Theor Appl Genet (1997) 94:773–781

N. M. Ayres · A. M. McClung · P. D. Larkin
H. F. J. Bligh · C. A. Jones · W. D. Park

Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germ plasm

Received: 31 July 1996 / Accepted: 22 November 1996

Abstract The *Waxy* gene (*Wx*) encodes the granule-bound starch synthase responsible for the synthesis of amylose in rice (*Oryza sativa*). Recently, a polymorphic microsatellite sequence closely linked to the *Wx* gene was reported. To determine whether polymorphism in this sequence correlates with variation in apparent amylose content, we tested an extended pedigree of 92 current and historically important long-, medium- and short-grain US rice cultivars representing the efforts of many breeders over more than 80 years. Seven *Wx* microsatellite alleles were identified which together explained 82.9% of the variation in apparent amylose content of the 89 non-glutinous rice cultivars tested. Similar results were also obtained with 101 progeny of a cross between low- and intermediate-amylose breeding lines. An additional, unique microsatellite allele, $(CT)_{16}$, was detected in one glutinous cultivar, CI 5309. However, the other glutinous cultivars, Calmochi 101 and Tatsumi mochi, were in the $(CT)_{17}$ class along with three other cultivars that contained 15–16.5% amylose.

We sequenced a 200-bp PCR-amplified fragment containing the CT microsatellite and the putative 5′ splice site of the *Wx* leader intron from a subset of 42 cultivars representing all eight microsatellite alleles. All of the cultivars with 18% or less amylose had the sequence AGTTATA at the putative leader intron 5′ splice site, while all cultivars with a higher proportion of amylose had AGGTATA. This single nucleotide substitution could also be assayed by *Acc*I digestion of the amplified fragment. Overall, this single nucleotide polymorphism could explain 79.7% of the variation in the apparent amylose content of the 89 non-glutinous cultivars tested. Interestingly, cultivars in the $(CT)_{19}$ microsatellite classes that differed substantially in amylose content still showed the correlation between this G-T polymorphism and apparent amylose content. The G-T polymorphism at this site was not, however, able to explain the very low amylose contents of the three glutinous cultivars tested, all of which had the sequence AGTTATA.

Key words Apparent amylose · Microsatellite · *Waxy* · Rice · RNA splicing

Communicated by G. E. Hart

N. M. Ayres · P. D. Larkin · W. D. Park (✉)
Crop Biotechnology Center, Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA

A. M. McClung
USDA-ARS, Route 7, Box 999, Beaumont, TX 77713, USA

H. F. J. Bligh · C. A. Jones
Department of Biochemistry, Queens Medical Centre, University of Nottingham, Nottingham NG7 2UH, UK

Introduction

Amylose content is a key determinant of the cooking and processing quality of rice (*Oryza sativa*) (Juliano 1985). Low amylose levels are usually associated with tender, cohesive, glossy cooked rice, while higher amylose cultivars tend to cook dry, be fluffy and separate (Juliano 1971). Apparent amylose level is also an important determinant of market class. In the US, for example, standard long-grain rice cultivars are specifically selected to contain 20–22% apparent amylose, while standard medium- and short-grain cultivars are specified to contain 14–19% amylose (Webb 1985).

Control of the apparent amylose content in rice has been the subject of numerous genetic studies. In most crosses, apparent amylose content has been found to be controlled by an allelic series at one locus with major effects and by one or more modifier genes with minor effects (Bollich and Webb 1973; McKenzie and Rutger 1983; Pooni et al. 1993a). This is consistent with the fact that, as in other plants, amylose in rice

kernels is produced by granule-bound starch synthase encoded by the *Wx* gene (Shure et al. 1983; Sano 1984; MacDonald and Preiss 1985). However, inheritance of apparent amylose content can sometimes be complex due to epistasis, cytoplasmic effects, and to the triploid nature of endosperm (Pooni et al. 1993 b). Stansel (1965), for example, was able to recover high-amylose transgressive segregants from a cross between a glutinous cultivar and the low-amylose cultivar Century Patna 231. Other investigators have postulated the presence of two dominant complementary genes to explain the inheritance of amylose content in certain crosses between low- and high-amylose cultivars (McKenzie and Rutger 1983). Genes such as dull and amylose extender can also modify apparent amylose content (Yano et al. 1985; Juliano et al. 1990; Asaoko et al. 1993). A further complication is that the "apparent amylose" measured in theses studies includes various levels of long-chain amylopectin (Takeda et al. 1987).

Using RFLP analysis, two alleles of the *Wx* locus have been identified (Sano et al. 1986). These alleles largely correspond to the *indica* and *japonica* subspecies of rice, however, and are not adequate to explain all of the observed variation in apparent amylose content among commercial rice cultivars. For example, most commercial rice cultivars grown in the US contain the *Wx*:EcoRV-A allele despite ranging from low (14–19%), intermediate (20–22%) to, in a few cases, high ($>24\%$) levels of amylose (Paul et al. 1995, 1996; W. Park, unpublished observations).

Wang et al. (1995) recently observed that apparent amylose content and the level of waxy protein in 31 rice cultivars from China were correlated with the cultivar's ability to excise the leader intron of the *Wx* transcript. High-amylose cultivars contained only the mature 2.3-kb *Wx* mRNA, while lower amylose cultivars contained varying ratios of both the mature 2.3-kb *Wx* mRNA and a 3.3-kb pre-mRNA. The mechanism responsible for this difference in splicing was not determined, but these authors noted a substitution at the putative leader intron 5′ splice site of the low-amylose cultivar, Hangfeng, that could have affected splicing efficiency.

A polymorphic microsatellite was recently identified in the *Wx* gene (Bligh et al. 1995) located 55 bp upstream of the putative 5′-leader intron splice site. However, the relationship of microsatellite classes to amylose content was not examined. To determine whether the *Wx* microsatellite classes were related to variation in apparent amylose content and whether the polymorphism at the leader intron splicing site was correlated with apparent amylose content, we examined 92 rice cultivars which represent an extended pedigree of current and historically important US rice germ plasm and 101 progeny of a cross between low-amylose and intermediate-amylose breeding lines.

Materials and methods

Materials

Rice cultivars and breeding lines were grown under field conditions at the Texas A&M University System Research and Extension Center located in Beaumont, Texas. For DNA isolation, the cultivars were grown under standard greenhouse conditions and leaves of 3–4-week-old seedlings were harvested.

Experimental procedures

Leaves from 10–20 seedlings were ground in a coffee grinder with dry ice. Twenty milliliters of extraction buffer [6.25 mM potassium ethyl xanthogenate (Fluka), 0.7 M Tris-HCl, pH 7.5, 0.7 M NaCl, 10 mM EDTA, pH 8.0] were added to each 15 ml of powdered tissue. The samples were mixed and were incubated at 65°C for 15 min to 1 h. The samples were extracted once with chloroform : isoamyl alcohol (24 : 1) and DNA was precipitated with 0.1 vol of 3 M sodium acetate, pH 5.2 and 2 vol of ethanol for 10–20 min at room temperature. The DNA was hooked on a glass hook, washed with 70% ethanol, air-dried and re-suspended in 0.5 ml 10 mM Tris-HCl, 1 mM EDTA (pH 8.0). The DNAs were quantitated on a Beckman Model LS50B fluorometer.

For the microsatellite assay, PCR reactions consisted of 20 µl containing 30 ng DNA, 0.2 mM deoxynucleotides, 2.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl, pH 9.0, 0.1% Triton X-100, 0.2 µM oligo 484 (5′-CTTTGTCTATCTCAAGACAC-3′), 0.2 µM oligo 485 (5′-TTGCAGATGTTCTTCCTGATG-3′) and 0.25 units *Taq* polymerase (Promega). Using a PE9600 thermocycler (Perkin Elmer Cetus), the PCR reactions were denatured at 95°C for 4 min, followed by 35 cycles of 94°C for 45 s, 55°C for 30 s and 72°C for 60 s. The final extension was at 72°C for 5 min. After PCR, 13 µl of loading buffer (98% formamide, 10 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol) were added, the samples were denatured at 70–80°C for 2 min and 4-µl aliquots were loaded on 5% polyacrylamide gels. The samples were electrophoresed for 2 h at 60 V. Bands were detected by silver staining (Fritz et al., submitted).

For sequence analysis, a larger fragment was amplified from a subset of 42 cultivars, using oligos 484 and W2R (5′-TTTCCAGCCCAACACCTTAC-3′). The amplified bands ranged from 182 to 206 bp. The PCR product was run on 1.2% agarose gels and the band was excised. The amplified DNA was purified by QIAquick columns (Qiagen) or by electroelution. The DNA was sequenced on an ABI Model 377 automatic sequencer using cycle sequencing according to the manufacturer's instructions (Applied Biosystems, Inc.).

To determine if the AGGTATA/AGTTATA polymorphism at the putative 5′-leader intron splice site could be detected by restriction endonuclease cleavage, 6 µl of 25 mM $MgCl_2$ and 2 units of *Acc*I (BRL) were added to 20 µl of PCR product (using oligos 484 and W2R) and the mixture was incubated for 1 h at 37°C. The samples were electrophoresed for 3 h at 50 V on 1.2% agarose gels.

Apparent amylose contents were determined by the USDA-ARS Rice Quality Laboratory in Beaumont, Texas, using a near-infrared reflectance spectroscopy method (Delwiche et al. 1995) which was developed based on a series of 50 diverse commercial rice samples assayed using the colorimetric assay for rice amylose (Williams et al. 1958) as modified by Juliano (1971) and Juliano et al. (1981).

Statistical analyses were performed using the StatView and SuperAnova programs from Abacus Concepts.

Results

Microsatellite analysis

We screened 92 rice cultivars representing a cross-section of current and historically important US germ plasm using primers flanking the *Wx* microsatellite (Table 1). These included high-, intermediate- and low-amylose cultivars as well as three glutinous cultivars that contain essentially no amylose. We detected all four classes of microsatellites reported by Bligh et al. (1995) and also four additional classes (Fig. 1). The amplified products ranged from 103 to 127 bp in length and contained $(CT)_n$ repeats of between $n = 8$ and $n = 20$. Two predominant classes of $(CT)_n$ repeats were identified; 37 of the cultivars tested contained the $(CT)_{20}$ allele and 27 had the $(CT)_{18}$ allele.

The $(CT)_{18}$ class had significantly lower levels of apparent amylose than all but one of the other classes, whereas the $(CT)_{11}$ class had significantly higher levels of amylose (Table 2). Analysis of variance demonstrated that the *Wx* microsatellites explained 82.9% (adjusted R^2) of the variation in apparent amylose content among the 89 non-glutinous cultivars evaluated ($P < 0.0001$). The *Wx* microsatellites were not, however, always able to distinguish glutinous cultivars. 'CI 5309' was present in a class, $(CT)_{16}$, different from any of the other cultivars tested. However, 'Calmochi 101' and 'Tatsumi mochi' were in the $(CT)_{17}$ class along with the low-amylose cultivars Rico1, Nortai, and Tainan Iku 487.

The inheritance of the *Wx* alleles can be traced through a cross-section of the US rice pedigree that includes the 92 cultivars examined, as shown in Fig. 2. In almost all cases, the *Wx* microsatellite alleles show the expected simple pattern of inheritance and can be traced back to the original introductions of foreign germ plasm. However, it should be noted that the $(CT)_{14}$ allele that is present in 'Katy' is not present in any of its reported ancestors. This may have been due to microsatellite instability during the development of 'Katy', but other causes such as accidental outcrossing can not be ruled out. It should also be noted that the $(CT)_{20}$ allele in the intermediate-amylose cultivar CI 9545 is not present in either 'Rexark' or 'Tainan Iku 487'. Interestingly, 'CI 9545' is the only medium-grain sample that contains the $(CT)_{20}$ allele and is also atypical among medium- grain US germ plasm in containing intermediate rather than low amylose.

The predominate microsatellite classes correspond closely with standard US amylose classes. Almost all of the cultivars in the $(CT)_{18}$ class contain 14–19% amylose, those in the $(CT)_{14}$ or $(CT)_{20}$ class contain 20–23% amylose, and those in the $(CT)_{11}$ class have greater than 23% amylose (Table 1). However, the $(CT)_{17}$ and $(CT)_{19}$ classes contain cultivars with divergent amylose contents. The $(CT)_{17}$ allele was present in two glutinous cultivars and also in three cultivars with apparent amylose contents from 15.1 to 16.5%. As shown in Fig. 2, the $(CT)_{17}$ allele in the glutinous cultivar, Calmochi 101, was derived from its glutinous parent, 'Tatsumi mochi'. The $(CT)_{17}$ allele in the low-amylose cultivars Nortai and Rico1 can be traced back to 'Tainan Iku 487'. However, these two groups of cultivars with the $(CT)_{17}$ allele are only distantly related. Similarly, the low-amylose cultivars Calrose, Calrose 76, CSM3 and their low-amylose ancestor Caloro are only distantly related to the $(CT)_{19}$ cultivars with higher amylose contents, Shoemed and Nira. Since microsatellites are highly polymorphic (Wu and Tanksley 1993), independent evolution of *Wx* alleles with 17 and 19 CT repeats in distantly related cultivars of rice would not be surprising.

The *Wx* microsatellite is polymorphic enough to distinguish most rice cultivars in different amylose classes, yet stable enough to be easily traced through multiple generations of the US rice pedigree. Thus it should be a useful tool to track the *Wx* gene in a wide array of crosses. To test this, we analyzed bulked F_5 progeny derived from 101 F_3 lines of a cross between two breeding lines, Pd1-28 and Pd1-66 (Fig. 3). The line Pd1-28 has intermediate apparent amylose (21.2%) and the $(CT)_{20}$ allele, while Pd1-66 has low apparent amylose (14%) and the $(CT)_{18}$ allele. Twenty eight of the bulked progeny (27.7%) were homozygous for the $(CT)_{20}$ allele, 22 (21.8%) were homozygous for the $(CT)_{18}$ allele, and the remaining 50 (49.5%) contained varying ratios of the two parental alleles (Fig. 3). Despite bulking the F_5 progeny to effectively represent F_3 families, and thus screening mixtures of segregating seed in some cases, the *Wx* microsatellite was still able to explain 81.9% of the variation in amylose content in this cross. Note that the heterozygous samples show partial dominance for high-amylose content. Similar results have also been obtained with three other crosses (data not shown).

Sequence analysis

To help identify the molecular basis for the strong correlation between *Wx* microsatellite class and the level of apparent amylose, we sequenced the region containing the microsatellite from a random set of 42 cultivars representing all eight microsatellite classes (Table 1). Interestingly, all 26 intermediate and high-amylose cultivars have the sequence AGGTATA in the putative 5′-leader intron splice site indicated by Li et al. (1995) and Wang et al. (1995). However, the 14 low-amylose cultivars all have the sequence AGTTATA at this position. This correlation between the G-T polymorphism and apparent amylose content extends even to the divergent $(CT)_{17}$ and $(CT)_{19}$ *Wx* classes and distinguishes all cultivars with greater than 18% amylose from those with 18% or less amylose (Fig. 4).

Table 1 Apparent amylose content and *Waxy* microsatellite class of 92 current or historically important US rice cultivars and breeding lines. Also shown is whether the putative 5′-leader intron splice site has the sequence AGGTATA or AGTTATA. In some cases, indicated by an *, the putative 5′-leader intron splice site was directly sequenced. In others, this single nucleotide substitution was assayed by *Acc*I digestion

Cultivar	Abbreviation	Type	Origin	% Apparent amylose	$(CT)_n$	G-T
Alan	ALAN	Long	AR	20.9	20	G
Arkrose	AROS	Medium	AR	15.4	18	T
Belle Patna	BLPT	Long	TX	20.9	20	G
Bellemont	BLMT	Long	TX	21.5	20	G*
Bengal	BNGL	Medium	LA	12.6	18	T
Blue Rose	BROS	Medium	USA	16.2	18	T*
Bluebelle	BBLE	Long	TX	21.0	20	G
Bluebonnet	BBNT	Long	TX	19.7	20	G*
Bluebonnet 50	BB50	Long	TX	21.3	20	G
Bonnet 73	BN73	Long	AR	23.0	20	G
Brazos	BRAZ	Medium	TX	15.6	18	T
Calmochi 101	CM101	Short	CA	1.4	17	T*
Caloro	CALO	Short	CA	15.3	19	T*
Calrose	CROS	Medium	CA	16.0	19	T
Calrose 76	CR76	Medium	CA	15.3	19	T
Carolina Gold	CRGD	Long	Madagascar	21.9	14	G*
Century Patna 231	CP31	Long	TX	12.0	18	T*
CI 5309	5309	Long	China	0.0	16	T*
CI 9122	9122	Long	TX	18.4	20	G
CI 9515	9515	Long	TX	21.9	14	G*
CI 9545	9545	Medium	TX	21.4	20	G
CI 9881	9881	Long	TX	19.8	20	G
CI 9902	9902	Long	LA	22.3	20	G
CI 9187	9187	Long	AR	20.0	20	G
CI 9453	9453	Medium	AR	18.0	18	T
CI 9722	9722	Long	AR	22.1	20	G
Colusa	COLU	Short	CA	16.9	19	T*
CP231/SL017	CPSLO	Long	TX	15.3	18	T
CSM3	CSM3	Medium	CA	15.8	19	T
Cypress	CPRS	Long	LA	21.2	20	G*
Dawn	DAWN	Long	TX	21.2	14	G*
Fortuna	FRTA	Long	LA	19.4	20	G*
Gulfmont	GFMT	Long	TX	21.7	20	G
Gulfrose	GROS	Medium	TX	16.3	18	T
Hill selection long grain	HLSL	Long	USA	13.2	18	T
Jefferson	JEFF	Long	TX	20.3	20	G
Jodon	JODN	Long	LA	25.1	20	G*
Jojutla	JJLA	Long	Mexico	24.0	11	G*
Katy	KATY	Long	AR	21.5	14	G*
Kaybonnet	KBNT	Long	AR	21.7	14	G
L-201	L201	Long	CA	19.5	20	G
L-202	L202	Long	CA	24.5	20	G*
LA 110	L110	Medium	LA	26.0	11	G
Labelle	LBLE	Long	TX	22.0	14	G*
Lacassine	LCSN	Long	LA	22.0	20	G
Lacrosse	LACR	Medium	LA	14.2	18	T
Lady Wright	LDWR	Long	USA	13.2	18	T*
LaGrue	LGRU	Long	AR	22.2	20	G
Leah	LEAH	Long	LA	22.4	20	G
Lebonnet	LBNT	Long	TX	21.0	20	G
Lemont	LMNT	Long	TX	21.4	20	G*
M-201	M201	Medium	CA	12.3	18	T*
Magnolia	MGNL	Medium	LA	13.0	18	T
Mars	MARS	Medium	AR	14.5	18	T*
Maybelle	MBLE	Long	TX	21.9	20	G*
Millie	MILL	Long	AR	21.2	20	G
Newbonnet	NBNT	Long	AR	22.9	14	G*
Newrex	NWRX	Long	TX	25.5	11	G*
Nira	NIRA	Long	LA	22.0	19	G*
Nortai	NTAI	Short	AR	16.5	17	T
Northrose	NROS	Medium	AR	15.4	18	T
Nova	NOVA	Medium	AR	14.1	18	T
Nova 66	NV66	Medium	AR	12.4	18	T
Orion	ORIN	Medium	AR	14.8	18	T*
Palmyra	PLMR	Medium	MO	16.5	18	T
Pecos	PCOS	Medium	TX	16.9	18	T

Table 1 Continued

Cultivar	Abbreviation	Type	Origin	% Apparent amylose	$(CT)_n$	G-T
PI 331581	T018	Long	TX	21.4	20	G*
Rexark	RXAR	Long	AR	15.8	18	T*
Rexmont	RXMT	Long	TX	25.4	11	G*
Rexoro	RXOR	Long	LA	21.4	20	G*
Rico1	RICO	Medium	TX	15.1	17	T*
Rosemont	RSMT	Long	TX	21.5	20	G*
Saturn	STRN	Medium	LA	16.0	18	T
Shoemed	SHMD	Short	Philippines	19.7	19	G*
Sri Lanka H4	SRLA	Medium	Sri Lanka	26.0	11	G
Sinawpagh	SNPG	Long	Philippines	19.8	20	G*
Skybonnet	SKBT	Long	TX	22.6	20	G
Starbonnet	STBN	Long	AR	22.1	20	G
Supreme Bluerose	SBRS	Medium	USA	16.2	18	T*
Taichung Native-1	TN-1	Long	Taiwan	24.0	11	G*
Tainan Iku 487	T487	Medium	Taiwan	16.0	17	T
Tatsumi mochi	TTSM	Short	Japan	0.0	17	T
Tebonnet	TBNT	Long	AR	22.9	14	G
Tetep	TETP	Medium	Vietnam	21.8	8	G*
Texas Patna	TXPT	Long	TX	22.3	20	G*
Texas Patna 49	TP49	Long	TX	21.2	20	G
Texmont	TXMT	Long	TX	20.9	20	G
Toro	TORO	Long	LA	15.5	18	T
Toro-2	TOR2	Long	LA	14.9	18	T*
Vegold	VGLD	Long	AR	18.5	20	G
Vista	VSTA	Medium	LA	13.3	18	T
Zenith	ZNTH	Medium	AR	18.0	18	T*

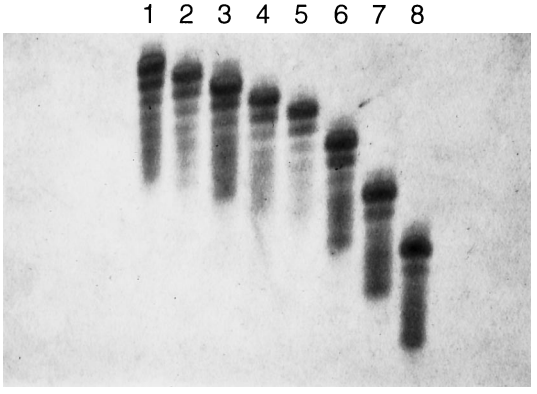


Fig. 1 Silver-stained polyacrylamide gel showing eight *Wx* microsatellite alleles. *Lane 1* – Lemont, $(CT)_{20}$; *lane 2* – Nira, $(CT)_{19}$; *lane 3* – Nova, $(CT)_{18}$; *lane 4* – Rico, $(CT)_{17}$; *lane 5* – CI5309, $(CT)_{16}$; *lane 6* – Katy, $(CT)_{14}$; *lane 7* – Sri Lanka H4, $(CT)_{11}$; *lane 8* – Tetep $(CT)_8$

Table 2 Comparison of the mean amylose contents of seven microsatellite classes of the *Wx* locus among 89 non-glutinous representatives of the US rice germ plasm pool

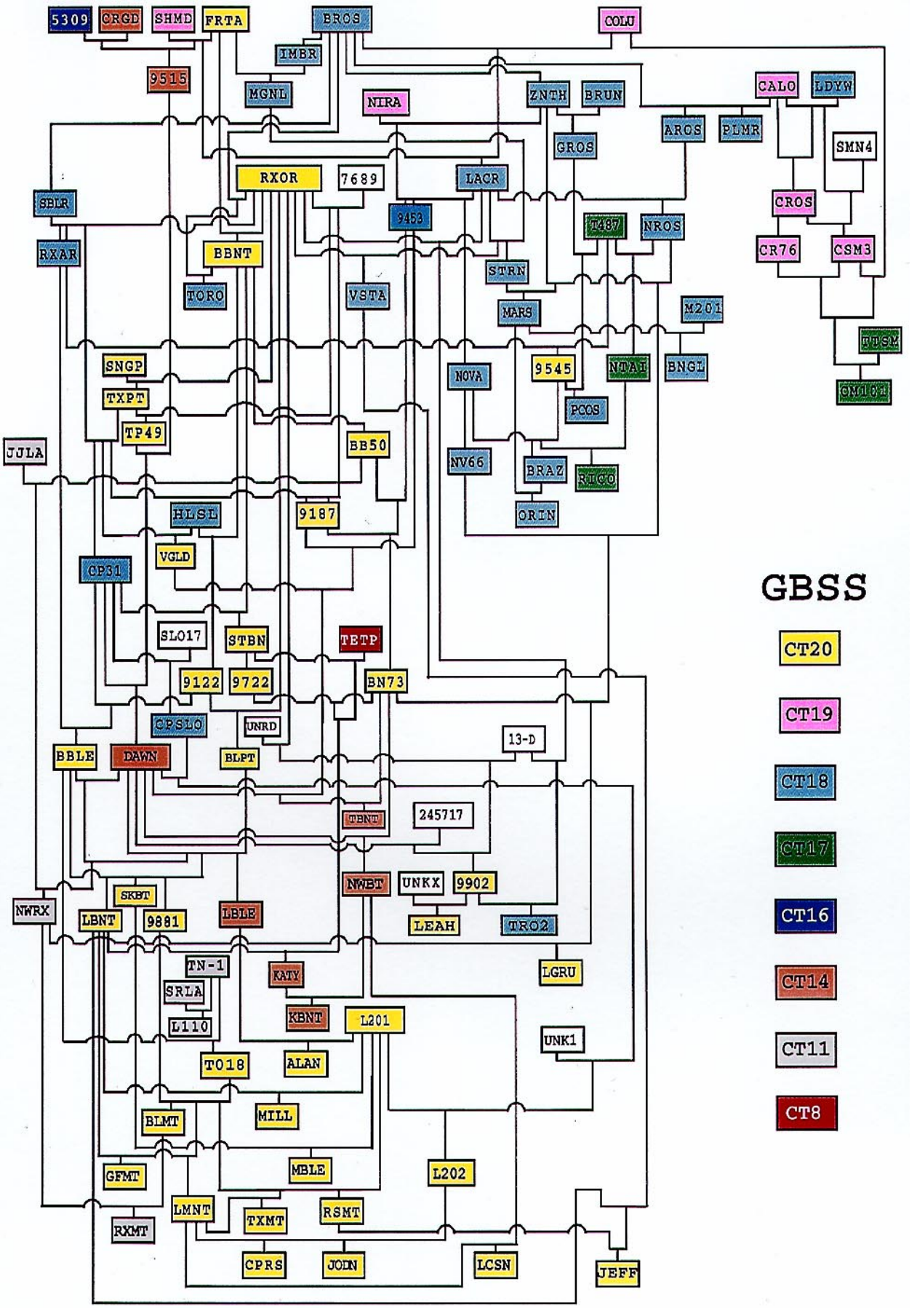
Microsatellite class	Number	Apparent amylose[a]
$(CT)_{18}$	27	14.9 a
$(CT)_{17}$	3	15.9 ab
$(CT)_{19}$	7	17.3 b
$(CT)_{20}$	37	21.3 c
$(CT)_8$	1	21.8 c
$(CT)_{14}$	8	22.0 c
$(CT)_{11}$	6	25.2 d

[a] Means having a different letter are significantly different ($P < 0.05$) using Fisher's Protected LSD

Overall, this single G-T polymorphism could explain 80.1% (adjusted R^2) of the total observed variation in apparent amylose content in the 40 non-glutinous cultivars that were directly sequenced. However, this polymorphism was not able to distinguish the glutinous cultivars Calmochi 101 and CI 5309, both of which contained the sequence AGTTATA, and was also not able to discriminate between intermediate- and high-amylose cultivars as did the CT microsatellite (Fig. 4).

We noted that this G-T base substitution occurred at a putative *Acc*I digestion site. To determine if the polymorphism could be detected without sequencing, we digested the amplified product with *Acc*I and ran the reaction mixture on agarose gels (data not shown). In all cases, amplified fragments containing the sequence AGGTATA were cleaved by *Acc*I and gave products of the expected size, while the corresponding amplified fragments from cultivars having AGTTATA at this position were not cleaved. Thus this key G-T polymorphism can easily be assayed without running sequencing reactions or polyacrylamide gels.

Based on data from all 89 non-glutinous cultivars tested, this single G-T polymorphism could explain 79.7% of the total variation in apparent amylose


5309
CRGD
SHMD
FRTA
BROS
COLU
IMBR
9515
MGNL
NIRA
ZNTH
BRUN
CALO
LDYW
AROS
PLMR
GROS
SMN4
RXOR
7689
LACR
SBLR
CROS
9453
T487
NROS
RXAR
BBNT
CR76
CSM3
STRN
TORO
VSTA
MARS
M201
TTSM
SNGP
NOVA
9545
NTAI
BNGL
TXPT
PCOS
TP49
BB50
JJLA
NV66
BRAZ
RICO
HLSL
9187
ORIN
VGLD
CP31
GBSS
SLO17
STBN
TETP
CT20
9122
9722
BN73
CT19
CPSLO
UNRD
13-D
BBLE
DAWN
BLPT
CT18
TBNT
245717
CT17
NWBT
UNKX
9902
SKBT
NWRX
LBNT
9881
LBLE
LEAH
TRO2
CT16
TN-1
KATY
LGRU
SRLA
CT14
KBNT
L201
L110
UNK1
CT11
T018
ALAN
CT8
BLMT
MILL
GFMT
MBLE
L202
LMNT
TXMT
RSMT
RXMT
CPRS
JOIN
LCSN
JEFF

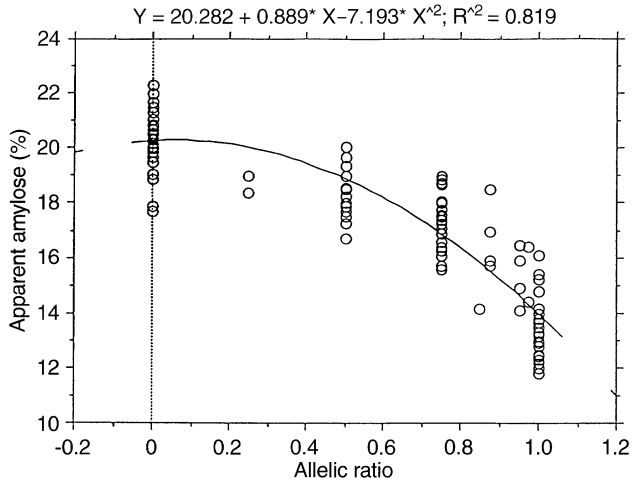


Fig. 3 Relationship between the *waxy* microsatellite allele and the apparent amylose content of 101 bulked F_5 samples representing 101 F_3 progeny of a cross between intermediate- and low-amylose breeding lines. Samples with an allelic ratio of "0" are homozygous for the $(CT)_{20}$ *Wx* allele from the intermediate-amylose parent Pdl-28, while those with an allelic ratio of "1" are homozygous for the $(CT)_{18}$ *Wx* allele from the low-amylose parent Pdl-66

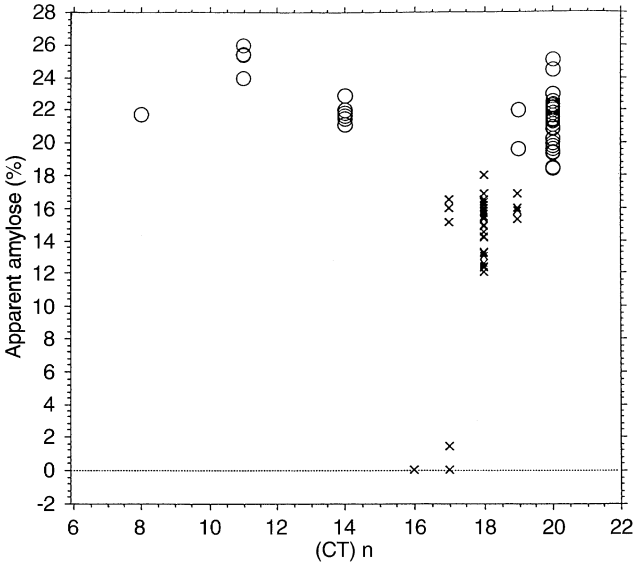


Fig. 4 Apparent amylose content, *Wx* microsatellite class and the putative 5′-leader intron splice site sequence of 92 short- medium- and long-grain US rice cultivars. All of the cultivars tested have either the sequence AGGTATA (O) or AGTTATA (×) at the 5′-leader intron splice site postulated by Wang et al. (1995)

content. Using both microsatellite class and the G-T polymorphism, 85.9% of the total variation in apparent amylose content of the 89 non-glutinous cultivars tested could be explained. This was slightly higher than the 82.9% of variation explained by the microsatellite class alone because the intermediate- and low-amylose members of the $(CT)_{19}$ class were resolved by the G-T polymorphism (Fig. 4).

Discussion

Previous studies using RFLPs have identified two alleles of the *Wx* gene: the Wx^b allele which predominates in *indica* cultivars and the Wx^a allele which predominates in *japonica* cultivars (Sano et al. 1986; Paul et al. 1995; 1996). However, these two alleles are not adequate to explain the large variation observed in amylose contents among US rice cultivars. For example, in a survey of world germ plasm that included seven of the cultivars shown in Fig. 2, the *Wx*:*Eco*RV-B allele was found in 'Jojutla', 'Rexmont', and 'Tetep' (Paul et al. 1995; 1996). These cultivars had apparent amylose contents ranging from 21.8 to 25.4%. However, the *Wx*:*Eco*RV-A allele was found in 'Carolina Gold', 'CP/SLO', 'Lemont', 'Shoemed', 'Texas Patna' and 'Toro', which ranged from 15.3 to 23.6% amylose. We found a similar lack of close correspondence between RFLP alleles and apparent amylose content in a separate study which included 21 of the cultivars shown in Fig. 2. The *Wx*:*Eco*RV-B allele was generally associated with high-amylose content, but the *Wx*:*Eco*RV-A allele was present in low- and intermediate-amylose cultivars, as well as in few high-amylose cultivars such as L-202 and Jodon (data not shown). In contrast to the two alleles seen using RFLP analysis, the CT microsatellite allows the *Wx* locus to be resolved into at least eight different alleles. Because of this increased resolution, more than 82% of the total variation in apparent amylose content among the 89 non-glutinous cultivars tested could be explained based on which *Wx* allele was present.

The presence of multiple alleles at the *Wx* locus which account for most of the variation in apparent amylose content agrees with previous genetic studies (Bollich and Webb 1973; McKenzie and Rutger 1983; Pooni et al. 1993 a). Most of the remaining variation can easily be accounted for by a combination of factors including postulated modifier genes, environmental effects (Sano et al. 1985) and analytical error. It should be noted that the germ plasm tested ranged from very early maturing cultivars such as Labelle and Maybelle that ripen in less than 115 days to longer season cultivars such as Rexoro that require more than 160 days for maturity. Thus, development of the starch granules was occurring under different environmental conditions. It should also be noted that the "apparent amylose" determined in these studies may be biased by the presence of various amounts of long-chain amylopectin (Takeda et al. 1987).

Fig. 2 Pedigree associations of *Wx* microsatellite alleles in current and historically important cultivars in the US rice gene pool. Note that multiple backcrosses with the same cultivar are not indicated

Based on previous genetic studies, the *Wx* microsatellite would have been expected to explain a large proportion of variation in apparent amylose content among the progeny from a wide range of crosses. However, one would not necessarily have predicted that it would explain such a large proportion of the variation in amylose content among a group of diverse cultivars produced by many different breeders working at several locations over a period of more than 80 years.

One reason for the very strong relationship we observed between *Wx* microsatellite alleles and apparent amylose content is that the US rice germ plasm base is very narrow (Dilday 1990). In many cases, individual *Wx* genes can be traced through multiple generations of the US rice pedigree as identical by descent (Fig. 2). For example, the intermediate amylose $(CT)_{14}$ allele can be traced as identical by descent from 'Carolina Gold', a long grain, firm texture rice extensively grown in the southeastern US in the 18th and 19th centuries, to the cultivar Labelle released in Texas in 1972 and to the cultivar Newbonnet which was released in Arkansas in 1983. The high-amylose $(CT)_{11}$ allele can be traced as identical by descent from the unadapted Mexican cultivar Jojutla, to the high-yielding semidwarf cultivar Rexmont released in 1986. Perhaps the most interesting case, however, is provided by the low-amylose $(CT)_{18}$ allele which is predominant in medium-grain cultivars and also found in specialty long-grain cultivars such as Toro-2. In most cases, the $(CT)_{18}$ allele can be traced as identical by descent to 'Blue Rose' a medium-grain cultivar selected in 1907 by S.L. Wright from a vigorous plant of unknown origin which was found in a field of Japanese rice growing near Jennings, Louisiana (Johnston 1958). The identity of the Japanese cultivar in which the progenitor of 'Blue Rose' was found is not known, but we have found that the $(CT)_{18}$ allele is present in a number of low-amylose Japanese cultivars such as Asahi (data not shown).

Pedigree analysis has proven to be very useful in situations, such as human genetics, where controlled crosses are not possible (reviewed in Lander and Schork 1994). However, even when controlled crosses are possible, pedigree analysis provides a way of extracting additional information and making useful predictions. For example, examining the relationship between *Wx* microsatellite alleles and amylose content across an extended pedigree directly indicates where differences in amylose content between cultivars can be expected to be due to differences in the *Wx* gene, and thus where the *Wx* microsatellite can be used for marker-assisted selection in future crosses. It also provides a direct way of identifying variation in amylose content that is not due to differences in the *Wx* gene. For example, Stansel (1965) was able to recover high-amylose transgressive segregants from a cross between a glutinous cultivar and 'Century Patna 231', but not from a similar cross with 'Toro'. This was not likely to have been due to differences in the *Wx* genes of 'Century Patna 231' and 'Toro' since they both contain a $(CT)_{18}$ allele that appears to be identical by descent from 'Blue Rose'. In contrast, the fact that 'L-202' and its progeny, 'Jodon', have 2–3% higher amylose contents than standard US long-grain cultivars and different amylographic paste viscosity profiles (Tseng et al. 1984; Linscombe et al. 1995) could be due to differences in the *Wx* gene since one of the parents of 'L-202' was an unknown mixed semidwarf cultivar which may have contained a different version of the $(CT)_{20}$ allele. Note that in this respect the G-T polymorphism at the putative *Wx* 5′-leader intron splice site subdivides the $(CT)_{19}$ microsatellite class based on amylose content, but does not separate cultivars such Tainan Iku 487, Nortai and Rico1, which would be predicted to contain genes that are identical by descent.

One of the most striking findings of this study was that a single G-T polymorphism at the putative 5′-leader intron splice site could explain a large proportion of the variation in apparent amylose content among non-glutinous cultivars. In the course of this work, we also identified a number of other polymorphisms in the *Wx* genes from US rice cultivars. However, most of these were more consistent with the RFLP classification into Wx^a and Wx^b alleles than with apparent amylose content and none was as closely associated with apparent amylose content as the AGGTATA/AGTTATA polymorphism in the putative 5′-leader intron splice site. Our results are consistent with the hypothesis of Wang et al. (1995) that substitution at this site might interfere with the binding of sn1 RNA and thus reduce the efficiency of RNA splicing in low-amylose cultivars. It should be noted that the sequence AGGTATA, which we found in all of the intermediate- and high-amylose cultivars tested, conforms better to the monocot splice 5′-consensus $A_{60}G_{80}G_{100}T_{100}$ $A_{69}G_5G_{50}$ (Goodall and Filipowicz 1991) than to the alternate allele. However, while the number of observations supporting the hypothesis of Wang et al. (1995) has increased from 2 to 91 with the current study, it should be noted that the hypothesis remains to be definitively proven and that the actual 5′-leader intron splice site in most cultivars has yet to be determined.

Other factors must also be involved in determining the amount and activity of granule-bound starch synthase because the glutinous cultivars tested have the same sequence at the putative 5′ splicing site as those containing 14–19% amylose. Also, unlike the CT microsatellite, the G-T polymorphism at the putative 5′-leader intron splice site does not distinguish intermediate- and high-amylose cultivars. Experiments are currently in progress to further define the mechanisms which control apparent amylose content in rice. We are particularly interested in identifying other polymorphisms which modify the amount of *Wx* gene transcription or which alter the specific activity and

specificity of the granule-bound starch synthase produced.

Acknowledgements Samples of 'Carolina Gold', 'Nira', 'Shoemed' and 'Sinawpagh' DNA were generously provided by Dr. Susan McCouch. We thank Grace Walker for excellent technical assistance, Dr. Hongyong Fu for performing some of the sequencing reactions, and Dr. Bill Webb for apparent amylose determinations and technical guidance. This work was funded by the USDA-ARS, the Texas Advanced Technology Development Program, the Texas Rice Research Foundation and the Texas Agricultural Experiment Station.

References

Asaoka M, Okuno K, Yano M, Fuwa H (1993) Effects of shrunken and other mutations on the properties of rice endosperm starch. Starch 45 : 385–387

Bligh HFJ, Till RI, Jones CA (1995) A microsatellite sequence closely linked to the *waxy* gene of *Oryza sativa*. Euphytica 86 : 83–85

Bollich CN, Webb BD (1973) Inheritance of amylose in two hybrid populations of rice. Cereal Chem 50 : 631–636

Delwiche SR, Bean MM, Miller RE, Webb BD, Williams PC (1995) Apparent amylose content of milled rice by near-infrared reflectance spectrophotometry. Cereal Chem 72 : 182–187

Dilday RH (1990) Contribution of ancestral lines in the development of new cultivars of rice. Crop Sci: 30 : 905–911

Goodall GJ, Filipowicz W (1991) Different effects of intron nucleotide composition and secondary structure on pre-mRNA splicing in monocot and dicot plants. EMBO J 10 : 2635–2644

Johnston TH (1958) Registration of rice varieties. Agro J 50 : 694–700

Juliano BO (1971) A simplified assay for milled-rice amylose. Cereal Sci Today 16 : 334–340

Juliano BO (1985) Criteria and test for rice grain quality. In: Juliano BO (ed) Rice chemistry and technology. American Association of Cereal Chemists, Inc. St. Paul, Minnesota, USA, pp 443–513

Juliano BO, Perez CM, Blakeney AB, Castillo T, Kongseree N, Laignelet B, Lapis ET, Murty VVS, Paule CM, Webb BD (1981) International cooperative testing on the amylose content of milled rice. Starch 33 : 157–162

Juliano BO, Perez CM, Kaushik R, Khush GS (1990) Some grain properties of IR36-based starch mutants. Starch 42 : 256–260

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science 265 : 2037–2048

Li YZ, Ma HM, Zhang JL, Wang ZY, Hong MM (1995) Effects of the first intron of rice waxy gene on the expression of foreign genes in rice and tobacco protoplasts. Plant Sci 108 : 181–190

Linscombe SD, Jodari F, McKenzie KS, Bollich PK, Groth DE, White LM, Dunand RT, Sanders DE (1995) Registration of Jodon rice. Crop Sci 35 : 1217–1218

MacDonald FD, Preiss J (1985) Partial purification and characterization of granule-bound starch synthases from normal and *waxy* maize. Plant Physiol 78 : 849–852

McKenzie KS, Rutger JN (1983) Genetic analysis of amylose content, alkali spreading score and grain dimensions in rice. Crop Sci 23 : 306–313

Paul E, Goto M, McCouch SR (1995) RiceGenes, an information system for rice research. Probe 4 : 26–27

Paul, E, Goto M, McCouch SM (Curators) (1996) RiceGenes. Database accessible via World Wide Webb, http://probe.nalusda.gov.8300

Pooni HS, Kumar I, Khush GS (1993 a) Genetical control of amylose content in a diallel set of rice crosses. Heredity 71 : 603–613

Pooni HS, Kumar I, Khush GS (1993 b) Genetical control of amylose content in selected crosses of *indica* rice. Heredity 70 : 269–280

Sano Y (1984) Differential regulation of *waxy* gene expression in rice endosperm. Theor Appl Genet 68 : 467–473

Sano Y, Maekawa M, Kikuchi H (1985) Temperature effects on the Wx protein level and amylose content in the endosperm of rice. J of Hered 76 : 221–222

Sano Y, Katsumata M, 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

Shure M, Wessler S, Federoff N (1983) Molecular identification and isolation of the *Waxy* locus in maize. Cell 35 : 225–233

Stansel JW (1965) Influence of heredity and environment on endosperm characteristics of rice (*Oryza sativa* L.) PhD thesis, Purdue University, University Microfilms, Ann Arbor, Michigan (Diss abstr 27 : 48B)

Takeda Y, Hizukuri S, Juliano BO (1987) Structure of rice amylopectins with low and high affinities for iodine. Carbohydr Res 168 : 79–88

Tseng ST, Carnahan HL, Johnson CW, Oster JJ, Hill JE, Scardaci SC (1984) Registration of L-202 rice. Crop Sci 24 : 1213–1214

Wang ZY, Zheng FQ, Shen GZ, Gao JP, Snustad DP, Li MG, Zhang JL, Hong MM (1995) The amylose content in rice endosperm is related to the post-transcriptional regulation of the *waxy* gene. The Plant Jour 7 : 613–622

Webb BD (1985) Criteria of rice quality in the United States. In: Juliano BO (ed) Rice chemistry and technology. American Association of Cereal Chemists, Inc. St. Paul, Minnesota, USA, pp 403–442

Williams VR, Wu T, Tsai HY, Bates HG (1958) Varietal differences in amylose content of rice starch. J Agric Food Chem 6 : 47–48

Wu K-S, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241 : 225–235

Yano M, Okuno K, Kawakami J, Satoh H, Omura T (1985) High-amylose mutants of rice, *Oryza sativa* L. Theor Appl Genet 69 : 253–257