ROLE OF GBSS ALLELIC DIVERSITY IN RICE GRAIN QUALITY

A Dissertation

by

MACAIRE DOBO

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2006

Major Subject: Molecular and Environmental Plant Sciences

ROLE OF GBSS ALLELIC DIVERSITY IN RICE GRAIN QUALITY

A Dissertation

by

MACAIRE DOBO

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	William D. Park
Committee Members,	Edwin C. Price
	Marla L. Binzel
	Jorge Cruz-Reyes
Chair of Intercollegiate Faculty,	Marla L. Binzel

December 2006

Major Subject: Molecular and Environmental Plant Sciences

ABSTRACT

Role of GBSS Allelic Diversity

in Rice Grain Quality. (December 2006)

Macaire Dobo, B.S., University of Abidjan;

M.S., University of Abidjan

Chair of Advisory Committee: Dr. William D. Park

Amylose content is generally the most important factor determining rice eating and cooking quality. Commercial rice varieties are, in fact, placed into market classes based on having "zero" (0-7%), low (10-20), intermediate (20-25%), or high (>25%) apparent amylose. This study demonstrates that the single-nucleotide polymorphisms (SNP) in the exon 1 (G→T) and in exon 6 (A→C) of GBSS can be used as markers to efficiently distinguish the amylose classes. These two SNPs accounted for 89.2% of the variation in apparent amylose content in a pedigree of 89 US rice varieties and 93.8% of the variation among 279 accessions in a European germplasm collection. All low amylose varieties had the T allele of exon 1. All intermediate amylose varieties had the G allele of exon 1 and the C allele of exon 6. All high amylose varieties had the G allele of exon 1 and the A allele of exon 6. In contrast to previous reports, the amylose content of rice varieties in West Africa was also largely determined by GBSS alleles, which accounted for 93.3% of the variation among 77 samples from West Africa Rice Development Association. GBSS gene from *O. glaberrima* was found to lack a transposon in exon 10 and have an additional polymorphism (G→A) in exon 12, but these do not significantly alter amylose content.

The study also shows that some GBSS genes from high amylose varieties contain an additional C→T polymorphism in exon 10. This SNP does not significantly alter

amylose content, but alters rice starch pasting properties. Traditional RVA analysis of starch pasting properties is complicated by differences in the shear forces between samples. However a simple method was developed to overcome this problem and it was shown that starch granules from rice varieties with the T allele of exon 10 are notably more shear resistant. Amylose and the SNP in exons 1, 6 and 10 of GBSS also played a key role in starch re-association. They accounted for 81 and 71.5%, respectively, of the variation in "gel hardness" of RVA samples which have been allowed to incubate at room temperature for 24 hours.

DEDICATION

This dissertation is dedicated to my wife Sydio Clay Ange Mireille Dobo for her love, contagious optimism and support which enabled me to achieve this goal; to my daughters, Dobo Ruth Ketura Elvyre and Dobo Ruth Keren for their appreciation, hugs and love; to my mother Gboko Sole Julie and to the memory of my father Dobo Dasset Roger; to all the good people who have helped me; and especially to my Lord Jesus Christ.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor, Dr. William D. Park, for his support and guidance. I wish to thank my Committee members: Dr. Cruz-Reyes, for his support and for having been an inspiring lecturer of molecular biology; Dr Marla L. Binzel for her advice, encouragement and also her great sympathy for graduate students in general; and Dr Edwin C. Price, whose assistance and encouragement have been constant throughout my career at A&M. Dr. Edwin played a key role in establishing the project, and all the African rice varieties used in this study were obtained through him.

I am thankful to Grace Walker, technician in Dr Park's lab, for her contribution in generating the RVA data. I, also, want to thank a former student of Dr. Park, Concetta Bormans (Connie) who taught me, with patience, the basic techniques of molecular biology. Special thanks to Dr. Nicola Ayres for her contribution in reading and editing my dissertation. Cathryn, your great assistance and kindness in difficult time will always be remembered.

I would like to express my deep gratitude to Walter De Man, for his assistance and key contribution in establishing the project.

The financial support of WARDA/Rockefeller Foundation is gratefully acknowledged.

LIST OF ABBREVIATIONS

A→C or A/C	Base change adenine to thymine in the exon 6 of GBSS gene
BD	Breakdown value on the RVA curve
C→T or C/T	Base change cytosine to thymine in the exon 10 of GBSS gene
CPV	Cool paste viscosity value on the RVA curve
CV	Consistency viscosity value on the RVA curve
G→A or G/A	Base change guanine to adenine in the exon 12 of GBSS gene
GAC	Guanine, Adenine and Cytosine (SNP in GBSS exons 1, 6 and 10)
GAC-glab	*O. glaberrima* version of GBSS GAC allele (Adenine in exon 12 & no transposon in intron 10)
GAC-sat	*O. sativa* version of GBSS GAC allele (Guanine in exons 12 & transposon in intron 10). This base is also found in other *O. sativa* GBSS alleles at the same position.
GAT	Guanine, Adenine and Thymine (SNP in GBSS exons 1, 6 and 10)
GCC	Guanine, Cytosine and Cytosine (SNP in GBSS exons 1, 6 and 10)
G→T or G/T	Base change guanine to thymine in the exon 1 of GBSS gene
GBSS	Granule bound starch synthase
HPV	Hot paste viscosity value on the RVA curve
NERICA	New Rice for Africa
SB	Setback viscosity value on the RVA curve
SNP	Single Nucleotide Polymorphism
Rice1-ext	Results obtained in RVA when the holding period at 95^{o}C was extended from 2.5 to 10 min.

LIST OF ABBREVIATIONS (Continued)

TAC	Thymine, Adenine and Cytosine (SNP in GBSS exons 1, 6 and 10)
TCC	Thymine, Cytosine and Cytosine (SNP in GBSS exons 1, 6 and 10)
WARDA	West Africa Rice Development Association

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

CHAPTER I

INTRODUCTION

Rice is the staple food for more than half of the world's population. It is an annual plant grown in every continent except the Artic and Antarctic. Rice belongs to the genus *Oryza* which consists of 20 well recognized wild species and two cultivated species, *O. sativa,* originally from Asia, and *O. glaberrima,* endemic to Africa (Viguier 1939). *O. glaberrima* can be distinguished from *O. sativa* by its short ligule and the lack of secondary branches on the panicles.

At maturity, cultivated rice produces seeds that are used as food for humans and animals. As with other cereals, the rice seed consists of a hull and a caryopsis. The caryopsis contains the embryo and the starchy endosperm and is surrounded by three membranes, the pericarp, the tegmen and the aleurone layer. These three membranes and the embryo constitute the rice bran which is removed during the milling process. Most of the seed's lipid and protein are found in the rice bran. The endosperm is essentially made of starch granules and represents 90% of milled rice (Smith et al. 1997).

Starch granules are composed of two distinct types of glucose polymer, amylose and amylopectin. Amylose is essentially a long linear chain of 10^2 - 10^4 1→4 linked α-D-glucose residues with relatively few 1→6 linked branches. Amylose usually contributes 20-30% to the total starch. Amylopectin, the major component of the granule, consists of 10^4 - 10^5 short linear chains of 1→4 linked α-D-glucose residues connected through more frequent (1→6)-α-linkage. Amylopectin contributes the remaining 70-80% of the total starch.

This dissertation follows the style of Cereal Chemistry.

The synthesis of amylose and the linear portion of amylopectin are catalyzed by starch synthases. Theses enzymes include granule bound starch synthase (GBSS), which alone controls the synthesis of amylose, and several soluble starch synthases (SS) which are involved in amylopectin biosynthesis.

During photosynthesis, CO_2 is reduced to triose phosphate by the Calvin cycle. These triose phosphates are used as the primary substrate in the synthesis of glucose-1-phosphate, which is used for the synthesis of ADP-glucose. The latter is a substrate in amylose and amylopectin synthesis. Starch synthases catalyze the transfer of a glucose residue from ADP glucose to the non-reducing end of the growing glucan chain (Figure 1.1). The branched structure of amylopectin is formed by the starch branching enzymes (SBEs), which generate α-(1$\rightarrow$6)-linkages by cleaving internal α-(1$\rightarrow$4) bonds within a pre-existing α-(1$\rightarrow$4) glucan, and transfer the released reducing end to C6 hydroxyls of other glucans. The position of branch linkages relative to one another is not random. Instead, they are organized in clustered structures (Figure 1.2). The A chains that do not carry any other chains are linked to chains B, which in turn are linked to a chain C, containing the single reducing group. The cluster organization contains a high branched region rich in A and B chains, the crystalline region, and a region nearly devoid of chains, the amorphous region.

Even though there are many different grain quality parameters that influence consumer acceptance, it is generally accepted that amylose content is the most important element determining cooking and processing quality (Juliano et al. 1983). Amylose content is directly related to volume expansion and water absorption during cooking and to the texture and cohesiveness of cooked rice. Low amylose varieties usually cook tender, cohesive and glossy. In contrast, high amylose varieties tend to be dry and fluffy when cooked, with firm separate grains.

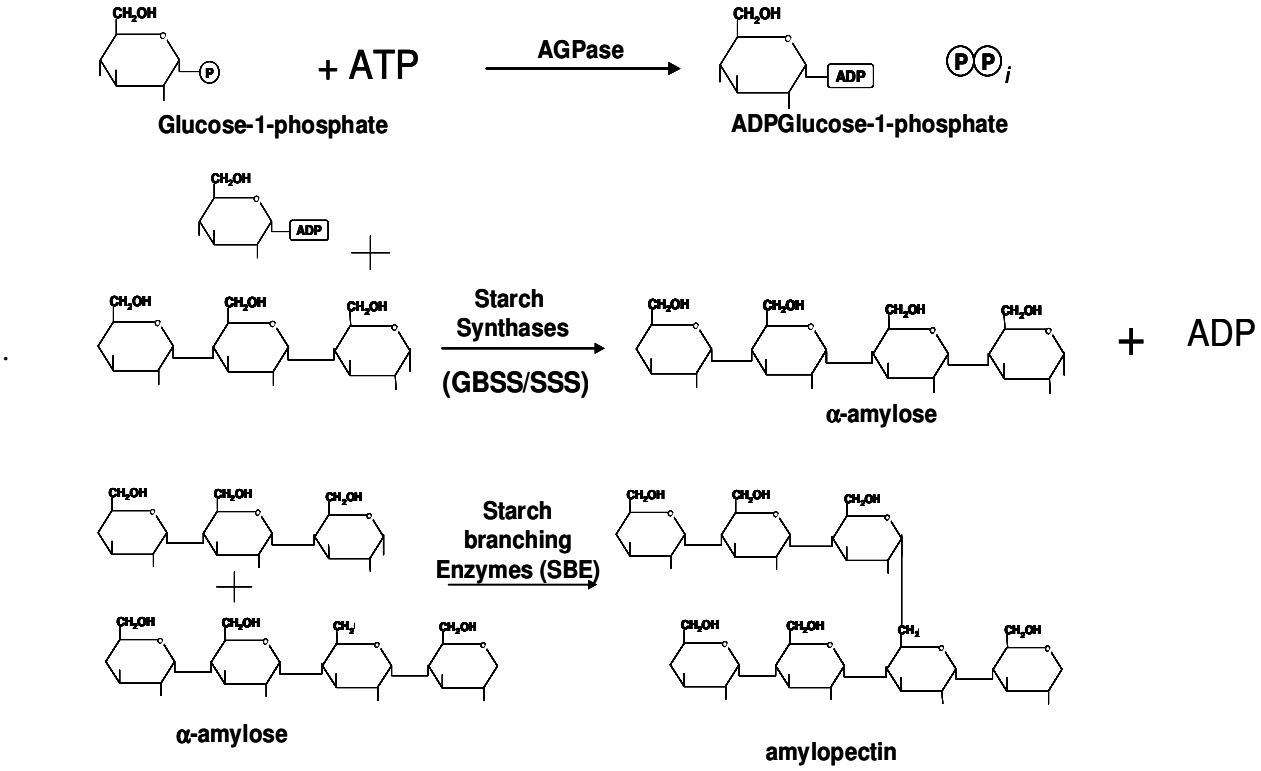

Figure 1.1. Primary reactions of starch synthesis in the endosperm.

GBSS ≈ Granule Bound Starch Synthase

SSS ≈ Soluble Starch Synthase

AGPase ≈ ADP Glucose Pyrophosphorylase

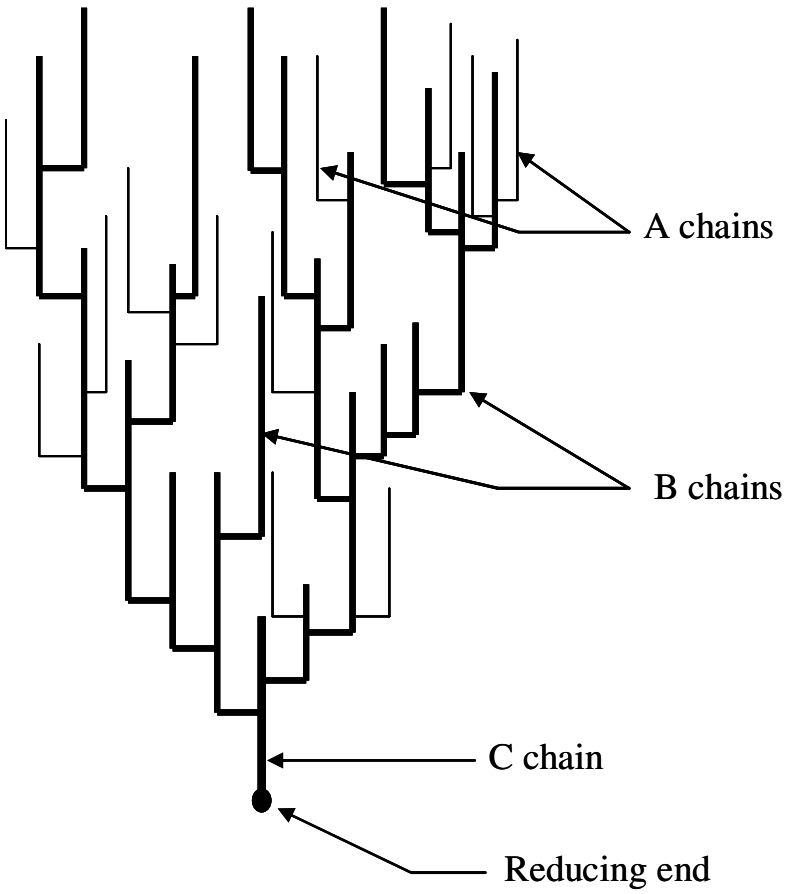

Figure 1.2. Amylopectin clustered structure

Amylose content can be easily assayed with iodine (Juliano 1971) or by Near-Infrared (NIR) Spectrophotometry (Wesley et al. 2003). However amylose assays are increasingly being replaced with DNA markers for granule-bound starch synthase (GBSS or waxy) (Ayres et al. 1997, Bergman et al. 2001, Larkin et al. 2003). Since the enzyme encoded by this gene directly produces amylose, the strong correlation between different alleles of GBSS and amylose content is not surprising.

In a survey of current and historically important US rice cultivars, Ayres et al. (1997) identified seven different alleles of GBSS based on a CT repeat in the 5' leader intron of the gene. These alleles explained more than 82% of the variation in apparent amylose content in the 89 non-glutinous varieties tested and were able to efficiently distinguish low, intermediate and high amylose market classes of US rice. Further work confirmed these results, including that the CT repeat explained 81% and 91% of the amylose variation in studies done on Chinese (Tan and Zhang 2001) and Korean (Shu et al. 1999) rice cultivars.

Strikingly, Ayres et al. (1997) also found that all 52 of the varieties tested with greater than 18% amylose contained the sequence AGGTAT at the GBSS 5' leader intron splice site, while all 37 of the varieties with less than 18% amylose contained the sequence AGTTATA at this location. This single G/T polymorphism was able to explain almost 80% of the variation among these varieties.

Changing the GBSS leader intron sequence from AGGTATA to AGTTATA disrupts the normal binding site for U1snRNP and has a dramatic impact on GBSS mRNA processing and stability. It causes alternate GBSS mRNA splicing at several different sites ranging from -93 to +10, some of which lead to production of upstream open reading frames out of register with the GBSS coding region (Bligh et al. 1995, Larkin and Park, 1999). Surprisingly, it can also lead to use of a non-consensus TT/AG splice

at the +1 site rather than the typical GT/AG seen in the vast majority of mRNAs (Cai et al. 1998).

As a result of the various possibilities for alternate splicing, the phenotypic effect of the G→T change is temperature dependent (Larkin and Park 1999). At 18°C, the G→ T change has only a small phenotypic effect. However, at 25°C and particularly at 32°C, the amount of mature GBSS mRNA in developing seeds is dramatically reduced. This is one of the reasons that low amylose varieties are particularly temperature sensitive and that their amylose content increases at lower temperatures (Asaoka et al. 1985, Hirano and Sano 1998, Larkin and Park, 1999). In contrast, in varieties with the sequence AGGTATA, the steady state level of GBSS mRNA is essentially identical over this temperature range.

The G/T SNP in exon 1 and the CT repeat have both been the subject of extensive studies. However, these are not the only polymorphisms found in the GBSS gene. Larkin and Park (2003) also identified SNPs in exon 6, 9 and 10 in US rice cultivars (Figure 1.3). The SNP in exon 9 does not lead to an amino acid substitution, but the SNPs in exons 6 (A→C) and 10 (C→T) result in amino acid changes serine→tyrosine and serine→proline that are associated with differences in apparent amylose content. Larkin and Park (2003) found that all the intermediate amylose varieties tested contained the cytosine allele of the SNP in both exons 6 and 10, while high amylose varieties generally had the adenine allele in exon 6 and either cytosine or thymidine in exon 10. However the studies were done with only 28 varieties, most of which were from the very narrow US rice pedigree.

Despite the importance of amylose content, it is well known that varieties with the same amylose content sometimes have substantial differences in both palatability and processing behavior (Juliano and Perez 1983, Kuo et al. 2001). Because of this, secondary screens for rice quality such as gelatinization temperature and temperature/viscosity profiles are often used.

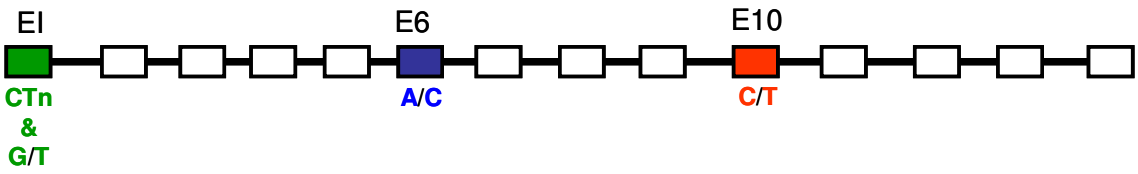

Figure 1.3. Key polymorphisms in rice GBSS gene. Rectangles correspond to exons interspaced by lines representing the introns. The letters below some boxes are the base changes in the gene. Exon 1 (EI) contains a cytosine and thymine repeat (CTn) and a base change guanine to thymine (G/T). Exon 6 (E6) has a base change adenine to cytosine (A/C). Exon 10 (E10) contains a base change cytosine to thymine (C/T). In the GBSS gene sequence of the cultivar Jodon (AF141955), for example, the base changes in EI, E6 and E10 are in positions 144, 2389 and 3381, respectively..

Gelatinization temperature (GT) is the temperature at which the starch granules start to lose crystallinity and birefringence by irreversible expanding. GT can be measured directly using a microscope with a heated stage, but is usually assayed indirectly as alkali spreading value (Juliano 1998). Using map-based cloning, gelatinization temperature was recently shown to be largely determined by amino acid substitutions in SSIIa, a form of soluble starch synthase (Gao et al. 2003). These substitutions alter the enzymatic activity of SSIIa, which in turn changes the fine structure of the crystalline lamellae of amylopectin and the GT of rice.

Analysis of pasting properties has also proved useful in the prediction of rice eating and cooking quality (Juliano 1996, Shu et al. 1998, Bao et al. 1999, Bao and Xia 1999). This analysis was initially done using the Brabender viscoamylograph. However, currently the Rapid Viscoanalyzer (RVA) is more often used due to its ability to quickly assay smaller samples. Many investigators have found correlations between RVA profiles and cooking quality. For example, in a study of Chinese varieties, Shu et al. (1998) found that the hardness of cooked rice was positively correlated to setback viscosity and negatively to breakdown viscosity. In the same manner, a study carried out by Kuo et al. (2001) on Taiwanese varieties revealed a positive correlation between peak viscosity and palatability score, and negative correlation for setback and palatability score. However, it should be noted that other investigators have not found RVA parameters to be correlated with human sensory evaluation of cooking quality (Champagne et al. 1999).

RVA analysis is able to detect differences in starch structure due to starch branching enzyme I and starch branching enzyme III (Liu et al. 2004 and Han et al. 2004), as well as differences due to protein and lipid content (Fitzgerald et al. 2003). However, some of the differences in temperature/viscosity profile between varieties with the same amylose content are apparently due to differences in exon 10 of GBSS (Larkin 1999, Larkin and Park 2003, Larkin et al. 2003). In particular, the C/T SNP in exon 10 of GBSS was able to distinguish US high amylose cultivars such as Rexmont, which have low breakdown

and high setback, from other high amylose cultivars such as L202. This is important because varieties with "Rexmont quality" have low solids loss and superior grain stability during processing that is not seen in other high amylose varieties (Bollich et al. 1990, Larkin and Park 2003, Larkin et al. 2003).

Recently Hirose and Terao (2004) reported eight other starch synthases in rice in addition to GBSS and SSIIa. GBSS and some of the soluble starch synthases (SSII-3 and SSIII-2) were expressed predominately in mid- to late-stages of grain filling. A second granule bound starch synthase, GBSSII, was detected in the early stages of grain filling, along with alternate forms of soluble starch synthases (SSII-2 and SS III-1). Other genes such as SSI, SSII-1, SSIV-1 and SSIV-2 were expressed at more constant levels during the entire period of grain filling (Huawu et al. 2003, Gao et al. 2003). The role of these other forms of starch synthase in cooking quality is not yet known.

While extensive rice cultivation is usually associated with Asia, rice is also the main dietary source of carbohydrate in Africa (FAO 1984). *O. glaberrima,* the other cultivated rice, has been grown in West Africa for over 3500 years (Viguier 1939). *O. sativa* was introduced in Africa only 450 years ago (Semon et al. 2005).

O. glaberrima has higher tillering ability than *O. sativa* and is more tolerant to biotic stresses (diseases, insect pests, weed pressures) and abiotic stresses (drought, soil acidity), (Jones et al. 1997, Dingkuhn et al. 1998). However, *O. glaberrima* is not widely grown for commercial purposes because of its susceptibility to lodging, its low yield, and spontaneous shattering.

Combining the stress resistance of *O. glaberrima* with the yield of *O. sativa* has long been a key goal of African rice breeding programs. However, such crosses are difficult due to sterility barriers. One indication of these sterility barriers is the low level of

intermixing between *O. sativ*a *and O. glaberrima,* even though they have long been grown side-by-side and are even intermixed in fields (Semon et al. 2005).

Recently, the West African Rice Development Association (WARDA) was able to overcome the sterility barriers in *O. sativa*/*O. glaberrima* crosses by extensive backcrossing or by the use of anther culture (Jones et al. 1997). While this has been technically demanding, the resulting "Nerica" varieties have succeeded in combining the high yield potential of *O. sativa* with the biotic and abiotic stress tolerance of *O. glaberrima* (Dingkuhn et al. 1998).

The Nerica varieties have had a transforming effect on rice production in Africa, particularly on upland rice cultivation. Indeed, with the Nerica varieties, many farmers are able produce enough rice to feed their families as well as excess to sell in the market. The impact of the Nerica varieties on African rice production is illustrated by the fact that Monty Jones, a key scientist involved their development, was recently awarded the World Food Prize (www.warda.org).

West Africa contains many cultural and ethnic groups with distinct needs and preferences. To facilitate the utilization of the Nerica rice varieties across the diverse groups, "participatory Varietal selection" (PVS) was carried out. Rather than having the breeder make the final selection before Varietal release, in PVS farmers are given access to a range of elite Nerica and *O. sativa* lines. The farmers then help evaluate the material and choose for themselves which varieties to grow.

PVS has facilitated rapid adoption of Nerica varieties in Africa, but has also demonstrated the importance of rice quality. In Burkina Fasso, for example, the upland rice most favored by producers was FKR 1. This variety yields only 1.5 t/ha, but had higher palatability scores than other varieties with yields up 2.7 t/ha (WARDA 2000). A similar situation was observed in Cote d'Ivoire where varieties such as CC9708

received high rating for palatability, but were not among those identified as having the best yield by producers (WARDA 2002). Thus, in most countries where PVS was conducted, it appears that varieties are selected by producers on the basis of both field performance and palatability.

As local and national markets continue to develop, it is likely that grain quality will become increasingly important. In fact, some African producers are already growing one variety targeted to feeding their families and other varieties which meet grain quality requirements of the market place to generate cash for their other needs. While grain quality might be viewed by some as a luxury for African people, it is actually a key trait for varietal acceptance by producers and for the development of rice markets.

O. glaberrima varieties generally have high amylose contents (Juliano 1993). Some studies suggest that the amylose content of *O. sativa*/*O. glaberrima* progeny is largely due to their particular GBSS alleles. For example, it was shown the GBSS CT repeat explained 79% of the variation in amylose content in a population derived from *O. sativa* x *O. glaberrima* (Aluko et al. 2004). Regions with minor effects were also found on chromosomes 3 and 12 (Li et al. 2004).

In contrast, other data suggest that grain quality may be regulated differently *in O. sativa/O. glaberrima* crosses. Watanabe et al. (2002) reported strong transgressive segregation in which progeny with as little as 15% amylose were found in a cross between an *O. sativa* with 20% amylose and an *O. glaberrima* with 26% amylose. This result is strikingly different than that seen in typical *O. sativa/O. sativa* crosses, but could possibly be explained by the segregation of functionally different forms of more of the enzymes involve in starch biosynthesis.

A study done by Heuer and Miezan (2003) on an *O. sativa*/*O. glaberrima* population suggested a high recombination frequency in an 800 bp region between the GBSS

promoter and structural gene. This suggests that recombination might also occur at high frequency within the GBSS structural gene and lead to the creation of novel alleles in the progeny of such crosses.

Thus far there are only two *glaberrima* GBSS sequences in public databases (Umeda et al. 1991). Both have a 139-bp deletion in intron 10 and as well as a SNP in exon 12 compared to the *O. sativa* variety L202 (accession AF515481). The G to A change seen in exon 12 of *glaberrima* causes a conservative amino acid substitution (asparagine to aspartic acid) in the GBSS protein. The effect of the SNP and deletion has been not clarified.

The goal of this study was to examine the relationship of GBSS alleles to rice quality in three germplasm collections. The specific objectives were to:

1. Further examine a larger set of US samples as well as 279 lines from a European rice germplasm collection to determine the relationship between GBSS alleles and amylose content as well as to determine to what degree GBSS alleles account for differences in pasting properties.
2. Examine the relationship of GBSS alleles with amylose content and temperature viscosity relationships in African rice germplasm. This included *O. sativa, O. glaberrima*, and *O. sativa/O. glaberrima* crosses grown in West Africa. This material was particularly interesting since previous data suggest that amylose content may be regulated differently in *O. sativa/O. glaberrima* crosses (Watanabe et al. 2002). In addition, a collection of *O. glaberrima* from the US germplasm collection was also analyzed.
3. Screen 1000 low amylose lines from the US rice germplasm to find those which lack the AGTTATA found in the 5' leader intron of GBSS in typical low amylose lines. This sequence is known to make GBSS mRNA processing temperature-dependent in typical amylose lines and thus makes grain quality particularly subject to environmental variation.

CHAPTER II

POLYMORPHISM IN THE GBSS GENE AFFECTS AMYLOSE CONTENT IN US AND EUROPEAN RICE GERMPLASM

Introduction

There are many different grain quality parameters that influence consumer acceptance. However, it is generally accepted that amylose content is the most important factor that determines cooking and processing quality of rice (Juliano et Perez 1983, Juliano 1985, Webb 1985, 1991). Amylose content is directly related to volume expansion and water absorption during cooking, and to the texture and cohesiveness of cooked rice. Low amylose varieties usually cook tender, cohesive, and glossy. In contrast, high amylose varieties tend to be dry and fluffy with firm, separate grains when cooked.

Amylose is synthesized by granule bound starch synthase (GBSS), which is encoded by a single gene sometimes called waxy (Echt and Schwartz 1981, Hirano and Sano 1991). Three major GBSS alleles have been identified in rice based on the proportion of amylose in the grain and the amount of GBSS protein on washed starch granules (Sano et al. 1986). Rice varieties with the Wx_a allele have approximately three-fold more GBSS protein than those with the Wx_b allele (Villareal and Juliano 1989). The Wx_b allele is often found in the *japonica* subspecies of *O. sativa*, while Wx_a is primarily found in the *indica* subspecies (Hirano and Sano 1991). Genotypes with the *wx* allele have opaque seeds that contain little if any true amylose and are often referred to as glutinous rice varieties. Despite their lack of true amylose, waxy varieties typically give a low level of residual staining with I_2-KI due to long chains of amylopectin (Takeda et al. 1987). To acknowledge this contribution, the results of the commonly used I_2-KI assays are more properly referred to as “apparent amylose”.

The apparent amylose content of milled rice can be easily assayed with I_2-KI (Juliano 1971) or by near-infrared (NIR) spectrophotometry (Wesley et al. 2003). However direct analysis of apparent amylose requires that one waits until the plant has set seed. In addition to the increase in efficiency of a seedling assay, DNA markers for GBSS also allow homozygous and heterozygous plants to be readily distinguished rather than requiring additional generations of plants to be grown for testing. Furthermore, direct analysis with DNA markers often provides a more definitive way of classifying GBSS alleles compared to apparent amylose assays since it avoids complications such as modifier genes, cytoplasmic factors, as well as environmental effects such as differences in temperature during grain development (Sano et al. 1985, McKenzie and Rutger 1983, Asaoka et al. 1985, Kumar and Khush 1987, Pooni et al. 1993, Larkin 1999). Avoiding complications due to environmental variation is particularly important when plants are grown without replication under non-optimal conditions, as is sometimes the case when screening large numbers of breeding lines.

One way to distinguish different alleles of the GBSS is to determine the length of a CT repeat that is naturally present in the 5' untranslated region of the gene (Bligh et al. 1995). Using this CT repeat, Ayres et al. (1997) identified seven different GBSS alleles in an extended pedigree of 89 non-glutinous US commercial rice varieties and were able to explain 82.9% of the variation in amylose content. Similar success in using this repeat to distinguish low, intermediate and high amylose varieties was seen in subsequent reports (Bligh et al. 1998, Bergman et al. 2001, Tan and Zhang 2001). While this CT repeat is useful, it is simply a very closely linked marker rather than the actual cause of variation in amylose content. Thus varieties with the same number of CT repeats sometimes have different amylose contents and varieties with different numbers of CT repeats sometimes have similar amylose content (Ayres et al. 1997).

Ayres et al. (1997) found that a single G/T polymorphism at the leader intron splice site explained almost 80% of the variation in amylose content of the 89 US commercial

varieties tested. All low amylose varieties tested have the sequence AGTTATA, while intermediate and high amylose varieties have AGGTATA. The key role of this single base change in controlling the processing and steady state levels of GBSS mRNA and the subsequent level of amylose was also seen by others in both US and Asian rice germplasm (Wang et al. 1995, Bligh et al. 1998, Cai et al. 1998, Isshiki et al. 1998, Hirano et al. 1998, Larkin and Park 1999).

While the G/T polymorphism in the GBSS 5' leader intron splice site cleanly differentiates low amylose rice varieties, it does not differentiate between intermediate and high amylose varieties. By sequence analysis, Larkin and Park (2003) identified an adenine/cytosine (A/C) polymorphism in exon 6 and an A/T polymorphism in exon 10 which resulted in non-conservative amino acid changes. In this study, only intermediate amylose varieties were found to have the allelic pattern GCC, *i.e.* the "G" allele of exon 1, the "C" allele of exon 6 and the C allele of exon 10. However, the study of Larkin and Park (2003) only included a relatively small number of samples (28 total) distributed across low, intermediate and high amylose levels.

The first goal of this project was thus to further examine the relationship between GBSS alleles and amylose content in a larger group of US samples, as well as in varieties from a European germplasm collection.

Materials and Methods

The experimental material consisted of the set of historical US rice varieties that was previously examined by Ayres et al. (1997), as well as a collection of 279 varieties and breeding lines from a European germplasm collection that was grown in Northern Italy during the summer of 2004.

DNA was extracted as described previously (Larkin and Park 2003). Briefly, ten milled rice seed were ground in 0.5 mL of 6.25 m*M* potassium ethyl xanthogenate, 100 m*M*

Tris-HCl, pH 7.5, 700 m*M* NaCl and 10 m*M* EDTA pH 8.0. After incubation for 1 hour at 65ºC, the samples were extracted with chloroform:isoamyl alcohol (24:1) and centrifuged at 16,000 x g for 10 min. DNA was precipitated with 0.1 volume of 3*M* sodium acetate (pH 5.2) and 2 volumes of 95% ethanol, and the samples were incubated at -20°C for 2 hours. DNA was recovered by centrifugation for 7 min at 16,000 x g, washed with 70% ethanol, air dried and resuspended in 0.2 mL 10m*M* Tris-HCl, 1 m*M* EDTA (pH 8.0).

The CT repeat and the G/T polymorphism in the leader intron were assayed as described by Ayres et al. (1997). The SNP in exon 6 was initially determined by PCR amplification of a 700 bp region followed by sequence analysis using the BigDye® Terminator V3.1 Cycle Sequencing Kit (Applied Biosystem). This had the advantage of also allowing other potential SNPs in exons 6-9 to be detected. However, this method was relatively expensive. Therefore, a second assay for the polymorphism in exon 6 was developed that used allele-specific forward primers which differed only in their 3' terminal base 5'-CCATACTTCAAAGGAACTT(A/C)-3'. These primers were incubated in separate PCR reactions along with a common reverse primer 5'-TTGTCCTTGCTAGGATCC-3'. Amplification only took place when the target DNA contained the base corresponding to the final one in the primer.

The A/C polymorphism in exon 10 was assayed using the primer pair 5'-CCATACTTCAAAGGAACTTC-3 and 5'-ATACTTCTCCTCCATGCTCT-3'. For both the exon 6 and exon 10 assays, 0.2u*M* each primer pair were combined with 5 to 10 nanograms of DNA, 0.2 m*M* deoxynucleotides, 2.5 m*M* $MgCl_2$, 10m*M* Tris-HCl, pH 9.0, 0.1% Triton X-100, and 0.25 units Mango Taq polymerase (Bioline, Randolph, MA). The reaction was denatured at 95ºC for 4 min, followed by 35 cycles of 94ºC for 45 sec, 55ºC for 30 sec, 72ºC for 1 min, and a final extension of 72ºC for 5 min.

For the exon 10 assay, PCR amplification was followed by addition of 2.5 unit of ApaI for a total 10ul reaction and incubation at 37ºC for 3 hr. The samples were then electrophoresed on 1.5% agarose gels for 2hr at 3V/cm and stained with ethidium bromide (0.01%).

The apparent amylose content of milled grain was measured by staining with I_2-KI and measuring absorbance at 620nM as described by Juliano (1971). Total starch granule-bound protein was examined by a modification of the method of Schwartz and Echt (1982). Briefly, ten milled seeds were soaked overnight in sterile water and were then homogenized for 20 seconds in 1 mL of buffer A (0.05M Tris-HCl, pH 6.8, 2.3% SDS, 5% 2-mercaptoethanol, 10% glycerol) using a FastPrep FP120 (QBIOgene, Irvine, CA). The homogenate was filtered through a single layer of Miracloth (EMD Biosciences, San Diego, CA) and centrifuged at 16,000 x g for one minute. After three additional washes in buffer A and two washes with acetone, the starch granules were dried under vacuum. Fifty mg of granules were then mixed with 0.1 mL of buffer A and heated in a boiling water bath for 5 minutes. After cooling, the slurry was centrifuged at 16,000 x g for 5 minutes and 25 μL of the resulting supernatant was electrophoresed on a 10% SDS polyacrylamide gel according to Laemmli (1970) and stained with Coumassie blue.

Statistical analyses were performed with StatView software (v. 5.0) from Abacus Concept. One way ANOVA (analysis of variance) and Fisher's Protected Least Significant Difference were used to define the significance of amylose at $p < 0.001$.

Results

Analysis of GBSS alleles in the US rice pedigree

To complete analysis of the extended US pedigree examined by Ayres et al. (1997), exons 6 and 10 were directly examined in all of the non-glutinous samples (Figure 2.1). The results are in agreement with previous work by Larkin and Park (2003) on the subset of the US pedigree, except that TN1 was found to have the "A" allele of exon 6 rather than the "C" allele reported by Larkin and Park (2003). This result was confirmed by analysis of multiple accessions of TN-1 from the US germplasm collection.

Among 89 non-glutinous accessions in the US pedigree, five allelic patterns were found: TCC, TAC, GCC, GAC and GAT (Figure 2.1). In the allelic pattern, the first letter corresponds to the G/T polymorphism in exon 1, the second letter corresponds to the A/C polymorphism in exon 6 and the third letter corresponds to the C/T polymorphism in exon 10.

As shown in Figure 2.2, these alleles clearly distinguish all three commercial amylose classes. As previously reported by Ayres et al. (1997), low amylose varieties all had the sequence AGTTATA in exon 1. In agreement with preliminary work by Larkin and Park (2003), all of the intermediate amylose varieties have the allelic pattern GCC. All of the high amylose varieties have either the GAC or GAT allele of GBSS.

As one might expect, the three SNPs in GBSS more accurately predict apparent amylose content than the CT repeat. For example, high amylose varieties with 20 CT repeats can be readily distinguished from intermediate amylose varieties with 20 CT repeats based on SNP in either exon 6 or exon 10. Together, the three SNP explained 89.2% of the variation in apparent amylose content in the 89 non-glutinous varieties tested, compared to 82.9% for the CT repeat alone and 85.9% for the CT repeat plus only the exon 1 SNP.

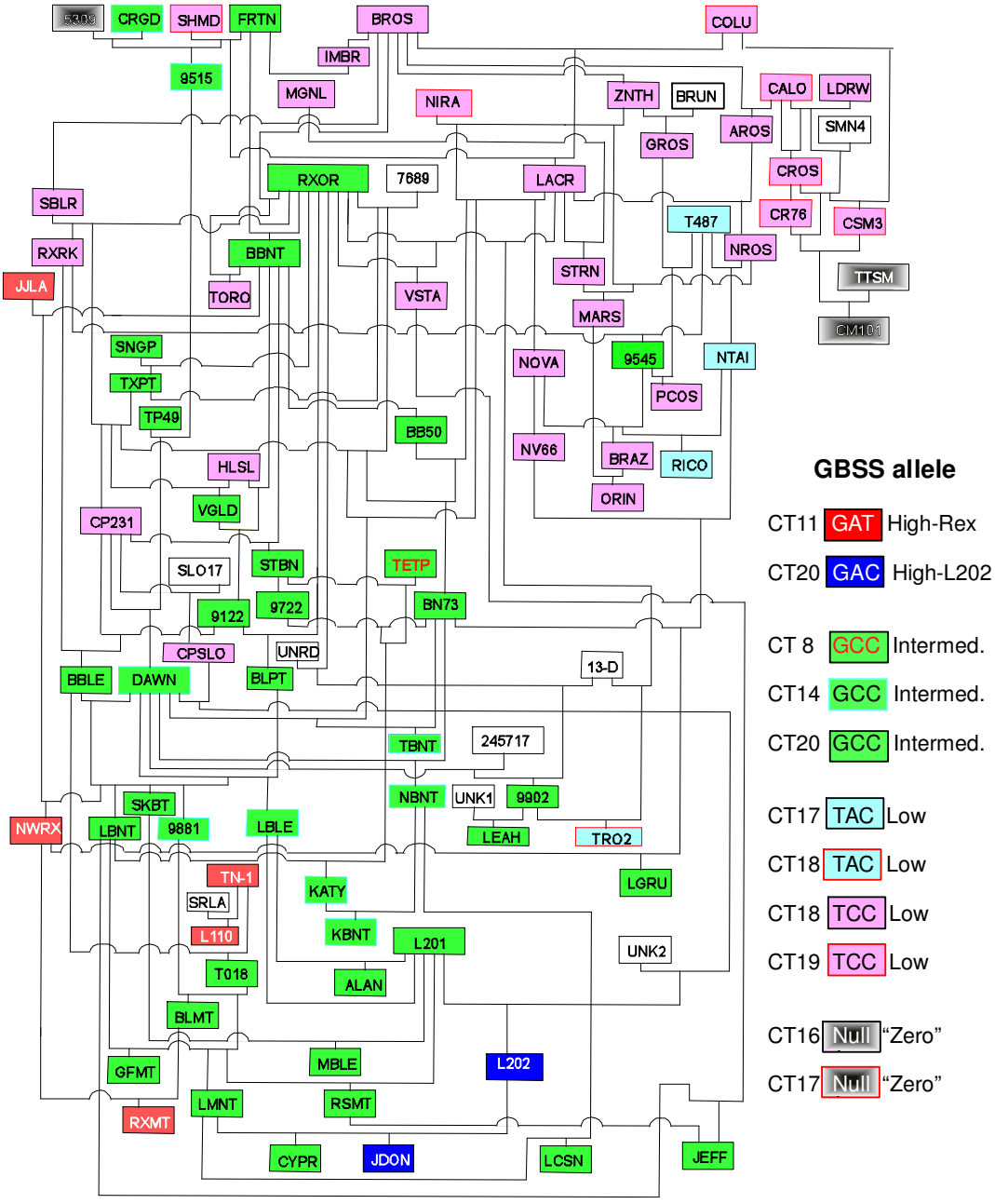

Figure 2.1. GBSS alleles in the US rice pedigree. Pedigree diagram, CT repeat, exon 1 and amylose data were adapted from Ayres et al. (1997).

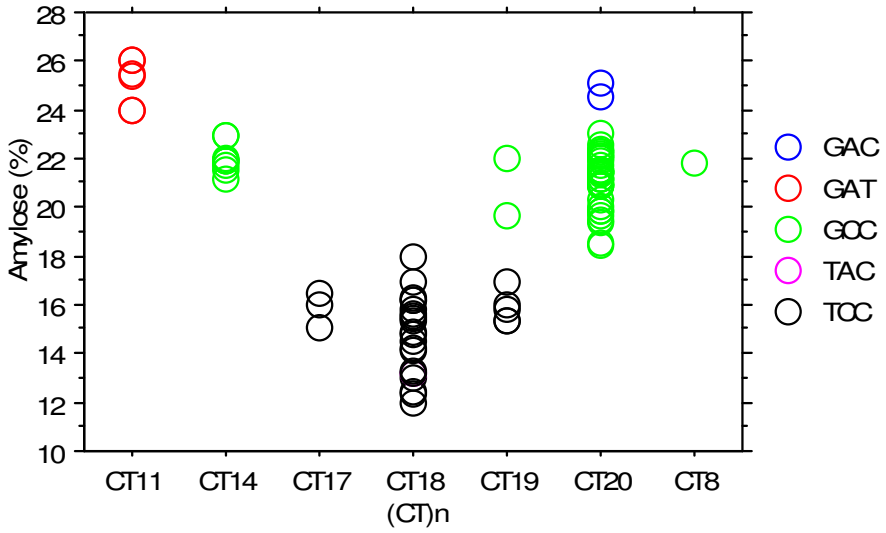

Figure 2.2. GBSS alleles and amylose content in US germplasm. CT repeat and historical amylose data are from Ayres et al. (1997)

As noted previously, the germplasm base of the US rice pedigree is very narrow (Dilday 1990). To determine whether the same simple relationship between GBSS alleles and amylose content was also seen in other germplasm pools, 279 diverse varieties from a European germplasm collection were also examined.

Since the CT repeat in the 5' untranslated region is technically simple to assay, it was examined first. Seven apparently homozygous GBSS alleles could be distinguished in the European collection (Figure 2.3). Five varieties had two different GBSS alleles, but these were hybrid varieties and thus the presence of two alleles was not surprising.

A close, but not exact, correspondence between amylose content and CT repeat was expected based on results from the US pedigree. As shown in Figure 2.3 and Table 2.1, this expectation was borne out. Some CT repeat alleles directly corresponded to low, intermediate and high amylose classes. With one exception, all of the varieties with the CT_{18} and CT_{19} alleles fell into the low amylose class, containing 15-20% apparent amylose. All varieties with the CT_{14} allele are in the intermediate amylose class, containing 21-24% apparent amylose, while all varieties with the CT_{11} allele were in the high amylose class, containing >25% apparent amylose.

While some CT repeat alleles directly corresponded to specific amylose classes, wide variation in amylose content was seen in others. Varieties with CT_{20} alleles contained 21.2-29.3% apparent amylose, spanning both the intermediate and high amylose classes. Varieties with 17 CT repeats also had a wide range of apparent amylose (17.4-24.5%). The CT repeats explained 81.2 % of the variation in apparent amylose, which is consistent with previous work (Ayres et al. 1997, Tan and Zhang 2001).

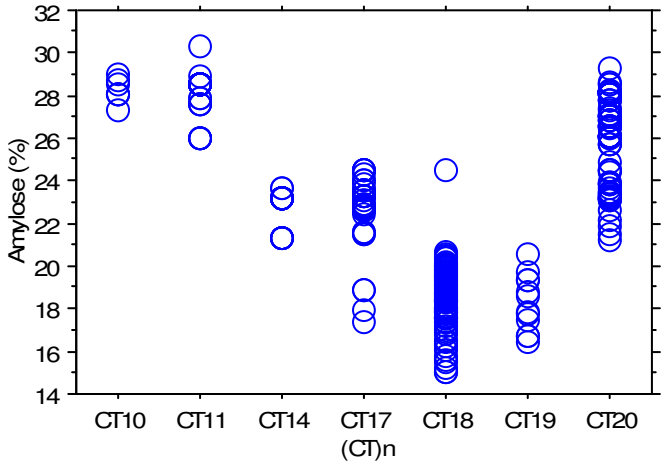

Figure 2.3. Amylose content and GBSS CT repeats in European collection.

Table 2.1. Mean of amylose content of GBSS CT repeat alleles in European collection. Data are included for 279 non-glutinous varieties and breeding lines.

Microsatellite class	Apparent Amylose (%)[a]
$(CT)_{10}$	28.2a
$(CT)_{11}$	28.2a
$(CT)_{20}$	26.3ab
$(CT)_{14}$	22.8b
$(CT)_{17}$	22.2b
$(CT)_{18}$	18.6c
$(CT)_{19}$	18.3c

[a]Values followed by the same letter are not significantly different ($P<0.05$)

Based on the results from the US pedigree, it was anticipated that defining GBSS alleles based on the SNP pattern in exons 1, 6, and 10 would give a better correlation with apparent amylose content. As shown in Figure 2.4, this was the case. The three SNPs explained 93.8% of the variation in apparent amylose among non-glutinous varieties tested, compared to 82.9% for the CT repeat.

As in the US pedigree, all varieties with the AGTTATA version of exon 1 from the European collection had low apparent amylose contents. However, in contrast to the predominance of the TCC allele that was seen in US low amylose varieties, most of the low amylose varieties in the European collection had the TAC allele.

All varieties with the GCC allele were in the intermediate amylose class, averaging 23.1% amylose. Also, all varieties with > 25% amylose had either the GAC or GAT allele. As expected, the average apparent amylose content of varieties with the GAC (28.4% amylose) and GAT (27.3%) was not significantly different.

Combining information about both the number of CT repeats and the three SNPs (Figure 2.5) shows the ability of exon 6 to distinguish between intermediate and high amylose varieties with 20 CT repeats and the ability of exon 1 to distinguish intermediate and low amylose varieties with 17 or 18 CT repeats. However, the variation in amylose content explained by the combination of CT repeats and the SNP pattern in exons 1, 6 and 10 of GBSS was not significantly greater than that explained by the three SNPs alone, 94% vs. 93.8%.

Combining data about the number of CT repeats and the three SNPs allowed the identification of several novel GBSS alleles. In particular, it revealed that a number of premium Italian risotto varieties such as ‘Carnaroli’ contain a GBSS allele with 17 CT repeats and the SNP pattern GCC that was not present in the extended US pedigree (compare Figure 2.1 and Figure 2.6).

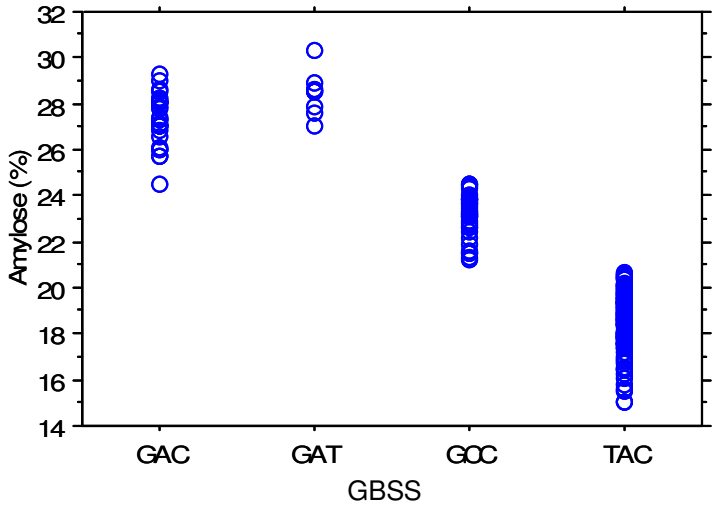

Figure 2.4. Amylose content and GBSS SNP in European collection

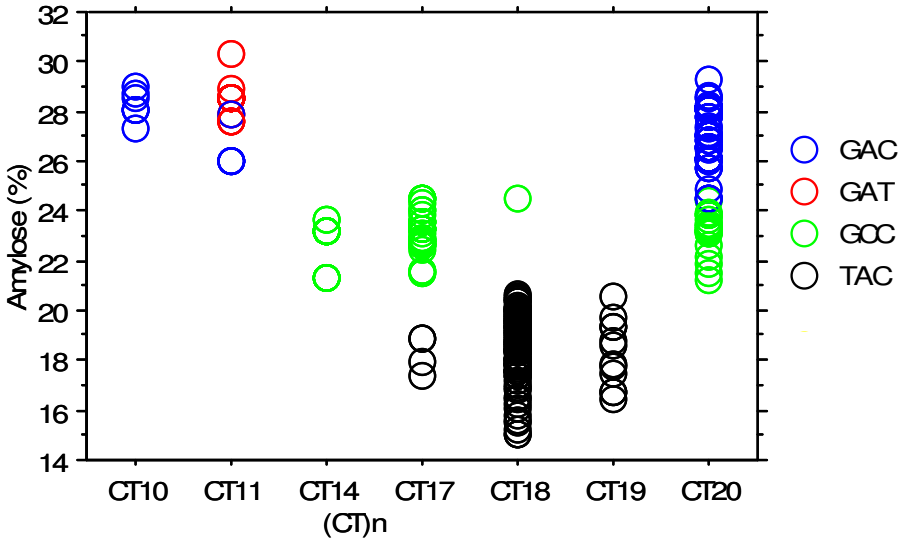

Figure 2.5. Amylose and GBSS CT repeat and SNP alleles in European collection

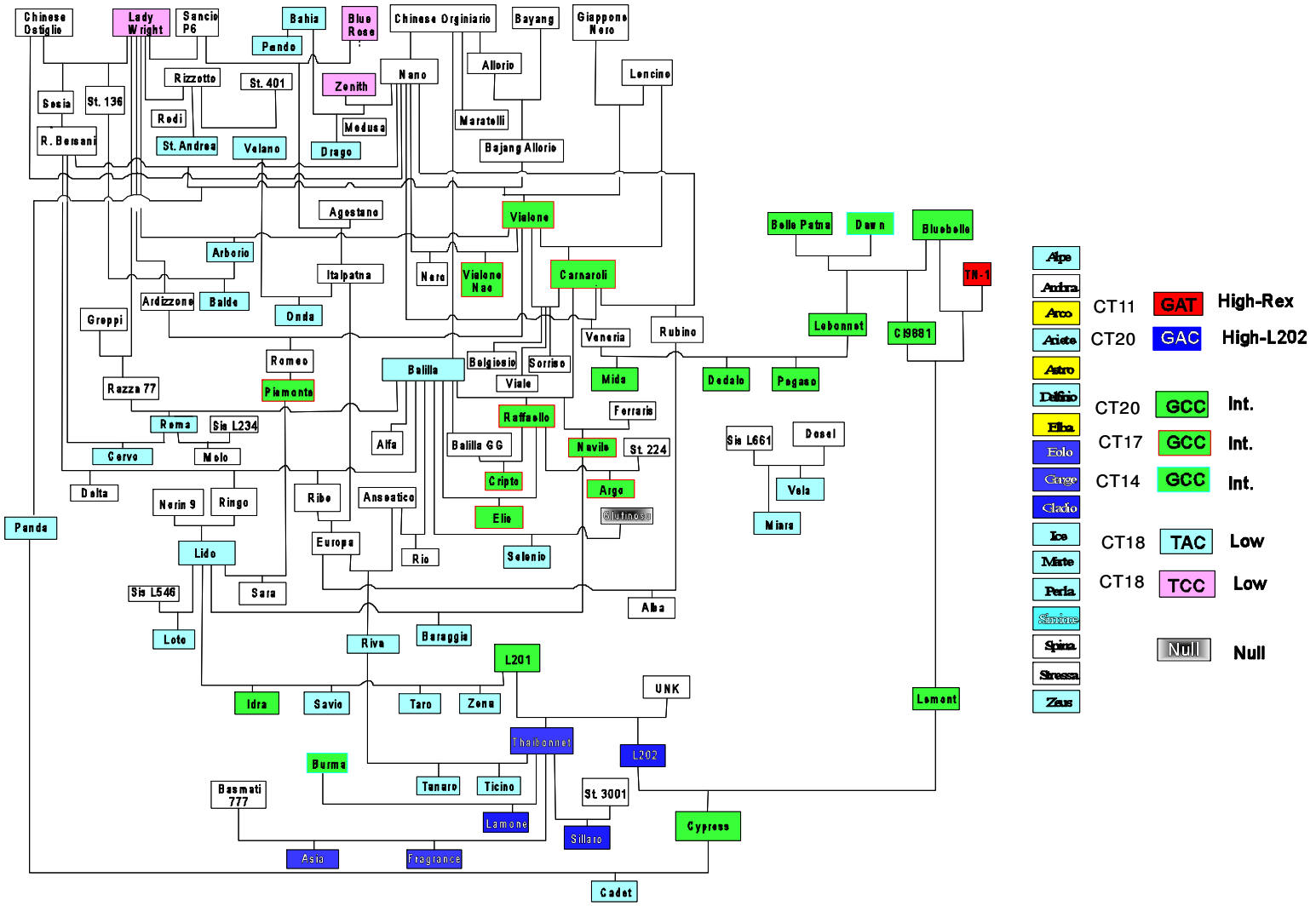

Figure 2.6. GBSS alleles in the Italian rice pedigree

To determine whether this allele contained other mutations which might be responsible for the cooking properties of risotto varieties, the GBSS gene from Carnaroli was sequenced. However the encoded protein was found to be exactly the same as that in typical CT_{20} GCC intermediate amylose varieties such as Lemont.

Relationship between GBSS protein and allelic classes

As discussed above, varieties with an intermediate level of apparent amylose (21-24%) could be reliably distinguished from those with higher apparent amylose based on a single amino acid change in exon 6. The simplest interpretation of this data is that this amino acid polymorphism changed either the amount of the GBSS protein in starch granules or that it changed the specific activity of the GBSS enzyme. Evidence for both possibilities existed. Reduction in the amount of GBSS protein in starch granules in intermediate amylose varieties grown at high temperatures had been noted previously (Larkin 1999) and there was also a previous report of differences in GBSS specific activity between Wx_a and Wx_b varieties (Villareal and Juliano 1989).

To address this question, protein was extracted from washed starch granules and electrophoresed on SDS gels. It has been shown previously that the primary protein on such washed granules is GBSS which has an apparent molecular weight of 60kDa (Schwartz and Echt 1982, Sano 1984).

The results of this analysis showed that the amount of the GBSS protein on washed starch granules correlated well with amylose content (Figure 2.7). Large amounts of GBSS protein were observed for high amylose varieties containing either the GAC or GAT allele and low amounts were seen in low amylose varieties containing the TAC allele, as would be expected due to the GBSS splicing defects associated with the 5' leader intron sequence AGTTATA. Most importantly, intermediate levels of GBSS protein were seen in all of the varieties with the GCC allele of GBSS and intermediate amylose content. Thus the amount of amylose produced by different GBSS alleles can

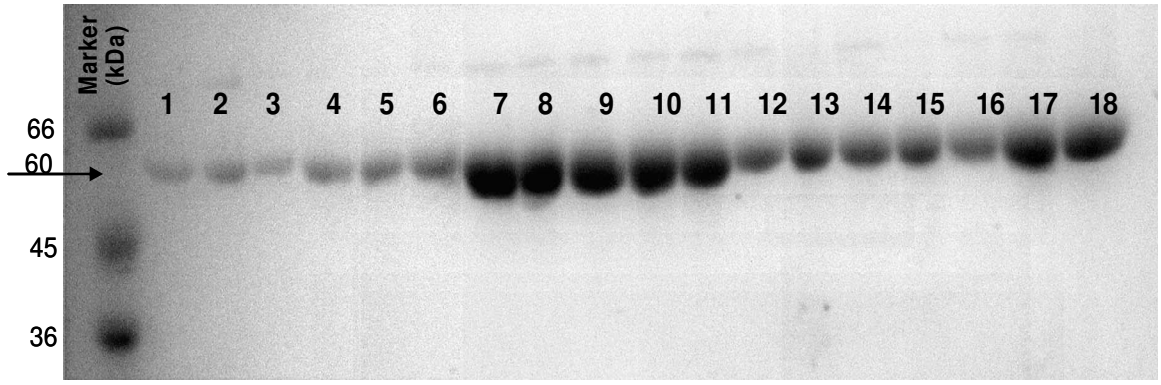

Figure 2.7. GBSS protein level in European collection. Total extractable protein from washed starch granules of varieties from the European germplasm collection was electrophoresed on SDS gels. Lane 1-6, varieties with TAC allele (low amylose); Lane 7-11, GAC allele (high amylose); Lane 12-16, GCC (intermediate amylose); Lane 17 & 18, GAT allele (high amylose). Also shown are molecular weight markers and an arrow representing the expected position of GBSS.

be explained based on differences in the amount of GBSS protein and does not require the existence of alleles with different specific activities.

To further examine the difference in GBSS specific activity between Wx_a and Wx_b postulated by Villareal and Juliano (1989), the varieties used in their study were obtained from the US germplasm collection. The pattern of GBSS SNPs in these seeds was then determined. The upper part of Figure 2.8 shows the data essentially as originally presented by Villareal and Juliano (1989). In the lower part of the figure, the same points are labeled based on GBSS alleles defined with the SNPs in exons 1, 6 and 10. None of the GBSS alleles defined based on the SNP in exons 1, 6 and 10 correspond directly to Wx_a and Wx_b identified by Villareal and Juliano (1989). The predominant allelic pattern, GCC, which differentiates intermediate amylose varieties, was found in a comparable ratio in samples labeled as "Wx_a" and "Wx_b" by Villareal and Juliano (1989). The TAC allele (low amylose), was also found in both classes. As also shown in the bottom panel, the data gives a reasonably good fit with a single regression line (R^2 = 0.689).

The re-analysis of the data of Villareal and Juliano (1989) should, however, be treated with caution since the samples obtained from the US germplasm collection may not actually correspond to those in used in the original study. This would not be surprising since multiple accessions of the same variety are often found to be different with molecular markers (Olufowote et al. 1997).

Irrespective of whether the samples used in the re-analysis actually correspond to those used by Villareal and Juliano (1989), analysis of the GBSS protein in these samples illustrates the large variation in GBSS protein that can be found in accessions with the GCC allele obtained from the US germplasm collection (Figure 2.9). Some varieties with GCC alleles have GBSS levels only slightly higher than those seen in TAC alleles, while others have levels only slightly lower than those seen in high amylose lines.

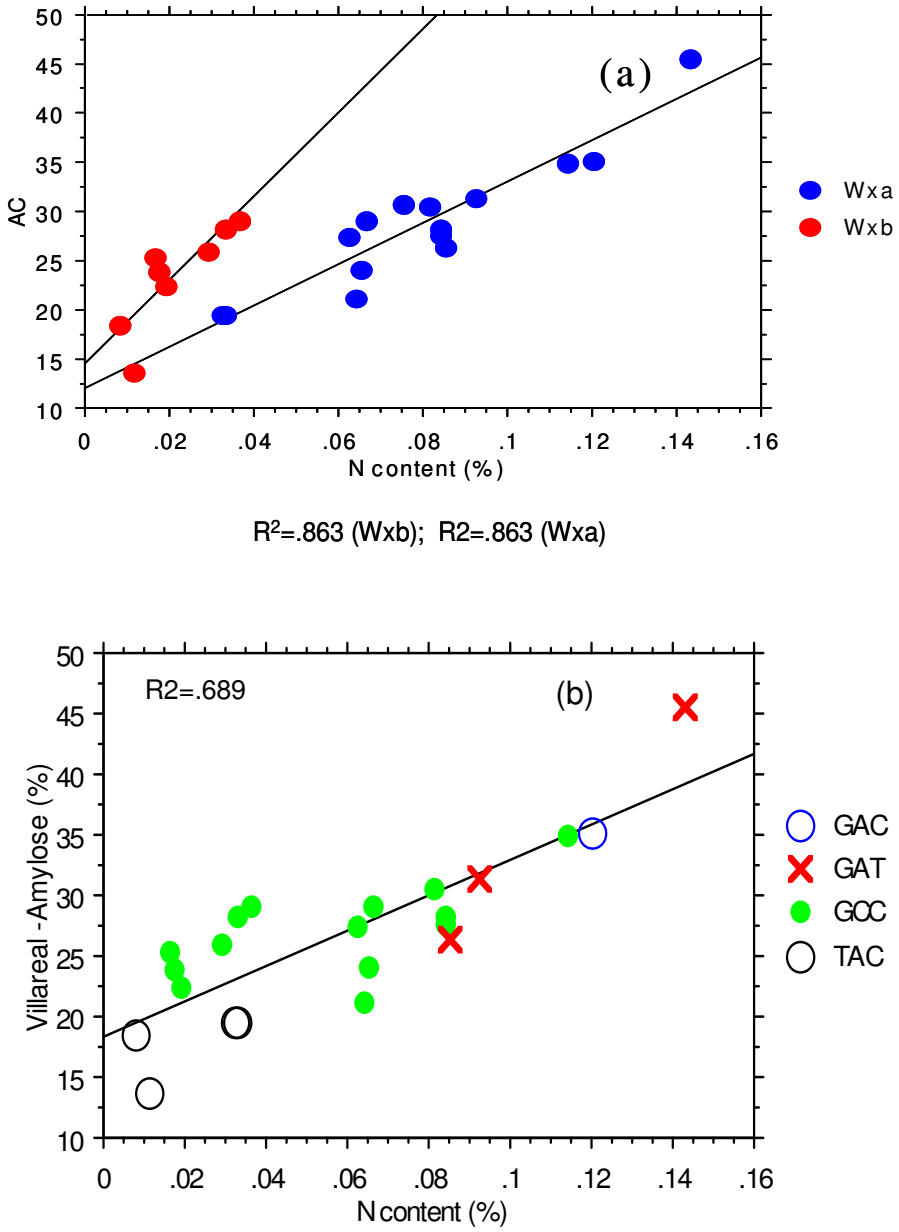

Figure 2.8 Relationship between Wx_a/Wx_b and GBSS allelic pattern. The upper panel was adapted from Villareal and Juliano (1989) and shows apparent amylose content vs. putative GBSS protein N content. The lower panel shows the GBSS SNP alleles for each sample and the fit to a single regression line.

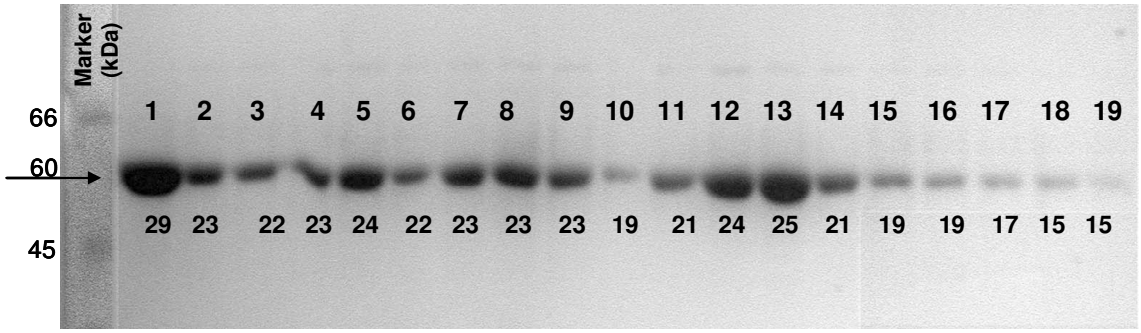

Figure 2.9. GBSS protein in varieties used by Villareal and Juliano (1989). Total extracted protein from putatively the same varieties used by Villareal and Juliano (1989) were electrophoresed on SDS gels. Lane 1, GAC allele (high amylose); Lanes 2-15, GCC allele (intermediate amylose); Lanes 16-19 TAC allele (low amylose). Also shown are molecular weight markers, an arrow representing the expected position of GBSS, and amylose number below each band.

Since Wx_a and Wx_b alleles are defined based on the amount of GBSS protein, it is easy to see how some GCC varieties might be classified as Wx_a and others classified as Wx_b.

Discussion

The CT repeat in exon 1 of GBSS has proven to be a very useful tool for identifying GBSS alleles and thus explaining differences in amylose content. Ayres et al. (1997) identified seven GBSS alleles and were able to explain 82.9% of the variation in apparent amylose in an extended pedigree of US germplasm. Similar successes have been obtained in many other studies. For example, the CT repeat explained 81% of the variation in apparent amylose in a collection of Chinese germplasm (Tang and Zhang 2001), 91% of the variation in Korean germplasm (Shu et al. 1999) and 88% of the variation among samples from the US Rice Regional Uniform Nursery (Bergman et al. 2001). Thus the explanation of 81.2% of the variation in apparent amylose content using the CT repeat observed in the current study with a European germplasm collection is consistent with previous work.

The CT repeat is very informative and is technically simple to assay, but it has limitations as a method for classifying GBSS alleles. As shown in Figure 2.5, for example, varieties with the same number of CT repeats sometimes have very different amylose levels. Furthermore, varieties with different numbers of CT repeats often have similar levels of amylose. This is to be expected since the CT repeat is simply a closely linked marker rather than the cause of variation in amylose content. Varieties with 17-19 CT repeats generally have low amylose content, while those with 10 or 11 CT repeats generally have high amylose content. However, there are many exceptions to the correlation between a low number of CT repeats and high amylose content. For example, many of the high amylose varieties in the European germplasm collection have 20 CT repeats, as do most of the intermediate amylose varieties in the US pedigree. Thus there is no evidence that the length of the CT repeat is functionally important in determining the amount of GBSS or amylose produced.

Identifying GBSS alleles based on the nucleic acid changes which directly affect GBSS mRNA processing or cause differences in amino acid sequence gives substantially better results. The present study shows that the combination of SNPs in exons 1, 6, and 10 can efficiently differentiate all three classes of apparent amylose: low, intermediate and high. This was seen both in 89 non-glutinous varieties from an extended pedigree of US rice germplasm that had been grown in Texas as well as in 279 non-glutinous varieties from a European germplasm collection that had been grown in northern Italy. The three SNPs explained 89.2% of the variation in the US collection and 93.8% of the variation in the European collection. This was a significant improvement over the results obtained using the CT repeats, which explained 82.9% and 81.2% of the variance respectively. In agreement with the idea that the CT repeat has no functional role in determining amylose content, combining CT repeat information with the three SNPs did not result in a significant increase in variance explained compared to the three SNPs alone.

These results demonstrate the utility of the three SNPs in defining different functional GBSS alleles. They also suggest that other factors such as modifier genes (Bollich and Webb 1973, McKenzie and Rutger 1983, Pooni et al. 1993) have only a minor effect on apparent amylose content in either the US or European germplasm.

The apparent amylose content in varieties with same GBSS allele was generally very similar between the US and European collection. Minor differences can likely be accounted by the fact that apparent amylose content in the European study was determined manually using I_2-KI whereas the apparent amylose values used for the US pedigree samples had been determined previously by the Rice Quality Lab at Beaumont, Texas using near infrared spectroscopy and a different set of standards.

However, there were substantial differences in the amylose content of the low amylose varieties in the US and European collections. For example, the average apparent

amylose content of varieties with 18 CT repeats was 14.9% for the US collection, but 18.6% in the European collection. This is not surprising since the germplasm collection was grown under cool conditions in Northern Italy compared to growth under the high temperatures of Beaumont, Texas. Low amylose varieties are known to be particularly sensitive to temperature during grain development, with lower temperatures leading to increased amylose content (Takeda 1988, Asaoka et al. 1985, Hirano & Sano 1998, Larkin et al. 1999). It should also be noted that TAC was the predominant GBSS allele among low amylose varieties in the European collection, while TCC predominated in the US collection. Since varieties with the GCC allele have lower amylose content than those with the GAC allele, differences in exon 6 might also contribute to the lower amylose content of the TCC varieties in the US collection compared to the TAC varieties that predominated in the European collection.

The reason for the low amylose content in varieties with exon 1 sequence AGTTATA is well understood and can be mainly attributed to inefficient GBSS mRNA processing, temperature dependent alternate splice site utilization and effects of a short premature open reading frame (Isshiki et al. 1998, Wang et al. 1995, Larkin 1999, Larkin and Park 1999).

In contrast, the reasons for the difference between intermediate and high amylose content were not well defined in previous work. Most previous work has focused on the reason for high vs. low amylose content rather than on intermediate amylose. While varieties with intermediate amylose are predominant among US long grains, they have traditionally been less important in Asia.

In the present study, all varieties with intermediate amylose in the germplasm collection were found to have the GCC allele of GBSS, while all high amylose varieties were found to have either the GAC or GAT allele. The C/A polymorphism in exon 6 causes a serine/tyrosine amino acid substitution. Thus the simplest interpretation of the data was

that this amino acid substitution changed either the amount or specific activity of the GBSS produced and thus resulted in reduced amylose production.

Direct analysis of washed starch granules from grain of the European germplasm collection clearly showed that intermediate amylose varieties have less GBSS on their starch granules than those of either GAC or GAT class of high amylose varieties. As one would expect based on defects in GBSS mRNA processing, starch granules from low amylose varieties contained substantially less GBSS than that seen in intermediate amylose varieties. Thus differences in the amount of GBSS on starch granules appear to provide a simple explanation for differences in the amount of amylose produced.

Tyrosine was reported to significantly contribute to protein stability through hydrogen bonding (Pace 2001). Therefore, the change of tyrosine to serine in intermediate amylose varieties may destabilize GBSS and lead to its degradation. This would be consistent with the relatively constant steady state level of GBSS mRNA, but reduced amount of GBSS protein and amylose seen in the intermediate amylose variety Lemont when it was grown at high temperature (Larkin 1999).

There was a claim in the literature (Villareal and Juliano 1989) for different in GBSS specific activity of Wx_a and Wx_b alleles. Thus it became of interest to determine the relationship between the classification of GBSS alleles as Wx_a and Wx_b compared to the classification based on the three SNPs.

The original classification of alleles as Wx_a or Wx_b was based on the amount of GBSS protein on washed starch granules (Sano 1984). Wx_b varieties predominant among the *japonica* subspecies of rice, generally having a low amylose content and the T allele of exon 1 (Sano 1984, Villareal and Juliano 1989, Cai et al. 1998, Isshiki et al. 1998, Larkin and Park 1999). Wx_a varieties predominate among the *indica* subspecies of rice and generally have high amylose content and the G allele of exon 1.

Applying this classification system to intermediate amylose varieties, however, is problematic. For example, as shown in Figure 2.1, the intermediate US varieties Labelle (LBLE) and Bluebonnet (BBNT) both have the GCC allele of GBSS. However, these varieties were scored by Villareal and Juliano (1989) as Wx_b based on amount of GBSS protein assay and were later classified as Wx_a by Isshiki et al. (1998) based on the G allele in exon 1.

To further examine this issue, the varieties used in the study by Villareal and Juliano which were claimed to show a difference in GBSS specific activity between Wx_a or Wx_b were re-examined. Most of the varieties with the T allele of exon 1 had low amylose content and were classified by Villareal and Juliano as Wx_b. All of the varieties the high amylose GBSS alleles GAC and GAT were classified by Villareal and Juliano as Wx_a. The intermediate amylose varieties, however, were split evenly between Wx_a and Wx_b in Villareal and Juliano's classification. Thus the putative difference in specific activity is not due to the SNP in exon 6. Whether there actually is a difference in specific activity is unclear. As shown in the bottom section of Figure 2.8, all of the data can be fit relatively well (R^2=0.689) with a single regression line, and thus without postulating differences in specific activity.

The reason it is sometimes difficult to classify varieties with the GCC allele is evident in Figure 2.9. The amount of GBSS in the starch granules of some of the intermediate amylose varieties was only slightly greater than that seen in low amylose varieties, while the level in others was only slightly less than that seen in high amylose varieties. This is sharp contrast to the very similar levels seen in varieties with the GCC allele from the European germplasm collection (compare Figures 2.7 and 2.9). That some varieties do not fit well within the traditional classification of Wx_a/Wx_b has been noted by others and led to the suggestion that an additional intermediate class was needed (Umemoto and Terashima 2002).

There are several possible explanations for the variation in GBSS levels in the GCC varieties in the European collection compared to the varieties examined by Villareal and Juliano (1989). The varieties examined by Villareal and Juliano (1989) may represent a wider range of germplasm and thus contain additional mutations not seen in the European samples. Another possibility is that the greater variability reflects differences in environmental conditions during grain development. All of the samples in the European collection were grown side-by-side under the cool conditions of Northern Italy. However, the samples previously examined by Villareal and Juliano were directly obtained from the US germplasm collection and may have developed under a variety of different conditions. The significant variation observed within the GCC in these samples may be due to high temperature stress, as suggested by Umemoto and Terashima (2002). This would be consistent with the reduced level of amylose and GBSS protein, but relatively constant levels of GBSS mRNA, seen when the GCC variety Lemont was grown at high temperatures (Larkin 1999). It should be noted that Larkin (1999) examined only a single intermediate amylose variety. However, a similar decline in the amylose content of other varieties with the GCC allele has also been observed and will be discussed in more detail in chapter V.

CHAPTER III

CONTROL OF AMYLOSE CONTENT IN AFRICAN GERMPLASM

Introduction

In the last decade the West African Rice Development Association (WARDA) has developed high yielding varieties, named 'Nerica" which stands for New Rice for Africa, by crossing *O. sativa* with the African species *O. glaberrima* (Jones et al. 1997). The Nerica varieties have succeeded in combining the high yielding potential of *O. sativa* with the biotic and abiotic stresses tolerance of *O. glaberrima* (Dingkuhn et al. 1998). As result, many farmers are able to produce enough rice to feed their families and have excess to sell on the market to generate cash for their other needs. However, as local and national markets continue to develop, it is likely that grain quality will become more important and Nerica varieties will be expected to meet increasingly stringent grain quality requirements.

As discussed in chapter II, amylose content and cooking quality of *O. sativa* appear to be primarily determined by three SNPs in GBSS exons 1, 6, and 10. *O. glaberrima* varieties generally have high amylose content and, as would be expected, give firm, separate grains when cooked (Juliano 1993). Thus far there are only two *glaberrima* GBSS sequences in public databases (Umeda et al. 1991), but both have the allelic pattern GAC, which is consistent with high amylose content. Also in agreement with results from *O. sativa*, it has been reported that amylose content in *O. sativa*/*O. glaberrima* progeny was largely due to their particular GBSS alleles. For example, the GBSS CT repeat explained 79% of the variation in amylose content in a population derived from *O. sativa*/*O. glaberrima* (Aluko et al. 2004). Regions with minor effects were also found on chromosomes 3 and 12 (Li et al. 2004).

In contrast, other data suggest that amylose content and grain quality may be regulated differently in *O. sativa/O. glaberrima* crosses. Watanabe et al. (2002) reported strong transgressive segregation in which progeny with as little as 15% amylose were found in a cross between an *O. sativa* cultivar with 20% amylose and an *O. glaberrima* accession with 26% amylose. This result is strikingly different from that seen in typical *O. sativa/O. sativa* crosses, but could possibly be explained by the segregation of functionally different forms of other enzymes involved in starch biosynthesis.

Moreover, a study done by Heuer and Miezan (2003) on an *O. sativa*/*O. glaberrima* population suggested that recombination occurred at high frequency in the 800 bp region between the GBSS promoter and structural gene. High levels of recombination within such a short region would be quite unusual, but could result in creation of novel GBSS alleles by shuffling SNPs in the parental GBSS alleles in *sativa/glaberrima* crosses. It should be noted that there are eight possible combinations of the SNPs in exons 1, 6, and 10, but only five of these combinations are seen in the US and European germplasm collection. Also, both of the *O. glaberrima* sequences in the database have an additional SNP in exon 12 (G→A), which is predicted to lead to a conservative amino acid change (asparagine to aspartic acid) (Umeda et al. 1991). Whether this SNP has any effect on cooking quality is unknown.

Determining how much of the variation in amylose content in *O. sativa/O. glaberrima* crosses is actually due to GBSS is very important. If amylose content and cooking quality are largely determined by GBSS, marker-assisted selection for GBSS can likely be used in the very near future as a cost effective tool to facilitate broader acceptance of higher yielding varieties across Africa and for development of local and regional rice markets. If the situation is more complex, as suggested by previous work from WARDA, marker-assisted selection will require further development. However, the other putative factors controlling cooking quality might prove to be very useful both for

development of new varieties for African as well as for production of novel high-value African rice products for the world market.

Materials and Methods

Experimental material consisted of 77 varieties, including 19 Nerica (progeny derived from *O. sativa/O. glaberrima*), 58 *O. sativa* and two *O. glaberrima* which were obtained from WARDA. In addition 48 *O. glaberrima* lines were obtained from the USDA-ARS National Small Grains Collection in Aberdeen, Idaho.

DNA extraction and analysis of the CT repeat and SNP in exon 1 of GBSS was done as described by Ayres et al. (1997). Apparent amylose content and the SNPs in exons 6 and 10 of GBSS were assayed as described in chapter II. The SNP in exon 12 was assayed using primers 5'-ATGATGATTTTCCTTGTTGA-3' and 5'-AAAACCATATAACCAGCAAC-3'. The reaction mixture (10uL) contained 0.2u*M* of each primer, 5 to 10 nanograms of DNA, 0.2m*M* deoxynucleotides, 2.5 m*M* $MgCl_2$, 10m*M* Tris-HCl, pH 9.0, 0.1% Triton X-100, 0.2u*M* and 0.25 units Taq polymerase. The reaction was denatured at 95ºC for 4 min, submitted to 35 cycles of 94ºC for 45 sec., 55ºC for 30 sec., 72ºC for 1 min, and a final extension of 72ºC for 5 min. The PCR product was mixed with 1 unit of the restriction enzyme SalI and incubated at 37ºC for 2 hr. Ten uL of the sample was then electrophoresed on a 1.5% agarose gel for 2h at 120 V (3V/cm) and was stained with ethidium bromide (0.01%). The *O. sativa* and *O. glaberrima* GBSS genes were also distinguished based on the lack of a 139 bp insertion in intron 10 of *O. glaberrima*. The intron 10 was assayed using PCR and primers pair 5'-CCATACTTCAAAGGAACTTC-3 and 5'-ATACTTCTCCTCCATGCTCT-3'.

Results

Examination of varieties from WARDA

To determine to what degree GBSS alleles determine amylose content in West African rice varieties, the CT repeat and SNPs were examined in 77 varieties from WARDA. This collection includes *O. sativa*, *O. glaberrima* and Nerica (*O. sativa/O. glaberrima*) varieties. As in the US and European germplasm collections, there is a clear relationship between GBSS alleles and apparent amylose content (Figure 3.1). Overall, the CT repeats explained 78.7% of the variation in apparent amylose content and the SNP in exons 1, 6, and 10 explained 93.3%.

For consistency, the three SNP nomenclature used in chapter II has been retained. However, it has been expanded to indicate whether varieties had the *O. sativa* or *O. glaberrima* type of GBSS as determined by the absence of the 139 bp insertion in intron 10 in *O. glaberrima* and the SNP in exon 12. In all cases, genes which lacked the insertion in intron 10 had the *O. glaberrima* type of exon 12.

As expected, all varieties with the AGTTATA form of exon 1 had low amylose content. Varieties with the allelic pattern TAC_sat had apparent amylose contents from 15.0 to 17.4% and the one variety with the TCC_sat allele had 14.6% apparent amylose. The average apparent amylose in varieties with the TAC_sat allele was substantially lower in the African samples (10.2%) than in the European germplasm collection (50%). This was expected since, as discussed above, the amylose content is known to decrease when grain develops at higher temperatures. All three low amylose varieties from WARDA also had 18 CT repeats, as do most of the low amylose varieties in the US and European collections.

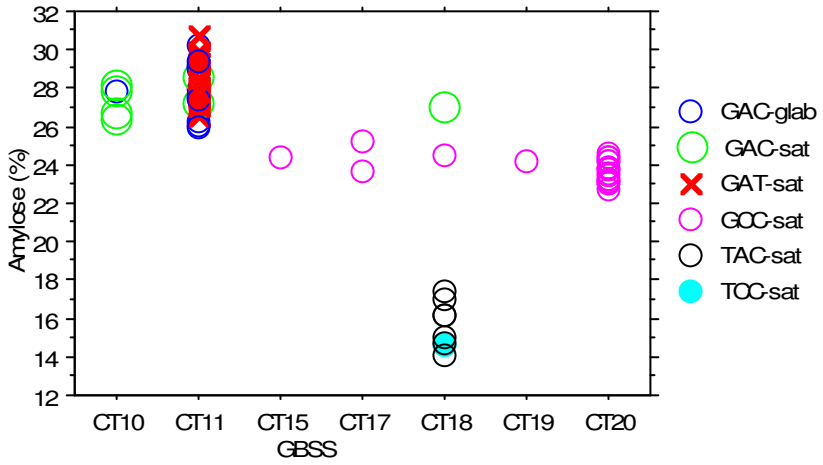

Figure 3.1. Amylose and GBSS alleles in African germplasm

As in the US and European samples, all varieties with the allelic pattern GCC_sat had intermediate apparent amylose content (23.0 to 24.5%) and most of them contained 20 CT repeats. However, GCC_sat genes with 15, 17, 18 and 19 CT repeats were also seen. Interestingly, none of the African samples had the GCC_sat allele with 14 CT repeats that was seen in both the US and Europe. The greater diversity in CT repeat length holds out the possibility that the African GCC_sat genes may also contain other types of potentially useful sequence diversity. However, there was no evidence for functional diversity based on apparent amylose content.

As expected, varieties with the GAC_sat and GAT_sat alleles had high levels of apparent amylose, (24.6-29.6% and 26.5%-30.0% respectively). As in the US and European collections, all of the varieties with the allelic pattern GAT_sat also had 11 CT repeats. Most of the high amylose varieties with the GAC_sat allele had 10 CT repeats rather than the 20 CT repeats seen most commonly in genes with this SNP pattern in the US and European collection. Interestingly, one of the African high amylose varieties with the SNP pattern GAC_sat had 18 CT repeats, a value more typically associated with low amylose content.

The African sample set also included 11 Nerica varieties which contain the *O. glaberrima* form of GBSS based on the lack of a transposable element in exon 10 (Umeda et al. 1991), as well as two of the parental *O. glaberrima* accessions. All of these had the GAC allelic pattern in exons 1, 6 and 10 and also contained the same SNP in exon 12 (G→A) that was observed in the two *O. glaberrima* GBSS sequences in the database (Umeda et al. 1991). The asparagine to aspartic acid change in exon 12 was not associated with a significant difference in apparent amylose content, 27.9% average apparent amylose for GAC_glab compared to 27.4% for GAC_sat (Table 3.1).

Table 3.1. Mean of amylose content of GBSS alleles in African germplasm

GBSS allele	**Amylose (%)**
GAC-gal	27.9a
GAT_sat	28.3a
GAC_sat	27.4a
GCC_sat	23.8b
TAC_sat	15.8c
TCC_sat	14.6c

[a]Values followed by the same letter are not significantly different (P<0.05).

Examination of "O. glaberrima" from the US germplasm collection

The WARDA collection discussed above only included two *O. glaberrima* accessions, CG14 and CG20. Furthermore, all of the Nerica lines which contained the *O. glaberrima* allele were derived from crosses with these accessions. To examine a wider range of *O. glaberrima* germplasm, 48 "pure *O. glaberrima"* accessions were obtained from the USDA-ARS National Small Grains Collection.

As expected, most of the "*O. glaberrima"* accessions from the US germplasm collection had high levels of apparent amylose and the GAC_glab form of GBSS as judged by the lack of the 139 insertion in exon 10 and the SNP in exon 12 (Figure 3.2). However, approximately 30% of the accessions had the *O. sativa* form of the GBSS as judged by both the presence of the 139 bp insertion and the SNP in exon 12. These varieties contained a wide range of apparent amylose levels and contained a diverse set of GBSS alleles based on the SNP in exons 1, 6 and 10. In all cases, however, the amylose level observed was consistent with expectations based on which GBSS allele was present.

The "*O. glaberrima*" accessions which had the *O. sativa* form of GBSS also had ligule and panicle architecture features characteristic of *O. sativa* rather than *O. glaberrima*. For example, *O. glaberrima* has a short ligule (Figure 3.3) and few secondary panicle branches compared to *O. sativa*. As shown in Figure 3.2, all of the accessions with the GAC_glab version of GBSS also have the short ligule (< 0.5 cm) characteristic of *O. glaberrima.*

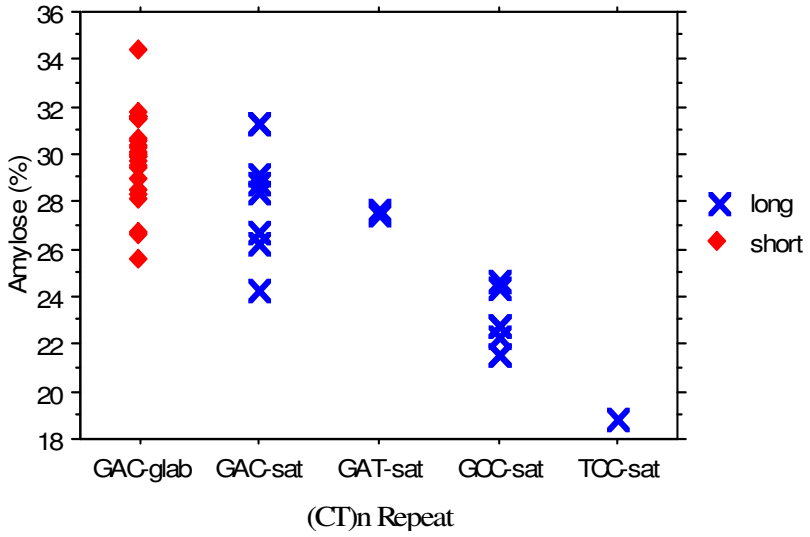

Figure 3.2. Amylose content and GBSS alleles in *O. glaberrima* varieties from US germplasm collection. Also shown is whether these accessions had long (>0.5cm) or short (< 0.5 cm) ligules.

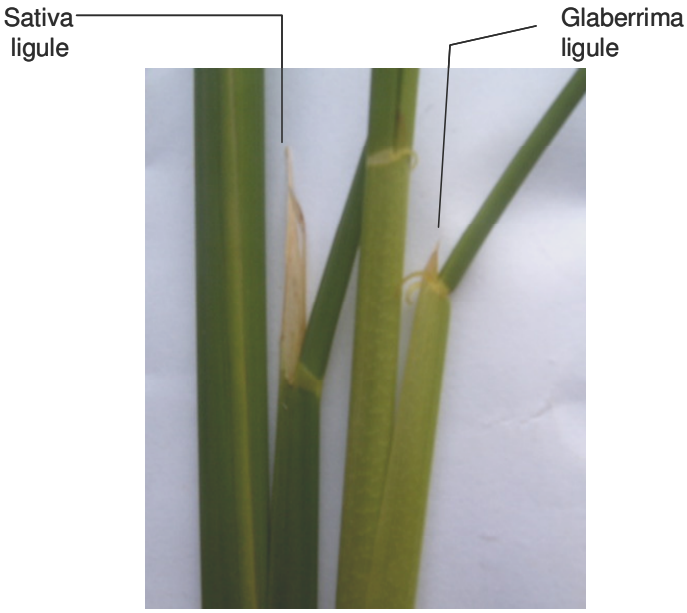

Figure 3.3. Ligule morphology of *O. sativa* and *O. glaberrima*

Putative low amylose progeny from O. sativa/O. glaberrima cross

Watanabe et al. (2002) reported progeny with low amylose content from a cross between an intermediate amylose *O. sativa* cultivar and a high *O. glaberrima* accession. To determine the origin of this material, the GBSS genes from one of the progeny with 15% apparent amylose and from both parents were analyzed.

This analysis showed that the low amylose "progeny" had the allelic pattern TCC_sat, rather than either the GCC_sat allele in the intermediate amylose parent or the GAC_glab allele in the high amylose parent. Thus, rather than reflecting a unusual genetics or a segregation of other genes which have a large effect on amylose content, the low amylose "progeny" observed by Watanabe et al. (2002) are most likely due to accidental outcrossing.

It should be emphasized that it is not unusual to find some level of non-parental alleles among the progeny of crosses. Other than deliberate crossing, the assumption is typically made in breeding programs that all seed are derived from self pollination. However, this is not always the case.

Discussion

The present study was conducted to determine whether amylose is regulated differently in African germplasm as suggested by previous studies (Watanabe et al. 2002, Heuer and Miezan 2003) rather than being largely determined by GBSS.

No evidence was found for unusual genetics or for other genes playing a dominant role in determining amylose content in African germplasm. On the contrary, the same three SNPs in exons 1, exon 6 and exon 10 of GBSS which explained 89% and 94%, respectively, of the variation in the US and European germplasm also explained >93% of the variation in apparent amylose content in the African germplasm. Moreover, the

apparent amylose level seen in African varieties containing a specific GBSS allele was comparable to that seen in US and European varieties with the same allele. The primary exception was the lower level of amylose seen in African varieties with the intron one 5' splice site sequence AGTTATA. However, lower amylose content in grain which developed at higher temperatures is to be expected due to the known temperature sensitivity of this GBSS mRNA splicing site.

Furthermore, the apparent transgressive segregation reported by Watanabe et al. (2002) in progeny derived from a cross between an intermediate amylose *O. sativa* and a high amylose *O. glaberrima* variety was not seen. Instead, GBSS analysis revealed that the low amylose progeny putatively due to transgressive segregation had as expected, TCC-sat allele, consistent with its low amylose content rather than either the GCC-sat or GAC-glab in the parent of the cross.

The primary difference found in African germplasm was the absence of a 139 bp insertion in intron 10 and a SNP in exon 12 in most *O. glaberrima* accessions. This work confirms the finding of Umeda et al. (1991) about the difference between the waxy genes in *O. sativa* and *O. glaberrima*. The GBSS genes from approximately 30% of the "*O. glaberrima*" accessions obtained from the USDA-ARS Small Grains Collection lack these features. However these accessions also lacked the short ligule and other morphological characteristics of *O. glaberrima*.

The SNP in exon 12 would lead to the replacement of asparagine with aspartic acid. However this conservative amino acid change does not appear to change the apparent amylose content. Thus as in the US and European germplasm collections discussed in chapter II, all low amylose varieties have the AGTTATA form of exon 1, all intermediate amylose varieties have the GCC allele, and all high amylose varieties have either the GAC or GAT allele.

Selection for GBSS alleles should thus provide a rapid, cost effective way for early generation seedling screens to breed for specific grain quality types in Africa. It should be particularly effective when coupled with anther culture technology to create high yielding Nerica varieties targeted both for specific African domestic markets and potentially for high value export markets.

CHAPTER IV

GBSS ALLELES AND TEMPERATURE/VISCOSITY PROFILES

Introduction

As emphasized in previous chapters, amylose content is generally the most important factor determining rice grain quality. However, varieties with the same amylose content sometimes have different cooking and processing quality. To deal with this issue, several methods have been developed to look at the functional properties of rice. The most common secondary criteria for rice quality is gelatinization temperature which can be either measured directly or estimated based on alkali spreading value (Little et al. 1958, Bhattacharya et Sowbhagya 1972, 1980, Halick and Kelly 1959, Halick et al. 1960).

While gelatinization temperature is clearly important, it is only one of the many factors that control the cooking and processing quality of rice. Thus, rice varieties with similar amylose contents and gelatinization temperatures can still have very different grain quality. For example, Jojutla, Newrex, and Rexmont are members of a family of high amylose varieties that have low solids loss and superior retention of grain integrity, size, and shape after cooking (Bollich et al. 1980, Bollich et al. 1990, Ayres et al. 1997, McClung et al. 1998, Larkin and Park 2003, Larkin et al. 2003). Varieties with "Rexmont quality" are superior for the manufacture of certain types of rice products, such as canned, cooked rice (Webb 1991). However other rice varieties with similar amylose contents, gel temperatures, and geometry such as L202 and Jodon do not have these favorable properties (Bollich et al. 1990, Larkin and Park 2003, Larkin et al. 2003).

To examine a wider range of factors controlling cooking and processing quality, methods have been developed to continuously measure the viscosity, or "pasting properties", of rice flour during a simulated cooking cycle (Mazurs et al. 1957, Juliano et

al. 1964, Shu et al. 1998, Bao et al. 1999, Bao and Xia 1999). This was initially performed using the Brabender Viscoamylograph, but is currently more commonly done using the Rapid Visco™ Analyzer (RVA) because of faster assay times and the smaller sample size. In a typical experiment, an accurately weighed aliquot of rice flour is mixed with a set amount of water and gelatinized by gradual heating to 95°C. The gelatinized sample is then held at 95°C for a set period and gradually cooled to 50°C, all with constant stirring. The typical parameters measured are the "peak viscosity" (PV), "hot paste viscosity" (HPV), and "cool paste viscosity" (CPV) (Figure 4.1). These parameters are used to compute "breakdown" (BD) which is the PV – HPV, "setback" which is PV–CPV and "consistency viscosity" (CV) which is HPV-CPV.

Analysis of pasting properties can often distinguish varieties which have different cooking quality. For example, Newrex and Rexmont were distinguished from typical US long grain rice varieties based on low BD and high SB (Bollich et al. 1980, Bollich et al. 1990). BD is generally negatively correlated with starch granule stability and integrity, while SB is generally considered an indicator of starch reassociation or retrogradation (Juliano et al. 1964, Mazurs et al. 1957). Thus the pasting properties of these varieties are consistent with enhanced starch granule stability and with their low solids loss during processing and high amylose content.

In some cases, good correlations have been seen between pasting properties and important sensory properties of cooked rice, such as texture. For example, in a study of Chinese varieties, Shu et al. (1998) showed that the hardness of cooked rice was positively correlated with SB and negatively correlated with BD. Negative correlations between SB and palatability score were also seen in Taiwanese varieties by Kuo et al. (2001). Several investigators have found a close association between amylographic characteristics and eating quality (Perez and Juliano 1979, Chang and Lii 1985, Lii et al. 1986, Ohtsubo et al. 1996). However other investigators have found only weak

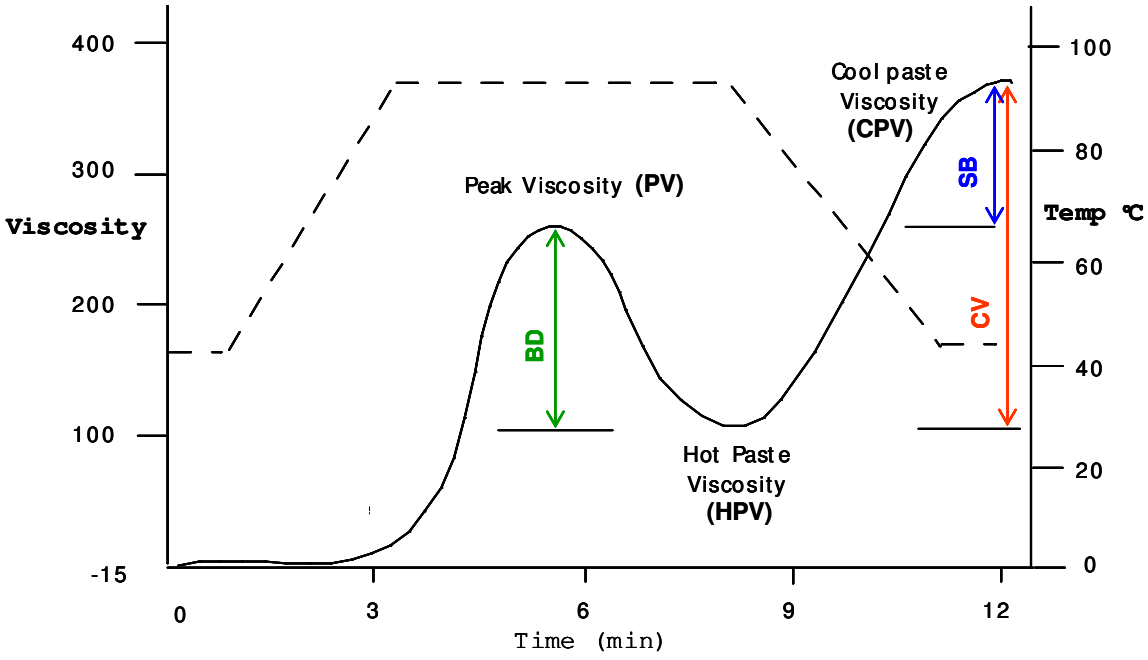

Figure 4.1. Rapid Viscoamylograph (RVA) starch pasting curve. Solid line represents change in viscosity.

Dashed line temperature. BD=Breakdown; SB=Setback; CV= Consistency Viscosity.

correlations between the pasting properties of rice and human sensory evaluation of rice cooking quality (e.g. Champagne et al. 1999).

There have been a number of studies of genes and chromosomal regions which control pasting properties (Bao et al. 1999, Bao et al. 2000, Bao et al. 2004, Gravois and Weeb 1997, Larkin and Park 2003). As one might expect, many different genes are involved in the various aspects of rice temperature/viscosity relationships. For example, Bao et al. (2004) found 20 different chromosomal regions involved in controlling pasting properties in a cross between the Chinese rice varieties Zhai-Ye-Qing8 and Jing-Xi 17. However, the primary factor controlling pasting properties has generally been found to be GBSS (He et al. 1999, Tan et al. 1999, Bao et al. 1999, Larkin et al. 2003, Bao et al. 2004, Bao et al. 2000). For example, all of the varieties in the US rice pedigree which have "Rexmont quality" have the GAT allele of GBSS, while high amylose varieties with the GAC allele do not have these favorable processing qualities (Larkin 1999, Larkin and Park 2003, Larkin et al. 2003, McClung et al. 2004).

As noted above, the germplasm base of the US rice pedigree is very narrow. Thus the goal of the current study was to determine to what degree GBSS alleles could explain the pasting properties in the European germplasm collection.

Materials and Methods

The same European germplasm collection grown in Northern Italy in 2004 that was used for the amylose analysis discussed in chapter II was also used in this study. Milled grain were ground to powder with a Cyclotec 1093 sample mill (Foss Tecator, Eden Prairie, MN). The resulting flour was then defatted by Soxhlet extraction overnight with 85% methanol. The defatted flour was allowed to dry at room temperature and then re-ground in the Cyclotec mill to remove aggregates.

A Rapid Visco™ Analyzer model 3D (RVA) (Newport Scientific, Warriewood, Australia) was used to determine starch pasting properties. Three grams of defatted rice flour, calculated based on 14% moisture content, were mixed with 25 g of accurately weighed water. Two different controlled heating and cooling cycles were used. Initial experiments used the Rice1 profile that was pre-programmed in the RVA. In these experiments, the mixture was held at 50°C for 1 min, heated to 95°C in 3.8 minutes, held at 95°C for 2.5 minutes, cooled to 50°C in 3.8 minutes and then held at 50°C for 1 min. After 10 seconds of rapid stirring to suspend the rice flour, the samples were stirred at 160 rpm for the remainder of the experiment. Two replicates per sample were used. In later experiments, the holding period at 95°C was extended to 10 minutes. This is referred to as the "Rice1-ext" profile.

Gel hardness was measured on a modification of the method of Beta and Corke (2001) and Bao et al. (2004). Immediately after completion of RVA analysis, the stirring paddle was removed and the aluminum canister was covered with Parafilm®. After the samples had been stored at room temperature for 24 hours, gel hardness was measured by back extrusion using a SMS Model TA-XT2i Texture Analyzer as the steady state force required to move a 3.4 mm piston at 0.2 mm/sec. Typically, force was recorded after the piston had moved 8 mm, but similar results were obtained at 6-10 mm of piston displacement.

Statistical analyses were performed with StatView™ software (v 5.0) from Abacus Concepts. One way ANOVA (analysis of variance) and Fifer's Protected Least Significant Difference were done to define the significance of the RVA parameters.

Results

Standard pasting parameters

All of the pasting parameters shown in Figure 4.1 were initially determined in the each of the 279 European germplasm samples using the standard Rice1 profile. As shown in Figure 4.2, there was substantial variation for each of the pasting parameters between varieties that have the same GBSS allele. In most cases, there was also substantial overlap in the range of values obtained for varieties with different GBSS alleles. However, particularly for HPV, CPV, CV and SB, there is a clear relationship with GBSS alleles. As shown in Table 4.1, GBSS alleles explained 68.5%, 73.1%, 54.4% and 59.7% of these parameters respectively. The ability of RVA to distinguish between varieties with the GAT and GAC allele is particularly noteworthy since they have similar amylose contents but different functional properties, as will be discussed further below.

While GBSS clearly is an important determinant of pasting properties, most of the variation was associated with the GAT allele. As shown in Figure 4.1, there is substantial overlap in the pasting properties of varieties which have the low amylose allele, TAC, the intermediate amylose allele, GCC, and the high amylose allele, GAC. When the GAT allele is excluded, the GBSS parameters generally explain less than 25% of the variation of the RVA parameters (Table 4.1). The two exceptions are HPV and CV which still only explain 32.4% and 27.6% of the variation in amylose respectively. Thus other than the ability to discriminate the GAT allele, there is not a particularly strong relationship between most pasting parameters and GBSS alleles.

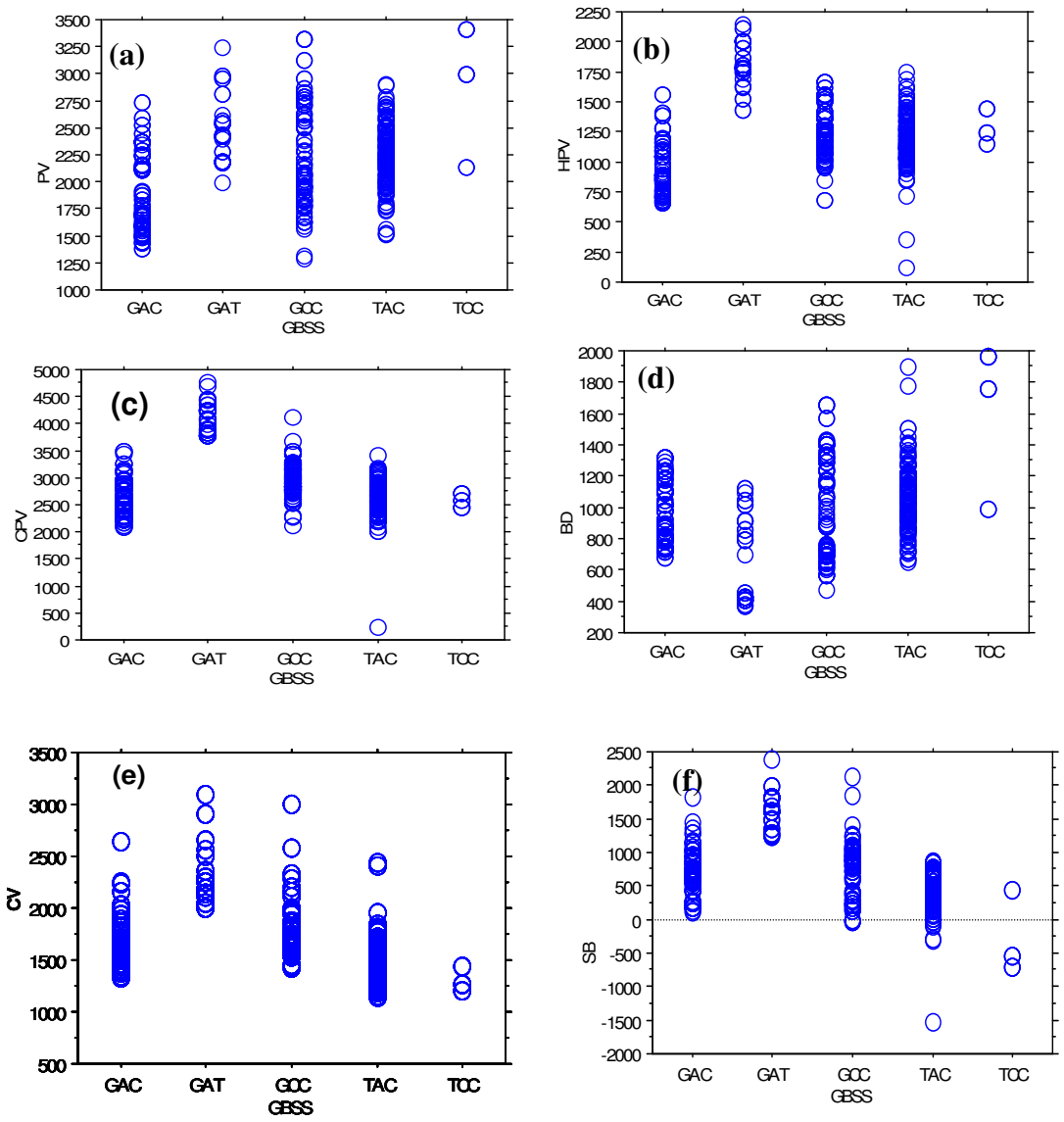

Figure 4.2. RVA parameters and GBSS alleles in European collection. PV=peak viscosity, CPV=cool paste viscosity, HPV=hot paste viscosity, BD=breakdown, SB= setback, CV=consistency viscosity.

Table 4.1. R^2 (in %) of pasting properties and GBSS alleles

Parameters	All GBSS alleles	Omitting GAT allele
PV	17.5 **	13.6**
HPV	68.5***	32.4***
CPV	73.1***	20.5***
BD	32.1***	3.7 ns
SB	59.7***	22.3***
CV	54.4***	27.6***

** and *** indicate significance at 0.01 and 0.001 level, respectively.

A strong relationship between the GAT allele and pasting properties was expected based on previous work on the US rice pedigree (Ayres et al. 1997, Larkin 1999, Larkin and Park 2003, Larkin et al. 2003). As noted above, Jojutla, Newrex, and Rexmont were initially distinguished from standard US long grain varieties based on low BD and high SB. The low BD of these varieties is consistent with greater starch granule stability and with their low solids loss during processing. Jojutla, Newrex and Rexmont all contain the GAT allele of GBSS. Thus one might expect that all of the varieties in the European germplasm collection with the GAT allele would also have more stable starch granules and low BD.

As will be discussed below, starch granules from varieties in the European germplasm collection which have the GAT allele are, in fact, substantially more resistant to shear. However, the relationship between GBSS alleles and granule strength was not immediately obvious from standard RVA data.

As expected, some cultivars with the GAT allele in the European germplasm collection have notably low BD (Figure 4.2). However, many varieties with the GAT allele do not have low BD. Furthermore, varieties in the European collection with high SB do not necessarily have low BD (Figure 4.3). This is inconsistent with the reported behavior of the US varieties Jojutla, Newrex and Rexmont, which had both high SB and low BD (Bollich et al. 1980, Bollich et al. 1990).

The more recent “Rexmont quality” varieties Dixiebelle and Bolivar have been reported to have significantly higher HPV and CPV than standard US cultivars (McClung et al. 1998, McClung et al. 2004). These cultivars also both contain GAT alleles. Dixiebelle, in fact, has exactly the same GAT allele that was inherited from Jojutla that is also found in Newrex and Rexmont.

Like Dixiebelle and Bolivar, most of the samples in the European collection with the GAT allele have high HPV and CPV (Figure 4.2). However, they do not necessarily have low BD or high SB (Figure 4.3). Thus from these results it is not at all clear whether the GAT allele is generally associated with more stable starch granules or with superior processing quality.

At this point, it was productive to stop and think about the factors which control breakdown in the RVA. During RVA analysis, samples are stirred at a constant rate of 160 rpm irrespective of their viscosity. If a sample has high peak viscosity, starch granules are thus exposed to high shear forces (Figure 4.4). Even if starch granules are very stable, they may disintegrate because of the high shear forces. If a sample has low peak viscosity, however, starch granules are exposed to only weak shear forces. Thus even relatively fragile starch granules may remain intact.

To solve this problem, a method was needed to look at BD as a function of the maximum shear force to which starch granules are exposed. There are several ways that this could be done, but the most straightforward was to simply plot BD as a function of PV. As shown in Figure 4.5, this works surprisingly well, particularly when the holding period at 95^{o}C is extended from 2.5 to 10 minutes to allow BD to approach equilibrium

As shown most clearly on the lower panel of Figure 4.5, BD is higher in varieties with high PV, irrespective of which GBSS allele they contain. This was expected since starch granules are more likely to disintegrate when exposed to higher shear forces. However, varieties with the GAT allele had notably higher shear resistance, i.e. lower BD at a given PV. For individual GBSS alleles, the proportion of variation in BD explained by PV ranged from 0.88 to 0.99 when the longer holding period was used.

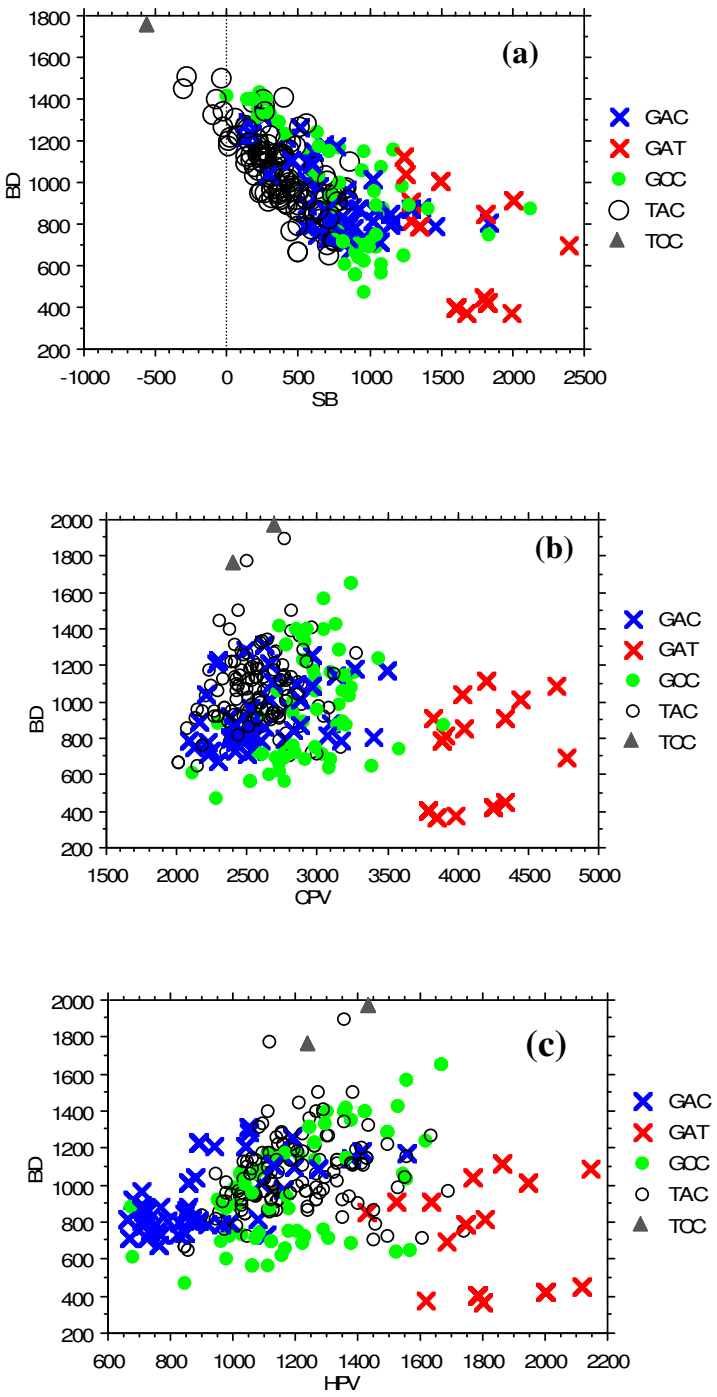

Figure 4.3. Relationship between BD and other pasting properties (SB, CVP and HPV)

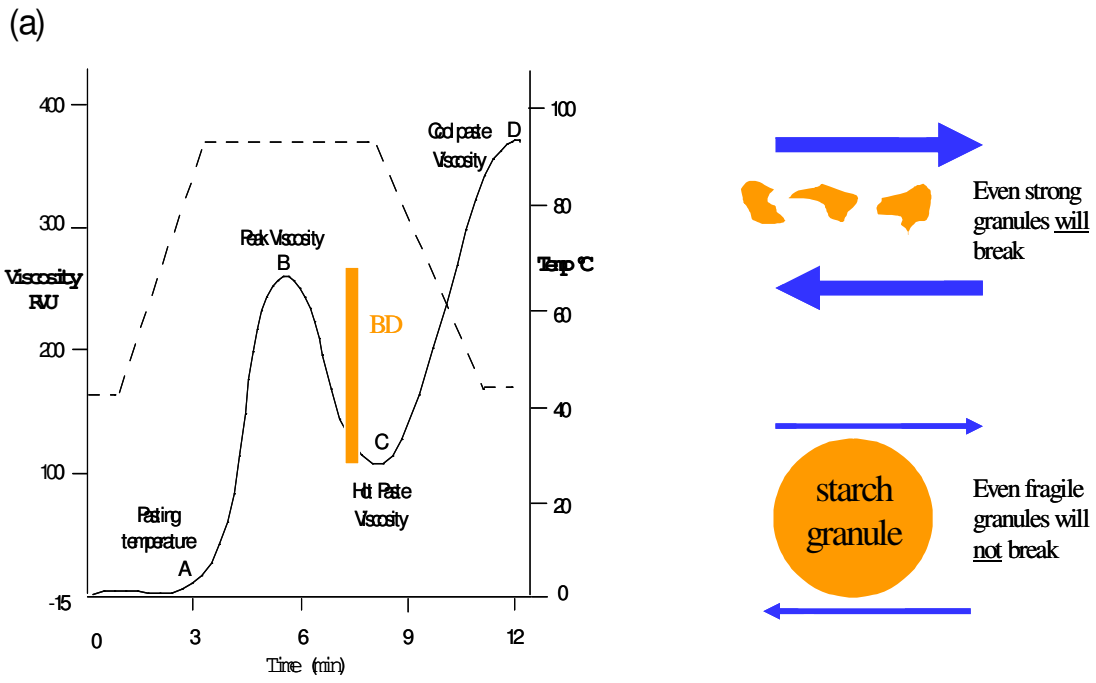

Figure 4.4 Relationship between shear force and breakdown. The amount of the breakdown depends on the strength of the shear forces generated when samples of different peak viscosity are stirred at 160 rpm while being hold at 95°C.

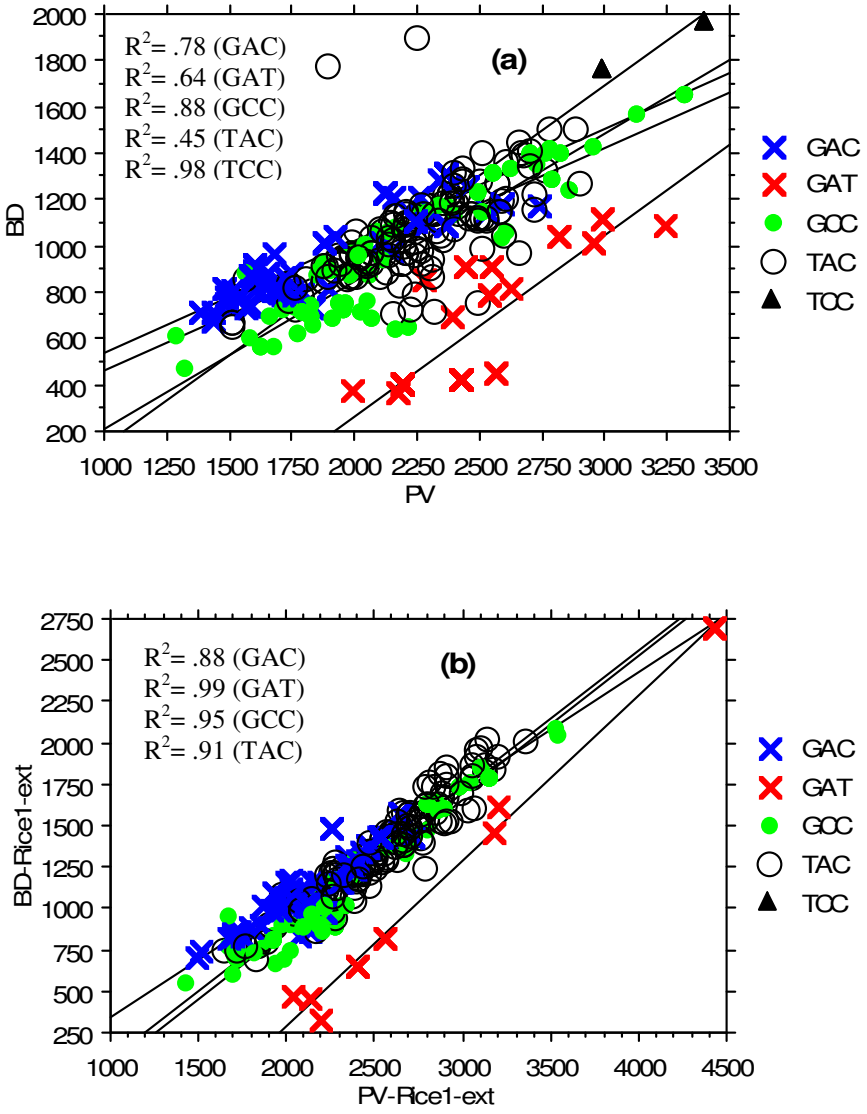

Figure 4.5. Breakdown (BD) vs. Peak viscosity (PV). The upper panel shows the results for the European germplasm collection obtained using the standard Rice1 RVA profile. Results obtained when the holding period at 95°C was extended from 2.5 to 10 min. The Rice1-ext profile is shown in the lower panel.

Varieties with the GAT allele were clearly more shear resistant (Figure 4.5b). However, there was surprisingly little variation in the shear resistance of varieties with other GBSS alleles. In fact, when varieties with the GAT allele were excluded from the analysis, a single regression line explained 89.4% of the variation in breakdown vs. peak viscosity ($p < 0.0001$). Thus the single amino acid change in exon 10 appeared to largely explain differences in shear resistance in the entire European germplasm collection.

Most of the variation (59.7%) in SB was explained by the GBSS alleles even under standard RVA conditions (Table 4.1). Plotting setback vs. peak viscosity further improved resolution of the varieties with the GAT allele (Figure 4.6)

Similarly, there was already a good relationship between the GAT allele and high HPV (R^2=0.685) and CPV (R^2=0.731) even with standard pasting parameters (Figure 4.2 & Table 4.1). As noted in chapter II, there was also a clear relationship with high amylose content. When these definitions are combined, however, the GAT allele again become even more clearly distinguished (Figure 4.7).

Thus apparent contradictions of the various definitions of "Rexmont quality" that are shown in Figure 4.3 can be reconciled. Not only do all the varieties with the GBSS GAT allele in the European germplasm collection have higher shear resistance, they also consistently have high SB, high HPV, high CPV and high amylose. While it remains to be directly proven, all of these varieties are thus also likely to have the low solids loss and superior retention of grain integrity after processing that is seen in US varieties with the GAT allele.

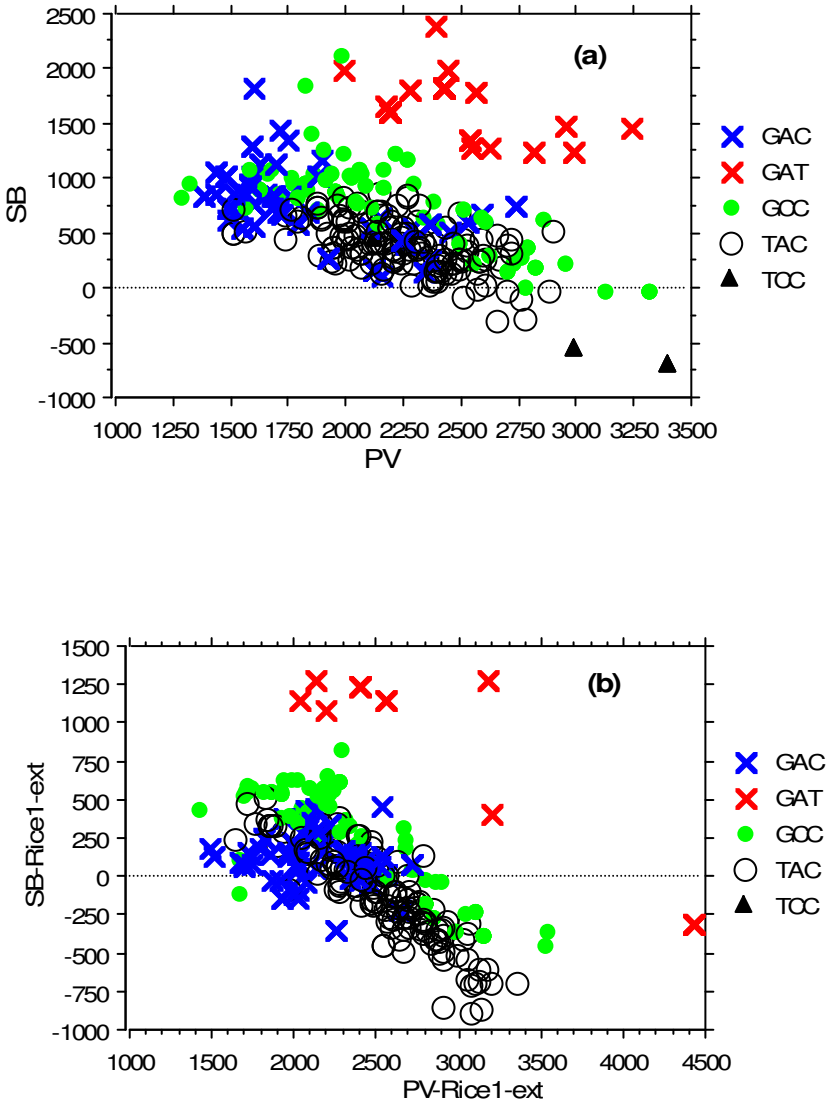

Figure 4.6. Setback (SB) vs. Peak viscosity (PV). Samples from the European germplasm collection were analyzed using both the standard Rice1 RVA profile (top panel) or the Rice1-ext profile in which the holding time at 95°C was increased to 10 min (bottom panel)

GBSS Alleles and Gel Hardness

To look at the relationship between GBSS alleles and starch reassociation to form gels, a modification of the "gel hardness" assay of Bao et al. (2004, 2006) was used. In this assay, the same samples used for RVA analysis were covered and stored at room temperature for 24 hours. The viscosity of the resulting starch gels was then measured by back extrusion.

A strong relationship (R^2=0.81) was seen between GBSS alleles and gel hardness (Figure 4.8). The hardest gels were formed by varieties which contain the high amylose GAT allele (mean = 2.11), while varieties with the low amylose TAC allele had the lowest values (mean = 0.91). Amylose content accounted for 71.5% in the variation of gel hardness (Figure 4.9). However, there was a large degree of overlap in the gel hardness values for varieties containing the GCC and GAC alleles that can not readily be explained by apparent amylose content (Figure 4.9).

In contrast, only weak correlations were observed between gel hardness and standard RVA parameters (Table 4.2). For example, CPV explained less than 18% of the total variance in gel hardness. A relatively low correlation between the final RVA viscosity and gel hardness might initially seem surprising since the gel hardness measurements were made on exactly the same samples, but 24 hours apart. However, very little actual gel formation occurs in the RVA since the sample is continuously being stirred at 160 rpm and is always at least 50^{o}C. Most of the increase in viscosity that is observed when RVA samples are cooled from 95^{o}C to 50^{o}C is actually due to the increase in viscosity of the water in the sample. Water has a viscosity of 0.287 cp at 95^{o}C and 0.547 cP at 50^{o}C. Thus even if there is no starch reassociation, the viscosity of the sample will almost double when it is cooled to 50^{o}C.

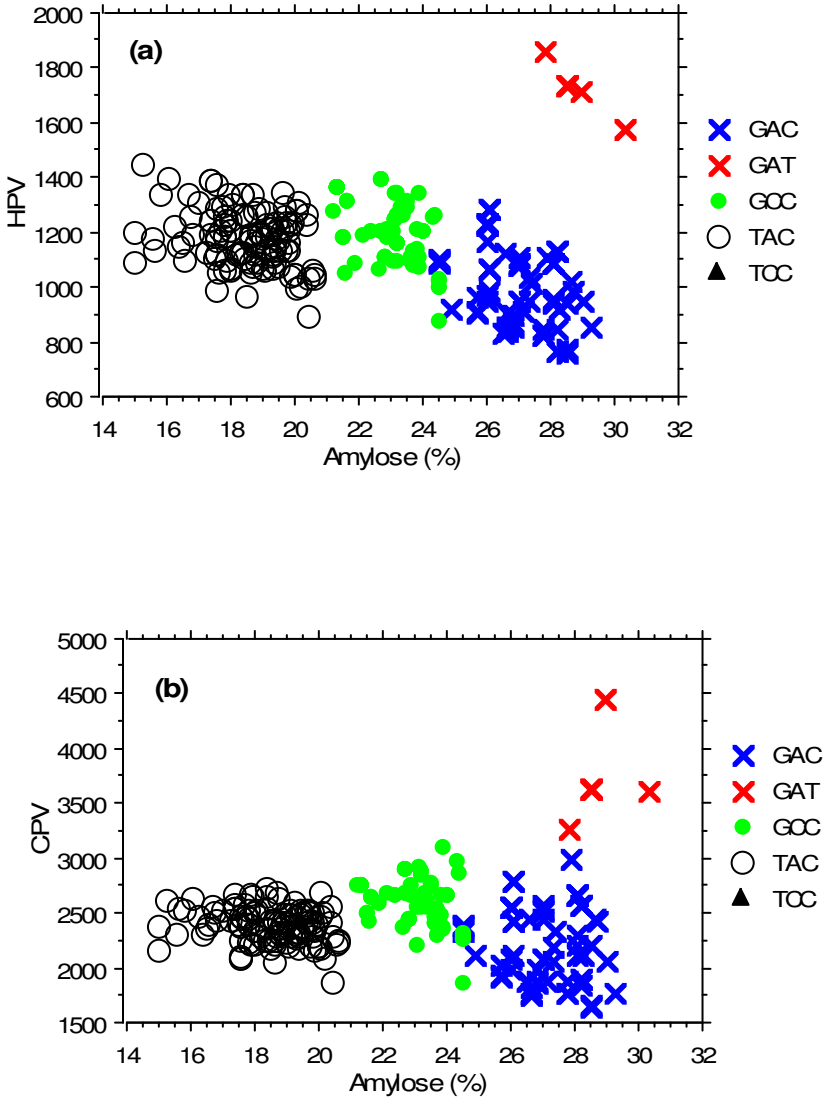

Figure 4.7 Relationship between two RVA parameters (HPV and CPV) and amylose. Panel A, Hot paste viscosity (HPV) vs. Amylose; Panel B, Cool Paste viscosity (CPV) vs. amylose.

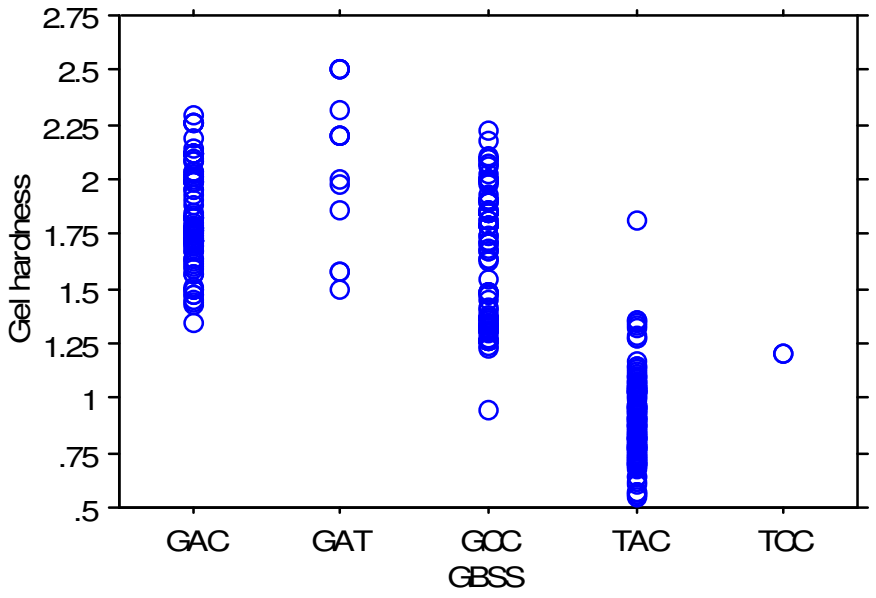

Figure 4.8. GBSS alleles and gel hardness

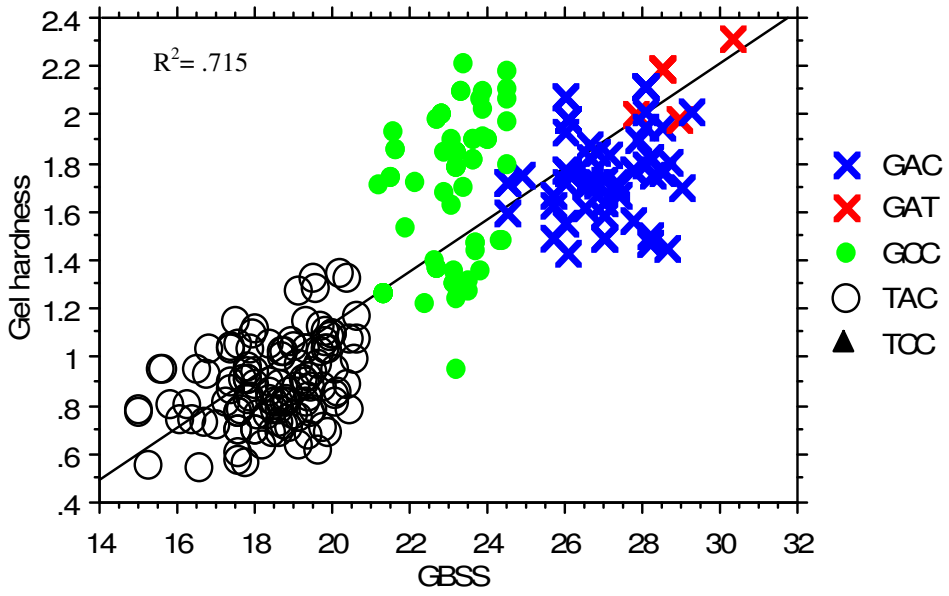

Figure 4.9. Gel hardness vs. amylose content and GBSS alleles

Table 4.2. Variation in the gel hardness explained by RVA parameters

RVA Parameters	**Gel Hardness (%)**
PV	4
HPV	0.3
CPV	17.9
BD	4.3
SB	22
CV	17.9

The importance of accounting for changes in water viscosity is illustrated by comparing BD and CV based on simple viscosity and when using specific viscosity (specific viscosity = sample viscosity/viscosity of water). Based on simple viscosity (Figure 4.10), CV appears comparable to BD, i.e. the apparent increase in viscosity upon cooling from 95°C to 50°C is roughly comparable to the decrease in viscosity that occurred while the sample was being held at 95°C. However, when the same data is expressed in terms of specific viscosity to account for changes in the viscosity of water, it becomes clear that the actual increase in viscosity upon cooling is quite small.

Accounting for changes in the viscosity of water improves the correlations between RVA parameters and gel hardness. For example, the proportion of variance in gel hardness explained by CV increases from 17.9 to 38.5% when specific viscosity was used. This is, however, still substantially less than the 71.5% variance explained by amylose content or the 81% variance explained by GBSS alleles.

Accounting for changes in the viscosity of water also provides insight into what is actually being measured by SB and why it was strongly correlated with BD, e.g. R^2 = 0.66 in Figure 4.3. SB is typically defined as CPV-HPV. However, this is equivalent to SB = CV-BD (Figure 4.1). In other words, SB compares the decrease in viscosity that occurs when samples are stirred continuously for a given period at 95°C (BD) with the apparent increase in viscosity when the sample is subsequently cooled to 50°C (CV).

As illustrated in Figure 4.10, CV and BD are roughly comparable when expressed based on simple viscosity. Thus the difference between these two values (SB) appears to provide a very different parameter than either component alone. However as discussed above, most of the change in viscosity upon cooling to 50°C is actually due to changes in the viscosity of water. Thus, while SB = CV – BD, the amount of starch

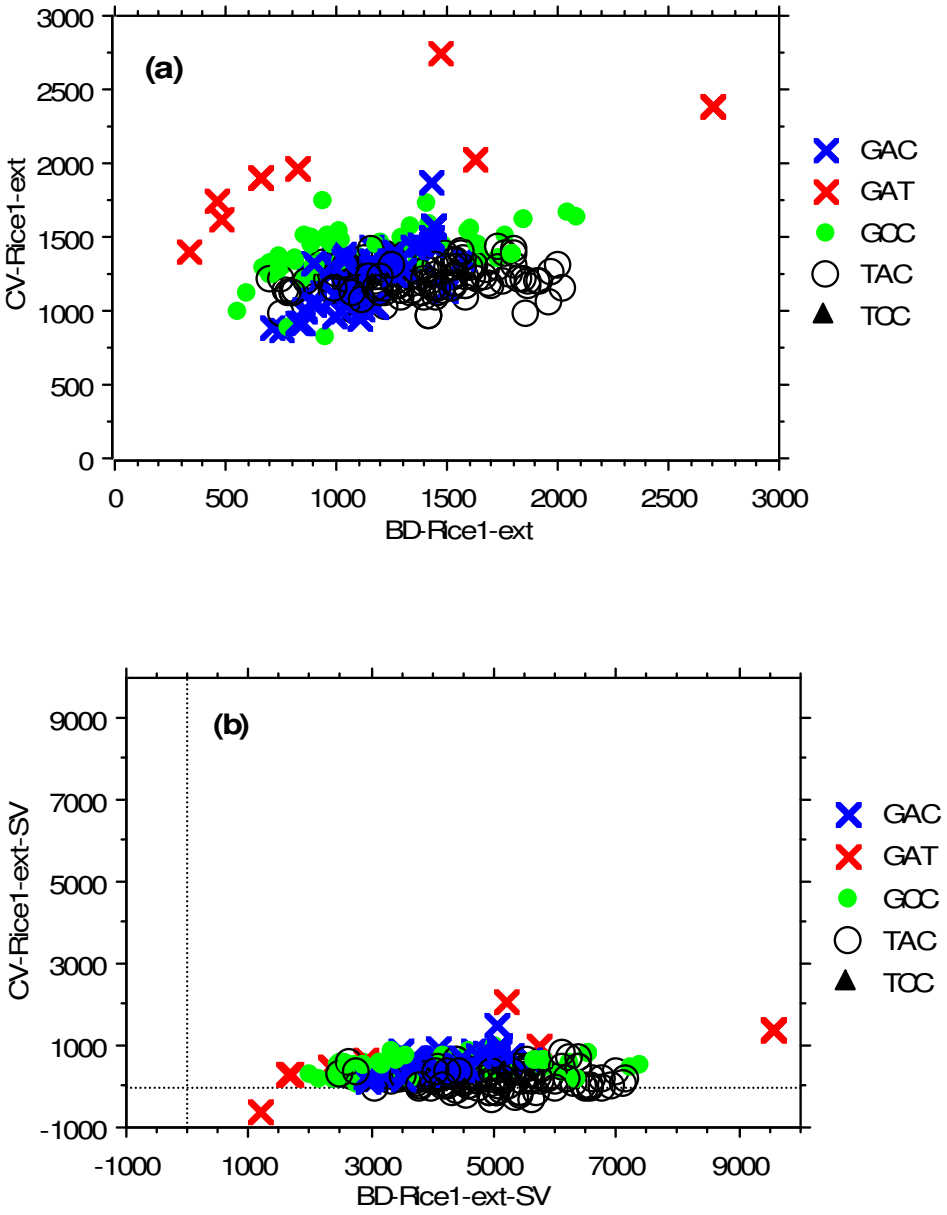

Figure 4.10. Consistency viscosity (CV) vs. Breakdown (BD). Results from the European collection using sample viscosity taken directly from the RVA (Panel A) or specific viscosity (Panel B) calculated as sample viscosity/viscosity of water at that temperature.

reassociation (CV) is actually very small (Figure 4.10). Other than differences in water viscosity, SB ≈ -BD. The very tight negative correlation (R^2 =0.945) seen when SB and BD are expressed based on specific viscosity is shown in Figure 4.11.

Discussion

Analysis of the European germplasm collection demonstrated two key roles for GBSS in determining the pasting properties of rice. First, as most clearly illustrated in the lower panel of Figure 4.5, varieties which contain the GAT allele of GBSS were distinctly more resistant to shear and thus had lower BD. These results are consistent with previous work demonstrating that varieties such as Jojutla, Newrex and Rexmont, which contain the GAT allele (Larkin 1999, Larkin and Park 2003), could be distinguished from standard varieties based on low BD (Bollich et al. 1980, Bollich et al. 1990). The relative shear resistance of varieties with all of the other GBSS alleles was surprisingly similar. Thus the single amino acid change serine/proline in exon 10 with GAT varieties appears to largely determine relative shear resistance.

The reason why the GAT allele of GBSS is associated with enhanced shear resistance is not completely clear. However, it may be related to the role of GBSS in the synthesis of long amylopectin chains (Denyer et al. 1996). GBSS null, *wx,* varieties lack the very long amylopectin chains seen in other varieties (Hizukuri et al. 1989). However, high amylose *indica* varieties contain a higher proportion (>14%) of long amylopectin chains. Long amylopectin chains along with amylose retard or reduce starch granule swelling and thereby maintain the integrity of swollen starch granules in contrast to short chains (Srichuwong et al. 2005).

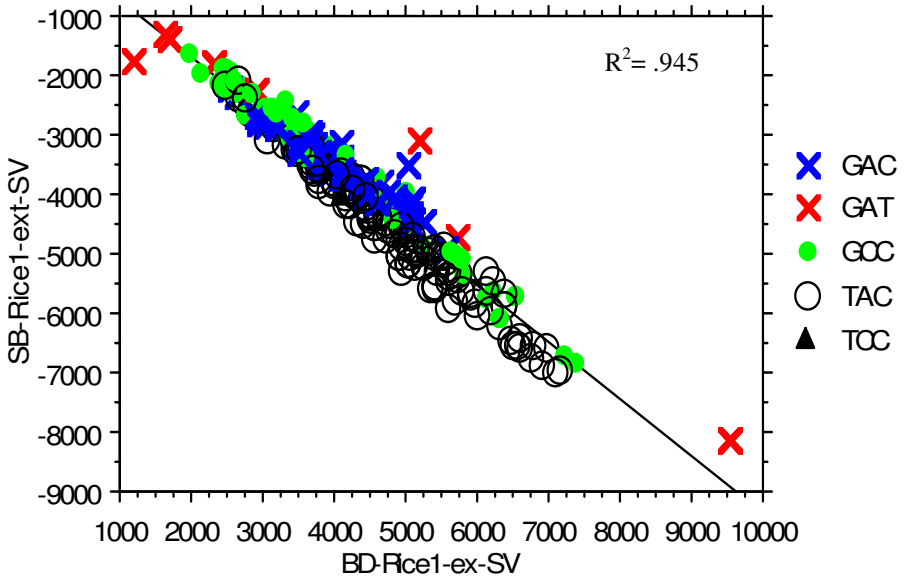

Figure 4.11. Breakdown specific viscosity vs. Setback specific viscosity

As discussed in chapter II, the traditional classification of GBSS alleles as Wx*a* = *indica* and Wx*b* = *japonica* is somewhat difficult to directly apply to the five classes of GBSS alleles, defined based on SNP in exon 1, 6 and 10 or the CT repeat. However, most of the high amylose *indica* varieties appear to have the CT_{11} GAT allele. For example Shu et al. (1999) found that most of the Chinese indica varieties had the CT_{11} allele of GBSS. Subsequent analysis in the Park lab has demonstrated that almost all of these contain the GAT SNP pattern (Walker, Dobo, and Park, unpublished). It is likely that the structure of amylopectin, particularly the ratio of long and short amylopectin chains, is responsible for the enhanced shear resistance seen in varieties with the GAT allele. However, this remains to be directly demonstrated.

A second key role of GBSS alleles is in controlling starch reassociation and gel hardness. This is most clearly demonstrated in Figure 4.9. These results are consistent with the results of Bao et al. (2000), who demonstrated that the region of chromosome 6 which contains GBSS largely controlled gel hardness in a segregating population derived from Zhai-Ye-Qing8/ Jing-Xi 17. It is also consistent with the fact that formation of an extended amylose matrix is a key aspect of formation of aged starch gels and that the nature and extent of this matrix largely determine the properties of starch gels (Eliasson and Bohlin 1982, Ring and Stainsby 1982, Christianson and Bagley 1983, Steeneken 1989).

The role of GBSS alleles in determining relative shear resistance in the RVA and their role in determining starch gel hardness appeared to be largely independent. As mentioned above, relative shear resistance of varieties with GBSS alleles other than GAT were all very similar. However, there were large differences in gel hardness among varieties with other GBSS alleles (Figure 4.9).

In contrast, GBSS alleles played only a small role in determining PV. As shown in Table 4.1, GBSS explained only 17.5% of the variation in PV. Similarly, apparent

amylose content only explained 19.3% of the variance in PV. These results are in contrast to the strong correlation (R^2 = -0.85) observed between apparent amylose and PV by Chen et al. (2004). However, they are consistent with previous studies by other investigators. For example, using an *indica/japonica* cross, Bao et al. (2000) found that GBSS accounted for from 27.8 to 57% of the variance in all of the standard RVA pasting parameters except PV. The only QTL for peak viscosity, however, were on chromosomes 2 and 12 and only accounted for 14 and 10.6% of the variation, respectively.

As illustrated in Figures 4.6 and 4.7, differences in peak viscosity can obscure the relatively simple relationship between GBSS alleles, shear strength, and BD. Differences in PV expose starch granules to different shear forces and thus lead to differences in BD. The resulting differences in BD also affect subsequent starch reassociation. If BD is very limited, the resulting starch gel will consist of largely intact starch granules embedded in a matrix of leached amylose. However, if PV is very high and BD is thus extensive, the resulting gel will largely consist of fragmented remnants of starch granules embedded in a matrix both of leached amylose and amylose/amylopectin released from the fragmented granules.

Previous investigators have noted the proportionality between PV and key parameters such as BD, and pointed out the difficulties in comparing BD in samples which have different PV (Mazurs et al. 1957; Bhattacharya and Sowbhagya 1978). However, the method developed in the current study is more straightforward and efficient than the concentration screening suggested by Bhattacharya and Sowbhagya (1978) and thus is more applicable to analysis of large numbers of samples.

The correlations observed between GBSS alleles and other pasting properties can be explained based on GBSS's role in shear strength and gel formation. As shown in Figure 4.2, varieties with the GAT allele had notably higher HPV and CPV. HPV is the

viscosity at the end of the holding period at 95°C. Since varieties with the GAT allele were more resistant to shear and had lower BD, it is not surprising that the HPV was also high. CPV is higher than HPV both because of the increase in viscosity of water upon cooling to 50°C and because of the limited amount of starch reassociation. As one might expect based on both their high amylose content and shear resistance, varieties with the GAT allele had both the highest CPV (Figure 4.2) and starch gel hardness values (Figure 4.9).

As discussed above, traditional measurements of starch pasting parameters do not account for changes in viscosity due to water. Water has a viscosity of 0.287 cp at 95°C and 0.547 cP at 50°C. While these viscosity values are both very low, changes in water viscosity cause a corresponding change in the viscosity of rice flour gels. Thus even if there is no starch reassociation, the viscosity of the sample will almost double when it is cooled to 50°C. As discussed above, changes in the viscosity of water account for the majority of the increase in viscosity upon cooling from 95 to 50°C that is described as CV.

A simple method to deal with changes in water viscosity is to express the data in terms of specific viscosity (viscosity/viscosity of water). Using specific viscosity substantially improves the correlation between CV and gel hardness (R^2=0.385) compared to using the simple viscosity numbers taken directly from the RVA (R^2 = 0.179).

These data suggest that the early stages of gel formation begin to occur at 50°C even though the samples are being stirred continuously at 160 rpm. Since gel hardness and amylose were tightly correlated with GBSS alleles (Figure 4.9), it was not surprising that CV was also correlated with GBSS alleles. As discussed above, SB = CV – BD. However, the actual increase in specific viscosity measured as CV is very small relative to BD (Figure 4.10). Thus when complications due to changes in the viscosity of water were removed, SB was almost identical to BD (R^2 = 0.945).

While all of the traditional RVA parameters can be correlated with GBSS alleles to at least some degree, most of them actually provided little, if any, new information. BD is a measure of shear resistance, but is complicated by variation in peak viscosity and thus the shear forces to which granules are exposed. HPV combines the variation in PV and BD, but provides no new information. The traditional measures of CPV, CV and SB ignore the 1.9 fold increase in viscosity that would be expected to occur simply due to the increase of water viscosity at 50^{o}C (0.547 cP) compared to 95^{o}C (0.287). This is particularly problematic since the traditional calculation of CV and SB both involve calculating the difference in viscosities measured at 95^{o}C and 50^{o}C. This type of calculation ignores the fact that the change of viscosity due to the effects of water will be in constant ratio rather than differing by a constant value. Thus the traditional values generated for CV and SB are somewhat artifactual.

Fundamental measurements such as shear resistance are more informative than traditional parameters such as BD. CV provides some measure of starch reassociation, particularly when expressed as specific viscosity. However, more direct information about the properties of starch gels can be obtained by allowing samples to cool and form consolidated starch gels rather than relying on measurements made on samples at 50^{o}C which are being continuously stirred at 160 rpm.

RVA can be used as a screening tool to efficiently identify samples which have the GAT allele, but it provides limited resolution of the other GBSS alleles. In most cases, it is likely to be simpler and easier to directly measure GBSS alleles using either the CT repeats or the three SNPs discussed in chapter II. The primary value of the RVA is to examine properties such as starch swelling and peak viscosity which are important for rice functionality, but which were largely controlled by genes other than GBSS.

.

CHAPTER V

SCREENING FOR LOW AMYLOSE VARIETIES WHICH ARE LESS TEMPERATURE SENSITIVE

Introduction

As discussed in previous chapters, amylose content is generally the primary factor controlling the cooking and processing quality of rice. Low amylose varieties are usually soft and sticky after cooking, while high amylose varieties usually have firm, separate grains. However, the amount of amylose and cooking quality can vary substantially depending on environmental conditions.

Low amylose varieties are known to be particularly temperature sensitive and have increased amylose content when grain develops at low temperature (Takeda 1988, Asaoka et al. 1985, Hirano and Sano 1998, Larkin 1999). Not only does this make it difficult to produce low amylose rice in cool climates, it also causes year-to-year variability in the quality of rice products.

As also discussed above, the amylose content of rice is largely determined by three single base changes in GBSS. Specifically, low amylose varieties have the sequence AGTTATA at the 5' leader intron splice site, while both intermediate and high amylose varieties have the sequence AGGTATA (Ayres et al. 1997). This single base change appears to interfere with the binding of the U1 SNRP to the 5' splice site, leading to inefficient mRNA splicing, accumulation of partially spliced intermediates, and to utilization of alternate splice sites (Bligh et al. 1998, Cai et al. 1998, Isshiki et al. 1998, Larkin and Park 1999).

The effect of the G/T polymorphism in the leader intron of GBSS is temperature dependent (Larkin and Park 1999). At 18°C, the G to T change has only a small

phenotypic effect. However, at 25°C and particularly at 32°C, the amount of mature GBSS mRNA in developing seeds is dramatically reduced. The reduction in GBSS mRNA level at higher temperatures is associated with increased utilization of splice sites which create a short upstream open reading frame (uORF) with a near-consensus ribosome binding site (Larkin and Park 1999). This uORF appears to capture ribosomes before they can reach the open reading frame that produces GBSS protein and also leads to GBSS mRNA degradation.

To find low amylose varieties in which cooking quality would be less variable in response to temperature differences during grain development, 1000 low amylose accessions were screened to see if any lacked the AGTTATA that is known to cause temperature sensitive GBSS mRNA splicing. This screen yielded only one accession. This variety contained approximately 14% amylose, but had the temperature stable AGGTATA 5' leader intron splice site typically seen in intermediate and high amylose varieties. As expected from its 5' leader intron splice site, the amylose content of this variety did not increase when grain developed at 18 or 25°C. However, the variety was not stable at high temperature, with its amylose level falling dramatically at 32°C.

A similar drop in amylose content between 25 and 32°C was seen in several standard intermediate amylose varieties. In fact, the profile of amylose vs. temperature in the atypical low amylose variety was parallel to that in the standard intermediate amylose varieties examined. A series of experiments were thus conducted to determine whether the unusual amylose vs. temperature profile in the atypical line was due to a novel GBSS allele or to a trans-acting factor(s) that could potentially be used to reduce the amylose level of a temperature stable GBSS allele.

Materials and Methods

One thousand low amylose rice accessions of diverse geographical origin were obtained from the USDA-ARS National Small Grains Collection in Aberdeen, Idaho with the generous assistance of Harold Bockelman. To determine whether any of these lacked the 5' splice sequence AGTTATA that causes temperature-sensitive GBSS mRNA processing, genomic DNA from each accession was screened using the READIT™ system as described by Bormans et al. (2002). To determine whether lines that lacked the sequence AGTTATA did, in fact, have low amylose content, milled grain were then screened using I_2-KI and the absorbance measured at 620 nm (Juliano 1971). Candidate accessions and controls were then grown in growth chambers at either 18, 25 or 32°C with a 16 hour light/8 hour dark cycle as described by Larkin and Park (1999).

To determine whether the low amylose content of the atypical variety "Kumbi" was due to a novel GBSS allele, Kumbi was crossed with the standard low amylose variety M103. Segregation of the GBSS genes in the cross was assayed using the CT repeat in the GBSS leader intron as described by Ayres et al. (1997). Selected progeny and parental controls were then grown in a greenhouse at Texas A&M University, College Station. At two to three days post-anthesis, plants were transferred to 18, 25 or 32°C growth chambers maintained under a 16/8 hour light/dark cycle.

In addition to assaying amylose content in the mature dried grain, total starch granule-bound protein was examined by a modification of the method of Schwart and Echt (1982) as described in chapter II.

The GBSS gene from Kumbi was sequenced using five pairs of PCR primers (Table 5.1). The resulting fragments were electrophoresed and excised from 1% agarose gels and purified with a QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) according to the manufacturer's instructions. The purified DNA was then sequenced using a

BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA).

Results

The screen of 1000 lines from the US germplasm collection that were labeled as having 14-18% amylose yielded several candidates that lacked the AGTTATA splice known to cause temperature dependent GBSS mRNA splicing. However as might be expected, most of these were found to actually have intermediate or high amylose content when the same seed used for DNA analysis were directly assayed with I_2-KI. Of the varieties tested, only the Indonesian accession "Kumbi" actually had low amylose levels when grain developed 18°C (Figure 5.1).

Both the READIT™ assay and direct DNA sequencing showed that Kumbi contains the 5' GBSS leader intron sequence, AGGTATA, which is found in typical intermediate and high amylose varieties. This sequence does not cause temperature-dependent GBSS mRNA splicing (Larkin and Park 1999). Thus, as expected, the amount of apparent amylose in Kumbi, and all other varieties with the AGGTATA splice site that were tested, remained essentially constant from 18 to 25°C. In contrast, the amount of apparent amylose in all of the varieties with the temperature sensitive AGTTATA splice site that were tested decreased over this temperature interval.

While the apparent amylose content of Kumbi was stable from 18 to 25°C, it was not stable at higher temperatures. At 32°C, the apparent amylose level of Kumbi decreased to approximately 5%. This level of apparent amylose is comparable to that seen in glutinous (GBSS null) varieties such as Calmochi 101 (Figure 5.1). The residual iodine bound by glutinous varieties such as Calmochi 101 is thought to reflect binding to the long chains of amylopectin rather than actual amylose (Takeda et al. 1987) and results in

Table 5.1. PCR primers used to sequence GBSS from Kumbi

Primer	Sequence
Wx1F/Wx1R	F: ACCATTCCTTCAGTTCTTTGTC
	R: TGAATTGTTTAAGGTTTGGTGAGCC
Wx2F/2R	F: GTTCTTGATCATCGCATTGG
	R: TACACATCCATCCAATGCGATGATCAAGAAC
Wx3F/3R	F: GTTCTTGATCATCGCATTGG
	R: TTGTCCTTGCTAGGATCC
Wx4F/4R	F: GAAGATCAACTGGATGAAGG
	R: GGCATGGTATAATAGGAACA

a red staining with I_2-KI rather than the characteristic blue color seen with amylose. As expected, grains of Kumbi which developed at 32°C also stained red with I_2-KI and had the opaque waxy appearance that is typical of Calmochi 101 and other glutinous rice varieties, while Kumbi grain which developed at 25°C stained blue and had the translucent appearance of standard rice varieties (Figure 5.2, A and B).

The decrease in apparent amylose of Kumbi seed which developed at 32°C rather than 25°C was similar to that seen in intermediate amylose varieties (Figure 5.1). The reason for the decline is not completely clear, but it does not appear to be due to differences in GBSS mRNA processing (Larkin and Park, 1999). These investigators showed that the amount of mature GBSS mRNA and the pattern of splice site utilization are essentially constant over this temperature range in Lemont and other intermediate varieties with the AGGTATA splice site.

To determine the genetic basis for Kumbi's unusual amylose vs. temperature profile, it was crossed with the standard low amylose variety M103. After fertilization, progeny from the cross and parental controls were then transferred to growth chambers maintained at a constant temperature of either 25 or 32°C, and seed were allowed to develop to maturity.

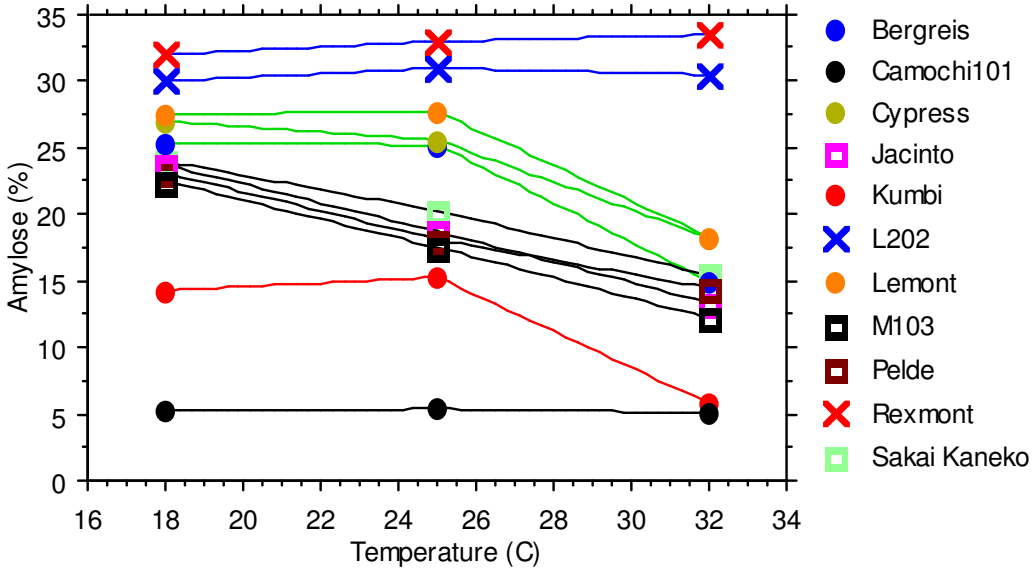

Figure 5.1. Effect of temperature on amylose content. After anthesis, plants were grown at the temperature indicated until the grain was mature. Results are shown for high amylose varieties with the GAT allele (Rexmont) and GAC allele (L202); the intermediate amylose varieties, Bergreis, Lemont and Cypress which all have the GCC allele; and the standard low amylose varieties Jacinto, Pelde, Sakai Kaneko and M103 which all have the sequence AGTTATA in exon 1. Results are also shown for the atypical low amylose variety Kumbi, and for Calmochi 101 a "zero" amylose variety.

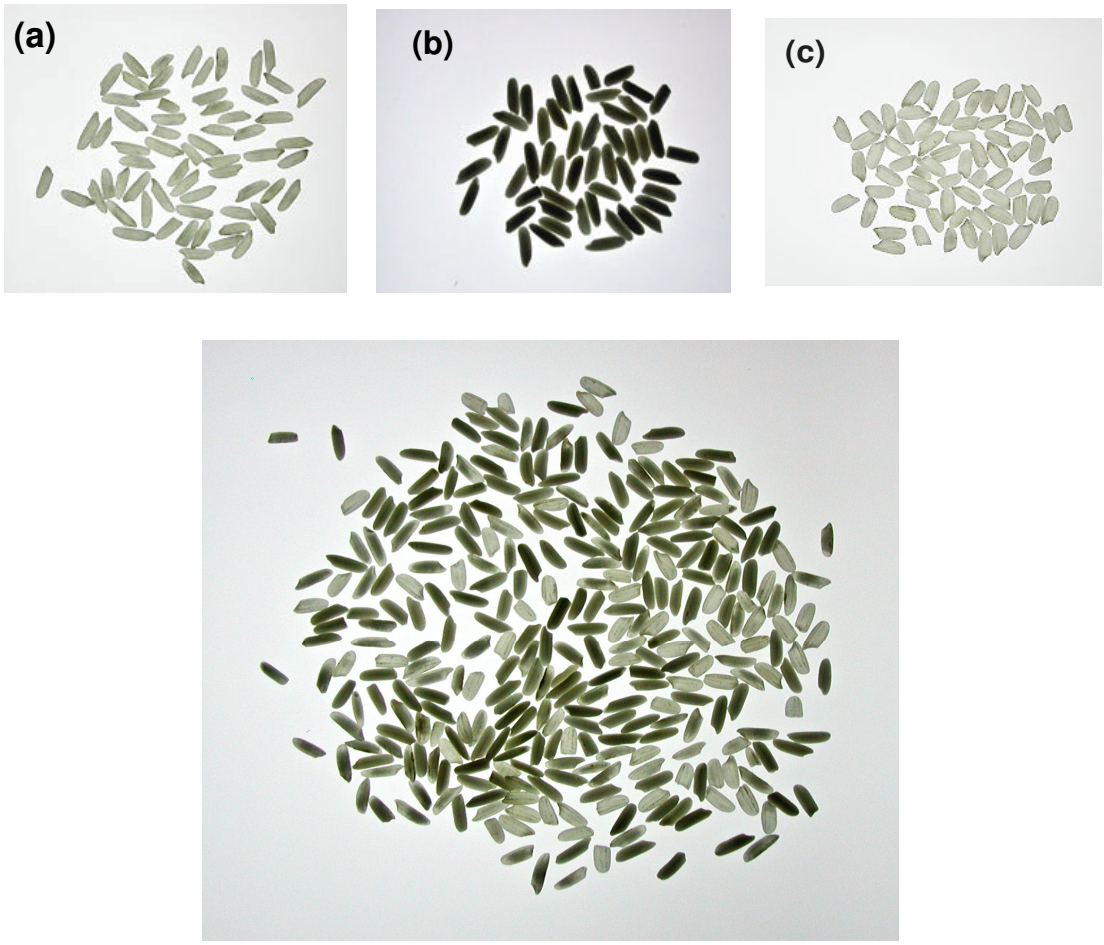

Figure 5.2. Kumbi x M103 F_2 progeny seed phenotypes. Seeds from Kumbi that were grown at 25°C (a) and 32°C (b). Seeds from M103 that were grown at 32°C. F_2 seeds from Kumbi x M103 that were grown at 32°C (d).

As mentioned above, all of the Kumbi grains which developed at 32°C had the opaque appearance typical of glutinous rice varieties (Figure 5.2). In contrast, all of the M103 seed were translucent at 32°C. Translucency thus provided a simple phenotypic assay for segregation of amylose content in the cross. As expected for segregation of a single co-dominant gene, three phenotypic classes were observed among the F_2 seed: opaque, intermediate and translucent (Figure 5.2). These three classes were present in the ratio of 153:276:144. This is in good agreement with the ratio of 1:2:1 expected for segregation of a single co-dominant gene ($\chi^2 = 1.05$, $p < 0.001$).

The key question at this point was whether the single co-dominant gene apparently responsible for the segregation of opaque, intermediate and translucent grain in the F_2 population from the Kumbi/M103 cross was a novel allele of GBSS from Kumbi or whether it was some other gene affecting the synthesis or activity of GBSS mRNA or protein. This question could be readily addressed since preliminary experiments showed that the GBSS genes from Kumbi and M103 can be distinguished based on the CT repeat that is present immediately upstream of the GBSS 5' leader intron splice site.

When individual F_2 seeds were subjected to DNA analysis, it was found that all of the opaque seeds had the Kumbi allele of GBSS, while all of the translucent seed had the M103 allele (Figure 5.3). As expected if the phenotype was due to segregation of co-dominant GBSS alleles, seed with the intermediate, semi-translucent phenotype were heterozygous and carried both alleles.

Co-segregation of the opaque phenotype and the GBSS allele from Kumbi suggested that the unusual amylose vs. temperature profile of Kumbi was due to a novel allele of GBSS. However, it did not exclude the possibility of involvement of other genes which are closely linked to GBSS.

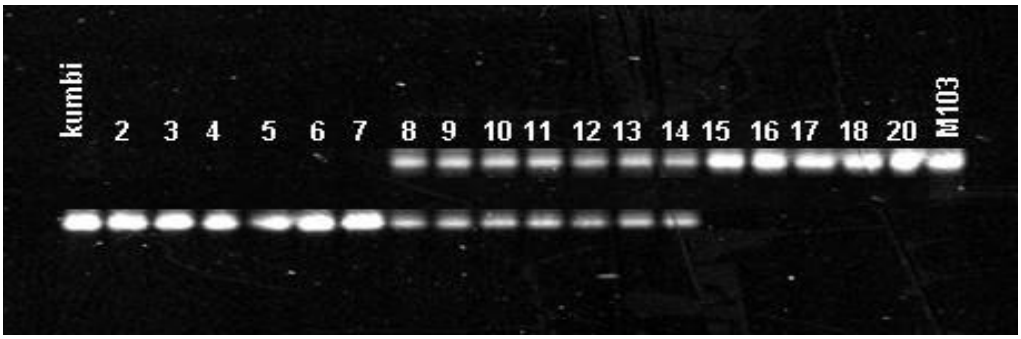

Figure 5.3. The CT repeats of GBSS alleles in Kumbi/M103 progeny. Lane 1, Kumbi which has the CT_{10} allele and opaque phenotype when grown at 32°C. Lanes 2 to 7, F_2 seeds with opaque phenotype when grown at 32°C. Lanes 8 to 14, F_2 seeds with intermediate phenotype at 32°C. Lanes 15 to 20, F_2 seeds with translucent phenotype. Lane 21, M103 which has the CT_{19} allele and is translucent when grown at 32°C.

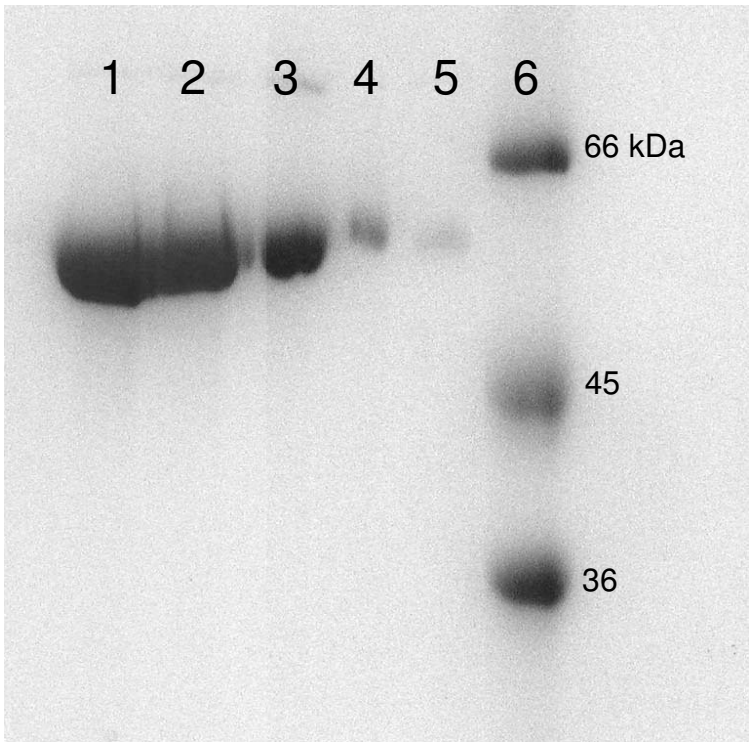

Figure 5.4. Effect of temperature on GBSS protein accumulation. Lanes 1 & 2, Kumbi GBSS protein at 25 & 32°C respectively. Lanes 3 & 4, M103 GBSS at 25 & 32°C respectively. Lane 5, CM101 a "zero amylose" variety, at 25°C. Lane 6, molecular weight markers (kDa).

To test this possibility and to further examine the molecular mechanisms involved, total granule-bound protein was extracted from seed that had developed at 25 and 32ºC. GBSS typically accounts for the majority of the protein in such washed starch granules (Schwart and Echt 1982, Sano 1984). As expected, the primary protein observed in these preparations had the apparent molecular weight of 60 kDa expected for GBSS (Figure 5.4). Also as expected, this band was almost completely absent in the glutinous variety Calmochi 101 (Figure 5.4, lane 5).

M103 has the temperature-sensitive AGTTATA GBSS mRNA splice site and shows the typical pattern of lower amylose content at high temperature (Figure 5.1). Thus as expected, starch granules from M103 that had developed at 32^{o}C contained much less GBSS protein than those grown at 25^{o}C (Figure 5.4, compare lanes 3 and 4). In contrast, Kumbi grain that developed at either 25ºC (lane 1) or 32ºC (lane 2) contained large amounts of GBSS protein. The essentially constant amount of GBSS protein in Kumbi at 25^{o}C and 32^{o}C is consistent with Kumbi having the temperature stable AGGTATA GBSS splice site. However, it does not explain Kumbi's dramatic low amylose content at 32^{o}C.

The simplest interpretation of the data is that Kumbi contains a GBSS allele for which mRNA and protein synthesis is temperature stable, but whose enzymatic activity is temperature sensitive. Consistent with this interpretation, an A to G change was found at position 2004 in Kumbi relative to the high amylose variety Rexmont (GenBank accession AF 141954). This single nucleotide change would result in changing aspartic acid to glycine. While the effect of this single amino acid change has not been directly assayed, it was noted that this region of GBSS is highly conserved between species and that the amino acid change observed in Kumbi only seven amino acids away from the position of a substitution that led to reduced amylose synthesis that was described by Sato et al. (2002).

Discussion

The vast majority of low amylose varieties contain the GBSS 5' leader splice site AGTTATA that is known to cause temperature-dependent differences in GBSS mRNA processing. The phenotypic effect of this splice site varies dramatically over the temperature range of 18°C to 32°C (Larkin and Park 1999). Since fluctuations over this temperature range often occur in rice fields during grain development, amylose production and thus cooking quality of the grain is variable.

A mutant named COI (cool temperature insensitive) was previously identified by Suzuki et al. (2002). The GBSS protein level in COI remained constant across moderate temperatures (21 to 28°C) and the amylose level in COI also was unaffected by temperature. However, the amylose level in COI is extremely low (3%) and thus it is not suitable for applications requiring the 14-19% amylose of standard low amylose varieties.

The goal of this project was to identify varieties with 14-19% amylose that are less subject to environmental variation by screening for those which lacked the temperature sensitive AGTTATA splice site. Other factors beside the GBSS splice site also contribute to temperature sensitivity, as illustrated by the drop in amylose at 32°C seen in typical intermediate amylose varieties such as Lemont, which contain the temperature stable AGGTATA splice site (Larkin and Park 1999; Figure 5.1). However, the possibility of identifying intermediate and low amylose varieties which are less subject to environmental variation was suggested by the stable amylose content across the entire temperature range of 18 to 32°C that was observed in the high amylose varieties L202 and Rexmont (Figure 5.1). These varieties contain the temperature stable GBSS 5' splice site and also have a single amino acid change in exon 6 that may affect GBSS protein stability (Larkin 1999, Larkin and Park 1999, Larkin and Park, 2003).

The initial screen of 1000 accessions from the US germplasm collection that were labeled as having 14-18% amylose yielded several accessions with the exon 1 sequence AGGTATA. This GBSS 5' leader intron splice site that is normally found only in intermediate and high amylose varieties. Most of the accessions found in the initial screen appear to have been mislabeled and were found to actually contain intermediate or high amylose. This was not surprising since variation is often observed within accessions of rice germplasm collections and between duplicate accessions of the same variety (Olufowote et al. 1997). The initial seed of a few other varieties contained 15-19% amylose, but had a temperature response pattern similar to the standard US long grain varieties "Lemont" and "Cypress" when grown under controlled conditions. These varieties have intermediate amylose content at 18°C or 25°C, but lower amylose content at 32°C. The low amylose content of some original seed from the germplasm collection may thus have been the result of growth under high temperatures.

Of the accessions examined, only the Indonesian variety Kumbi had low amylose content when grain developed at 18°C. This variety had the temperature stable AGGTATA GBSS 5' leader intron splice site that is characteristic of intermediate amylose and high amylose varieties. Interestingly, its temperature profile was parallel to that of standard intermediate amylose varieties such a Lemont and Cypress (Figure 5.1).

It was hoped that Kumbi might contain a mutation in a trans-acting factor which could be used to reduce the amylose content of temperature stable varieties such as L202 or Rexmont. Trans-acting "dull" mutations have previously been identified which decrease amylose content by up to 10% (Yano et al. 1988, Tomita and Nakagahra 1990). However the dull mutations only reduce the amylose content of varieties with the GTTATA GBSS leader intron splice site, and thus these varieties are still affected by temperature fluctuations.

Results of three types of analysis indicated that the atypical temperature profile of Kumbi is due to a novel allele of GBSS encoding a temperature sensitive enzyme rather than to a trans-acting factor. First, the very low levels of amylose observed in Kumbi grain which developed at 32°C and the corresponding opaque appearance and red staining with iodine were found to co-segregate with the Kumbi allele of GBSS in a cross with the standard low amylose variety M103 (Figure 5.3). Second, while the amount of amylose produced by Kumbi was much lower at 32°C than at 25°C, Kumbi had high and essentially constant levels of GBSS protein at both 25 and 32°C (Figure 5.4). A similar phenomenon was also observed in the progeny of the Kumbi/M103 cross that are homozygous for the Kumbi allele of GBSS (data not shown).

Finally, direct sequence analysis is also consistent with Kumbi containing a GBSS allele that produces relatively constant amounts of protein at different temperatures, but which encodes an enzyme which is temperature sensitive. Consistent with the essentially constant amount of GBSS protein produced by Kumbi at 25°C and 32°C, Kumbi contains the same AGGTATA leader intron splice site as the temperature stable high amylose varieties L202 and Rexmont. Kumbi also contains the same sequence in exon 6 that distinguished L202 and Rexmont from intermediate amylose varieties such as Lemont and which was postulated to play a role in GBSS protein stability (Larkin 1999).

The only difference in the sequence of L202 and Kumbi is a single base mutation in exon four. This mutation results in a substitution of glycine for aspartic acid in Kumbi. Glycine is neutral amino acid with only a single H as a side chain, while aspartic acid has a large negatively charged side chain. This non-conservative amino acid change occurs in a region of the GBSS protein that is highly conserved in both monocots and dicots, and is only seven amino acids away from an amino acid substitution associated with reduced enzymatic activity that was identified by Sato et al. (2002). The GBSS gene examined by Sato et al. (2002) contained an additional amino acid change in exon five and it was not shown which amino acid change affected the enzymatic activity.

However the clear difference in the amount of amylose produced by Kumbi at 25 and 32^{o}C compared to the essentially constant amount of GBSS protein suggests that this region of exon four plays a key role in the enzymatic activity of GBSS.

In conclusion, after screening 1000 accessions, only one was found which has low amylose content at 18^{o}C but lacks the AGTTATA splice site that makes GBSS mRNA processing temperature sensitive. While the amount of GBSS protein produced by this variety is less temperature stable, its enzymatic activity decreases dramatically at 32^{o}C due to a single amino acid change in a conserved region of exon four. Thus rather than identifying a useful exception to the tight linkage between low amylose content and the temperature sensitive AGTTATA splice site, this study demonstrated just how tight this linkage actually is.

CHAPTER VI

SUMMARY AND CONCLUSION

It has long been known that rice eating and cooking quality are mainly determined by amylose content. Commercial rice varieties are, in fact, classified into market classes based on having "zero" (0-7%), low (10-20), intermediate (20-25%), and high (>25%) apparent amylose (Juliano et al. 1981). The gene that makes amylose, GBSS, was identified many years ago (Echt and Schwartz 1981, Hirano and Sano 1991). However until very recently, it was difficult to clearly differentiate between the market classes of rice in a seedling assay. Due to the effects of environmental variation, modifier genes, and dominance relationships, it was sometimes difficult to unambiguously distinguish amylose classes even when grain was available.

Ayres et al. (1997) showed that the CT repeat in exon 1 of GBSS could often distinguish between market classes. However, some varieties with different amylose contents had the same number of CT repeats. Conversely, varieties with the same amylose content often had different numbers of repeats. A single nucleotide polymorphism in exon 1 of GBSS was also shown to clearly distinguish low amylose varieties. However, the exon 1 SNP could not differentiate between varieties with intermediate vs. high amylose (Ayres et al. 1997). Varieties with intermediate amylose were also difficult to fit into the classification of GBSS alleles as *wx*, Wx_a and Wx_b.

Another major question at the outset of this project was which genes control the pasting properties of rice and by what mechanism they act. Analysis of pasting properties during simulated cooking cycles has been widely used to distinguish between rice varieties which have similar levels of apparent amylose and similar gelatinization temperature, but different cooking properties. Many different genes and chromosomal regions have been shown to influence the pasting properties of rice. Surprisingly

however, several studies had found that GBSS was the most important factor controlling pasting parameters (Bao et al. 1999). Furthermore, definitions of the same type of cooking quality based on pasting parameters sometimes appeared contradictory.

In this project, substantial progress was made in both areas. The SNP in exons 1, 6, and 10 were found to account for 89.2% of the variation in apparent amylose content in the extended pedigree of 89 US rice varieties that had previously been examined Ayres et al. (1997) for exon 1. In particular, the addition of data from exon 6 or 10 allowed intermediate amylose varieties to be clearly distinguished from those with high amylose. In previous work using differences in the CT repeat to define GBSS alleles, it had not been possible to distinguish most of the intermediate amylose US cultivars from certain high amylose cultivars such as L202 and Jodon since they all have 20 CT repeats. However, this was not a problem using the three SNPs. All intermediate amylose varieties have the GCC allele and all high amylose varieties have either the GAC or GAT allele.

The second objective of the project was to determine whether there was also a very close relationship between GBSS alleles and apparent amylose content in other, more diverse, germplasm collections. These included both a large European germplasm collection which had been grown in northern Italy, a set of diverse samples from WARDA which had been grown in West Africa, and also a collection of the native African rice species *O. glaberrima* that were obtained from the US germplasm collection. In all cases, a strong relationship was found between apparent amylose content and SNP in exons 1, 6 and 10 of GBSS. The three SNP accounted for 93.8% of the variation in apparent amylose content among the 279 accessions in the European collection, 93.3% of the variation in apparent amylose content in the 77 samples from WARDA, and 77% of the variation in apparent amylose in the 48 samples of *O. glaberrima* from the US germplasm collection. The ability of the three GBSS SNPs accurately to distinguish low, intermediate and high amylose varieties across a wide range of rice germplasm has

also been reported recently by (Traore at al 2004) and thus appears to be generally applicable and extremely robust.

Most accessions of *O. glaberrima* and some progeny of *O. glaberrima*/*O. sativa* crosses from WARDA contain a novel GBSS allele, GAC_glab. Consistent with the suggestion that *O. glaberrima* represents a very ancient type of rice (Semon et al. 2005), the GAC_glab allele lacks a transposable element in exon 10 that is present in both the *indica* and *japonica* subspecies of *O. sativa.* Consistent with the two *O. glaberrima* GBSS genes currently in the public sequence database, the GAC_glab allele in all the varieties examined in the current study also contains an additional SNP in exon 12. However, the conservative amino acid substitution that would result from this SNP does not appear to affect apparent amylose content.

The close relationship between GBSS alleles and apparent amylose content in African material that was observed in the current study is in sharp contrast to previous reports of dramatic transgressive segregation of amylose content in the progeny of *O. sativa/O. glaberrima* crosses (Watanabe et al. 2002). Direct examination of low amylose progeny from this experiment, generously provided by WARDA, revealed the presence of non-parental alleles. However, even the off-type progeny that was tested still had the amylose content which would have been expected based on its GBSS allele. Thus the intermediate amylose *O. sativa* parent had the expected GCC allele and the high amylose *O. glaberrima* parent had the expected GAC_glab allele and the low amylose "progeny" had a TAC allele consistent with its low amylose content. Rather than transgressive segregation, the low amylose progeny from this cross were likely due to accidental outcrossing.

It should be emphasized that the presence of off-types in breeding populations is not just a problem in Africa. The Park lab has also detected significant numbers of off-types in breeding populations from Asia, Europe and the US (Walker and Park, unpublished

observations). Off-types and other types of confusion have also been seen in foundation seed from Asia, Europe and the US (Walker and Park, unpublished observations). Accessions from germplasm collections also often contain off-types and different accessions of the same variety sometimes differ (Olufowote et al. 1997). For example, as noted in chapter III, some of the "*O. glaberrima*" that was obtained from the US germplasm collection did not have the short ligule and other morphological features expected for actual *O. glaberrima*. The GBSS genes from these accessions also lacked the SNP in exon 12 that is characteristic of *O. glaberrima* and contained the transposon in intron 10 that was expected to be absent in *O. glaberrima*. In addition it should also be noted that, as this is being written, there is great concern about recent detection of low levels of genetically modified rice that are present as contaminants in traditional US rice varieties and in rice varieties from China.

The key conclusion from chapter III is that marker-assisted selection for GBSS alleles should provide a quick and cost effective method to develop African rice varieties with specific cooking and processing properties. Analysis of GBSS SNP is technically simple and allows early generation selection for specific cooking and processing properties. It should be particularly useful for further development of Nerica varieties since plantlets regenerated from anther culture can be directly tested to see which GBSS allele they contain. Since cooking and processing quality are largely determined by GBSS allele, plantlets which do not have the appropriate allele for the target producers can be discarded without having to invest additional time or other resources. In other cases, both GBSS alleles from a cross may be useful but to different end-users. In this case, GBSS allele information could determine the selection scheme into which regenerated plants should be placed.

The strength of correlation between GBSS alleles and apparent amylose was demonstrated by the attempt to find low amylose varieties which lack the sequence AGTTATA which is known to cause temperature dependent GBSS mRNA splicing that is described in chapter V. After screening 1000 accessions from the US germplasm

collection that were listed as having low amylose, only one variety was found which had low amylose content at 18°C, but the exon sequence AGGTATA normally characteristic of intermediate and high amylose varieties. Even the unusual amylose vs. temperature profile of this variety, however, proved to be due to a mutation in GBSS.

Many genes besides GBSS are involved in synthesis of starch granules and other aspects of rice grain development. The best characterized of these is SSIIa, which largely controls gelatinization temperature. However, varieties with the same apparent amylose content and gelatinization temperature can sometimes have very different cooking and processing quality.

The most common method currently used to examine rice pasting properties during a simulated cooking cycle is the Rapid Visco™ Analyzer (RVA). RVA provides information about gelatinization and starch granule swelling and stability and also a limited amount of information about starch reassociation upon cooling. Notably it can distinguish rice varieties which have low solids loss and greater retention of grain integrity during processing from varieties with similar grain geometry, amylose content and gelatinization temperature which lack these favorable properties.

The initial relationship seen between GBSS alleles and pasting properties shown in Figure 4.2 was not particularly strong. Varieties with the GAT allele had higher HPV and CPV, but did not have the notably lower SB that was expected based on descriptions of the US varieties Jojutla, Newrex, and Rexmont. Thus it was not at all clear whether the GAT allele was correlated with stronger starch granules or whether the varieties in the European collection with the GAT allele would have the favorable processing quality of “Rexmont quality” varieties from the US.

The simplest interpretation of the data was that the poor correlation between GBSS alleles and BD was due to the segregation of other genes, perhaps reflecting wider

genetic diversity of the European germplasm collection. At this point, one could have embarked upon making crosses between the varieties with extreme phenotypes and then used QTL analysis to identify the genes and chromosomal regions involved, as done by other investigators (e.g. Bao et al. 2004).

However, rather than using more powerful genetic tools, it was more productive to think mechanistically about what factors control BD during the RVA assay. As discussed in chapter IV, the standard RVA assay for BD suffers from two fundamental problems. First, samples are typically stirred at 160 rpm irrespective of their maximum viscosity. Thus starch granules are exposed to higher shear forces in samples with high peak viscosity than in samples with low peak viscosity. Second, in standard RVA analysis, samples are only held at 95^{o}C for 2.5 minutes. This is not enough time for BD to approach equilibrium.

Extending the holding time at 95^{o}C and plotting BD vs. PV solved these problems and revealed a surprisingly simple relationship between shear strength and GBSS alleles. Other than varieties with the GAT allele of GBSS, all of the other varieties in the collection had remarkably similar shear resistance. In fact, peak height could explain 90% of the variation in BD for all of the varieties with other GBSS alleles, irrespective of amylose content. Varieties with the GAT allele, however, showed considerably higher shear resistance. Thus relative shear resistance appears to be almost completely determined by the single proline to serine amino acid substitution in exon 10 of varieties with the GAT allele.

It should be emphasized that other genes do actually determine the extent of BD that occurs in a given sample. Indeed, there is a wide range in the amount of DB observed in samples with the same GBSS allele, depending on their peak viscosity. Peak viscosity correlated poorly with GBSS alleles and thus largely reflects the effect of other genes and/or environmental interactions. Thus there are two advantages of looking at relative

shear strength rather than directly at BD. First, relative shear resistance is more of a fundamental measure of the physical properties of rice starch granules, whereas BD is subject to complications due to differences in the magnitude of the shear forces encountered in a given experiment. Second, removing variance due to other factors such as peak viscosity reveals that relative shear strength has a remarkably simple genetic basis and that is largely determined by a single SNP in exon 10 of GBSS.

Simplistically one might expect that the increase in viscosity observed when RVA samples are cooled from 95°C to 50°C is largely due to starch reassociation and thus would provide a good measure of the properties of cooked rice. However, as also discussed in chapter IV, this is not the case. The majority of the increase in viscosity upon cooling to 50°C is actually due to changes in the viscosity of water. Thus traditional pasting properties such as CV and SB are somewhat artifactual. In fact, when specific viscosity (viscosity/viscosity of water) was used to account for changes in water viscosity, SB was revealed to be essentially equivalent to BD (R^2=0.945).

The best assay to examine the relationship between GBSS alleles and the properties of starch gels was the gel hardness assay used by Bao et al. (2004, 2006). Rather than attempting to assay starch reassociation in samples at 50°C which are being continuously stirred at 160 rpm, RVA samples were stored at room temperature for 24 hours to allow firm gels to form. The viscosity of these samples was then examined by back extrusion using a piston which fit into the RVA canister. Consistent with previous work by Bao et al. (2004) on segregating populations, gel hardness was largely explained by GBSS alleles or amylose, which explained 81 and 71.5% of the variance, respectively.

Whereas relative shear strength was largely determined by the SNP in exon 10, the viscosity of starch gels was strongly correlated with amylose content. In fact, when varieties with the GAT allele were removed from the analysis to eliminate the effects of relative shear strength, the proportion of variance in gel hardness explained by amylose

only decreased from 71.5% to 71%. Thus the two primary effects of GBSS alleles on pasting parameters appear to be largely independent.

Thus two major conclusions can be drawn from the RVA analysis described in chapter IV. First, GBSS alleles strongly influence gel hardness and the GAT allele is associated with increased shear resistance. While GBSS did play a significant role in determining PV (R2= 0.17), it was not the dominant factor. The second major conclusion was the value of thinking mechanistically about pasting properties and how they are measured in the RVA rather than blindly following standard protocols.

LITERATURE CITED

Aluko, G., Martinez, C., Tohme, J., Castano, C., Bergman, C., and Oard, J. H. 2004. QTL mapping of grain quality traits from the interspecific cross Oryza sativa x O-glaberrima. Theoretical and Applied Genetics 109: 630-639.

Asaoka, M., Okuno, K., and Fuwa, H. 1985. Effect of environmental-temperature at the milky stage and amylose content and fine-structure of amylopectin of waxy and non-waxy endosperm starches of rice (*Oryza sativa* L.). Agricultural and Biological Chemistry 49: 373-379.

Ayres, N. M., McClung, A. M., Larkin, P. D., and Park, W. D. 1997. Microsatellites and single nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US germplasm. Theoretical and Applied Genetics 94:773-781.

Bao, J. S., Shen, S. Q., Sun, M., and Corke, H. 2006. Analysis of genotypic diversity in the starch physicochemical properties of nonwaxy rice: apparent amylose content, pasting viscosity and gel texture. Starch-Starke 58:259-267.

Bao, J. S., Sun, M., Zhu, L. H., and Corke, H. 2004. Analysis of quantitative trait loci for some starch properties of rice (*Oryza sativa* L.): thermal properties, gel texture and swelling volume. Journal of Cereal Science 39:379-385.

Bao, J. S, Zheng, Y. W., Xia, W. Y., He, P., Shu, Y. Q., Chen, Y., and Zhu, H. L. 2000. QTL mapping for the paste viscosity characteristics in rice (*Oryza sativa* L.) Theoretical and Applied Genetics 100: 280-284.

Bao, J. S, He, P., Xia Y. W, Chen, Y., and Zhu, L. H. 1999. Starch RVA profile parameters of rice are mainly controlled by Wx gene. Chinese Science Bulletin 44:2047-2051.

Bao, J. S., and Xia, Y. W. 1999. Genetic control of paste viscosity characteristics in indica rice (Oryza sativa L.) Theoretical and Applied Genetics 98:1120-1124.

Bergman C. J., Delgado J. T., McClung A. M., and Fjellstrom R. G. 2001. An improved method for using a microsatellite in the rice waxy gene to determine amylose class. Cereal Chemistry 78:257-260.

Beta, T., and Croke, H. 2001. Genetic and environmental variation in sorghum starch properties. J. Cereal Sci. 2001. 34:261-268.

Bhattacharya, K. R., and Sowbhagya, C. M. 1980. On the alkali degradation type of rice kerels. J. Sci. Food Agric. 31:615-618.

Bhattacharya, K. R. and Sowbhagya, C. M. 1978. On viscograms and viscography, with special reference to rice flour. Journal of Texture Studies 9:341-351.

Bhattacharya, K. R. and Sowbhagya, C. M. 1972. An improved alkali reaction test for rice quality. J. Food Technol. 7:323-331.

Bligh, H. F. J., Larkin, P. D., Roach, P. S., Jones, C. A., Fu, H., and Park, W. D. 1998. The use of alternate splice sites in granule-bound starch synthase mRNA from low-amylose rice varieties. Plant Molecular Biology 38: 407-415.

Bligh, H. F. J., Till, R. I., and Jones, C. A. 1995. A microsatellite sequence closely linked to the waxy gene of *Oryza sativa* L. Euphytica 86: 83-85.

Bollich, C. N., Webb, B. D., Marchetti, M. A., and Scott, J. E. 1990. Registration of Rexmont rice. Crop Science 30: 1160-1160.

Bollich, C. N., Webb, B. D., Marchetti, M. A., and Scott, J. E. 1980. Registration of Newrex rice. Crop Science 20:286-287.

Bollich, C. N., and Webb, B. D. 1973. Inheritance of amylose in 2 hybrid populations of rice. Cereal Chemistry 50:631-636.

Bormans, C. A., Rhodes, R. B., Kephart, D. D., McClung, A. M., and Park, W. D. 2002. Analysis of a single nucleotide polymorphism that controls the cooking quality of rice using a non-gel based assay. Euphytica 128: 261-267.

Bollich, C. N., Webb, B. D., Marchetti, M. A., and Scott, J. E. 1980. Registration of Newrex rice. Crop Science 20:286-287.

Bollich, C. N., Webb, B. D., Marchetti, M. A., and Scott, J. E. 1990. Registration of Rexmont rice. Crop Science 30:1160.

Cai, X., Wang, Z., Xing, Y., Zhang, J., and Hong, M. 1998. Aberrant splicing of intron 1 leads to the heterogeneous 5' UTR and decreased expression of waxy gene in rice cultivars of intermediate amylose content. The Plant Journal 14: 459-465.

Champagne, E. T., Bett, K. L., Vinyard, B. E., McClun, A. M., Barton, F. E., Moldenhaur, K., Linscombe, S., McKenzie, K. 1999. Correlation between cooked rice texture and Rapid Visco Analyser Measurement. Cereal Chemistry 76:764-771.

Chang, S. M., and Lii, C. Y. 1985. The pasting behaviors and the eating qualities of some rices in Taiwan. Bull. Inst. Chem. Academia Sinica. 32:57-61.

Chen, M. H., Bergman, C., and Fjellstrom, R. B. 2004. Genetic variation at the waxy locus associated with starch pasting properties in international rice germplasm. Page 168, in Plant and Animal Genome Conference Proceedings. New Orleans, Louisiana.

Christianson, D. D., and Bagley., B. E. 1983. Apparent viscosities of dispersions of swollen cornstarch granules. Cereal Chemistry 60:116-121.

Denyer, K., Clarke, B., Hylton, C., Tatge, H. and Smith , A. M. 1996. The elongation of amylose and amylopectin chains in isolated starch granules. The Plant Journal 10:1135-1143.

Dilday, R. H. 1990. Contributions of ancestral lines in the development of new cultivars of rice. Crop Science 30:905-911.

Dingkuhn, M. and Jones, M., and Johson, D. E. 1998. Growth and yield potential of *Oryza sativa* and *O. glaberrima* upland rice cultivars and their interspecific progenies. Field Crops Research 57:57-69.

Echt, C. S., and Schwartz, D. 1981. Evidence for inclusion of controlling elements within the structural gene at the waxy locus in maize. Genetics 99:275-284.

Eliasson, A. C., and Bohlin, L. 1982. Rheological properties of concentrated wheat starch gels. Starch-Starke 34:231-233.

Food and Agriculture Organization (FAO), 1994. Le riz dans la nutrition humaine: International Rice Research Institute, Los Baños, Philippines.

Fitzgerald, M. A, Martin, M., Ward, R. M., Park, W. D., Shead, H. J. 2003. Viscosity of rice flour: A rheological and biological study. Journal of Agricultural and Food Chemistry 51:2295-2299.

Gao, Z., Zeng, D., Cui, X, Zhou, Y., Yan, M., Huang, D., Li, J., and Qian, Q. 2003. Map-based cloning of the ALK gene, which controls the gelatinization temperature of rice. Science in China 46:661-668.

Gravois, K. A., and Webb B. D. 1997. Inheritance of long grain rice amylograph viscosity characteristics. Euphytica 97:25-29.

Halick, J. V., Beachell, H. M., Stansel, J. W., and Kramer, H. H. 1960. A note on the the determination of gelatinization temperatures of rice varieties. Cereal Chemistry 37:670-672.

Halick, J. V., and Kelly, V. J. 1959. Gelatinization and pasting characteristics of rice varieties as related to cooking behavior. Cereal Chemistry 36:91-98.

Han, Y., Xu M., Liu, X., Yan, C., Korban, S. S., and Chen, X., Gu, M. 2004. Genes coding for starch branching enzymes are major contributors to starch viscosity characteristics in rice (*Oryza sativa* L.). Plant Sciences 166: 357-364.

He, P., Li, S. G., Qian, Q., Ma, Q. Y., Li, Z. J., Wang, M. W., Chen, Y. and Zhu, L. H. 1999. Genetics analysis of rice grain quality. Theoretical and Applied Genetics 98:502-508.

Heuer, S., and Miezan K. M. 2003. Assessing hybrid sterility in *Oryza glaberrima* x *O-sativa* hybrid progenies by PCR marker analysis and crossing with wide compatibility varieties. Theoretical and Applied Genetics 107:902-909.

Hirano, H. Y., Eiguchi, M., and Sano, Y. 1998. A single base change altered the regulation of the Waxy gene at the posttranscriptional level during the domestication of rice. Molecular Biology and Evolution 15:978-987.

Hirano, H. Y., and Sano, Y. 1998. Enhancement of Wx gene expression and the accumulation of amylose in response to cool temperatures during seed development in rice. Plant and Cell Physiology 39:807-812.

Hirano, H. Y., and Sano, Y. 1991. Molecular characterization of the waxy locus of rice (*Oryza sativa* L.). Plant and Cell Physiology 32:989-997.

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

Hizukuri, S., Takeda, Y., Maruta, N., and Juliano, O. B. 1989. Molecular structures of rice starch. Carbohydrate Research 189:227-235.

Huawu, J., Weimin, D., and Ping, W. 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.

Isshiki, M., Morino, K., Nakajima, M., Okagaki, R. J., Wessler, S. R., Izawa, T., and Shimamoto, K. 1998. A naturally occurring functional allele of the rice waxy locus has a GT to TT mutation at the 5 ' splice site of the first intron. The Plant Journal 15:133-138.

Jones M. P., Dingkuhn, M., Aluko, G. K., and Mande, S. 1997. Interspecific *Oryza sativa x Oryza glaberrima* S. Progenies in upland rice improvement. Euphytica 92:237-246.

Juliano, B. O. 1998. Varietal impact on rice quality. Cereal Foods World 43:207+

Juliano, B. O. 1996. Rice functional properties and rice food products. Food Reviews International 12:71.

Juliano, B. O. 1993. Grain Quality Evaluation: International Rice Research Institute, Manila, Philipines.

Juliano, B. O. 1985. Cooperative tests on cooking properties of milled rice. Cereal Foods World 30:651-656.

Juliano, B. O., and Perez, C. M. 1983. Major factors affecting cooked milled rice hardness and cooking time. Journal of Texture Studies 14:235-243.

Juliano, B. O., Perez, C. M., Blakeney, A. B., Castillo, D. T., Kongseree, N., Laignelet, B., Lapis, E. T., Murty, V. V. S., Paule, C. M., and Webb, B. D. 1981. International cooperative testing on the amylose content of milled rice. Starch-Starke 33: 157-162.

Juliano, B. O. 1971. A simplified assay for milled-rice amylose. Cereal Sci. Today 16: 334-340.

Juliano, O. B., Cagampang, L. J., Cruz, J. L., and Santiago, G. R. 1964. Some physicochemichal properties of rice in southeast Asia. Cereal Chemistry 41:275-286.

Kumar, I., and Khush, G. S. 1987. Genetic-analysis of different amylose levels in rice. Crop Science 27:1167-1172.

Kuo, B. J., Hong, M. C., and Thseng, F. S 2001. The relationship between the amylographic characteristics and eating quality of japonica rice in Taiwan. Plant Production Science 4:112-117.

Laemmli, U. K. 1970. Cleavage of structural proteins during assembly of head of bacteriophage-t4. Nature 227: 680-685.

Larkin, P. D. and Park W. D. 2003. Association of waxy gene single nucleotide polymorphisms with starch characteristics in rice (*Oryza sativa* L.). Molecular Breeding 12:335-339.

Larkin P. D, McClung M. A., Ayres, M. N., and Park W. D. 2003. The effect of the waxy locus (Granule bound starch synthase) on pasting curve characteristics in specialty rices (*Oryza sativa* L.). Euphytica 131:243-253.

Larkin, P. D. 1999. Identification and post-transcriptional regulation of genes controlling grain quality in rice (*Oryza sativa* L.). Phd dissertation, Texas A&M University: College Station.

Larkin, P. L., and Park, W. D. 1999. Transcript accumulation and utilization of alternate and non consensus splice sites in rice granule bound starch synthase is temperature-sensitive and controlled by a single nucleotide polymorphism. Plant Molecular Biology 40:719-727.

Li, J. M., Xiao, J. H., Grandillo, S., Jiang, L., Wan, Y., Deng, Q., Yuan, L., and McCouch, R. S. 2004. QTL detection for rice grain quality traits using an interspecific backcross population derived from cultivated Asian (*Oryza sativa* L.) and African (*Oryza. glaberrima* S.) rice. Genome 47:697-704.

Lii, C. Y., Chang, S. M., and Yang, H. L. 1986. Correlation between the physicochemichal properties and eating quality of milled rice in Taiwan. Bull. Inst. Chem., Academia Sinica. 33:55-62.

Liu, X., Gu, M., Han, Y., Ji, Q., Lu, J., Gu, S., Zhang, R., Li, X., Chen, J., Korban, S. S., and Xu, M. 2004. Developing gene-tagged molecular markers for functional analysis of starch-synthesizing genes in rice (*Oryza sativa* L.). Euphytica 135:345-353

Little, R. R., Hilder, G. B., and Dawson, E. H. 1958. Differential effect of dilute alkali on 25 varieties of milled white rice. Cereal Chemistry 35:111-126.

Mazurs, E. G., Schoch, T. J, and Kite F. E. 1957. Graphical analysis of brabender viscosity curves of various starches. Cereal Chemistry 34:141-157.

McClung, A. M., Bergman, C. J., Fjellstrom, R. G., Bormans CA, Park, W. D., and Marchetti, M. A. 2004. Registration of 'Bolivar' rice. Crop Science 44: 353-355

McClung A. M., Marchetti M. A., and Webb B. D. 1998. Registration of Dixiebelle rice. Crop Sciences 38: 898.

McKenzie, K. S. and Rutger, J. N. 1983. Genetic analysis of amylose content, alkali spreading score and grain dimension in rice. Crop Sciences 23:306-313.

Ohtsubo, K., Toyoshima, H. and Okadome, H. 1996. Quality evaluation of rice in Japan. in: Crop Research in Asia. Achievements and Perspective. R. Ishii and T. Horie eds. Proc. 2nd Asian Crop Sci. Conf., Fukui, Japan: 94-97.

Olufowote, J. O., Xu, Y., Chen, X., Park, W. D., Beachell, H. M., Dilday, R. H., Goto, and M, McCouch, S. R. 1997. Comparative evaluation of within-cultivar variation of rice *(Oryza sativa* L.) using microsatellite and RFLP markers. Genome 40:370-378.

Pace, C. N., Hebert, E. J., Bechert, J., Shaw, K., Sevcik, J., Urbanikova, L., Scholtz, J. M., and Horn, G. 2001. Tyrosine hydrogen bonds make a large contribution to protein stability. J. Mol. Biol. 312:393-404.

Perez, C. M. and Juliano, B. O. 1979. Indicators of eating quality of non-waxy rice. Food Chem. 4:185-195.

Pooni, H. S., Kumar, I., and Khush, G. S. 1993. Genetic-control of amylose content in a diallele set of rice crosses. Heredity 71:603-613.

Ring, G. S., and Stainsby, R. G. 1982. Filler reinforcement of gels. Progr. Food Nutr. Sci. 6:323-329.

Sano, Y., Katsumata, M., and Okuno, K. 1986. Genetic-studies of speciation in cultivated rice .5. interspecific and intraspecific differentiation in the waxy gene-expression of rice. Euphytica 35:1-9.

Sano, Y, Maekawa, M., and Kikuchi, H. 1985. Temperature effects on the wx protein level and amylose content in the endosperm of rice. Journal of Heredity 76:221-222.

Sano, Y. 1984. Differential regulation of waxy gene-expression in rice endosperm. Theoretical and Applied Genetics 68:467-473.

Sato H., Suzuki, Y., Sakai, M. and Imbe, T. 2002. Molecular characterization of Wx-mq, a novel mutant gene for low-amylose content in endosperm of rice (*Oryza sativa* L.) Breeding Science 52: 131-135.

Schwartz D., and Echt C. S. 1982. The effect of Ac dosage on the production of multiple forms of wx protein by the wxm-9 controlling element mutation in maize. Molecular and General Genetics 187:410-413.

Semon, M., Nielsen, R., Jones P. M., and McCouch, R. S. 2005. The population structure of African cultivated rice *Oryza glaberrima* (Steud.): evidence for elevated levels of the LD caused by admixture with *O. sativa*. Genetic 169: 1639-1647.

Shu, Q. Y., Wu, D. X., Xia, Y. W., Gao, M. W., Ayres, N. M., Larkin, P. D. and Park, W. D. 1999. Microsatellite polymorphism on the waxy gene locus and their relationship to amylose content in indica and japonica rice, *Oryza sativa* L. Acta Genetica Sinica 26: 350-358.

Shu, Q., Wu, D., Xia, Y., Gao, M., and McClung, A. 1998. Relationship between RVA profine character and eating quality in *Oryza sativa* L. Scientia Agriculture Sinica 31: 25-29.

Smith, A. M., Denyer, K., and Martin, C. 1997. The synthesis of the starch granule. Annu. Rev. Plant Physiol. Plant Molecular Biology 48:67-87

Srichuwong S, Sunarti TC, Mishima T, Isono N, Hisamatsu M. 2005. Starches from different botanical sources II: Contribution of starch structure to swelling and pasting properties. Carbohydr Polym. 62: 25-34.

Steeneken, M. A. P. 1989. Rehological properties of aqueous suspensions of swollen starch granules. Carbohydr. Polym. 11:23-42.

Suzuki Y., Sano Y., and Hirano H. Y. 2002. Isolation and characterization of rice mutant insensitive to cool temperatures on amylose synthesis. Euphytica 123:95-100.

Takeda K. 1988. Temperature response of amylose content in rice varieties of Hokkaido. Japanese Journal of Breeding 38:357

Takeda Y., Hizukuri, S., and Juliano, B. O. 1987. Structures of rice amylopectins with low and high affinities for iodine. Carbohydrate Research 168:79-88.

Tan, Y. F., and Zhang, Q. F. 2001. Correlation of simple sequence repeat (SSR) variants in the leader sequence of the waxy gene with amylose content of the grain in rice. Acta Botanica Sinica 43:146-150.

Tan, Y. F., Li, X. J., Yu, B. S., Xing, Z. Y, Xu, G. C. and Zhang Q. 1999. The three important traits for cooking and eating quality of rice grains are controlled by a single locus in an elite rice hybrid, Shanyou 63. Theoretical and Applied Genetics 99: 642-648.

Tomita, K., and Nkagahra M. 1990. Genetic control for low amylose content content of rice endosperm found in indogenous Asian cultivars. Jpn. J. Breed. 40:248.

Traore, K., Fjellstrom, R. G., and McClung, A. M. 2004. Genetic diversity among West African rice varieties for grain quality traits using chemical and DNA marker analyses. Proc. 30th Rice Tech. Working Group, New Orleans, LA..

Umeda, M., Ohtsubo, H. and Ohtsubo, E. 1991. Diversification of the rice waxy gene by7 insertion of mobile DNA elements into introns. Jpn. J Genet 66:569-586.

Umemoto, T., and Terashima, K. 2002. Activity of granule bound starch synthase is an important determinant of amylose content in rice endosperm. Fund. Plant. Biol 29:1121-1124.

Viguier, P. 1939. La riziculture indigene au soudan français. Larose: Paris.

Villareal, C. P. and Juliano, B. O. 1989. Comparative levels of waxy gene-product of endosperm starch granules of different rice ecotypes. Starch-Starke 41:369-371.

Wang, Z. Y., Zheng, F. Q., Shen, G. Z., Gao J. P., Snustad, D. P., Li, M. G., Zhang, J. L., and Hong, M. M., 1995. The amylose content in rice endosperm is related to the post-transcriptional regulation of the waxy gene. The Plant Journal 7: 613-622.

Walker, G., Dobo, M., and Park, W. D. 2006. Unpublished data. College Station, Texas.

Walker, G., and Park, W. D. 2000. Unpublished observations. College Station, Texas.

Watanabe, H., Futakuchi, K., Jones, M. P., Teslim, I., and Sobambo, B. A. 2002. Brabender viscogram characteristics of interspecific progenies of Oryza glaberrima steud and *Oryza sativa L.* Nippon Shokuhin Kagaku Kogaku Kaishi 49:155-165.

Webb, B. D., 1991. Rice quality and grades. In: Rice. Vol. II. Utilization (2nd Ed.), pp. 89-119. AVI Books: New York .

Webb B. D. 1985. Criteria of rice quality in the United States. Pages 403-442 in: Rice: Chemistry and Technology. B. O. Juliano, ed. Am. Assoc. Cereal Chemistry: St. Paul, MN.

Wesley, I. J., Osborne, B. G., Anderssen, R. S., Delwiche, S. R., Graybosch, R. A. 2003. Chemometric localization approach to NIR measurement of apparent amylose content of ground wheat. Cereal Chemistry 80: 462-467.

West Africa Rice Development Association (WARDA). 2000. Participatory varietal selection. Bouake, Cote D'Ivoire

Yano, M., Okuno, K., Satoh, H., and Omura, T. 1988. Chromosomal location of genes conditioning low amylose content of endosperm starches in rice, *Oryza sativa* L. Theoretical and Applied Genetics 76:183-189.

VITA

Macaire Dobo completed his B.S. and M.S. degrees in 1991 and 1994 respectively at the University of Abidjan. He worked as a high school lecturer of biological sciences for two years and as research assistant at West Africa Rice Development Association (WARDA) for four years. He joined Texas A&M in 2002 to begin his Ph.D. program on developing rice grain quality molecular markers in the Department of Crops and Soil Sciences.

Mr. Dobo may be reached through Glenda Kurten in the department of Crops and Soil Sciences, MEPS Program, 2474 TAMU, College Station, TX 77843-2474. His email address is mccaire2006@yahoo.fr.