Profiling the Expression of Grain Quality Related Genes in Developing Hybrid Rice Seeds

A thesis submitted to the graduate division of the Chinese University of Hong Kong in partial fulfillment of the requirements for the degree of

Master of Philosophy

in

Biology

By

Duan Meijuan

August 2003

Thesis committee

Prof. S. S. M. Sun (Supervisor)

Prof. H. M. Lam (Internal examiner)

Prof. L. W. Jiang (Internal examiner)

Prof. N. S. Wong (External examiner)

Statement

All the experimental work reported in this thesis was performed by the author, unless specially stated otherwise in the text.

Signed----------

Duan Meijuan

Acknowledgements

First of all, I would like to express my sincere gratitude and respect to my supervisor, Prof. S. S. M. Sun, both for giving me a precious opportunity to learn and study in the field of plant molecular biology and for his patient guidance, valuable advice, considerate concern, and great emotional encouragement throughout my study in Hong Kong.

I would like to express appreciation to Prof. H.M. Lam and Prof. L.W. Jiang for being my internal examiners and their valuable suggestions, critical comments, great support and encouragement. I am also thankful to my external examiner, Prof. N. S. Wong for his patience to review my thesis and comments.

Besides, I would like to express my thanks to my labmates (Gundam Wong, Joun-Juon Lee, Dr. Chen Daming, Dr. Tu Jumin, Dr. Deng Qiyun, Dr. Wu Xiaojin, Lau On Sun, Stanley Cheung, Kenneth Chik, Walter Wong, Wendy Chan, Eric Suen, Jennifer Wai, Ruby Ng, Kingston Xiao, Abby Yu, Avon Li, Marilyn Yu and Vivian Li) for their research suggestions and technical advice.

In addition, my appreciation should also go to my dear friends (Dr. Liu Qiaoquan, Dong biao, Angela Yu, Zhao min, Robin Li, Zeng Naiyan, Qi Wei, Rainbow Phang, Helen Tsai) for their kind help, support, caring, encouragement and available

suggestions.

Particularly, my special gratitude should go to my dearest husband, Yuan Dingyang. His endless love, unconditional support, spiritual encouragement and experimental discussion gave me strength to finish my project.

Last but no least, I would like give my deepest appreciation to my dearest parents, parents-in-law, elder sister and elder brother for their love, everlasting support and eternal caring in my life.

Abstract

Rice is one of the most important crops in the world and a primary food for more than half of the world's population, particularly in the developing countries in Africa and Asia. China is the first country in the world to exploit the heterosis in rice. The success of the hybrid rice led to a major breakthrough in rice production, providing an effective approach to increase rice yields on a large scale. The grain quality of super hybrid rice, however, is not ideal. Improvement of grain quality, particularly in nutritional and eating and cooking quality, is thus a most important research priority in China, as the living standards of the country are improving.

Protein and starch are the two major substances in rice grains, accounting for 5-10% and more than 80% by dry weight, respectively. According to solubility, rice seed storage proteins are classified into four types: glutelin, prolamin, albumin, and globulin. These storage protein fractions are encoded by multigene families. Starch plays a central role in the cooking and eating quality of rice, which consists of two types of polysaccharides, amylose and amylopectin. Amylose synthesis is catalyzed by granule-bound starch synthase (GBSS), whereas starch branch enzyme (SBE) mainly controls the synthesis of amylopectin. Moreover, ADP–glucose pyrophosphorylase (AGPase), soluble starch synthase (SSS) and starch debranching enzyme (SDE) also play important roles in the process of starch synthesis.

Many aspects of the biology of rice seeds have been studied including gene cloning, transformation and regulation at DNA, RNA and protein levels. However,

there is no systematic profiling of the expressed genes involved in grain quality in the maturing rice seeds, especially hybrid rice.

In this study, the parental lines and hybrid F1 were used as research materials to study the gene expression patterns during seed development of super hybrid rice. The maturing seeds from 3, 6, 10, 15 and 20 days after fertilization were collected. Profiling the expression of genes encoding the four rice storage proteins, starch synthesis enzymes and key enzymes involved in lysine synthesis in the aspartate family pathway were performed by northern blot and RT-PCR. A total of 36 major genes related to grain quality have been studied and their expression can be grouped into several distinct patterns, based on the observed variations in expression time and level. Most genes show high expression levels at mid maturation stage (10 DAF) before declining at later stages. Some genes are active mainly during early development stage (6-day-old seeds) while others at late maturation (from 10 DAF to 20 DAF), with an expression peak at 15 DAF. Yet a few genes exhibit strong expression at very early stage (3 DAF). The heredity of the expression level of F1 hybrid generally showed four types, the same, higher, in between and lower than those of their parents. Another observation is that the expression level of female parental line (P64S) is generally higher than that of the male parental line (9311), based on the expression patterns of most of the genes under current study.

Expression patterns of the four main storage protein genes were also studied at protein level by SDS-PAGE and 2-D gel electrophoresis. Results from the SDS-PAGE analysis showed that there are variations in the synthesis and accumulation of the four

storage proteins in the parental lines and F1 hybrid. In female parental line, the four storage proteins generally appear at 6 DAF and reach the highest level at 15 DAF; for the male parental line, proteins appear at 6 DAF and remain accumulating in later stages; whereas accumulation of proteins in F1 is under various degrees of influence by their parents. The expression patterns by 2-D gel also revealed clear variations among protein species between the female and male parental lines and their hybrid, and these proteins are mainly of the 13-KD proteins with pIs between 5 and 9. In addition, variations in the expression profiles between mRNA and protein levels were also observed, demonstrating the importance of proteomics studies.

This study systematically dissected the expression profiles of the major 36 genes involving in determining the grain quality during hybrid rice seed formation. The results revealed a complex network of genes and their expression during rice seed formation. The complexity is manifested by the diverse, variations, and interrelated functions of these seed quality genes, usually of multigene families in nature. This information and the available genetic materials (genes) will provide an entry into the functional genomics of rice grain quality, allowing our understanding of how this important seed quality trait is controlled and our efforts towards its manipulation in the future.

摘要

水稻是世界上最重要的作物之一，也是世界上半數以上人口尤其是亞洲和非洲人民的主要食物來源. 中國是世界上第一個將雜種優勢應用於水稻的國家. 雜交水稻的成功帶來了水稻生產的巨大突破，它提供了一條在大面積上提高水稻產量的有效途徑. 然而，雜交水稻的米質不是太理想. 隨著人們生活水平的不斷提高，改良和提高稻米的品質尤其是其營養和蒸煮食味品質日益受到重視.

蛋白質和澱粉是稻米的兩大主要組成部分，分別占米粒幹重的 5-10%和 80%以上. 根據可溶性的不同，水稻種子貯藏蛋白分爲四大類: 穀蛋白，醇溶蛋白，白蛋白和球蛋白. 這四種水稻種子貯藏蛋白均由多基因家族編碼. 澱粉對稻米蒸煮和食味品質起著舉足輕重的作用. 它由直鏈澱粉和支鏈澱粉兩種結構不同的多糖組成. 直鏈澱粉經顆粒凝結型澱粉合成酶後(GBSS) 催化形成，而澱粉分支酶 (SBE)主要參與支鏈澱粉的合成. 此外，ADPG 焦磷酸化酶 (AGPase)，可溶性澱粉合成酶 (SSS)和澱粉去分支酶 (SDE)在澱粉合成過程中也起到非常重要的作用. 水稻種子生物學的許多方面都做了研究，包括基因克隆，轉化和在DNA,RNA和蛋白質水平的調控机制. 然而，對於系統地研究與稻米品質相關的基因在水稻種子，尤其是雜交水稻種子形成過程中的表達特性還沒有先例.

本研究以超級雜交稻的親本和 F1 雜種一代爲材料，研究超級雜交稻種子形成過程中基因的表達模式. 親本和雜種 F1 代受精後 3, 6, 10, 15, 20 天的未成熟種子採集後，用 northern 雜交和逆轉錄 PCR 分析四種種子貯藏蛋白，澱粉合成酶和天冬氨酸家族賴氨酸合成途徑的關鍵酶基因的表達模式. 總共 36 個相關基因被研究，結果顯示這些基因的表達模式依據表達時間和表達水平的差異可分爲幾種類型. 大部分基因在受精後第 10 天達到最高水準，之後逐漸減弱；一些基因在

種子形成早期(受精後第 6 天)表現活躍, 之後逐漸減弱; 一些基因主要在種子形成後期(受精後第 10 天至 20 天)表達, 伴隨著受精後第 15 天出現表達高峰. 此外, 有少數基因的最高表達水平出現在種子形成的最早期(受精後第 3 天). 而且, 基因表達水平的遺傳在雜種中表現出相等, 高於, 居中和低於雙親表達水平的四種類型. 大部分基因在母本(P64S)中的表達水平高於在父本(9311)中的表達水平.

同時, 通過十二烷基硫酸鈉-聚丙烯酰胺凝膠電泳 (SDS-PAGE)和雙向凝膠電泳 (2-D gel electrophoresis)對四種種子貯藏蛋白的表達也做了蛋白质水平的研究. SDS-PAGE 的分析結果表明四種種子貯藏蛋白在親本和雜種 F1 代的積累模式是不同的. 在母本中, 這四種種子貯藏蛋白均在受精後第 6 天開始積累到受精後第 15 天達到沈積高峰, 受精後第 20 天有所下降; 在父本中, 這四種種子貯藏蛋白從受精後第 6 天到受精後第 20 天都呈現出持續累積的模式; 而種子貯藏蛋白在雜種 F1 代的表達受到雙親不同程度的影響. 2-D gel 的結果亦揭示種子貯藏蛋白在兩個親本和雜種中的表達存在明顯不同. 這些蛋白質主要爲分子量在 13kDa 和等電點位於 5 到 9 的蛋白質. 另外, 所觀察到的這些基因的表達模式在 mRNA 與蛋白質水平的差異性亦表明了蛋白組學研究的重要性.

本研究系統地剖析了 36 個主要與稻米品質有關的基因在雜交水稻種子發育過程中的表達模式, 結果揭示了這些基因在決定稻米品質的過程中形成了一個錯綜複雜的網路. 這種複雜性是由於這些基因在自然界中通常以多基因家族存在, 而且這些基因之間存在著多種多樣的相互關係. 本研究所獲得的有用的資訊和遺傳材料(基因)將成爲稻米品質功能基因組學研究的起點, 從而進一步瞭解和調控這些重要的種子特性, 以實現提高和改良雜交水稻的營養和蒸煮食味品質的目標.

List of Tables

List of Figures

List of Abbreviations

AA	Amino acid
ABA	Abscisic acid
ACHCA	alpha-cyano-4-hydroxy cinnamic acid
ACN	Acetonitrile
ADP	Adenosine diphosphate
ADP-G	ADP-glucoase
AGPase	ADP-Glc pyrophosphorylase
AK	Aspartate kinase
A line	Male sterile line
Anti-DIG-AP	Anti-Digoxigenin-alkaline phosphatase
ATP	Adenosine triphosphate
BiP	Binding protein
BLAST	Basic local alignment search tool
B line	Maintainer line
bp	Base pair
cDNA	Complementary DNA
CDS	Coding sequences
C-ER	Cisternal ER
CMS	Cytoplasmic male sterility
CSIT	Critical sterility inducing temperature
CSPD	Disodium 3-(4-methoxyspirol{1,2-dioewtane-3,2'-(5'-chloro)tricyclo[3.3.1.1^{3,7}]decan}-4-yl) phenyl phosphate
Cys	Cysteine
2-D gel	Two-dimension eletrophoresis gel
DAF	Days after fertilization
DEPC	Diethylpyrocarbonate
DHPS	Dihydrodipicolinate synthase
DI	Deionized
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DP	Degree of polymerization
DTT	Dthiothreitol
EtBr	Ethidium bromide
EDTA	Ethylenediaminetetra-acetic acid
ESI	Electrospray ionization
ER	Endoplasmic reticulum
EST	Expressed sequence tag
G-1-P	Glucose-1-phosphate

g	Gram
Gb	Globulin
GBSS	Granule-bound starch synthase
Gln	Glutamine
GluA	Glutelin A
GluB	Glutelin B
GT	Gelatinization temperature
GUS	β-glucuronidase
ha	Hectare
HHRRC	Hunan Hybrid Rice Research Center
His	Histine
hr	Hour
HSD	Homoserine dehydrogenase
HVR	Hypervarable region
IAA	Iodoacetamide
IEF	Isoelectric focusing
IgE	Immunoglobulin E
IPG	Immobilized pH gradient
IRRI	International Rice Research Institute
kb	Kilobase
kDa	Kilodalton
L	Liter
Leu	Leucine
L-type	Long-type
LMW	Low molecular weight
LS	Large subunit
Lys	Lysine
MALDI-ToF	Matrix assisted laser desorption ionization-time of flight
MB	Membrane-bound
Met	Methionine
min	Minute
ml	Milliliter
mM	Millimole
M-MLV	Moloney murine leukemia virus
MOPS	3-(N-Morpholino) Propanesulfonic acid
mRNA	Messenger RNA
MS	Mass spectrometry
MW	Molecular weight
N	Nitrogen
NB	Northern blot
NCBI	The National Center for Biotechnology Information
ng	Nanogram
P64S	Pei'ai 64S
P64S/9311	Pei'ai 64S cross 9311

3-PGA	3-phosphoglycerate
PAGE	Polyacrylamide gel electrophoresis
PB I	Protein body I
PB II	Protein body II
PB-ER	Protein body-forming ER
PCR	Polymerase chain reaction
PPi	pyrophosphate
PS	Position shift
PSD	Post-source decay
PSLs	Position shift loci
PTGMS	photo- and thermo-sensitive genic male sterility
R line	Restorer line
RNA	Ribonucleic acid
RNase	Ribonulease
rpm	Revolution per minute
rRNA	Ribosome RNA
RT-PCR	Reverse transcription-polymerase chain reaction
SBE	Starch branching enzyme
SDE	Starch debranching enzyme
SDS	Sodium dodecyl sulfate
sec	Second
Seq.	Sequence
SNP	Small nuclear protein
SP	Signal peptide
SS	Small subunnit
SSS	Starch soluble synthase
S-type	Short-type
TFA	Trifluoroacetic acid
Tris	Tris(hydroxymethyl)amminomethane
TRRM	Terminator ready reaction mix
U	Unit
ug	Microgram
ul	Microliter
UTR	Untranslational region
vac	Vaccum
vhr	Voltage with hour
VR	Variable region
v/v	Volume by volume
W	Watt
WC	Wide compatibility
WHO	World Health Organization
w/v	Weight by volume

Tables of Contents

Chapter 1 General introduction

Rice is the second most important staple food for the world population following wheat. 95% of the world's rice is grown by developing countries, mostly in Asia (IRRI, 1995). The production and consumption of rice remain stable in Asia over the past 30 years (Duff, 1991). In order to increase rice yield for the needs of the increasing population, hybrid rice was developed first in China in 1964. Hybrid rice normally has a yield advantage of 20-30% over the non-hybrid rice cultivars. From 1976 to 1997, hybrid rice technology enabled rice production to increase by more than 312 million tons. Recently, record yields of 11.2 t/ha from the combination *Liangyou Peite* in a single hybrid crop on a large scale (1,000 ha) and of 17.1 t/ha in a small plot (0.1 ha) were reported (Bai and Luo, 1996). Another super hybrid combination, *Liangyou Peijiu*, has been maintaining an average yield of 12 t/ha in large scale for the last several years. The breeding of elite hybrid rice combinations has made great progress in China, however, research on the grain quality of hybrid rice is lacking.

Rice grain is the main organ of a rice plant and the product people consume because of its abundant nutrients for human diet including starch, protein, sugar, fiber, lipid, vitamins and minerals etc. Thus, rice seeds play a significant role in supporting the life of people who use rice as their staple food.

Rice grain quality is measured by its appearance quality, milling quality, cooking and eating quality as well as nutritional quality. Nutritional quality is largely determined by the quality and quantity of seed proteins and their amino acid

composition. Cereal proteins, however, are deficient in the essential amino acids lysine and tryptophan. Several methods to improve the amino acid balance of the cereal crops were proposed including transgenic expression of homologous or heterologous genes encoding essential amino acid-rich proteins and modification of protein sequence (Tu, 1998; Hoffman et al., 1987; Shaul and Galili, 1992; Cheng, 1999 and Liu, 2002). The cooking and eating quality of cereal grains is largely determined by starch content and by amylose to amylopectin ratio. Regulation of starch synthesis has been an area of research for many scientists, including attempts to increase or decrease the starch content, and to modify the amylose to amylopectin ratio through sense and antisense RNA technology (Stark et al., 1992; Müller-Rober et al., 1992 and Visser et al., 1991).

However, systematically profiling and analyzing the many important genes determining the grain quality in maturing hybrid rice seeds is lacking. Information on how the rice seeds are formed, how their important ingredients (e.g. storage proteins and starch) are laid down, what and how genes collectively control their synthesis and deposition during grain development, and how these genes pass on to the next generation hybrid rice seeds will be critical in our effort to improve the quality of rice.

This research aims at profiling the expression of major grain quality related genes for seed storage proteins, starch synthesis enzymes, and key enzymes involved in lysine synthesis in the aspartate amino acid pathway during grain development in elite hybrid rice lines. We hope to provide this basic information in our effort to seek new and more efficient ways to improve the quality of rice in the future.

Chapter 2 Literature Review

2.1 Hybrid rice: Genetics and breeding

Rice is the most important crop for more than half of the world population, providing various essential nutrients for human diet, especially for people in developing country including poverty countries in Asia and Africa. To maintain their basic diet requirements, increasing output of rice is one of the most important means. The success achieved in the development of hybrid rice is a major breakthrough in rice breeding and a technological innovation in rice production, providing an effective approach to increase rice yields on a large scale.

China is the first country in the world to utilize heterosis in rice to produce hybrid rice seeds for commercial use. Research work on hybrid rice was originated in 1964 (Yuan, 1966) when Yuan observed a nature-crossing hybrid rice plant with tall and greater vegetative vigor in the rice field in Anjiang Agricultural School, Qianyang County, Hunan Province, China. However, hand emasculation is a limitation in producing commercial F1 hybrid seeds on a large scale. Until in 1973, the discovery of the wild abortive (WA) male sterile cytoplasm in the wild species *(O. rufipogon Griff* or *O. sativa f. Spontanea*) in Hainan Island provided a valuable pathway to exploit heterosis in rice. Therefore, the three-line system including male sterile (A line), maintainer (B line) and restorer (R line) was successfully developed in 1973 (Yuan and Virmani, 1988). The first hybrid combination with great heterosis was fulfilled in 1974, and the corresponding F1 hybrid seed production techniques were

established in 1975. Since then, hybrid rice production has gained popularity in large areas. The planted area of hybrid rice in China has been increasing year after year, from 0.14 million ha in 1976 to 15.3 million ha in 1997. Moreover, the yield of hybrid rice is also going up steadily. It has been proven practically that hybrid rice varieties have about 20% yield advantage over the best conventions (Lin and Yuan, 1980; Shen, 1980). In 1998, the super hybrid combination (P64S/9311) was successfully developed, its yield is 15-20% higher than that of the general hybrid combination.

However, the yield level of the three-line hybrid rice has obtained a plateau since 1980s (Yuan, 1994d and 1997a) because more than 85% of all A lines originated from the single wild abortive (WA) male sterile cytoplasm type and their yield potentials had been extensively used. In addition, this single cyto-sterility system may be vulnerable to destructive pests or diseases. To further improve output of hybrid rice and enhance heterosis utilization in rice Yuan (1987) proposed three breeding strategies at three levels: three-line system at intercultivar hybrids level to two-line system [PTGMS (photo-thermo sensitive genic male sterility line) and restorer line] at intersubspecific hybrids level to one-line system (apomictic system) at distant hybrids level (interspecific hybrids).

In the present study, P64S, 9311 and their F1 hybrid (P64S/9311) were used as experimental materials. P64S is maternal line with PTGMS for hybrid production (p64S/9311), which belongs to *javanica* subspecies (Yuan 1997a) and its sterile gene was transferred from Nongken 58S [the first PTGMS line discovered by Shi in 1973 (Shi 1981; 1985; Shi and Deng, 1986)]. The critical sterility inducing temperature

(CSIT) of P64S is 23.5 °C, that is to say, P64S will convert into fertile if the temperature is below 23.5 °C, at its sterility sensitive stage.

Line 9311 is the paternal line for hybrid production (P64S/9311). It belongs to *indica* subspecies and is a major planting variety in South China. It has high rice yield [14.4% over control (Shanyou 63)], good grain quality (first class standard) and multiple resistances to blast, blight and lodging (Dai and Zhao et al., 1997).

P64S/9311 is the super hybrid rice combination, commercially named LYP9, which is the typical two-line intersubspecies hybrid with super high yield at 12 tons/ha, 15-20% higher than other hybrids. Furthermore, the rice grain quality of LYP9 reached second-class national standard and the morphology is excellent.

2.1.1 Classification in rice

There are about 20 species in rice, of which *Oryza. sativa* (the Asian cultivated species) and *Oryza. Galberrima* (the West Afirican cultivated species) are the two cultivated species. According to morphological and ecological characteristics *Oryza. Sativa* can be divided into two subspecies: *indica* and *japonica* (Kato et al., 1990). *Indica* has compact tillers and greater height with longer, slender grain and generally grow in the tropical and subtropical regions such as Southern China, India, Thailand and Vietnam. By contrast, *japonica* has cluster tillers, short, plump grain and is grown in temperate climates such as Northern China, Japan, Korea, Australia and California. However, some scientists suggested that there should be a third rice group, the Indonesian cultivars with tall and long-panicles, named *javanica* by Körnicke (1885)

and tropical *japonica* by Oka (1983). Moreover, Morinaga and Kuriyama considered *javanica* an intermediate type (1958).

2.1.2 Heterosis in rice

2.1.2.1 Performance of heterosis in rice

In 1865 Mendel observed hybrid vigor in his pea experiments. Whereas the term of heterosis was first created by Shull in 1914 (Shull, 1952). Up to now, the term heterosis refers to the phenomenon in which the F1 population derived from the crossing of two genetically dissimilar parents shows superiority to both parents in growth vigor, vitality, reproductive capacity, stress resistance, adaptability, grain yield, grain quality and other traits. The commercial utilization of heterosis by developing and planting F1 hybrids so as to increase economic returns is called exploitation of heterosis in crops (Yuan, 1985). Hybrid corn is the first crop developed with heterosis utilzation in the 1930s (Pingali et al., 1997). Since then, Jones and Clarke (1943) invented the CMS system for hybrid onion production, which provided an important strategy for developing the three-line hybrid rice by agronomists in China.

Generally speaking, heterosis in rice appears in many morphological and physiological features including vegetative heterosis, reproductive heterosis and resistance heterosis to unfavorable environments. For vegetative heterosis, F1 hybrids appeared to have a stronger root system including the numbers (Li et al., 1982), distribution (Yuan and Chen, 1988; Bai and Xiao, 1988; Lu et al., 1988), support and absorb activity, higher lodging resistance (Yuan and Chen, 1988), larger leaf area (Li

et al., 1982; Singh, 1997), efficient nutrient transportation, lower respiration intensity (Lin and Yuan, 1980) as well as higher photosynthetic efficiency (Li, Wang and Liu, 1990). Reproductive heterosis was apparent by the larger panicles, heavier grains (Zeng et al., 1979) and higher yield (Sun and Cheng, 1994) even longer growth duration (Song et al., 1990). For resistance heterosis, several researches showed that F1 hybrids have excellent resistance to diseases, pests, drought (Yab and Chang, 1976; Tian et al., 1980), cold, barren soil, saline soil (Akbar and Yabuno, 1975) and marshland. The hybrids thus possess greater adaptability to conquer various adverse environmental conditions.

2.1.2.2 Genetic mechanism of heterosis in rice

Heterosis is a complex genetic and physiological phenomenon. The occurrence of heterosis is attributed to the differentiation at some degree of the sexual factors in both parents. However, the mechanism of its action is very complicated. Based on the previous theories on heterosis (Bruce, 1910; Gustafsson, 1946; Hull, 1945; Castle, 1946 and Hayes, 1952), Prof. Yuan put forward the assumption that heterosis in rice is caused by various effects of gene interaction including allelic nuclear genes interaction, nonallelic nuclear genes interaction and interaction between the nucleus and cytoplasm (Yuan, 1985).

In general, interaction between alleles in nucleus will induce two kinds of effects: dominance and over-dominance effects. For dominance effects, according to previous hypothesis (Davenport, 1908; Jones, 1917), available features are controlled by

dominant genes and the unavailable characteristics are controlled by recessive genes, and the F1 hybrid heterosis arises from inhibition of unavailable recessive genes effects and accumulation of dominant genes from two parents. For over-dominance effects, the F1 heterosis is based on heterogeneous characteristics of alleles in nucleus. It's assumed that the interaction effect between heterozygote is greater than that between homozygote. Once the two parents with heterozygous alleles are crossed, F1 hybrid combines the heterozygotes from both parents and will be more vigor than it parents (Shull, 1908). However, this hypothesis can not elucidate the reason why some traits in rice hybrid are inferior to their parental lines. Many researchers doubted that the over-dominance effect plays an important role in F1 heterosis (Crow, 1997).

The interaction between nonallelic genes at various loci and on different chromosomes is also a causal force for expression of heterosis. This interaction may divide into three effects: additive effects, which accumulate effects of genes determining the same trait; epistatic effects, which is the inhibition of nonallelic genes; and recombination effect. Yu et al. (1997) reported that there was little correlation between marker heterozygosity and trait expression, but digenic interactions frequently existed in the F3 progeny derived from "Zhenshan 97 x Minhui 63". This suggested that epistasis significantly affected the performance of heterosis in rice.

Besides heterosis is controlled by nuclear genes, the cytoplasmic genes and the interaction between nuclear genes and cytoplasmic genes also govern the expression of heterosis. This is the reason why F1 hybrid of reciprocal cross showed different expression of heterosis in the same combination and the same nuclear genome

incorporating into different cytoplasmic backgrounds also exhibits different levels of heterosis. Sheng (1987) reported that the eight cytoplasms negatively affected many features such as plant height, numbers of effective tillers, seed setting percentage, 1000-grain weight and yield.

Generally speaking, the effect of nuclear genes, in particular, interaction between nucleus alleles is the major force for F1 hybrid heterosis. The interaction between nonalleles in nuclear and dominance effects are related to specific combining ability and general combing ability, respectively.

2.1.3 Utilization of heterosis in rice

Up to now, the breeding approaches for ultilizaiton of heterosis in rice include three types: three-line method [cytoplasmic male sterility (CMS) system] at the intercultivar heterosis level, two-line method [photo- or thermo-sensitive genic male sterility (PTGMS)] at the intersubspecific heterosis level, and one-line method (apomictic system) at distant heterosis.

Three-line system contains the male sterile line (A line), maintainer line (B line) and restorer line (R line). The A line refers to lines whose anthers are abortive but pistils are normal and can fertilize with normal pollen. The B line is a variety with normal pollen and can produce progeny maintaining male sterility when crossed by A line, whose function is to retain the male sterility of A line. The R line is also normal pollen donor to A line and produce the hybrid with male fertile (normal anthers) and apply in practice by selfcrossing. Currently, most combinations are intercultivar

hybrids (*indica* by *indica* or *Japonica* by *Japonica*) by using three-line system, which lead to 20-30% higher yields than conventional rice varieties. However, the output of this kind of hybrids has reached the maximum level due to narrowing parental germplasm diversity (Luo and Yuan, 1990). Moreover, the breeding and seed production procedures of this hybrid type are complicated and more costly. For this reason, the three-line system is being replacing with simpler systems such as the two-line method. Two-line method includes only PTGMS line and R line because the PTGMS line can maintains the male sterility under specific conditions such as high temperature [> critical sterility inducing temperature (CSIT)] or longer light period (>14 hr). Intersubspecific hybrids derived from crosses between subspecies such as an *indica* and *japonica* cultivar possess strong yield heterosis with theoretical yielding 30-50% higher than the intercultivar hybrids. Further, the application of wide compatibility (WC) gene and subspecies *Javanica* had also made great progress to improve the utilization of interspecific heterosis (Ikehashi, 1982 and Yuan, 1991a,b). The two-line system is genetically controlled by nuclear genes and thus there is no negative effect from the cytoplasm and no risk of unilateral cytopalsmic breakdown. Thus there is more opportunity to develop elite rice hybrids using the two-line system than using the three-line system. Furthermore, omission of the B line used for maintaining male sterility in the three-line system reduces seed costs. Therefore, the two-line system is a powerful methodology to increase rice yield and to utilize heterosis in rice. In the future, a new method, one-line system, will be developed, which will result in no deterioration of heterosis with year-after year seed production

because apomixes is asexual reproduction via seeds without genetic segregation occurrence.

2.2 Grain quality in rice

2.2.1 Composition of rice grain quality

Rice grain quality includes appearance quality, milling quality, cooking and eating quality, and nutritional quality, which determines the rice economical value and consumption in international rice market.

2.2.1.1 Appearance quality

Rice appearance quality comprises many parameters including grain size, shape (length/width ratio), color, gloss, translucency and uniformity, which are used to assess rice grains putting into market as commercial products.

2.2.1.2 Milling quality

Brown rice becomes milled rice after hull, bran and germ removal. Milled rice is the commercial product consumed by human. To minimize the breakage of endosperm is a very important element in milling quality in rice market as the grain breakage during milling process is a serious problem in rice growing countries. Cracking and chalk area in rice is the two major causes of grain breakage. Cracking is easy to form when harvest is delayed or when too rapid drying of the crop takes place. Chalk, the layers of cells with loose packing of starch granules in rice grain, also leads

to easy breakage during milling process. The formation of chalkiness is controlled by genetic and environmental conditions during rice seed ripening. Li et al. (1989) reported that positive correlation was found between chalkiness and daily mean temperature, average daylight hours and the relative humidity in the atmosphere, of which daily mean temperature is the most sensitive environmental factor to influence rice grain quality. Other causes that contribute to milling quality include moisture content in rice grain, seed hardness and shape, seed infestation as well as the design of milling and harvesting machines. Many scientists had attempted to find a method to reduce breakage of rice grain until parboiling was discovered by Garibaldi in 1974. Parboiling, precooking rice grains before milling, not only enhances rice grain hardness but also seals internal cracks and chalkiness resulting in reducing grain breakage during milling. Moreover, Parboiling also leaches B group vitamins from the aleurone layers into the endosperm, increasing the nutritional value of the milled grain.

2.2.1.3 Nutritional quality

Nutritional quality of rice grain is a very important factor in human health, especially for people who use rice as staple, since the balance of nutrition is a requirement to maintain body development and growth. Therefore, protein content and composition of amino acid are two major elements in evaluation of rice nutrition quality. High protein would contribute to the nutritive value of rice as a food. Moreover, protein content correlates with cooking and eating quality.

Protein is the second most abundant storage matter following starch in rice endosperm. It accounts for 5-10% by dry seed weight, which is the lowest among cereals (Table 2.1). The protein content in brown rice of 17,587 cultivars collected all over the world by International Rice Research Institute (IRRI) ranged from 4.3% to 18.2% with a mean of 9.5% and most entries have 8% protein (Gomez, 1979). According to function, rice protein can be divided into two classes: storage proteins and housekeeping proteins, whose functions are as storage nutrition for seed germination and to maintain cell structure and metabolism for plant growth, respectively. Seed storage proteins are the major protein and can be sorted into four groups based on solubility, which are acidic or basic solution soluble glutelin, salt-soluble globulin, water-soluble albumin, and alcohol-soluble prolamin (Juliano, 1972), accounting for 80%, 10%, 5%, 5%, respectively (Chavan and Duggal, 1978). The distribution of different solubility groups in rice grain is different. In general, albumin and globulin are concentrated in the embryo and aleurone layer and decrease toward the center of the grain, while glutelin is the major fraction in the endosperm of rice grain. Moreover, the protein content of milling rice is lower than that of brown rice since the protein-rich husk, bran and aleurone layer are removed during the milling process.

The amino acid composition of the total storage protein in rice seed is relative balanced, especially with relatively higher lysine content comparing with the other cereals. But lysine content in rice grain is about 3.8 g/16.8 g N (Table 2.1), which is much lower than the recommended content (5.5 g/16.8g N) by WHO (World Health

Organization, 1973). Therefore, lysine is still the first limiting factor in rice quality improvement as with the other cereals. Liu et al. (2002) transformed rice with a winged bean lysine-rich protein gene and were able to significantly enhance the lysine content in transgenic rice. Threonine is the second limiting essential amino acid in rice and methionine and tryptophan are also lower than the WHO standards. Utilizable protein proportion and true digestibility of rice protein are also higher than the other cereals, especially its very high biological value (74.0). It is known that rice protein is easy to digest and absorb by human body, providing various nutrients to human diet.

In addition, some lipids, minerals and vitamins are accumulated in rice grains. Unsaturated fatty acids are the major fraction in rice lipid. It accounts for 2.6% by dry weight of brown rice, distributing mainly in embryo and outer layer while only small amount remains in the milled rice after the bran is removed during milling process. Vitamins are also generally present in higher levels in brown than in milled rice. Vitamin B is the major group and little or no vitamin A, C (ascorbic acid) and D present in the rice grains. Early researches reported that over 80% of the thiamine is in the scutellum plus aleurone layers, 11% in embryo and 8% in the endosperm (Juliano, 1985). Potassium, calcium, magnesium, phosphorus and silicon are present also in large amounts in brown rice as well as the microelements such as iron and zinc which are essential for human nutrition but are not readily available from vegetable source. The mineral composition of brown rice is accounted for around 0.9% by dry weight and varied considerably under various cultivars and soil nutrients during growth.

Knudsen et al. (1981) studied ten samples from Indonesia and Bangladesh and showed that the brown rice contains per gram 0.6-1.4 mg of potassium and silicon, 1.7-4.3 mg of phosphorus, 1.7-2.8 mg of magnesium, <15-65 μg of calcium, 103-254 μg of aluminum, 2.1-10.2 μg of iron and 6.0-12.1 μg of zinc.

Table 2.1 The protein content and nitrogen and energy balance data of cereal crops (Juliano, 1985).

Properties	Brown rice	Wheat	Corn	Barley	Millet	Sorghum	Rye	Oat
Protein (Nx6.25)%	7.3	10.6	9.8	11.0	11.5	8.3	8.7	9.3
Lysine g/16gN	3.8	2.3	2.5	3.2	2.7	2.7	2.7	4.0
True N digestibility %	99.7	96.0	95.0	88.0	93.0	84.8	77.0	84.1
Biological value %	74.0	55.0	61.0	70.0	60.0	59.2	77.7	70.4
Net protein utilization %	73.8	53.0	58.0	62.0	56.0	50.0	59.0	59.1
Utilizable protein %	5.4	5.6	5.7	6.8	6.4	4.2	5.1	5.5
Digestible energy %	96.3	86.4	87.2	81.0	87.2	79.9	85.0	70.6

Notes:

Nitrogen and energy balance studied in growing rats.

Net protein utilization: digestibility times biological value divided by 100

Utilizable protein: protein content times net protein utilization divided by 100

Biological value: absorbed protein amount per 100 g protein intaken

2.2.1.4 Cooking and eating quality

Amylose content and gelatinization temperature are the two most major physiochemical properties relating to rice cooking and eating quality. Starch is composed of 80-90% of dry substances in rice grain and amylopectin comprises over 80% of total starch. In fact, cooking process gelatinizes the rice starch. Amylose content has been shown to be very important in rice cooking quality. Also, it's reported that correlation was found between amylose and eating quality. Cooked rice with high amylose content is flaky, dry, hard and faint; by contrast, rice with low amylose content is sticky, moist, soft and glossy.

2.2.2 Genetics and breeding for high nutritional quality rice

In general, a good deal of characters relate to quality such as amylose content, gelatinization temperature, protein content and amino acid composition, which appear in endosperm and could be called as the endosperm traits. The expression of endosperm traits is under triploid genetic control and shows complex genetic behavior characteristics. Previous studies concluded that protein content is quantitative characteristics controlled by minor multiple genes (Singh et al., 1977). Shi et al. (1996) reported that inheritance of protein content is governed by triploid endosperm nucleic gene, cytoplasm gene and diploid maternal nucleic gene, whose inheritance effect is 44.3%, 34.4% and 4.9%, respectively. As far as the heritability of protein content, different results were observed by different researchers. Hille et al. (1972) found that the hereditary capacity of protein content is 0.130-0.372 resulting in low efficient

selection in early generation, but 0.59 and 0.71 in researches of Tsuzuk and Furusho (1986) and Shenoy et al. (1991), which are useful in the selection of high protein content rice line at early generation. Considerable negative correlation was observed between rice yield and protein content (Gomez and De Datta, 1975). Correlation was also found between 1000-grain weight, plant height, growth duration and protein content. In addition, environmental factors affected greatly the protein content in rice grain, particularly, the enhancing effect of N fertilizer usage at heading stage.

Lysine content has a close correlation with protein content. Some researchers reported that their relationship is positive correlation (Juliano, 1985) and some found that they have a negative correlation (Eppendorfer et al., 1983). Huang et al. (1990) pointed out that the contents of serine, proline, glutamic acid, histidine, arginine, alanine and threonine had a close positive correlation with eating quality of rice grain in early season varieties, of which glutamic acid and serine were important for increasing the taste and arginine decreasing the taste. While for later season rice, cystine and glutamic acid had more significant enhancing effect on the taste.

2.2.3 Structural, physiological and biochemical changes during rice seed development

Generally speaking, rice is a strictly self-pollinating crop, whose pollination occurs before or as the flowers open. Rice flower is composed of six stamens and one pistil. When pollens are dispersed from anthers and come into contact with stigmas, the fertilization is completed in 30 min. After that, a serial of changes in structural,

physiology and biochemistry take place during grain development including the following aspects (Juliano, 1985):

At 3-4 days after fertilization (DAF), caryopsis attains full length; the coleoptile, coleorhiza, scutellum and aleurone layer starts to differentiate; a multinucleated layer of endosperm is formed at the periphery of the embryo sac and cell membrane is formed; compound starch granules are detected in the endosperm cells; and large spherical protein body (PB) is found.

At 5-7 DAF (termed milky stage), the plumule and coleoptile are present and the embryo attains the capacity to germinate; aleurone layer is fully developed; the embryo sac becomes filled with endosperm cells; crystalline protein body appears.

At 9-10 DAF (dough stage), cell division is essentially completed and the size of the starch granule increases most rapidly.

At 14-15 DAF (yellow stage), caryopsis reaches the maximum width; dry matter is actually accumulated; compound starch granule gains the maximum size; potassium, calcium and manganese are accumulated rapidly in the hull.

At 20-21 DAF (mature stage), the differentiation of embryo is fully completed; the caryopsis attains the maximum thickness and the pericarp reaches its mature state; aleurone layer matured; and starch granules are formed.

2.3 Molecular biological characteristics of rice seed storage proteins

2.3.1 Property, classification and structure

2.3.1.1 Property and classification

Rice seed storage proteins are the major storage substance in rice endosperm. They can be classified into four types based on solubility, diluted acidic or alkaline glutelin, salt-soluble globulin, alcohol-soluble prolamin and water-soluble albumin (Osborne, 1924). Of them, glutelin is the most abundant fraction, which shows considerable difference from the other cereals with prolamin as the most abundant group.

Like other cereals rice seed storage proteins have some characteristics including (Shewry, 1995): 1) They are expressed and accumulated at special developing stages and in specific organ (endosperm) and the process is easily influenced by plant nutritional condition; 2) They are comprised of several kinds of polypeptides with different amino acid composition and physiochemical features; 3) Each type of storage protein is heterogeneous, which consists of many polypeptides and controlled by multiple gene family. For example, glutelin was encoded by at least 6 genes involved in A and B subfamily, respectively; 4) All of them are deposited in the protein body (PB). For instance, glutelin and globulin deposit in protein body II (PBII) with irregular shape, bigger size and deep-color while prolamin is deposited in protein body I (PBI) with regular round, smaller size and light-color (Yamagata et al., 1982); 5) They have no enzymic activity; and 6) They are decomposed and released nutrients

such as amino acid, nitrogen and sulphur during seed germination.

2.3.1.2 Composition and structure

2.3.1.2.1 Glutelin

Glutelin is the main form of storage protein in rice grain, accumulating greatly in the middle and late seed development (Yamagata et al., 1982). Although its solubility differs from the 11S globulin in legume they share similarity in amino acid composition, protein structure and synthesis pathway (Zhao et al., 1983; Wen and Luthe, 1985). Glutelin is of molecular weight in rice grain, the 57-kDa precursor, 37-39-kDa acidic subunit (large subunit) and 22-23-kDa alkaline subunit (small subunit), of which acidic and alkaline subunit are formed by precursor processing. Mature polypeptides all show heterogeneous. For example, the large subunit comprises at least 12 polypeptides and the small subunit consists of at least 9 polypeptides (Wen and Luthe, 1985). Glutelin is classified into two subfamilies, termed A and B subfamily, based on amino acid sequence. Each subfamily is controlled by multiple genes. The homology in amino acid sequence of among the intra-two subfamilies is 80-98%, whereas that of inter-subfamilies is lower than 65% (Takaiwa et al., 1991a). Analysis of the deduced amino acid sequence from cDNA showed that a typical signal peptide composed of 24 AA in the N-terminal of glutelin precursor, which contains a conserved leucine-rich hydrophobic core. This signal peptide (SP) plays an important role in leading glutelin precursor into the ER and further transporting to lumen of ER (Okita et al., 1989; Takaiwa et al., 1991a). Seven

cysteine-residuals exist in precursor, of which the –122 and –315 positions are responsible for production of bisulfate bond between acidic subunit and alkaline subunit. There are 5 variable regions (VR) in the different glutelin precursors; three of them are located in the C-terminal of signal peptide (SP), acidic subunit and alkaline subunit, respectively. Similar hypervariable regions (HVR) also exist in the storage proteins of legumes. The two others VRs present in the middle of acidic subunit. HVR is regarded as the optimal position to accept relatively large heterogeneous polypeptide sequence, and therefore it might act as the best location to modify this kind of protein and improve nutritional quality (Agros et al., 1985).

2.3.1.2.2 Prolamin

Generally speaking, prolamin is the most abundant protein in cereals. However, it is scarce in rice grain, which only accounts for 5-10% of the total rice protein. The molecular of weight of prolamin is relatively small but the variety is large. Prolamin is classified as three types on the basis of molecular weight, 16-kDa, 13-kDa and 10-kDa prolamin (Yamagata et al., 1982; Kim and Okita, 1988b). The 16-kDa mature polypeptide is composed of 130-140 amino acids while its the N-terminal of precursor contains a SP with 18-19 AA (Kim and Okita, 1988b). The 13-kDa prolamin has at least 7 polypeptide groups with different AA composition, whose SP in the N-terminal of precursor consists of 18-19 AA (Kim et al., 1988a; Sha et al., 1996; Hibino et al., 1989). The 10-kDa polypeptide is comprised of 110 AA and its precursor SP is 24 AA, longer than that of the two others (Masumara et al., 1989). In

comparison of the protein and DNA sequences, there is 70%-95% homology shared in different polypeptides inside the same molecular weight (MW), while the homology is low between different MW types. For instance, 47% homology is shared between the 13 kDa and 16-kDa prolamin, whereas the 10-kDa polypeptide has no homology with the 13-kDa and 16-kDa (Sha et al., 1996). Like the other cereals, all of the prolamin in rice contains a palindromic unit – QQQCCQQL octapolypetide. Moreover, Rice prolamin, gliadin (wheat) and zein (corn) have no significant homology in DAN and protein sequence, indicating that they possess individual origin in biological evolution (Kim and Okita et al., 1988b).

2.3.1.2.3 Globulin and Albumin

Some salt-soluble globulin and water-soluble albumin are considered as allergenic proteins due to specific reaction with immunoglobulin E (IgE) in sera from people allergenic to rice (Shibasaki et al., 1979; Matsuda et al., 1991). They all belong to α–amylase/trypsin inhibitor family (Limas et al., 1990, Izumi et al., 1992; Adachi et al., 1993; Alvarez et al., 1995). Albumin in rice is also heteogenous and contains several polypeptides with 14-16 kDa MW. Albumin is present in the aleurone protein bodies, which differ drastically in composition from the endosperm protein bodies. In endosperm protein bodies, prolamin mainly exists in the spherical protein body (PBI). Glutelin and globulin are formed in the crystalline protein body (PBII), which are more readily digested by pepsin and by humans and rats and more readily degraded during seed germination (Wu and Chen, 1978; Tanaka et al., 1980; Tanaka et al., 1978;

Horikoshi and Morita, 1982). The precursor of albumin consists of 150-170 AA and a SP with 26-27 AA in the N-terminal. The homology of different polypeptides is 70-95% (Adachi et al., 1993).

Reports on globulin are rare. α-globulin is the major fraction in rice globulin and is controlled by single gene. Its MW is about 26 kDa. The precursor is composed of 186 AA and has a SP with 22 AA in the N-terminal (Shorrosh et al., 1992).

2.3.2 Structure, expression and regulation of genes encoding rice seed storage proteins

The genes encoding rice seed storage proteins are of multigene family. Their expression shows tissue and temporal specificity, that is, their protein products are accumulated at certain stage in endosperm.

2.3.2.1 Structure

2.3.2.1.1 Glutelin gene family

Glutelin genes can be divided into two subfamilies according to DNA and AA sequence homology, subfamily A (GluA) and subfamily B (GluB). There are 60-65% homology in DNA sequence between GluA and GluB, but higher than 80% homology between the genes belong to the same subfamily. Several genes involved in each subfamily and each gene is multi-copy in rice genome (Takaiwa et al., 1991a).

Until now, several glutelin genes have been cloned and isolated in rice grain (Okita et al., 1989; Takaiwa et al., 1987a, b; Takaiwa et al., 1991a; Takaiwa and Oono, 1991). GluA subfamily comprises at least three genes, named GluA-1, GluA-2 and

GluA-3, respectively, by Takaiwa and Oono (1991), which correspond to the names of Gt2, Gt1 and Gt3 by Okita (1989). Each of them has 5-8 copies in rice haploid genome. One pseudogene namely GluA4 was found (Takaiwa and Oono, 1991). The homology of GluA-1 and GluA-2 is very high (95%), while 81% homology with GluA-3. They shared several conserved regions in DNA sequence, for example, the same transcription initiation site (-41 to –32 bp site), TATA box, CAAT box, and ploy(A) tail. However, there are differences in the distant sequence (–183 bp upstream of the 5' start condon between GluA-3 and GluA-1 (or GluA-2), which might cause the expression pattern of GluA-3 differing from GluA-1 (Okita et al., 1989).

In GluB subfamily, Takaiwa et al (1991a) isolated three genes, called GluB-1, GluB-2, and GluB-3, of which GluB-3 is a pseudogene. Recently, Jiang et al. (2002) identified a new rice glutelin gene, GluB-4. Each of them contains 5-8 copies in rice haploid genome. The homology of coding sequence is 80-88% among the three members. Similarly, the conserved transcription initiation site, TATA box, CAAT cassette, and poly (A) tail exist in family member genes. Interestingly, it's found that glutelin gene family varies obviously between *Indica* and *Japonica* rice cultivars (Takaiwa et al., 1991a; Takaiwa and Oono, 1991).

Furthermore, Okita et al. (1989) reported that 32-35% homology in DNA and protein sequences presents between rice glutelin gene and legume 11S globulin genes and the numbers and position of introns in coding sequence are similar, which reveal that they might come from the similar precursor. Since they share no similarity in promoter region, which might the reason why 11S globulin genes are expressed in

embryo (or cotyledon) but rice glutelin gene products are accumulated in endosperm (Okita et al., 1989).

2.3.2.1.2 Prolamin gene family

Rice prolamin gene family can be classified into at least three types according to MW, namely the 16-kDa, 13-kDa and 10-kDa, of which 13-kDa is the most abundant and is classified into 13-kDa-a and 13-kDa-b (Yamagata et al., 1986). Many cDNA and genomic clones have been obtained in these years, most of which belong to the 13 kDa group. Based on cDNA and AA sequence homology, the 13-kDa polypeptide is further classified into three classes: Class I, Class II and Class III. Class I contains Prol 7, Prol 14, RM2, S23, X24 and RP3 (Kim and Okita, 1988a; Yamagata et al., 1992; Shyur and Chen, 1990). There is a 88-97% homology between member genes. Class I prolamins are composed of 148-150 AA rich in glutamine (Gln) and hydrophobic amino acids such as leucine (Leu) but deficient in lysine (Lys), histine (His) and sulfur-containing AA such as methionine (Met) and cysteine (Cys). Class II prolamins include only Prol 17, which share 75% homology with Class I. they encode 149 AA and are rich in Met and Cys; Class III prolamins contain S18, RM17, RP5, RP6 and NRP33, which share 91% homology between member genes. These genes encode 156-158 AA, whose AA composition is similar to that of Class I (Shyur et al., 1992; Mausmura et al., 1990; Wen et al., 1993; Sha et al., 1996). The genes encoding the 10-kDa polypeptide have been reported, including cDNA clone RP10, and genomic clones X17074 (Feng et al., 1990) and Trp10. Homology in the coding

sequences (CDS) is above 95%. These genes encode 134 AA abundant in sulfur-containing AA.

Analysis of DNA sequence of rice prolamin genes revealed that they not only have normal regulation elements such as TATA box, CAAT box, poly (A) tail but also possess several characteristics including: 1) there is no intron in CDS of the rice prolamin, which is similar to the other cereals and legumes; 2) Low homology exists between different types of prolamin genes; 3) Mature RNA contains a relatively long 5'-untranslation region (UTR). The function of this region is unclear, but Muench and Okita (1997) speculated that a intron may be involved in this region; 4) There is low or no homology between the 5'- and 3'- UTR of different prolamins, indicating different expression patterns for different prolamins; 5) More short repeated sequences present in the 5'-UTR such as ATT repeat, which might act as enhancer function (Sha et al., 1996); and 6) There are no obvious homology in proalmins between rice and other cereals, revealing possibly a different evolutional pathway of the rice prolamin (Kim and Okita et al., 1988b; Sha et al., 1996).

The copy number of rice prolamin in rice haploid genome is 80-100, much higher than that of glutelin but similar to gliadin and zein (Kim and Okita et al., 1988b; Muench and Okita, 1997). But Shyur and Chen (1990) reported that the 13-kDa prolamin contains 30-50 copy numbers in rice haploid genome.

2.3.2.1.3 Albumin and globulin gene family

Albumin gene family is classified into four subfamilies, namely RA5, RA14,

RA16 and RA17 according to cDNA sequence homology. They all have been cloned and called RA5, RA14, RAG1 and RAG2, respectively (Adachi et al., 1993). There is 70-95% homology between different genes (Izumi, et al., 1992). The copy number of albumin is much less than that of prolamin.

Four globulin cDNAs gene and one genomic gene have been cloned up to now (Shorrosh et al., 1992; Nakase et al., 1996b). α-globulin (Glb gene) is a major fraction with 26-kDa in MW and encoded by a single gene (Shorrosh et al., 1992). Moreover, α-globulin shares similarity of the 5'-flanking region with those of the genes encoding wheat high-molecular weight glutelin and barley hordein (Nakase et al., 1996b).

2.3.2.2 Expression of storage proteins during rice seed development

Previous expression profiling focused more on glutelin and prolamin genes. Kim et al. (1993) reported that glutelin was expressed early at 5 days after flowering (DAF). The expression level and pattern differed between the GluA and GluB subfamily. In GluA, the expression pattern of Gt1 and Gt2 is similar, which starts at 5 DAF and gradually increases during rice ripening. Gt3 gene showed the diverge pattern, which reaches the highest expression level at 10 DAF and then gradually decreased at the later stages. In GluB, gene expression can be detected at 6 DAF, obtained the expression peak at 14 DAF and then deduced during the later stages (Takaiwa et al., 1991a).

It was reported that the expression patterns of prolamin genes are alike. Kim and Okita (1988b) observed that Prol 14, a member of Class I 13-kDa prolamin, was

expressed beginning at 5 DAF with slight amount and showed abundant level at 10 DAF, then increased during seed ripening. The similar result was reported by Shyur et al. (1992). The expression level of Prol 17, a member of Class II 13-kDa prolamin, is much higher than that of Prol 14, but similar after 25 DAF. Class III 13-kDa group gene (S23) was showed to have similar expression pattern to these of Class I and II (Masumura et al., 1990). These indicated that prolamin proteins are accumulated abundantly at the later stages during grain development. Similarly, immunoblot analysis showed that the ratio of glutelin to prolamin was 1.7 at 10 DAF and this ratio steadily decreased to 1.2 at 25 DAF, revealing that the synthesis and accumulation of prolamin was greatly increased at latter stages (Li and Okita, 1993).

Adachi et al. (1993) found that albumin could not be detected in leaf, root, and stem except seeds, showing strong tissue-specificity. During seed development, albumin gene expressed starting at 10 DAF and reached the highest level at 15-20 DAF, then declined until seed matured.

Up to now, the expression profiles of globulin genes are still unclear. Report on this aspect is rare.

2.3.2.3 Regulation of expression of seed storage protein genes

2.3.2.3.1 Regulation at transcriptional level

The expression of genes encoding seed storage proteins is mainly regulated at the transcriptional level. The promoter regions of total rice seed storage protein genes contain several regulatory elements that relate to spatial and temporal specific

expression of these genes. Some promoters with transcriptional activity have been identified for Gt1 (Takaiwa et al., 1991b; Zheng et al., 1993), Gt2 (Kim and Wu, 1990), Gt3 (Zhao et al., 1994; Croissant-Sych and Okita, 1996; Yoshihara et al., 1996; Wu et al., 1998a), GluB-1 (Wu et al., 1998a, 2000; Washida et al., 1999), NRP33 (Wu et al., 1998a), RP3 (Yang et al., 2000), RP5 (Su et al., 2001), RAG1 (Wu et al., 1998a) and α-globulin (Wu et al., 1998a). These promoters have a –300 element (or endosperm box), which is important in regulation of spatial and temporal-specific expression and organ-specific expression of seed storage protein genes in monocotyledon (Muench and Okita, 1997). For example, the activity of GUS was reduced 20-fold than when the 5.1-kb Gt1 completed promoter was excised into 1.8 kb; GUS activity was unchanged if 1.8 kb was further reduced to 507bp or 399bp; while GUS activity was completely lost if the 399bp promoter was reduced to 214bp (Zheng et al., 1993).

Until now, several cis-elements related to endosperm-specific expression have been identified and cloned, including the AACA motif (AACAAACTCTATC), GCN4 motif (TGAGTCA), ACGT motif (GTACGTG), GCAA motif (GCAAAATGA) and prolamin box [TG(t/a/c/)AAA(g/t)], of which AACA and GCN4 motifs are most common in the promoter regions of seed storage protein genes, particularly in glutelin. These motifs control endosperm-specific expression of promoters as well as expression quantity. Combination of AACA and GCN4 motifs is more efficient on enhancing the expression level (Yoshihara et al., 1996).

A few trans-regulatory factors, nucleoprotein binding to cis-elements, were

identified. For instance, OSMYB5 specifically binds to AACA motif in glutelin (Suzuki et al., 1998); bZIP nucleoprotein binds to GCN4 motif including RISBZ1 (Onodera et al., 2001), and Opaque-2 (Wu et al., 1988b). Furthermore, it has been reported that some nucleoproteins can bind to 16-kDa albumin, 13-kDa prolamin and glutelin promoter regions at the same time (Nakase et al., 1996a) and a novel rice bZIP protein (REB) which specifically binds to α-globulin promoter was isolated (Nakase et al., 1997).

2.3.2.3.2 Regulation at post-transcriptional level

Previous works showed that in addition to regulation at transcriptional level, the post-transcriptional level control is also involved in the expression of rice seed storage protein multigene families. Nuclear run-on transcription assays indicated that the transcriptional activity of Gt3 is 4-fold higher than that of Gt1/Gt2 at 5 DAF, but the amount of mature mRNA is similar among the three genes. Moreover, similar results were observed in the prolamin multigene family. The level of Prol 14 transcripts is obvious higher than that of Prol 17 at 5 to 15 DAF whereas the mature mRNA amount of the former is much less than that of the latter (Kim et al., 1993).

2.3.2.3.3 Regulation at translational level

The ER-membrane-associated translational regulation plays important role in the expression of seed storage protein multigene family. Kim et al. (1993) indicated that the prolamin transcripts are abundant in the mid- to late seed developing stages, which

are 40% excess over the glutelin transcripts. Interestedly, the ultimate prolamin products are much less than glutelin. Further analysis showed that the glutelin mRNAs bind to membrane-bound polysome (MB polysome) with high efficiency, which aids to form translation initiation complex. In contrast, significant lower amounts of prolamin mRNAs were observed in MB polysome and a large proportion was present in membrane-free polysome (F polysome). Therefore, the translation efficiency of glutelin is much higher than that of prolamin.

2.3.3 Synthesis, processing and deposition of rice seed storage proteins

It's well known that seed storage proteins are deposited in specific protein bodies (PB), which provide some advantages to protein storage. For example, protein can be rapidly accumulated and stored quickly to avoid degradation and cytological interference. However, synthesis of seed storage proteins is a complicate biological process, whose mechanism is not clearly understood.

Seed storage protein (glutelin and prolamin) mRNAs are transcribed in the nucleus and then released into the cytosol and localized on the distinct ER membrane to synthesize protein. Previous researches found that there are two kinds of ERs exist in the endosperm. One is the cisternal ER (C-ER), consisting of a layer of pellicle and is distributed in the endosperm cells. Glutelin mRNAs are localized in this kind of ER. Another is the protein body-forming ER, which is the place where prolamin mRNA is localized (Li, et al., 1993 a, b).

To investigate the mechanism of specific subcellular localization of seed storage proteins, many studies have been carried out in these years. Choi et al. (2000) reported that 3'-untranslational region (3'-UTR) plays a significant role in localization of prolamin mRNA in PB-ER. It was found that signal peptide (SP) might recognize small nuclear protein (SNP), which helps mRNA to localize in distinct ER (PB-ER or C-ER), for instance, QQQCCQQL in prolamin SP and CXXLLCXGS in glutelin SP (Takaiwa, et al., 1991a). In addition, the chaperone, belonging to binding protein (BiP) in ER lumen, may also play an important role in processing of mRNA specific-localization. Li et al. (1993b) observed that many BiP appeared in PBI. BiP can bind to mature prolamin to form protein complex, which is responsible for translocating prolamin from the outside of ER into its lumen, followed by folding in the lumen, assembling, and formation of PBI. It has been suggested that this protein complex benefits the extension of translation complex binding in the ER and favors the localization of mRNA in PB-ER (Okita et al., 1994).

mRNAs of storage proteins are localized on the ER to synthesize protein precursor with SP. Under the leading of SP, precursor translocates into ER lumen to process, including disulfide formation, folding into trimer and oligomerization (Li et al., 1993b; Müntz, 1998). At the same time SP is removed from precursor, which is necessary to complete the latter process (Dickinson et al., 1989; Coleman et al., 1995)

There are at least two types of PB in endosperm. The origin, structure and content are different between them. PBI contains prolamin and PBII is composed of glutelin and globulin. Therefore, the biosynthesis of these two PBs is different. PBI is directly

derived from endosplasmic reticulum (ER) membrane while PBII comes from vacuole (Krishnan et al., 1986; Shewry et al., 1995). Polypeptide trimers are secreted outside of ER and translocated into vacuole through Golgi apparatus. Inside the vacuole, the trimers are processed into small and large subunit, and then folded into hexamer, which are accumulated in the vacuole. Finally, these hexamers render the vacuole lyses into small pieces to form PBII (Müntz, 1998).

2.4 Molecular characteristics of starch in rice grain

2.4.1 Property of rice starch

Starch is the most abundant carbohydrate component of rice grain, making up around 90% constituent of milled rice by dry weight. It is present in rice endosperm and cluster into starch compound granules within an amyloplast with spherical to ellipsoidal crystal shape (Hoshikawa, 1968; Hayakawa et al., 1980).

Starch is composed of two kinds of molecular components, amylose and amylopectin, both are high-molecular weight polysaccharides. Amylose is the minor fraction of starch granules with a linear molecular structure linked by α-1,4-glucosidic bonds. Nonwaxy milled rice has 8-37% amylose, depending on species, in its starch, staining purple-blue to blue with iodine. Glutinous rice (Waxy rice) starch has lower amylose content (0.8-1.3%) and is located at the center of the granule, staining red or brown color with iodine. According to amylose content, milled rice (white rice) can be classified into four classes: waxy, 1-2% amylose; low, 7-20%; intermediate, 20-25%; and high, >25% (Juliano, 1979). Amylopectin is a branched fraction with

both α-1,4-glucosidic bonds and α-1,6-glucosidic linkage like glycogen in animals and bacteria. Approximately every 20-25th glucose residue has a branch point at which a chain of α-1,4-glucosidic bonds is linked by an α-1,6-glucosidic bond to an other α-1,4-glucose chain. On the basis of sequential enzymatic hydrolysis and investigation of products of degradation, the amylopectin is regarded as non-randomly branched and has a high-degree of structure organization. It contains three types of chains: A chains, composed of glucose linked α-1,4 chains; B chains, consisted of glucose linked α-1,4 and α-1,6 chains; C chains, made up of glucose with α-1,4 and α-1,6 linkages plus a reducing group in the molecule. The A chain/B chain ratio is used for accounting for the enzymatic degradation of amylopectin and its characteristics. The starch granule is a mixture of amylose and amylopectin, but how they associate together in a granule is still unclear, possibly through hydrogen bonding.

Starch granules immersed in cold water at room temperature will swell and their diameter will be increased by about 20%. Heating the starch suspension to higher temperature results in gelatinization. At the gelatinization temperature (GT) the energy level is sufficiently high to dissociate the relatively weak hydrogen bonding in the starch granules and the amylose and amylopectin molecules are considered to be dissolved. Nonwaxy rice and Waxy rice starch granules have a similar range of final GT, 55-79 °C (Juliano, 1985; Hizukuri et al., 1983a). Final GT may be divided into four grades as low, <69.5 °C; intermediated, 70-74 °C; and high, >74 °C (Juliano, 1979). High GT has so far been confided mainly to waxy and low amylose starch.

Amylose is thought primarily responsible for gelatinization and may form firm gels in concentrations as low as 1.5%. It plays a dual role in gel formation: as the chief component forming the gel and entrapping the water, and as binding material linking intact starch granules or fragments of granules. Which of the two roles is more significant is determined by the degree of dispersing of the starch granules. The amylose/amylopectin ratio determines many of the properties of boiled rice (Juliano, 1979). Higher amylose content improves the capacity of the starch granule to absorb water and expand in volume without collapsing because of the greater capacity of amylose to hydrogen bond or retrograde. This observations support the hypothesis that the physical structure of the starch granule determines the physical properties of the endosperm and the rheological behavior of the starch suspension as it takes up water.

As the starch granule size increases during grain development, the amylose content of nonwaxy rice starch increases slightly by 6.4% in IR28 (Singh and Juliano, 1977) and 1.8-7.9% in other seven lines (Blakeney, 1980), while in waxy rice, amylose content changes from 2.5 to 1.5% in IR19 (Singh and Juliano, 1977). Final GT decreases in both waxy and nonwaxy rice during seed development, by 2% in IR8 and by 1.8-4.3 in seven nonwaxy and two waxy lines (Briones et al., 1968; Blakeney, 1980). In general, amylose content tends to increase and amylopectin tends to decrease during grain development. However, some researches have also found no notable changes in the amylose and amylopectin content in developing IR8 seeds (Briones et al., 1968).

Environmental conditions greatly influence the amylose content and the final GT. Samples of the same cultivars may give a range of amylose content up to six percentage points variation and a range of final GT of more than 10℃. The major environmental factor is to do with the ambient temperature during grain maturation (Suzuki and Murayama, 1967; Nikuni et al., 1969; Resurreccion et al., 1977). Low temperature tends to lower the GT and to increase the amylose content of starch. In contrast, high temperature induces higher GT and decreases the amylose content.

Starch granules contain small amount of residual nitrogen, phosphorus and lipids. Most of the nitrogen is the *Wx* protein product (Sano, 1984; IRRI, 1985) and perhaps starch synthase and its glucoprotein intermediates (Tandecarz et al., 1975). Phosphorous appears mainly (93-98%) as 6-phosphoglucose phosphorous in the waxy starch granules but 75-89% as phospholipid in the nonwaxy rice (Hizukuri et al., 1983b). Major starch lipids are free fatty acids, lysophosphatidycholine, and lysophosphatidylethanolamine (Choudhury and Juliano, 1980b). Starch lipids probably are complexed with amylose and are either free fatty acids or monoacyl lipids. High amylose rice has slightly lower bound lipids than intermediate amylose rice (Choudhury and Juliano, 1980b; Russell and Juliano, 1983).

2.4.2 Starch biosynthesis in rice

Adenosine diphosphate (ADP) -glucose starch synthase pathway is now regarded as the major pathway of starch biosynthesis instead of phosphorylase (Murata et al., 1964b and Figure 2.1). Firstly, Adenosine diphosphate-glucose (ADP-G) was formed

from glucose-1-phosphate (G-1-P) and adenosine triphosphate (ATP) by ADP-glucose pyrophosphorylase (AGPase) catalysis. By adding an ADP-G to the nonreducing end of an α-1.4 glucan primer to generate linear α-1,4 glucosyl chains, which is catalyzed by starch synthase. The linear α-1,4 glucans are used as the substrate for the next enzyme in the biosynthetic pathway. Amylose is produced by the catalytic action of granular-bound starch synthase (GBSS). In the other hand, starch branching enzyme (SBE) introduces α-1,6 glucosyl linkages between linear chains and starch debranching enzyme (SDE) removes the surplus branching (α-1,6 glucan) to synthesize amylopectin.

The biosynthesis of starch is a tissue and subcellular compartment-specific process. In leaf, starch is synthesized during the day directly from photosynthetically fixed carbon dioxide in the stroma of chloroplasts, where it serves as a short-term carbohydrate reserve termed transitory starch. During the night, this pool of starch is degraded to provide a carbon supply for sucrose synthesis and export, and for respiration (Zeeman and Rees, 1999). In seed, starch is synthesized in amyloplasts as a long-term storage form for carbohydrates. The starch biosynthetic pathway starts with the conversion of glucose-1-phosphate in to ADP-glucose, a key step catalyzed by ADP-glucose pyrophosphorylase (AGPase). The ADP-glucose then serves as a glucosyl donor for α-glucan synthesis by the action of starch synthases and starch-branching enzymes (Bechles et al., 2001 and Slattery et al., 2000).

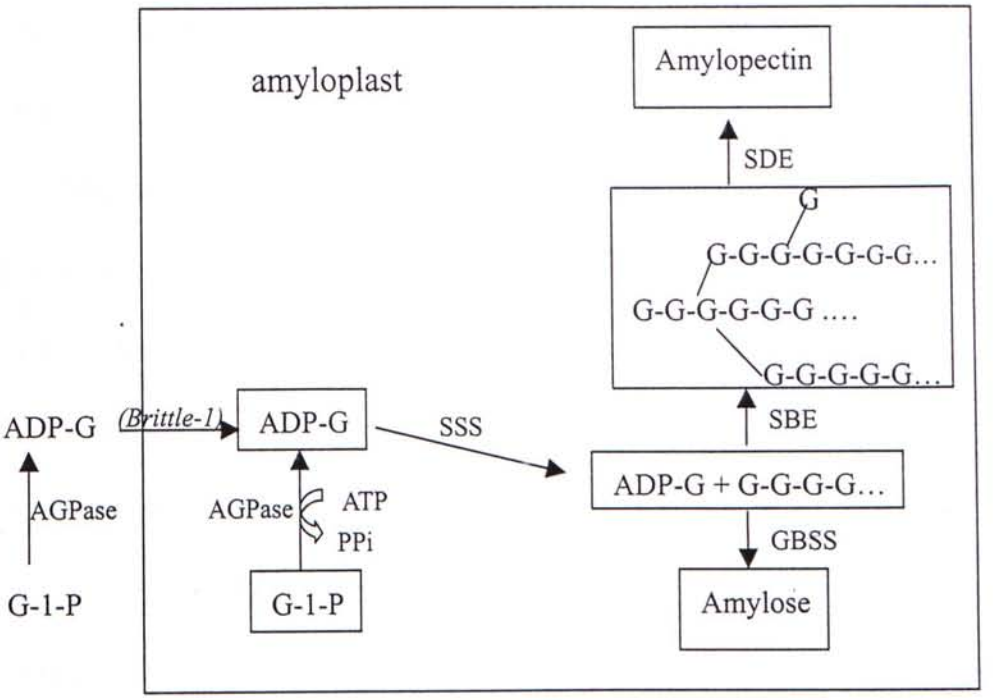


Figure2.1. The biosynthesis of starch in rice grain development

G-1-P: Glucose-1- phosphate; ADP-G: ADP-glucose; AGPase: ADP-glucose pyrophosphorylase; SSS: soluble starch synthase; GBSS: granular-bound starch synthase; SBE: starch branching enzyme; SDE: starch debranching enzyme; *Brittle-1*: a protein with capable of transporting ADP-G from the cytosol into plastid in maize.

2.4.3 Enzymes involved in starch biosynthesis in rice

2.4.3.1 ADP-glucose pyrophosphorylase (AGPase)

AGPase catalyzes the first unique step in the starch biosynthesis pathway, which produces the activated glucosyl donor ADP-G and pyrophosphate (PPi) from G-1-P and ATP. AGPase is present in both source and sink tissues. In photosynthetic organs such as leaf, AGPase is restricted to the plastids and regulated positively by 3-phosphoglyceric acid (3-PGA) but inhibited by orthophosphate (Hannah, 1997; Preiss et al., 1991). In cereal endosperm, recent studies suggested that AGPase is

largely extra-plastidial (85-95% cytosolic) and is the major form of this enzyme. But this was found not the case in non-cereal starch storage organs (Beckle et al., 2001; Denyer et al., 1996). Recent researches suggested that distinct cytosolic and plastidial forms of AGPase are encoded by separate large- and small-subunit genes exist in all cereal endosperm (Thorbjornsen et al., 1996; Denyer et al., 1996; Beckles et al., 2001; Burton et al., 2002). The cytosolic AGPase may assist the partitioning of carbon from sucrose to starch when sucrose is abundant in the endosperm (Beckle et al., 2001). The transportation of ADP-G from cytosol into plastids is carried out by the activation of the *brittle-1* protein in maize, which is a small inner envelope protein and has an adenylate translocator with a common ADP-glucose-binding domain (Sulllivan, 1995; Shannon et al., 1998).

AGPase is heterotetramer, consisting of two large subunits and two small subunits, each of which is encoded by a single gene. The AGPase small subunit gene sequences from various dicots and monocots differ mainly in exon 1, indicating that cytosolic AGPase might have evolved more than once (Hannah et al., 2001). The large subunits are mainly responsible for modulating the sensitive of small subunits to allosteric regulation by Pi and 3-PGA while the major function of small subunit is catalytic (Fu et al., 1998).

Previous researches showed that the activities of enzymes related to the conversion of sucrose to ADP-glucose including AGPase remain high expression throughout starch accumulation during grain development and reach the highest level 8-18 DAF (Perez et al., 1975; Lee and Su, 1982). AGPase plays a key role in

regulating starch biosynthesis in cereal seeds and is likely the most important determinant of the sink strength of the seeds. Smidansky et al. (2002) reported that an increase in AGPase activity within the endosperm of wheat seeds should enhance the development of the sink strength and overall plant productivity including seed yield, and total plant biomass.

2.4.3.2 Soluble starch synthase (SSS)

Starch synthase comprised of two distinct classes, soluble starch synthase (SSS) and granule-bound starch synthase (GBSS). As suggested by its name, SSS can be found in the soluble phase of starch biosynthesis. SSS catalyzes the chain-elongation reaction of linear α-1,4 glucosyl chains by transferring a glucan from ADP-G to non-reducing end of the bonds. SSS consists of at least four isoforms, termed SSS1, SSSIIa, SSSIIb, and SSSIII. Wu et al. (2001) recently submitted some putative new SSS genes to the genbank (NCBI) including SSSII-1, -2, -3, SSSIV-1, and -2. Studies indicated that the entire carboxy-terminal region of SSS1 is required for starch binding, especially for the generation of the shortest amylopectin chains, capable of polymerization of 10 glucosyl units or less. Furthermore, the SSSI binding affinity is positively correlated with the length of linear substrate chains and is negatively correlated with the catalytic capability of SSSI. For further extension to synthesize longer chains, it is carried out by isoforms SSSII and SSSIII (Commuri et al., 2001; Lloyd et al., 1999; Edwards et al., 1999). The activities of these isoforms vary in different plant species and tissues (Smith et al., 1997). Recent researches on genetic

mapping of the *japonica* and *indica* rice varieties indicated that the structure of amylopectin can be divided into two types, L-type and S-type, which can be distinguished between *japonica* and *indica*. *Japonica* cultivars are abundant in S-type amylopectin (short chains with DP<10) in comparison with the *indica* rice. SSSIIa gene plays a key role in producing different types of amylopectin because this enzyme is responsible for elongating the S-type amylopectin to generate L-type amylopectin (intermediate chains with DP 13-22), which typically form double helices and determines the function and structure of starch. The activity of this enzyme is inhibited in *japonica* rice. Edwards and his colleagues (1999) utilized antisense inhibition of SSSII and SSSIII in potato tubers and suggested that SSSIII is the primary catalyst of producing long B1 and B2 chains of amylopectin.

Although the precise role of individual SSS isoform has not been identified, it's believed that the function of these isoforms is especially in amylopectin synthesis and different isoforms perform unique role in determining the distribution of amylopectin chain lengths; for example, SSSI, SSSIIa and SSSIII particularly contribute to the formation of very short chains, intermediate chains and long chains, respectively (Tsai, 1974; Commuri et al., 2001; Wang et al., 1993; Gao et al., 1998; Harn et al., 1998).

2.4.3.3 Granular-bound starch synthase (GBSS)

GBSS is the granule-bound starch synthase isoform (or insoluble starch synthase) involved in starch synthesis, which is encoded by the *Waxy* (*Wx*) gene in cereal and has the function is specifically to produce amylose. Tsai (1974) reported that the *Wx*

mutation lacking GBSS in maize only produced the amylopectin but no amylose, providing the good evidence to support that amylose is synthesized by GBSS catalysis. GBSS contains two forms in maize and pea, a 59-kDa GBSSI and a 77-kDa GBSSII, of which GBSSI is more commonly called the "*Wx* protein" (Dry et al., 1992; MacDonald and Preiss, 1983 and 1985; Smith, 1990; Sano, 1984; Konishi et al., 1985). GBSS belongs to starch synthase, which elongates the linear α-1,4 glucosyl chains by adding the glucan from ADP-G to the non-reducing end of the linkages. All plants possess the granule-bound isoform GBSSI that are conserved broadly in evolution. Some researches revealed that GBSSI activity and the *Wx* gene dosage are linearly proportional, but the *Wx* gene dosage is not proportional to amylose content in maize and wheat. This suggested that amylose synthesis might be influenced by elements besides GBSSI (Tsai, 1974; Fujita et al., 2001). Sano (1984) found that there are two naturally functional types of alleles in rice *Wx* locus, Wx^a and Wx^b, which are defined based on the amount of the *Wx* protein accumulated in the rice endosperm. The activity of Wx^a is 10-fold higher than that of Wx^b at both protein and mRNA levels (Isshiki et al., 1998). All *indica* cultivar rice (Oryza sativa), *O. rufipogon*, and *O. glaberrima* possess the Wx^a with normal sequence of GT at the 5' splice site of the first intron, suggesting a high expression level of the *Wx* transcripts in the endosperm. On the other hand, Wx^b has a G to T mutation (TT sequence) at the same position as Wx^a in all *japonica* varieties, which leads to inefficient splicing of intron 1 and low expression level of the *Wx* transcripts and *Wx* proteins (Umemoto et al., 1995; Isshiki et al., 1998; Hirano et al., 1998). Wang et al. (1995) reported that cultivars with high

amylose content possessed only the mature 2.3 kb *Wx* mRNA; cultivars with intermediated amylose content not only contained the 2.3 kb matured *Wx* mRNA but also the 3.3 kb pre-*Wx* mRNA, indicating that the first intron still present; and the waxy rice (glutinous rice) had no amylose, no mature *Wx* mRNA but only the 3.3 kb pre-mRNA resulting from incompletely splicing. These results indicated that the amylose content in rice endosperm is regulated by the post-transcriptional level of the *Wx* gene and amylose content is positively correlated with the ability of the cultivars to remove the intron 1 from the leader sequence of the *Wx* transcripts.

Furthermore, the ambient environmental factors greatly influence the activity of starch synthase, especially temperature during seed development. It has been reported that the low temperature specifically enhanced the expression of the *Wx* gene and resulted in increasing amylose content in rice endosperm, whereas the activities of AGPase, SBE and SSS decreased at low temperature (Umemoto et al., 1995).

2.4.3.4 Starch branching enzyme (SBE)

Starch branching enzyme (or Q-enzyme) is a key enzyme in the production of amylopectin polymers in rice endosperm. It is responsible for the formation of α-1,6-glusocidic linked branches on the already synthesized and elongating amylose chains, and also in generation of new nonreducing ends to provide additional α–glucan acceptor in the α–1,4 glucan elongation system (Yamanouchi et al., 1992; Mizuno et al., 2001). SBEs are also present as multiple isoforms. Burton et al. (1995) reported that there are two families involved in SBE, family A and family B. Family

A contains SBEII in maize, SBEI in pea, SBE3 and SBE4 in rice. The others belong to family B including maize SBEI, pea SBEII and rice SBE1. The two families differ in their C-terminal region sequences of almost 50 amino acids, which might play an important role in determination of the substrate specificity and reactivity of SBE (Fisher et al., 1993; Bhattacharyya et al., 1990; Mizuno et al., 1993; Baba et al., 1991; Burton et al., 1995; Mizuno et al., 1992). Furthermore, Mizuno et al. (2001) suggested that the members of family A can be divided into two-subgroups, SBE3 and SBE4 types according to expression specificity. In rice, it's generally believed that SBE comprised of SBE1, SBE3 (SBEII b) and SBE4 (SBEII a) based on their amino acid sequence homology (Mizuno et al., 1992). Rice mature forms of SBE1, SBE3 and SBE4 are 82-, 87- and 84-kDa protein, respectively, and all of them belong to α–amylase family (Båga et al., 2000, Mizuno et al., 2001). Although the mature form of SBE4 shares 47% and 80% identity with SBE1 and SBE3 in amino acid sequence, they play distinct roles in starch synthesis and possess a specific tissue and stage expression. For example, SBE1 is more active toward amylose to branching amylopectin and prefers to transfer long and intermediate chains (DP16-37). By contrast, SBE3 and SBE4 tend to increase branching amylopectin and prefer to transfer short glucose units (DP6). In addition, genes encoding SBE4 and SBE1 are expressed in both leaf and developing seeds, whereas SBE3 shows seed-specific expression. The expression of SBE4 begins at 3 DAF, which is earlier than that of SBE1 and SBE. Moreover, SBE4 gene reaches the expression peak at 5 to 7 DAF, earlier than the two others (Mizuno et al., 2001).

Investigation on the possible functions of these three SBE isoforms in starch synthesis was performed by mutation studies, which provided a clearer understanding of the amylopectin biosynthesis pathway. The *ae* mutation lacking SBE3 resulted in increasing amylose content and decreasing 50% of the SSSI activity, indicating that SBE3 plays an important role in amylopectin production (accounting for about 15-20% of the total SBE activity) and that it also influences the SSSI expression (Nishi et al., 2001). SBE1 mutation resulted in the synthesis of more short and less long and intermediate chains in the amylopectin structure, suggesting that SBE might play an important role in production of B1 and B2-B3 chains. In addition, although SBE4 mutation in rice and maize showed no change on the composition and structure of starch, it was suggested that SBE4 contributes to the generation of short chains amylopectin in starch synthesis based on the observation of SBEIIa *Mutator* insertional maize mutants (Nakamura et al., 2002; Blauth et al., 2001).

2.4.3.5 Starch debranching enzyme (SDE)

Starch debranching enzyme (SDE) is also an important enzyme in amylopectin synthesis, as it directly catalyzes the hydrolysis of α-1,6 bonds. There are two types of SDE isoforms in plants, isoamylase and pullulanase (R-enzyme or limit-detrinase), with distinct enzymatic and structural characteristics (Nakamura, 1996; Beatty et al., 1999; Myers et al., 2000). The main difference between the two isoforms lies in that isoamylase excises α-1,6 bonds in the denatured amylopectin, glycogen and limited dextrins, but not pullulan, while pullulanase specifically cleaves α-1,6 branches in

pullulan but inactive to glycogen (Myers et al., 2000).

For the genes encoding the two SDE isoform enzymes, isoamylase is encoded by *Sugary-1 (su1), isa-1, iso1, Dbe1* locus on chromosomes 8 in maize and rice, barley, wheat and Arabidopsis, respectively. Pullulanase is encoded by *ZPU1* on chromosomes 4 in maize (James et al., 2003, Fujata et al., 1999; Nakamura et al., 1996).

Rahman et al. (1998) reported that isoamylase is a 79-kDa protein. Sequence comparisons revealed that the two SDE isoforms share less homology while the same isoform compared among different species show 60%-80% identity, suggesting that these two enzymes might have distinct functions in glucan metabolism. The *su1* mutation with deficient isoamylose and pullulanase in rice coincides with speculation. In this mutation, isoamylase activity was found to decrease more significantly than that of pullulanase and to increase in the accumulation of photoglycogen but with reduced starch content. It's generally believed that isoamylase has a predominant function in amylospectin synthesis. Although pullulanase plays a minor role, it is essential and can compensate for the role of isoamylase in the construction of the amylopectin multiple-cluster structure. Furthermore, pullulanase exists not only in starch accumulating organs but also in germinating seeds and leaves, showing degradative function in the endosperm, while isoamylase present widely in starch accumulation tissues (Kubo et al., 1999).

With this overall review on the enzymes related to starch synthesis pathway, we believe that all of these genes are essential and play distinct roles during starch

production. GBSS catalyzes the amylose synthesis while SSS, SBE and SDE involve in generation of amylopectin. However, how do these enzymes function together to determine starch content and structure, and the correlations between them in starch synthesis are still unclear up to now. More detailed researches are thus needed in the future.

2.5 Aspartate family amino acids biosynthetic pathway in rice

2.5.1 Introduction

Human can not synthesize 8 out of 20 kinds of amino acids including lysine, methonine, tryptophan, threonine, lecucine, isoleucine, valine, and phenylalanine, therefore need to obtain these from daily diet. Cereals are deficient in lysine while lugume is lack of methonine, so lysine and methonine are regarded as the most limiting essential amino acids for most of the people in the world who depend on crops aas dietary protein source. Because of the nutritional importance of lysine and methonine, the regulations of their metabolism and enhancement of their contents have always been focus points by researchers. It's well known that the synthesis of lysine and methonine involves in distinct branches of aspartate family amino acids biosynthetic pathway catalyzing by several regulatory enzymes including aspartate kinase (AK), dihydrodipcolinate synthase (DHPS) for lysine synthesis and homoserine dehydrogenase (HSD) for methonine synthesis (Galili, 1995; Bryan, 1990; Gengenbach et al., 1992) (Figure 2.2).

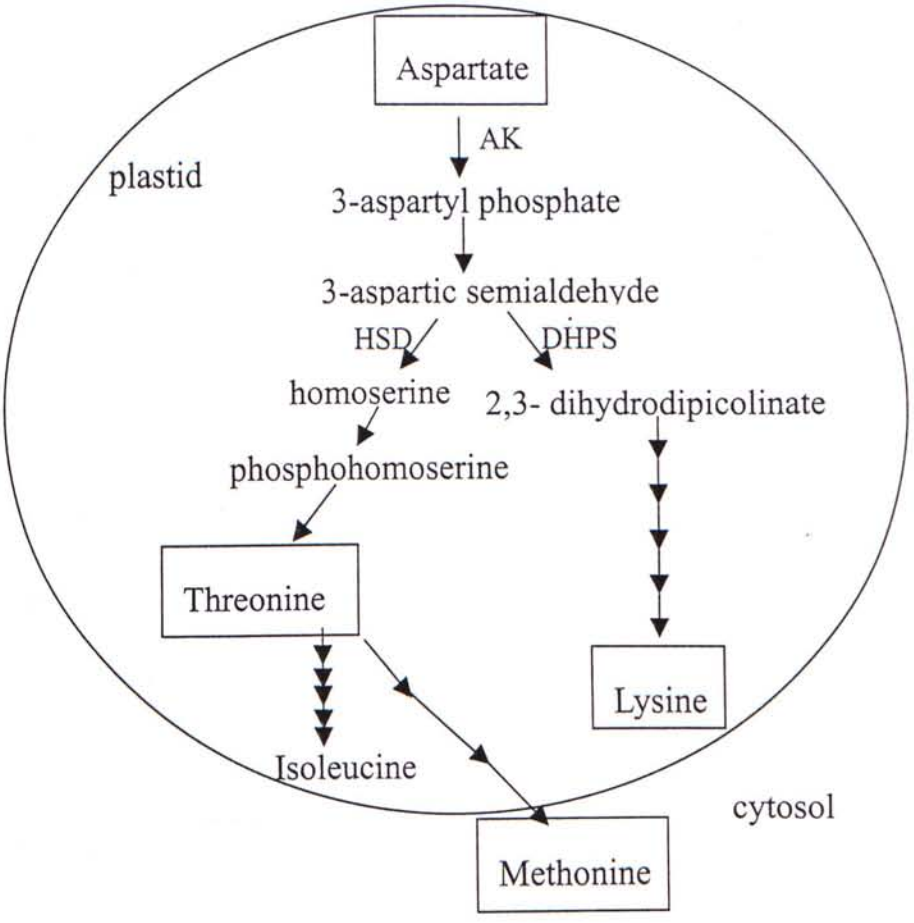


Figure 2.2 Aspartate family amino acids biosynthetic pathway in plant (after Galili, 1995).

AK: aspartate kinase; HSD: homoserine dehydrogenase; DHPS: dihydrodipicolinate synthase.

This pathway begins with aspartate as the common substrate through AK catalysis. The branches of lysine and threonine synthesis occur in the plastid, while methonine biosythesis takes place in the cytosol (Bryan, 1990). Furthermore, the aspartate family amino acids biosynthetic pathway is regulated by several end products feedback inhibition (Cohen and Saint-Girons, 1987).

The aspartate family amino acids biosynthetic pathway is functional in both vegetable organs (leaf and stem) and storage tissues (seed). Karchi et al. (1993, 1994)

reported that seed-specific expression of the bacterial feedback-insensitive AK resulted in a great enhancement in free threonine in mature seeds. Likely, expressions of the less lysine-sensitive DHPS driven by seed specific promoter significantly raise the accumulation of free lysine (Falco et al., 1995).

Like the other cereals, rice is also rich in methonine but lack of lysine. Therefore, increasing lysine content is the most important in improving the rice grain quality.

2.5.2 Two key regulatory enzymes involved in lysine synthesis pathway

2.5.2.1 Aspartate kinase (AK)

AK catalyzes the first step of the aspartate family amino acid pathway, which is necessary for the synthesis of all the end products. AK is responsible for phosphorylation of aspartate to 3-aspartyl phosphate. In higher plants, AK activity is feedback inhibited by lysine, threonine as well as S-adenosyl methionine (SAM) (Rognes et al., 1980; Dotson et al., 1989; Azevedo et al., 1992). Lysine-feedback-insensitive bacterial AK is encoded by *E.coli lysC* gene (Falco et al., 1995). As in *E.coli*, there are at least two or three AK isozymes in plants. Different isozymes possess individual feedback inhibition. For example, in barley, AKI isozyme is sensitive to threonine while AKII and AKIII are sensitive to lysine (Rognes et al., 1983). Some AK isozymes are found to have bifunctional property, with both AK and HSD activities (Wilson et al., 1991; Azevedo et al., 1992). Complementary DNA (cDNA) of AK and AK-HSD were isolated and cloned from

several plant species such as *Arabidopsis. thaliana* (Frankard et al., 1997; Ghislain et al., 1994), *Oryza sativa* (Fukuzawa et al., 2000; Kiyota and Sakano, 1996, Genbank), and maize (Doston et al., 1989; Muehlbauer et al., 1995). Expression of bacterial less lysine-inhibition sensitive AK resulted in significant overproduction of free threonine. Therefore, AK is regarded as the most limiting factor in threonine synthesis, whereas HSD is the second important element in threonine synthesis.

2.5.2.2 Dihydrodipicolinate synthase (DHPS)

DHPS is the critical functional enzyme in lysine synthesis branch pathway in that it catalyzes the first reaction in this pathway, that is, production of 2,3-dihydrodipicolinate by condensation of 3-aspartic semialdehyde with pyruvate. Unlike AK, DHPS activity is inhibited by lysine with $I_{0.5}$ only between 10 to 50μM, which is most sensitive to lysine feedback inhibition comparison with the other enzymes involved in AK pathway. For example, DHPS is about 10 folds higher sensitive to lysine inhibition than that to AK in plant. Falco et al. (1995) reported that expressing lysine-feedback-insensitive bacterial DHPS led to more than 2-fold increase in total lysine content in canola seeds. Previously, Negrutiu et al. (1984) had reported that overproduction of lysine could result from the expression of less lysine-sensitive DHPS. These results indicate that DHPS plays a key regulatory role in lysine biosynthesis in plants. This is opposite to that in *E.coli,* because DHPS is 20-100 fold more sensitive to lysine inhibition in plant than in *E.coli.* DHPS gene had been identified and cloned from rice (Suh et al., 2000, Genbank), maize (Frisch et al.,

1991), wheat (Kaneko et al., 1990), pea (Dereppe et al., 1992), sorghum (Piryns et al., 1988) and tobacco (Ghislain et al., 1990).

2.5.2.3. Regulation of lysine and the other end products of AK pathway

Lysine, methionine, threonine and isoleucine are the end products of individual branched pathways involved in AK family amino acids biosynthetic pathway. They share aspartate as their common substrate and AK enzyme activity. Therefore, their synthesis also asserts inhibitory regulation between each other. Frankard et al. (1992) and Galili (1994) reported that increase in lysine is accompanied by a significant decrease in accumulation of threonine in transgenic plants expressing both a less lysine-inhibition AK and DHPS gene. This suggested that the balance of different amino acid products in biosynthesis is determined by competition between corresponding distinct key functional enzymes for limiting amount of their common substrate. On the other hand, since DHPS is most sensitive to lysine feedback inhibition, lysine synthesis is regulated by most stringent condition. If DHPS is insensitive to lysine inhibition, the branch driving the lysine synthesis becomes the major pathway in comparison to other branches, thus the lysine content is higher than that of the other amino acids (Galili, 1995).

The activity of AK pathway enzymes appeared mainly at middle stage during seed development and declined greatly at the late stages, such as DHPS and LKR (lysine-ketoglutarate reductase) (Karchi et al., 1994). Further studies indicated that the GCN4 motif in promoter region is a major factor to regulate the amino acid

biosynthesis in yeast, maize and barley AK pathway (Hinnebusch, 1988; Mauri et al., 1993; Müller and Knudsen, 1993). The GCN4 box is also conserved in the promoters of storage protein genes. It's suggested that some similar promoter elements exist between amino acid biosynthesis genes and storage protein genes.

2.6 Proteomics in plants

Proteomics has been defined as the systematic analysis of the protein population in a tissue, cell, or subcellular compartment (Wijk, 2001). Zivy and de Vienne (2000) pointed out that proteomics play a link role between genomics, genetics and physiology. In recent years, more and more researchers are interested in functional genomics research on sequenced genes, therefore proteomics is becoming a field in great demand with a large impact on plant biology. Application of proteomics contains identification of polypeptides, analysis of differential protein expression, characterization of posttranslational modifications, and protein-protein interactions.

Proteomics is based on two-dimension electrophoresis (2-D gel) technique, mass spectrometry (MS) and bioinformatics (Figure 2.3). In 2-D gel, proteins are separated by pI in the first dimension and with molecular weight in the second dimension (O'Farrell, 1975). Matrix-assisted laser desorption-ionization-time of flight (MALD-ToF) provides high precision measurement to identify peptide masses after digestion by an endoprotease such as trypsin and pepsin. A well-constructed and available genomic database will greatly facilitate successful identification of proteins.

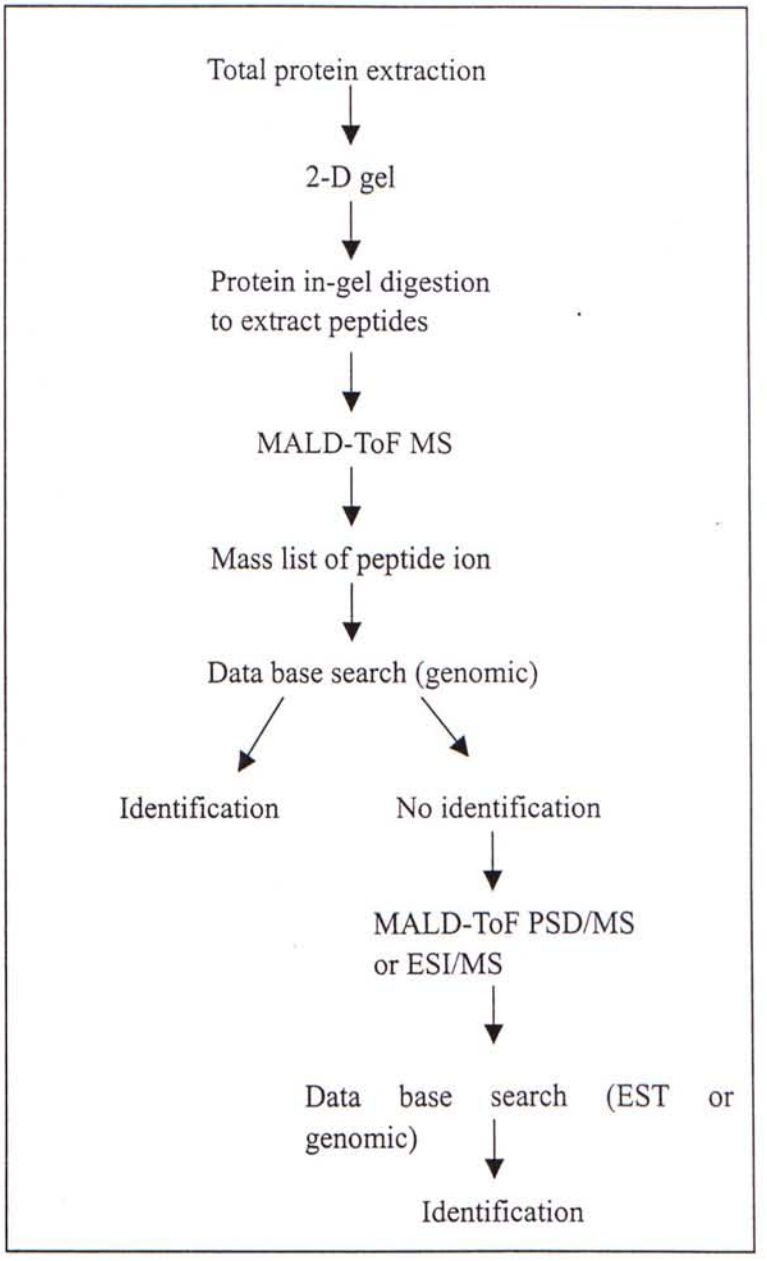


Figure2.3 Procedure for analysis of proteomics by mass spectrometry (revised from Wijk, 2001).

2-D gel: two-dimension electrophoresis; MALD-ToF: Matrix-assisted laser desorption-ionization-time of flight; PSD: post-source decay; ESI: electrospray ionization; EST: expressed sequence tag.

2-D gel has been applied in genetic maps of maize, barley, pea and pine by PS (position shifts) of proteins. PS is the allelic form of the same protein and can be used

to localize position shift loci (PSLs) on genetic map (de Vienne et al., 1996; Touze et al., 1995). A monogenic and co-dominant of PS were observed in all cases. It is likely that PS is resulted from polymorphisms in the genes encoding the variable proteins and PSLs on the genetic maps coincide with the location of the structural genes of the proteins. Moreover, 2-D gel is an abundant and less expensive source of variable genetic markers. It also provides a means for knowing if a gene at a particular locus is translated or not. For example, de Vienne et al. (1999) reported that the PGMA (phosphoglycerate mutase) gene is translated in 3 organs by 2-D gel analysis. Polymorphisms revealed by 2-D gel and isozymes are about the same but higher levels of polymorphisms are only revealed by RFLP due to lower conservation of no-coding sequence. This indicates that 2-D gel markers are exclusively for expressed genes and the only polymorphisms leading to the modification of pI or apparent MW of the proteins can be visualized by 2-D gel (Burstin et al., 1994). In addition, the identity of proteins revealed on 2-D gels is an useful means to examine the metabolic changes induced by treatments or any environmental variations. This approach is used to study not only the effects of drought on protein expression in maize (Riccardi et al., 1998), potato (Rey et al., 1998) and pine (Costa et al., 1998) but also the effects of anoxia tolerance and hormone treatment (ABA) on protein expression and to isolate novel cDNAs in maize (Chang et al., 2000) and rice (Moons et al., 1997).

Proteomics is thus an excellent tool for functional genomics study. Parallel study of proteomics and transcriptomics should make it possible to understand the relationship between mRNA and protein level and the regulation of genome

expression from transcription and post-translational processes. However, the database of peptide sequence tag is limited in non-sequenced plant species. Thus to develop an overall searching database is necessary.

2.7 Approaches for grain quality improvement in rice

2.7.1 Improvement of nutritional quality

Because of the important values of seed storage protein and its AA composition, increasing their contents specially lysine is the key point in grain nutritional quality improvement. Creation and selection of protein mutants is the major method to improve the grain nutritional quality by traditional breeding (Kumamaru et al., 1988; Ogawa et al., 1989). However, the limitations of this method include: low efficiency of mutant selection with and the accompanying undesirable traits (Iida et al., 1993; Schaeffer and Sharper, 1990) and its minor effects on protein content and lysine content enhancement.

In recent years, molecular technologies provide more effective ways to improve the grain nutritional quality. Several means at molecular level have been reported, which include modification of protein sequence to elevate essential amino acid content, overexpression of elite homogeneous protein genes, transformation and overexpression of elite heterogeneous protein genes, synthesis of novel protein gene, and increase of free amino acid content (Beach and Ballo, 1992; Sun et al., 1992, 1993; Mattews and Hughes, 1993; Habben and Larkins, 1995). These methods have been utilized widely in nutritional quality improvement of various plant species such

as corn (Wallace et al., 1988; Ohtani et al., 1991), potato (Tu et al., 1998), rape (Guerehe et al., 1990; Faloc et al., 1995), soybean (Kho and Lumen, 1988), wheat (Singh et al., 1993), and tobacco (Shaul and Galili, 1992) while reported on rice grain nutritional quality improvement is rare.

In general, the amino acid composition in rice is unbalanced, especially lack of lysine. Of the major four storage proteins, glutelin contains relative higher proportion of lysine but prolamin and albumin are low in lysine. Protein body I (PBI) containing prolamin is difficult to digest by human and animal with simple stomach. By contrast, PBII with glutelin and globulin is easy to digest and absorb. Therefore, glutelin possesses higher nutritional value than prolamin (Tanaka et al., 1975, 1978; Ogawa et al., 1987). Furthermore, albumin and globulin are allergen to sensitive people and they belong to the α-amylase/trypsin inhibitor family (Alvarez et al., 1995).

To enhance the nutritional quality of rice grains, there are several feasible strategies: 1) to increase glutelin and decrease prolamin content; 2) to reduce albumin and globulin synthesis to diminish allergic proteins in rice; 3) to enhance lysine content to improve amino acid balance through regulation of key enzymes involve in lysine biosynthesis pathway and transformation with lysine-rich protein (Gao et al., 2000; Liu, 2002).

2.7.2 Improvement of eating and cooking quality

As review above, amylose/amylopectin content and ratio are the most important elements to determine eating and cooking quality in rice grains. In recent years, more

and more scientists have been attempting to improve starch quality by means of plant genetic engineering, including to further understand the starch biosynthesis pathway, the function and relationship between enzymes involved in this pathway, and to modify starch quantity and quality by regulation of enzymes expression. Researches mainly focus on manipulation of AGPase and GBSS. Stark et al. (1992) transformed AGPase gene from *E.coli* driven by patatin promoter into tobacco and potato, and obtained transformants with increased starch content. Visser et al. (1991) introduced anti-sense GBSS into potato, decreased activity of GBSS and 70-100% lower level of the transformants than the control were observed. By contrast, addition of GBSS transgene resulted in increased amylose content in potato with no amylose (Van der Leij, 1991). Similar observation occurred in transgenic rice and cassava (Salehuzzman, 1993). These studies indicated that manipulation of GBSS and AGPase appears to be an effective method to grossly modify the starch quantity and quality. Up to now, reports on manipulation of SSS and SBE to improve amylopectin content and quality are rare.

2.8 Objectives of my project

Many aspects of the biology of rice seeds have been studied including gene cloning, transformation and regulation at DNA, RNA and protein levels (Bechtel, 1980; Gao, et al., 2001; Kim, et al., 1993; and Li, et al., 1993). However, there is no systematic profiling of the expressed genes, including their individual family members, in the maturing rice seeds, especially in relation to rice grain quality and

hybrid rice.

The aim of my project is to study the gene expression patterns during seed development of super hybrid rice, using the parental lines and hybrid F1 as research materials. The maturing seeds from 3, 6, 10, 15 and 20 days after fertilization (DAF) will be collected. Profiling the expression of genes encoding the four rice storage proteins, starch synthesis enzymes and key enzymes involved in lysine biosynthesis branch pathway, with emphasis on individual family gene members, will be performed by northern blot and RT-PCR for the parental and the hybrid generations. Expression patterns of these four main storage protein genes will also be studied at the protein level by tricine-SDS-PAGE and 2-D gel.

This basic molecular knowledge will provide useful information and genetic materials (genes) to understand the molecular events involved in rice seed formation, especially for the hybrid rice seeds, in our effort to enhance their grain quality.

Chapter 3 Materials and Methods

3.1 Materials

3.1.1 Chemicals

Most chemicals used were analytical reagent and purchased from Ridedl-deHaen (Germany), Roche Diagnostics Corporation (USA), Bio-Rad Co. (USA), Invitrogen Corporation (New Zealand), or Sigma-Aldrich Chemical Co. (USA). The Taq DNA polymerase, 10X buffer, 25mM $MgCl_2$, DTT, dNTPs, M-MLV-reverse transcriptase, 5X M-MLV buffer, DTT, RNAase inhibitor were obtained from Promega Bioscience (USA).

3.1.2 Apparatus and commercial kits

Protein, DNA, and RNA gel electrophoresis were carried out using electrophoresis apparatus and power supply purchased from Bio-Rad Company. Southern blotting was performed using the VacuGene XL vacuum blotting system (Pharmacia Biotech AB, Sweden), cross linker (GS Gene linker™, UV chamber), and model 1012 hybridization oven [China scientific (HK) Ltd.]. The EttanTM IPGphor™ Isoelectric Focusing System for first-dimension IEF and Ettan DALT*twelve* Large Format Vertical System for second-dimension SDS-PAGE were come from Amersham Bioscience (Sweden). The DIG-labeling kit and the DIG luminescent detection kit for nucleic acid analysis were purchased from Roche Diagnostics Co. (USA). The PCR product purification for DNA sequencing was executed using Wizard PCR preps DNA purification kit from Promega Bioscience (USA). The

Plusone™ 2-D Clean-Up kit, Plusone™ 2-D Quant kit, 24 cm IPG strips, IPG buffer (pH 3-10), precast polyacrylamide gels (255 x 196 x 1 mm) and Plusone silver staining kits were all obtained from Amersham Bioscience (Sweden).

3.1.3 Plant materials

The super hybrid rice seeds (P64S, 9311 and F1 hybrid) were provided by Hunan National Hybrid Rice Research and Developing Center (HNHRRDC). Pei'ai 64S (P64S) is a female parental line of hybrid rice with photo- and thermo-sensitive genic male sterility (PTGMS). It has a typical morphology of super hybrid rice and a high combining ability with different store lines. 9311 is the male parental line and a good restorer line with a better rice quality. P64S/9311 is the F1 hybrid, which is the strongest elite combination in China having national second-class rice quality and the highest production at 12 tons/ ha up to now.

3.1.4 DNA sequencing

DNA sequencing was performed using Perkin Elmer ABI Prism™ 3100 Genetic Analyzer and ABI Prism™ dRhodamine Terminator Cycle Sequencing Kit according to the manufacturer's protocol (Fostercity, CA).

3.1.5 Software

Search for protein and DNA sequences was carried out by BLAST program (http://www.ncbi.nlm.gov/BLAST). Comparison and alignment of homologous

protein and DNA sequences were performed using BioEdit program. Specific Oligonucleotide primers were designed with IDT Biotools OligoAnalyzer 3.0 program (http://207.32.43.70/biotools/oligocalc/oligocalc.asp). Mass spectrometry was carried out with Ettan MALDI-ToF software version I.II SPI (Amersham Bioscience).

3.2 Methods

3.2.1 Search for protein and DNA sequences of all genes

Search for protein and DNA sequences of all genes in relation to rice seeds storage protein (glutelin, prolamin, albumin and globulin), starch synthase and aspartate kinase pathway enzymes was performed by BLAST programme. Details of the total 36 genes were showed in the following tables.

3.2.1.1 Genes encoding rice glutelin family

Table 3.1 Cloned genes encoding rice glutelin family and their genbank accession numbers

Subfamily	Gene	Accession No.	Protein seq. (AA)	mRNA seq. (bp)
GluA	GluA-1 (Gt2)	L36819	499	1649
	GluA-2 (Gt1)	M28156	499	1675
	GluA-3 (Gt3)	X54313	496	>1522
	GluA-4 (Gt4)	pseudogene		
GluB	GluB-1	X54314	499	1672
	GluB-2	X54192	496	>1535
	GluB-3	pseudogene		
	GluB-4	AF537221	500	1619

3.2.1.2 Genes encoding rice prolamin family

Table 3.2 Cloned genes encoding rice prolamin family

Subfamily		Gene	Accession No.	Protein seq. (AA)	mRNA seq. (bp)
10kDa prolamin			X15231 AF29458	134	562
13kDa prolamin	ClassI	Prol 7	Y00747	150	720
		Prol14	M23744	148	717
		RP3	X71981	102	>306
	ClassII	Prol17	M23743	149	650
	ClassIII	RP5	AF156714	156	>471
		RP6	X65064	156	>564
		NRP33	D63901	156	>471
16kDa prolamin			X60979	158	612

3.2.1.3 Genes encoding rice albumin family

Table 3.3 Cloned genes encoding rice albumin family

Gene	Accession No.	Protein sequence (AA)	CDS (bp)
RA5	D11430	157	474
RA14	D11432	165	498
RA16 (RAG1)	D11433	166	501
RA17 (RAG2)	D11434	162	489

Note: CDS=coding sequences.

3.2.1.4 Genes encoding rice globulin family

Table 3.4 Cloned genes encoding rice globulin family

Gene	Accession No.	Protein sequence (AA)	mRNA sequence (bp)
26kDa	D50643	186	561
Globulin1	AF051153	>174	>524
Globulin2	AF051154	>136	>408
LMW globulin	X62091	165	498

3.2.1.5 Genes encoding rice starch synthesis enzymes

Table 3.5 Cloned genes encoding rice starch synthesis enzymes

Gene	Accession No.	Protein seq. (AA)	CDS seq. (bp)
GBSS	AF031162	609	1830
AGPase (small subunit)	AY028315	500	1503
AGPase (large subunit)	AY028314	518	1557
SSS1	D38221	626	1881
SSSII-1	AF383878	749	2250
SSSII-2	AF395537	694	2085
SSSII-3	AF419099	810	2433
SSSIII-2	AY100469	1553	4662
SSSIV-1	AY100470	962	2889
SSSIV-2	AY100471	919	2760
SBE-1	AF136268	820	2463
SBE-3	D16201	825	2478
SBE-4	E14723	841	2526
Pullulanase	AB012915	988	2967
Isoamylase	AB015615	>733	>2202

3.2.2 Alignment of homologous DNA sequences between family member genes

3.2.2.1 Storage protein gene families of rice seeds

According to solubility, rice seed storage proteins are classified into four types: acidic or basic solution soluble glutelin, salt-soluble globulin, water-soluble albumin, and alcohol-soluble prolamin. These storage protein fractions are mixtures of components that exhibit polymorphism arising from the presence of multigene families (Shewry, 1995). In order to profile the individual member gene expression pattern, it's necessary to analyze and compare the DNA sequences of family member genes one by one using designed specific primers. Alignment of homology between different members of a storage protein gene family was listed in Tables 3.6 to 3.9.

Table 3.6 Sequence homology in glutelin gene family

Identity (%)	Protein sequence						mRNA sequence					
	GtA1	GtA2	GtA3	GluB1	GluB2	GluB4	GtA1	GtA2	GtA3	GluB1	GluB2	GluB4
GtA1		95	81	62	62	30		95	79	60	65	37
GtA2			82	62	61	30			79	59	65	37
GtA3				62	64	29				60	66	38
GluB1					91	41					85	42
GluB2						42						47
GluB4												

Note: GluA4 and GluB3 are pseudogenes that cannot express during the rice seed developing stages. GluA2 and GluB2 were partial mRNA sequence.

Table 3.7 Sequence homology in globulin gene family

Identity (%)	Protein sequence				CDS			
	26kDa globulin	Gb 1	Gb 2	LMW Gb	26kDa globulin	Gb 1	Gb 2	LMW Gb
26kDa globulin		12	9	18		0	2	0
Gb 1			14	5			1	1
Gb 2				13				1
LMW Gb								

Note: Globulin 1 and globulin 2 are partial amino acid and mRNA sequences. CDS=coding sequences, Gb=globulin.

Table 3.8 Sequence homology in prolamin gene family

Identity (%)	Protein sequence							mRNA sequence						
	10 kDa	RP5	RP6	Pro 7	Prol 14	Prol 17	16 kDa	10 kDa	RP5	RP6	Pro 7	Prol 14	Prol 17	16 kDa
10kDa		23	21	21	21	23	23		2	1	0	0	0	3
RP5			93	71	69	74	96			91	61	60	55	90
RP6				71	71	73	89				62	61	55	82
Pro7					98	61	68					88	59	61
Prol14						60	66						62	60
Prol17							72							54
16kDa														

Note: RP5 and RP6 are of partial mRNA sequence.

Table 3. Sequence homology in albumin gene family

Identity	Protein sequence				CDS			
(%)	RA5	RA14	RA17	RAG2	RA5	RA14	RA17	RAG2
RA5		74	68	75		77	77	80
RA14			78	95			80	95
RA17				68				82
RAG2								

3.2.2.2 Rice starch synthase gene family

Starch is composed of two types of molecule, namely amylose and amylopectin. Amylose is an essentially linear molecule composed of alpha-1,4 linked glucosidic chains, whereas amylopectin is a highly branched glucan with alpha-1,6 linked glucosidic bonds that connect to the linear chains. Amylose is synthesized by ADP-Glc pyrophosphorylase (AGPase) and granule-bound starch synthase (GBSS), which is encoded by the *Waxy* gene. Amylopectin has a distinct fine structure called multiple cluster structure and is jointly produced by AGPase, soluble starch synthease (SSS), starch branching enzyme (SBE) and starch debranching enzyme (SDE), which are multiple subunits or isoforms except GBSS in rice. Alignments of these multiple isoforms were performed by Bioedit software and listed below in tables.

Table 3.10 CDS comparison between soluble starch synthase (SSS) isoforms

Identity (%)	Protein sequence							CDS						
	SSS 1	SSS II-1	SSS II-2	SSS II-3	SSS III-2	SSS IV-1	SSS IV-2	SSS 1	SSS II-1	SSS II-2	SSS II-3	SSS III-2	SSS IV-1	SSS IV-2
SSS1		0.45	0	0	0	0	0		0.32	0.35	0.33	0.1	0.19	0.18
SSSII-1			0.60	0.54	0	0	0			0.51	0.48	0.12	0.20	0.21
SSSII-2				0.62	0	0	0				0.54	0.11	0.19	0.20
SSSII-3					0	0	0					0.12	0.21	0.22
SSSIII-2						0	0						0.18	0.18
SSSIV-1							0.69							0.64
SSSIV-2														

Table 3.11 Sequence homology in SBE gene family

Identity (%)	Protein sequence			CDS		
	SBE1	SBE3	SBE4	SBE1	SBE3	SBE4
SBE1		46	45		50	48
SBE3			75			77
SBE4						

3.2.3 Primer design

According to the sequence homology alignment information, primer design for each gene family member was carried out using the software as mentioned in Section 3.1.5. Specific primers that could be designed for the gene members were listed in the following table.

Table 3.12 Specific primers for rice genes

Genes	Forward primer	Reverse primer	Amplified product length (bp)
Gt-1	5'-caa gag agc att atc aag aag g-3'	5'-tct tga ggg atg tcc tta gc-3'	735
Gt-2	5'-cag ttt gct gca gcc ata tc-3'	5'- atc atc att agt cga aag act gg-3'	824
Gt-3	5'-ccg agc aag acc aac aat tg-3'	5'- cca tag tct ctc gat tgc acc-3'	475
GluB1	5'-caa caa caa tcg ggc tca ac-3'	5'- cgc cga tac tag ttc tca c-3'	926
GluB2	5'- atg atg gtg atg cat cgg-3'	5'- agg ttt act ctg gta gca tc-3'	544
GluB4	5'- acg tta atg ccc ata gcc-3'	5'- ccg cca caa agt ttc aca-3'	467
10kDa prolamin	5'-atg gca gca tac acc agc-3'	5'-gca cac gat agt atg caa cac c-3'	460
RP5	5'- tac cag gcc att agt agc-3'	5'-atc aag ttt caa ctg tca cg-3'	268
RP6	5'-act gca gca gtt tag tgg-3'	5'-cct aag ttt caa cag tca ca-3'	229
Prol 7	5'-aat gga caa gcg caa tgg-3'	5'-acg tta aaa gcc aac aga gc-3'	466

Prol 14	5'-atc agc tgc agt cgc ctg-3'	5'-tta caa gac acc gcc aag g-3'	356
Prol 17	5'-aga tgc tta gcc cat gcg g-3'	5'-ctt ggg ctt gag ctt gag c-3'	249
26 kDa globulin	5'-gtg agc gag tcg gag atg-3'	5'-aca gca acg gtg tga tcc-3'	538
Gb1	5'-tac aac atc ctg agc ggc-3'	5'-ttc tgg ctg ttg agg acg-3'	326
Gb2	5'-gtt ctt cag cgc gtt tag c-3'	5'-ttc ggc gca atc att cga g-3'	343
LMW Gb	5'-cgt tgc tcc tca tca tcg-3'	5'-ctc cct gta gat gcc tcc-3'	293
RA5	5'-ctc agt gtt gct tct cgc-3'	5'-tag cag aca cca cct ccg-3'	331
RA14	5'-cac gag tcg aaa caa aaa gcc-3'	5'-tcc ctg tag atg cct cca-3'	341
RA17	5'-gct cga tca cat gct gtc-3'	5'-tta gca acc cca cat agt-3'	218
SBE1	5'-atg ctg tgt ctc acc tcc tc-3'	5'-cct cat aga tgc gtg gag c-3'	751
SBE3	5'-aag agg cag cat gca agt-3'	5'-tgg cag ttc gaa cgg tac-3'	260
SBE4	5'-aac tca agg gag ttg gtg-3'	5'-gcc aac cga tat cgt tat g-3'	266
GBSS	5'-aac agc tag aca acc acc-3'	5'-tag tcc tcg aaa gcg aag-3'	660
Pulluanase	5'-aag gtc tct gat tgc ctg-3'	5'-cag ctg agc ctc atc aac-3'	1026
Isoamylase	5'-ggt gac tga gga ggt tcc-3'	5'-tct cat tac cct cgg ctg-3'	689
SSS1	5'-ttg tgg cta gtg agc agg-3'	5'-gac tgg cac aag act ggc-3'	527
SSSII-1	5'-cag gtt tga cca aac cgg-3'	5'-agg tac atc tgt agg tca tag-3'	324
SSSII-2	5'-tgg aat cga agg ctg cac-3'	5'-ttg cca tca ccg tag acg-3'	548
SSSII-3	5'-ccg gtg acc ata acg aag ct-3'	5'-cga atc gtc atc ctg gtc g-3'	381
SSSIII-2	5'-gtg atg gat gaa gct aag gc-3'	5'-ccg agc ttc atc caa cac-3'	525

SSSIV-1	5'-ggc aat cca gca cat ttg tg-3'	5'-gct gat cca aca atg cag cat ac-3'	358
SSSIV-2	5'-gct gaa ggg cag ttc ggt-3'	5'-gtc cca gaa cca tag cag ca-3'	344
AGPase SS	5'-cag cac gag tgt tct tgg g-3'	5'-cct ttg ccc tct cag tat caa g-3'	617
AGPase LS	5'-gaa agc act gag gaa gag gt-3'	5'-tag tca gaa gct cgg ctc c-3'	611
AK	5'-gtc tac acg gat ttt tgc tgc-3'	5'-tca tgt gaa ggg ctt ctt cct-3'	1046
DHPS	5'-ggt gct gaa ggt gta ata gtg gg-3'	5'-aac cct tgc ttc ttt ctg acc c-3'	768

Note: SS=small subunit; LS=large subunit.

3.2.4 Collection of developing hybrid rice seeds

Because the female parental hybrid rice line (P64S) is sterile at higher temperature (>23 °C) and fertile at low temperature (<23 °C), the selfcrossing seeds of P64S were collected in the experimental field treated with cold water (19-21 °C average) from the "Xiu Mountain" reservoir in Liuyang county, Changsha city, Hunan province. This cold treatment induces and converts male sterility into fertility at the fertility sensitive stage, i.e. during pollen mother cell formation. Male parental hybrid rice line 9311 is relatively easy to selfcross seeds in different stages due to its fertility, bigger stigma, and vigor pollen activity. The F1 hybrid seeds were acquired by artificial pollination. In the first day afternoon, we removed all opened flowers of female parental line to make sure all the opened flowers in the second day are pollinated by male parental line. In the second day afternoon we cut off all the unopened florets so that same day pollinated flowers could be obtained. Finally,

wearing gloves and facemasks, we colleted maturing seeds at the 3, 6, 10, 15, 20 days after fertilization (DAF). All the harvested seeds were packed in aluminum foil paper and submerged into liquid nitrogen for short time storage and put into -80°C freezer for long-term storage. A total of 90,000 seeds (3 x 5 x 6,000) were labeled and collected for three rice lines (the two parents and their hybrid) and 5 different maturing stages (3, 6, 10, 15, 20 DAF) by assistant workers and myself in Hunan Hybrid Rice Research Center field. The following pictures recorded flower labeling and seeds collection at different developing stages.

Figure 3.1 Labeling, artificial pollination and bagging of the flowers of F1 hybrid (P64S/9311).

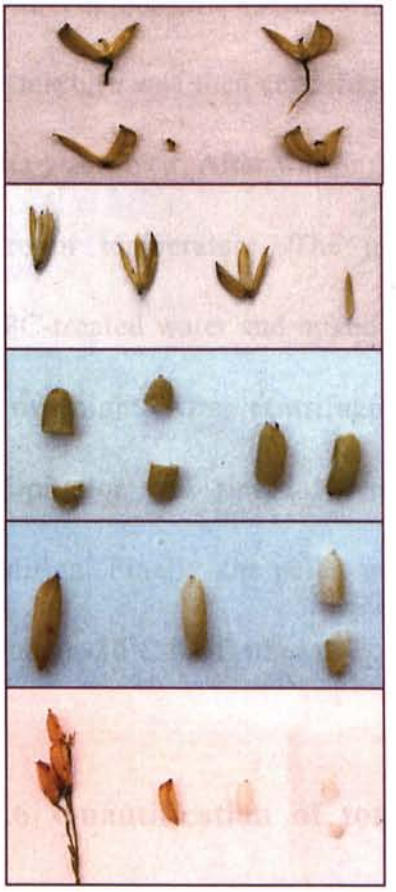
Figure 3.2 Labeling the florets of the male parental line (9311).

Figure 3.3 labeling the florets of the female parental line (P64S).

The right panel showed a labeling tag with line name and collection date.

Figure 3.4 Collection of developing rice seeds at appropriate stage with gloves and facemask.

Figure 3.5 Putting the collected seeds into liquid nitrogen for short-term storage.



Seeds collected 3 days after fertilization (7/8/2001~10/8/2001)

Seeds collected 6 days after fertilization (7/8/2001~13/8/2001)

Seeds collected 10 days after fertilization (7/8/2001~17/8/2001)

Seeds collected 15 days after fertilization (7/8/2001~22/8/2001)

Seeds collected 20 days after fertilization (7/8/2001~27/8/2001)

Figure 3.6 Developing rice seeds at 3, 6, 10, 15, 20 DAF.

3.2.5 Total RNA extraction

Total RNA was extracted from developing rice seeds using the cold-phenol method of Feiqi Zhen (1993). About 10 g of the developing seeds at 3, 6, 10, 15, or 20 DAF were ground into fine powder in liquid nitrogen with homogenizer (IKA ultra turrax T18 basic, USA). And 0.5 ml extraction buffer (0.05 M Tris-HCL, pH 9.0, 0.02 M NaCl, 0.002 M EDTA, 1% SDS) and equal volume of saturated phenol in Tris buffer (pH 8.0) were added. The mixture was vortexed for 10 min on ice. Then the mixture was centrifuged at 4000 rpm for 20 min at 4°C. The supernatant was transferred to a new falcon tube and two volumes of cold 100% ethanol were added. The mixture was inverted for several times and put on ice for 30 min. The mixture was then centrifuged at 4000 rpm for 20min at 4°C again and the pellet was saved. After washing one time by 70% ethanol, the pellet was dried at room temperature. The pellet was then dissolved with 1 ml cold DEPC-treated water and mixed with 1/3 volume of 8 M LiCl and kept at 4°C for overnight. After centrifuging at 4000 rpm for 10 min at 4°C, the RNA precipitation was rinsed twice with 70% ethanol and dried under natural condition. Finally, the pellet was dissolved in 2 ml DEPC-treated water and stored at -80°C for further use.

3.2.6 Quantification of total RNA and determination of internal control

The quality of total RNA used for Northern blotting and RT-PCR was mixed with

master-mix solution [1 x MOPS [3-(N-Morpholino) Propanesulfonic acid], 50% formamide, 17.5% formaldehyde, 0.02 μg EtBr and 1 x loading buffer] and checked by gel electrophoresis in a 1% agarose/formaldehyde gel (100 ml gel solution contained 1 g agarose, 87 ml DEPC-treated H_2O, 3 ml formamide and 10 ml 10 x MOPS) and the quantification of total RNA was carried out by a gel densitometer. In this study, an internal control was used to further confirm that even quantity of total RNA samples from developing rice seeds (3, 6, 10, 15, 20 DAF) of different materials (P64S, 9311, F1 hybrid) was used. In order to determine which gene can act as an internal control for this study I tested the actin gene and 18S ribosome RNA (18S rRNA). Results showed that the actin gene is unfit for internal control in this project because it was expressed at different levels during the developing rice seed stage. The expression level of the 18S ribosome RNA, however, was found to remain constant at different developing stage. PCR cycles are also an important factor in comparison of RT-PCR products. Therefore, selection of optimum PCR cycles was carried out using different volumes of diluted 1st strand cDNA samples (0.5μl, 1 μl, 2 μl, 3 μl and 5 μl) as template and 5 PCR cycles (5, 10, 15, 20, 30). The RT-PCR products were checked by 1% agarose gel and transferred onto positive charge nylon membrane with a VacuGene XL vacuum blotting system (Pharmacia Biotech AB, Sweden). After UV (254 nm) cross linking and prehybridization in hybridization buffer [50% formamide, 5 x SSC, 2% blocking reagent, 50 mM Na-phosphate (pH 7.0), 0.1% N-lauroylsarcosine, 7% SDS] at 42°C for 3 hr, hybridization was carried out with antisense single-strand DIG-labeled 18S rRNA probe for at least 16 hr. Then washed,

blocked and chemiluminescent detected by anti-DIG-AP and CSPD as substrate.

3.2.7 RT-PCR (Reverse-transcription polymerase chain reaction)

One μg quantitated total RNA from developing rice seeds at 3, 6, 10, 15, 20 DAF of different varieties (P64S, 9311 and F1) was treated with 1 unit RNase-free DNaseI in 10 μl reaction mixture at room temperature for 10 min to remove DNA contamination. The treated total RNA sample was denatured by heating at 65°C for 5 min in a 10 μl reaction mixture containing 1 μg total RNA and 1 μl Oligo$(dT)_{15}$ primer and then cooled on ice immediately. The 10 μl Moloney murine leukemia virus (M-MLV) reverse transcriptase mixture containing 20 units M-MLV reverse transcriptase, 0.5 x M-MLV reaction buffer, 0.5 mM dNTPs, 10 mM DTT, and 2 units rRNasin RNase inhibitor was added into the above template/primer mixture and the first-strand cDNA synthesis was performed at 42°C for 1 hr. Then the reaction was stopped by putting on ice. The completed first-strand cDNA mixture was diluted 10 times with distilled water. After that, the second-strand cDNA synthesis was carried out in a 25 μl reagent containing 1x PCR reaction buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.8 ng forward primer, 0.8 ng reverse primer, 0.4 units Taq DNA polymerase, 5 μl diluted first-strand cDNA as template and appropriate volume of DEPC-treated distilled water. PCR conditions were varied depending on different genes. Amplification of most genes was carried out using conditions as follow: denaturation at 95°C for 5 min, then 30 cycles at 94°C for 50 sec,

50°C for 50 sec, 72°C for 50 sec, followed by 1 cycle of 72°C for 10 min. Some fragments were amplified by 25 cycles or 50 cycles and annealing temperature set at 55°C. The results of RT-PCR were detected with Gel-Dol 1000 (Bio-Rad) and quantited by molecular analysis (MA) program.

3.2.8 Northern blotting analysis

The anti-sense DIG-labeled DNA probe was prepared according to the DIG DNA Labeling Kit (Boehringer Mannheim) through PCR amplification. The first round PCR was performed in 25 µl PCR reaction mixture containing 50 ng RT-PCR product as template, 1x PCR reaction buffer, 0.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.8 ng forward primer, 0.8 ng reverse primer, and 0.4 units Taq DNA polymerase. The PCR programme was carried out as described in Section 3.2.7 except the number of amplification cycle was 55. In the second round PCR, the single-stranded DIG-labeled probe was produced by amplifying in 100 µl reaction of the followed reagent containing 50 ng the first round PCR product, 1 x reaction buffer, 0.5 mM $MgCl_2$, 0.02 mM DIG-labeled dNTPs, 0.8 ng reverse primer, 1 unit Taq DNA polymerase and appropriate volume distilled water. The PCR condition was the same as the first round PCR. The yield of single-stranded DIG-labeled DNA probe was determined by spot test using DIG-labeled control DNA.

Five µg quantified total RNA was separated by 1% agarose/formaldehyde gel electrophoresis and transferred onto the positive charge nylon membrane (Boehringer Mannheim, F.R.G) using capillary action for at least 16 hr. After that, the transferred

total RNA was fixed onto the membrane by UV (250nm) crossing linker. This was followed by prehybridization at 50°C for 3 hr and hybridization with the anti-sense DIG-labeled probe at 50°C for at least 16 hr in hybridization buffer as described in Section 3.2.6. Washing, blocking and signal detection using the Anti-Digoxigenin-AP (alkaline phosphatase) were preformed according to the method described in the DIG Nucleic Acid Detection Kit (Boehringer Mannheim). The results of Northern blotting were evaluated with Model GS-690 Imaging Densitometer (Bio-Rad). For each gene profiling expression, repeated RT-PCR results for at least 3 times. The curve graph of gene expression pattern was carried out by the average value.

3.2.9 DNA sequencing

DNA sequencing was performed using ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq-FS. Sixty ng of RT-PCR product was used as template and mixed with 1.6 μM forward or reverse primer, 4 μl terminator ready reaction mix [TRRM (from ABI company)] and trace amount of deionied water to 10 μl in a 0.2 ml PCR tube. The mixture was amplified in a PCR machine by sequencing PCR cycles as follow: the thermal cycler was pre-warmed at 96°C, followed by 25 cycles at 96°C for 10 sec, 50°C for 5 sec and 60°C for 4 min. The sequencing PCR products were transferred to a new 1.5 ml microtube and vortexed with 1 μl 3 M sodium acetate, pH 5.2, 25 μl 95% ethanol, 0.5 μl glycogen (20 μg/μl). The mixture was left in -20°C freezer for 10 min and centrifuged at 15,000 rpm for 30

min. The supernatant was carefully removed as much as possible without disturbing the visible pellets and the pellets were washed with 200 µl 75% ethanol. After spinning at 15,000 rpm for 5 min, the pellets were vacuum dried and dissolved in 12 µl Hi-Di formamide. The dissolved sample was heated at 99°C for 2 min and the entire sample was transferred to a 96 wells sequencing sample plate. Finally, the plate was loaded into ABI 3100 genetic analyzer. The sequencing machine has the capacity of 3 hr for 16 x samples.

3.2.10 Protein extraction

3.2.10.1 Extraction of four kinds of rice seed storage proteins

Extraction of four kinds of proteins was carried out following the procedure of Hiroshi Yamagata (1982). The rice seeds of five developing stages from P64S (female parental line), 9311 (male parental line) and P64S/9311 (F1 hybrid), respectively, were dried in a freeze dry system [China scientific (HK) Ltd.] for 24-48 hr. After removing the husk, the seeds were grounded into fine powder with small mortars and pestles. For each sample, 0.0100 g (10 mg) fine power was used. 100 µl albumin extraction buffer containing 10 mM Tris-HCl, pH 7.5, 1 mM EDTA was added and albumin was extracted by shaking at 1,400 rpm at 25°C for 1-3 hr. The extract was centrifuged at 15,000 rpm at 4°C for 20 min and the supernatant (about 95 µl) was recovered as much as possible without disturbing the precipitate and transferred to a new eppendorf tube. The precipitate was used for another extraction, which was mixed with 100 µl globulin extraction buffer (0.5 M NaCl, 10 mM Tris-HCl, pH 7.5,

1 mM EDTA) and stirred at 1,400 rpm at 25°C for 1-3 hr. The mixture was centrifuged as the same conditions as described above. After removing the upper phase into a new microtube, 100 µl prolamin extraction buffer containing 70% ethanol, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA was added into the retained precipitate. This was followed by vigorously stirring at 1,400 rpm at 25°C for 1-3 hr and centrifuge at 4°C at 15,000 rpm for 20 min. The remaining steps were performed according to procedure as described above to obtain prolamin. The precipitate was vortexed with 100 µl glutelin extraction buffer (0.02N NaOH) and glutelin was obtained using the protocol as above. The same method was used to acquire the total protein, with the total protein extraction buffer (125 mM Tris-HCl, pH 6.8, 4 M urea, 4% SDS, 5% β-mercaptoethanol) adding into 10 mg rice seed fine powder. All the protein samples were stored at -20°C for tricine SDS-PAGE electrophoresis.

3.2.10.2 Extraction of the *Wx* protein

Extraction of the *Wx* protein was carried out according to the procedure of Sano (1985). The maturing seeds from 3, 6, 10, 15, or 20 DAF was dried at 60 °C oven for 5 hr and ground into fine powder with pestle and mortar. All fine powders were washed for 3-4 times by SDS buffer containing 50mM, Tris-HCl, pH6.8, 2.8% SDS, 10% glycerol, 5% β-mercaptoethanol. For each time, the mixture was centrifuge at 11,000 rpm for 10 min and the precipitate was saved. Then the precipitate was purified with acetone for 3 times. This was followed by vacuum dry to get dried powder. Five mg treated powder was mixed with 50 µl SDS buffer and heated for 5

min at 99 °C. After cooled it down, 100 µl SDS buffer was added and the mixture was centrifuged at 11,000 rpm for 10 min. The supernatant was transferred to a new microtube and stored at -20 °C for tricine SDS-PAGE electrophoresis.

3.2.11 Tricine SDS-PAGE

Investigation of the expression pattern of all types of protein was performed by Tricine SDS-PAGE (Schagger, 1987). Ten µl protein (about 10 µg) was mixed with equal volume of 2 x sample loading buffer containing 24% glycerol, 0.1 M Tris-base, 8% SDS, 0.005% bromophenol blue and 4% β-mercaptoethanol and treated at 99°C for 10 min. The treated samples were spinned down and loaded into the wells of 1.0 mm slab tricine SDS-PAGE gel. Then the protein samples were separated by 16.5% polyacrylamide separating gel with a 10% polyacrylamide spacer gel and a 4% polyacrylamide stacking gel. Electrophoresis was run on a Pharmacia/LKB electrophoresis system with 35 voltages for stacking gel and 100 voltages for separating gel at room temperature. The anode buffer containing 0.2 M Tris-base (pH 8.9) and cathode buffer containing 100 mM Tris-base, 100 mM tricine, 0.1% SDS (pH 8.25) were used as outer and inner running buffer, respectively. After about 4 hr electrophoresis, the tricine SDS-PAGE gels were removed from gel tank and stained with staining solution (0.1% Coomassie brilliant blue G-250 and 1% methanol) overnight, then destained in destaining solution (methanol: 100% glacial acetic acid: water, 20: 6: 55).

3.2.12 Determination of crude protein and amylose content in P64S, 9311 and F1 hybrid

3.2.12.1 Determination of crude protein

Determination of crude protein in cereals was performed with semi-micro kjeldahl method according to Chinese national standard (1983) at the laboratory of rice quality analysis in Hunan National Hybrid Rice Engineering and Research Center. About 20 g of removed husk rice grain were dried in 60-65°C oven for more than 8 hr. The dried seeds were ground with grinder and selected by 40-mesh sieve. Then two samples of 0.1000 g fine powder, one for determination of crude protein and another for detection of water content were weighed. The test samples were transferred into 50 ml Kjeldahl flask and stirred with 0.5 g accelerator (10 g $CuSO_4.5H_2O$ mixed with 100 g K_2SO_4) and ground into fine powder with mortar and pestle. The powder was passed through a 40-mesh sieve and mixed with 3 ml solution (H_2O_2: H_2SO_4: H_2O, 3:2:1). The sample in the Kjeldahl flask with small swan-neck funnel slanting ways was put on electronic oven for boiling at 175 voltages for 5 min and 200 voltages for 25 min. The boiled sample in the reaction tube of semi-micro distillation system was then transferred and mixed with 10 ml NaOH, followed by distilling and condensing in the Erlenmeyer flask (about 50 ml distillates) with 10 ml H_3BO_3-indicator mixed solution [0.5 g bromocresol green and 0.1 g methyl red were dissolved with 95% ethanol firstly and water was added to 100 ml. The indicator was then mixed with 2% (w/v) H_3BO_3 at 1: 100, pH 4.5, color in gray-purple]. After that, the 50 ml distillate was titrated by 0.2 N HCL standard solutions until blue-green changed into

gray-purple color. Finally, the crude protein percentage was calculated using the following formula:

$$\textbf{Crude protein (\%)} = \frac{(V_2\text{-}V_1) \times N \times 0.0140 \times K \times 100}{W \times (100\text{-}X)} \times 100$$

V2: volume of consumed acid standard solution for titration of test sample (ml)

V1: volume of consumed acid standard solution for titration of blank (ml)

N: normality of acid standard solution

K: coefficient of nitrogen converse into crude protein

W: weight of the test sample

X: water content of the test sample

0.0140: N_2 g per equivalent mg
In this study, 0.1 g sugar was used as blank.

3.2.12.2 Determination of amylose content

0.1000 g rice fine powder was passed through a 100-mesh sieve and put into 100 ml volumetric flask. 1.0 ml 95% ethanol was added firstly and 9.0 ml 1M NaOH was then slowly added and boiled for 10 min. After the mixture was cooled down to room temperature the volume was fixed volume to 100 ml with dH_2O. 0.5 ml sample solution was took out and mixed with 1.0 ml 1 M acetic acid, 1.5 ml iodine solution (0.2% iodine and 2% iodine potassium) in a new 100 ml volumetric flask. The volume was fixed to 100 ml and kept static situation for 20 min. This sample solution was detected absorbance value at 620 nm by spectrophotometer. Another 100 ml fixed

volume containing 5 ml 0.09 M NaOH, 1.0 ml 1 M acetic acid and 1.5 ml iodine solution was used as blank for zero adjustment. The standard curve was described by amylose content of standard sample as y-axis and corresponding absorbance value as x-axis. The experiment was repeated once for each test sample, The amylose content of test sample was calculated by following formula:

$$Y = a + bx$$

Y: amylose content of test sample (%)
a: intercept of standard curve
b: slope of standard curve
x: absorbancy value of test sample

3.2.13 Two-dimension gel electrophoresis

3.2.13.1 Cleaning up of protein sample for 2-D gel

The total protein was prepared according to method described in Section 3.2.10. 100 µl total protein sample were mixed with 300 µl precipitant (PlusOne™ 2-D Clean-up kit, Amersham) and incubated on ice for 15 min, 300 µl co-precipitant (PlusOne™ 2-D Clean-up kit, Amersham) was added into the mixture of protein and precipitant and centrifuged at 12,000 rpm for 5 min. After removing the supernatant as much as possible without disturbing the pellet 25 µl DI water was added on top of pellet and votexed for 5-10 s. One ml of wash buffer (pre-chilled for at least 1 hr at -20°C) and 5 µl wash additive were added and votexed until the protein pellet was fully dispersed. Following by incubating the tubes at -20°C for at least 30 min, the content was vortexed for 20-30 s once every 10 min and centrifuged at 12,000 rpm for 5 min. The supernatant was discarded and air dried the white visible pellet. Finally,

the pellet was resuspended in appropriate volume of rehydration solution containing 8 M urea, 2% (w/v) CHAPS, 0.5% (v/v) carrier ampholyte, 40 mM DTT, and 0.002% bromophenol blue.

3.2.13.2 Quantification of protein samples

BSA standard solution was used for standard curve. 500 µl precipitant (PlusOne™ 2-D Clean-up kit, Amersham) was added into 50 µl cleaned-up total protein sample (including the standard solution) and incubated for 2-3 min at room temperature. Then added 500 µl co-precipitant (PlusOne™ 2-D Clean-up kit, Amersham) and centrifuged at 10,000 rpm for 5 min. Followed by decanting the supernatants and adding 100 µl of copper solution and 400 µl of DI water. After dissolving the precipitated protein, 1 ml of working color reagent (reagent A: reagent B, 100:1) was added and incubated for 15-20 min at room temperature. At last, read the absorbance of each sample and standard at 480 nm using water as the reference (PlusOne™ 2-D Clean-up kit, Amersham).

3.2.13.3 First-dimension IEF (isoelectric focusing)

Fifty µg of quantitative protein were mixed with 450 µl rehydration solution including 8 M urea, 0.5% CHAPS, 0.2% DTT, 0.5% IPG buffer (pH3-10), 0.002% bromophenol blue and applied into the strip holder. Then generally removed the protective cover foil from the immobilized pH gradient (IPG) strip starting at the acidic (pointed) end, position the IPG strip with the gel side down and the pointed end

of the strip directed toward the pointed end of the strip holder. Be careful not to trap air bubbles under the IPG strip, and then applied IPG cover fluid to cover half of the IPG strip. The first-dimension isoelectric focusing was performed by rehydration for 12 hr at 30 voltages, followed by holding at 500 voltages for 1 hr (500 vhr) and 1000 voltages for 1 hr (1000 vhr), finally IEF at 8000 voltages for 8-10 hr (64,000 vhr-80,000 vhr).

3.2.13.4 IPG strips equilibration

The IPG strips in individual tubes with the gel facing up were mixed with 10 ml SDS equilibration solution containing 50 mM Tris-HCl, pH 8.8, 6 M urea, 30%(v/v) glycerol, 2%(w/v) SDS, 0.002%(w/v) and 100 mg DTT (fresh prepared) and let equilibrate for 15 min. The solution was decanted and 10 ml the second SDS equilibration buffer with 250 mg iodoacetamide (IAA) instead of 100 mg DTT was added to equilibrate for 15 min again. These treated IPG strips were ready for the second-dimension SDS PAGE.

3.2.13.5 Second-dimension SDS PAGE

The second-dimension SDS PAGE was carried out using the Ettan DALT*twelve* system. 12.5% polyacrylamide precast gel was inserted into the Ettan DALT precast gel cassette with the roller to press out any bubbles or liquid from between the gel and the glass, and then gently pushed the equilibrated IPG strips down to contact the gel slab with a thin plastic ruler. This was followed by sealing the IPG strips using

agarose sealing solution [SDS electrophoresis buffer (25mM Tris-base, 192 mM glycine, 0.1% SDS), 0.5% agarose, 0.002% bromophenol blue]. The second-dimension SDS PAGE was performed under following conditions: 2.5 W/gel for 30 min, then 17 W/gel for 4.5 hr at 25°C until the dye front is approximately 1 mm from the bottom of the gel.

3.2.13.6 Silver staining of 2-D gel

The gels were fixed with fixation solution (40% ethanol, 10% acetic acid and 50% distilled water) twice, for 60 min each time. Then sensitized the gels by sensitizing solution without glutardialdehyde [30% ethanol, 0.2% sodium thiosulfate, 6.8%(w/v) sodium acetate] for 60 min. This was followed by washing the gels with distilled water five times, for 8 min every time (total 40 min) and silver reaction in a 0.25% silver nitrate solution without formaldehye for 60 min. After rinsed them four times (1 min each time) the gels were developed in developing solution (2.5% sodium carbonate, 0.148‰ formaldehyde) until appropriate protein spots were found. The gels was then transferred to the stop solution (1.46% EDTA-Na_2) for 45 min and washed with distilled water two times, 20 min every time. Finally, the gels were preserved in preservation solution including 8.7% glycerol.

3.2.14 MALDI-ToF mass spectrometry (Matrix Assisted Laser Desorption Ionization-time of Flight)

3.2.14.1 Sample destaining

The available spots were picked and cut from the silver staining gels with sterilized knife. The protein spots were split into small pieces and put into a new microtube, 20 µl mixture solution (30 mM potassium ferricyanide: 100 mM sodium thiosulphate, 1:1) was added. Once the dark stain had been removed, the spots were washed with distilled water for several times. The protein spots were then equilibrated in 20 mM ammonium bicarbonate 10 min for 2 times.

3.2.14.2 In-gel digestion with trypsin enzyme

The destained protein spots were dehydrated with 25 µl acetonitrile (ACN) 3 times, for 10 min each time and dried in speed vac (Gene Co. Ltd.) for 5 min until the spots appeared "dust like". These spots were rehydrated with 10 µl trypsin enzyme in buffer (40ng/ul trypsin in 50 mM ammonium bicarbonate) and incubated on ice for 30 min. More 50 mM ammonium bicarbonate was added to cover the hydrated spots and digested at 30°C for overnight. In the second day, 20 µl 50 mM ammonium bicarbonate was added into the digested spots and sonicated for 10 min with ultrasonic cleaner [China Scientific (HK) Ltd.]. Then 20 µl of the second extraction buffer [ACN: 5% trifluoroacetic acid (TFA) = 1:1] was added and sonicated for 10 min again. After removing the supernatant into a new microtube extraction was repeated for 7 times with the second extraction buffer and the supernatant extracts were combined. Finally, dried the supernatant extracts with speed vac.

3.2.14.3 Desalination of the digested sample with Zip Tip

The vacuum–dried sample was resuspended in 10 µl 0.1% TFA. The Zip Tip was wetted by aspirated 10 µl wetting solution (50% ACN) into the tip and dispensed to waste for twice. The wetted Zip Tip was equilibrated by aspirating 10 µl equilibration solution (0.1% TFA) into the tip and dispensed to waste twice. After binding the sample into Zip Tip by aspirating and dispensing 10 cycles in the eppendorf tube the Zip tip was washed with 10 µl washing solution (0.1% TFA) by aspirating and dispensing into waste for 5 times. At last, the digested peptides were eluted with 2 µl elution solution by aspirating and dispensing for 5 times in a clean 0.5 ml microtube.

3.2.14.4 Mass spectrometry

The matrix [ACHCA (alpha-cyano-4-hydroxy cinnamic acid)] powder was mixed with buffer A (50% ACN, 0.5% TFA) for 30 s in a microtube, then the tube was sit on the bench for a few min until all matrix has settled at the bottom, the supernatant was transferred to a new microtube. Three µl treated samples were spotted on the slice and incubated at room temperature until air dried. Then 0.4 µl matrix was loaded on the dried sample. After matrix was air dried the slice was loaded into the MALDI-ToF machine (Amersham Biosecience) and mass spectrometry was evaluated by Ettan MALDI-ToF software version I.II SPI (Amersham Bioscience).

Chapter 4 Results

4.1 Quantification of the total RNA from seeds at different developing stages

Since profiling gene expression of rice strains (varieties) at different developing stages can be carried out through RT-PCR and northern blot analysis using equal amount of total RNA as template, accurate quantification of total RNA from each variety or stage is thus a prerequisite for the present study. In this work, 15 total RNA samples, extracted from developing seeds 3, 6, 10, 15, or 20 DAF, of three varieties, P64S, 9311, and their F1 hybrid, were quantified together as described in Section 3.2.6. After several adjustments through checking with 1% agrose/formamide gel, accurate concentration of total RNA samples (about 1 μg/μl) could be obtained (Figure 4.1).

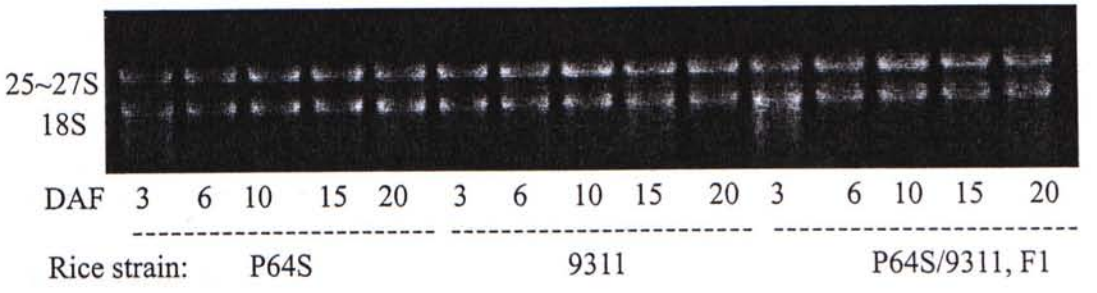


Figure 4.1 Confirmation of equal concentration of total RNA from seeds at various developing stages of the three rice strains by electrophoresis

One μg total RNA of each sample was loading in each lane. The numbers represent days after fertilization (DAF).

4.2 Determination of internal control

An internal control serving as a basis in comparing the level of gene expression at different tissues and at developing stages is another requirement in this study. In general, housekeeping genes, such as actin and 18S rRNA, are used as internal control because their expression levels do not changed significantly in/at different varieties, tissues, or developing stages. In this study, the rice 18S rRNA is used as an internal control. Equal concentration of the 15 RNA samples was confirmed based on the amount of 18S rRNA by optimized RT-PCR conditions.

To optimize RT-PCR conditions, appropriate PCR cycle number was first determined for synthesizing desirable amount of cDNA. The 1st strand of cDNA was synthesized using about 1 μg quantified total RNA and a reverse primer of 18S rRNA gene (5'-AACCAGACAAAGCGCTCCAC-3'). Then, double strand cDNA was produced by PCR amplification using various cycles and different volumes of diluted 1st strand of cDNA (0.5, 1, 2, 3, and 5 μl) as template, in the presence of forward primer (5'-AAACGGCGACCACATCCAAG-3') and reverse primer of the rice 18S rRNA gene. Equal volume of synthesized cDNA from different RNA samples was used for gel electrophoresis analysis, and transferred onto a positive charge nylon membrane for Southern blot analysis to select the optimum number of PCR cycles. Results showed that cDNA products could not be or only a weakly detected on a 1% agarose gel after PCR amplification for fewer cycles (10 and 15 cycles), but detectable when the number of PCR cycles was increased (20, 25, and 30 cycles) (Figure 4.2-A). This was confirmed by Southern blot analysis (Figure 4.2-B). For 20

cycles, when high amounts of the 1st strand cDNA were used as template, more double stranded cDNAs were detected. But the cDNA products were almost saturated with different concentrations of templates when 25 and 30 cycles were used (Figure 4.2-B). It thus can be concluded that the amount of cDNA product is positively correlating to the amount of template under 20 cycles of PCR. Therefore, 20 cycles were optimum for the further RT-PCR analysis.

With these optimized RT-PCR conditions, equal amount of 1st strand cDNA as template for 20 cycles, was performed to further confirm the equal quality of the total RNA from all the 15 samples. The results of RT-PCR demonstrated that the constant RT-PCR products were amplified at different varieties and stages (Figure 4.3), indicating that the 18S rRNA can serve as internal control.

The ability to measure equal amount of total RNA from 15 samples was further confirmed by northern blot using DIG-labeled anti-sense18S rRNA as a probe. One μg of total RNA from each sample was separated on 1% agarose gel containing formamide and transferred onto the positive charge nylon membrane by capillary absorption. Northern blot analysis was carried out according to the method as described in 3.2.8. Results (Figure 4.4) strongly indicated that the expression level of 18S rRNA gene was near parallel in deferent materials and at different developing stages, further proving that the total RNA concentration from the 15 samples was very close if not equal.

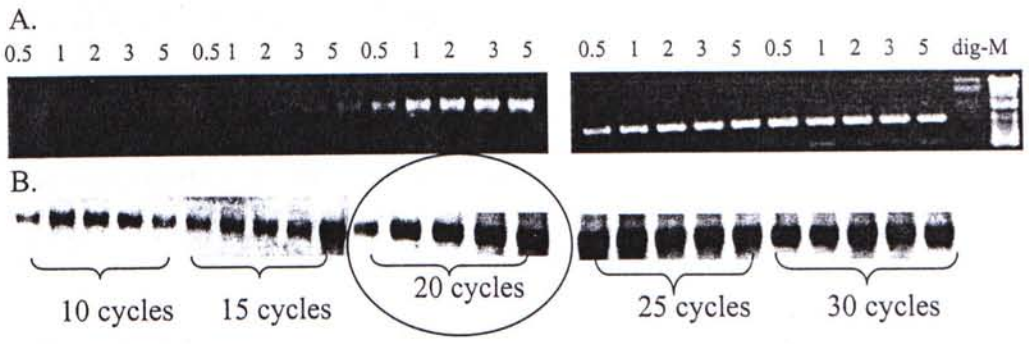


Figure 4.2 Electrophoresis and Southern blot analysis to show cDNA amounts after PCR amplification with different cycles and template concentrations.

A: RT-PCR products using different concentrations of template and PCR cycles. B: Corresponding Southern blot analysis.

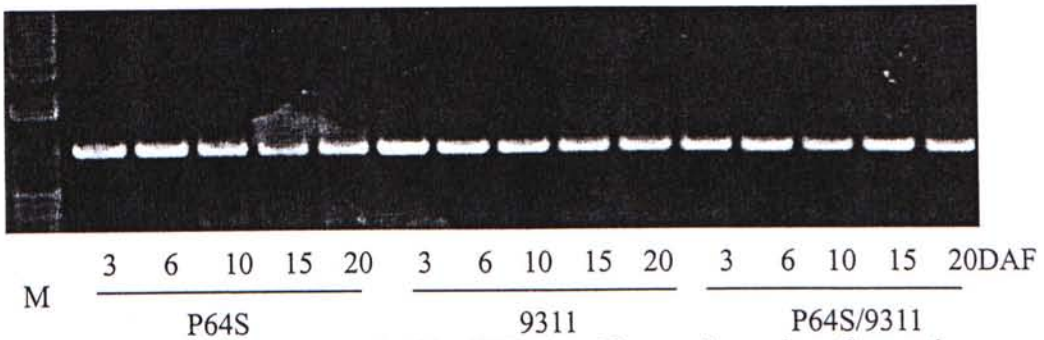


Figure 4.3 RT-PCR analysis of 18S rRNA in 15 samples using the optimum number of cycles.

Five μl of 10 times diluted 1st cDNA were used as template for PCR with 18S rRNA primers. 10 μl of RT-PCR products for each sample were checked by 1% agarose gel electrophoresis. M, 1kb plus DNA Marker; the numbers mean the days after fertilization (DAF).

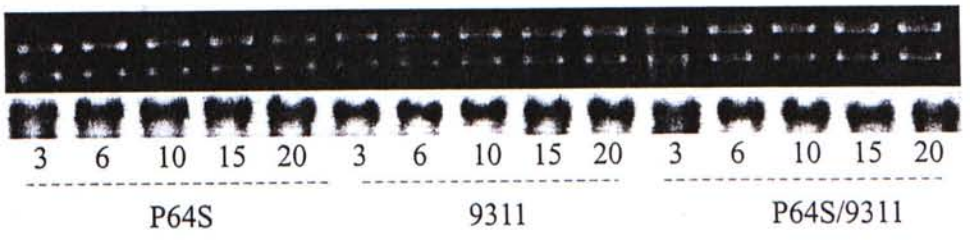


Figure 4.4 Northern blot analysis of 18S rRNA in 15 samples at different developing stages.

The upper image shows electrophoresis of the total RNA (1 μg/lane) from parental lines and F1 hybrid at different developing stages on 1% agrose/formamide gel. The lower picture represents Northern blot analysis using 1 μg total RNA.

4.3 DNA sequence analysis

To confirm the identity of the RT-PCR product for each gene, DNA sequencing was performed by the ABI 3100 analysis system. All the RT-PCR products were thus identified authentic with all the genes under study. Details of all the genes included in this study were as listed in Table 4.1.

Table 4.1 DNA sequence analysis of RT-PCR products for individual genes under study.

Gene Family	Subfamily	Genes	Sequence identity (%)	Family	Subfamily	Genes	Sequence identity (%)
Glutelin multigene family	A	Gt1	98	Albumin multigene family		RA5	95
		Gt2	98			RA14	99
		GluA-3	100			RA17	99
	B	GluB-1	98	Granular-bound starch synthase		GBSS	99
		GluB-2	97	Starch branch enzyme		SBE-1	99
		GluB-4	99			SBE-3	100
Prolamin multigene family	10kDa prolamin	10kDa prolamin	100			SBE-4	97
	13kDa prolamin	RP5	100	Soluble starch synthase		SSS1	99
		RP6	99			SSSII-1	100
		Prol 7	99			SSSII-2	98
		Prol 14	99			SSSII-3	99
		Prol 17	98			SSSIII-2	99
Globulin multigene family		26kDa Gb	99			SSSIV-1	99
		Gb 1	97.5			SSSIV-2	98
		Gb 2	100	ADP-G-Pyrophosphorylase		AGPase LS	99
		LMW Gb	98			AGPase SS	99
Aspartate family amino acids biosynthesis pathway		AK	99	Starch debranch enzyme		isoamylase	99
		DHPS	98			pullulanase	99

4.4 Profiling the expression of genes encoding rice seed storage proteins

4.4.1 The glutelin genes

Rice glutelin is encoded by many genes, which are classified into at least two subfamilies, GluA & GluB, according to their homologous DNA and deduced amino acid sequences. For example, at least three genes, namely GluA-1 (Gt2), GluA-2 (Gt1), and GluA-3 (Gt3), belong to the GluA subfamily; while GluB subfamily contains GluB-1, GluB-2 and GluB-4. As the DNA sequence identity among most of the glutelin genes is high (85-95%), to profile the expression of most of these genes thus had to employ RT-PCR using specific primers, except the genes encoding GluB-4 and GluA-3, which were investigated by northern blot analysis because of their relatively low DNA sequence homology.

RT-PCR was carried out using the equal amount of total RNA for the 15 samples simultaneously with three duplicates for each sample. Ten μl of RT-PCR products of each sample were checked on the same agarose gel and quantified using Model GS-690 Imaging Densitometer (Bio-Rad). RT-PCR analysis of each gene was repeated at least 3 times, while 2 times for these by northern blot analysis. The average values of these samples were used to generate the curve graphs for the expression patterns. PCR cycle number and annealing temperature were adjusted for different genes for optimum quantitation (Figs. 4.5-4.6).

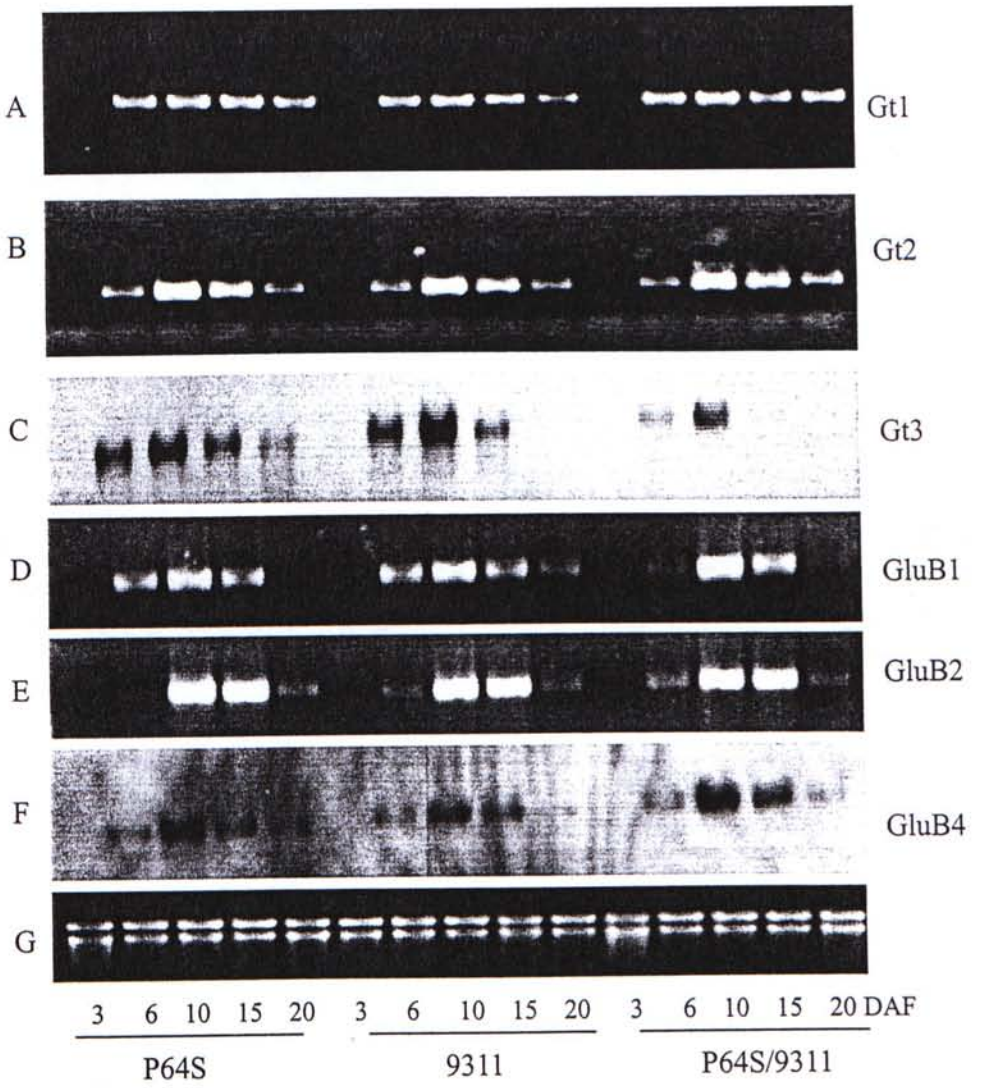


Figure 4.5 RT-PCR and northern blot analysis of the expression patterns of glutelin gene family

Profiling of the glutelin gene family by RT-PCR and northern blot analysis for two parental lines and F1 hybrid during seed development was carried out. A) RT-PCR products of Gt1 gene using specific primers and 20 PCR cycles; B) RT-PCR products of Gt2 gene using Gt2 specific primers and 25 PCR cycles; C) Northern blot analysis of GluA-3 gene using GluA-3 reverse specific probe; D) RT-PCR products of GluB-1 gene using GluB-1 specific primers and 30 PCR cycles; E) RT-PCR products of GluB-2 gene using GluB-2 specific primers and 30 PCR cycles; F) Northern blot analysis of GluB-4 gene using GluB-4 antisense specific probe. G) Agarose/formamide gel electrophoresis of total RNA from three rice lines and five developmental stages.

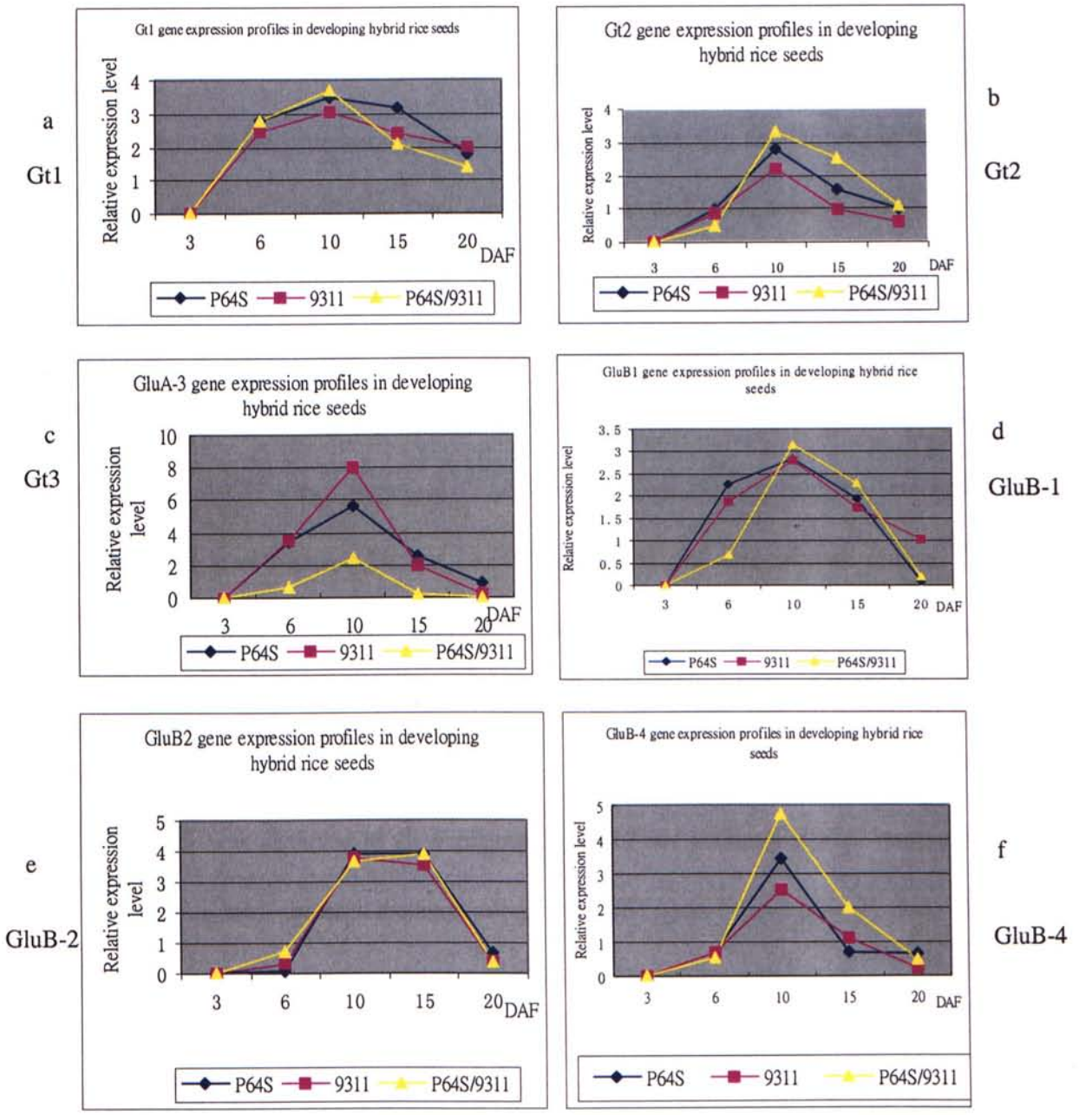


Figure 4.6 Expression profiles of glutelin family genes in developing hybrid rice seeds.

Figures were prepared according to the average densitometer values for individual samples based on Figure 4.5. a) Gt1 gene; b) Gt2 gene; c) GluA-3 gene; d) GluB-1 gene; e) GluB-2 gene; and f) GluB-4 gene.

4.4.1.1 The Gt1 (GluA-2) gene

The expression profiles of rice Gt1 gene were shown in Figure 4.5-A. The results indicated that the expression patterns of this gene were also similar among the two parents and their hybrid rice. Their expression began at 6 DAF, achieving the highest level at 10 DAF, then decreased through the remaining developing stages. The decrease, however, was not as sharp as in Gt2 (e.g. 15 DAF). These results were not coincided with those reported by Kim et al. (1993), which showed that the expression of Gt1 and Gt2 genes increased throughout the seed development and reached a maximum level between 20 and 25 DAF using Gt1 and Gt2 mixture probes. The difference between these two results might be due to the different rice strains involved and/or analysis method used, for example, the gene family member specific probing in this study while mixed cDNA hybridization probing in Kim's study. No marked difference was observed in the expression level of Gt1 gene among the two parents and their hybrid, although the expression level of F1 hybrid was relatively lower than in the parents at the later stages of seed development (Figure 4.6 a).

4.4.1.2 The Gt2 (GluA-1) gene

RT-PCR analysis of the expression of rice Gt2 gene in developing seeds of rice parental lines P64S and 9311 and their F1 hybrid showed that the expression patterns of rice Gt2 gene were similar among the two parents and their F1 hybrid. Expression could not be detected at 3 DAF, but with a relative low level at 6 DAF, reaching the highest level at 10 DAF and then gradually reduced towards 20 DAF. Although the

expression level of this gene appeared to be a little higher in the developing seeds of hybrid rice than that of the two parents, the difference were not significant (Figures 4.5-B and 4.6 b).

4.4.1.3 The Gt3 (GluA-3) gene

The overall expression profile of Gt3 gene was somewhat differed from that of Gt1 gene but similar to that of Gt2 gene (Figures 4.6 a, b and c). However, the expression level of F1 hybrid appeared to be significantly lower than its parents during the whole period of seed development.

4.4.1.4 Comparison of the expression profiles of GluA subfamily genes

The expression patterns in the developing seeds were generally similar among the three GluA subfamily members, Gt1, Gt2 and GluA-3 in the three rice strains. For example, the mRNA of all of the three genes was firstly detected at 6 DAF, reached their maximum levels at 10 DAF, and then declined during the later stage of seed development. One notable difference was that the expression level of Gt1 gene was relatively high during most of developmental stages and declined more gradually.

4.4.1.5 The GluB-1 gene

The GluB-1 gene showed an overall bell shape development pattern with its maximum level expression at 10 DAF. However, when comparing the three rice lines, the male parental line 9311 maintained a relatively higher level of GluB-1 mRNA at

20 DAF (Figures 4.5-D and 4.6 d).

4.4.1.6 The GluB-2 gene

RT-PCR analysis showed that the expression of GluB-2 gene also began at 6 DAF and reached maximum at 10 DAF. Differing from other glutelin gene members, it maintained a high level of expression through 15 DAF and then declined sharply at 20 DAF (Figures 4.5-E and 4.6 e). Similar expression profiles were seen in the two parental lines as well as their F1 hybrid.

4.4.1.7 The GluB-4 gene

The expression profile of GluB-4 gene in the developing seeds of the three rice strains was performed by northern blot analysis using equal amount of RNA sample, the results were shown in Figures 4.5-F and 4.6 f. The GluB-4 gene was expressed with a pattern generally similar to that of the GluA-3 gene. However, during seed development, the expression of GluB-4 gene was maintained at a higher level in the hybrid rice than that in its two parental lines (Figure 4.6 f).

4.4.1.8 Comparing of the expression profiling of GluB subfamily genes

The expression patterns of the three gene members in the GluB subfamily during seed development differed from each other. GluB-1 was expressed in a bell shape pattern with its peak level at 10 DAF; GluB-2 maintained its high expression from DAF 10 to 15; while GluB-4 expression peaked at 10 DAF with steeper slopes of

expression at both earlier and later development stages.

4.4.1.9 Profiling the expression of glutelin family genes in developing hybrid rice seeds

Based on the expression patterns of individual members in the glutelin gene family in female parental line P64S (Figure 4.7), 4 general expression patterns could be recognized: 1) A shaper peak pattern: this included Gt2, Gt3, and GluB-4, in which an expression maximum was peaked at 10 DAF with rather sharper declined in expression level at both early and later stages; 2) the bell shape pattern: this is represented by GluB-1 gene, in which the expression peaked at 10 DAF and gradually declined at both early and later development stages; 3) the plateau maximum pattern: this was shown by GluB-2 gene with its expression level maintaining at maximum between 10 and 15 DAF; and 4) the half bell shape pattern: Gt1 showed such an expression pattern between 6 DAF and 20 DAF with its mRNA level remained high while declining at 20 DAF.

4.4.1.10 Profiling glutelin gene expression in developing rice seeds at protein level

Results from tricine SDS-PAGE gel for glutelin (Figure 4.8) showed that little mature glutelin could be detected at 3 DAF except the 57-kDa glutelin precusor; the 22-23-kDa glutelin small subunit and 37-39-kDa large subunit started to accumulate at 6 DAF; and in the subsequent developing stages, these subunit polypeptides became dominant. The expression patterns of glutelin gene members at protein level

showed striking difference from those at mRNA level during the later developmental stages, i.e. from 10 to 20 DAF. While at mRNA level, these were a general decline during seed maturation, the storage protein glutelin continued to accumulate.

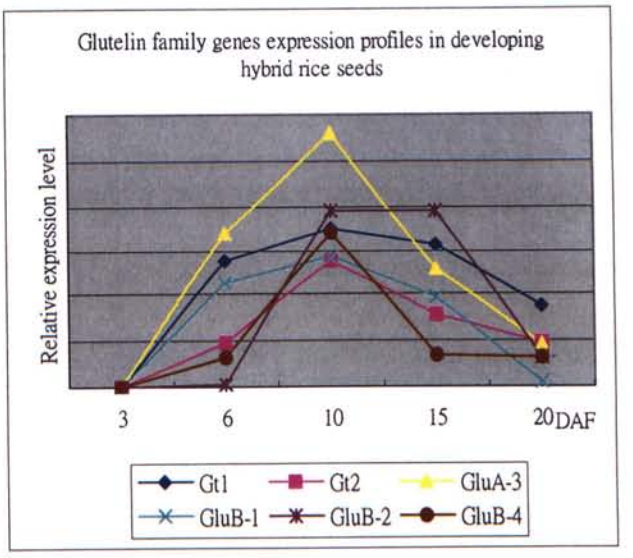


Figure 4.7 Summary of the glutelin family gene expression profiles in developing hybrid seeds.

The figure was prepared based on comparison of the RT-PCR products using the equal amount total RNA template from P64S and gene specific primers.

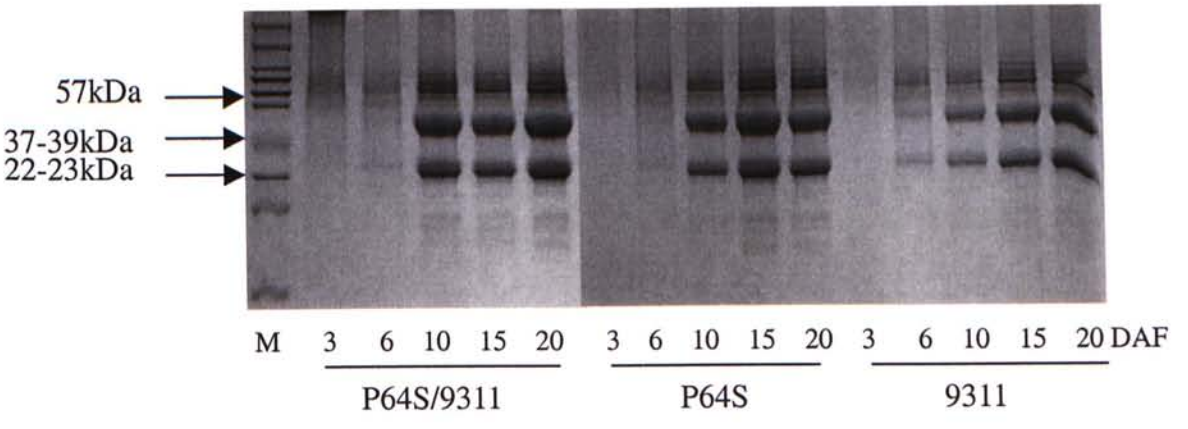


Figure 4.8 Tricine SDS-PAGE of glutelin protein from developing hybrid rice seeds.

4.4.2 Profiling the expression of genes encoding the prolamin family proteins in developing hybrid rice seeds

4.4.2.1 The 10- kDa prolamin gene

Because the 10-kDa prolamin gene shows low DNA sequence homology with other members of the prolamin family, the expression profiling of 10-kDa prolamin gene was studied by northern blot analysis as described in Section 3.2.8.

Northern blot results (Figure 4.9 A) showed that the expression of 10-kDa prolamin gene started at 6 DAF in the two parental lines and their F1 hybrid. The expression pattern differed somewhat between different lines. In P64S the highest expression peak appeared at 6 DAF, earlier than 9311 and F1 hybrid. While in 9311 and F1 hybrid, the expression peaked at 10 DAF. At later developmental stages, the level of 10-kDa prolamin mRNA in 9311 gradually declined, while in both P64S and the hybrid its mRNA remained at higher levels.

These results also revealed that the expression level of 9311 during the whole seeds developing stage is at lower level comparison to P64S and F1 hybrid (Figure 4.10 a).

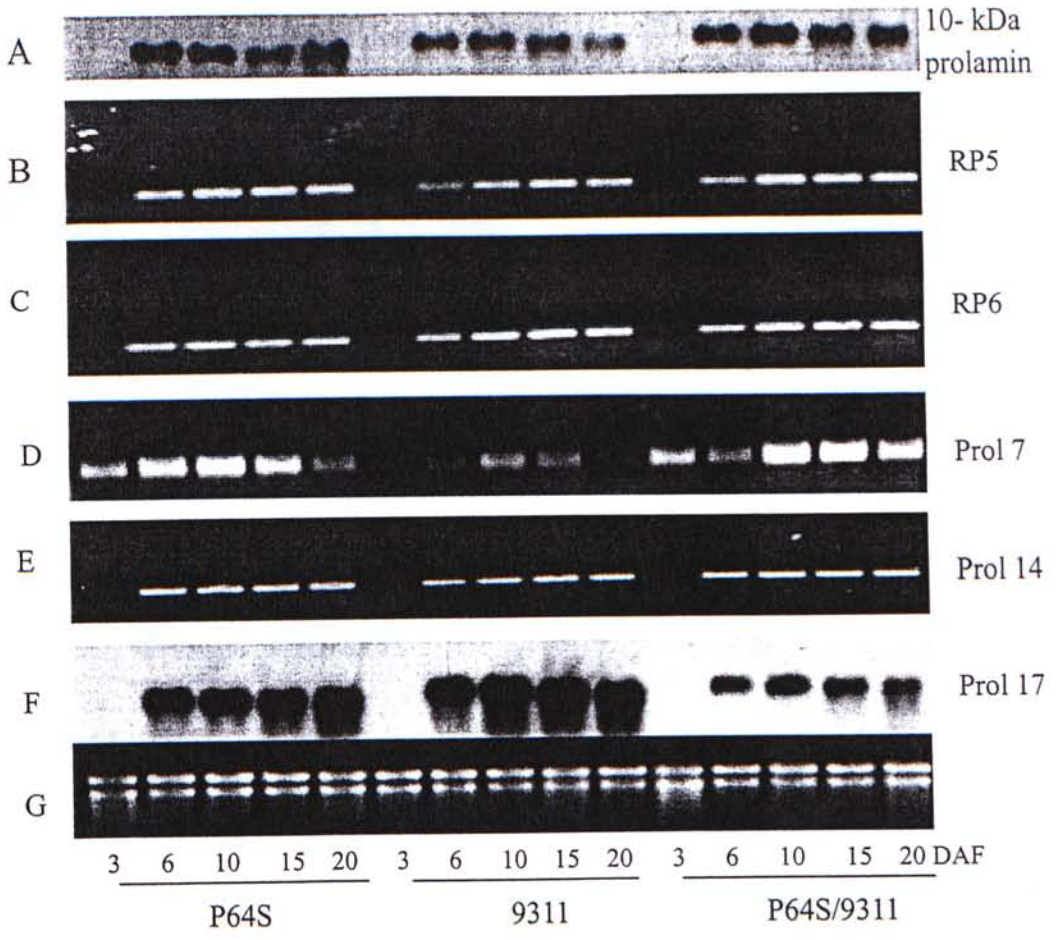


Figure 4.9 RT-PCR and northern blot analysis of prolamin family genes in developing hybrid rice seeds.

Prolamin genes expression was profiled by RT-PCR and Northern blot analysis for two parental lines and their F1 hybrid during seed development. A): Northern blot analysis of the 10-kDa prolamin gene using specific antisense probe; B): RT-PCR products of RP5 gene using specific primers and 20 PCR cycles; C) RT-PCR products of RP6 gene using specific primers and 25 PCR cycles; D): RT-PCR products of Prol 7 gene using specific primers and 30 PCR cycles; E): RT-PCR products of Prol 14 gene using specific primers and 25 PCR cycles; F): Northern blot analysis of Prol 17 gene using antisense specific probe; and G): Agarose/formamide gel electrophoresis of total RNA from the three lines and five developing stages.

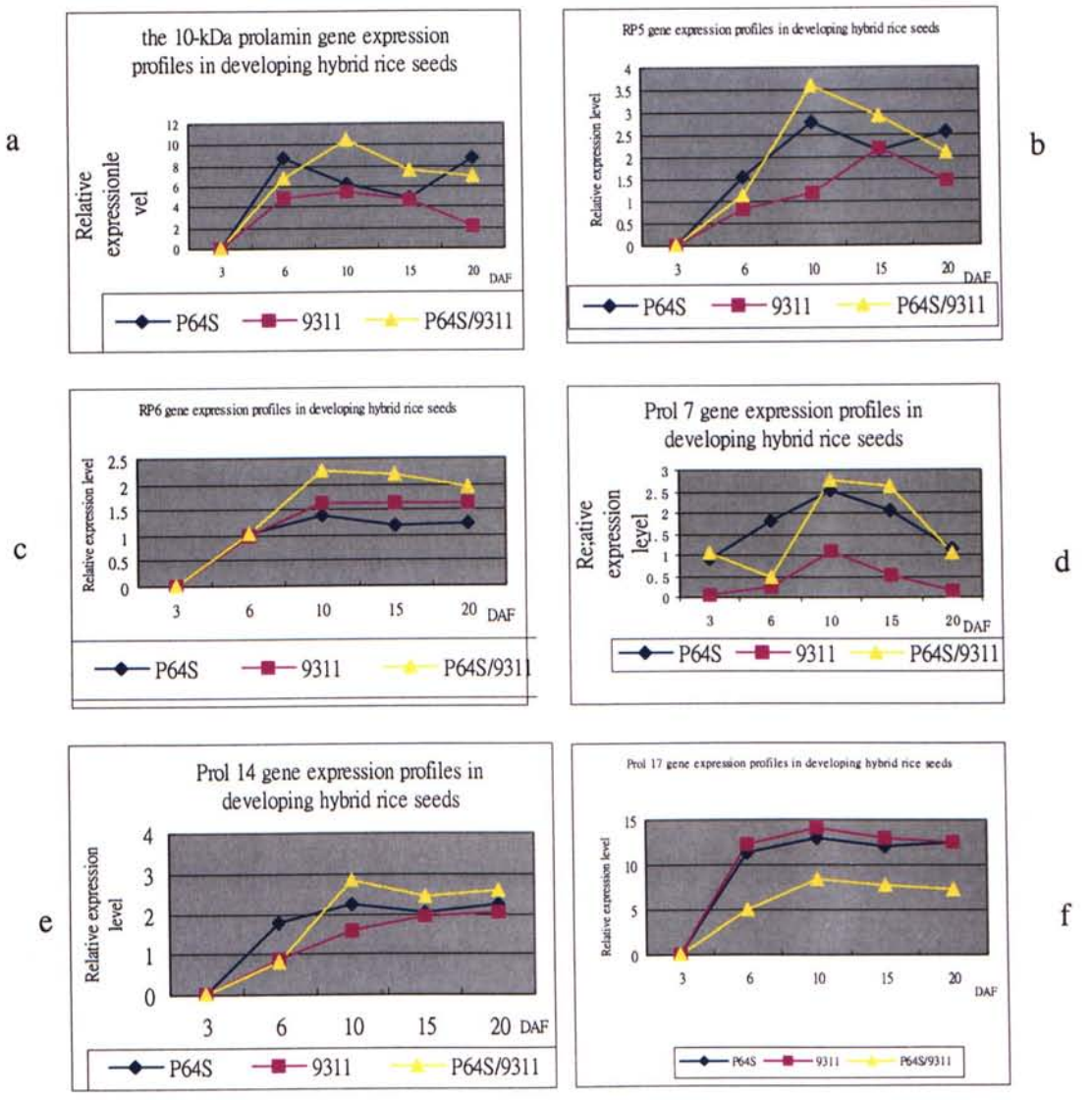


Figure 4.10 Prolamin gene expression porofiles in developing hybrid rice seeds.

Graphs were prepared based on the average densitometer values of samples in Figure 4.9. a) 10-kDa prolamin gene, by RT-PCR products; b) RP5 gene, by RT-PCR products; c) RP6 gene, by RT-PCR products; d) Prol 7 gene, by RT-PCR products; e) Prol 14 gene, by RT-PCR products; and f) Prol 17 gene, by northern blot analysis.

4.4.2.2 The RP5 gene

The expression of the RP5 gene was detected at mRNA level at developmental stages from 6 to 20 DAF, with different levels. There was notable difference among the expression patterns of the two parental lines and their F1 hybrid. In P64S and F1 hybrid, the peak expression level appeared at the stage of 10 DAF, while 15 DAF in 9311. The expression pattern of RP5 gene in F1 hybrid was similar to its female parent P64S, while the level of its expression in F1 hybrid was higher than the two parental lines during the later developmental stages from 10 to 20 DAF (Figures 4.9-B and 4.10 b).

4.4.2.3 The RP6 gene

Figure 4.9-C showed that the expression of the RP6 gene remained high at the later development stages (from 10 DAF), even at mature stage (20 DAF). Comparison with the expression pattern of RP5 gene, RP6 gene was expressed at a lower level in P64S, while higher level in 9311; this is contrary to the situation in RP5 gene, indicating that the expression level of storage protein genes may be variety dependant. Expression level of RP6 gene in F1hybrid is higher than the two parental lines during developing rice seeds. Further analysis showed that RP5 gene was expressed at higher level than the RP6, because RT-PCR product for RP5 was carried out only for 20 cycles, while 25 cycles for RP6 gene (Figure 4.10 c)

4.4.2.4 The Prol 7 gene

The striking characteristic of the Prol 7 gene expression pattern is that this gene was expressed at very early stage (3 DAF) with relative high expression level, especially in F1 hybrid and P64S (Figure 4.9-D). The highest expression level appeared at 10 DAF and then gradually declined until mature (20 DAF). The expression pattern of F1 hybrid obviously differed from those of its parental lines for its relative high level of expression at 3 DAF and 15 DAF. Data also revealed that the expression level of 9311was the lowest among the three rice strains.

4.4.2.5 The Prol 14 gene

The Prol 14 gene appeared to express beginning at 6 DAF and reach and maintained high level at later stages (from 10 DAF to 20 DAF) without a clear expression peak (Figure 4.9-E). This result is similar to previous studies (Kim et al., 1993). Similar expression patterns were detected in the two parental lines and F1 hybrid, while the expression level of 9311 was relatively lower than that of P64S, and F1 hybrid showed the highest expression level (Figure 4.10 e).

4.4.2.6 The Prol 17 gene

The transcript for Prol 17 was detected in total RNA fraction at 6 DAF at very high level, which was maintained throughout the subsequent developmental stages. The patterns of the two parental lines and F1 hybrid were similar (Figure 4.9-F and 4.10 f). But the expression level of F1 is much lower than that of two parents.

4.4.2.7 Expression profiles of prolamin family genes

A general expression profile of the prolamin family genes is drawn by using P64S' expression profiles (Figure 4.11). The following results were observed:

1. Most family genes were expressed beginning from 6 DAF to mature stage (20 DAF).
2. The 10-kDa prolamin gene showed a distinct expression pattern. Its mRNA transcript reached high level at 6 DAF, then declined somewhat till 15 DAF, and increased to a higher level at 20 DAF.
3. The Prol 7 gene is also special because it was expressed at very early stage (at 3 DAF). After reaching peak level at 10 DAF, its expression declined at subsequent stages.
4. The RP6 and Prol 14 genes showed similar expression patterns during seed development. The mRNA transcript of them appeared at 6 DAF, reached the highest level at 10 DAF, and then slightly declined at 15 DAF, while remained relative high level till 20 DAF.
5. The Prol 17 gene maintained a high expression level from 6 DAF till 20 DAF.
6. For the tow parental lines and F1 hybrid, the expression patterns and levels of prolamin family genes showed some variations. The expression level of F1 hybrid was higher than these of two parental lines for most of the genes except Prol 17 gene; while the expression pattern of F1 hybrid is more similar to P64S (female parental line) except Prol 7 gene, where its expression pattern of F1 hybrid is different from that of two parental lines.

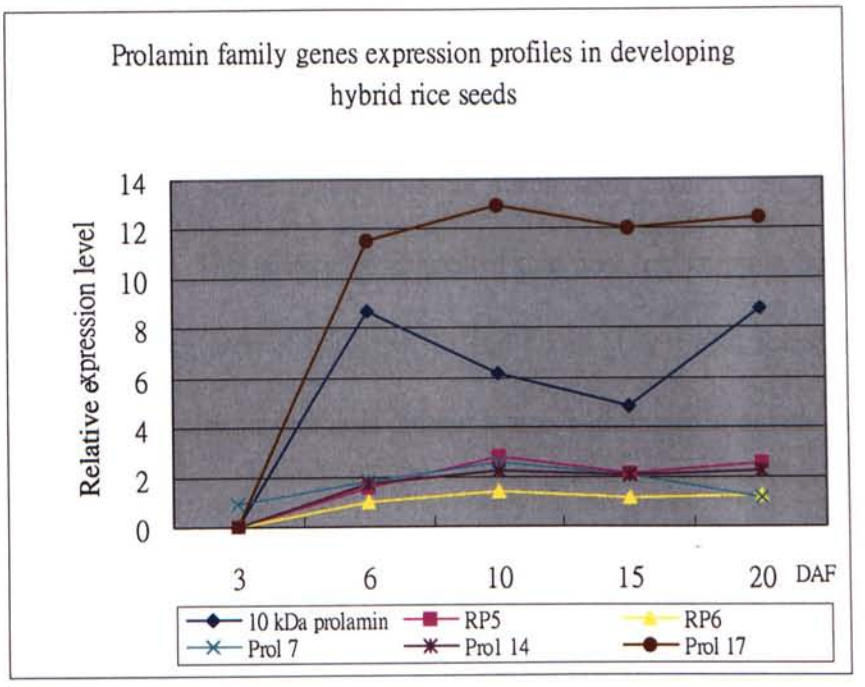


Figure 4.11 Summary of the prolamin family genes expression profiles in developing hybrid rice seeds.

4.4.2.8 Expression profiles of prolamin genes in developing hybrid rice seeds at protein level

By the established extraction procedure used in this study, the 13-kDa and 16-kDa prolamins could be detected by tricine SDS-PAGE, but not the 10-kDa prolamin (Figure 4.12). The expression pattern of 13-kDa prolamin was prominent and increased with time, starting at 6 DAF until 20 DAF. The expression level of 9311, however, was the lowest among the three lines, while that of F1 hybrid is the similar to or even higher than the P64S'. The 16-kDa prolamin appeared as a more heterogeneous protein fraction, with a similar accumulation pattern as that of the 13-kDa prolamin.

Comparison of expression patterns of prolamin genes at RNA level and protein level revealed that prolamin was starting to accumulate at 6 DAF, which was coincided with the mRNA transcript level at the same stage except for Prol 7 gene. mRNA transcript of Prol 7 gene appeared earlier during seed development, at 3 DAF, with relative high level. This difference suggested that post-transcription might play an important role during protein synthesis. Li and Okita (1993) also found that the expression pattern of prolamin synthesis did not match with the accumulating mRNA transcripts.

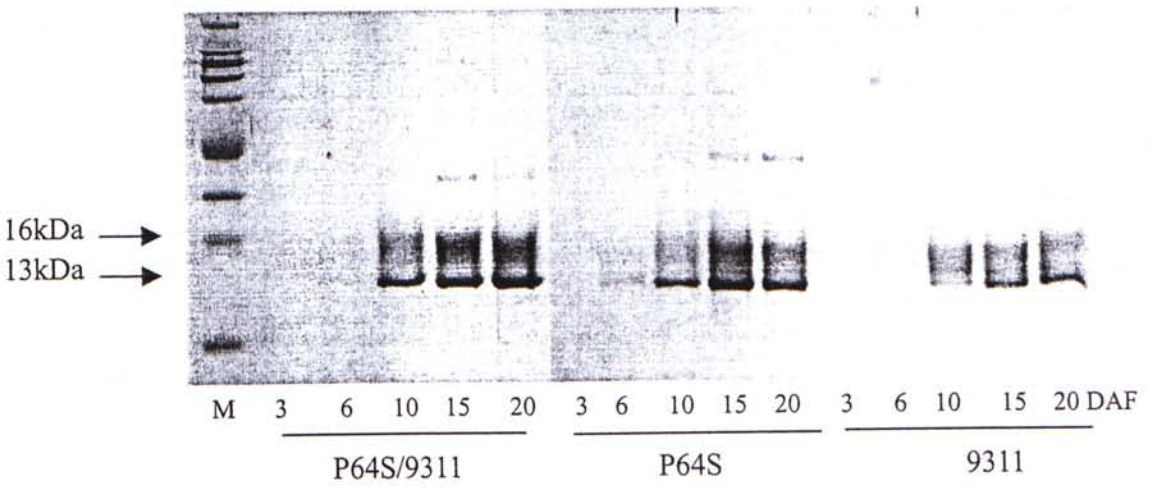


Figure 4.12 Tricine SDS-PAGE gel of prolamin in developing hybrid rice seeds.

4.4.3 Profiling the expression of genes encoding globulin family in developing hybrid rice seeds

4.4.3.1 The 26-kDa globulin (alpha-globulin) gene

The expression profile of the 26-kDa globulin gene was performed by northern blot analysis using specific DNA probe as the DNA sequence homology among the family members is low. Figure 4.13-A showed that the expression of the 26-kDa globulin gene could not detected at 3 DAF but at 6 DAF with relative high level. The expression remained at high expression level throughout the later stages without an obvious expression peak activity. Similar patterns were seen in the parental lines and the F1 hybrid. Moreover, the expression level of F1 hybrid is similar to its two parental lines (Figure 4.14 a).

4.4.3.2 The Globulin 1 gene

Figure 4.13-B showed that the globulin 1 gene was expressed at 10 DAF with its maximum level, and then decreased gradually at 15 DAF and by 20 DAF, the mRNA level was quite low. Similar expression patterns were observed in the parental lines and the F1 hybrid. This result revealed that globulin 1 gene was expressed later in the seed development than the other genes. Moreover, its mRNA transcripts were accumulated in a short period, mainly around 10 DAF. The expression level is similar between parental lines and F1 hybrid except relatively higher level at 15 DAF in P64S.

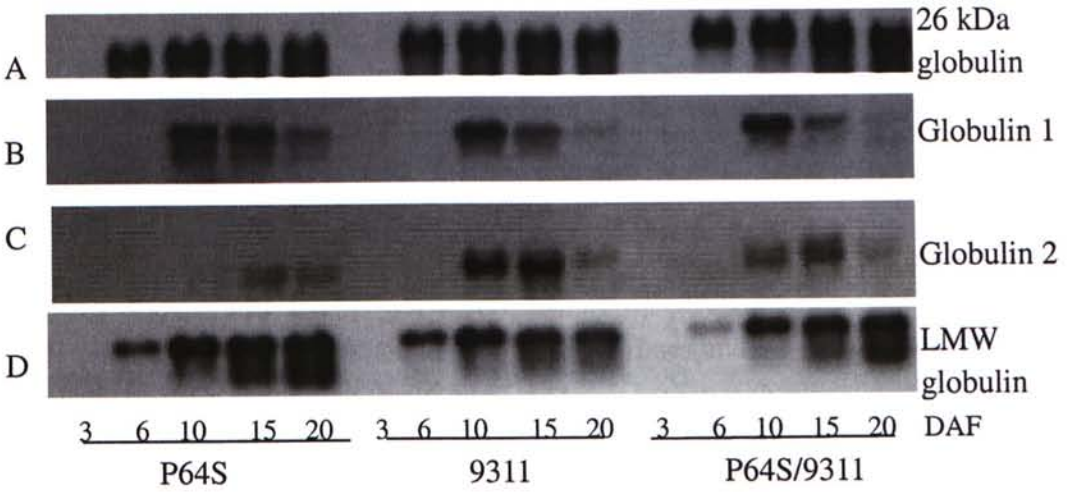


Figure 4.13 Northern blot analysis of the globulin family genes in developing hybrid rice seeds.

A): The 26-kDa globulin gene; B): The globulin 1 gene; C): The globulin 2 gene; and D): The LMW globulin gene.

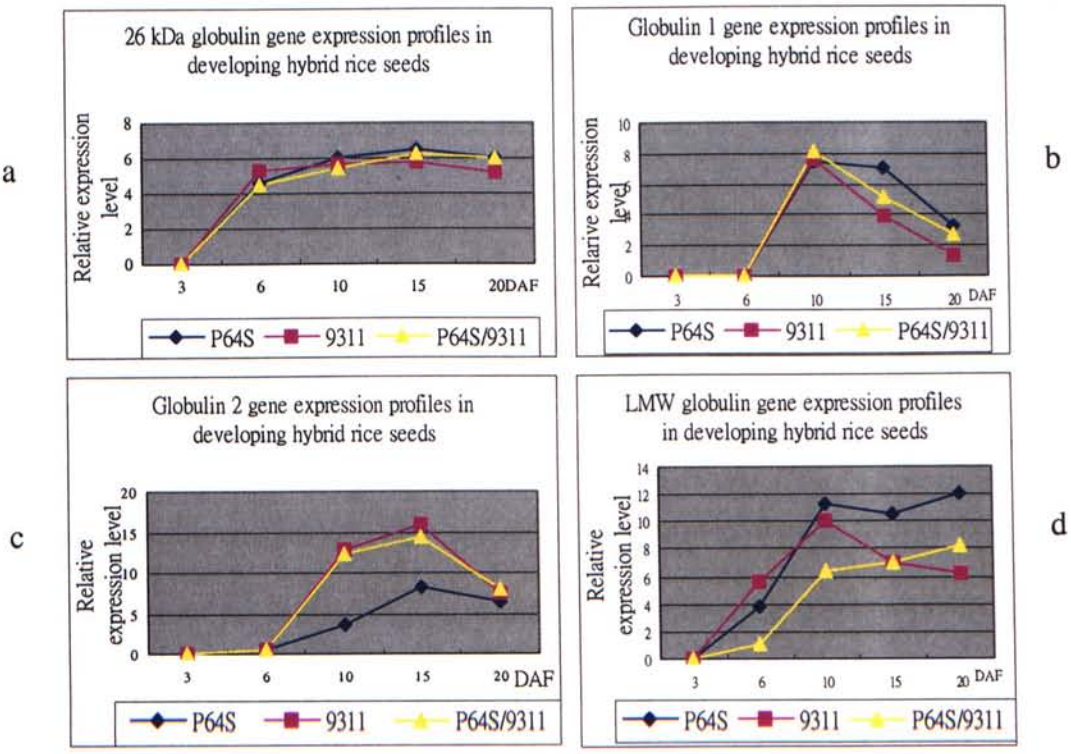


Figure 4.14 Expression profiles of the globulin family genes

a) The 26-kDa globulin gene; b) The globulin 1 gene; c) The globulin 2 gene; and d) The LMW globulin.

4.4.3.3 The Globulin 2 gene

Similar to the globulin 1 gene, the globulin 2 gene was expressed strongly around 10 DAF. However, the globulin 2 gene reached its expression peak at 15 DAF, not 10 DAF as in globulin 1 and decreased towards 20 DAF (Figure 4.13 C). The expression patterns showed little difference between parental lines and F1 hybrid, whereas, the expression level varied, P64S with the lowest, F1 hybrid in between, and 9311 with significantly higher level at 10 and 15 DAF (See Figure 4.14 c).

4.4.3.4 The Low molecular weight (LMW) globulin gene

Figure 4.13-D showed that the LMW globulin gene was expressed around at 6 DAF and the level increased with time until 20 DAF. The expression peaked at 10 DAF in 9311, but for P64S and the F1 hybrid, it appeared that the levels remained high towards 20 DAF. This revealed that the expression pattern of F1 hybrid was close to that of the female parental line (Figure 4.14 d).

4.4.3.5 Profiling the expression of the globulin family genes

The expression profiles of the globulin family genes were obtained by northern blot analysis using 5 μg total RNA and specific antisense probes for each of the genes. Comparison of the expression patterns of all the genes in the globulin family in the P64S rice line was showed in Figure 4.15. Results indicated:

1. All the genes were not expressed at the stage of 3 DAF.
2. Each family member showed a distinct expression pattern although they

all belong to the same gene family.

3. The 26-kDa globulin and the LMW globulin genes exhibited higher expression level at 6 DAF than the other members. Moreover, these two genes maintained relatively high expression levels at late developmental stages, from 10 DAF till 20 DAF. These results indicated that the LMW globulin, like the 26-kDa globulin, represents another major fraction in the globulin gene family.

4. Globulin 1 gene was expressed at later development stages among all the globulin genes, starting at 10 DAF, at which is also its highest expression point. At 20 DAF, globulin 1 gene was expressed very weakly. Thus globulin 1 gene was mainly expressed at 10 DAF, extending to 15 DAF, within a short developmental period in P64S.

5. Globulin 2 gene was expressed earlier than globulin 1 as weak mRNA could be detected at 6 DAF. Its expression peaked at 15 DAF, later than globulin 1. In brief, globulin 2 was expressed at low level during seed development except in 9311 that showed relatively stronger expression at 10 to 15 DAF.

6. The globulin expression pattern in P64S during seed development was general by similar to those of 9311 and F1 hybrid.

7. The expression levels of F1 hybrid were between those of two parental lines except for the lowest expression level in LMW globulin gene at early stages (6 DAF to 10 DAF).

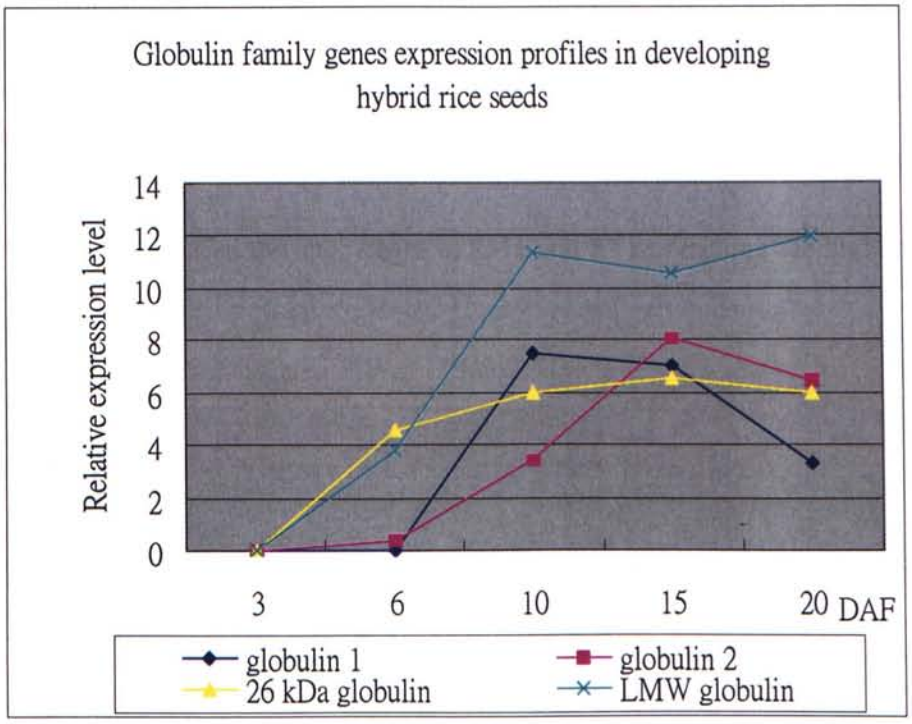


Figure 4.15 Summary of expression profiles of the globulin family genes in developing hybrid rice seeds.

4.4.3.6 Expression profiles of globulin proteins in developing hybrid rice seeds at protein level

Seed powder (0.0100 g) from different rice strains at different developing stages was extracted for globulin with 100 μl of extraction buffer as described in section 3.2.10. The protein samples were analyzed by tricine SDS-PAGE gel using 10 μl globulin extract. Tricine SDS-PAGE gel showed that the 26-kDa globulin was the major fraction (Figure 4.16) and its accumulation began at 10 DAF, continued towards 20 DAF. This expression pattern is similar to that of prolamin. The expression levels in the three lines are generally comparable.

Further analysis showed that at mRNA transcriptional level, globulin expression was not matching with the protein accumulation pattern. For example, the 26-kDa globulin was expressed at high level of mRNA at 6 DAF and remained so throughout the later stages, while almost no protein was detected at 6 DAF and linear increase in protein level appeared at the later stages in 9311 and F1 or reaching the highest level at 15 DAF in P64S.

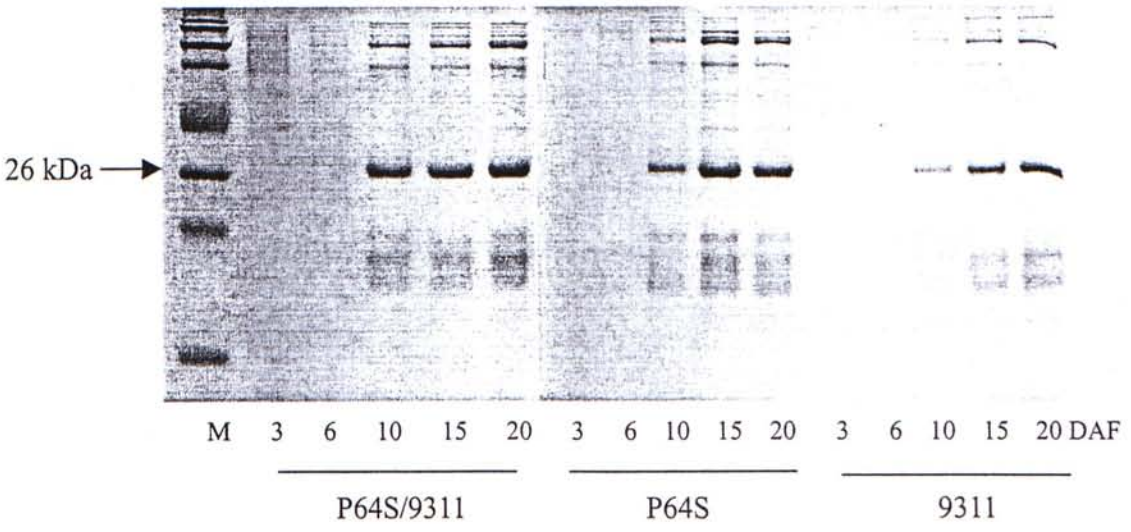


Figure 4.16 Tricine SDS-PAGE gel analysis of globulin proteins.

4.4.4 Profiling the expression of genes encoding rice albumin family in developing hybrid rice seeds

4.4.4.1 The RA5 gene

Figure 4.17-A showed that the RA 5 gene was expressed around 6 DAF at low level and the expression reached high level at later stages from 10 DAF till 20 DAF. There was not an obvious peak of expression during seed development. The

expression patterns between the two parental lines and F1 hybrid were similar. The expression level of F1 hybrid was higher than those of the two parental lines while the expression level of and that of P64S was relative lower than that of 9311 (see Figure 4.18 a).

4.4.4.2 The RA 14 gene

The RA 14 gene was expressed weakly from 10 DAF to 20 DAF. Relative higher levels appeared at 15 DAF in P64S and F1 hybrid, while at 10 DAF in 9311. The expression pattern of F1 was similar to the female parental line (P64S). The expression level of F1 appeared to be higher than that of the parental lines at 15 DAF (Figure 4.17-B and 4.18 b).

4.4.4.3 The RA 17 gene

The expression pattern of RA17 gene was found similar to that of the RA5 gene. It was expressed around 6 DAF and with increasing levels through 10 DAF to 20 DAF. The expression patterns of F1 and parental lines are almost the same. Their expression levels are also similar (Figure 4.17-C and 4.18 c).

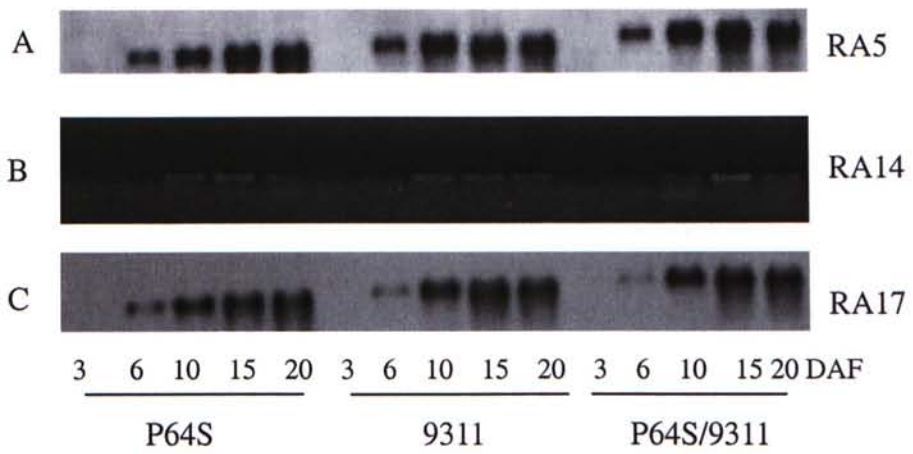
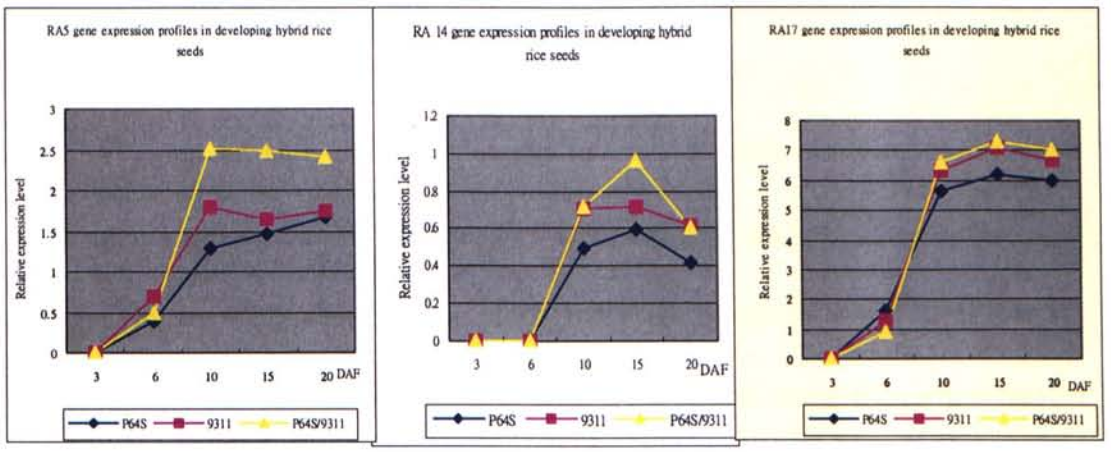


Figure 4.17 Northern blot and RT-PCR analysis of albumin family genes.

A) Northern blot analysis of the RA5 gene with antisense probe; B) RT-PCR products of the RA14 gene (30 cycles); and C) Northern blot analysis of the RA17 gene.



a b c

Figure 4.18 Expression profiles of albumin family genes

Graphs were prepared based on data from Figure 4.17. a) The RA5 gene; b) The RA14 gene; and c) The RA17 gene.

4.4.4.4 Profiling the expression of the albumin family genes in developing hybrid rice seeds

Comparison of the expression profiles of individual genes involved in albumin family (Figure 4.19) revealed:

1. The RA 5 and RA 17 genes showed similar expression patterns, both starting at 6 DAF with relative low mRNA level and maintaining a higher expression level through 10 DAF to mature, although they share relative low DNA sequence homology (77%).
2. The RA 14 gene was expressed around at 10 DAF and remained relatively weak level of expression during the seed development.

4.4.4.5 Profiling albumin gene expression in developing hybrid rice seeds at protein level

A group of high molecular weight (50 to 150-kDa) water-soluble proteins (albumins) appeared at 3 DAF through 6 DAF and then gradually decreased at later stages (Figure 4.20), while a group of low molecular weight (about 15-kDa) albumin polypeptides started to accumulate at 10 DAF towards seed maturation, which may be a group of allergic albumin. Accumulation of a major albumin polypeptide (26-kDa) was observed at 10 DAF and its expression remained obvious through seed maturation. The accumulation level of major albumins appeared to be relative lower than those of the other storage proteins.

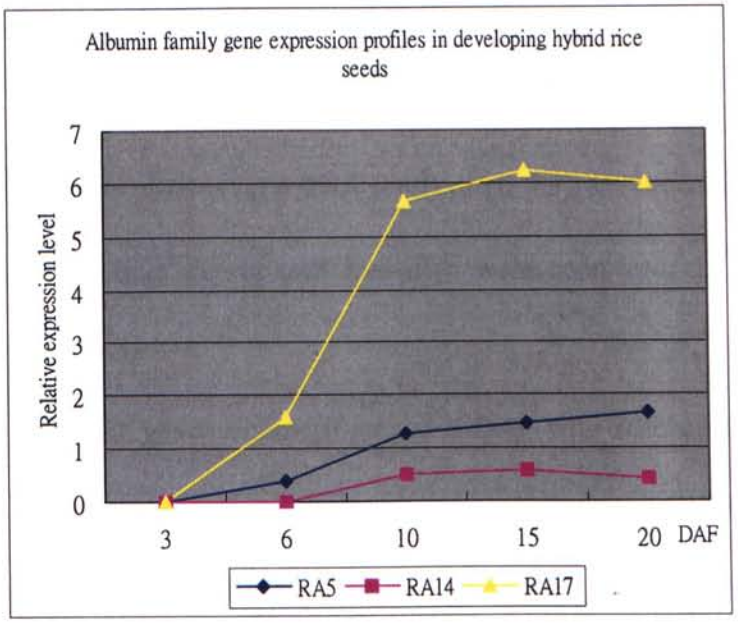


Figure 4.19 Summary of expression profiles of the albumin family genes in developing hybrid rice seeds

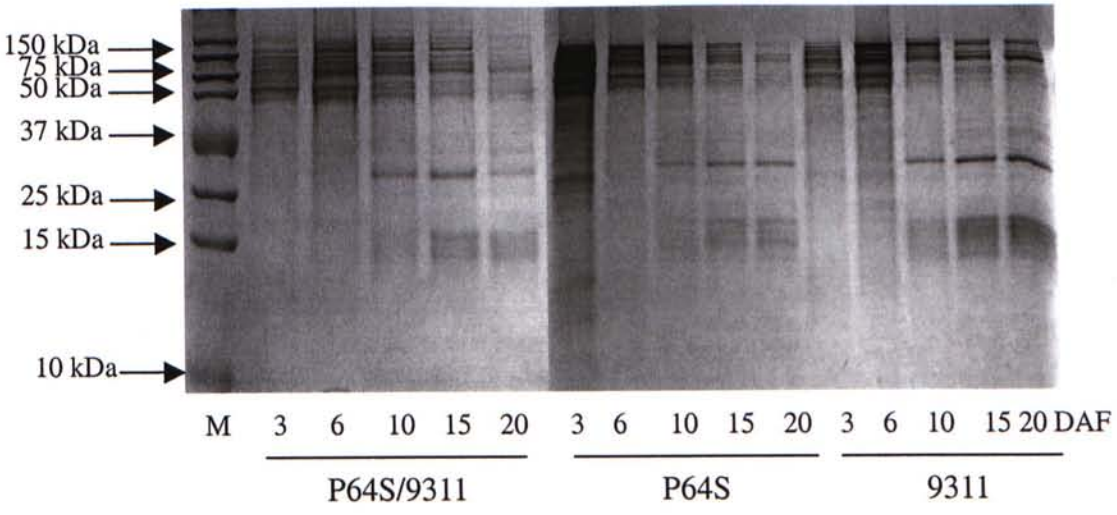


Figure 4.20 Tricine SDS-PAGE gel for albumin from developing hybrid rice seeds

4.4.5 Comparison of the expression patterns of all genes encoding seed storage proteins in developing hybrid rice seeds

When the expression patterns (at mRNA level) of all the genes encoding storage proteins in hybrid rice lines during seed formation were compared, the following observations could be made:

1. Almost all of the genes expressed around 6 DAF with an exception Prol 7 gene (13-kDa prolamin) began to express at an earlier stage (3 DAF) though its expression level was not high. On the other hand, globulin 1 and RA14 albumin were expressed at later stage (10 DAF) during seed development.
2. Five major expression patterns could be generally recognized (Figure 4.21 b):
    a) The genes were expressed early on and remained active during course of development, e.g. Prol 17, Prol 14, LMW globulin, 26-kDa globulin, RA5, RA17;
    b) The expression pattern is similar to a) except after reaching the peak expression, the activity slowly declined during seed maturation (bell shape pattern), e.g. Gt1,GluB-1 and Prol 7 (in P64S);
    c) The expression patter is generally similar to b) but the rise and fall of the expression was steep, with a sharp and clear peak activity, e.g. Gt2, GluA-3 and GluB-4;
    d) This pattern is similar to c) except that the peak expression remained for several days (a flat peak) before declining, e.g. GluB-2 and globulin 1 (in P64S);

e) Variations from the other general patterns were also observed. For example, the 10-kDa prolamin showed a flat-top expression profile during seed development.

3. In connection to expression patterns, different expression peaking times during seed formation were also observed. For example, 6 DAF and 20 DAF for the 10-kDa prolamin; 10 DAF for Gt1, Gt2, Gt3, GluB-1, GluB-2, GluB-4, and Gb1; and 15 DAF for Gb2, 26-kDa Gb and GluB-2.

4. The relative expression levels of the storage protein genes under this study varied greatly, from high mRNA level (e.g. Prol 17, 26-kDa globulin), to low level (e.g. RA14), to various levels in between (e.g. 10-kDa prolamin, GluB-2).

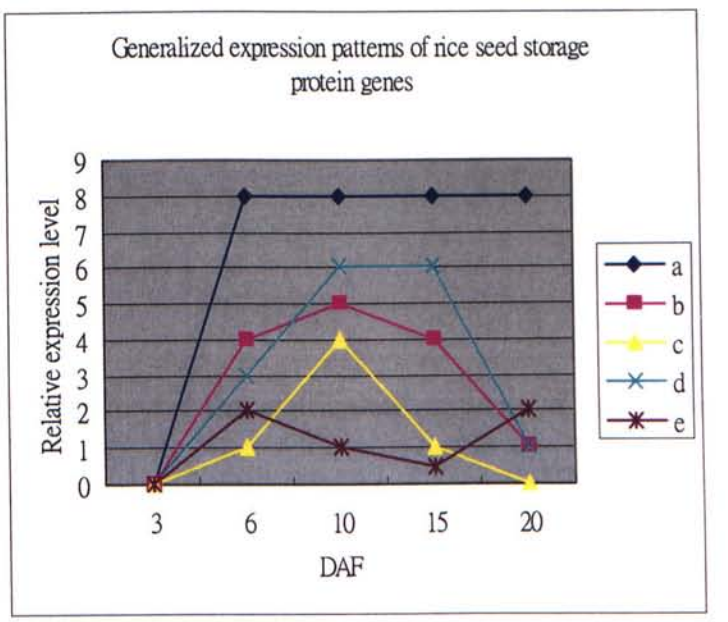


Figure 4.21 b) Generalized expression patterns of rice seed storage protein genes

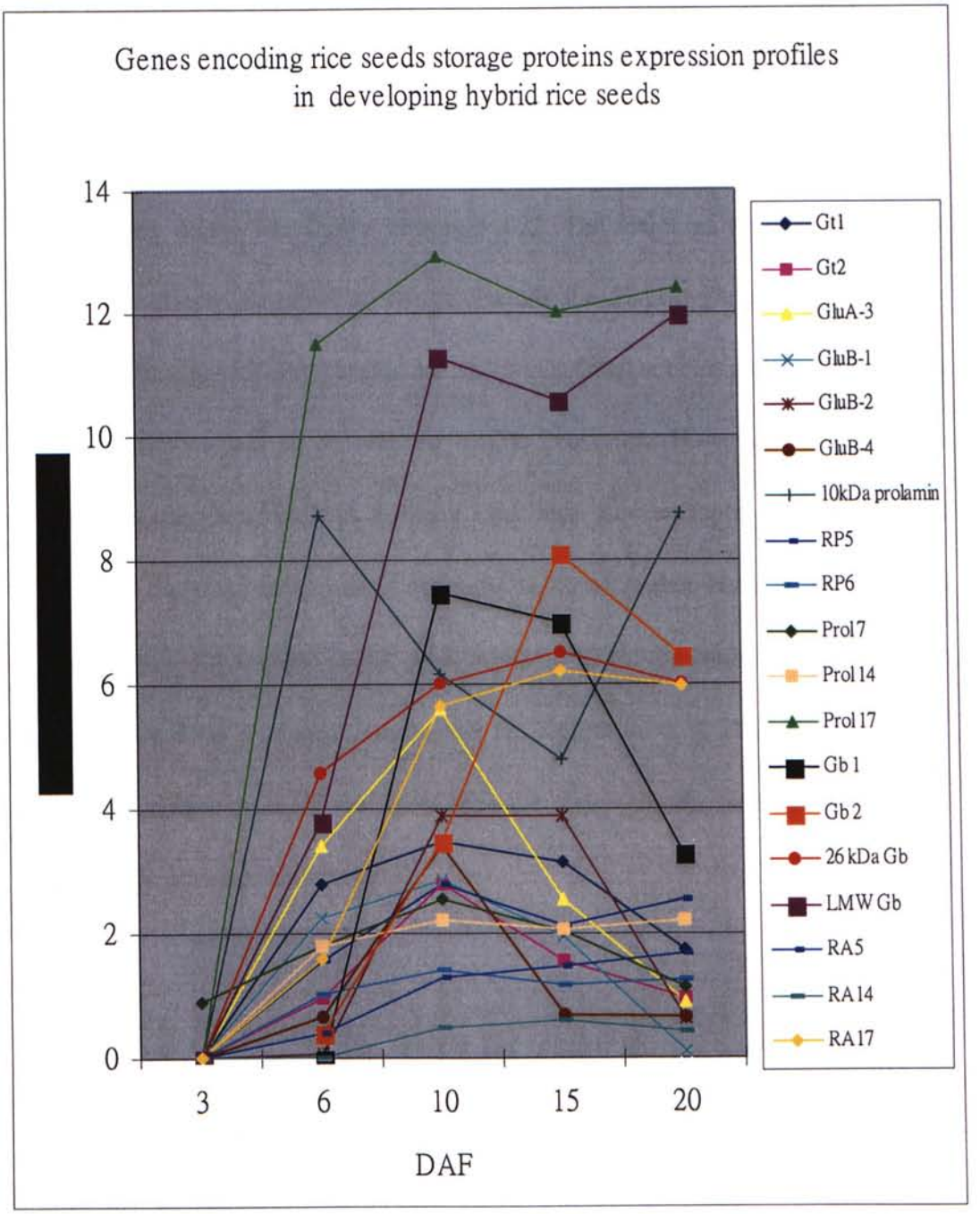


Figure 4.21 a) A summary of the expression patterns of genes encoding storage proteins in developing hybrid rice seeds

4.4.6 Profiling the total proteins in developing hybrid rice seeds

Tricine SDS-PAGE gel of the total protein from three rice lines at different developing stages was shown in Figure 4.22. The major polypeptide bands that could be clearly identified on the gel include the 57-kDa glutelin precursor, the 37-39-kDa large subunit glutelin, the 22-23-kDa small subunit glutelin, the 13-kDa prolamin (13 a and 13 b), and the 10-kDa prolamins. However, the expression pattern varied somewhat in different rice lines. For example, in P64S (female parental line), the total protein appeared to be of higher level at 15 DAF and declined to some extent at 20 DAF, while in 9311, the male parental line, the expression level increased and reached the highest level at 20 DAF. The overall protein profiles revealed that glutelin is the major storage protein and prolamin follows in abundance in developing rice.

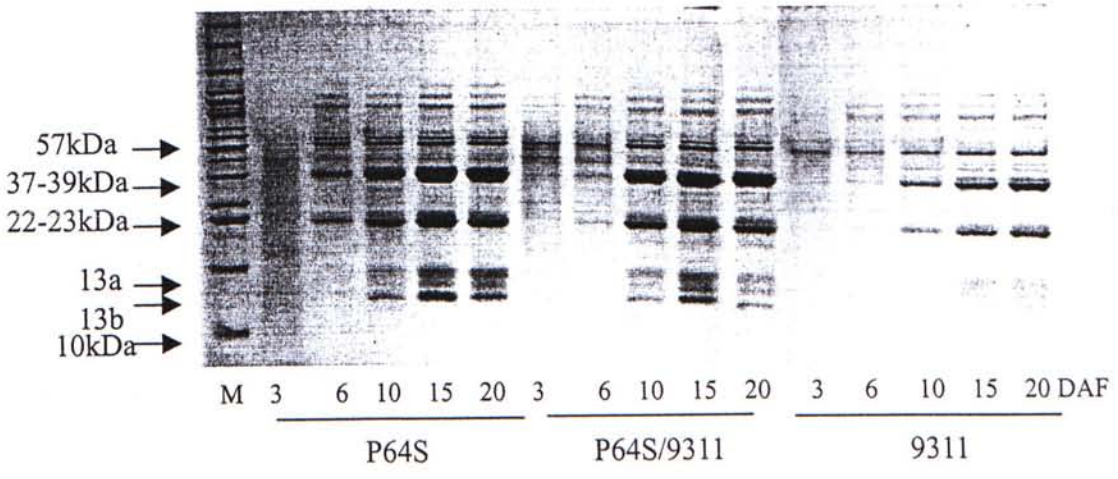


Figure 4.22 Tricine SDS-PAGE of total protein in developing hybrid rice seeds.

The protein content of the rice lines under this study was determined using 20 DAF seeds at Hunan Hybrid Rice Research Center (HHRRC, Changsha, Hunan

Province, China). The results showed that the protein content of F1 hybrid is 12.9% by seed weight, which is 1.8% lower than that of P64S (14.7%) but 2.1% higher than that of 9311 (10.8%). These results are generally in agreement with the previous studies which showed that the protein content of most rice hybrids was between the two parents (near to the average value of parents) and exhibited positive incomplete dominance (Liu et al., 1990).

4.5 Profiling the expression of genes encoding rice starch synthases in developing hybrid rice seeds

4.5.1 Rice ADP-glucose pyrophosphorylase (AGPase) genes

4.5.1.1 The AGPase large subunit gene

Figure 4.23-A showed that the mRNA of AGPase large subunit gene could be detected at 3 DAF and its level reached highest at 6 DAF. The level then declined during the subsequent developing stages but remained relative high. The expression pattern of F1 hybrid was similar to those of the two parental lines with an expression level between the two parents (Figure 4.24 a).

4.5.1.2 The AGPase small subunit gene

The expression pattern of AGPase small subunit gene was found similar to that of AGPase large subunit except that its expression level appear to decline more rapidly at the later development stages (15 to 20 DAF). The expression patterns of P64S, 9311 and F1 hybrid were almost the same, while the expression level of F1 was lower

and higher than those of two parental lines at early and later stage, respectively (Figure 4.24-b).

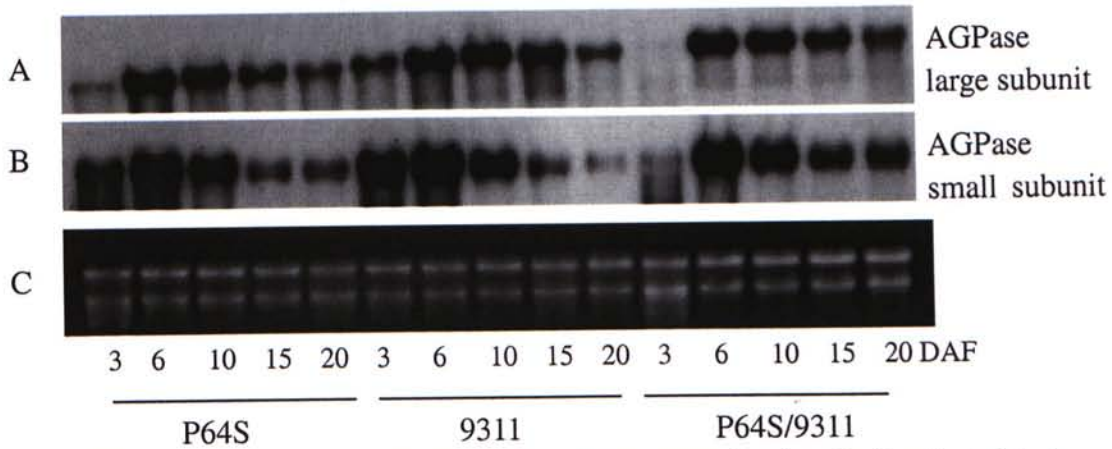
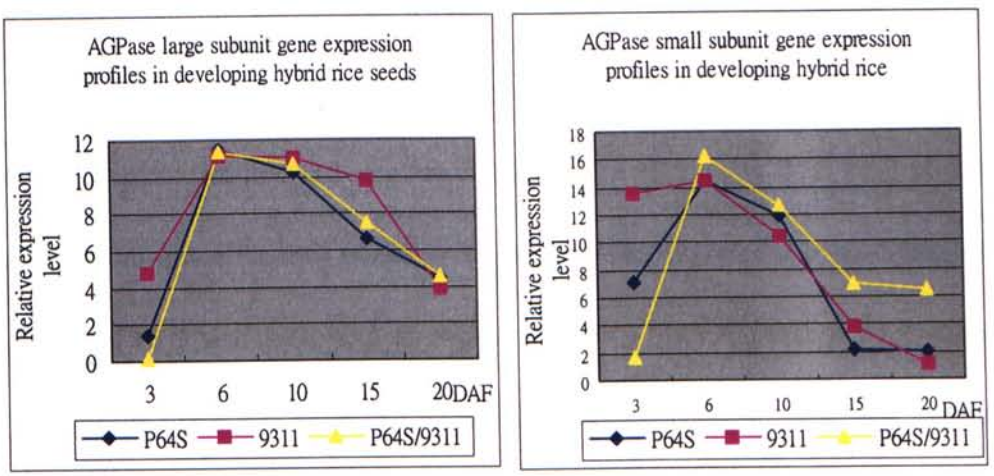


Figure 4.23 Northern blot analysis of the AGPase gene in developing hybrid rice seeds.

A) The AGPase large subunit gene; and B) The AGPase small subunit gene.



a b

Figure 4.24 Expression profiles of the AGPase genes in developing hybrid rice seeds.

a) The AGPase large subunit gene; and b) The AGPase small subunit gene.

4.5.2 The *Wx* (Granule bound starch synthase, GBSS) gene

Granule bound starch synthase (GBSS) is encoded by the *Waxy (Wx)* gene, which is required for the synthesis of amylose in rice endosperm. Previous researches revealed that amylose content in rice is correlated with the ability to excise intron 1 from the leader sequence of the *Wx* transcripts (Wang and Zheng et al., 1995). If intron 1 can not be excised, the 3.3 kb *Wx* pre-mRNA will appear in northern blot analysis besides the 2.3 kb mature mRNA. Therefore, the efficiency of intron 1 auto-splicing in cultivars will directly affect the amylose content in endosperm.

Northern blot analysis of the *Wx* gene transcript was shown in Figure 4.25A. Results revealed that the expression of the *Wx* gene started at 3 DAF (lowest level) and remained at relative high level at later developmental stages, from 10 DAF to 20 DAF. The expression pattern of F1 hybrid is similar to that of P64S (female parental line) but somewhat different from that of 9311, in which the 3.3 kb *Wx* pre-mRNA could be clearly identified and an obvious decrease in expression level after 10 DAF, resulting from low efficiency of intron 1 excision. The expression level during seed formation of 9311 was the lowest among the three rice lines and that of F1 is slightly lower than the P64S (Figure 4.25B). This result is similar to the mature *Wx* protein synthesis during seed formation, as shown by tricine SDS-PAGE gel (Figure 4.26). The *Wx* protein was accumulated at low level at 3 DAF and then maintained at higher level from 6 to 20 DAF for P64S and F1 hybrid, while for 9311, the highest accumulation level appeared at 6 DAF and then declined linearly at later stages. Moreover, the accumulation level of *Wx* protein in 9311 was significant lower than

that of the F1 hybrid (medium level) and more so than the P64S (highest level).

According to the rice amylose contents performed by HHRRC in Changsha, Hunan, the amylose content of 9311 is 14.9%, which is the lowest among the three lines. This content is coincided with the low level of *Wx* mature mRNA because of the inefficient excision of intron 1. The amylose content of F1 hybrid (21.4%) is 0.9% less than that of P64S (22.3%) but 6.5% more than that of 9311 (14.9%) (see Table 4.2). These contents are similar to Dr. Liu Qiaoquan's earlier findings (Liu, et al., unpublished), i.e. 24.9% and 15.0% for P64S and 9311, respectively.).

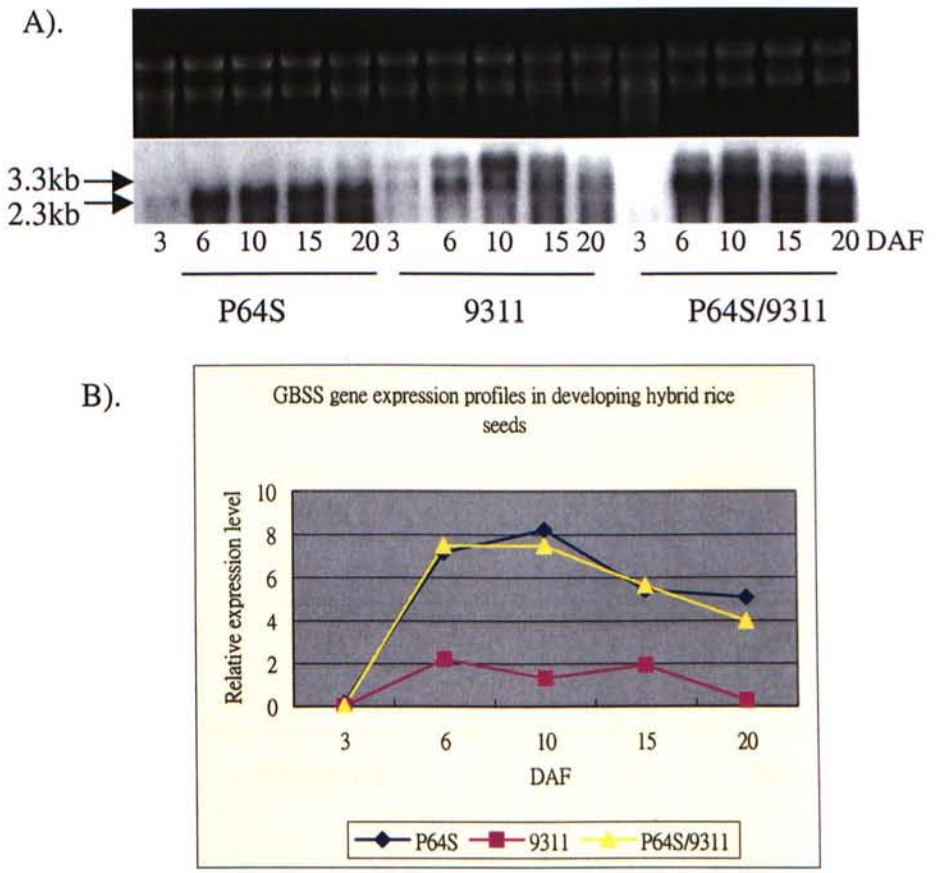


Figure 4.25 The *Wx* protein gene expression profiles in developing hybrid rice seeds.

A) Northern blot analysis of the *Wx* gene. The 3.3 kb band represents *Wx* pre-mRNA containing intron 1; the 2.2 kb fragment represents the mature mRNA. B) Expression patterns of the *Wx* gene, based on northern blot analysis of the *Wx* gene (Figure 4.25 A).

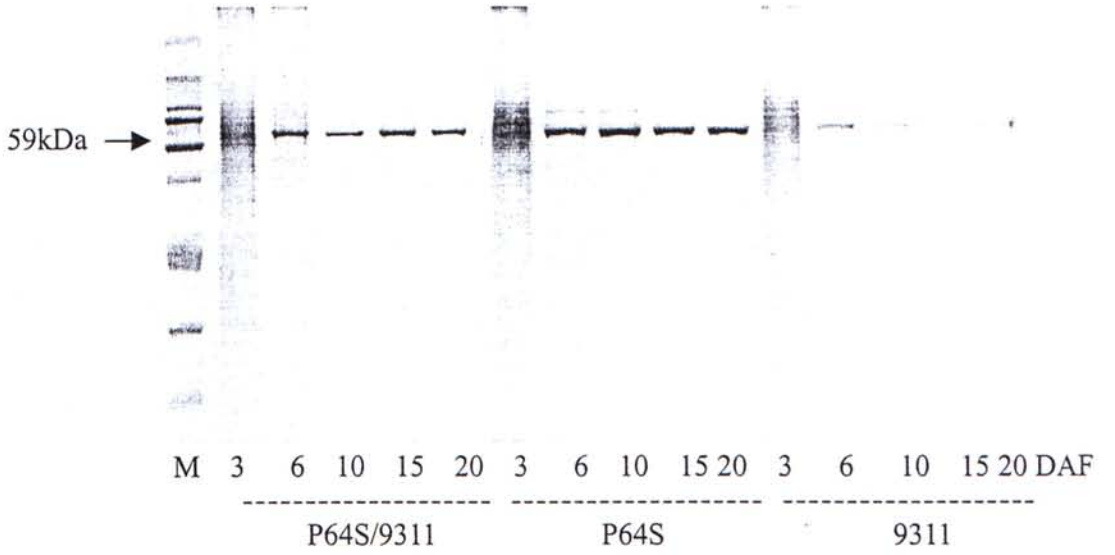


Figure 4.26 Tricine SDS-PAGE electrophoresis of the *Wx* protein from developing rice seeds.

Rice varieties	Amylose content (%)
P64S	22.3
9311	14.9
P64S/9311	21.4

Table 4.2 Amylose content of P64S, 9311 and P64S/9311 at 20 DAF seeds.

4.5.3 Genes encoding rice SSS (Soluble starch synthase) family

4.5.3.1 The SSS1 gene

The SSS1 gene was expressed at around 3 DAF with the lowest mRNA level, after reaching the highest level at 10 DAF, and then its expression declined at later stages. The expression peaked at 10 DAF in P64S, 9311 and F1. The expression patterns in all three lines were similar in general (Figure 4.27-A), while the relative expression level of P64S is the highest and that of F1 is the lowest (Figure 4.28 a).

4.5.3.2 The SSS II-1 gene

The SSS II-1 gene was expressed mainly at the early stages during seed development, with the highest level at 3 DAF and gradually declined toward maturation. At 20 DAF, there was little amount of mRNA present (Figure 4.27-B). The expression patterns of the two parental lines and F1 hybrid were similar and the expression level of F1 hybrid is between its parents (Figure 4.28 b).

4.5.3.3 The SSS II-2 gene

Figure 4.27-C showed that the expression pattern of SSS II-2 gene was similar to that of SSS II-1 although they share relative low DNA sequence homology (60%). Its expression was detected at 3 DAF with the highest level and then linearly declined during subsequent stages. The expression levels of the two parental lines and F1 hybrid were alike (Figure 4.28 c).

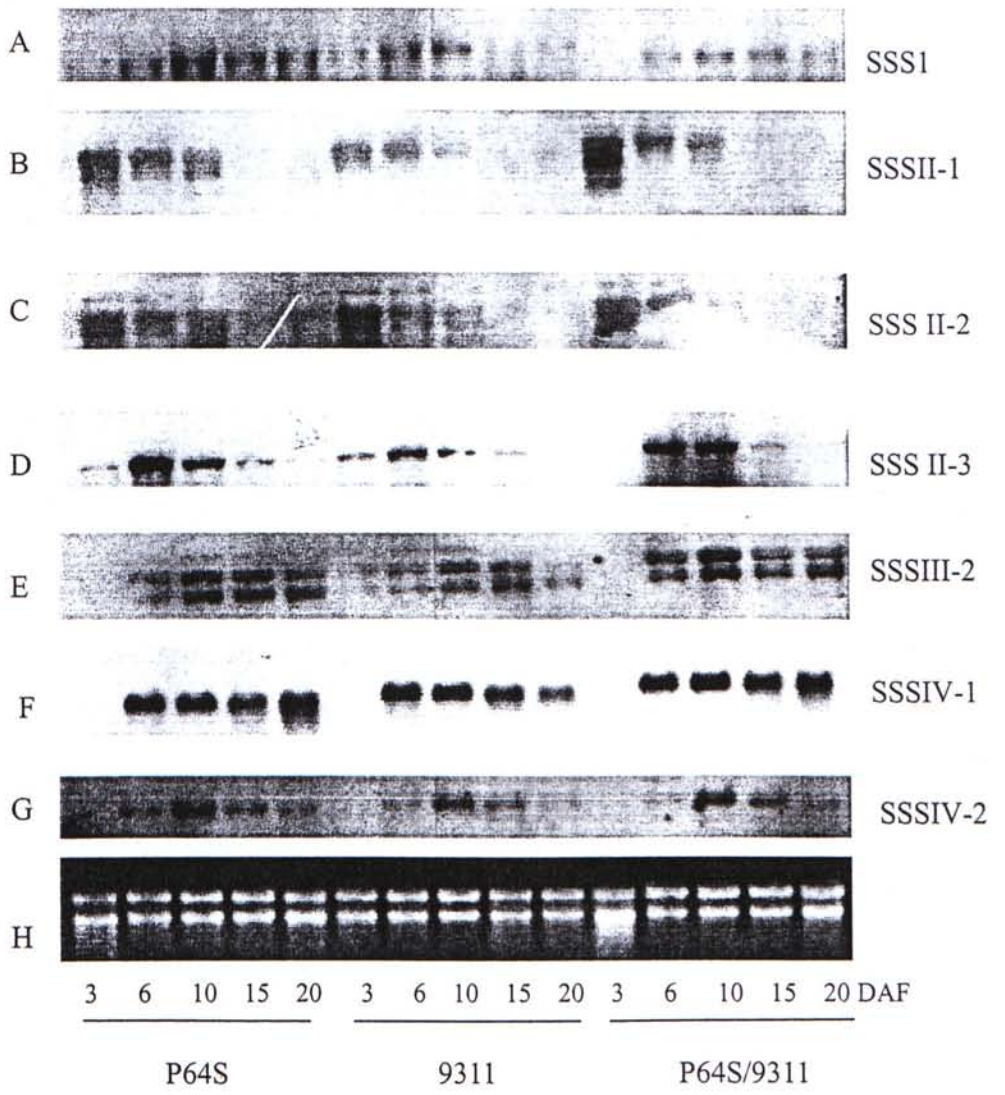


Figure 4.27 Northern blot analysis of starch soluble synthase genes during rice seed development

A) SSS 1 gene; B) SSSII-1 gene; C) SSSII-2 gene; D) SSS II-3 gene; E) SSS III-2 gene; F) SSSIV-1 gene; G) SSS IV-2 gene; and H) Electrophoresis of the total RNA from different lines at different developing stages on 1% agarose/ formamid gel.

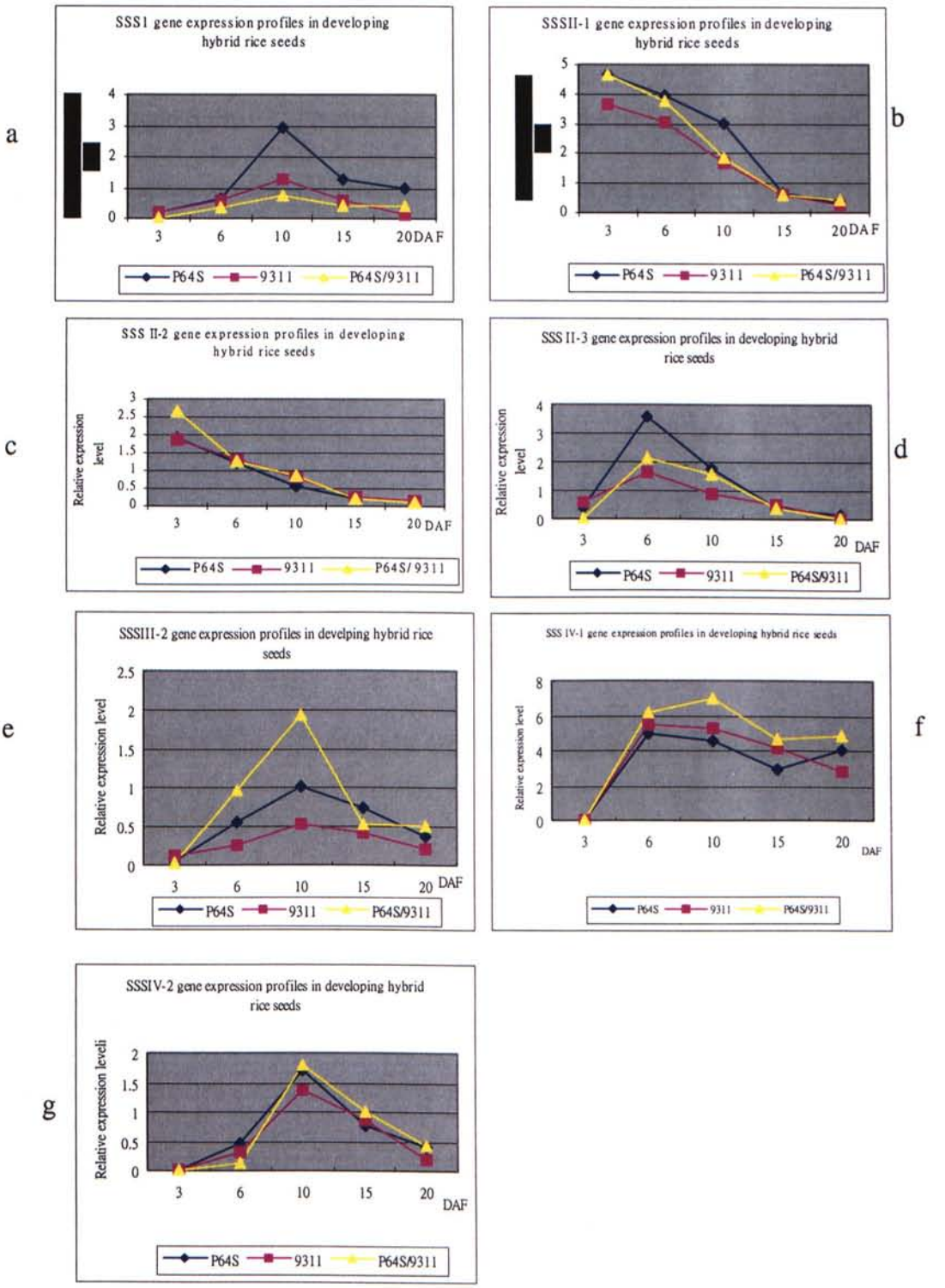


Figure 4.28 Expression profiles of the starch soluble synthase genes

a) SSS1 gene; b) SSSII-1 gene; c) SSSII-2; d) SSSII-3; e) SSSIII-2; f) SSSIV-1 gene; and g) SSSIV-2 gene.

4.5.3.4 The SSS II-3 gene

The expression pattern of SSSII-3 gene was found markedly different from those of SSSII-1 and SSS II-2 although they all belong to the SSS II family. The SSS II-3 gene was expressed early at 3 DAF, after reaching a strong expression peak at 6 DAF, its level gradually declined until 20 DAF where trace amount of mRNA remained. The expression level in P64S reached the highest at 6 DAF but declined to similar level as 9311 and F1 at later stages.

4.5.3.5 The SSS III-2 gene

The SSSIII-2 gene was expressed with a peak activity at 10 DAF with the highest level in P64S and F1 hybrid. A low level of mRNA was detected at 3 DAF in 9311, indicating that the SSS III-2 gene was expressed earlier in 9311 than the P64S and F1 (Figure 4.27 E and 4.28 e).

4.5.3.6 The SSS IV-1 gene

mRNA transcripts of the SSSIV-1 gene could not be detected at 3 DAF, this gene was thus expressed around 6 DAF with relative high level and remained high activity at 10 DAF, then gradually declined towards 15 and 20 DAF (Figure 4.27-F and 4.28 f).

4.5.3.7 The SSS IV-2 gene

Figure 4.27-G showed that the expression of the SSSIV-2 gene peaked at 10 DAF.

Its expression began at low level at 6 DAF, after reaching peak level at 10 DAF, the mRNA level declined gradually through late maturation stages. The levels of expression are comparable in the three rice lines.

4.5.3.8 Profiling the expression of the SSS family genes

In comparison of the expression patterns of the seven genes in the SSS family (Figure 4.29) we found that:

1. The SSS1, SSSII-1, SSSII-2 and SSSII-3 genes were expressed at early stage (3 DAF). Of all these genes, SSSII-1 and SSSII-2 showed the highest expression level at this stage, indicating that these genes play predominant function of SSS at early stage.
2. Expression patterns of SSS1, SSSIII-2 and SSSIV-2 were similar with a sharp expression peak in the mid developmental stage and then gradually declined during subsequent stages.
3. The expression pattern of the SSSIV-1 gene is unique in this group of genes in that it maintained relatively high expression level toward seed maturation after reaching peak activity at 6 DAF. Moreover, its increase expression level appeared at mature stage (20 DAF), indicating that its promoter might be different from the others, which is being active at late stages.
4. The SSSII-1 and SSSII-2 genes share a very similar expression pattern, both of them were expressed at highest level at 3 DAF and then deduced gradually through the whole developing stages.

5. The expression pattern of SSSII-3 is distinctive due to its expression peak at 6 DAF and then decreased gradually at late stages.

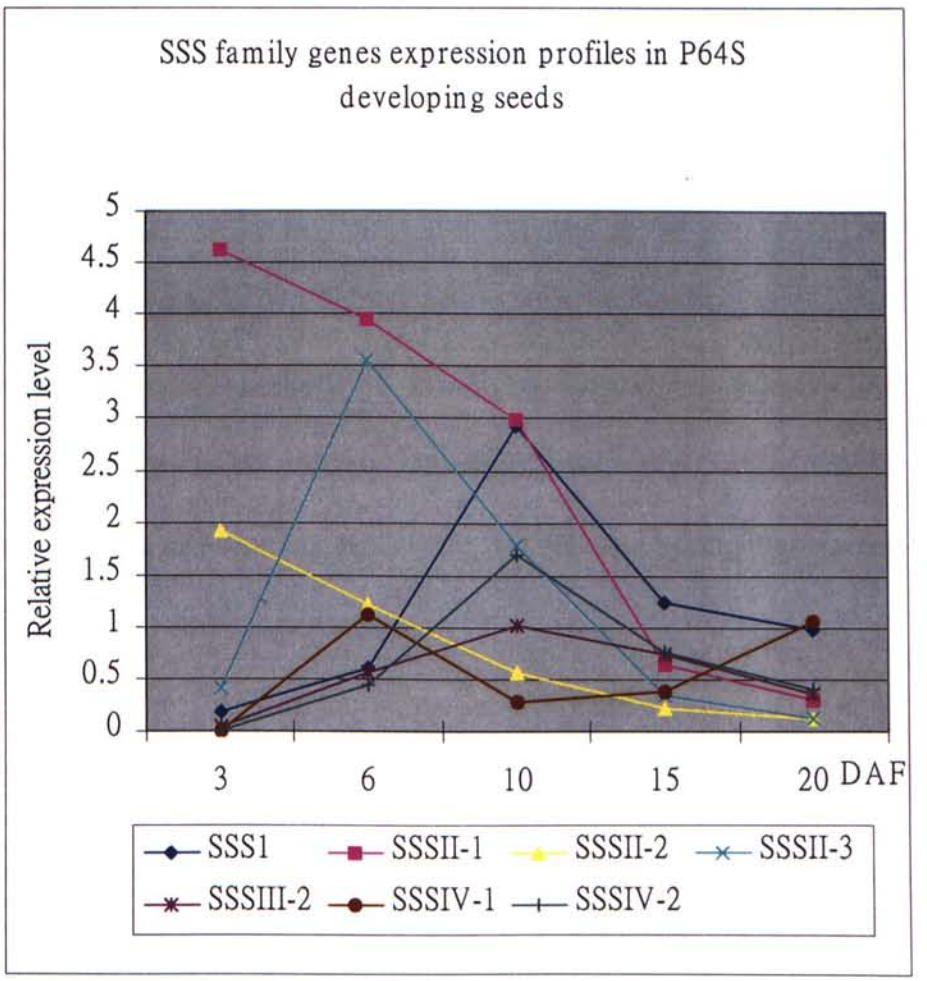


Figure 4.29 Summary of expression profiles of SSS family genes in developing hybrid rice seeds

4.5.4 Genes encoding rice starch branching enzyme (SBE) family

4.5.4.1 The SBE-1 gene

Figure 4.30-A showed that the SBE-1 gene was expressed around 6 DAF, with increasing level at 6 and 10 DAF, and then declined afterward. The highest expression level was at 10 DAF. All the three lines of rice showed similar expression pattern (Figure 4.31 a).

4.5.4.2 The SBE-3 gene

The mRNA transcript of the SBE-3 gene was found to present at lowest level at 3 DAF and 20 DAF, with peak activity at 6 DAF and relative high level at 10 DAF. The expression patterns and levels in P64S, 9311 and F1 hybrid (P64S/9311) were very similar (Figure 4.30-B and Figure 4.31 b).

4.5.4.3 The SBE-4 gene

The SBE-4 gene showed a distinct expression pattern. The gene was expressed strongly at early stage (3DAF); then its mRNA level dropped sharply to very low level at 15 DAF; and finally the level went up somewhat at 20 DAF. All the three rice lines showed similar expression pattern and level (Figure 4.30-C and Figure 4.31 c).

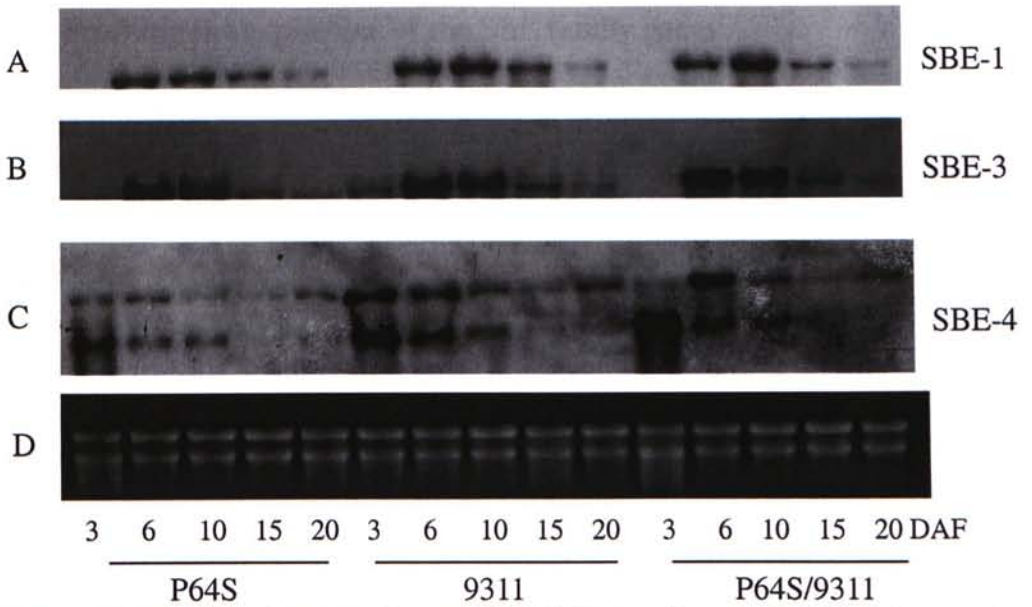
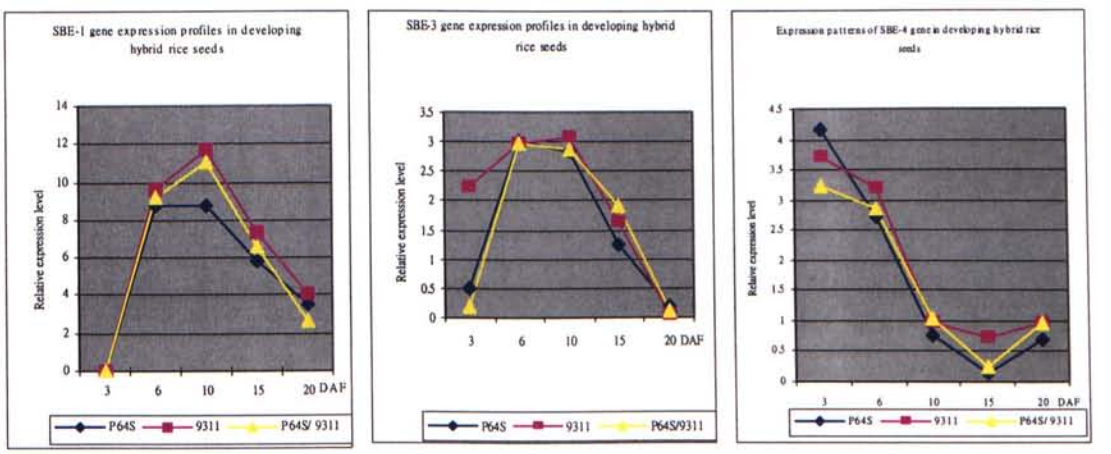


Figure 4.30 Northern blot analysis of the SBE family genes in developing hybrid rice seeds.

A) SBE-1 gene; B) SBE-3 gene; C) SBE-4 gene; and D) Electrophoresis of total RNA on 1% agrose/ formamide gel.



a b c

Figure 4.31 Expression profiles of the SBE family genes

a) SBE-1 gene; b) SBE-3 gene; and c) SBE-4 gene.

4.5.4.4 Profiling the expression of the SBE family genes

In comparison of the expression patterns of the genes encoding the SBE family, namely, SBE-1, SBE-3 and SBE-4 (Figure 4.32), we observed the followings:

1. The SBE-4 gene was expressed with the highest mRNA level at 3 DAF, while SBE-1 and SBE-3 showed no or little amount of mRNA at 3 DAF. Moreover, the expression level of SBE-4 gene was observed to increase again at late maturation stages after reaching the lowest level at 15 DAF.
2. The SBE-3 gene was expressed with the highest level at 6 DAF and relative high level at 10 DAF, while the SBE-1 gene appeared to have expression peak at 10 DAF.

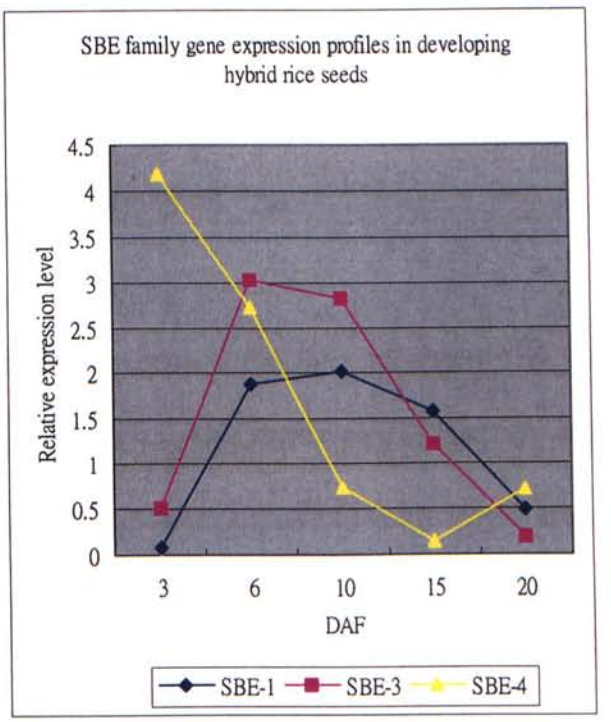


Figure 4.32 Summary of expression profiles of the SBE family genes in developing hybrid rice seeds.

4.5.5 Genes encoding rice starch debranching enzyme (SDE) family

4.5.5.1 The isoamylase gene

The isoamylase gene showed maximum expression between 6 DAF and 10 DAF. Low expression levels were detected at 20 DAF and 3 DAF except a relative higher level was observed at 3 DAF in 9311. The expression pattern and level of F1 hybrid are similar to those of two parental lines (Figure 4.33-A and Figure 4.34 a).

4.5.5.2 The pullulanase gene

The expression pattern of the pullulanase gene was different from that of isoamylase gene in that it showed a clear expression peak at 10 DAF. The expression pattern of F1 is similar to those of the two parental lines and the expression level of F1 is between those of two parents (Figure 4.33-B and 4.34 b).

4.5. 5.3 Difference in expression profile between isoamylose and pullulanase

1). The isoamylase gene was observed to express earlier than the pullulanase gene.

2). The isoamylase gene maintained its high mRNA level at both 6 and 10 DAF, while the pullulanase gene showed an obvious expression peak at 10 DAF.

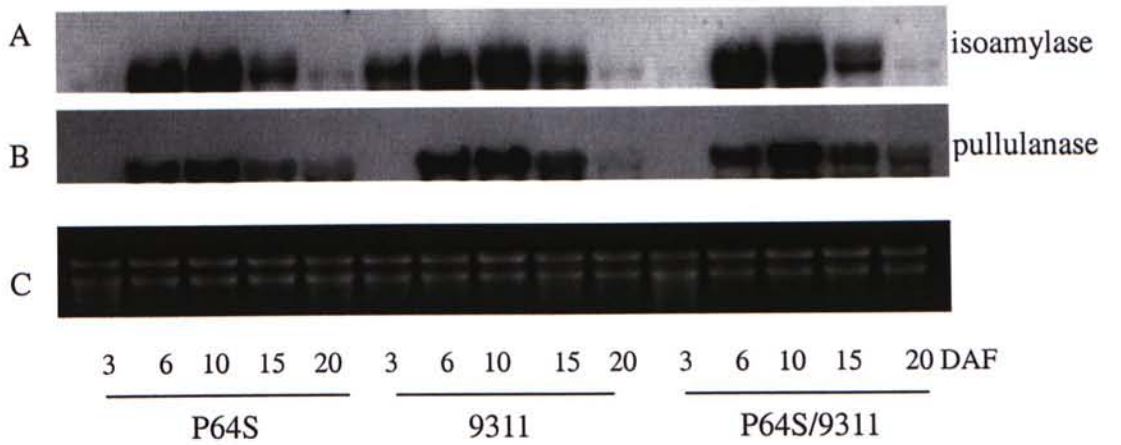


Figure 4.33 Northern blot analysis of the SDE genes in developing hybrid rice seeds

A) Northern blotting of isoamylase gene; B) Northern blotting of pullulanase gene; C) Gel electrophoresis of total RNA on 1% agarose/formamide gel.

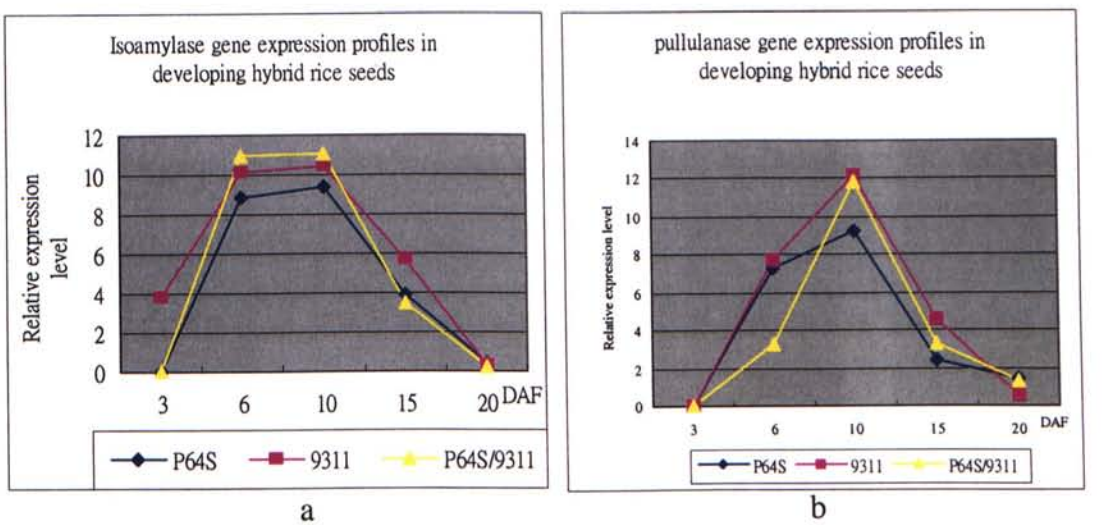


Figure 4.34 Expression profiles of the SDE family genes

a) The isoamylase gene and b) The pullulanase gene.

4.5.6 Comparison of the expression patterns of genes encoding enzymes involved in starch synthesis in developing hybrid rice seeds

Figure 4.35 summarized the expression patterns of all genes encoding the enzymes involved in the complex starch synthesis system.

1. The GBSS gene expressed at very high level at 6 DAF and remained at this level through later stages, especially at 20 DAF.
2. Most of the genes were expressed around 3 DAF except for the SSSIV-1, SSSIV-2 and pullulanase gene (not detectable). Furthermore, GBSS, AGPase, SBE-4, SSSII-1 and SSSII-2 were expressed at relative high level at this stage.
3. Most of the genes reached its highest expression level at 6 DAF, such as AGPase, SBE-3, isoamylase, SSSII-3 and SSSIV-1. In addition, the AGPase, SBE-3, SSSIV-1 and isoamylase genes remained at high level until 10 DAF.
4. A few genes reached its expression peak at 10 DAF, for instance, SSS1, SSSIV-2, SSSIII-2, SBE-1 and pullulanase.
5. At 15 DAF, most of the genes were with decreasing expression level except the AGPase and GBSS gene, still with relatively high level.
6. At 20 DAF, all the genes were at their low expression level except the SSSIV-1, AGPase and GBSS, still maintained relative higher level at this stage.

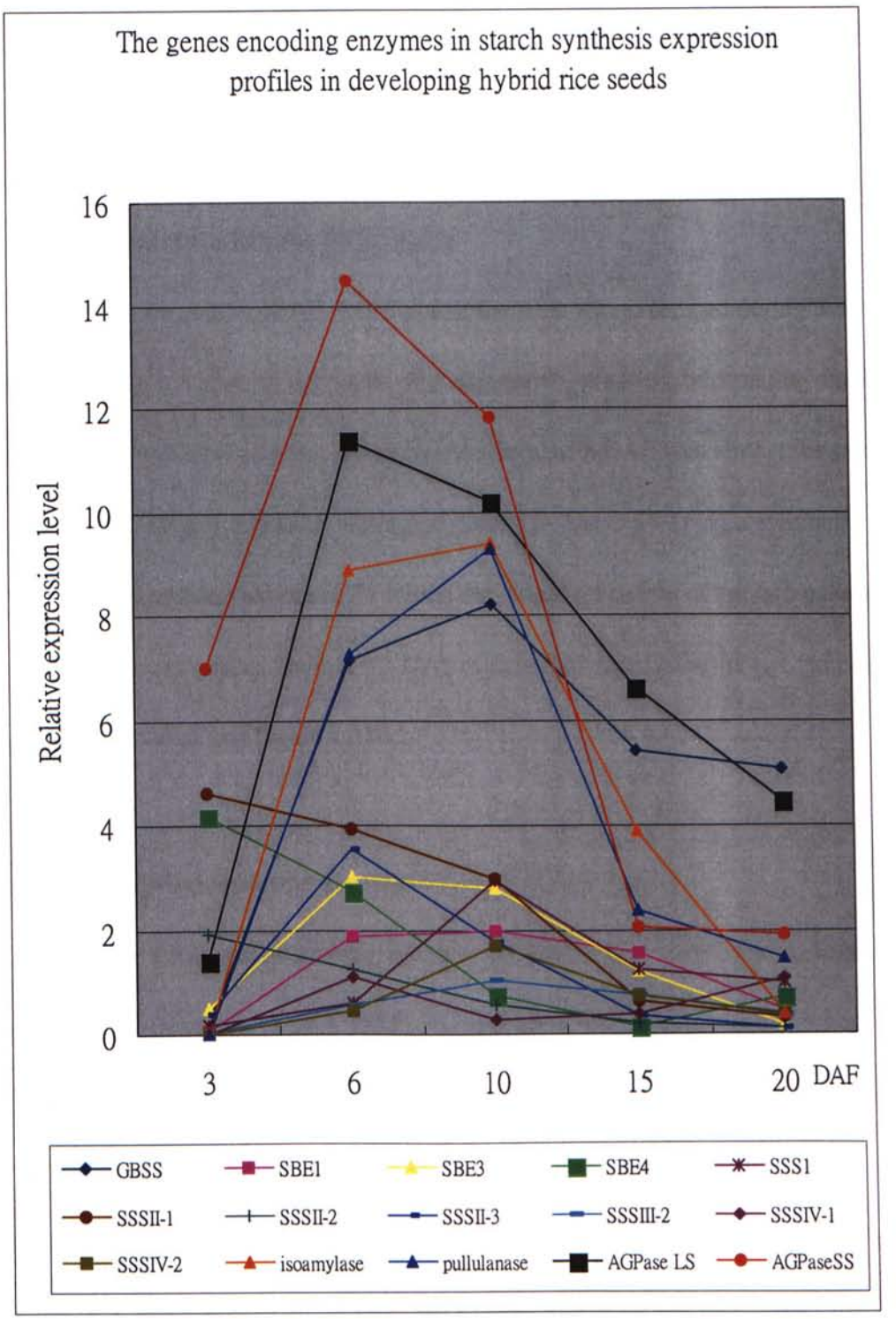


Figure 4.35 Summary of expression profiles of the genes encoding enzymes in starch synthesis during hybrid rice seed development.

4.6 Profiling the expression of genes encoding aspartate family amino acid biosynthetic pathway during developing hybrid rice seed development

4.6.1 Rice aspartate kinase (AK) gene

The expression pattern of AK revealed that the gene was expressed during 3 DAF. After reaching its peak at 10 DAF, the expression gradually decreased during maturation. It was observed that the expression level at 3 DAF was almost the same as that at 20 DAF and at 6 DAF was again similar to that at 15 DAF, a symmetrical pattern. The expression pattern of F1 hybrid was identical to that of the two parental lines while its expression level at 10 DAF was higher than those of two parental lines (Figure 4.36-A and Figure 4.37 a).

4.6.2 The dihydeodipicolinate synthase (DHPS) gene

The DHPS gene was expressed through out the whole developmental stages under study, i.e. from 3 DAF to 20 DAF. Its expression was peaked at 6 DAF, earlier than that of AK gene, and gradually decreased at 10 DAF and 15 DAF. The expression level, however, increased somewhat at 20 DAF. The expression pattern of F1 was similar to that of two parental lines but its expression level was lower than the two parents at later stages (from 10 to 15 DAF) (Figure 4.36-B and Figure 4.37 b).

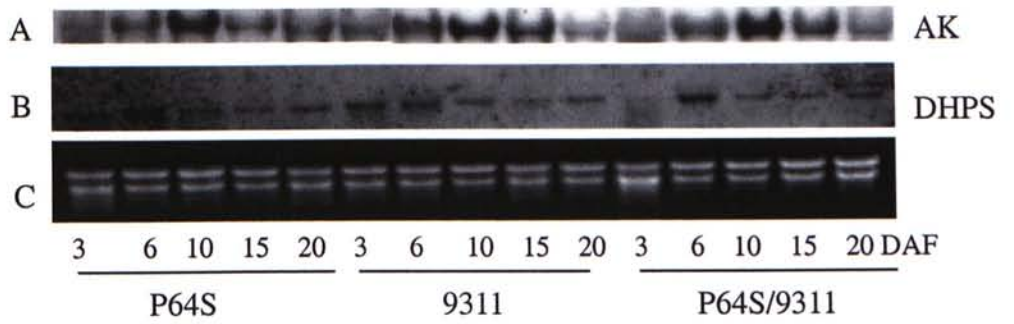


Figure 4.36 Northern blot analysis of the aspartate family amino acid biosynthesis pathway enzymes in developing hybrid rice seeds

A) The aspartate kinase gene; B) The DHPS gene; and C) Gel electrophoresis of total RNA on 1% agarose/formamide gel

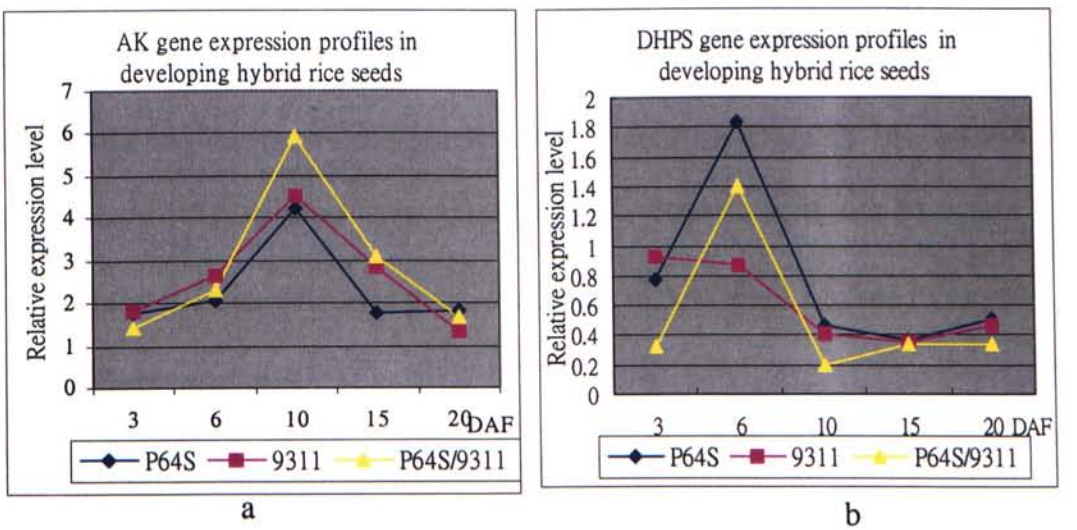


Figure 4.37 Expression profiles of the aspartate family amino acid biosynthesis pathway enzymes in developing hybrid rice seeds

a) The AK gene; b) The DHPS gene.

4.7 Two-dimensional gel electrophoresis and MALDI-ToF seed proteins analysis of rice

Two-dimensional gel electrophoresis and MALDI-ToF were performed to determine the identity of some selected proteins that showed differences between the two parental lines and F1 hybrid. The results from 2-D gel electrophoresis showed that the major differences in protein between the two parental lines and F1 hybrid were mainly of the 13 -16 kDa proteins with pI between 5 and 9 (Figure 4.38).

Fifteen protein spots showed different patterns were picked from the two-dimension gels of P64S, 9311 and F1 hybrid and named as F1, F2, F3, F4, F5, M1, M2, M3, M4, M5, H1, H2, H3, H4 and H5, respectively (Figure 4.39). Of them, M1 showed a much stronger staining signal than that of H1, while F1 lacked this protein spot, suggesting that this protein of 15-kDa and pI 5 was expressed at highest level in 9311 and little or no expression in P64S. On the other hand, M2, M3, M4 and M5 showed very weak or no signal when compared with the P64S and F1 hybrid, indicating that these proteins were expressed mainly in P64S and F1 hybrid (inherited from female parental line).

Mass spectrometer of these fifteen protein spots was carried out. After searching the available mass spectrometer database, most of the protein spots were found no or little homology except F3 and H3, which were determined as the 13-kDa prolamin (Figure 4.40, 4.41 and Table 4.3). Since this spot was present in P64S and F1 hybrid but not 9311, it thus provides interesting and useful information for further

investigation. Moreover, this result was also observed by using tricine SDS-PAGE gel (Figure 4.22), which showed that the 13-kDa polypeptide was expressed only in P64S and F1 hybrid but not 9311. These results strongly suggest that the 13-kDa prolamin is maternally inherited. Similar conclusion was obtained in pervious researches (Kumamaru et al., 1990).

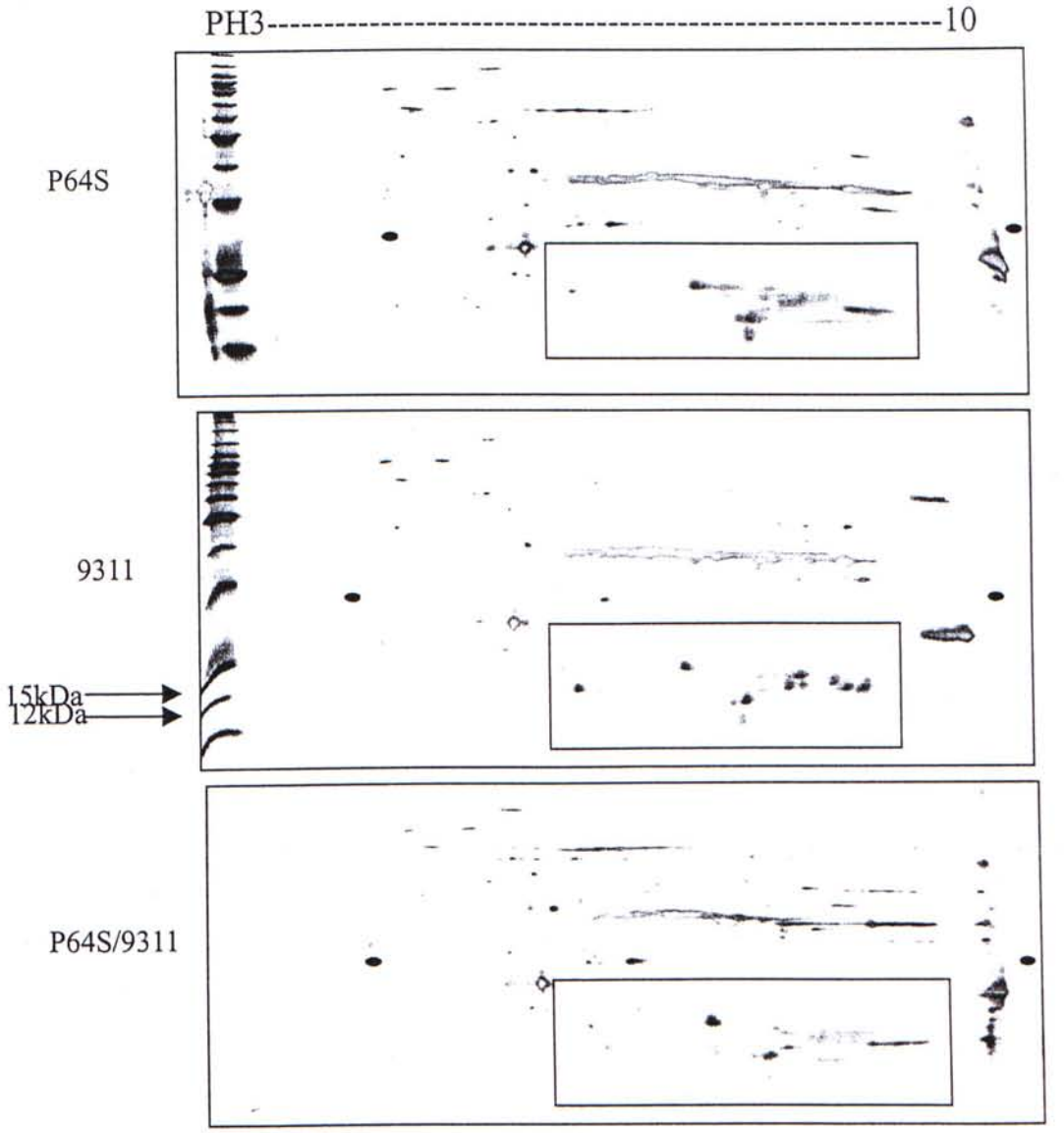


Figure 4.38 Two-dimension gel electrophoresis of the total protein from P64S, 9311 and P64S/9311 at 10 DAF.

The rectangles represent the regions showed different protein spots between the parents and F1 hybrid lines.

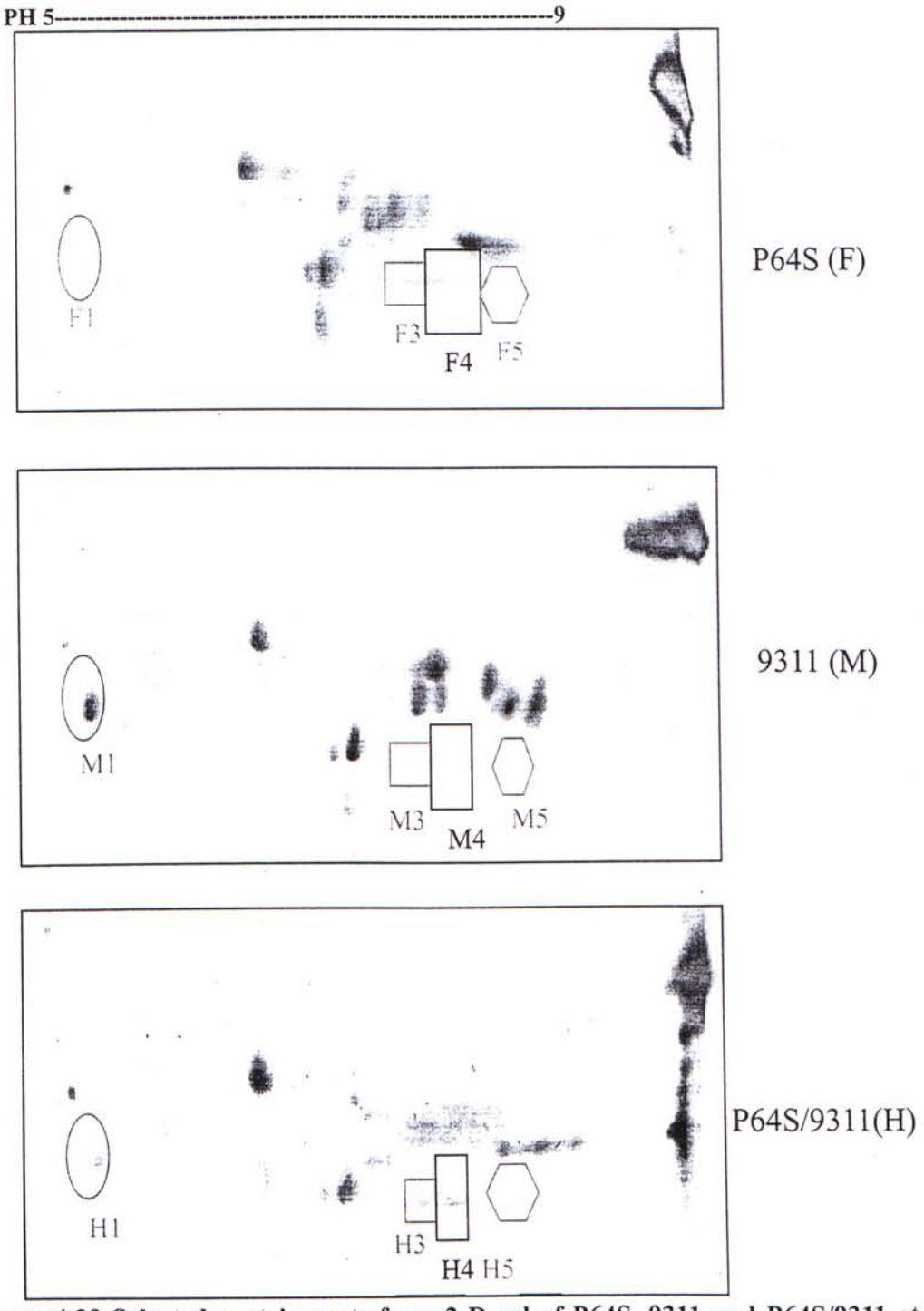


Figure 4.39 Selected protein spots from 2-D gel of P64S, 9311 and P64S/9311 at 10 DAF

F: female parental line; M: male parental line; H: F1 hybrid. The numbers represent different protein spots.

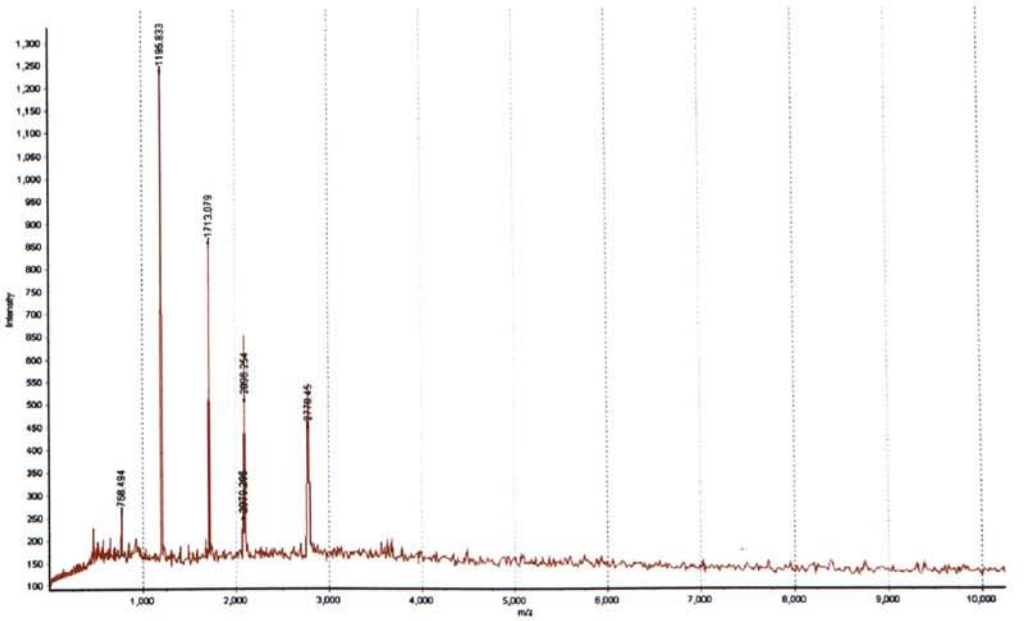


Figure 4.40 Mass spectrometer of F3 protein spot from female parental line (P64S).

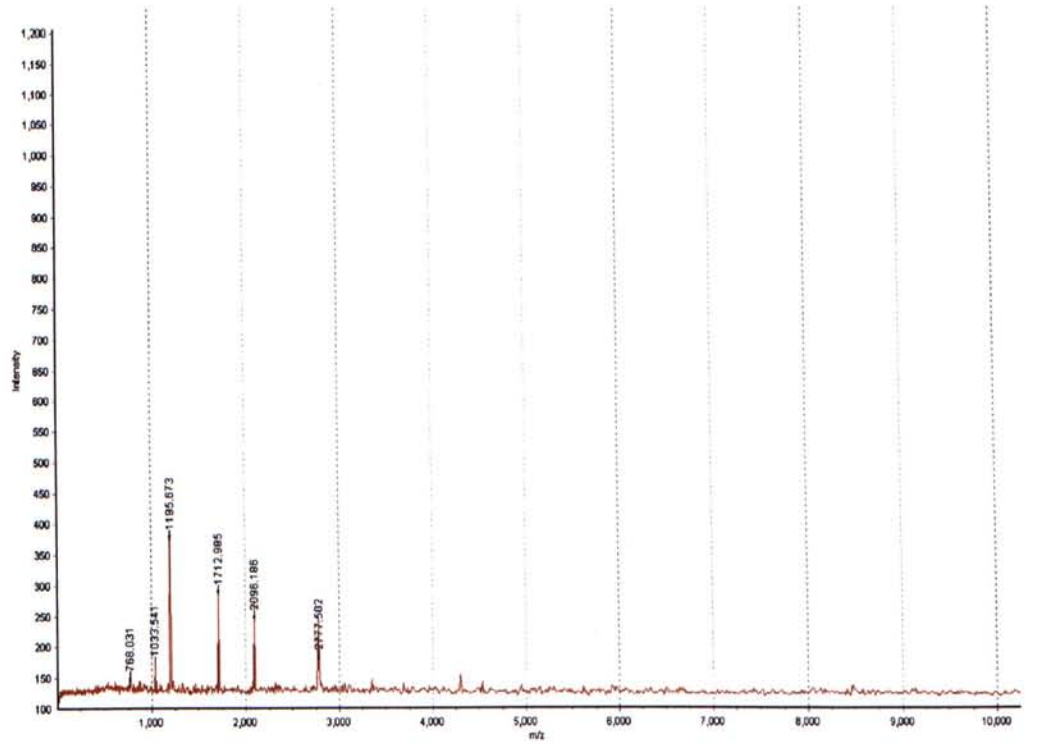


Figure 4.41 Mass spectrometer of H3 protein spot from F1 hybrid.

Table 4.3 MALDI-ToF of selected protein spots from P64S, 9311 and P64S/9311.

No. of picked spots	P64S (F)	9311 (M)	P64S/9311 (H)
1	no	unknown	unknown
2	unknown	unknown	unknown
3	13 kDa prolamin	no	13 kDa prolamin
4	unknown	unknown	unknown
5	unknown	unknown	unknown

Chapter 5 Discussion

5.1 Super hybrid rice as experimental materials and its significance

High yield, good quality and multi-resistance are the main targets for rice improvement worldwide. In the past, high yield is the first criterion for evaluation of rice cultivars in China because of shortage of food supply for the huge Chinese population. However, this unbalanced development led to the ignoring of rice grain quality. Most of the Chinese rice hybrids in the 1970s and early 1980s had poor grain quality. Rice grain with poor quality not only can no longer satisfy the Chinese people with improving living standards, but also greatly affects the export of rice grain from China. Since 1980s, China has been attempting to develop rice hybrids with superior grain quality. For instance, aromatice rice hybrids 'Xiangyou 63', 'Xingxiangyou 77' and 'Xingxiangyou 80' were developed with good quality and high yield potential (Zhou and Liao, 1995 and 1997; Chen et al., 1997a). Nevertheless, it's found that the heredity of nutritional quality traits and eating and cooking quality properties is complicated and greatly influenced by environmental factors. Furthermore, a negative correlation was found to exist between yield and protein content in rice grain in certain range. It is also time-consuming and technical difficult (e.g. linkage of unexpected traits) by using traditional breeding practice in grain quality improvement. Recently, with the advances in molecular biotechnology such as molecular marker, gene transformation, rice genome sequencing, the combination of molecular

biotechnology with traditional breeding technique provides marvelous prospects for rice quality improvement.

Super hybrid rice combination Liangyou Peijiu (P64S/9311) was developed in 1998 by Hunan Hybrid Rice Research Center (HHRRC) and Jiangsu Agricultural Academy in China. This rice variety has a yield of average 12 t/ha in a large-scale production but its grain quality only reaches the second-grade or even third-grade in certain traits. Hence, research on gene expression during grain development for this combination is necessary and significant in improving its grain quality. Hopefully now modified cultivars with super high yield and excellent grain quality can be developed in the future.

5.2 RT-PCR and northern blot as methods to profile gene expression

Several methods are available to profile gene expression under certain conditions, such as RT-PCR, northern blotting, nuclear run-on transcription assays, microarray at mRNA level and SDS-PAGE gel, 2-D gel and western blot at protein level. Among these methods, microarray might be the most efficient way but it's expensive; RT-PCR is fit for distinguishing the members in multigene family with high DNA sequence homology; northern blot analysis is suitable in profiling the members of multigene family with low DNA sequence homology; SDS-PAGE provides a method to identify protein with molecular weight; 2-D gel electrophoresis separates proteins by MW as well as by isoelectric point (pI); and Western blot can identify proteins

with specific antibody.

In this study, both RT-PCR and northern blot were used for genes with high DNA sequence homology in the same multigene family and the ones with low DNA sequence homology, respectively. The results of northern blot are quantitative and accurate, therefore it is preferred in mRNA profiling. However, RT-PCR is necessary in this study to profile mRNAs with high sequence homology in the same gene family. Since RT-PCR is sensitive to template concentrate, temperature, cycle number, reaction volume, PCR machine types, and even different wells in the same PCR machine, a reliable procedure must be established for its accuracy in identification and quantitation of a specific mRNA. Much effort and time was thus devoted in establishing such assay system with satisfaction. To illustrated the quality of the established assay in this study, for example, the results from RT-PCR agreeed with those by northern blot in profiling the expression of GluB-1, 26-kDa globulin, GBSS and SBE-3 (Figure 5.1).

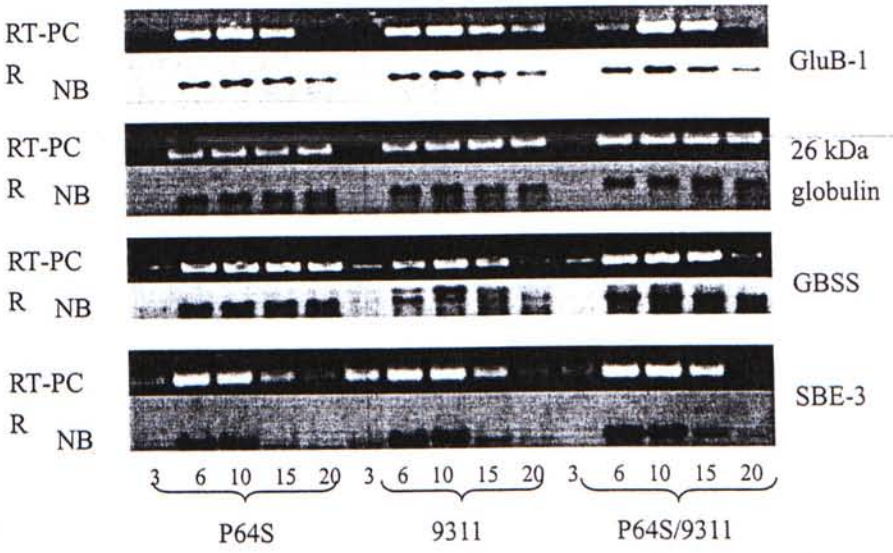


Figure 5.1 RT-PCR and northern blot analysis of GluB-1, 26-kDa globulin, GBSS and SBE-3 gene expression.

5.3 Regulation of genes related to nutritional quality in rice

5.3.1 Storage protein genes

Most of the major rice seed storage proteins, namely glutelin, globulin, prolamin and albumin, are encoded by multigene families. While the expression profiles between different storage proteins differ in time and level during seed formation, as revealed by RT-PCR, northern blot and 2-D gel analysis, the members in a given multigene family, however, also show distinct expression patterns, in time as well as level. Thus, during rice seed maturation, a complex network of storage protein genes is at work in laying down the final heterogeneous storage protein complements. Since storage proteins function as amino acid reserves, without enzymatic activities and particular function, accumulation of sequence mutations and duplications is tolerable and thus their heterogeneity and existence of multigene family. However, it requires further study and understanding as to why the different family members, presumably all serving as amino acid reserves, differ in their expression time and level. Nevertheless, these differences in expression profiles will be regulated by the cis-regulatory elements, such as AACA motif and GCN4 motif, and other elements including prolamin box and the TATA box, in their 5' flanking sequences. These motifs control temporal-specific expression and expression quantity (Muench and Okita, 1997; Zheng et al., 1993). The AACA and GCN4 motifs are most common in promoter regions of seed storage protein genes, particularly in glutelin. Combination of these two motifs is more efficient in enhancing the expression level (Yoshihara et

al., 1996). The results obtained from this study on the diverse expression patterns of various storage proteins and their family members thus provide very useful information and genetic materials for utilizing their regulatory sequences to express and manipulate target genes for rice improvement. For example, based on the 8 groups of expression patterns we can isolate several types of promoters to manipulate gene expression at different time and level in the future.

In this study, we found that the glutelin genes are expressed early at 6 DAF and most of their family members reach the expression peak at 10 DAF except GluB-2 at 10-15 DAF and then gradually decline at late stages. These results are coinciding with the previous works of Kim et al. (1993). However, in their report, Gt1 and Gt2 were increasingly expressed during rice ripening, which was opposite to the results from this study. The reason for this discrepancy may be due to different rice varieties or different probes, Gt1/Gt2 as mixed probe in Kim study and Gt1 and Gt2 individual probe in this study, were used.

The expression of the storage protein genes was found to differ between mRNA and protein levels. For instance, glutelin genes are commonly expressed at highest level at 10 DAF, while its protein accumulation reaches the highest level at 15 DAF in P64S and at 20 DAF in 9311 (Figure 4.5 and 4.8). Most prolamin, globulin and albumin genes are expressed mainly between 10 to 20 DAF, but their proteins accumulated at highest level at 15 DAF in P64S and at 20 DAF in 9311 (Figure 4.9, 4.12, 4.13, 4.16, 4.17 and 4.20). While storage proteins are generally stable and rarely turn over during seed formation, so their continuous accumulation and

difference from their mRNA profiles are understandable. However, this may not explain all the differences observed. Biosynthesis of rice seed storage proteins involves regulation at transcription, post-transcriptional and translational levels, as has been reported in other monocot and dicot storage protein genes (Kim et al., 1993; Reeves et al., 1986; Sorensen et al., 1989; Beach et al., 1985; Walling et al., 1986). For example, in cereal ER membrane-associated translational controls, previous studies suggested that the accumulation of protein is correlated to the efficiency of mRNA located in the distinct ER membrane-bound polysome. Prolamin transcripts were found to be more than that of glutelin at the seed developing later stage but the amount of prolamin was much lower than that of glutelin in mature seeds because 85-100% of the total glutelin transcripts could be recovered in the polysome and post-polysome supernatant fractions, whereas only 65-70% of the total prolamin mRNAs were detected in these fractions and the rests were prone to degrade in cytoplasm (Kim et al., 1993).

As mentioned in chapter 2, prolamin is difficult to digest by human and glutelin contains relative higher proportion of lysine, therefore, to decrease prolamin and increase glutelin content are the feasible strategies to enhance the nutritional quality of rice grain. In this study, we found that GluB-4 gene is expressed higher in F1 hybrid than those in the two parents, indicating that the positive dominant effect exist in this gene. We can use cross-breeding to obtain the rice varieties with high glutelin content. Similarly, Prol 17 gene is expressed at much lower level in F1 hybrid than those in the two parents, indicating the negative overdominant effect exist in Prol 17

gene. We can use the crossing technology to acquire low prolamin content rice varieties.

5.3.2 Lysine synthesis enzymes

As reviewed in chapter 2, previous works on aspartate kinase pathway have been focused on the following aspects: biochemical, genetic, developmental and environmental regulation as well as improvement of amino acid content by transformation of branch point enzymes (Frankard et al., 1992; Karchi et al., 1994; Galili, 1995; Falco et al., 1995; Liu et al., 2002; Gao et al., 2001; Tu, 1998; Hoffman et al., 1987; Shaul and Galili, 1992 and Cheng, 1999). In this study, the expression patterns of genes enconding aspartate kinase (AK) and DHPS, two key enzymes, involved in lysine synthesis pathway during rice seed maturation, were targeted for investigation. We found significant differences between the expression patterns of these two genes. The expression level of DHPS is much lower than that of AK while its expression peaks earlier (6 DAF) than the AK (10 DAF) during seed maturation. At 10 DAF, DHPS mRNA level has decreased to very low level while AK is at its peak activity (Figure 4.37). Since DHPS is the first branch enzyme in lysine synthesis and is competing for the common substrate (3-aspartic semialdehyde from AK) with HSD (branching enzyme in isoleucine and methionine synthesis, Figure 2.2), its late expression may render disadvantage to lysine synthesis. To extend the expression of DHPS by using an appropriate promoter may be a practice method, for example, the promoters with constant strong expression at the later developmental stages of group I

gene, such as Prol 17 and 26 kDa globulin or the promoter from AK gene with its highest expression level at 10 DAF, to match the AK gene expression pattern during seed development if the half life of DHPS mRNA is short. On the other hand, to reduce the free lysine feedback inhibition to DHPS and AK, introduction of lysine-rich protein, lysine-insensitive DHPS and antisense to block the LKR enzyme should be performed at the same time. The profiling information from this study thus is with critical importance to our effort in increasing the lysine content through molecular manipulation.

5.4 Regulation of genes related to cooking and eating quality in rice

Regulation of genes related to starch biosynthesis has been studied in the past (Chapter 2 for review), including the increase and decrease of amylose content by regulation of AGPase and GBSS activities, indicating both GBSS and AGPase play very important role in starch biosynthesis and in its property. However, little information was available in the modification of amylopectin content by regulation of SSS and SBE gene. In this research, the expression profiles of all the major starch synthesis genes and their family members were studied. Diverse expression profiles differed in expression time and level during seed formation were observed (Figure 4.35). This diversity provides important insights in our understanding of starch biosynthesis in maturing rice seeds and how their starch property, thus rice cooking and eating quality, is determined at molecular level. This background information and

its associated genetic materials, including regulatory sequences (promoters) that control a range of expression time and level and coding sequences encoding distinct enzymes/isozymes determining the quality and ratio of amylose and amylopectin, will provide options in our effort to manipulate rice starch property. For example, while the AGPase large and small subunit genes are coordinately expressed during seed development (Figure 4.24) so that efficient formation of heterotetramer AGPase can be achieved, the *Wx* (GBSS) gene in the male parental line (9311) produces two populations of mRNA differed in size (2.3 and 3.3 kb; Figure 4.25), as a result of different efficiency in intron I splicing, leading to lower amylose content in its endosperm. This finding allows us to develop strategy to manipulate the splicing site for starch property engineering (Umemoto et al., 1995; Isshiki et al., 1998 and Hirano et al., 1998). In this connection, we also observed that SSSII-3, SSSIII-2 (Figure 4.27) and SBE-4 (Figure 4.30) produced two sizes of mRNAs in northern blot analysis, which were not reported before. Their biochemical significance and possible involvement of auto-splicing of intron such as in the *Wx* gene, require further elucidation. Moreover, based on the difference in expression level between the two parents and F1 hybrid, we can manipulate their genes expression. For instance, GBSS is expressed at medium level in F1 hybrid, that is, lower than female parental line and higher than male parental line, so we can use high amylose content variety cross low content variety to obtain intermediate content variety.

The genes encoding the isoforms of SSS (Figure 4.29), SBE (Figure 4.32) and SDE (Figure 4.34) showed diverse expression profiles in time and level during rice

seed development. As they possess distinct activities in determining the starch formation, chain length, and amount of amylopectin, manipulation of their expression should enhance our capability in improving rice cooking and eating quality. For example, *japonica* rice generally has more short chain (S-type) amylopectin, due to its lower SSSII activity, but better cooking and eating quality than *indica* rice. It's suggested that the short chain amylopectin is correlated to better grain quality. Thus to obtain more S-type amylopectin in *indica* rice by inhibition of SSSII activity could be an approach. As for SBE isoforms, SBE-4 is expressed earlier than SBE-1 and SBE-3, but the mRNA level of SBE-4 is much lower than that of the two others. This result agrees with the previous work of Mizuno et al. (2001). In this study, we found that the expression patterns in F1 hybrid and two parents are similar but their expression levels are different.

5.5 Heredity of genes expression in F1 hybrid

This study also identified several types of heredity in genes expression between the F1 hybrid with its two parents. In type 1, the expression level is the same between F1 hybrid and its two parents, as observed in most of the genes under study; in type 2, the F1 hybrid has higher expression level than those of the two parents, as in Gt2, GluB-4, 10 kDa prolamin, RP6, prol 7, Prol 14 RA5, RA14, SSSIII-2, SSSIV-1 and AK; in type 3, the F1 hybrid contains medium expression level between the two parents, as in Globulin 1, Globulin 2, GBSS, SSSII-1, SSSII-3, pullulanase and DHPS; and in type 4, the F1 hybrid has lower expression level than those of two parents, as in GluA-3,

prol 17 and SSS1. While this diverse inheritance patterns, in relation to heterosis, if any, appears obscure, information from SDS-PAGE, crude protein, and amylose content analysis revealed that crude protein and amylose content in F1 hybrids are between those of the two parents, i.e. higher than 9311, but lower than P64S, indicating that the heterosis exhibiting in this super hybrid combination (P64S/9311) may be mainly in morphological traits and yield properties, such as excellent plant posture, higher lodging resistance, larger panicles, and more spikelets. To understand and utilize, if possible, the genes with overdominance to improve storage protein content and quality and to modify the amylose content in hybrid rice seeds are certainly with basic interest and challenging; Further, it has great application value.

5.6 Application of 2-D gel electrophoresis

Two-D gel electrophoresis is a technique for identifying proteins and a basic method for large-scale analysis of proteome variations, offering insights into the quantity and quality of the final gene products, the proteins. Comparison between the results from northern blot and SDS-PAGE analysis, we found that the amounts and changes in proteins are not correlated to those of the mRNA in some of the genes under study, which may be due largely to the differences between mRNA and protein turn over, accumulation, and to the fact that many proteins undergo post-translational modifications. Hence, the study of proteins (namely proteomics) can provide further insights into the expression and function of seed quality genes. In this study, analysis of 2-D gel electrophoresis of the total protein of two parents and F1 hybrid showed

obvious differences among them. However, only one (13-kDa prolamin) of these different protein spots can be identified by MALDI-ToF method due to the lack of available protein database information. Thus, further development of efficient protein identification (e.g. mass spectrometer) methods and establishment of search database, in this case, for cereal proteins, are important.

5.7 Future perspectives

In the current study, we have chosen the parents lines and its F1 hybrid of an elite super hybrid rice combination (P64S/9311) as materials to profile the expression of genes related to grain nutrition and cooking and eating quality at mRNA and protein levels. Results from this study have provided the basic information and useful genetic materials in our understanding of the complex network of gene expression during rice grain development in relation to its final quality.

In the future, studies should be focus on selected genes of importance and use the information and genetic materials (e.g. promoters) available, to manipulate the expression of key genes determining grain quality for hybrid rice improvement.

Chapter 6 Conclusion

Super hybrid rice combination P64S/9311 as well as its female (P64S) and male (9311) parental lines were chosen as model plants to study the expression profiles of genes related to grain quality at both transcript and protein levels during seed formation. Thirty-six genes encoding seed storage proteins, starch biosynthesis enzymes and lysine synthesis pathway key enzymes, were included in the study. The following conclusion can be drawn:

1. Families of genes encoding rice storage proteins and starch biosynthesis enzymes are functioning in the hybrid rice F1 and parental lines during seed formation. Their expression patterns vary between families and within family, at both time and level, exhibiting a complex network of gene actions.
2. Several types of expression patterns, however, can be recognized, as shown in Figure 6.1:
   a) Group I: Some genes reach their highest expression level at 6 DAF and remain so until maturity (20 DAF), such as α-globulin, Prol 17, GBSS and SSSIV-1. A variation is noted in this group in that some genes show their maximum expression level later at 10 DAF and remain so till 20 DAF, such as RP6, Prol 14, RA5, RA17 and LMW globulin.

b) Group II: The bell shape pattern. In this group, the genes begin to express at 3 DAF, after reaching the peak activity gradually, their expression level starts to decline during seed maturation. The Gt1, GluB-1 and SBE-1 genes are in this group.

c) Group III: The sharp peak pattern. The genes in this group are expressed throughout seed development with a pattern like group II; however, the rise and fall of its expression level is steep, with a sharp peak, such as the AK, Prol 7, SSS1 and SSSIII-2 genes. Some genes show minor variation from this pattern, i.e. their expression starts later at 6 DAF, for example, the Gt2, Gt3, GluB-4, SSSIV-2 and pullulanase genes while globulin 1 gene expression starts at 10 DAF.

d) Group IV: The plateau maximum pattern. A few genes maintain their top expression level for a few days (6 to 10 DAF or 10 to 15 DAF) during the rise and fall of its level, e.g. isoamylase, SBE-3 and GluB-2.

e) Group V: The expression of some genes peaks at early development (6 DAF) and then declines gradually at the later stages, e.g. the AGPase small subunit, AGPase large subunit, SSSII-3 and DHPS genes.

f) Group VI: The expression of a few genes peaks later in development, at 15 DAF, e.g. the globulin 2 and RA14 genes.

g) Group VII: A few genes are expressed at high level at early stage but linear show decrease in expression level during seed maturation, for example, the SSSII-1 and SSSII-2 genes. Expression pattern of SBE-4 is similar to this group except it shows a slightly increase at 20 DAF.

h) Group VIII: The 10-kDa prolamin gene shows a two-peak expression profiles at 6 DAF and 20 DAF.

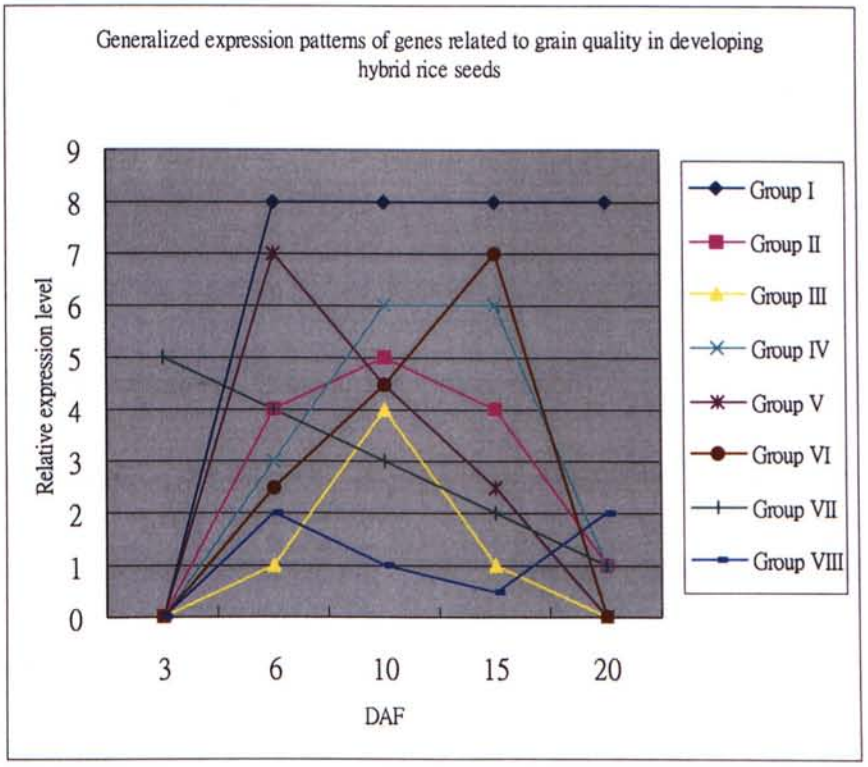


Figure 6.1 Generalized expression patterns of genes related to grain quality in developing hybrid rice seeds

This graph was prepared according to RT-PCR results and northern blot analysis.

3. Although the genes between families and within a family vary greatly in their

expression profiles during seed development, some genes are expressed in a coordinated manner such as the large- and small-subunit genes of AGPase.

4. The relative expression levels of the grain quality related genes differ greatly during seed formation, from very low (e.g. RA14) to very high (e.g. Prol 17), amounting some 12 folds in difference (Figure 4.21).

5. Most enzyme genes involved in starch and lysine synthesis peak their expression at 6 DAF, earlier than that (10 DAF) of the genes encoding seed storage proteins, suggesting that starch and lysine synthesis is highly active at early stage during seed development, while the storage proteins begin to accumulate mainly after mid-maturation stage.

6. Relationships of the expression level between the F1 hybrid and its two parental lines also vary among the different genes under study.

   a) Most genes in the F1 hybrid are expressed at the same level as those of the two parents, such as GluB-2, RA17, 26 kDa globulin, SBE-3, SBE-4, SSSII-2, SSSIV-2, AGPase large subunit and isoamylase.

   b) Expression levels of the GluB-4, 10 kDa prolamin, RA5, SSSIV-1, SSIII-2 and AK genes are higher in the hybrid than in the two parents during seed development. Some genes in the F1 hybrid exhibit higher level than the two parents, but only during 15-20 DAF, e.g. the Gt2, RP6, Prol 7, Prol 14 and

AGPase small subunit genes.

c) Six genes in the F1 hybrid show expression levels in F1 hybrid in between the two parents, namely the globulin 2, GBSS, SSSII-1, SSSII-3, pullulanase and DHPS genes.

d) A few genes are expressed at much lower level in the hybrid than in the two parents, such as the GluA-3, Prol 17, LMW globulin and SSS1 genes.

7. At the protein level, accumulation of the four storage proteins in parental lines and F1 hybrid also vary. In female parental line (P64S), the four storage proteins generally appear at 6 DAF and reach the highest level at 15 DAF; for the male parental line (9311), proteins appear at 6 DAF and remain accumulating in later stages; whereas accumulation of proteins in F1 is under various degrees of influence by their parents. For example, glutelin expression pattern in F1 is distinct from the two parents; globulin in F1 is similar to 9311; albumin in F1 is more like the P64S; and prolamin, however, is strongly correlated to female parental line in its expression quantity but closely similar to the male parental line in its expression pattern. The expression patterns of some seed proteins at the mRNA level were found not in agreement with those at the protein level. This could be resulted from difference in protein accumulation and other post-translational modifications and protein processing, documenting the

importance of proteomics study.

8. Comparing the relative gene expression level between the two parents, it's found that most genes are expressed with higher level in P64S than in 9311, with the exception of GluA-3, RP6, Prol 17, globulin 2, pullulanase and SBE-1 where they are expressed higher in 9311.

9. In sum, this study revealed that the 36 genes related to grain quality are expressed with diverse developmental profiles, resulting from regulation at transcriptional, translational, and post-translational levels. Though complex, this information provides an entry point into our understanding of how these genes at work collectively in determining the grain quality during seed maturation, so that further research and development can be focused and pursued, for example, the effects of genotype and environment, in rice quality improvement.

References

Adachi, T., Izumi, H., Yamada, T., Tanaka, K., Takeuchi, S., Nakamura, R. and Matsuda, T. (1993) Gene structure and expression of rice seed allergenic proteins belonging to the alpha-amylase/trypsin inhibitor family. *Plant Mol. Biol.* 21: 239-248.

Akbar, M. and Yabuno, T.T. (1975) Breeding for saline-resistant varieties to rice. III. Response of F1 hybrids to salinity in reciprocal crosses between Jhona 349 and Magnolia. *Japan. J. Breed.* 25: 215-220.

Alvarez, A.M., Adachi, T., Nakase, M., Aoki, N., Nakamura, R. and Matsuda, T. (1995) Classification of rice allergenic protein cDNAs belonging to the α-amylase/trypsin inhibitor gene family. *Biochem. Biophysi. Acta.* 1251: 201-204.

Argos, P., Narayana, S.V.L. and Nielsen, N.C. (1985) Structural similarity between legumin and vicilin storage proteins from legumes. *EMBO J.* 4: 1111-1117.

Azevedo, R.A., Smith, R.J. and Lea, P.J. (1992) Aspartate kinase regulation in maize: Evidence for co-purification of threonine-sensitive aspartate kinase and homoserine dehydrogenase. *Phytochemistry* 31: 3371-3374.

Baba, T. Kimura, K., Mizuno, K., Etoh, H., Ishida, Y., Shida, O. and Arai, Y. (1991) Sequence conservation of the catalytic regions of amylolitic enzymes in maize branching enzyme-I. *Biochem. Biophys. Res. Commun.* 181: 87-94.

Båga, M., Nair, R.B., Repellin, A., Scloes, G.J. and Chibbar, R.N. (2000) Isolation of a cDNA encoding a granule-bound 152-kilodalton starch-branching enzyme in wheat. *Plant Physiol.* 124: 253-263.

Bai, D.L. and Luo, X.H. (1996) Liangyou Peite, a new two-line hybrid rice released in China. *Int. Rice Research Notes* 1: 42-43.

Bai, S.N. and Xiao, Y.H. (1988) A study of the root growth and respiration of hybrid rice. *Acta Agron. Sinica* 1: 53-59.

Beach, L.R., Spencer, D., Randall, P.J. and Higgins, T.J.V. (1985) Transcriptional and

post-transcriptional regulation of storage protein gene expression in sulfur-deficient pea seeds. *Nucl. Acids Res.* 13: 999-1013.

Beach, L.R. and Ballo, B. (1992) enhancing the nutritional value of seed crops. In *Biosynthesis and Molecular Regulation of Amino Acids in Plants.* pp 229-236, Singh, B.K., Flore, H.E. and Shannon, J.C. (ed.), American Society of Plant Physiology.

Beatty, M.K., Rahman, A., Cao, H., woodman, W., Lee, M., Myers, A.M. and Jamnes, M.G. (1999) Purification and molecular genetic characterization of ZPU1, a pullulanase-type starch-debranching enzyme form maize. *Plant Physiol.*119: 255-266.

Bechles, D.M., Smith, A.M. and Rees, T.A. (2001) A cytosolic ADP-Glucose Pyrophosphorylase is a feature of graminaceous endosperms, but not of other starch-storing organs. *Plant Physiol.* 125: 818-827.

Bechtel, D.B. and Juliano, B.O. (1980) Formation of protein bodies in the starchy endosperm of rice (*Oryza sativa* L.): a re-investigation. *Ann. Bot.* 45: 503-509.

Beckles, D.M., Smith, A.M. and Ree, T. (2001) A cytosolic ADP-glucose pyrophosphorylase is a feature of graminaceous endosperms, but not of other starch storing organs. *Plant Physiol.*125: 818-827.

Bhattacharyya, M.K., Smith, A.M., Ellis, T.H.N., Hedley, C. and Martin, C. (1990) The wrinkled-seed character of pea described by Mendel is caused by a transposon-like insertion in a gene encoding starch-branching enzyme. *Cell* 60:115-122.

Blakeney, A.B. (1980) Studies on rice starch. The accumulation and properties of starches in the stem, pollen and grain of some rice varieties grown in southern New South Wales. *M. S. Agric. Thesis*, Univ. of Sydney, Australia, pp140.

Blauth, S.L., Yao, Y., Klucinec, J.D., Shannon, J.C., Thompson, D.B. and Guiltinan, M.J.(2001) Identification of *mutator* insertional mutants of starch- branching enzyme 2a in corn. *Plant Physiol.* 125: 1396-1405.

Briones, V.P., Magbanua, L.G. and Juliano, B.O. (1968) Changes in physicochemical properties of starch of developing rice grain. *Cereal Chem.* 45: 351-357.

Bruce, A.B. (1910) The mendelian theory of heredity and the augmentation of vigour. *Science* 32: 627-628.

Bryan, J.K. (1990) Advances in the biochemistry of amino acid biosynthesis. *In: the Biochemistry of Plants*, Vol.5, pp 403-452, B.J.Milflin, ed. (New York Academic Press).

Burstin, J., de Vienne, D., Dubreuil, P. and Damerva, C. (1994) Molecular markers and protein quantities as genetic descriptors in maize. I. Genetic diversity among 21 inbred lines. *Theor. Appl. Genet.* 89: 943-950.

Burton, R.A., Bewley, J.D., Smith, A.M., Bhattacharyya, M.K., Tatge, H., Ring, S., Bull, V., Hamilton, W.D.O. and Martin, C. (1995) Starch branching enzymes belonging to distinct enzyme families are differentially expressed during pea embryo development. *Plant J.* 7: 3-15.

Burton, R.A., Johnson, P.E., Beckles, D.M., Fincher, G.B., Jenner, H., Naldrett, M.J. and Denyer, K. (2002) Characterizaiton of the genes encoding the cytosolic and plastidial forms of ADP-glucose pyrophosphorylase in wheat endosperm. *Plant Physiol.*130: 1464-1475.

Castle, W.E. (1946) Genes which divide species or produce hybrid vigor. *Proc. Nat. Acad. Sci. (USA)* 32: 145-149.

Chang, W.W., Huang, L., Sheng, M., Webster, C., Burlingame, A.L. and Roberts, J.K. (2000) Patterns of protein synthesis and tolerance of anoxia in root tips of maize seedlings acclimated to a low-oxygen environment, and identification of proteins by mass spectrometry. *Plant Physiol.* 122: 295-318.

Chavan, J.K and Duggal, S.K. (1978) Studies on the essential amino acid composition, protein fractions and biological value of some new varieties of rice, *J. Sci. Food Agric.* 29: 235.

Cheng, S.M.K. (1999) Expression of LRP in *Arabidopsis* Plants. M. Phil. thesis, Department of Biology. The Chinese University of Hong Kong.

Choi, S.B., Wang, C., Muench, D.G., Ozawa, K., Franceschi, V.R., Wu, Y. and Okita,

T.W. (2000) Messenger RNA targeting of rice seed storage proteins to specific ER subdomains. *Nature* 407: 765-767.

Choudhury, N.H. and juliano, B.O. (1980a) Lipids in developing and mature rice grain. *Phytochemistry* 19: 1063-1069.

Cohen, G.N. and Saint-Girons, I. (1987) Biosynthesis of threonine, lysine and methionine. In *Escherichia coli and Salmonella typhimurium*: *Cellular and Molecular Biology*, pp 429-444, F.C. Neidhardt, ed. (Washington, DC: American Society for Microbiology).

Coleman, C.E., Lopes, M.A., Gillikei, J.W., Boston, R.S. and Larkins, B.A. (1995) A defective signal peptide in the maize high-lysine mutant fleury-2. *Proc. Nat. Acad. Sci. (USA)* 92: 6828-6831.

Commuri, P.D. and Keeling, P.L. (2001) Chain-length specificities of maize starch synthase I enzyme: studies of glucan affinity and catalytic properties. *Plant J.* 25:475-486.

Costa, P., Bahrman, N., Frigerio, J.M., Kremer, A. and Plomion, C. (1998) Water-deficit-responsive protein in maritime pine. *Plant Mol. Biol.* 38: 587-596.

Croissant-Sych, Y. and Okita, T.W. (1996) Identification of positive and negative cis-elements of the rice glutelin Gt3 promoter. *Plant Sci.* 116: 27-35.

Crow, J.F. (1997) Dominance and overdominance. In: *The genetics and exploitation of heterosis in crops.* pp10. Coors, J.G., Pandey, S., et al. (ed.). Proc. Int. CIMMYT symp., ASA, CSSA, SSSA, Madison, WI (in press), 17-22 August 1997, CIMMYT, Mexico.

Dai, Z.Y., Zhao, B.H., Liu, X.J., Xia, G.H., Tan, C.L., Zhang, B.Q. and Zhang, H.X. (1997) 優質高產多抗中秈稻新品種揚稻 6 號. *江蘇農業科學* 4: 13-15.

Davenport, C.B. (1908) Degeneration, albinos, and inbreeding. *Science* 28: 454-455.

Denyer, K., Dunlap, F., Thorbjornsen, T., Keeling, P. and Smith, A.M. (1996) The major form of ADP-glucose pyrophosphorylase in maize endosperm is extra-plastidial. *Plant Physiol.* 112: 779-785.

Dereppe, C., Bold, G.., Ghisalba, O., Ebert, E. and Schar, H.P. (1992) Purification and characterization of dihydrodipicolinate synthase from pea. *Plant Physiol.* 98: 813-821.

Dickinson, C.D., Hussein, E.H.A. and Nielsen, N.C. (1989) Role of posttranslational cleavage in glycinin assembly. *Plant Cell* 1: 459-469.

Doston, S.B., Somers, D.A. and Gengenbach, B.G. (1989) Purification and characterization of lysine-sensitive aspartate kinase from maize. *Plant Physiol.* 91: 1602-1608.

Dry, I., Smith, A., Edwards, A., Bhattacharyya, M., Dunn, P. and Martin, C. (1992) Characterization of cDNAs encoding two isoforms of granule-bound starch synthase which show differential expression in developing storage organs of pea and potato. *Plant J.* 2: 193-202.

Duff, B. (1991) Trends and patterns in Asian rice consumption, In *Rice Grain Marketing and Quality Issues,* selected paper from the International Rice Research Conference 27-31 August, 1990. International Rice Research Institure, Manila, pp1-22.

Edwards, A. (1999) A combined reduction in activity of starch synthases II and III of potato has novel effects on the starch of tubes. *Plant J.* 17:251-261.

Eppendorfer, W.H, Bille, S.W., Prabuddham, P. and Patipanawattana, S. (1983) Fertilizer and water stress effects on yield, mineral composition, protein production and protein quality of tropical foods crops. *Asian Inst. Technol. Bangkok Res.* 157-187.

Falco, S.C., Guida, T., Mauvais, J., Sanders, C., Ward, R.T. and Webber, P. (1995) Transgenic canola and soybean seeds with increased lysine. *Biotechnology* 13: 577-582.

Feng, G.H., Wen, L., Huang, J.K., Shorrosh, B.S., Muthukrishnan, S. and Reeck, G.R. (1990) Nucleotide sequence of a cloned rice genomic DNA fragment that encodes a 10 kDa prolamin polypeptide. *Nucleic Acids Res.* 18: 683.

Fisher, D.K., Boyer, C.D. and Hannah, L.C. (1993) Starch branching enzyme II from

maize endosperm. *Plant Physiol.* 102:1045-1046.

Frankard, V., Ghislain, M. and Jacobs, M. (1992) Two feedback insentitive enzymes of the aspartate pathway in *Nicotiana sylvestris. Plant Pysiol.* 99: 1285-1293.

Frankard, V., Vauterin, M. and Jacobs, M. (1997) Molecular characterization of an *Arabidopsis thaliana* cDNA coding for a monofunctional aspartate kinase. *Plant Mol. Biol.* 34: 233-242.

Frisch, D.A., Tommey, A.M., Gengenbach, B.G. and Somers, D.A. (1991) Direct genetic selection of a maize cDNA for dihydrodipicolinate synthase in an *Escherichia coli dapA*$^-$ auxotroph. *Mol. Gen. Genet.* 228: 287-293.

Fujita, N., Kubo, A., Francisco, P.B.J., Nakakita, M., Harada, K., Minaka, N. and Nakamura, Y. (1999) Purification, characterization, and cDNA structure of isoamylase from developing endosperm of rice. *Planta* 208:283-293.

Fu, Y., Ballicora, M.A. and Preiss. J. (1998) Mutagenesis of the Glucose-1-Phosphate-binding site of potato tuber ADP-Glucose pyrophosphorylase. *Plant physiol.* 117: 989-996.

Fujita, N., Hasegawa, H. and Taira, T. (2002) The isolation and characterization of a waxy mutant of diploid wheat (Triticum monococcum L). *Plant Sci.* 160:595-602.

Fukuzawa, H., Uchimiya, H., Tagawa, M. and Arai, S. (2000) Oryza sativa AK - a gene promoter region and partial ORF. (Unpublished). Only published in Database.

Galili, G., Shaul, O., Perl, A. and Karchi, H. (1994) Synthesis and accumulation of the essential amino acids lysine and threonine in seeds. In *Seed Development and Germination,* pp811-831, Kigel, H. and Galilli, G. (eds.), (New York: Marcel Dekker).

Galili, G. (1995) Regulation of lysine and threonine synthesis. *Plant Cell* 7:899-906.

Gao, M., Wanat, J. Stinard, P.S., James, M.G. and Myers, A.M. (1998) Characterization of *dull 1*, a maize gene coding for a novel starch synthase. *Plant Cell* 10: 399-412.

Gao, Y.F., Jing, Y.X., Shen, S.H., Tian, S.P., Kuang, T.Y. and Sun, S.S.M. (2001) Transfer of Lysine-rich protein gene into rice and production of fertile transgenic plants. *Acta Botamica Sinica* 43: 505-511.

Garibaldi, F. (1974) Rice Parboiling. *Food and Agriculture Organization of the United Nations,* Rome.

Genenbach, B.G., Somers, D.A., Keith, R.A., Muehlbauer, G.J., Sellner, J.M., Bittel, D.C. and Shaver, J.M. (1992) Cellular and molecular genetic regulation of the synthesis of lysine, methionine, and threonine – the aspartate family amino acids. *Biosynthesis and Molecular Regulation of Amino Acids in Plants*. Singh, B.K., Flores, H.E., and Shannon, J.C., (eds.) American Society of Plant Physiologists, pp 195-207.

Ghislain, M., Frankard, V. and Jacobs, M. (1990) Dihydrodipicolinate synthase of *Nicotiana sylvestris*, a chloroplast-localized enzyme of the lysine pathway. *Planta* 180: 480-486.

Ghislain, M., Frankard, V., Vandenbassche, D., Matthews, B.F. and Jacobs, M. (1994) Molecular analysis of the aspartate kinase-homoserine dehydrogenase gene from *Arabidopsis thaliana. Plant Mol. Biol.* 24: 835-851.

Gomez, K.A. (1979) Effect of environment on protein and amylose content of rice. *Proc. Workshop on Chemical Aspects of Rice Grain Quality*, pp56-68. Int. Rice Res. Inst., Los Banos, Languna, Phillippines.

Gomez K.A. and De Datta S.K (1975) Influence of environment on protein content of rice. *Agron.J.* 67:565-568.

Gustafsson, A. (1946) The effect of heterozygosity on variability and vigour. *Hereditas* 32: 263-286.

Habben, J.E. and Larkins, B.A. (1995) Genetic modification of seed proteins. *Current Opinion in Biotechnology* 6: 171-174.

Hannah, L.C., Shaw, J.R., Giroux, M.J., Reyss, A., Prioul, J.L., Bae, J.M. and Lee, L.Y. (2001) Maize genes encoding the small subunit of ADP-glucose pyrophosphorylase. *Plant Physiol.* 127:173-183.

Hannah, L.C. (1997) Starch synthesis in the maize endosperm. In *advances in cellular and molecular biology of plants.* Vol 4. Edited by Larkins, B.A., Vasil I.K. Dordrecht., the Netherlands: Kluwer Academic Publishers: 375-405.

Harn, C., Knight, M., Ramakrishnan, A., Guan, H., Keeling, P.L. and Wasserman, B.P. (1998) Isolation and characterization of the zSSIIa and zSSIIb starch synthase cDNA clones form maize endosperm. *Plant Mol. Biol.* 37:639-649.

Hayakawa, T., Seo, S.W. and Igaue, I. (1980) Electron microscopic observation of rice grain I. Morphology of rice starch. *J. Jpn. Sov. Starch Sci.* 27:173-179.

Hayes, H.K. (1952) Development of the heterosis concept. In: *Heterosis.* pp49-65. Gowen, J.W. (ed.). Iowa State College Press. Ames.

Hibino, T., Kidzu, T., Masumura, T., Ohtsubo, K., Tanaka, K., Kawabata, K. and Fujii, S. (1989) Amino acid composition of rice prolamin polypeptides. *Agri. Biol. Chem.* 53:313-318.

Hille, R. (1972) Heritability of protein content and its relationship to agronomic characters in rice. *Proc 14th Rice Tech Working Group,* Univ. California, Davis.

Hinnebusch, A.G. (1988) Mechanisms of gene regulation in the general control of amino acid biosynthesis in *Saccharomyces cerevisiae. Microbiol. Rev.* 52: 248-273.

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. *Mol. Biol. Evol.* 15: 978-987.

Hizukuri, S., Kaneko, T. and Takeda, Y. (1983a) Measurement of the chain length of amylopectin and its relevance to the origin of crystalline polymorphism of starch granules. *Biochim. Biophy. Acta* 760: 188-191.

Hizukuri, S., Shurasaka, K. and Juliano, B.O. (1983b) Phosphorous and amylose branching in rice starch granules. *Staerke* 35: 348-350.

Hoffman, L.M., Donaldson, D.D., Bookland, R., Rashka, K. and Herman, E.R. (1987) Synthesis and protein body deposition of maize 15-kDa zein in transgenic tobacco seeds. *EMBO J.* 6: 3213-3221.

Horikoshi, M. and Morita, Y. (1982) Changes in ultrastructure and subunit composition of protein body in rice endosperm during germination. *Agric. Biol. Chem.* 46: 269-274.

Hoshikawa, K. (1968) Studies on the development of endosperm in rice II. Development of starch granules in endosperm tissue. *Nippon Sakumotsu Gakkai Kiji* 37:207-216.

Huang, C.W. and Huang, Y.S. (1990) Preliminary studies of free amino acid of milled rice in relation to taste of cooked rice in *indica* varieties. *穀類作物品質性狀遺傳研究進展* pp58-63. 江蘇科學技術出版社. 南京.

Hull, F.H. (1945) Recurrent selection for specific combining ability in corn. *J. Am. Soc. Agron.* 37: 134-145.

Ikehashi, H. (1982) Prospects for overcoming barriers in the utilization of *indica-japonica* crosses in rice breeding. *Oryza* 19: 69-77.

Iida, S., Amano, E. and Nishio, T. (1993) A rice (Oryza sativa L.) mutant having a low content of glutelin and a high content of prolamine. *Theor. Appl. Genet.* 87: 374-378.

IRRI. (1985) Annual report for 1984. *Int. Rice Res. Inst.*, Los Baños, Laguna, Phillippines.

IRRI (1995) IRRI Rice Facts, *International Rice Research Institute*, Los Baños, Laguna, Phillippines.

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 ha a GT to TT mutation at the 5' splice site of the first intron. *Plant J.* 1: 133-138.

Izumi, H., Adachi, T., Fujii, N, Matsuda, T., Nakamura, R. and Tanaka, K. (1992) Clone encoding a major allergenic protein in rice seeds-homology of the deduced amino acid sequence with members of α-amylase/trypsin inhibitor family. *FEBS Lett.* 302: 213-216.

James, M.G., Denyer, K. and Myers, A. (2003) Starch synthesis in the cereal endosperm. *Current opinion in Plant Biology.* 6: 215-222.

Jiang, S., Wan, L. and Xu, L. (2002) New rice glutelin gene GluB-4. *Rice Research Institute, Najing Agricultural University*, Weigang, Nanjing, China.

Jones, D.F. (1917) Dominance of linked factors as a means of accounting for heterosis. *Genetics* 2: 466-479.

Jones, H.A. and Clarke, A.E. (1943) Inheritance of male sterility in the onion and the production of hybrid seed. *Proc. Am. Soc. Hort. Sci.* 43: 189-194.

Juniano, B.O., Antonio, A.A. and Esmama, B.V. (1972) Variability in protein content of rice. *Kalikasan* 1: 74-81.

Juliano, B.O. (1979). The chemical basis of rice grain quality. in: *Proc.Workshop on Chemical aspects of rice grain quality*. pp 69-90. Int. Rice Res. Inst., Los Banos, Laguna, Philippines.

Juliano, B.O. (1985) *Rice: Chemistry and Technology*. American Association of Cereal Chemists, St Paul, Minnesota, pp1-174.

Kaneko, T., Hashimoto, T., Kumpaisal, R. and Yamada, Y. (1990) Molecular cloning of wheat dihydrodipicolinate synthase. *J. Biol. Chem.* 265: 17451-17455.

Karchi, H., Shaul, O. and Galili, G. (1993) Seed specific expression of a bacterial desensitized aspartate kinase increases the production of seed threonine and methionine in transgenic tobacco. *Plant J.* 3: 721-727.

Karchi, H., Shaul, O. and Galili, G. (1994) Lysine synthesis and catabolism are coordinately regulated during tobacco seed development. *Proc. Nat. Acad. Sci. (USA)* 91: 2577-2581.

Kato, H., et al. (1990) Temperature response and inheritance of a thermosensitive genic male sterility in rice. *Japan J. Breed.* 40 (suppl. 1): 352-369.

Kim, S.Y. and Wu, R. (1990) Multiple protein factors bind to a rice glutelin promoter region. *Nucleic Acids Res.* 18(23): 6845-6852.

Kim, W.T. and Okita, T.W. (1988a) Nucleotide and primary sequence of a major rice prolamin. *FEBS Lett.* 18: 6845-6852.

Kim, W.T. and Okita, T.W. (1988b) Structure, expression, and heterogeneity of the rice seed prolamins. *Plant Physiol.* 88: 646-655.

Kim, W.T., Li, X. and Okita, T.W. (1993) Expression of storage protein multigene families in developing rice. *Plant Cell Physiol.* 34: 595-603.

Kiyota, S. and Sakano, K. (1996) Cloning and characterization of aspartate kinase from rice (Oryza sativa). (Unpublished), only published in Database.

Knudsen D., Clark R.B, Denning J.L. and Pier P.A. (1981) Plant analysis of trace elements by X-ray. *J. Plant Nutr.* 3: 61-75.

Konishi, Y., Nojima, H., Okuno, K., Asaoka, M. and Fuwa, H. (1985) Characterization of starch granules from waxy, non-waxy and hybrid seeds of Amaranths Hpochrandriaces L. *Agric, Biol. Chem.* 49: 1965-1971.

Körnicke, F.U.W. (1885) Handbuch des Getreidebaues, V. 2. Die Sorten und der Anbau des Getreides, Berlin, Germany.

Krishnan, H.B., Franceschi, V.R. and Okita, T.W. (1986) Immunochemical studies on the role of the Golgi complex in protein-body formation in rice seeds. *Planta* 169: 471-480.

Kubo, A., Fujita, N., Harada, K., Matsuda, T., Satoh, H. and Nakamura, Y. (1999) The starch-debranching enzymes isoamylase and pullulanase are both involved in amylopectin biosynthesis in rice endosperm. *Plant Physiol.* 121:399-409.

Kumamura, T., Satoh, H., Iwata, N., Ogawa, M. and Tanaka, K. (1988) Mutants for rice storage proteins. I. Screening of mutants for rice storage proteins of protein bodies in the starchy endosperm. *Theor. Appl. Genet.* 76: 11-16.

Lee, P.D. and Su, J. C. (1982) Sucrose-starch transforming system in rice grains-A trancer feeding study. *Proc. Natl. Sci. Counc. Rep.* China B6: 189-196.

Li, P., Wang, Y. and Liu, H. (1990) Physiological bases of high yielding heterosis in *indica*-type F1 hybrid rice. *Sci. Agro. Sinica*: 39-44.

Li, Z.B., Xiao, Y.H., Zhu, Y.G., Li, R.Q., Liu, C.L. and Wan, J.M. (1982) Study and practice of hybrid rice. *Shanghai Sci. Tech. Press*, Shanghai, China.

Li, X., Gu, M.H. and Pan, X.B. (1989) The effects of environmental factors on rice qualities during grain filling. *江蘇農學院學報* 10: 7-12.

Li, X. and Okita, T.W. (1993) Accumulation of prolamines and glutelins during rice seed development: a quantitative evaluation. *Plant Cell Physiol.* 34: 385-390.

Li, X., Franceschi, V.R. and Okita, T.W. (1993a) Segregation of storage protein mRNAs on the rough endosplasmic retivulum membranes of rice endosperm cells. *Cell* 72: 869-879.

Li, X., Wu, Y., Zhang, D.Z., Gallikin, J.W., Franceschi, V.R. and Okita, T.W. (1993b) Rice prolamin protein body biogenesis: a BiP mediated process. *Science* 242: 1054-1056.

Limas, G.G., Salwas, M., Monel, I., Fischer, S., Whittman-Liebold, B. and Mendez, E. (1990) Purificatio and characterizationof ten new rice NaCl-soluble proteins: identification of four protein synthesis inhibitors and two immunoglobulin-binding proteins. *Planta* 181: 1-9.

Lin, S.C. and Yuan, L.P. (1980) Hybrid rice breeding in China. In: *Innovative approaches to rice breeding*. pp33-51. Int. Rice. Res. Inst., Manila, Philippines.

Liu, Q.Q. (2002) Genetically Engineering rice for increased lysine. Doctoral Dissertation. Yanzhou University. Yangzhou, Jiangsu.

Lloyd, J.R. (1999) Simultaneous antisense inhibition of two starch synthase isoforms in potato tuber leads to accumulation of grossly modified amylopectin. *Biochem. J.* 338: 515-521.

Lu, Z.T., Tang, S.Q., Xun, Z.M., Min, S.K. and Sheng, Z.D. (1988) Analyses of yield components in conventional and hybrid rice. *Zhejiang Agri. Sci.* 4: 156-158.

MacDonald, F.D. and Preiss, J. (1983) Solubilization of the starch granule-bound starch synthase of normal maize kernels. *Plant Physiol.* 73:175-178.

MacDonald, F.D. and Preiss, J. (1985) Partials purification and characterization of granule-bound starch synthase from normal and waxy maize. *Plant Physiol.* 78: 849-852.

Masumara, T., Shibata, D., Hibino, T., Kato, T., Kawabe, K., Takeba G., Tanaka, K. and Fujii, S. (1989) cDNA cloning of an mRNA encoding a sulfur-rich 10 kDa prolamin polypeptide in rice seeds. *Plant Mol. Biol.* 12: 123-130.

Masumura, T., Hibino. T, Kidzu, K., Mitsukawa, N, Tanaka, K. and Fujii, S. (1990) Cloning and characterization of a cDNA encoding a rice 13 kDa prolamin. *Mol. Gen. Genet.* 221: 1-7.

Mattews, B.F. and Hughes, C.A. (1993) Nutritional improvement of the aspartate family of amino acid in edible crop plants. *Amino acids* 4: 21-34.

Mauri, I., Massimo, M., Lohmer, S., Motto, M., Salamini, F., Thompson, R. and Martegani, E. (1993) Functional expression of the transcriptional activator Opaque-2 of *Zea mays* in transformed yeast. *Mol. Gen. Genet.* 241: 319-326.

Mizuno, K., Kawasaki, T., Shimada, H., Satoh, H., Kobayashi, E., Okumura, S., Arai, Y. and Baba, T. (1993) Alteration of the structural properties of starch components by the lack of an isoform of starch branching enzyme in rice seeds. *J. Biol. Chem.* 268: 19084-19091.

Mizuno, K., Kimura, K., Arai, Y., Kawasaki. T., Shimada, H. and Baba, T. (1992) Starch branching enzymes from immature rice seeds. *J. Biochem. (Tokyo)* 112: 643-651.

Mizuno, K., Kobayashi, E., Tachibana, M., Kawasaki, T., Fujimura, T., Funane, K., Kobayashi, M. and Baba, T. (2001) Characterization of an isoform of rice starch branching enzyme, REB4, in developing seeds. *Plant Cell Physiol.* 42: 349-357.

Moons, A., Gielen, J., Vandekerckhove, J., Van der Straeten, D., Gheysen, G. and Van Montagu, M. (1997) An abscisic-acid- and salt-stress-responsive rice cDNA from a novel plant gene family. *Planta* 202: 443-454.

Morinaga, T. and Kuriyama, H. (1958) Intermediate type of rice in the subcontinent of India and Java. *Japan J. Breed.* 8: 253-259.

Muehlbauer, G.J., Somers, D.A., Matthews, B.F. and Gengenbach, B.G. (1995) Molecular genetics of the maize *(Zea mays L.)* aspartate kinase-homoserine dehydrogenaase gene family. *Plant Physiol.* 106: 1303-1312.

Müller-Rober B., Sonnewald, U. and Willmitzer, L. (1992) Inhibition of the AGPase transgenic potatoes leads to sugar-sorting tubers and influences tuber formation and expression of tuber storage protein genes. *EMBO J.* 11: 1229-1238.

Müller, M. and Knudsen, S. (1993) The nitrogen response of a barley C-hordein promoter is controlled by positive and negative regulation of the GCN4 and endosperm box. *Plant J.* 4: 343-355.

Murata, T., Dugiyama, T. and Akazawa, T. (1964b) Enzymic mechanism of starch synthesis in ripening rice grains. II. Adenosine diphosphate glucose pathway. *Arch. Biochem. Biophys.* 107: 92-101.

Müntz, K. (1998) Deposition of storage proteins. *Plant Mol. Biol.* 38: 77-99.

Myers, A.M., Morell, M.K., James, M.G. and Ball, S.G. (2000) Recent progress toward understanding biosynthesis of the amylopectin crystal. *Plant Physiol.* 122: 989-997.

Nakamura, Y. and Yamanouchi, H. (1992) Purification of two forms of starch branching enzyme (Q-enzyme) from developing rice endosperm. *Plant Physiol.* 99:1265-1266.

Nakamura, Y. (1996) Some properties of starch debranching enzymes and their possible role in amylopectin biosynthesis. *Plant Sci.* 121: 1-18.

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

Nakase, M., Yamada, T., Kira, T., Yamaguchi, J., Aoki, N., Nakamura, R., Matsuda, T.

and Adachi, T. (1996a) The same nuclear proteins bind to the 5'-flanking regions of genes for the rice seed storage protein: 16 kDa albumin, 13 kDa prolamin, and typeII glutelin. *Plant Mol. Biol.* 32: 621-630.

Nakase, M., Hotta, H., Adachi, T., Aoki, N., Nakamura, T., Masumura, T., Tanaka, K. and Matsuda, T. (1996b) Cloning of the rice seed alpha-globulin-encoding gene: sequence similarity of the 5'-flanking region to those of the genes encoding wheat high-molecular weight glutelin and barley D hordein. *Gene* 170: 223-226.

Negrutiu, I., Cattoir-Reynearts, A., Verbruggen, I. and Jacobs, M. (1984) Lysine overproducer mutants with an altered dihydropicolinate synthae fromprotoplast culture of *Nicotiana sylvestris.* (Spegazzini and Comes). *Theor. Appl. Genet.* 68: 11-20.

Nikuni, Z., Hizukuri, S., Kumagai, K., Hasegawa, H., Moriwaki, T., Fukui, T., Doi, K., Nara, S. and Maeda, I. (1969) The effect of temperature during the maturation period on the physico-chemical properties of potato and rice starches. *Mem. Inst. Sci. Ind. Res.* Osaka Univ. 26:1-27.

Nishi, A., Nakamura, Y., Tanaka, N. and Satoh, H. (2001) Biochemical and genetic analysis of the effects of *amylose-extender* mutation on rice endosperm. *Plant Physiol.* 127: 459-472.

O'Farrell, P.H. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250: 4007-4021.

Ogawa, M., Kumamaru, T., Satoh, H., Iwata, N., Omura, T., Kasai, Z. and Tanaka, K. (1987) Purification of protein body-I of rice seed and its polypeptide composition. *Plant Cell Physiol.* 28: 1517-1527.

Ogawa, M., Kumamaru, T., Satoh, H., Omura, T., Park, T., Shintaku, K. and Baba, K. (1989) Mutants for rice storage protein 2. Isolation and characterization of protein bodies from rice mutants. *Theor. Appl. Genet.* 78: 305-310.

Oka, H. I. (1983) The *indica-japonica* differentiation of rice cultivars: A review. *SABRAO J.* Malaysia.

Okita, T.W., Hwang, Y.S., Hnilo, J., Kim, W.T., Aryan, A.P., Larson, R. and Krishnan,

H.B. (1989) Structure and expression of the rice glutelin multigene family. *J. Biol. Chem.* 264: 12573-12581.

Okita, T.W., Li, X. and Roberts, M.W. (1994) Targeting of mRNAs to domains of the endoplasmic reticulum. *Trends Cell Biol.* 4: 91-96.

Perez, C.M., Perdon, A.A., Resurreccion, A.P., Villareal, R.M. and Juliano, B.O. (1975) Enzymes of carbohydrate metabolism in the developing rice grain. *Plant Physiol.* 56: 579-583.

Pingali, P. L., Hossain, M. and Gerpacio, R.V. (1997) Asian rice bowls: The returning crisis? pp208-233. *CAB Int. and Int. Rice Res. Inst.*, U.K.

Piryns, I., Vernaillens, S. and Jacobs, M. (1988) Inhibitory effects of aspartate derived amino acids and aminoethylcysteine, a lysine analog, on the growth of sorghum seedlings: relation with three enzymes of the aspartate pathway. *Plant Sci.* 57: 93-101.

Preiss, J. et al. (1991) Starch biosynthesis and its regulation. *Biochem. Soc. Trans.* 19: 539-547.

Rahman, A., Wong, K.S., Jane, J.L., Myers, A.M. and James, M.G. (1998) Characterization of SU1 isoamylase, a determinant of storage starch structure in maize. *Plant Physiol.* 117: 425-435.

Ressell, P.L. and Juliano, B.O. (1983) Differential scanning calorimetry of rice starches. *Staerke* 35: 382-386.

Resurreccion, A. P., Hara, T., Juliano, B.O. and Yoshida, S. (1977) Effect of temperature during ripening on grain quality of rice. *Soil. Sci. Plant Nutr.* Tokoy 23: 109-112.

Revves, C.D., Krishna, H.B. and Okita, T.W. (1986) Gene expression in developing wheat endosperm: accumulation of gliadin and ADP-glucose pyrophosphorylase messenger RNAs and polypeptides. *Plant Physiol.* 82: 34-40.

Rey, P., Pruvot, G., Becuwe, N., Eymery, F., Rumeau, D. and Peltier, G. (1998) A novel thioredoxin-like protein located in the chloroplast is induced by water deficit in *Solanum tuberosum L.* plants. *Plant J.* 13: 97-107.

Riccardi, F., Gazeau, P., de Vienne, D. and Zivy, M. (1998) Protein changes in response to progressive water deficit in maize. Quantitative variation and polypeptide identification. *Plant Physiol.* 117: 1253-1263.

Rognes, S.E., Lea, P.J. and Miflin, B.J. (1980) S-Adenosyl methionine – A novel regulator of aspartate kinase. *Nature* 287: 357-359.

Rognes, S.E., Bright, S.W.J. and Miflin, B.J. (1983) Feedack insensitive aspartate kinase isoenzymes in barley mutants resistant to lysine plus threonine. *Planta* 157: 32-38.

Salehuzzaman, S.N.I.M. (1993) Isolation and characterization of a cDNA encoding granule-bound starch synthesis in cassava and its antisense expression in potato. *Plant Mol. Biol.* 23: 947-967.

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

Schaeffer, G.W. and Sharpe, F.T.Jr (1990) Modification of amino acid composition of endosperm proteins from in vitro-seleted high lysine mutants in rice. *Theor. Appl. Genet.* 80: 841-846.

Sha, S., Sugiyama, Y., Mitsukawa, N., Masumura, T. and Tanaka, K. (1996) Cloning and swquencing of a rice gene encoding the 13-kDa prolamin polypeptide. *Biosci. Biotechnol. Biochem.* 60: 335-337.

Shannon, J.C., Pien, F.M., Cao, H. and Liu, K.C. (1998) Brittle-1, an adenylate translocator, facilitates transfer of extraplastidial synthesized ADP-glucose into amyloplasts of Maize endosperms. *Plant physiol.* 117:1235-1252.

Shaul, O. and Galili, G. (1992) Increased lysine synthesis in tobacco plants that express high levels of bacterial dihydrodipicolinate synthase in their chloroplasts. *Plant J.* 2: 203-209.

Shen, J.H. (1980) Rice breeding in China. In: *Rice improvement in China and other Asian countries*. Int. Rice Res. Inst., Manila, Philippines.

Sheng, X.B. (1987) Genetic effects of cytoplasm on agronomic characteristics in hybrid rice. *Chinese J. Rice Sci.* 3: 155-170.

Shewry, P.R., Napier, J.A. and Tatham, A.S. (1995) Seed storage proteins: structures and biosynthesis. *Plant Cell* 7: 945-956.

Shibisaki, M., Suzuki, S., Nemoto, H. and Kuroume, T. (1979) Alllergenicity and lymphocyte-stimulating property of rice protein. *J. allergy. Clin. Immunol.* 64: 259-265.

Shi, C.H, Xue, J.M., Yu, Y.G., Yang, X.E. and Zhu, J. (1996) Analysis of genetic affects for nutrient quality traits in *indica* rice. *Theor. Appl. Genet.* 92: 1099-1102.

Shorrosh, B.S., Wen, L., Zen, K.C., Huang, J.K., Pan, J.S., Hermodson, M.A, Tanaka, K., Muthukrishnan, S. and Reeck, G.R. (1992) A novel cereal storage protein: molecular genetics of the 19 kDa globulin of rice. *Plant Mol. Biol.* 18: 151-154.

Shull, G.H. (1908) The composition of a field of maize. *Am. Breed. Assoc. Rep.* 4: 296-301.

Shull, G. H. (1952) Beginnings of the heterosis concept. In: *Heterosis*. pp14-48. J.W.Gowen (ed.). Iowa State College Press, Ames.

Shyur, L.F. and Chen, C.S. (1990) Nucleotide sequence of two rice prolamin cDNAs. *Nucleic Acids Res.* 18: 6683.

Shyur, L.F., Wen, T.N. and Chen, C.S. (1992) cDNA clonig and gene expression of the major prolamins of rice. *Plant Mol. Biol.* 20: 323-326.

Singh, R. (1997) Studies on heterosis for certain physiological and biochemical parameters in rice. In: *The genetics and exploitation of heterosis in crops.* pp164-165. Coors, J.G., Pandey, S., Grinkel, M.V. Hallauer, D.C et al. (ed.), Proc. Int. CIMMYT symp. 17-22 August, 1997. ASA, CSSA, SSSA, Madison, WI, CIMMYT. Mexico (Abstr.).

Singh, R. and Juliano, B.O. (1977) Free sugars in relation to starch accumulation in developing rice grain. *Plant physiol.* 59: 417-421.

Slattery, C.J., Kavakli, I.H. and Okita, T.W. (2000) Engineering starch for increased quantity and quality. *Trends Plant Sci.* 5: 291-298.

Smidansky, E.D., Clancy, M., Meyer, F.D., Lanning, S.P., Blake, N.K., Talbert, L.E. and Giroux, M.J. (2002) Enhanced ADP-glucose pyrophosphorylase activity in wheat endosperm increases seed yield. *PNAS* 99: 1724-1729.

Smith, A.M. (1990) Evidence that the waxy protein of pea in not the major starch granule-bound starch synthase. *Planta* 182:599-604.

Smith, A.M., Denyer, K. and Martin, C. (1997) The synthesis of starch granule. *Annu. Rev. Plant Physiol. Plant Mol.Biol.* 48: 67-87.

Song, X.F., Xiong, Z.M. and Min, S.K. (1990) Biological characteristics of competitive heterosis in *indica/japonica* intersubspecific rice hybrids. In: *Selected papers on the development of P(T)GMS rice lines and utilization of intersubspecific heterosis*. pp129-134. China's Ministry Agr., Beijing, China.

Sorensen, M.B., Cameron-Mills, V. and Brandt. A. (1989) Transcriptional and post-transcriptional regulation of gene expression in developing barley endosperm. *Mol. Gen. Genet.* 217: 195-201.

Stark D.M., David, M. and Timmerman, K.P. (1992) Regulation of the amount of starch in plant tissues by ADP-glucose pyrophosphorylase. *Science* 258: 287-292.

Su, P., Yu, S.M. and Chen, C.S. (2001) Spatial and temporal expression of a rice prolamin gene RP5 promoter in transgenic tobacco and rice. *J. Plant Physiol.* 158: 247-254.

Suh, M.C., Cho, S.J., Joung, H., Liu, J.R., Takaiwa, F. and Lim. Y.P. (2000) Isolation and characterization of a rice genomic clone for dihydrodipicolinate synthase. (unpublished) only published in NCBI database.

Sullivan, T.D. (1995) The maize brittle 1 gene encodes amyloplast membrane polypeptides. *Planta* 196:477-484.

Sun, S.S.M., Zuo, W.N. and Tu, H.M. (1992) Molecular approaches for enhancing amino acid quality of plant proteins. In *Biosynthesis and Molecular Regulation of*

Amino Acids in Plants (Singh, B.K., Flore, H.E., and Shannon, J.C.eds), American Society of Plant Physiology. pp 208-126.

Sun, S.S.M. and Larkins, B.A. (1993) Transgenic plants for improving seed storage proteins. In *Transgenic Plants vol.1*, Kung, S.D., and Wu, R. ed., Academic Press, pp339-371.

Sun, Z.X. and Cheng, S.H. (1994) Hybrid rice breeding: from three-line system, two-line system to one-line system. *China Agri. Sci. Tech. Press.* Beijing, China.

Suzuki, H. and Murayama, N. (1967) Effect of temperature on the rice grains and rice starches. *Int. Rice Comm. Newsl. (spec.issu,)* pp82-92.

Takaiwa, F., Ebinuma, H., Kikuchi, S. and Oono, K. (1987a) Nucleotide sequence of a rice glutelin gene. *FEBS Lett.* 206: 33-35.

Takaiwa, F., Kikuchi, S. and Oono, K. (1987b) A rice glutelin family- a major type of glutelin mRNAs can be divided into two classes. *Mol. Gen. Genet* 208: 15-22.

Takaiwa, F. and Oono, K. (1991) Genomic DNA sequences of two new genes for new storage protein glutelin in rice. *Jpn. J. Genet.* 66: 161-171.

Takaiwa, F., Oono, K., Wing, D. and Kato, A. (1991a) Sequence of three members and expression of a new major subfamily of glutelin genes from rice. *Plant Mol. Biol.* 17: 875-885.

Takaiwa, F., Oono, K. and Kato, A. (1991b) Analysis of 5'-flanking region responsible for the endosperm-specific expression of a rice glutelin chimeric gene in transgenic tobacco. *Plant Mol. Biol.* 16: 49-58.

Tanaka, K., Sugimoto. T., Ogawa, M. and Kasai, Z. (1980) Isolation and characterization of two types of protein bodies in the rice endosperm. *Agric. Biol. Chem.* 44: 1633-1639.

Tanaka, Y., Hayashida, S. and Hongo, M. (1975) The relationship of the feces protein particles to rice protein bodies. *Agri. Biol. Chem.* 39: 515-518.

Tanaka, Y., Resurreccion, A.P., Juliano, B.O. and Bechtel, D.B. (1978) Properties of

whole and undigested fraction of protein bodies of milled rice. *Agric. Biol. Chem.* 42: 2015-2023.

Thorbjornsen, T., Villand, P., Kleczkowski, L.A. and Olsen, O.A. (1996) A single gene encodes two different transcriots for the ADP-glucose pyrophosphorylase small subunit from barley (*Hordeum vulgare*). *Biochem J.* 313: 149-154.

Tian, C., Cheng, X. and Liang, Z. (1980) Several views on popularization of Xian (*indica*) hybrid rice. *Yunan Agri. Sci. Tech.* 2: 12-18.

Tsai, C.Y. (1974) The function of the Waxy locus in starch synthesis in maize endosperm. *Biochem Genet.* 11: 83-96.

Touzet, P., Morin, C., Damerval, C., Le Guilloux, M., Zivy, M. and de Vienne, D. (1995) Charaterizing allelic proteins for genome mapping in maize. *Electrophoresis* 16: 1289-1294.

Tu, H.M., Godfrey, L.W. and Sun, S.S.M. (1998) Expression of the Brazil nut methionine-rich protein and mutants with increased methionine in transgenic potato. *Plant Mol. Biol.* 37: 829-838.

Umemoto, T., Nakamura, Y. and Ishikura, N. (1995) Activity of starch synthase and the amylose content in rice endosperm. *Phytochemistry* 40: 1613-1616.

Van der Leij, F.R., et al. (1991) Complementation of the amylose-free starch mutant of potato (*Solanum tuberosum*) by the gene encoding granule-bound starch synthase. *Theor. Appl. Genet.* 82: 289-295.

de Vienne, D., Burstin, J., Gerber, S., Leonardi, A., Le Guilloux, M., Murigneux, A., Bechert, M., Bahrman, N., Damerval, C. and Zivy, M. (1996) Two-dimensional electrophoresis of proteins as a source of monogenetic and codominant markers for population genetics and mapping in the expressed genome. *Heredity* 76: 166-177.

de Vienne, D., Leonardi, A., Damerva, C. and Zivy, M. (1999) Genetics of proteome variation for QTL characterization: application to drought stress responses in maize. *J. Exp. Bot.* 50: 303-309.

Visser, R.G.F., et al. (1991) Inhibition of the expression of the gene for granule-bound

starch synthase in potato by antisense constructs. *Mol. Gen. Gent.* 225: 289-296.

Walling, L., Drews, G.N. and Goldberg, R.B. (1986) Transcriptional and post-transcriptional regulation of soybean seed protein mRNA levels. *Proc. Natl. Acad. Sci. USA* 83: 2123-2127.

Wang, Z.Y., Zheng, F.Q., Shen, G.Z., Gao, J.P., Snustad, 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. *Plant J.* 7: 613-622.

Wang, Y.J., White, P., Pollack, L. and Jane, J.L. (1993) Characterization of starch structures of 17 maize endosperm mutant genotypes with Oh43 inbred line background. *Cereal Chem.* 70: 171-179.

Washida, H., Wu, C.Y., Suzuki, A., Yamanouchi, U., Akihama, T., Harada, K. and Takaiwa, F. (1999) Identification of cis-regulatory elements required for endosperm expression of the rice storage protein glutelin gene GluB-1. *Plant Mol. Biol.* 40: 1-12.

Wen, T.N. and Luthe, D.S. (1985) Biochemical characterization of rice glutelin. *Plant Physiol.* 78: 172-177.

Wen, T.N., Shyur, L. F., Su, J.C. and Chen, C.S. (1993) Nuceotide sequence of a rice (*Oryza sativa*) prolamin storage protein gene, RP6. *Plant Physiol.* 101: 1115-1116.

WHO. 1973. Energy and protein requirements. WHO Tech. Rep. Ser. 522. World Health Org., Gnenva.

Wijk, K.J.V. (2001) Challenges and prospects of plant proteomics. *Plant Physiol.* 126: 501-508.

Wilson, B.J., Gray, A.A. and Matthews, B.F. (1991) Bifunctional protein in carrot contains both aspartate kinase and homoserine dehydrogenase activities. *Plant Physiol.* 97: 1323-1328.

Wu, H.K. and Chen, J.T. (1978) Protein bodies in developing rice grain. *Proc. Natl. Sci. Council* TaiWan 2:281-192.

Wu, C.Y., Washida, H., Onodera, Y., Harada, K. and Takaiwa, F. (2000) Quantitative

nature of the prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal cis-element requirements for endosperm-specific gene expression. *Plant J.* 23: 415-421.

Wu, C.Y., Adachi, T., Hatano, T., Washida, H., Suzuki, A. and Takaiwa, F. (1988a) Promoters of rice seed storage protein genes direct endosperm-specific gene expression in transgenic rice. *Plant Cell Physiol.* 39: 885-889.

Wu, C.Y., Suzuki, A., Washida, H. and Takaiwa, F. (1988b) The GCN4 motif in a rice glutelin gene is essential for endosperm-specific gene expression and is activated by Opaque-2 in transgenic rice plants. *Plant J.* 14: 673-683.

Wu, P., Jiang, H.W. and Dian, W.M. (2001) Direct Submission to NCBI genbank internet. Institute of Plant Science, College of Life Sciences, Zhejiang University, Kuaixuan Road 268, Hangzhou, Zhejiang Province 310029, P.R. China.

Yab, T.C. and Chang, S.C. (1976) Heterosis of rice hybrids under wetland and dryland conditions. *SABRAO J.* 1: 35-40.

Yamagata, H., Sugimoto, T., Tanaka, K. and Kasai. Z. (1982) Biosynthesis of storage proteins in developing rice seeds. *Plant Physiol.* 70:1094-1100.

Yamagata, H. and Tanaka, K. (1986) The site of synthesis and accumulation of rice storage proteins. *Plant Cell Physiol.* 27: 135-145.

Yamagata, H., Nomura, T., Arai, S., Tanaka, K. and Iwasaki, T. (1992) Nucleotide sequence of a cDNA that encodes a rice prolamin, *Biosci. Biotechno. Biochem.* 56:1537.

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

Yang, C.Y, Chen, H.J., Chen, Y.F., Chen, M.H. and Chen, C.S. (2000) Seed-specific expression of a rice prolamin RP3/GUS chimeric gene in nicotiana tabacum. Cv. Wisconsin 38. *Plant Physiol.* 156: 100-105.

Yoshihara, T. and Takaiwa, F. (1996) Cis-regulatory elements responsible for quantitative regulation of the rice seed storage protein glutelin GluA-3 gene. *Plant*

Cell Physiol. 37: 107-111.

Yoshihara, T., Washida, H. and Takaiwa, F. (1996) A 45-bp proximal region containing AACA and GCN4 motif is sufficient to confer endosperm-specific expression of the rice storage protein glutelin gene, GluA-3. *FEBS Lett.* 383: 213-218.

Yu, S.B., Li, J.X., Xu, C.G., Tan, Y.F., Gao, Y.J., Li, X.H., Zhang, Q.F. and Maroof. M.A. (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *PNAS* 94: 9226-9231.

Yuan, L.P. (1966) A preliminary report on male sterility in rice. *Sci. Bul.* 4: 32-34.

Yuan, L.P. (1985) A concise course in hybrid rice. *Hunan Sci. Tech. Press*, Changsha, China.

Yuan, L.P. (1987) The strategy on hybrid rice breeding. *Hybrid Rice* 1: 1-3.

Yuan, L.P. and Chen, H.X. (1988) Breeding and cultivation of hybrid rice. *Hunan Sci. Tech. Press,* Changsha, China.

Yuan, L.P. and Virmani, S.S. (1988) Status of hybrid rice research and development. In: *Hybrid rice* (Proc. Int. Symp. On Hybrid Rice, October 6-10, 1986, Changsha, Chnia) p 7-24. Int. Rice Res. Inst., Manila, Philippines.

Yuan, L.P. (1991a) Breeding for intersubspecific heterosis in rice. Program Report. China National Hybrid Rice Res. Develop. Center.

Yuan, L.P. (1991b) Outlook on the development of hybrid rice breeding. In: *Prospects of rice farming for 2000*, pp205-211. Zhejiang Publ. House Sci. Tech., Hangzhou, China.

Yuan, L.P. (1994d) Increasing yield potential in rice by exploitation of heterosis. In: *Hybrid rice: new developments and future prospects*. pp1-6. Virmani, S.S. (ed.) Int. Rice Res. Inst., Manila, Philippines.

Yuan, L.P. (1997a) Exploiting crop heterosis b y two-line system hybrids: Current status and future prospects. In: *Proc. Int. Symp. On Two-line System Heterosis Breeding in Crops*. pp1-7. Changsha, Hunan, China, September 6-8, 1997.

Zeeman, S.C. and Rees, T. A. (1999) Changes in carbohydrate metabolism and

assimilate export in starch-excess mutants of *Arabidopsis*. *Plant Cell Environ.* 22:1445-1453.

Zeng, S.J., Liang, C.Y., Liang, J.K. and Chen, B.Y. (1997) Studies on the heterosis of two-line hybrid rice. I. Performance of main agronomic charaters. *J. Trop. Subtrop. Bot.* 4: 56-61.

Zhao, W.M., Gatehouse, J.A. and Boulter, D. (1983) The purification and partial amlino acid sequence of a polypeptide from the glutelin fraction of rice grains: homology to pea legumin. *FEBS Lett.* 162: 96-102.

Zhao, Y., Leisy, D.J. and Okita, T.W. (1993) Tissue-specific expression and temporal regulation of the rice glutelin Gt3 gene are conferred by at least two spatially separated cis-regulatory elements. *Plant Mol. Biol.* 25: 429-436.

Zheng, Z., Kawagoe, Y., Xiao, S., Li, Z., Okita, T., Hau, T.L., Lin, A. and Murai, N. (1993) 5' distal and proximal cis-acting regulator elements are required for developmental control of a rice seed storage protein glutelin genes. *Plant J.* 4: 357-366.

Zivy, M. and Vienne, D.D. (2000) Proteomics: a link between genomics, genetics and physiology. *Plant Mol. Biol.* 44: 575-580.