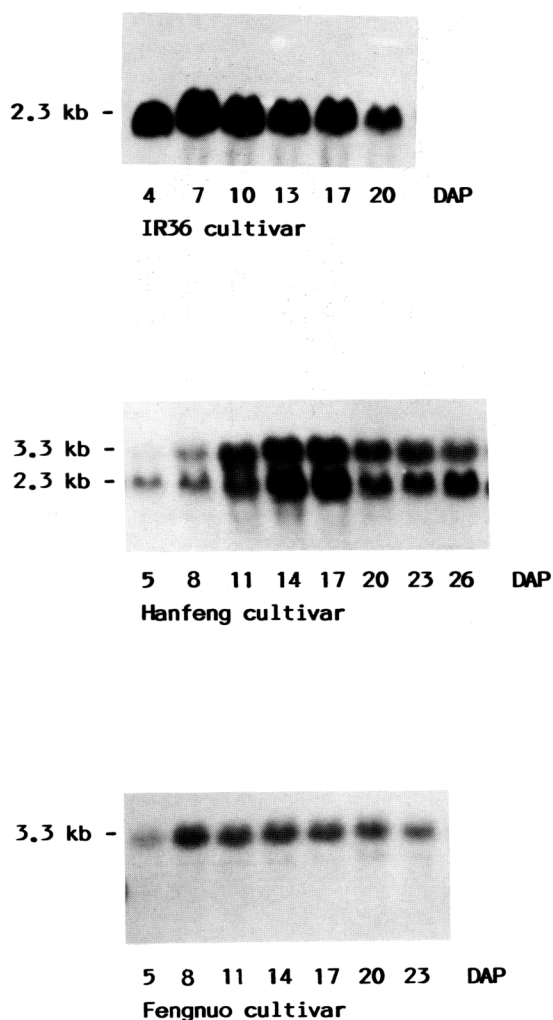


Figure 3. *Wx* RNAs present in rice seeds of three cultivars at different stages of development.

Total RNAs were isolated from the IR36 cultivar, the Hanfeng cultivar, and the Fengnuo cultivar. The RNAs on the membrane blots were hybridized to ^{32}P -labeled probe EB 1.6 (see Figure 1a).

intron 1 and strongly suggest that the 3.3 kb RNA represents a *Wx* transcript from which intron 1 has not been excised.

Poly(A)⁺ tail. To investigate whether the 3.3 kb RNA is an intact transcript with a poly(A)⁺ 3' terminus, total RNAs isolated from developing seeds of the Hanfeng cultivar were adsorbed on an oligo(dT)-cellulose column, and the poly(A)⁺ RNA fractions were eluted and re-examined by RNA gel blot analysis using either *Wx* DNA restriction fragments EB 1.6 (*Wx* intron 1 plus 5' coding fragment) or EBg (intron 1 only) as probes (see Figure 1a). The results shown in Figure 5 demonstrate that total RNAs and poly(A)⁺ RNAs yield the same hybridization patterns with these two probes. Both probes hybridized with a 3.3 kb RNA species present in the total RNA and the poly(A)⁺ RNA samples. These results demonstrate that both the

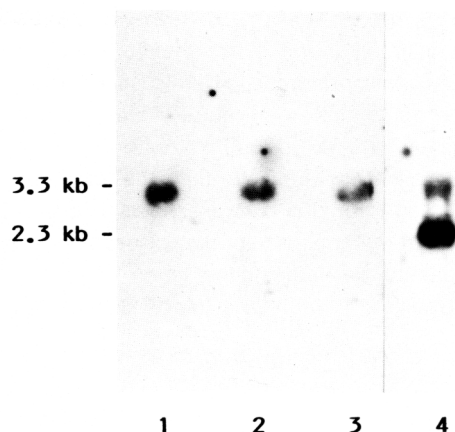


Figure 4. Identification of intron I sequences in the 3.3 kb *Wx* RNA species. Total RNAs were isolated from the Hanfeng cultivar at 12 dap. The probes used were SBg, derived from the left end of intron 1 (lane 1), BgE, derived from the middle segment of intron 1 (lane 2), EBg, derived from the right end of intron 1 (lane 3), and EB 1.6 (lane 4) (see Figure 1a). All RNA samples contained 20 μg of total RNA.

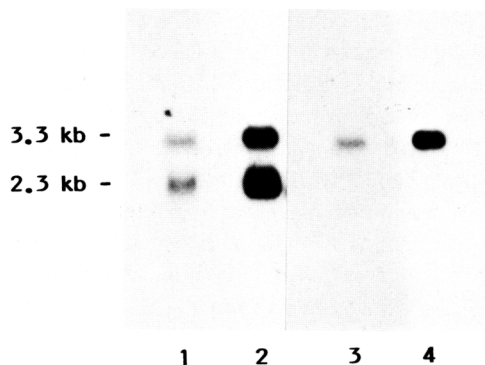


Figure 5. RNA gel blot analysis of *Wx* sequences in total RNA and Poly(A)⁺ RNA.

All RNAs were isolated from seeds of the Hanfeng cultivar harvested at 12 dap. Total RNAs (20 μg samples, lanes 1 and 3) and poly(A)⁺ RNAs (8 μg samples, lanes 2 and 4) were separated on agarose gels containing formaldehyde, and RNAs on the membrane blots were hybridized to ^{32}P -labeled probe EB 1.6 (lanes 1 and 2) or EBg (lanes 3 and 4). See Figure 1(a) for probe locations.

large 3.3 kb *Wx* RNA species and the mature *Wx* mRNA are polyadenylated.

Restriction map of *Wx* cDNA clone cWx6. To investigate the structure of the 3.3 kb *Wx* RNA further, we isolated and characterized a cDNA clone, designated cWx6, equivalent to the 3.3 kb *Wx* transcript. The cWx6 cDNA clone was isolated from a library constructed by using poly(A)⁺ RNAs of developing seeds of the Hanfeng cultivar. The library was screened for cDNAs equivalent to the 3.3 kb *Wx* RNA by using a restriction fragment from within intron 1 (fragment EBg; see Figure 1a) as the hybridization probe. The insert in clone cWx6 is 3.3 kb in size. A restriction map of the insert in cWx6 was constructed and compared with the composite restriction map of all 14 exons of rice *Wx*

gene (Wang *et al.*, 1990). The results of this comparison are shown in Figure 1(b); the restriction maps of the cWx6 cDNA and the 14 exons of the Wx gene are identical except for the extra 1.1 kb segment in the leader region of cWx6 clone.

Partial sequence of Wx cDNA clone cWx6. To examine the structure of the extra 1.1 kb of DNA in Wx cDNA clone cWx6 in greater detail, the 5' and 3' termini of the cDNA insert in clone cWx6 were sequenced. In addition, the putative exon-intron junctions of intron 1 in cDNA cWx6 were sequenced and compared with the corresponding sequences of Wx genomic clones isolated from libraries constructed from DNAs of the indica 232 (group I) and japonica Hanfeng (group II) cultivars (Wang *et al.*, 1990, 1994). The results (Figure 1c) verify that the extra 1.1 kb segment present at the 5' terminus of the insert of clone cWx6 contains the 5' and 3' terminal sequences of intron 1 of rice Wx gene. In addition, the sequence of the 3' non-coding region of the cWx6 cDNA was found to be identical to the sequence downstream from the translation termination codon in the Hanfeng Wx genomic clone (data not shown).

Distribution of Wx transcripts in 31 rice cultivars

To determine whether the 3.3 kb Wx RNA is also present in other rice cultivars, we used RNA gel blot hybridizations to examine the Wx transcripts of 31 rice cultivars, including non-glutinous indica and japonica rice and glutinous japonica rice. Wx restriction fragment EB 1.6 (see Figure 1a) was used as the hybridization probe. Based on the results shown in Figure 6, the 31 cultivars can be divided into three groups based on observed patterns of hybridization to the Wx probe. Cultivars of group I (indica rice) contained only the mature 2.3 kb Wx mRNA, just like the 232 and IR36 cultivars (see Figures 2 and 3, respectively). Cultivars of group II (indica and japonica rice) yield results similar to that observed with the Hanfeng cultivar (see Figure 2); group II cultivars produced both the 3.3 kb Wx RNA and mature 2.3 kb Wx mRNA. All of the glutinous cultivars tested were in group III; only the 3.3 kb pre-mRNA was detected in these cultivars. Also, the results clearly showed that the steady-state level of mature Wx mRNA was higher in developing seeds of group I cultivars than in developing seeds of group II cultivars.

Wx protein level in mature seeds of the three groups of rice cultivars

To determine the levels of Wx protein in various rice cultivars, Wx protein was extracted from the mature seeds of 31 different cultivars and analyzed by SDS-polyacrylamide gel electrophoresis. Relative concentrations of 59–

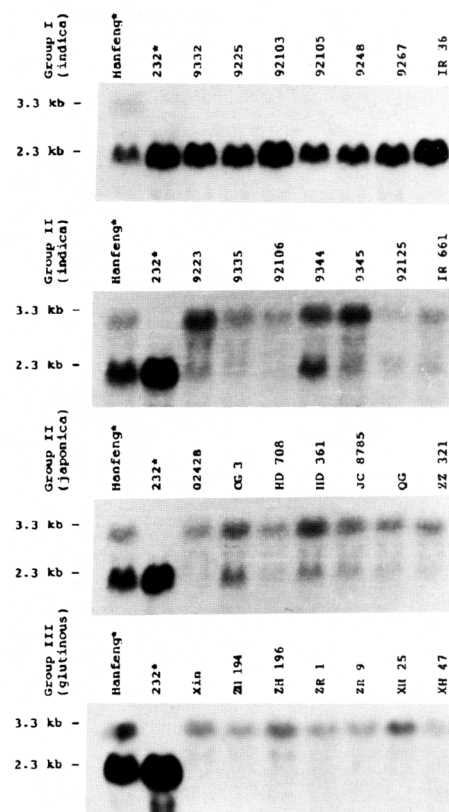


Figure 6. RNA gel blot survey of transcript patterns in 30 rice cultivars. All RNA samples contained 20 µg of total RNA isolated from developing seeds at 12 dap. RNAs on the blots were hybridized to probe EB 1.6 (see Figure 1a). The asterisk (*) indicates that the Hanfeng and 232 cultivars were used as a control on all blots.

60 kDa Wx protein were determined by densitometry (490 nm) after staining with Coomassie brilliant blue. The results showed that the seeds of group I cultivars contained the highest levels of Wx protein, the seeds of group II cultivars contained lower levels of Wx protein, and no Wx protein was detected in the seeds of group III cultivars (Table 1). These results also showed that the level of Wx protein is positively correlated with the level of Wx mRNA and is negatively correlated with the level of the large 3.3 kb Wx pre-mRNA in all 31 cultivars examined in this study.

Amylose content in mature seeds of the three groups of rice cultivars

The amylose content in mature seeds of all 31 rice cultivars was determined by the method of Juliano (1971). As the data in Table 1 show, large variations in amylose content occur among the various cultivars. Starch in the seeds of group I cultivars contains over 20% amylose, and that of group II cultivars contains 6–16% amylose, whereas no amylose could be detected in the seeds of group III cultivars.

Table 1. Levels of Wx RNAs, Wx protein and amylose content in 31 rice cultivars

Wx gene expression product	Group I (indica rice)							
	Cultivar							
	232	9332	9225	92103	92105	9248	9267	IR36
pre-RNA ^a	0	0	0	0	0	0	0	0
mRNA ^a	476	458	352	637	287	310	410	605
Wx protein ^b	430	816	697	832	832	832	832	816
AC (%) ^c	20.0	27.8	24.8	27.3	27.7	26.7	26.0	24.5
	Group II (indica rice)							
	Cultivar							
	9223	9335	92106	9344	9345	92125	IR661	
pre-RNA ^a	46	14	18	44	55	13	8	
mRNA ^a	31	41	16	106	40	27	16	
Wx protein ^b	33	179	33	35	60	23	69	
AC (%) ^c	6.7	12.3	6.8	10.2	10.0	9.2	11.2	
	Group II (japonica rice)							
	Cultivar							
	Hanfeng	02428	CG 3	HD708	HD361	JC8785	QG	ZZ321
pre-RNA ^a	38	32	43	16	33	19	23	48
mRNA ^a	120	18	116	24	63	38	21	42
Wx protein ^b	66	28	24	62	51	71	11	14
AC (%) ^c	16.0	11.1	14.3	12.5	15.3	9.3	9.2	— ^d
	Group III (glutinous rice)							
	Cultivar							
	Fengnuo	Xin	ZH194	ZH196	ZR 1	ZR 9	XH25	XH47
pre-RNA ^a	18	16	16	35	22	24	35	10
mRNA ^a	0	0	0	0	0	0	0	0
Wx protein ^b	0	0	0	0	0	0	0	0
AC (%) ^c	0	0	0	0	0	0	0	0

^a Relative radioactivity (relative amount of total RNAs were loaded in each lane for each cultivar, and probing of these blots with a ³²P-labeled p(dT)₁₅ (Boehringer, Cat. No. 814270) probe confirmed that all lanes contained similar levels of poly(A)⁺ RNA (data not shown)).

^b Relative amount.

^c Amylose content.

^d Not detected.

Discussion

The amylose content of rice seeds is an important agronomic trait, affecting both yield and nutritional quality of one of the world's most important foods. Thus, an understanding of the genetic control of amylose content in rice is a major goal of basic agricultural research. Information about the molecular mechanisms that control amylose content in rice endosperm will be important in designing breeding and genetic engineering strategies for improving both yield and nutritional quality in this prominent food crop. Therefore, research in several laboratories has focused on elucidating the genetic control of Wx gene expression and its relationship to amylose content in rice endosperm.

An important objective of our research is to identify the regulatory elements that control Wx gene expression and amylose content in rice seeds. Achieving this objective will be one step toward the goal of understanding the molecular mechanisms that control the quantities of gene products produced under a given set of environmental conditions. Such quantitative regulation of gene expression can occur either at the transcriptional level or at post-transcriptional levels (Brunelle and Chua, 1993; Gallie, 1993). The results reported here indicate that post-transcriptional events play an important role in regulating the Wx protein and amylose content of 31 different rice cultivars.

Our results support and extend the results of Sano (1984) and Vilareal and Juliano (1989). Both of these investigations showed that, in addition to wx mutant alleles that result

in little or no Wx protein or amylose, the various rice cultivars contain two types of wild-type Wx genes, referred to as Wx^a and Wx^b alleles, that are correlated with high and low Wx protein and amylose content, respectively. Vilareal and Juliano (1989) reported that Wx genes of type Wx^a and type Wx^b were well differentiated, with distinct regression equations relating the level of Wx protein and amylose content in various non-waxy cultivars. Our results also are consistent with those of Bollich and Webb (1973), who concluded that the amylose content of rice endosperm is controlled by one major gene and several modifier genes. This conclusion is supported by the results of Kumar and Khush (1988), who analyzed the amylose content of seeds of F₁, F₂, and backcross progeny produced from reciprocal crosses of low-amylose cultivars. Collectively, the results of Sano (1984), Vilareal and Juliano (1989), and the present study leave little doubt that the major gene controlling amylose content in rice is the Wx gene.

The results of the present study demonstrate the presence of three distinct patterns of Wx gene transcript accumulation in the 31 rice cultivars examined. Within a given cultivar, the Wx transcript pattern remains the same from the time of pollination through seed maturation.

Seed starch of the eight group I cultivars that we investigated contained from 20 to 28% amylose. All of these high amylose cultivars contained high levels of both Wx protein and mature 2.3 kb Wx mRNA in developing seeds. The 15 group II cultivars examined in this study exhibited more variability in seed starch composition, with amylose content ranging from about 6 to 16%. These intermediate-level amylose cultivars contained about 10-fold lower levels of Wx protein and mature Wx mRNA than the group I cultivars. In addition, they contained a large 3.3 kb Wx RNA. Seeds of the eight group III (glutinous rice) cultivars that we investigated contained no amylose, no Wx protein, and no mature 2.3 kb Wx mRNA; only the large 3.3 kb Wx RNA species was present in group III cultivars.

The size of the 3.3 kb Wx RNA equals the combined size of the 14 exons plus intron 1 of the rice Wx gene (Figure 1; Wang *et al.*, 1990, 1994). Restriction fragments from the 5' end, middle, and 3' end of intron 1 all hybridized with the 3.3 kb Wx RNA species, but not with mature 2.3 kb Wx mRNA. Moreover, a cDNA clone, cWx6, equivalent to the 3.3 kb Wx RNA of the group II Hanfeng cultivar, was isolated and partially characterized. The restriction map of the cWx6 cDNA corresponds precisely to the restriction map constructed by combining the maps of intron 1 and the 14 exons of the rice Wx gene. Finally, the exon 1–intron 1 and intron 1–exon 2 junctions of cDNA clone cWx6 were sequenced and shown to be identical to the corresponding regions of a Wx genomic clone prepared using DNA from the Hanfeng cultivar (Figure 1c; Wang *et al.*, 1990). These results provide strong evidence for the presence of the

intron 1 sequence in the 3.3 kb Wx RNA species of group II and group III rice cultivars.

Okagaki (1992) reported that Nato rice contained two Wx transcripts of sizes 2.4 and 4 kb. The 2.4 kb RNA presumably is the mature Wx mRNA, and the 4 kb RNA species probably corresponds to the 3.3 kb Wx transcript characterized in our study. Okagaki proposed that the 4 kb RNA might have resulted from the failure to excise an intron from the Wx gene transcript. Our results support his proposal and demonstrate that the unexcised intron is the Wx leader sequence intron (intron 1 of the rice Wx gene).

The intron 1 splice sites shown in Figure 1(c) were identified by reverse transcript polymerase chain reactions (RT-PCR) using RNAs isolated from indica rice cultivars 232 and IR36 as templates. The splice sites were verified by comparing the resulting reverse transcript sequences with the corresponding Wx genomic DNA sequences (Figure 1c and data not shown). In the Hanfeng cultivar, intron 1 is 1124 bp in length. In all three cultivars analyzed to date, intron 1 conforms to the GT-AG rule of Breathnach and Chambon (1981). Since intron 1 lies entirely within the 5' non-coding leader sequence of the rice Wx gene, its involvement in the regulation of Wx gene expression should not be surprising.

In the 31 cultivars investigated, the level of Wx protein and amylose content both were positively correlated with the level of the 2.3 kb Wx mRNA and were negatively correlated with the level of the 3.3 kb Wx transcript. This suggests that the lower levels of Wx protein and amylose in seeds of group II cultivars result from lower levels of mature Wx mRNA, which, in turn, result from inefficient excision of intron 1 from the Wx gene transcript. Although the 3.3 kb Wx RNA is polyadenylated, it clearly is not translated because no Wx protein is synthesized in group III cultivars.

The presence of the 3.3 kb Wx transcript in the seeds of group III (glutinous rice) cultivars demonstrates that the Wx gene is transcribed in endosperm cells of these plants. The presence of the 3.3 kb Wx transcript, but no mature 2.3 kb Wx mRNA or Wx protein, in the seeds of group III cultivars demonstrates that Wx gene expression is controlled, at least in part, post-transcriptionally. Moreover, our results suggest that the excision of intron 1 from the 5' leader region of the Wx transcript plays an important role in regulating Wx gene expression. Nevertheless, we can not exclude the possibility that the observed patterns of Wx gene transcripts result from differential rates of degradation of the 2.3 and 3.3 kb Wx RNA species in the various cultivars. No information is available on the stability of these Wx RNAs. However, given the similarity between the nucleotide sequences of the Wx genes of the group I and group II cultivars (Wang *et al.*, 1994), the total absence of the 3.3 kb Wx RNA containing intron 1 in the group I cultivars, and the total absence of the 2.3 kb Wx mRNA in

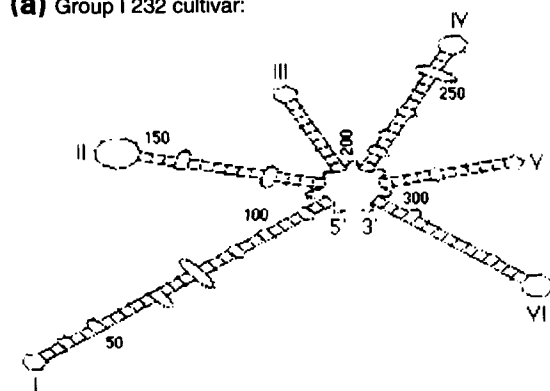
the group III cultivars, differential turnover probably is not responsible for the three distinct patterns of *Wx* transcripts observed in the 31 cultivars investigated in this study.

Unfortunately, we still do not know why intron 1 fails to be excised from the *Wx* gene transcripts of group III cultivars and is less efficiently excised from the *Wx* transcripts of group II cultivars than from the *Wx* transcripts of group I cultivars. However, the results of Sano (1984) showed that the difference in *Wx* protein level in rice strains C8005 (high level of *Wx* protein) and T65 (intermediate level of *Wx* protein) was controlled by *cis*-acting elements. F_2 seeds produced by reciprocal crosses of these two strains showed a direct relationship between allele dosage and endosperm *Wx* protein level, segregating 1:1:1:1, $Wx^a Wx^a Wx^a: Wx^a Wx^a Wx^b: Wx^a Wx^b Wx^b: Wx^b Wx^b Wx^b$. If genes encoding *trans*-acting regulatory factors were responsible for the differences in *Wx* protein level and amylose content in these rice strains, dominance should have been detected in analyzing the progeny of these crosses. The observed lack of dominance indicates that the segregating factors controlling *Wx* protein level in the rice strains studied by Sano (1984) were *cis*-acting elements. Our results suggest that these *cis*-acting elements are involved in the excision of intron 1 from the leader region of the *Wx* transcript. The *cis*-acting elements that affect the efficiency of excision of intron 1 from the rice *Wx* transcript remain unknown, but experiments to identify them are in progress.

During the past decade, research on yeast and Hela cells has provided considerable information about the molecular details of RNA splicing reactions (reviewed by Green, 1986). The introns of nuclear genes are excised within the nucleus on complex spliceosomes, which are composed of small nuclear RNA-protein complexes (snRNPs). However, numerous soluble factors also are involved in RNA splicing reactions (Lamm and Lamond, 1993). Although consensus splice sequences and branch point sequences can be defined for the introns of nuclear genes, only the 5' GT, the 3' AG, and the branch point A are absolutely required for intron excision (Hanley and Schuler, 1988).

The intron 1 nucleotide sequences of the *Wx* genes of group I cultivar 232 and group II cultivar Hanfeng differ by 10 single bp substitutions (seven of which occur within the transposon-like element near the middle of the intron), one single bp addition/deletion, and one addition/deletion of a 4 bp tandem repeat present six times in the group I 232 cultivar and in five copies in the group II Hanfeng cultivar. In addition, there is a single G addition/deletion at the exon 1-intron 1 junction, which results in a G at the 3' terminus of exon I in group I cultivar 232 and an A at this position in group II cultivar Hanfeng. This last difference might affect the efficiency of splicing since, at least in animals, the consensus 5' splice site contains a G at this position,

(a) Group I 232 cultivar:



(b) Group II Hanfeng cultivar:

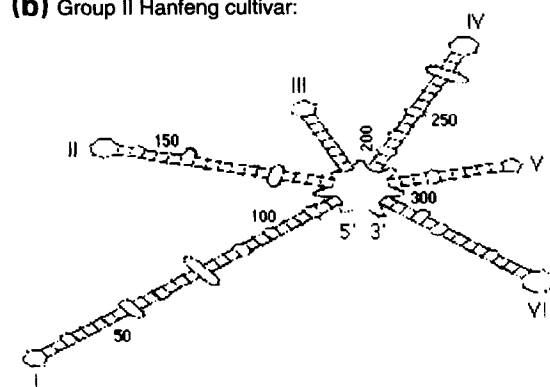


Figure 7. Computer prediction of RNA secondary structures in the 5' terminal region of *Wx* primary transcripts from the indica 232 cultivar (group I) (a) and the japonica Hanfeng cultivar (group II) (b). The squiggles are drawn by using default MFOLD and PLOTfold programs of the Genetics Computer Group (GCG). The arabic numbers indicate nucleotide positions.

which base-pairs with a C in the complementary sequence of the U1 snRNA of the spliceosome.

In an attempt to determine which of the other differences between the sequences of intron 1 of group I cultivar 232 and group II cultivar Hanfeng might be important, we used the Genetics Computer Group (GCG) squiggles program to examine possible secondary structures of the 5' terminus of the rice *Wx* transcript. As shown in Figure 7, the 5' terminus of the *Wx* primary transcript is rich in potential secondary structure. The 350 bp at the 5' end of the *Wx* transcript has the potential to form six cruciforms. However, since this region is very AT-rich, none of the potential cruciforms is particularly stable; the free energy of the structures shown in Figure 7 is only about $-50 \text{ kcal mol}^{-1}$ (-46 and $-56 \text{ kcal mol}^{-1}$ for the *Wx* transcripts of the Hanfeng and 232 cultivars, respectively). If these cruciforms are numbered I–VI starting from the 5' terminus, cruciforms II and III differ between the *Wx* transcripts of the group I and group II rice cultivars (see Figure 7). Cruciform III is formed by the 4 bp tandem repeats of *Wx* intron 1, and the longer stem formed by the extra copy of this repeat in

the 232 cultivar could be important in the efficient excision of this intron. We should emphasize that, at present, there is no information as to which of the observed sequence differences, if any, might affect the efficiency of excision of intron 1 from the rice *Wx* gene transcript. The important *cis*-acting regulatory sequences might be located either upstream or downstream from the *Wx* structural gene. Experiments designed to localize the *cis*-acting elements that regulate intron 1 excision are in progress.

Experimental procedures

Plant material

The rice cultivars used in this study were obtained from the China National Rice Research Institute and Jiangsu Agricultural College.

RNA extractions

Total RNA was isolated from immature rice seeds by the cold-phenol method. Plant material was homogenized in 50% phenol, 25 mM Tris (pH 9.0), 2 mM EDTA, 0.5% SDS at 4°C for 2 min and then centrifuged at 10 000 *g* for 10 min at 4°C. The aqueous layer was re-extracted with 1:1 phenol-chloroform (v/v). The RNAs were precipitated with 0.1 vol. of 3 M sodium acetate and 2.5 vol. of ethanol; the RNAs in the pellet were then dissolved in water. Lithium chloride (8 M) was added to the RNA solution to a final concentration of 2 M. After storage overnight at 4°C, the RNA solution was centrifuged at 4000 *g* at 4°C for 10 min. The pellet fraction was washed twice with 70% ethanol, dried at room temperature, and the precipitated RNAs were dissolved in H₂O at 4°C. The long incubation time was necessary to allow the RNAs to dissolve; yields were poor when shorter incubation times were used.

Construction of a rice endosperm cDNA library

Poly(A)⁺ RNAs were purified from total RNAs of immature seeds of the Hanfeng cultivar as described by Sambrook et al. (1989). Oligo(dT)-cellulose was purchased from BRL (Gaithersburg, MD). A cDNA library was constructed from these poly(A)⁺ RNAs by using the SuperScript Lambda cDNA Synthesis and Cloning System (BRL) according to the instructions of the manufacturer. This library was screened by *in situ* plaque hybridization (Sambrook et al., 1989) using ³²P-labeled restriction fragment EBg (see Figure 1a) of the rice *Wx* gene as probe. The probe was labeled using the random-primer method of Feinberg and Vogelstein (1983). [α -³²P]dATP was purchased from Amersham (Arlington Heights, IL).

RNA gel blot hybridization analysis

RNA gel blot hybridizations were carried out as described by Sambrook et al. (1989). Selected ³²P-labeled DNA restriction fragments of the *Wx* gene (see Figure 1a) were used as hybridization probes. Total RNAs (20 µg samples) were separated on 1% agarose gels containing formaldehyde and transferred to nylon membranes. The RNAs on the membranes were hybridized with labeled probe in a solution containing 50% formamide, 6×SSC, 1% SDS, 5% dextran sulfate, and 1% Denhardt's solution at 42°C. The RNA

gel blots were then washed in 0.1×SSC and 0.1% SDS at 65°C. The autoradiography was established by exposing the membrane for 24–48 h to X-ray film at –70°C with an intensifying screen.

The relative radioactivity of bands on Northern blots was measured by adding a few milliliters of toluene-based scintillation fluid to the dried membrane containing radioactivity cut from hybridized membrane and counting in the ³²P channel of a liquid scintillation counter.

DNA sequencing

The longest *Wx* cDNA clone isolated (cWx6) was subcloned into pBluescript KS, and the 5' and 3' ends of the cDNA were sequenced. DNA sequencing was done by the dideoxynucleotide chain termination method (Sanger et al., 1977) using Taq DNA polymerase in an automated fluorescent sequencer (ABI 373 model).

Electrophoresis of Wx protein

The rice *Wx* protein was extracted from mature seeds as described by Sano (1985). The protein samples were electrophoresed on 10% polyacrylamide gels containing SDS, and the proteins were stained with Coomassie brilliant blue.

Determination of amylose content

Amylose content in mature seeds was determined as described by Juliano (1971). Amylose and amylopectin were purchased from Sigma (St Louis, MO).

Acknowledgments

We thank Dr Shan-Jiong Shen for encouragement and Dr Zhen-E. Xu for critical reading of the manuscript. This work was supported by the Rockefeller Foundation, the High Technology Development Program of China, and the Plant Molecular Genetics Institute of the University of Minnesota.

References

- Ainsworth, C.C., Gale, M.D. and Baird, S. (1983) The genetics of β -amylase isozymes in wheat. I. Allelic variation among hexaploid varieties and intrachromosomal gene locations. *Theor. Appl. Genet.* **66**, 39–49.
- Bollich, C.N. and Webb, B.D. (1973) Inheritance of amylose of two hybrid populations of rice. *Cereal Chem.* **50**, 631–636.
- Breathnach, R. and Chambon, P. (1981). Organization and expression of eucaryotic split gene coding for proteins. *Ann. Rev. Biochem.* **50**, 349–383.
- Brunelle, A.N. and Chua, N. (1993) Transcription regulatory protein in higher plants. *Curr. Biol.* **3**, 254–258.
- Dry, I., Smith, A., Edwards, A., Bhattacharyya, M., Dunn, P. and Martin, C. (1992) Characterization of cDNAs encoding two isoforms of granule-bound starch synthase which show differential expression in developing storage organs of pea and potato. *Plant J.* **2**, 193–202.
- Echt, S. and Schwartz, D. (1981) Evidence for the inclusion of controlling elements within the structural gene at the waxy locus in maize. *Genetics*, **99**, 275–284.
- Feinberg, A.P. and Vogelstein, B. (1983) A technique for

- radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6–13.
- Gallie, D.R.** (1993) Post-transcriptional regulation of gene expression in plants. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **44**, 77–105.
- Green, M.R.** (1986) Pre-mRNA splicing. *Ann. Rev. Genet.* **20**, 671–708.
- Hanley, B.A. and Schuler, M.A.** (1988) Plant intron sequences: evidence for distinct groups of introns. *Nucl. Acids Res.* **16**, 7159–7176.
- Hovenkamp-Hermelink, J.H.M., Jacobsen, E., Ponstein, A.S., Visser, R.G.F., Vos-Scheerkeuter, G.H., Bijmolt, E.W., de Vries, J.N., Witholt, B. and Feenstra, W.J.** (1987) Isolation of an amylose-free mutant of the potato (*Solanum tuberosum*). *Theor. Appl. Genet.* **75**, 217–221.
- Hsieh, J.S.** (1988) Genetic studies on the Wx gene of sorghum (*Sorghum bicolor* L. Moench). 1. Examination of the protein product of the waxy locus. *Bot. Bull. Acad. Sinica*, **29**, 293–299.
- Juliano, B.O.** (1971) A simplified assay for milled rice amylose. *Cereal Sci. Today*, **16**, 334–338.
- Konishi, Y., Nojima, H., Okuno, K., Asaoka, M. and Fuwa, H.** (1985) Characterization of starch granules from waxy, non-waxy and hybrid seeds of *Amaranthus hypochondriaces* L. *Agric. Biol. Chem.* **49**, 1965–1971.
- Kumar, I. and Khush, G.S.** (1988) Inheritance of amylose content in rice (*Oryza Sativa* L.). *Euphytica*, **38**, 261–269.
- Lamm, G.M. and Lamond, A.I.** (1993) Non-snRNP protein splicing factors. *Biochim. Biophys. Acta*, **1173**, 247–265.
- MacDonald, F.D. and Preiss, J.** (1983) Solubilization of the starch-granule-bound starch synthase of normal maize kernels. *Plant Physiol.* **73**, 175–178.
- MacDonald, F.D. and Preiss, J.** (1985) Partial purification and characterization of granule-bound starch synthase from normal and waxy maize. *Plant Physiol.* **78**, 849–852.
- Nelson, O.E. and Rhines, H.W.** (1962) The enzymatic deficiency in the waxy mutant of maize. *Biochem. Biophys. Res. Commun.* **9**, 297–300.
- Okagaki, R.J.** (1992) Nucleotide sequence of a long cDNA from the rice waxy gene. *Plant Mol. Biol.* **19**, 513–516.
- Okagaki, R.J. and Wessler, S.R.** (1988) Comparison of non-mutant and mutant waxy genes in rice and maize. *Genetics* **120**, 1137–1143.
- Okuno, K. and Sakaguchi, S.** (1982) Inheritance of starch characteristics in endosperm of *Amaranthus hypochondriaces*. *J. Hered.* **73**, 467.
- Preiss, J.** (1990) Biology and molecular biology of starch synthesis and its regulation. *Oxford Surv. Plant Mol. Cell Biol.* **7**, 59–114.
- Rohde, W., Becker, D. and Salamini, F.** (1988) Structural analysis of the waxy locus from *Hordeum vulgare*. *Nucl. Acids Res.* **16**, 7185.
- Sambrook, J., Fritsch, E.F. and Maniatis, T.** (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Sanger, F., Nicklen, S. and Coulson, A.R.** (1977) DNA sequencing with chain terminating inhibitors. *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
- Sano, Y.** (1984) Different regulation of waxy gene expression in rice endosperm. *Theor. Appl. Genet.* **68**, 4567–4573.
- Sano, Y.** (1985) Gene regulation at the waxy locus in rice. In *Gamma-field Symposia*, No. 24 (Institute of Radiation Breeding, eds), pp. 63–77.
- Sano, Y., Katsumata, M. and Okuno, K.** (1986) Genetic studies of speciation in cultivated rice. 5: Inter- and intra-specific differentiation in the waxy gene expression in rice. *Euphytica*, **35**, 1–9.
- Smith, A.M.** (1990) Evidence that the waxy protein of pea is not the major starch granule-bound starch synthase. *Planta*, **182**, 599–604.
- Sprague, G.F., Brimhall, B. and Hoxon, R.M.B.** (1943) Some effects of the waxy gene in corn on properties of endosperm starch. *J. Am. Soc. Agron.* **35**, 817–822.
- Tsai, C.Y.** (1974) The function of the waxy locus in starch synthesis in maize endosperm. *Biochem. Genet.* **11**, 83–96.
- Villareal, C.P. and Juliano, B.O.** (1989) Comparative levels of waxy gene product of endosperm starch granules of different rice ecotypes. *Starch*, **41**, 369–371.
- Visser, R.G.F., Somhorst, I., Kuipers, G.J., Ruys, J., Feenstra, W.J. and Jacobson, E.** (1991) Inhibition of the expression of the gene for granule-bound starch synthase in potato by antisense constructs. *Mol. Gen. Genet.* **225**, 289–296.
- Wang, Z., Wu, Z., Xing, Y., Zheng, F., Gao, J. and Hong, M.** (1990) Nucleotide sequence of rice waxy gene. *Nucl. Acids Res.* **18**, 5898.
- Wang, Z., Zheng, F., Wang, X., Guo, X., Zhang, W. and Hong, M.** (1994) Characterization of two transposon like element in rice Wx gene. *Sci. China Series B*, **37**, 437–447.