Provitamin A Biofortified Rice Event GR2E

Application for Amendment to Standard 1.5.2 — Food Produced Using Gene Technology

---

SUBMITTED TO
Food Standards Australia New Zealand
Ground Floor, Boeing House
55 Blackall Street
BARTON ACT 2600

SUBMITTING ORGANIZATION
International Rice Research Institute
Pili Drive, UPLB, Los Baños, 4031, Laguna
Philippines

CONTACT INFORMATION
Donald J. MacKenzie, Ph.D.
Regulatory Affairs and Stewardship Leader
[redacted]

SUBMISSION DATE
November 11, 2016

Application Checklist

Requirement	Section in this Application	Page Number
General Requirements (Application Handbook Section 3.1)		
3.1.1 Form of application ☑ Executive summary ☑ Relevant sections of Part 3 identified ☑ Pages sequentially numbered ☑ Electronic plus two hard copies ☑ Electronic and hard copies identical ☑ Hard copies capable of being laid flat ☑ Relevant references provided		3
3.1.2 Applicant details	1.1	17
3.1.3 Purpose of the application	1.2	17
3.1.4 Justification for the application	1.3	18
3.1.5 Information to support the application	A–D	20–89
3.1.6 Assessment procedure ☑ General ☐ Major ☐ Minor	1.5	18
3.1.7 Confidential commercial information	1.6	19
3.1.8 Exclusive capturable commercial benefit	1.7	19
3.1.9 International and other national standards	1.8	19
3.1.10 Statutory declaration	provided with application cover letter	
Foods Produced Using Gene Technology (Application Handbook Section 3.5.1)		
A.1 Nature and identity of the GM food	A.1	20
A.2 History of use of host and donor organisms	A.2	22
A.3 Nature of the genetic modification	A.3	27
B.1 Characterization and safety assessment	B.1	39
B.2 New proteins	B.2	54
B.3 Other (non-protein) new substances	B.3	72
B.4 Novel herbicide metabolites	B.4	75
B.5 Compositional analyses	B.5	76
C Nutritional impact of the GM food	C	85
D Other Information	D	87

Executive Summary

Rice event GR2E (IR-ØØGR2E-5) was developed through the use of recombinant-DNA techniques to express elevated levels of provitamin A (mainly β-carotene) in the rice endosperm, which is converted in the body to vitamin A. GR2E rice is intended to complement existing efforts to mitigate vitamin A deficiency by supplying consumers in societies whose diet is primarily rice-based with a portion of the estimated average requirement for vitamin A.

GR2E rice was produced by *Agrobacterium tumefaciens*-mediated transformation of embryogenic rice calli with plasmid pSYN12424 resulting in the introduction of the *phytoene synthase (psy1)* gene from *Zea mays* (*Zmpsy1*), the *carotene desaturase I (crtI)* gene from *Pantoea ananatis*, and the *phosphomannose isomerase (pmi)* gene from *Escherichia coli* as a selectable marker. The PMI protein, encoded by the *pmi* gene, is also expressed in a number of previously authorized maize lines, including MIR604, MIR162, 3272, and 5307.

Molecular characterization of the introduced DNA within event GR2E confirmed the presence at a single insertion site of one copy of the transfer-DNA (T-DNA) region derived from plasmid pSYN12424 that was stably inherited over multiple generations as a single genetic locus according to Mendelian rules of inheritance. In addition, nucleotide sequencing of the entire inserted DNA, including portions of the 5' and 3' flanking rice genomic sequence, confirmed that the T-DNA was inserted without modifications, deletions, or rearrangements, except for small truncations at the 5' and 3' termini of 23 bp and 11 bp, respectively. There were also no new novel open reading frames created as a consequence of the DNA insertion that would have the potential to encode a protein with any significant amino acid sequence similarity to known and putative toxins or allergens.

As predicted, expression of the *Zm*PSY1 and CRTI proteins was limited to the rice endosperm as assessed by western immunoblot analysis, while expression of the PMI protein was detected in all tissue types, including grain, bran, hulls, stem, leaves, and roots. In order to estimate potential human and animal dietary exposure to the *Zm*PSY1, CRTI, and PMI enzymes expressed in GR2E rice, the concentrations of these proteins in grain and straw, which represent the only potential pathways of dietary exposure, were determined in samples collected from four field locations over two growing seasons. For each protein, the highest measured concentrations were in samples of dough-stage grain, ranging between *ca*. 308–359 ng/g and between *ca*. 54–68 ng/g for *Zm*PSY1 and CRTI, respectively. Across the four locations, the highest concentrations of *Zm*PSY1 and CRTI measured in samples of mature grain were *ca*. 245 ng/g and 30 ng/g, respectively. The highest concentrations of PMI in samples of mature GR2E rice grain and straw were *ca*. 1891 ng/g and *ca*. 796 ng/g fresh weight tissue, respectively.

The maximum potential human daily dietary exposures to *Zm*PSY1, CRTI, and PMI proteins from GR2E rice were estimated to be less than *ca*. 4.5, 0.85, and 30 μg/kg body weight, respectively, based on the highest concentrations of these proteins determined in dough-stage grain and a maximum rate of rice consumption of 12.5 g/kg body weight, as reported for children in Bangladesh.

A tiered "weight-of-evidence" approach was followed in assessing the safety of the *Zm*PSY1, CRTI, and PMI proteins expressed in GR2E rice.

The *Zm*PSY1 and CRTI proteins did not display significant amino acid sequence similarity with known allergens nor were there any primary sequence structural alerts for potential toxicity based on similarity searches against a database of known and putative protein toxins. Both *Zm*PSY1 and CRTI were rapidly and completely digested *in vitro* in the presence of simulated gastric fluid containing pepsin, and the enzymatic activity of both proteins was completely destroyed following treatment at temperatures well below those used during cooking.

A tier-1 assessment of potential hazards associated with the *Zm*PSY1 protein, which considered the food crop source of the gene, lack of significant amino acid sequence similarity with known toxins and allergens, susceptibility to heat inactivation, and rapid digestibility concluded that further hazard characterization by animal toxicity testing was unnecessary.

Due to the non-food source of the *crtI* gene, acute oral toxicity testing of CRTI protein in mice was conducted as a further assurance of safety, which demonstrated a lack of any observable adverse effects at a dose of 100 mg/kg body weight, which represents at least a 115,000–fold margin of exposure relative to any realistically conceivable human dietary intake.

Based on its presence in a wide range of foods derived from genetically engineered maize lines, and on the extensive history of prior regulatory reviews, additional characterization of the PMI protein was unnecessary. Previously submitted safety studies reviewed in the context of other genetically engineered plant events are directly applicable to the safety assessment of PMI protein expressed in GR2E rice.

The genetic modification resulting in GR2E rice was only intended to increase levels of provitamin A (primarily β-carotene) in the rice endosperm. To confirm the intended effect and the lack of any meaningful unintended consequences of the genetic modification, compositional parameters were compared between GR2E rice and control, unmodified, rice. Compositional analyses were performed on samples of rice grain and straw obtained from PSB Rc82 rice containing event GR2E and near-isogenic control PSB Rc82 rice that was grown in side-by-side trials at four separate sites in the Philippines during 2015 and again in 2016. The compositional assessment included analyses for proximates, fibre, and minerals in samples of straw, and analyses for proximates, minerals, vitamins, amino acids, fatty acids, vitamins, and key anti-nutrients in grain samples. Samples of processed bran derived from GR2E and control rice were also analyzed for proximates, fibre, and minerals.

Among the 69 compositional components that were assessed in samples of GR2E and control PSB Rc82 rice grain, and 10 components that were assessed in straw samples, the only statistically significant difference observed from the multi-year combined-site analysis was for stearic (C18:0) acid, a minor fatty acid component, measured in grain samples (not including the intended difference in provitamin A levels). With the exception of β-carotene and related carotenoids, the compositional parameters measured in samples of

GR2E rice, including stearic acid, were within or similar to the range of natural variability of those components in conventional rice varieties with a history of safe consumption. Overall, no consistent patterns emerged to suggest that biologically meaningful changes in composition or nutritive value of the grain or straw had occurred as an unexpected, unintended consequence of the genetic modification.

The purpose of this evaluation of GR2E rice was to determine whether the use of GR2E rice in food could raise any new safety concerns relative to conventional rice, and was not intended to address questions related to the efficacy of GR2E rice in helping combat vitamin A deficiency (VAD) in affected population sub-groups.

Collectively, the data presented in this submission have not identified potential health and safety concerns, and support the conclusion that food derived from provitamin A biofortified GR2E rice is as safe and nutritious as food derived from conventional rice varieties.

Study Catalogue

A listing of submitted study reports and their correlation with relevant sections of the application dossier is shown in the following table.

Study ID	Dossier Section	Citation
IR2015-07003	A.3.4.1, A.3.4.2, A.3.4.3, A.3.4.5	Cueto, M., Salcedo, M., Oliva, N., and Trijatmiko, K. (2016). Southern hybridization characterization of event IR-ØØGR2E-5 rice. Technical report, IR2015-07003 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2015-08001	A.3.4.4	Trijatmiko, K., Oliva, N., Alexandrov, N., Slamet-Loedin, I., and Chadha-Mohanty, P. (2015). Nucleotide sequence analysis of the inserted DNA and host genomic flanking regions in rice event IR-ØØGR2E-5. Technical report, IR2015-08001 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-07001	A.3.4.6	Swamy, M. and Samia, M. (2016). Stability of the elevated beta-carotene trait across multiple generations of rice event IR-ØØGR2E-5. Technical report, IR-2016-07001 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2015-08003	A.3.4.7	Swamy, M. and Samia, M. (2015). Segregation of the inserted DNA within multiple generations of rice event IR-ØØGR2E-5. Technical report, IR2015-08003 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-04003	B.1.3	Oliva, N., Cueto, M., and Salcedo, M. (2016b). Western immunoblot analysis of ZmPSY1, CRTI, and PMI expression in different plant tissues from rice event IR-ØØGR2E-5. Technical report, IR2016-04003 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2015-08004	B.1.3	Oliva, N., Cueto, M., and Salcedo, M. (2016a). Concentrations of phytoene synthase (ZmPSY1), phytoene desaturase (CRTI), and phosphomannose isomerase (PMI) in grain and straw harvested from event IR-ØØGR2E-5 rice grown in the Philippines in 2015. Technical report, IR2015-08004 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-04004	B.1.3	Oliva, N., Cueto, M., and Salceo, M. (2016c). Concentrations of phytoene synthase (ZmPSY1), phytoene desaturase (CRTI), and phosphomannose isomerase (PMI) in grain and straw harvested from event IR-ØØGR2E-5 rice grown in the Philippines in 2016. Technical report, IR2016-04004 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-02003	B.1.4	MacKenzie, D. J. (2016d). Estimated daily dietary exposure to ZmPSY1, CRTI, and PMI proteins expressed in rice event IR-ØØGR2E-5. Technical report, IR2016-02003 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-02005	B.1.5	MacKenzie, D. J. (2016c). Characterization of ZmPSY1 protein (Lot Number M20452-05) derived from a microbial expression system. Technical report, IR2016-02005 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-02004	B.1.6	MacKenzie, D. J. (2016a). Characterization of CRTI protein (Lot Number M20454-02) derived from a microbial expression system. Technical report, IR2016-02004 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-01005	B.2.2.1	Oliva, N. and MacKenzie, D. J. (2016d). Amino acid sequence similarity search between Zea mays phytoene synthase (ZmPSY1) and known and putative protein toxins. Technical report, IR2016-01005 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-01004	B.2.2.2	Oliva, N. and MacKenzie, D. J. (2016b). Amino acid sequence similarity search between Pantoea ananatis phytoene desaturase (CRTI) and known and putative protein toxins. Technical report, IR2016-01004 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-01002	B.2.3.1	Oliva, N. (2016). Characterization of the in vitro pepsin digestibility of phytoene synthase protein (ZmPSY1). Technical report, IR2016-01002 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-07003	B.2.3.2	Oliva, N. and Cueto, M. (2016). Characterization of the in vitro pepsin digestibility of phytoene desaturase protein (CRTI) using western blot analysis. Technical report, IR2016-07003 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-04001	B.2.4.2	Mukerji, P. (2016). CRTI: Acute oral toxicity study in mice. Technical report, IR2016-04001 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-03002	B.2.4.2	MacKenzie, D. J. (2016b). Characterization of CRTI protein (Lot Number M20603) derived from a microbial expression system. Technical report, IR2016-03002 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-02002	B.2.6.1	Oliva, N. and MacKenzie, D. J. (2016c). Amino acid sequence similarity search between Zea mays phytoene synthase (ZmPSY1) and known and putative protein allergens. Technical report, IR2016-02002 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines

Study ID	Dossier Section	Citation
IR2016-02001	B.2.6.2	Oliva, N. and MacKenzie, D. J. (2016a). Amino acid sequence similarity search between Pantoea ananatis phytoene desaturase (CRTI) and known and putative protein allergens. Technical report, IR2016-02001 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2015-12002	B.2.7.1	Welsch, R. and Beyer, P. (2016). Characterization of the heat stability of phytoene synthase protein (ZmPSY1). Technical report, IR2015-12002 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines
IR2015-12001	B.2.7.2	Schaub, P. and Beyer, P. (2016). Characterization of the heat stability of phytoene desaturase protein (CRTI). Technical report, IR2015-12001 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2015-07001	B.5	Swamy, M., Samia, M., Boncodin, R., Rebong, D., Ordonio, R., and MacKenzie, D. J. (2016a). Nutrient composition of event IR-ØØGR2E-5 and non-transgenic control rice grown during the rainy season in 2015 in the Philippines. Technical report, IR2015-07001 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines
IR2016-05001	B.5	Swamy, M., Samia, M., Boncodin, R., Rebong, D., Ordonio, R., Miranda, R., Rebong, A., Alibuyog, S., Marundan, S., Adeva, C., and MacKenzie, D. J. (2016b). Nutrient composition of event IR-ØØGR2E-5 and non-transgenic control rice grown during the dry season in 2016 in the Philippines. Technical report, IR2016-05001 (unpublished) International Rice Research Insstitute, Los Banos, Laguna, Philippines
IR2016-07004	B.5.4	Samia, M. and Swamy, M. (2016). Concentrations of beta-carotene and other carotenoids in grain samples from rice event IR-ØØGR2E-5. Technical report, IR2016-07004 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines

Abbreviations

ADF	acid detergent fibre
AOAC	Association of Official Analytical Chemists
AP	alkaline phosphatase
BLAST	basic local alignment search tool
BSA	bovine serum albumin
CI	confidence interval
CIM	callus induction medium
COA	certificate of analysis
crtI	carotene desaturase I
CRTISO	carotene *cis-trans* isomerase
CTAB	cetyltrimethyl ammonium bromide
DALY	disability-adjusted life year
DIG	digoxigenin
DMAPP	dimethylallyl diphosphate
EAR	estimated average requirement
EDTA	ethylenediamine tetraacetic acid
EFSA	European Food Safety Authority
ELISA	enzyme-linked immunosorbent assay
FAD	flavin adenine dinucleotide
FAO	Food and Agriculture Organization of the United Nations
FARRP	Food Allergy Research and Resource Program
FASTA	FAST All sequence alignment tool
FDA	Food and Drug Administration
FPIES	food protein-induced enterocolitis syndrome
FPP	farnesyl diphosphate
FWT	fresh weight tissue
GGPP	geranylgeranyl diphosphate
GluA-2	glutelin promoter
GPP	geranyl diphosphate
HPLC	high-pressure liquid chromatography
HRP	horseradish peroxidase
ILSI	International Life Sciences Institute
IPP	isopentenyl diphosphate
IRRI	International Rice Research Institute
LAAO	L-amino acid oxidase
LB	Left Border
LS	least squares
MAb	monoclonal antibody
MS	Murashige and Skoog basal salt mixture
NARES	National Agricultural Research and Extension System
NCT	National Cooperative Test
NDF	neutral detergent fibre

NOS	nopaline synthase
OD	optical density
OECD	Organisation for Economic Cooperation and Development
ORF	open reading frame
OsPDS	*Oryza sativa* phytoene desaturase
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDS	phytoene desaturase
pmi	phosphomannose isomerase
psy1	phytoene synthase
RAE	retinol activity equivalent
RB	Right Border
RDA	recommended daily allowance
RUBISCO	ribulose-1,5-bisphosphate carboxylase
SDS	sodium dodecylsulfate
SGF	simulated gastric fluid
SPT	skin prick test
SSC	saline sodium citrate
SSU	small sub-unit
T-DNA	transfer-DNA
TCEP	tris(2-carboxyethyl)phosphine
TDF	total dietary fibre
TMB	3,3',5,5'–tetramethylbenzidine
VAD	vitamin A deficiency
WHO	World Health Organization
Z-ISO	ζ-carotene *cis-trans* isomerase
ZDS	ζ-carotene desaturase

Contents

List of Tables

List of Figures

I. Application Handbook Section 3.1

1. General Information on the Application

1.1. Applicant Details

(a)	Applicant's name/s	Donald J. MacKenzie, Ph.D.
(b)	Company/organization name	International Rice Research Institute (IRRI)
(c)	Address	Pili Drive, UPLB, Los Baños, 4031, Laguna Philippines
(d)	Telephone number	+1 (202) 695-0436
(e)	Email address	d.mackenzie@irri.org
(f)	Nature of the applicant's business	IRRI is the world's premier research organization dedicated to reducing poverty and hunger through rice science; improving the health and welfare of rice farmers and consumers; and protecting the rice-growing environment for future generations. IRRI is an independent, nonprofit, research and educational institute, founded in 1960 by the Ford and Rockefeller foundations with support from the Philippine government. The institute, headquartered in Los Baños, Philippines, has offices in 17 rice-growing countries in Asia and Africa, and more than 1,000 staff.
(g)	Details of other individuals, companies or organizations associated with the application	Not applicable

1.2. Purpose of the Application

The purpose of this submission is to make an application to amend the Australia New Zealand Food Standards Code to allow for the inclusion of provitamin A biofortified rice event GR2E in Standard 1.5.2 — Food Produced Using Gene Technology.

Rice is a staple for nearly half of the world's seven billion people and is used in various forms including whole and milled grain, flour and bran. The husks may be used for fertilizers and animal feed as well as for fibre production. Over 90 percent of rice production and consumption is in Asia, with around five percent from the Americas, three percent from Africa and another one percent from Europe and Oceania.

Rice varieties containing event GR2E are intended for cultivation and use in a number of south and southeast Asian countries. It is therefore anticipated that raw agricultural

commodity and/or food products derived from GR2E rice varieties will enter the Australian and New Zealand food supply via imports from countries of production.

1.3. Justification for the Application

Rice event GR2E was produced through genetic engineering techniques and contains the *psy1* gene from *Zea mays*, the *crtI* gene from the common soil bacterium, *Pantoea ananatis* (syn. *Erwinia uredovora*), and the *pmi* gene from *Escherichia coli* as a selectable marker. Production of the *Zm*PSY1 and CRTI enzymes is targeted to the rice endosperm and results in a functional β-carotene biosynthetic pathway within the grain endosperm (Paine et al., 2005; Ye et al., 2000).

The intended nutritional effect of GR2E rice is to complement existing VAD control efforts by supplying up to 30–50 percent of the estimated average requirement (EAR) for vitamin A for preschool age children and pregnant or lactating mothers in high-risk countries, including Bangladesh, Indonesia, and the Philippines.

Various studies have examined the public health burden associated with VAD. According to calculations from Stein et al. (2006), the impact of VAD in India amounted to 2.3 million disability-adjusted life years (DALYs) lost each year, of which 2.0 million DALYs were lost due to child mortality alone. Similar studies in Bangladesh and the Philippines reported VAD-associated health burdens of 262,000 and 432,000 DALYs annually, respectively (Zimmermann and Ahmed, 2006; Zimmermann and Qaim, 2004). In 2004, it was estimated that VAD was responsible, globally, for *ca*. 5.3 percent of DALYs lost and 668 thousand deaths among children under five years of age (Black et al., 2008).

The introduction of GR2E rice in countries at high risk for VAD will not eliminate the problems of VAD, such as blindness or increased mortality rates. GR2E rice is a complementary approach and not a substitute for alternative interventions, and is expected to reduce the VAD-related health burden significantly.

1.4. Costs and Benefits and Impact on Trade

GR2E rice is not intended for commercialisation in Australia or New Zealand, and the inclusion of GR2E rice in Standard 1.5.2 is not anticipated to impact rice exports.

In 2014, total Australian imports of milled rice were 156,000 metric tonnes, primarily from Thailand (52%), India (19%), Pakistan (11.5%), the United States (7%), and Vietnam (5%) (USDA-FAS, 2015). The intended production of GR2E rice in south and southeast Asia is not expected to impact the volume rice imports. Inclusion of provitamin A biofortified GR2E rice in Standard 1.5.2 will reduce the likelihood of any trade disruptions, especially in the case of eventual cultivation in countries that export significant quantities of milled rice to Australia and/or New Zealand.

1.5. Assessment Procedure

IRRI considers that the appropriate assessment for this application is the General Procedure.

1.6. Confidential Commercial Information

IRRI is not asking that any part of this application be treated as confidential commercial information as defined in section 4 of the *FSANZ Act, 1991*.

1.7. Exclusive Capturable Commercial Benefit

This application will not confer an exclusive capturable commercial benefit upon the International Rice Research Institute. Rice varieties containing event GR2E are being developed for humanitarian purposes to aid in the control of VAD and their distribution within target countries will not accrue any commercial benefit to IRRI. Additionally, GR2E rice varieties are not intended for commercial propagation in Australia.

1.8. International and other National Standards

1.8.1. International Standards

The IRRI reports and studies included in the information supporting this application have been conducted according to international standards. In the food safety assessment of genetically engineered plant products, IRRI refers primarily to the Codex Alimentarius Commission Guideline for the Conduct of Food Safety Assessment of Foods Derived from Recombinant-DNA Plants (Codex, 2003).

In addition, the composition analyses were conducted in accordance with the Organisation for Economic Cooperation and Development (OECD) consensus document on new rice varieties (OECD, 2016).

1.8.2. Other National Standards or Regulations

As of the date of this submission, GR2E rice is not approved for use in food or livestock animal feed or for general (unconfined) environmental release in any country.

This submission is being filed concurrently with an application to Health Canada pursuant to the Novel Foods provisions of the *Food and Drug Regulations* (Division 28).

IRRI is undertaking a premarket biotechnology consultation with the United States Food and Drug Administration (FDA) and, together with its National Agricultural Research and Extension System (NARES) partners, planning regulatory submissions in Bangladesh, the Philippines, and Indonesia.

Additional regulatory submissions in other jurisdictions will be made on an as-needed basis.

II. Application Handbook Section 3.5.1

A. Technical Information on the GM Food

A.1. Nature and Identity of the GM Food

A.1.1. GR2E Rice as a Complementary VAD Control Strategy

Improving the availability and intake of vitamin A through dietary diversification is viewed as one of the most sustainable approaches to enhancing overall nutritional status of the population. This requires nutrition education to change dietary habits, as well as providing better access to vitamin A or provitamin A-rich foods, such as meat and dairy products, mangoes, papaya, or dark green leafy vegetables. And, it requires improving the economic status of populations to enable increased access to a diversified diet.

A second approach to increasing the dietary intake of vitamin A is through conventional fortification of staple foods or condiments with vitamin A. Although many food items such as fats, oils, margarine, and cereal products have long been fortified with vitamin A in high income countries, few other vitamin A fortification programmes with national reach currently exist in lower income countries.

Supplementation is regarded as an effective strategy and is widely practiced to control VAD in low-income countries. The periodic dispensing of potent supplements (200,000 IU of vitamin A) to preschool age children (< 5 years), or 100,000 IU of vitamin A to infants aged 6–11 months, has shown promise to control VAD in many developing countries in South Asia. While periodic vitamin A delivery in the community has been shown to reduce the risks of xerophthalmia and mortality in young children, it has been less effective in sustainably raising population serum retinol concentrations.

A complementary intervention to existing strategies for reducing VAD in the highest-risk countries is the biofortification of staple food crops with provitamin A through conventional plant breeding. This approach has been successfully used with crops such as maize, cassava, and sweet potato where existing germplasm with elevated levels of provitamin A expression has been introduced into breeding programmes (La Frano et al., 2013; De Moura et al., 2014).

Rice leaves, and indeed the photosynthetic tissues of all higher plants, produce and accumulate β-carotene. However, rice in its milled form, as it is usually consumed, is characterized by the complete absence of provitamin A carotenoids. The milled rice kernel consists exclusively of the endosperm, as the carotenoid-containing embryo and the aleurone layer are removed during milling of the rice grain.

The accumulation of provitamin A carotenoids in rice endosperm cannot be achieved through conventional breeding approaches. Germplasm screening has not revealed any rice

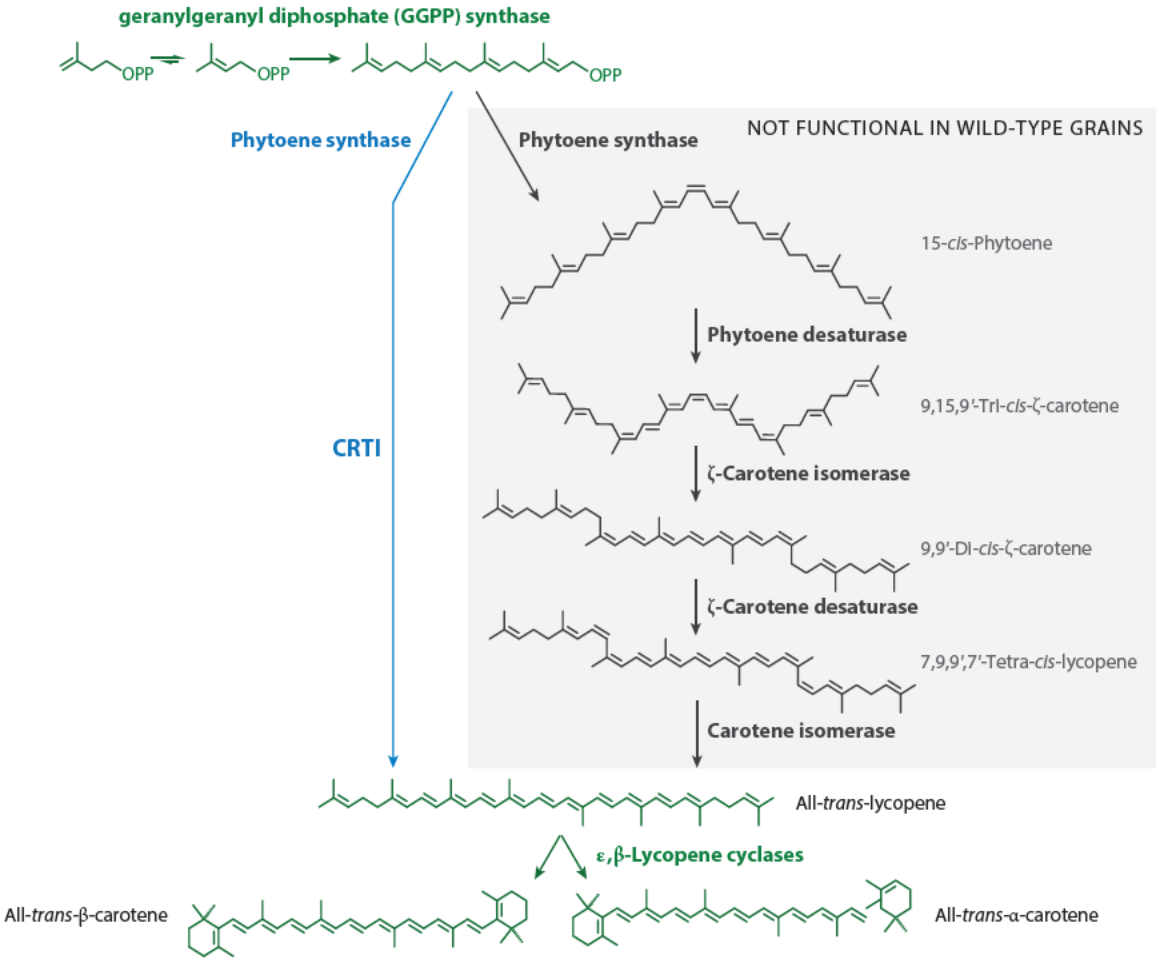


Figure 1. The green structures denote activities in the wild-type endosperm. The latest precursor available is geranylgeranyl diphosphate (GGPP). The subsequent steps (gray) are not functional in wild-type endosperm and two introduced enzymes, phytoene synthase and CRTI, are required to bridge the gap toward lycopene, which is the substrate of two competing cyclases: ε-cyclase and β-cyclase, which acting together lead to the formation of α-carotene, whereas the action of β-cyclase alone forms β-carotene.

cultivars capable of accumulating β-carotene in the grain that could be used as a parental line for further breeding.

The carotenoid biosynthetic pathway in immature rice endosperm functions up to the synthesis of geranylgeranyl diphosphate (Figure 1). While this compound is not solely devoted to carotenogenesis, it can be used as a substrate to produce the uncoloured carotene, phytoene, by expressing the heterologous enzyme phytoene synthase in rice endosperm (Burkhardt et al., 1997). Completion of the pathway leading to the synthesis of all-*trans*-lycopene via expression of the *crtI* gene from *Pantoea ananatis* was described by Beyer et al. (2002). Initial proof-of-concept work utilized the *psy* gene from *Narcissus pseudonarcissus* (daffodil), which was found to be rate-limiting for the accumulation of β-carotene. Paine et al. (2005) described improved constructs incorporating the *Zea mays* phytoene synthase gene (*Zmpsy1*) that yielded transformants, including event GR2E (SGR2E1), that produced up to 30 $\mu g/g$ total carotenoids in the endosperm, of which *ca.* 80 percent were mixed isomers of β-carotene.

A.1.2. Name, Number or other Identifier

In accordance with OECD's "Guidance for the Designation of a Unique Identifier for Transgenic Plants" GR2E has been assigned the unique identifier IR-ØØGR2E-5.

A.1.3. Name the Food will be Marketed Under (if Known)

Rice containing the transformation event GR2E will be produced primarily in countries in south and southeast Asia. There are currently no plans to produce this product in Australia and New Zealand. IRRI has not as yet determined a trade name for the product.

A.1.4. Types of Products Likely to Include GR2E Rice

Rice is a staple consumed by nearly half the world's population and is used in various forms, including whole and milled grain, flour, and bran. Rice byproducts, such as bran and straw, are used in livestock feed. GR2E rice may be utilized in the same manner and for the same uses as conventional rice.

A.2. History of Use of the Host and Donor Organisms

A.2.1. Host Organism

Rice is the common name for the plant *Oryza sativa* L., which has a long history of use as food dating back at least 4000 years. Rice is used in various forms including whole and milled grain, flour and bran. The husks may be used for fertilizers and animal feed as well as for fibre production. Numerous varieties of rice have been developed from subspecies indica, japonica, and javanica. Over 90 percent of rice production and consumption is in Asia, with around five percent from the Americas, three percent from Africa and another one percent from Europe and Oceania. The crop is well adapted to diverse growing conditions from cool climates to deserts (with irrigation) and is able to perform well in areas with saline, alkaline, or acid-sulphate soils.

The biology of rice has been extensively described in various publications from international organizations (OECD, 1999) and government agencies (Tripathi et al., 2011; OGTR, 2005) and will not be further reviewed here.

A.2.1.1 Consumption Patterns

Rice is a staple for nearly half of the world's seven billion people. However, more than 90 percent of this rice is consumed in Asia, where it is a staple for a majority of the population. Rice provides up to 50 percent of the dietary caloric supply and a substantial part of the protein intake for about 520 million people living in poverty in Asia (Muthayya et al., 2014).

With the success of the Green Revolution in the early 1960s, there was a steady rise in Asia's *per capita* rice consumption from 85 to nearly 103 kg/year in the early 1990s, and globally, an increase in *per capita* consumption from 50 to 65 kg/year (Mohanty, 2013). The rising *per capita* consumption plus the growing population more than doubled global rice consumption during this period from 150 to 350 million tons.

Since the 1990s, strong economic growth, particularly in China and India, has led to consumers diversifying their diet from rice to high-value products such as meat, dairy products, fruits, and vegetables, resulting in a flattening trend in *per capita* rice consumption. While *per capita* consumption in countries such as Bangladesh and Cambodia continues to rise, consumption is on a downward trend in Thailand, Vietnam, and Malaysia.

A.2.1.2 Processing

Parboiling is a traditional method of rice processing widely used in some South Asian countries (e.g., Bangladesh, India, Sri Lanka) and in some rice exporting countries. Parboiled rice is prepared by soaking, draining, heating (most often steaming; sometimes under pressure), then drying. Parboiling treatment induces various physiochemical changes in paddy, including gelatinization of the rice starch, and improves the hardness of the rice upon drying, which minimizes breakage loss during subsequent milling.

Paddy or parboiled rice is dehusked (hulled) to make brown rice, which is then further processed by milling to remove all or most of the bran that consists primarily of the seed coat, aleurone layer, and germ. Milled rice is processed by polishing to remove residual bran on the surface to give a smoother finish, and may be further polished to obtain the inner part of rice grain (i.e., polishings) containing less protein for further processing into rice flour and starch (Figure 2).

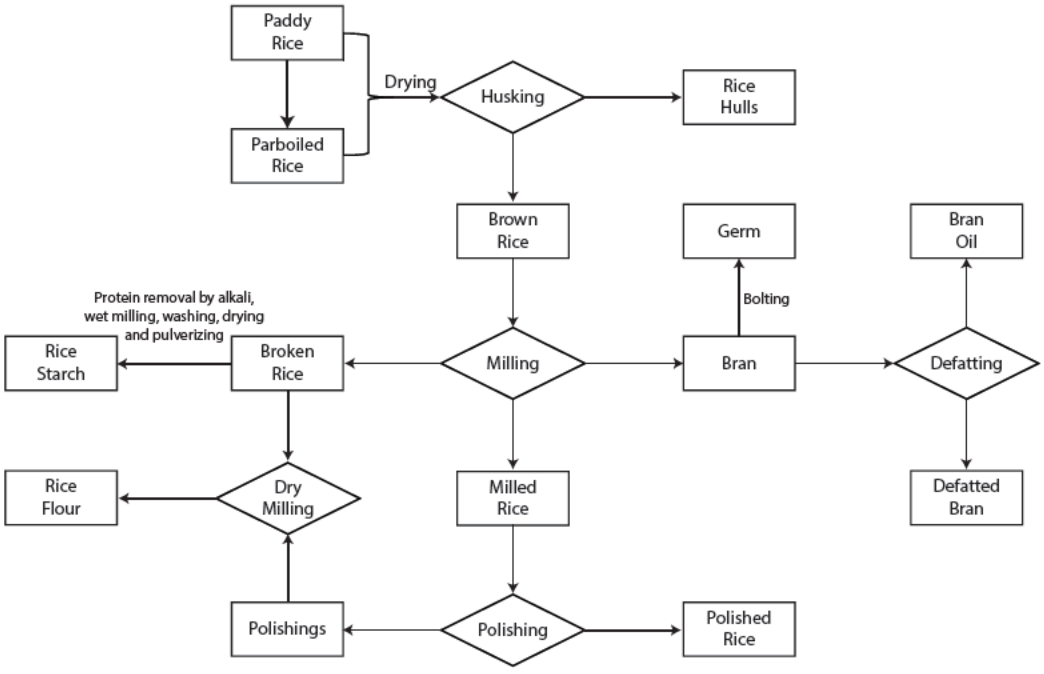


Figure 2. Schematic representation of rice processing steps. Adapted from OECD (2004).

Most of the rice used for food is milled rice. Rice flour is a pulverized product of the outer part or the whole milled rice. Rice bran oil is made from rice bran, which is used as cooking oil.

A.2.1.3 Key Nutrients and Anti-Nutrients

Brown, milled, polished, and parboiled rice are the major rice products consumed by humans in the form of grain after being cooked. Rice products, such as rice flour, are also used in food manufacturing including in cereals, baby food, and snacks.

The main nutrients provided by rice are carbohydrates and protein. The carbohydrate component of milled rice consists mainly of starch, which is comprised of amylose and amylopectin fractions, and small amounts of free sugars and non-starch polysaccharides. Most of the dietary fibre is contained in the hull and rice bran and germ, and is lost by hulling, milling, and polishing.

The principal rice storage proteins present in the outer layer and the inside of milled rice are glutelin and prolamin, accounting for 60–65 and 20–25 percent of total protein, respectively. Compared to other cereals such as wheat and maize, rice has a more complete and balanced amino acid composition. Other proteins, albumin and globulin, are found mainly in the outer layer of brown rice, thus the protein composition of bran and germ differ significantly from that of milled rice.

The fat content of rice grain is concentrated primarily in the germ, aleurone layer, and sub-aleurone layer, mostly as triglycerides in which glycerol is esterified with three fatty acids, oleic, linoleic, and palmitic acid. Additionally, free fatty acids, sterol, and diglycerides are found in smaller amounts together with lipid-conjugates like acylsterolglycoside and sterolglycoside, glycolipids and phospholipids. The fatty acid composition of rice grain is dependent on growing conditions, especially by temperature during the ripening stages, and by the genetics of different varieties.

Rice is not considered a significant source of micro-nutrient minerals, such as iron and zinc, and in high rice-consuming countries zinc deficiency and iron deficiency anaemia affect about 38 percent of pregnant women and 43 percent of preschool children.

Rice grain (paddy) contains water-soluble vitamins, including thiamine (B1), riboflavin (B2), niacin (B3), pyridoxine (B6), cyanocobalamin (B12), and fat-soluble vitamin E (tocopherols) but does not contain significant amounts of other fat-soluble vitamins such as vitamin A, D, and K. Vitamins are mainly present in the outer (bran) layers and are significantly reduced upon milling. Enriched white rice is fortified with iron, thiamine, and niacin (and with folic acid in the United States since January 1998) to control beriberi and iron deficiencies. Enriched rice, however, is not available to the large numbers of rice consumers who grow their own rice or eat locally produced rice.

Rice is generally not considered a significant source of anti-nutritional factors and these have not historically been present in rice-based foods at levels that would pose safety concerns. These anti-nutritional factors, which are mainly concentrated in the bran, are phytic acid, digestive enzyme inhibitors (e.g., trypsin inhibitors, alpha-amylase inhibitors, and the cysteine proteinase inhibitor, oryzacystatin), and lectins. With the exception of phytic acid, the other anti-nutritional factors are proteinaceous in nature and subject to heat denaturation and inactivation during cooking.

The OECD has published a consensus document on key food and feed nutrients and anti-nutrients found in rice (OECD, 2004), which was recently updated (OECD, 2016). Compositional parameters recommended for analysis in new rice varieties are shown in Table 1.

Table 1. OECD suggested nutritional and compositional parameters to be analyzed in rice matrices for food or feed use

	Food Use	Feed Use	
Parameter	**Paddy Rice or Brown Rice**	**Paddy Rice**	**Straw**
Proximates[a]	✓	✓	✓
Vitamins[b]	✓		
Amino acids	✓	✓	
Fatty acids	✓		

[a] Proximates includes moisture, protein, fat, ash, carbohydrates, and total dietary fibre. For feed use, the fibre components include acid detergent fibre (ADF) and neutral detergent fibre (NDF).
[b] Includes B vitamins, namely thiamine (B1), riboflavin (B2), niacin (B3), pantothenic acid (B5), and pyridoxine (B6), and vitamin E (α-tocopherol).

A.2.1.4 Allergy

Rice is not considered by allergists to be a common allergenic food. Although nearly one-half of the world population consumes cooked rice on a daily basis, published evidence indicates that food allergy to rice is rare. However, rice allergy has been reported in countries of Asia including Japan, Malaysia, Thailand, and Indonesia and in some European countries including Finland, France, Spain, Sweden, Denmark, Estonia, Lithuania, and in Russia (Besler and Tanabe, 2001; Kumar et al., 2007). The prevalence of IgE-mediated rice allergy is approximately 10 percent in atopic subjects in Japan (OECD, 2016; Suvarna, 2008), and is more prevalent in adults than children. Symptoms of rice allergy include atopic dermatitis, eczema, and food protein-induced enterocolitis syndrome (FPIES) (Hoffman, 1975; Ikezawa et al., 1992; Mehr et al., 2009).

If specific IgE binding or skin prick tests (SPTs) are used as the primary measures of allergy to rice, higher incidences have been reported, but these appear to be due primarily to low-affinity IgE binding in many subjects who do not suffer from allergy upon ingestion of cooked rice, or SPT cross-reactivity that might be associated with sensitization to rice or grass pollen, or inhalation of rice flour dust.

There are rare cases of FPIES associated with rice consumption in young children and rare cases of IgE mediated food allergy. However, most pediatricians still recommend rice as a very safe food that can be introduced into the diets of very young infants as they begin dietary supplementation following breast feeding and as they are weaned.

There are two putative rice food allergens, *Oryza* trypsin alpha-amylase inhibitors (14–16 kDa) and *Oryza* glyoxalase I (33 kDa), which are included in the Food Allergy Research and Resource Program database (http://www.allergenonline.org).

A.2.2. Donor Organisms

A.2.2.1 Zea mays

Zea mays (maize) was the source of the *psy1* gene (Buckner et al., 1996) and the polyubiquitin promoter (Christensen et al., 1992). Maize is a grass that was first domesticated in Mexico more than 8000 years ago and is now cultivated worldwide as the third most planted crop after wheat and rice. Maize has a long history of safe use and represents a staple food for a significant proportion of the world's population. No significant endogenous toxins are reported to be associated with the genus *Zea* (IFBC, 1990). Food allergy to maize is relatively rare, and the only significant reported food allergen is a nonspecific lipid transfer protein (Pastorello et al., 2000).

A.2.2.2 Pantoea ananatis

Pantoea ananatis (formerly *Erwinia uredovora*; family *Enterobacteriaceae*) was the source of the *crtI* gene (Misawa et al., 1990) and is found in a wide range of natural environments, including water, soil, as part of the epi- and endo-phytic flora of various plant hosts (Coutinho and Venter, 2009). Many members of the *Enterobacteriaceae*, belonging to the genera *Serratia*, *Enterobacter*, *Pantoea*, *Proteus*, and *Hafnia*, often contribute to meat spoilage (Ercolini et al., 2006). The ubiquity of *P. ananatis* suggests that it has adapted to proliferate in a wide range of environments, and its isolation from both plant and animal hosts indicate it has adapted for cross-Kingdom colonization and pathogenesis (De Maayer et al., 2014).

Although strains of *P. ananatis* have been found to be pathogenic on a broad range of plant hosts as well as humans, a recent analysis of the *P. ananatis* genome to identify potential molecular determinants of its underlying pathogenicity revealed the absence of many of the factors that are central to the pathogenicity and virulence arsenal of related plant and animal pathogens, including animal toxins, hemolysins, phytotoxins, and their associated effectors (De Maayer et al., 2014).

A.2.2.3 Pisum sativum

Pea (*Pisum sativum* L.) is the source of the ribulose-1,5-bisphosphate carboxylase (RUBISCO) small sub-unit (SSU) transit peptide signal sequence (Coruzzi et al., 1984) that is fused to the N-terminus of the CRTI protein to direct transport to the chloroplasts. Pea is an important legume grown and consumed extensively worldwide, and is a rich source of proteins, carbohydrates and vitamins. Pea is the fourth leading legume in terms of consumption in the world and is the second most important legume after common bean.

There are a few reports of allergy to garden peas, with clear IgE binding to a few major proteins (Sanchez-Monge et al., 2004; Sell et al., 2005), but no reports of IgE binding to the chloroplast transit peptide sequence or the full-length or mature RUBISCO.

A.2.2.4 Escherichia coli

Escherichia coli (family *Enterobacteriaceae*) strain K12, a non-pathogenic strain, was the source of the *pmi* gene (Miles and Guest, 1984) and is a gram-negative, motile, facultatively anaerobic rod-shaped bacterium. Certain serotypes are enteropathogenic and are known to

cause diarrhoea in infants. Some strains also cause diarrhoea in adults. *E. coli* is a normal inhabitant of the intestinal flora of humans and animals, where it generally does not cause disease.

A.2.2.5 Other Donor Organisms

Other donor organisms, including rice (*Oryza sativa*) and *Agrobacterium tumefaciens* were used as sources of gene regulatory sequences that are not expressed in the transformed plant. These sequences include the *glutelin promoter (GluA-2)* from rice and the nopaline synthase (NOS) 3' untranslated regions from *A. tumefaciens*. Since none of these sequences encode expressed products in GR2E rice, their donor organisms are of little relevance to assessing potential toxicity or allergenicity.

A.3. Nature of the Genetic Modification

A.3.1. Transformation Method

The *Agrobacterium*-mediated transformation of the japonica rice cultivar Kaybonnet using plasmid pSYN12424 was previously described by Paine et al. (2005). Briefly, embryogenic cultures were established from mature embryos on Murashige and Skoog basal salt mixture (MS)-callus induction medium (CIM) (4.3 g/l MS salts, 5 ml/l B5 vitamins, 30 g/l sucrose, 500 mg/l proline, 500 mg/l glutamine, 300 mg/l casein hydrolysate, 2 mg/l 2,4-D, 3 g/l Phytagel, pH 5.8). Embryogenic calli (3–4 mm) were inoculated with *Agrobacterium tumefaciens* harbouring plasmid pSYN12424 and incubated at 22°C for two days, followed by transfer onto MS-CIM media containing ticarcillin (400 mg/l), and then onto mannose selection media (MS-CIM containing 17.5 g/l mannose, 5 g/l sucrose, and 300 mg/l ticarcillin) for five weeks in the dark. Proliferating colonies were transferred to regeneration medium (MS-CIM with 0.5 mg/l indole acetic acid, 1 mg/l zeatin, 200 mg/l ticarcillin, 20 g/l mannose, 30 g/l sorbitol, no sucrose), grown in the dark for 14 days and then moved to light at 30°C for 14 days. Shoots were transferred to MS medium containing 20 g/l sorbitol for two weeks and then to soil.

A.3.2. Description of the Potentially Introduced Genetic Material

Agrobacterium-mediated transformation was performed using plasmid pSYN12424 (Figure 3), which contains three gene expression cassettes within the T-DNA (Figure 4). These gene expression cassettes are briefly described below, and summarized in Table 2.

The first cassette contains a copy of the *crtI* gene from *Pantoea ananatis* (Misawa et al., 1990) that is fused in-frame at the 5' terminus with the pea (*Pisum sativum*) RUBISCO SSU transit peptide encoding sequence (Coruzzi et al., 1984). Transcription of the *crtI* gene is controlled by the rice *GluA-2* promoter (Takaiwa et al., 1987) for targeted expression in the rice endosperm. Transcription termination of the *crtI* gene is provided by the polyadenylation signal and 3' untranslated region from the NOS gene of *Agrobacterium tumefaciens* Ti-plasmid pTiT37 (Depicker et al., 1982). The CRTI enzyme catalyzes the conversion of 15-*cis*-phytoene to all-*trans*-lycopene (Schaub et al., 2012).

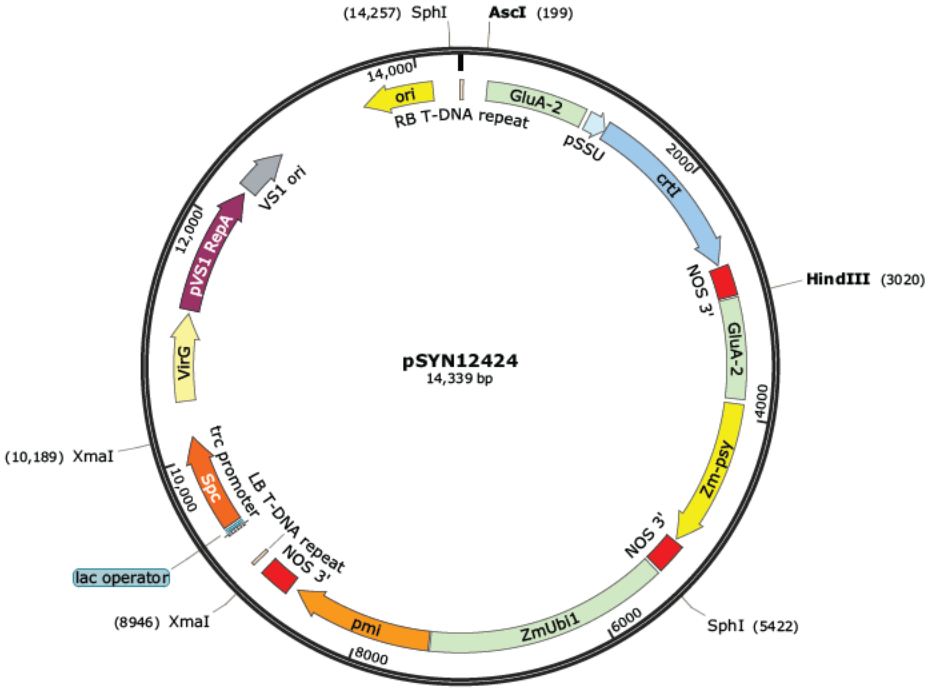


Figure 3. Schematic map of pSYN12424 indicating the organization of genetic elements. The Right Border (RB) and Left Border (LB) regions flank the T-DNA (Figure 4) that is intended for integration within the host plant genome during *Agrobacterium*-mediated transformation.

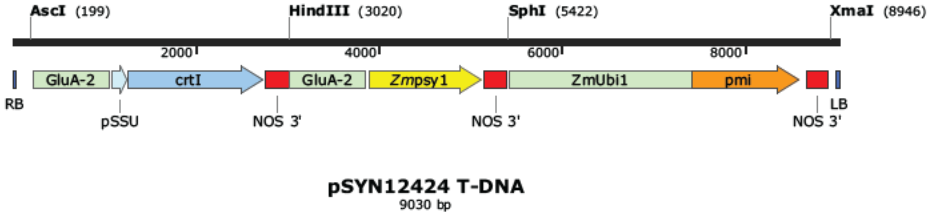


Figure 4. Schematic diagram of the pSYN12424 T-DNA indicating the *crtI*, *Zmpsy1*, and *pmi* genes along with their respective regulatory elements, including the rice glutelin promoter (*GluA-2*), pea RUBISCO SSU chloroplast transit peptide signal sequence, maize polyubiquitin promoter and intron region (*ZmUbi1*), and NOS 3' terminator. The size of the T-DNA is 9030 bp

The second cassette contains the *psy1* gene, isolated from *Zea mays* (Buckner et al., 1996), under the control of the rice *GluA-2* promoter with termination sequences derived from the 3' untranslated region of the *A. tumefaciens* NOS. Expression of phytoene synthase (*Zm*PSY1) in the rice endosperm catalyzes the conversion of geranylgeranyl diphosphate to phytoene.

The third and final cassette contains a copy of the phosphomannose isomerase (*pmi*) gene from *Escherichia coli* (Miles and Guest, 1984). Expression of the *pmi* gene is controlled by the maize polyubiquitin promoter (Christensen et al., 1992), providing constitutive expression of the PMI protein in rice. This region also includes the 5' untranslated region (UTR) and intron associated with the native polyubiquitin promoter. Transcription termination is via the NOS 3' untranslated region from *A. tumefaciens*. Expression of PMI catalyzes the reversible isomerization of mannose-6-phosphate to fructose-6-phosphate,

allowing positive selection of transformed calli and plantlets on mannose-containing medium (Negrotto et al., 2000).

A.3.3. Breeding of GR2E Rice

The breeding of GR2E rice proceeded as indicated in Figure 5 to produce specific generations for the characterization and assessments conducted, as described in Table 3. Event GR2E in japonica Kaybonnet background (inbred, direct line of descent from the original T0 transformant) was crossed into three different indica rice backgrounds, PSB Rc82, BRRI *dhan* 29, and IR64. During the breeding and development of GR2E rice, seed was produced for multiple purposes, including product evaluation, research testing, and regulatory testing. Due to seed limitations, seed was sourced from multiple generations of GR2E rice in different genetic backgrounds depending on the study, and where applicable, studies utilized the appropriate near-isogenic conventional comparator.

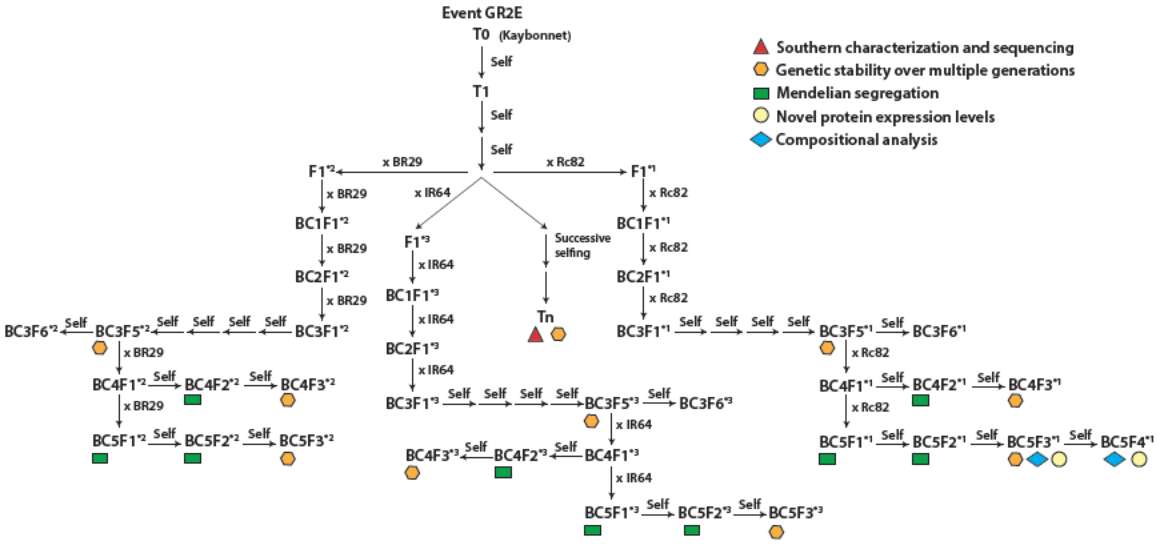


Figure 5. Breeding diagram for GR2E rice and generations used for analyses

A.3.4. Molecular Genetic Characterization of GR2E Rice

Southern blot analyses (Southern, 1975) were performed to investigate the number of sites of insertion of the pSYN12424 T-DNA, the integrity of introduced genetic elements, the absence of plasmid backbone sequences, and the multi-generational stability of the inserted DNA (Cueto et al., 2016). A complete description of the methods used, including locations and identities of various hybridization probes, is presented in Appendix A, beginning on page 103.

A.3.4.1 Insert Copy Number within the GR2E Rice Genome

The restriction enzymes *Hin*dIII and *Sph*I have unique recognition sites at positions 3020 and 5422 within the pSYN12424 T-DNA (Figure 6) and were used to provide information about the number of copies of the introduced DNA integrated within the GR2E rice genome.

Table 2. Description of the genetic elements within the pSYN12424 T-DNA

Gene Expression Cassette	Location on pSYN12424 T-DNA (bp)	Genetic Element	Size (bp)	Description
	1–25	Right border	25	Right border region of T-DNA from *Agrobacterium tumefaciens* nopaline Ti plasmid (GenBank accession no. J01826). Short direct repeat that flanks the T-DNA and is required for the transfer of the T-DNA into the genome of the host plant (Yadav et al., 1982).
	26–219	Ti plasmid region	194	Non-functional sequence from Ti plasmid of *A. tumefaciens*.
crtI gene expression cassette	220–1047	*GluA-2* promoter	828	Glutelin *GluA-2* promoter from *Oryza sativa* (Takaiwa et al., 1987) GenBank accession no. D00584.
	1048–1086	Intervening sequence	39	DNA sequence used for cloning.
	1087–1257	pSSU	171	RUBISCO SSU transit peptide coding sequence from pea (*Pisum sativum*) fused in frame with *crtI* coding sequence (Coruzzi et al., 1984); GenBank accession no. X00806.
	1258–2736	*crtI* Gene	1479	Phytoene desaturase gene from *Pantoea ananatis* (Misawa et al., 1990); GenBank accession no. D90087; with change A3992 to G.
	2737–2756	Intervening sequence	20	DNA sequence used for cloning.
	2757–3009	NOS 3'	253	Termination sequences of nopaline synthase (*nos*) gene from *Agrobacterium tumefaciens* (Depicker et al., 1982); GenBank accession no. V00087.
	3010–3025	Intervening sequence	16	DNA sequence used for cloning.
*Zm*psy1 gene expression cassette	3026–3853	*GluA-2* promoter	828	Glutelin *GluA-2* promoter from *Oryza sativa* (Takaiwa et al., 1987) GenBank accession no. D00584.
	3854–3892	Intervening sequence	39	DNA sequence used for cloning.
	3893–5125	*Zmpsy1* Gene	1233	Phytoene synthase gene from *Zea mays* (Buckner et al., 1996); GenBank accession no. U32636, B73 allele, A1031C. Phytoene synthase catalyzes the conversion of geranylgeranyl diphosphate to phytoene.
	5126–5145	Intervening sequence	20	DNA sequence used for cloning.
	5146–5398	NOS 3'	253	Termination sequences of nopaline synthase (*nos*) gene from *Agrobacterium tumefaciens* (Depicker et al., 1982); GenBank accession no. V00087.
	5399–5423	Intervening sequence	25	DNA sequence used for cloning.
pmi gene expression cassette	5424–7416	*ZmUBI* Promoter	1993	Promoter region and first intron from the *Zea mays* polyubiquitin gene (Christensen et al., 1992); GenBank accession no. S94464, with changes A160G, addition of C at 813, and deletion of C at 1012.
	7417–7428	Intervening sequence	12	DNA sequence used for cloning.
	7429–8604	*pmi* Gene	1176	Phosphomannose isomerase encoding gene from *Escherichia coli* (strain K12) (Miles and Guest, 1984); GenBank accession no. M15380. Catalyzes the isomerization of mannose-6-phosphate to fructose-6-phosphate (Negrotto et al., 2000).
	8605–8664	Intervening sequence	60	DNA sequence used for cloning.
	8665–8917	NOS 3'	253	Termination sequences of nopaline synthase (*nos*) gene from *Agrobacterium tumefaciens* (Depicker et al., 1982); GenBank accession no. V00087.
	8918–9005	Ti plasmid region	88	Non-functional sequence from Ti plasmid of *A. tumefaciens*.
	9006–9030	Left border	25	Left border region of the T-DNA from *A. tumefaciens* nopaline Ti plasmid (Zambryski et al., 1982); GenBank accession no. J01825. Short direct repeat that flanks the T-DNA and is required for the transfer of the T-DNA into the genome of the host plant (Yadav et al., 1982).

Table 3. Generations and comparators used for analysis of GR2E rice

Analysis	Generation(s)[a]	Experimental Control and Comparators
Molecular genetic characterization	Inbred Kaybonnet germplasm containing event GR2E (direct line of descent from T0 original transformant)	Non-transgenic (unmodified) Kaybonnet rice
Stability of the inserted DNA over multiple generations	Tn (Kaybonnet), BC_3F_5*1,*2,*3, BC_4F_3*1,*2,*3, and BC_5F_3*1,*2,*3	Not applicable
Segregation of the inserted DNA within multiple generations	BC_4F_2*1,*2,*3, BC_5F_1*1,*2,*3, and BC_5F_2*1,*2,*3	Not applicable
Protein expression analysis; *Zm*PSY1, CRTI, PMI	BC_5F_3*1 and BC_5F_4*1	Non-transgenic (unmodified) PSB Rc82 rice
Nutrient composition analysis	BC_5F_3*1 and BC_5F_4*1	Non-transgenic (unmodified) PSB Rc82 rice

a Generations of plant material designated with the superscripts "*1", "*2", or "*3", represent breeding lines where the recurrent back-crossed parent is PSB Rc82, BRRI *dhan* 29, or IR64, respectively.

The predicted fragment sizes following *Hin*dIII digestion are >3020 bp, corresponding to the 5' proximal region spanning the junction of the pSYN12424 T-DNA insert and the rice genome, and >6010 bp, corresponding to the 3' proximal region spanning the junction of the T-DNA insert and the rice genome. Thus, the insertion of a single copy of the pSYN12424 T-DNA in GR2E rice should result in a single detectable hybridization fragment of >3020 bp following *Hin*dIII digestion and hybridization with the *pSSU-crtI* gene probe, and >6010 following hybridization with either the *Zmpsy1* or *pmi* probes (Table 4).

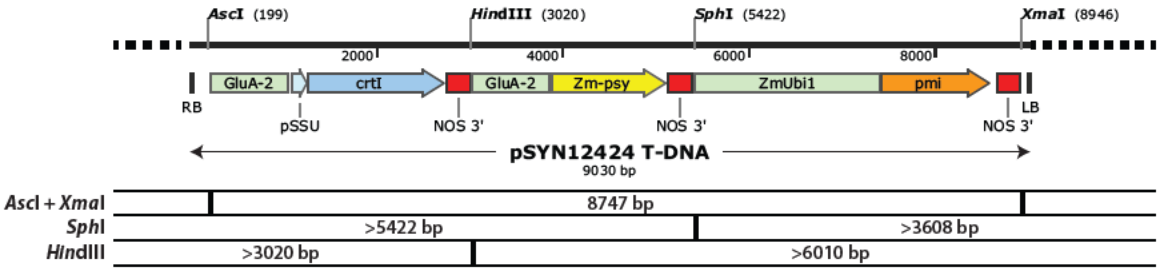


Figure 6. Schematic map of the pSYN12424 T-DNA showing the *Asc*I, *Hin*dIII, *Sph*I, and *Xma*I restriction endonuclease sites used for Southern hybridization characterization of GR2E rice. The flanking rice genome is represented by the horizontal dotted line. Based on the insertion of a single copy of the T-DNA at a single site, the predicted sizes of restriction fragments following digestion with *Asc*I plus *Xma*I (8747 bp), *Sph*I (>6818 bp and >5355 bp), or *Hin*dIII (>4985 bp and >7787 bp) are shown. Numbering shown on the map is relative to the pSYN12424 T-DNA.

The predicted fragment sizes following *Sph*I digestion are >5422 bp, corresponding to the 5' proximal region spanning the junction of the T-DNA insert and the rice genome, and >3608 bp, corresponding to the 3' proximal region spanning the junction of the T-DNA insert and the rice genome. Insertion of a single copy of the pSYN12424 T-DNA in GR2E rice should result in a single detectable hybridization fragment of >5422 bp following *Sph*I digestion and hybridization with the *pSSU-crtI* or *Zmpsy1* probes, and >3608 bp following hybridization with the *pmi* probe (Table 4).

Hybridization of *Hin*dIII-digested genomic DNA from individual plants of event GR2E in four genetic backgrounds (Kaybonnet, PSB Rc82, BRRI *dhan* 29, and IR64) with the *Zmpsy1* or *pmi* probes resulted in detection of a single fragment of *ca*. 7900 bp, and hybridization with the *pSSU-crtI* probe resulted in detection of a single *ca*. 7200 bp fragment (Table 4; Figure 7, lanes 16–19). Hybridization of *Sph*I-digested GR2E genomic DNA with the

Table 4. Correlation of predicted and observed fragment sizes based on a single site of insertion of the pSYN12424 T-DNA within GR2E rice

		Event GR2E Genomic DNA Fragment Size (bp)		Plasmid pSYN12424 Fragment Size (bp)	
Probe	**Enzyme**	**Predicted**	**Observed[a]**	**Predicted**	**Observed**
Zmpsy1	*Asc*I + *Xma*I	8747	8747	–[c]	–
	*Sph*I	>5422	~6900 ~5600[b]	5504	5504
	*Hin*dIII	>6010	~7900 ~4900[b]	–	–
pSSU-crtI	*Asc*I + *Xma*I	8747	8747	–	–
	*Sph*I	>5422	~6900	5504	5504
	*Hin*dIII	>3020	~7200	–	–
pmi	*Asc*I + *Xma*I	8747	8747	–	–
	*Sph*I	>3608	~5500	8835	8835
	*Hin*dIII	>6010	~7900	–	–

[a] Observed fragment sizes were approximated from the digoxigenin (DIG)-labelled DNA molecular size marker VII fragments on the Southern blots. Due to the inability to determine exact sizes on the blot, all values are approximations to the nearest 200 bp. The exception to this was when the observed fragment had equal migration as the control plasmid hybridization fragment, in which case the predicted size was assigned (e.g., 5504 bp in the case of *Sph*I digestions probed with DIG-labelled *Zmpsy1* or *pSSU-crtI*).
[b] These fragments were due to hybridization with sequences from the endogenous rice *psy* gene, which is *ca.* 83 percent identical to the *Zmpsy1* gene (Figure 32, page 108).
[c] Enzymatic digestion of plasmid DNA not performed.

Zmpsy1 or *pSSU-crti* probes resulted in detection of a single fragment of *ca*. 6900 bp, and hybridization with the *pmi* probe resulted in detection of a single fragment of *ca*. 5500 bp (Figure 7, lanes 6–9; Table 4).

Weak hybridization between the *Zmpsy1* probe and sequences derived from the endogenous rice *psy1* gene was detected for restriction enzyme digests of control Kaybonnet and event GR2E genomic DNA samples (Figure 7, panel A). Hybridizing fragments of *ca*. 5600 bp and *ca*. 4900 bp were detected in Southern blots of *Sph*I- and *Hin*dIII-digested DNA samples, respectively. This was not an unexpected finding considering the high degree of sequence identity, *ca*. 83 percent, shared between the *Zmpsy1* and *Oryza sativa psy1* genes (Figure 32, page 108).

A.3.4.2 Integrity of the Inserted DNA in GR2E Rice

Southern blot analyses of *Asc*I+*Xma*I digested GR2E rice genomic DNA were used to confirm the integrity of the pSYN12424 T-DNA insert containing the *Zmpsy1*, *pSSU-crtI*, and *pmi* gene cassettes. The pSYN12424 T-DNA contains a single *Asc*I restriction site located at position 199 and a single *Xma*I site at position 8946 (Figure 6). Insertion of an intact copy of the pSYN12424 T-DNA was predicted to result in the detection of a 8747 bp *Asc*I+*Xma*I fragment with the *Zmpsy1*, *pSSU-crtI*, and *pmi* probes (Table 4).

The results of Southern analyses (Figure 7, lanes 11–14) demonstrated that the predicted 8747 bp fragment was detected for each of the hybridization probes tested (Table 4).

A.3.4.3 Absence of Plasmid Backbone DNA in GR2E Rice

Southern blot analysis of *Asc*I+*Xma*I-digested genomic DNA obtained from event GR2E in Kaybonnet germplasm was performed to demonstrate the lack of integration of any sequences derived from the pSYN12424 plasmid backbone. For this analysis, five different

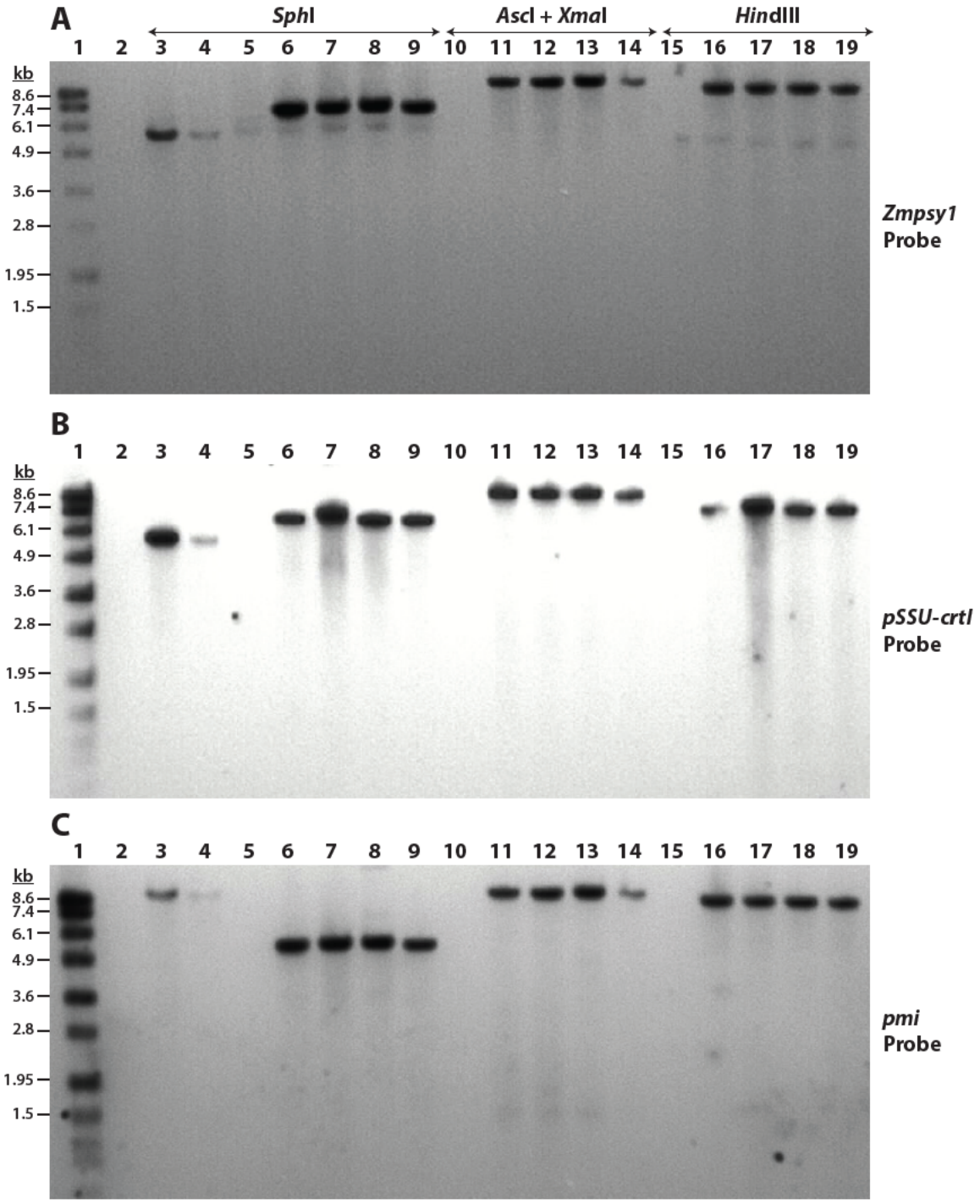


Figure 7. Samples of genomic DNA (*ca.* 7 µg) from individual plants of event GR2E in Kaybonnet (T_n; lanes 6, 11, and 16), BR29 (BC_5F_3; lanes 7, 12, and 17), IR64 (BC_5F_3; lanes 8, 13, and 18), and PSB Rc82 (BC_5F_3; lanes 9, 14, and 19) germplasm backgrounds; negative control DNA from Kaybonnet rice (lanes 5, 10, and 15); and negative control Kaybonnet rice containing either *ca.* one (lane 3) or 0.2 (lane 4) copy equivalents of pSYN12424 plasmid DNA, were subjected to restriction endonuclease digestion with *Sph*I (lanes 3–9), *Asc*I plus *Xma*I (lanes 10–14), or *Hin*dIII (lanes 15–19) followed by agarose gel electrophoresis and transfer onto nylon membrane. Blots were hybridized with DIG-labelled probes specific for *Zmpsy1* (panel A), *pSSU-crtI* (panel B), or *pmi* (panel C). Following washing, hybridized probes and DIG-labelled molecular weight markers VII (lane 1) were visualized using a chemiluminescent detection system followed by electronic image capture. Lane 2 was blank on all gels.

probes, which together corresponded to the entire plasmid backbone region (Figure 30; Table 4), were utilized.

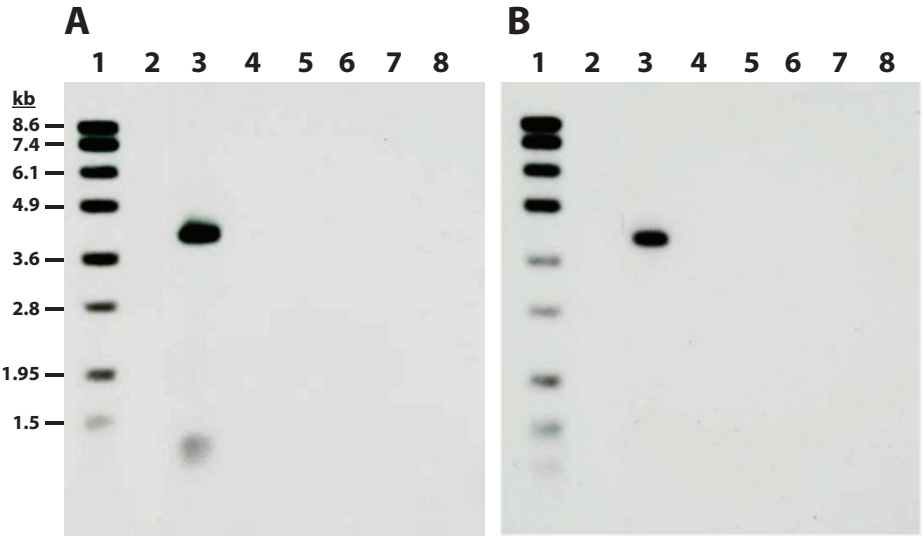


Figure 8. Samples of genomic DNA (*ca.* 7 µg) from control Kaybonnet rice (lanes 4 and 5) , event GR2E in Kaybonnet (T_n; lanes 6–8), and negative control Kaybonnet rice containing three (lane 3) copy equivalents of pSYN12424 plasmid DNA, were subjected to restriction endonuclease digestion with *Asc*I plus *Xma*I followed by agarose gel electrophoresis and transfer onto nylon membrane. Blots were hybridized with the mixture of DIG-labelled backbone probes 1–4 (panel A) or backbone probe 5 (panel B). Following washing, hybridized probes and DIG-labelled molecular weight markers VII (lane 1) were visualized using a chemiluminescent detection system followed by electronic image capture. Lane 2 was blank on all gels.

Hybridization of backbone probe 5 with *Asc*I+*Xma*I-digested pSYN12424 plasmid DNA would be expected to yield a single hybridizing fragment of 4349 bp, while hybridization of the mixture of backbone probes 1–4 with the same digested plasmid DNA would be expected to yield two hybridizing fragments of 1243 bp (due to hybridization with backbone probe 1) and 4349 bp.

Hybridizing fragments were not detected when the backbone probes were tested against samples of *Asc*I+*Xma*I-digested GR2E rice genomic DNA (Figure 8, lanes 6–8, panels A and B), confirming the lack of integration of any plasmid backbone sequences. Positive control samples of non-transgenic Kaybonnet genomic DNA spiked with pSYN12424 plasmid DNA did result in detection of the expected-size 4349 bp fragment using backbone probe 5 (Figure 8, panel B, lane 3) and two fragments of 1243 bp and 4349 bp, respectively, using the mixture of backbone probes 1–4 (Figure 8, panel A, lane 3).

A.3.4.4 Nucleotide Sequence Analysis of the Inserted DNA and Flanking Regions

The nucleotide sequence of the entire pSYN12424 T-DNA insert present in GR2E rice, including a portion of the 5' and 3' flanking host genomic region, was determined in order to demonstrate overall integrity of the insert, contiguity of the functional elements, and to detect any individual base-pair changes (Trijatmiko et al., 2015). In addition, an open reading frame (ORF) analysis was conducted to investigate the possibility of creating any new start-to-stop ORFs spanning the 5' or 3' junctional regions that could potentially encode sequences homologous to known allergens or toxins.

The nucleotide sequence of the plasmid pSYN12424 T-DNA, together with preliminary sequence information from the 5' and 3' flanking genomic DNA (Syngenta), was used to design seven sets of oligonucleotide primers that were used to amplify the insert and flanking regions from GR2E rice genomic DNA as seven individual overlapping fragments (Figure 9 and see Appendix B, page 109).

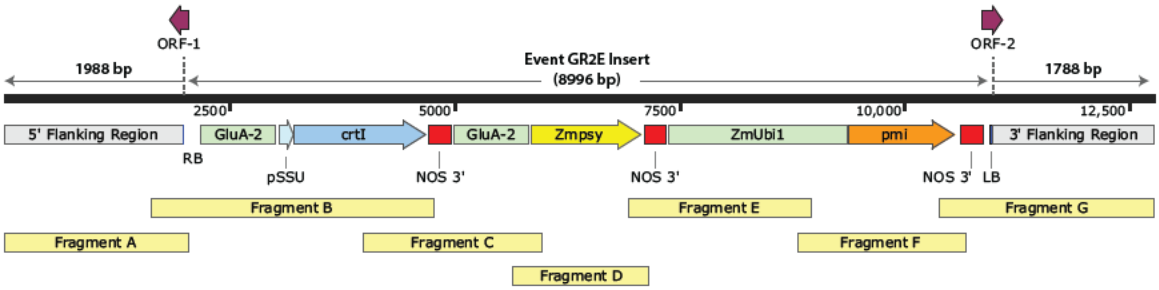


Figure 9. Schematic diagram of the GR2E T-DNA insert indicating the *crtI*, *Zmpsy1*, and *pmi* genes along with their respective regulatory elements. Approximate locations of the seven overlapping polymerase chain reaction (PCR) fragments subjected to nucleotide sequencing are indicated as A through G. The location of two putative ORFs spanning the 5' and 3' junctions of the T-DNA insert and the rice genomic DNA are indicated as ORF-1 and ORF-2, respectively.

In total, 12,772 bp of GR2E rice genomic sequence were confirmed, comprising 1,988 bp of the 5' genomic border sequence, 1,788 bp of the 3' genomic border sequence, and 8,996 bp of the inserted T-DNA (Figure 9). The inserted T-DNA in GR2E rice was found to have a 23 bp deletion on the right border (RB) end and a 11 bp deletion on the left border (LB) end, which is not uncommon for *Agrobacterium*-mediated transformation events (Kim et al., 2007). All remaining sequence was intact and identical to that of the T-DNA region of plasmid pSYN12424.

Basic local alignment search tool (BLAST) searches using the 5' and 3' flanking region sequences as queries against the *O. sativa* (japonica cultivar-group, Nipponbare) genome (MSU Rice Genome Annotation Project Release 7) identified that the site of insertion of the T-DNA was located on chromosome 3 within the intergenic region between LOC_Os03g43980 (3' proximal) and LOC_Os03g43990 (5' proximal) (Figure 10).

Bioinformatics analyses of the two ORFs spanning the 5' and 3' junctions of the T-DNA insert and the rice genomic DNA were conducted to investigate amino acid sequence similarities with known toxins and allergens (section B.1.7, page 52).

A.3.4.5 Stability of the Inserted DNA across Multiple Generations

The stability of the inserted DNA across multiple generations of GR2E rice was assessed by Southern blot analyses of genomic DNA samples prepared from the T_n generation (Kaybonnet) and the BC_3F_5, BC_4F_3, and BC_5F_3 generations of GR2E in BRRI *dhan* 29, IR64, and PSB Rc82 germplasm. Restriction enzyme digestions with *Hin*dIII, *Sph*I, and *Asc*I+*Xma*I were separated by agarose gel electrophoresis and blots were probed with DNA probes specific for the *Zmpsy1*, *pSSU-crtI*, or *pmi* genes, respectively (Figure 11).

Single hybridizing fragments of ~7900 bp, ~6900 bp, or 8747 bp were detected using the *Zmpsy1*, *pSSU-crtI*, or *pmi* probes, respectively, in corresponding Southern blots of *Hin*dIII, *Sph*I, or *Asc*I+*Xma*I digests of genomic DNA from each generation of GR2E rice.

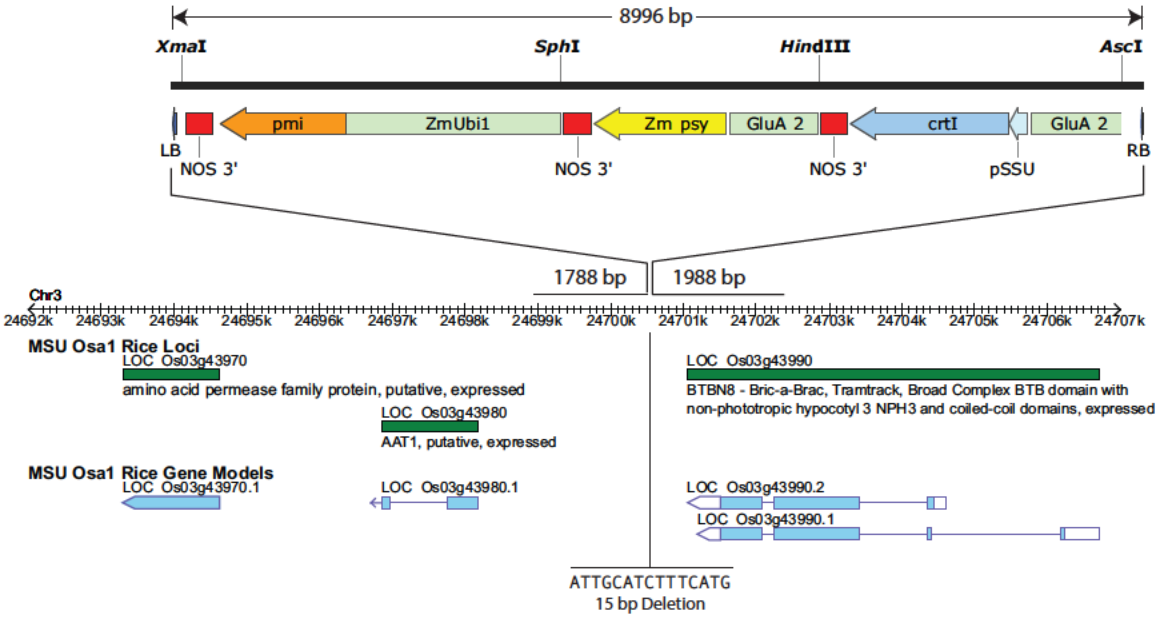


Figure 10. Map position is indicated according to the MSU Rice Genome Annotation Project Release 7 (Nipponbare). The locations of the LB and RB flanking sequences correspond to positions 24,698,762–24,700,549 and 24,700,565–24,702,552, respectively. The insertion of the pSYN12424 T-DNA was within an intergenic region between loci LOC_Os03g43980 and LOC_Os03g43990, and resulted in the deletion of 15 bp of host genomic DNA in addition to truncations of the LB and RB regions of 11 bp and 23 bp, respectively.

The observation of consistent hybridization patterns for each tested probe across each of the plant generations confirmed stable integration and inheritance of the inserted DNA in GR2E rice.

A.3.4.6 Stability of the Introduced Trait across Multiple Generations

In order to demonstrate the stability and inheritance of the elevated β-carotene trait in GR2E rice, the concentrations of total carotenoids were measured in samples of grain harvested from different breeding generations and germplasm backgrounds (Swamy and Samia, 2016).

Concentrations of total carotenoids were determined in rice grain samples collected from event GR2E plants in Kaybonnet (T_n) germplasm, the BC_3F_5 generations in PSB Rc82 and IR64 backgrounds, and the BC_4F_3 and BC_5F_3 generations in PSB Rc82, IR64, and BRRI *dhan* 29 germplasm backgrounds (Table 5).

Table 5. Concentrations of total carotenoids in different generations and germplasm backgrounds of GR2E rice

	Total Carotenoids (µg/g FWT)[†]			
Breeding Generation	**Kaybonnet**	**PSB Rc82**	**IR64**	**BRRI *dhan* 29**
T(n)	30.50 ± 2.49	–	–	–
BC_3F_5	–	12.84 ± 3.32	20.11 ± 8.71	ND[a]
BC_4F_3	–	9.09 ± 2.26	19.40 ± 2.00	24.50 ± 2.85
BC_5F_3	–	14.12 ± 0.11	13.23 ± 1.16	29.33 ± 3.14

[†] Values shown are mean values ± SD (standard deviation) determined spectrophotometrically for total carotenoids in grain samples after one day of storage at 16°C. All concentrations are on a fresh weight of tissue (FWT) basis, not corrected for moisture content.
[a] ND = Not determined. Due to poor quality of remnant seed from the BC_3F_5 generation of BRRI *dhan* 29 containing event GR2E, there was no seed germination and plants could not be produced for grain sampling.

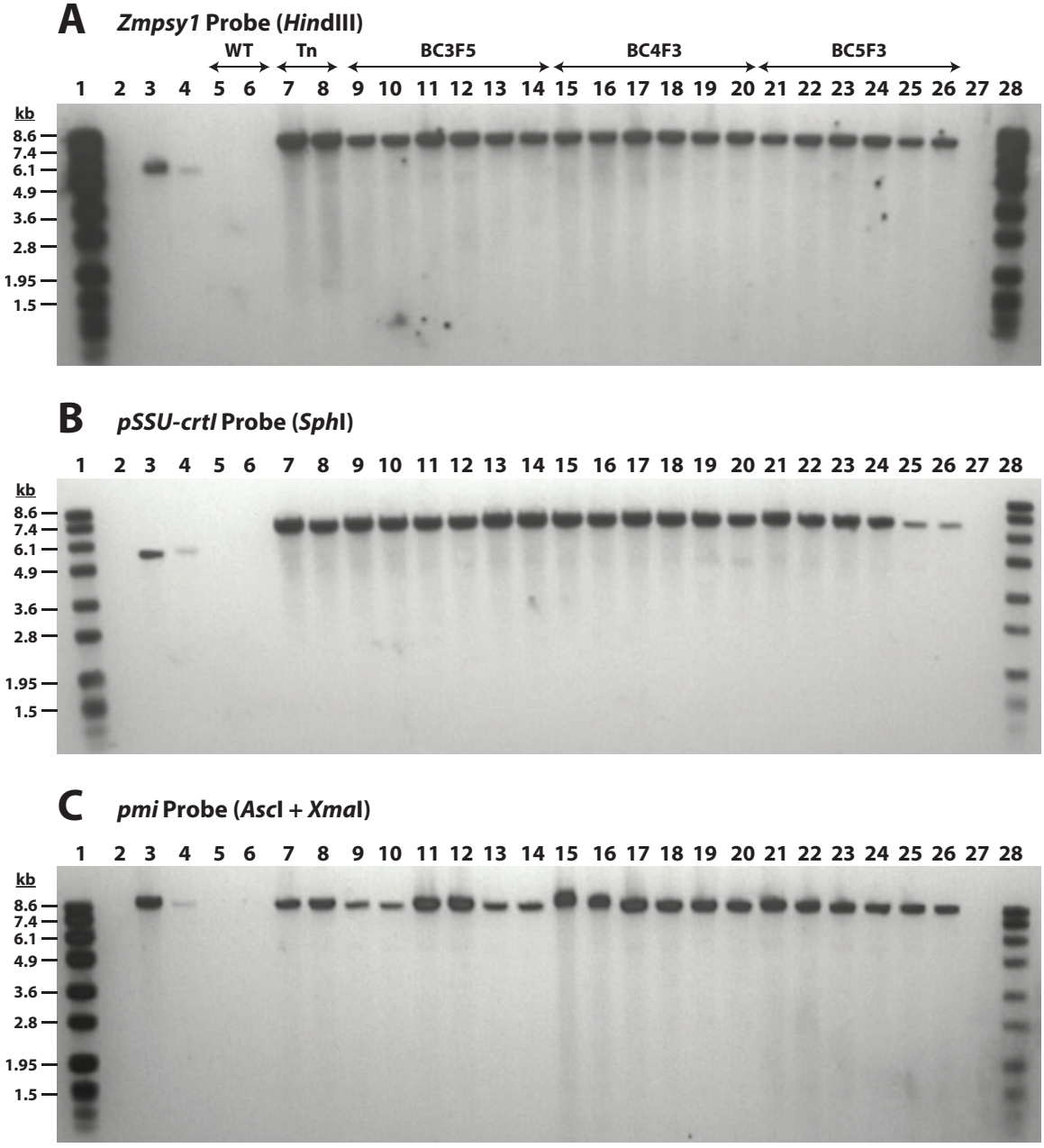


Figure 11. Samples of genomic DNA (*ca.* 5 µg) from individual plants of event GR2E in Kaybonnet (T_n; lanes 7–8), BR29 (BC_3F_5, lanes 9–10; BC_4F_3, lanes 15–16; and BC_5F_3, lanes 21–22), IR64 (BC_3F_5, lanes 11–12; BC_4F_3, lanes 17–18; and BC_5F_3, lanes 23–24), and PSB Rc82 (BC_3F_5, lanes 13–14; BC_4F_3, lanes 19–20; and BC_5F_3, lanes 25–26) germplasm backgrounds; and negative control DNA from Kaybonnet rice (lanes 5–6) were subjected to restriction endonuclease digestion with *Hin*dIII (panel A), *Sph*I (panel B), or *Asc*I plus *Xma*I (panel C) followed by agarose gel electrophoresis and transfer onto nylon membrane. Positive control samples consisted of negative control Kaybonnet rice containing either *ca.* one (lane 3) or 0.2 (lane 4) copy equivalents of pSYN12424 plasmid DNA that were digested with *Sph*I (panels A and B) or *Asc*I+*Xma*I (panel C), Blots were hybridized with DIG-labelled probes specific for *Zmpsy1* (panel A), *pSSU-crtI* (panel B), or *pmi* (panel C). Following washing, hybridized probes and DIG-labelled molecular weight markers VII (lanes 1 and 28) were visualized using a chemiluminescent detection system followed by electronic image capture. Lanes 2 and 27 were blank on all gels.

Carotenoid expression in the rice endosperm was positively correlated with the presence of the pSYN12424 T-DNA insert as previously established by Southern blot characterization of genomic DNA samples prepared from the same generations and germplasm backgrounds of GR2E rice (Cueto et al., 2016). Some variation in the concentrations of total carotenoids was observed depending on the germplasm background, with the highest concentrations found in Kaybonnet and in BRRI *dhan* 29 containing event GR2E.

The demonstration of carotenoid accumulation in the endosperm of GR2E rice samples collected from different germplasm backgrounds across four breeding generations (T_n, BC_3F_5, BC_4F_3, and BC_5F_3) provided further evidence that the introduced phenotypic trait was stable and inherited across multiple generations.

A.3.4.7 Mendelian Inheritance of the Inserted DNA

The inheritance pattern of the pSYN12424 T-DNA insert within GR2E rice was investigated using a PCR-based zygosity test to determine segregation of the insert within three segregating generations ($BC_4F_2^{*1,*2,*3}$, $BC_5F_1^{*1,*2,*3}$, and $BC_5F_2^{*1,*2,*3}$) (Swamy and Samia, 2015). Generations identified with the superscript "*1", "*2", and "*3" designations represented different GR2E breeding lines where the recurring non-transgenic parent was PSB Rc82, BRRI *dhan* 29, and IR64, respectively (Figure 5 and see Appendix A, page 106).

Chi-square (X^2) analysis of the data found no statistically significant differences between the observed and expected segregation ratios for the three segregating generations of GR2E rice in PSB Rc82, BRRI *dhan* 29, and IR64 genetic backgrounds (Table 6).

Table 6. Mendelian inheritance of the pSYN12424 T-DNA insert within multiple generations of GR2E rice

		Expected		Observed[b]			
Generation[a]	Total	Present	Absent	Present	Absent	Chi Square[c]	p-Value
BC4F2*1 (3:1)	64	48	16	44	20	1.333	0.2482
BC4F2*2 (3:1)	51	38.25	12.75	38	13	0.007	0.9356
BC4F2*3 (3:1)	63	47.25	15.75	43	20	1.529	0.2162
BC5F1*1 (1:1)	59	29.5	29.5	32	27	0.424	0.5151
BC5F1*2 (1:1)	49	24.5	24.5	25	24	0.020	0.8864
BC5F1*3 (1:1)	127	63.5	63.5	68	59	0.638	0.4245
BC5F2*1 (3:1)	100	75	25	74	26	0.053	0.8174
BC5F2*2 (3:1)	99	74.25	24.75	72	27	0.273	0.6015
BC5F2*3 (3:1)	100	75	25	70	30	1.333	0.2482

[a] Generations identified with the superscript "*1", "*2", and "*3" designations represented different GR2E breeding lines where the recurring non-transgenic parent was PSB Rc82, BRRI *dhan* 29, and IR64, respectively.

[b] Multiplex PCR-based zygosity testing was performed on each plant from each generation.

[c] The analysis tested the hypothesis that the introduced DNA was segregating in a Mendelian fashion. The critical value to reject the hypothesis at the five percent level is 3.84. Since the Chi-square (X^2) square value was less than 3.84 within each generation, the observed differences were not statistically significant.

The results from these analyses were consistent with Southern hybridization data indicating the stable integration of the pSYN12424 T-DNA at a single site within the GR2E rice genome and segregation of the introduced DNA as a single genetic locus according to Mendelian rules of inheritance.

A.3.4.8 Antibiotic Resistance Markers

No antibiotic resistance marker genes are present in rice event GR2E. Analysis for the presence of plasmid backbone sequences (section A.3.4.3) demonstrated that no plasmid backbone was incorporated into the rice genome during transformation.

A.3.4.9 Conclusions from Molecular Characterization

Based on a comprehensive molecular characterization of GR2E rice, one copy of the pSYN12424 T-DNA was introduced at a single site within the rice genome and stably inherited over multiple generations as a single genetic locus according to Mendelian rules of inheritance. In addition, nucleotide sequencing of the entire inserted DNA, including portions of the 5' and 3' flanking rice genomic sequence, confirmed that the T-DNA had been inserted without modifications, deletions, or rearrangements, except for small truncations at the 5' and 3' termini of 23 bp and 11 bp, respectively. There were also no new novel open reading frames created as a consequence of the DNA insertion that would have the potential to encode a protein with any significant amino acid sequence similarity to known and putative toxins or allergens.

B. Information Related to the Safety of the GM Food

B.1. Characterization and Safety Assessment of New Substances

B.1.1. Biochemical Function of Newly Expressed Proteins

B.1.1.1 ZmPSY1 Protein

Carotenoid compounds play essential roles in plants such as protecting the photosynthetic apparatus and in hormone signalling. Coloured carotenoids provide many of the yellow, orange and red colours observed in plant tissues, as well as offering nutritional benefit to humans and animals. The plant carotenoid biosynthetic pathway is localized in the plastid and has been molecularly elucidated, as reviewed by DellaPenna and Pogson (2006). It diverges from C_3 carbon metabolism by the action of the enzyme deoxyxylulose phosphate synthase, followed by a series of enzymes of the methylerythritol 4-phosphate pathway, yielding isopentenyl diphosphate (IPP), C_5, and its isomer dimethylallyl diphosphate (DMAPP). The addition of IPP units to DMAPP generates prenyl diphosphate molecules of increasing size such as geranyl diphosphate (GPP), C_{10}, farnesyl diphosphate (FPP), C_{15}, and GGPP, C_{20}. The condensation of two molecules of geranylgeranyl diphosphate to yield the first carotenoid, C_{40} 15-*cis*-phytoene, is catalyzed by phytoene synthase (PSY) (Figure 12).

Phytoene synthase plays a pivotal role in the carotenoid pathway as the first committed step and acts to control flux through the pathway (Welsch et al., 2010; Cazzonelli and Pogson, 2010). Phytoene undergoes consecutive modifications such as desaturation and isomerization to form lycopene, which is later cyclized to α- and β-carotene.

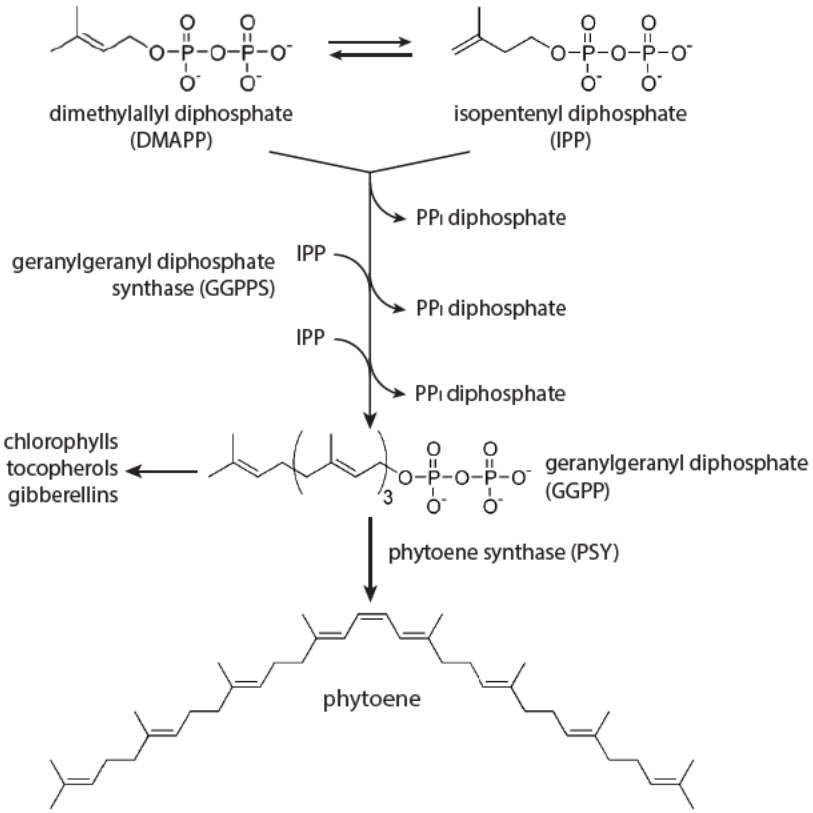


Figure 12. The phytoene biosynthetic pathway.

B.1.1.2 CRTI Protein

Carotenoids are a structurally diverse class of isoprenoids synthesized by all photosynthetic organisms, where they serve in light harvesting and photoprotection, and by many nonphotosynthetic organisms, such as certain species of bacteria, fungi, and archaea. As judged by sequence homology, carotene desaturation evolved at least twice (Figure 13) (Schaub et al., 2012). Cyanobacteria and plants employ a complex, multi-component pathway relying on two desaturases, namely phytoene desaturase (PDS) and ζ-carotene desaturase (ZDS), that form specific poly-*cis* configured carotene intermediates necessitating the participation of two *cis-trans* isomerases, ζ-carotene *cis-trans* isomerase (Z-ISO) and carotene *cis-trans* isomerase (CRTISO).

However, in nonphotosynthetic bacteria and fungi, carotenoid desaturation is completed solely by carotene desaturase I (*crtI*; syn. phytoene desaturase), which has the property of catalyzing the *cis*-to-*trans* conversion of carotenes (Sandmann, 2009). In plants, lycopene is the substrate of two competing cyclases: ε-cyclase (LCY-e) and β-cyclase (LCY-b), acting together on the two ends of the molecule, lead to the formation of α-carotene, whereas the action of LCY-b alone forms β-carotene. The bacterial homolog of LCY-b is CRTY (Figure 13).

The structure and function of a bacterial CRTI derived from *Pantoea ananatis*, over-expressed in *Escherichia coli*, was described by Schaub et al. (2012). The crystal

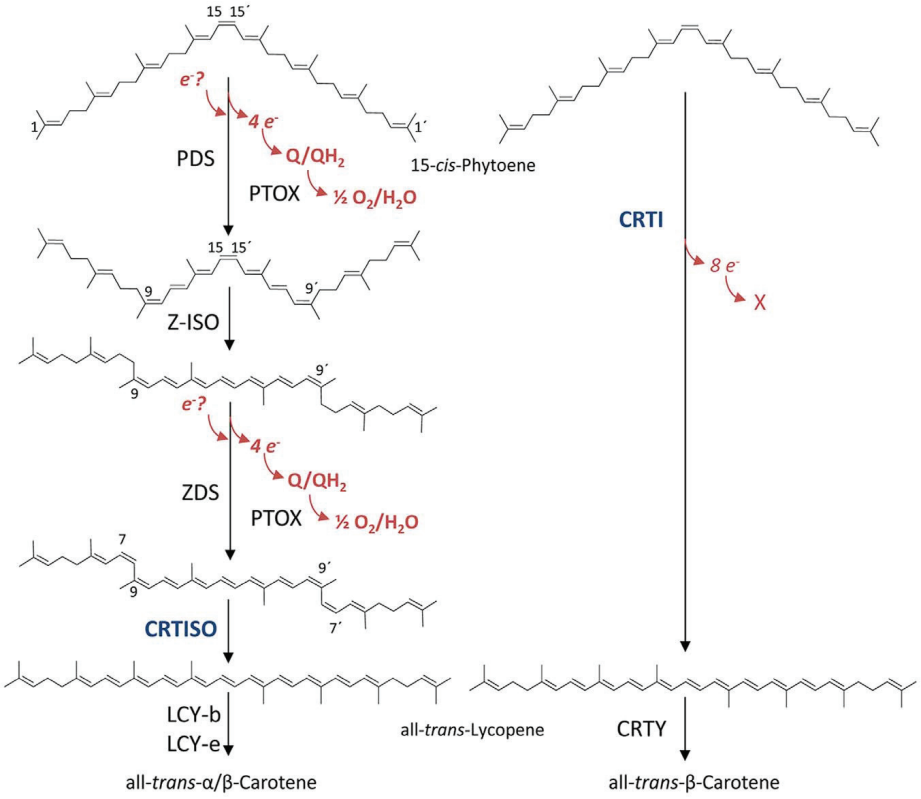


Figure 13. **Left**, the plant cyanobacterial system consisting of the two desaturases, phytoene desaturase (PDS) and ζ-carotene desaturase (ZDS). The pathway involves specific poly-*cis*-intermediates and results in the formation of 7,9,9′7′-tetra-*cis*-lycopene (i.e., prolycopene). *Cis-trans* isomerases act at the 9,15,9′-tri-*cis*-ζ-carotene (Z-ISO) and prolycopene (CRTISO) stage, the latter forming all-*trans*-lycopene, the substrate for lycopene cyclases.
Right, CRTI-mediated phytoene desaturation encompassing all four desaturation steps and one *cis-trans* isomerization step to form all-*trans*-lycopene. The desaturase CRTI and the isomerase CRTISO share amino acid sequence similarity. Figure from Schaub et al. (2012).

structure of apo-CRTI shows that it belongs to the flavoprotein superfamily comprising protoporphyrinogen IX oxidoreductase and monoamine oxidase. The *E. coli*-expressed CRTI exhibited high enzymatic activity with the lipophilic substrate 15-*cis*-phytoene contained in phosphatidyl-choline liposomal membranes and had an absolute requirement for flavin adenine dinucleotide (FAD) as the sole cofactor effective in CRTI-mediated phytoene desaturation.

Although CRTI from *Pantoea ananatis* and *Oryza sativa* phytoene desaturase (OsPDS) share only 22 percent amino acid sequence similarity and 11 percent identity, much of which is attributed to the FAD-binding Rossmann fold (Rao and Rossmann, 1973) common to both, and although there are clear differences in catalysis (e.g., differential roles of oxygen, quinones, stereochemistry and the number of double bonds introduced), the overall protein folds are quite similar (Gemmecker et al., 2015). Thus CRTI and OsPDS are members of the same structural grouping along with monoamine oxidases and protophorphyrinogen oxidases based on sequence comparisons of extended Rossmann fold domains (Dailey and Dailey, 1998) and homology modelling (Suarez et al., 2014). The two desaturases have

likely evolved divergently or convergently resulting in two different approaches to achieving similar catalytic goals (Schaub et al., 2012).

B.1.1.3 PMI Protein

Phosphomannose isomerase (PMI, EC 5.3.1.8) catalyzes the reversible interconversion of mannose-6-phosphate and fructose-6-phosphate. The enzyme has been successfully used as a selectable marker for the transformation of many plant species (Reed et al., 2001; Todd and Tague, 2001) as expression of the *E. coli pmi* gene (also referred to as *manA* gene) allows the cells to survive on medium containing only mannose as a carbon source. The *pmi* gene has been cloned from several bacteria and yeast species, as well as from humans (Miles and Guest, 1984; Proudfoot et al., 1994a) and the enzyme has been purified and characterized from yeast, bacteria, pigs, and humans (Proudfoot et al., 1994b,a). Lack of the active enzyme in humans is associated with carbohydrate-deficient glycoprotein syndrome whose symptoms include hereditary fructose intolerance, galactosaemia, and hyperinsulinemic hypoglycemia (Keir et al., 1999; de Lonlay and Seta, 2009).

B.1.2. Prior History of Consumption

B.1.2.1 ZmPSY1 Protein

In maize, *PSY1*, encoded by the *Yellow1* (*Y1*) locus, is required for endosperm carotenoid accumulation, which has also been correlated with the presence of *Zm*PSY1 transcripts in the endosperm (Buckner et al., 1996; Gallagher et al., 2004). Using antiserum developed against *E. coli*-expressed *Zm*PSY1, Li et al. (2008) demonstrated the accumulation of *Zm*PSY1 enzyme and localization in maize endosperm (Figure 14).

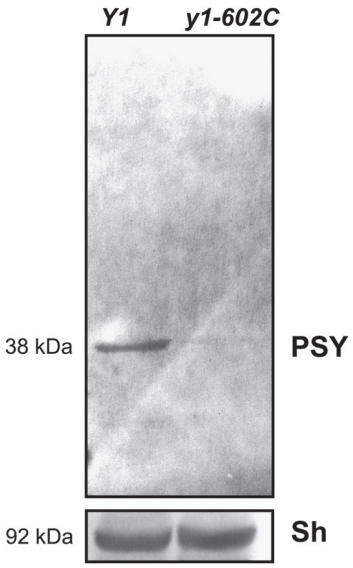


Figure 14. Western analysis demonstrating specificity of anti-PSY1 antiserum for PSY in maize endosperm. Proteins were extracted from yellow (*Y1*) and white (*y1*) endosperm at 20 days after pollination. Immunoblots were probed by anti-PSY1 antiserum (top) or anti-Shrunken1 (Sh) antiserum (bottom). Figure reproduced from Li et al. (2008).

Maize kernels can be different colours ranging from white to yellow to red to black. Most of the maize grown in the United States is yellow, whereas people in Africa, Central America, and the southern United States prefer white maize (Ranum et al., 2014). The nutritional value of yellow and orange maize is higher than white maize due to its higher levels of lutein, β-carotene, and β-cryptoxanthin, and is a healthy source of pro-vitamin A (Muzhingi et al., 2011). Conventional breeding programs to select maize lines with elevated levels of β-carotene (*ca*. 18 μg/g) are being pursued to combat VAD in sub-Saharan Africa (Gannon et al., 2014).

Thus, there is a history of likely human and animal dietary exposure to *Zm*PSY1 protein through maize consumption.

B.1.2.2 CRTI Protein

The source of the CRTI-encoding gene was *Pantoea ananatis*, a ubiquitous bacterium that is sometimes associated with meat spoilage (Ercolini et al., 2006) and found on fresh fruit and vegetables (Leff and Fierer, 2013). Consequently, there is the possibility of some dietary exposure to CRTI protein at low levels as a consequence of adventitious presence of *P. ananatis* on foods. However, a history of dietary exposure is difficult to establish.

B.1.2.3 PMI Protein

It is conceivable, and indeed likely, that small amounts of PMI proteins from various sources have always been present in the food and feed supply due to the ubiquitous occurrence of PMI proteins in nature, including food plants and animals. PMI proteins have been found in such diverse plant species as tobacco, walnut (Malvolti et al., 1993), and *Brassica* species (Chen et al., 1989), as well as in seeds of soybeans and other legumes (Lee and Matheson, 1984). Human PMI, which has significant structural and functional similarity to the *E. coli*-derived PMI expressed in GR2E rice, exists in several body tissues and is an essential enzyme. PMI is present in naturally occurring *E. coli* strains and other gram-negative bacteria that colonize the human intestine, and these PMI proteins are expected to be identical or highly homologous to the PMI present in GR2E rice.

The PMI protein expressed in GR2E rice is identical in amino acid sequence to the PMI protein expressed in maize events MIR162, 3272, and 5307.[1] The safety of the PMI protein was first reviewed by the US Food and Drug Administration as part of a new protein consultation (FDA, 2005). Since then, PMI has been reviewed multiple times by regulatory authorities in at least 21 countries in the context of reviewing the food and/or feed, or environmental, safety of a number of genetically engineered maize events, including MIR604, MIR162, 3272, and 5307 (Table 7). In some countries, additional reviews of PMI have occurred during regulatory reviews of breeding stacks containing one or more of these single events.

Based on the extensive history of prior regulatory review of the safety of the PMI protein (Table 7), new data and information are not directly provided with this submission.

[1] As a consequence of the genetic modification resulting in event MIR604, two nucleotide sequence changes occurred within the *pmi* gene resulting in two amino acid changes; a valine to alanine substitution at position 61 and a glutamine to histidine substitution at position 210 (FSANZ, 2006). Because of these changes, the amino acid sequence of the PMI protein in GR2E rice was not identical to the corresponding sequence of the PMI protein expressed in MIR604 maize.

Table 7. Regulatory authorizations for single events that included a review of the safety of PMI protein

	Food and/or Feed[†]				Environment			
Country/Region	**MIR604**	**MIR162**	**3272**	**5307**	**MIR604**	**MIR162**	**3272**	**5307**
Argentina	2012	2010			2012	2011		
Australia	2006	2009	2008	2012				
Brazil	2014	2009			2014	2009		
Canada	2007	2010	2008	2013	2007	2010	2008	2012
China	2008	2014	2013					
Colombia	2012	2012/2010[a]	2013*	2014/2013	2014			
European Union	2009	2012						
Indonesia	2011	2011	2011					
Japan	2007	2010	2010	2013	2007	2010	2010	2013
Korea	2007/2008	2010	2011	2013*				
Malaysia	2016	2016	2016	2016				
Mexico	2007	2010	2008	2013				
Paraguay						2014		
Philippines	2007	2010	2008	2015				
Russian Federation	2007/2008	2011/2012	2010	2014				
South Africa	2007/2011	2014						
Taiwan	2007	2009	2010	2012				
Thailand	2013							
Turkey	2015*	2015*						
United States	2007	2008	2007	2012	2007	2010	2011	2013
Vietnam		2014						

[†] Values shown are the year in which authorization was granted for food and/or feed use. Some countries provide separate food and feed authorizations, in other countries a single authorization covers both types of use, and some countries provide authorization only for food use. Information on regulatory authorizations was obtained from public databases maintained by CERA (2016) and ISAAA (2016).
[a] For entries with split years, the first year refers to food use authorization and the second refers to the year in which authorization was granted for feed use.
* Years indicated with an asterisk (*) indicate authorization for feed use only.

Previously submitted safety studies reviewed in the context of other genetically engineered plant events are directly applicable to the safety assessment of PMI protein expressed in GR2E rice.

B.1.3. Tissue Specificity and Concentrations of Newly Expressed Proteins

The tissue specificity of *Zm*PSY1, CRTI, and PMI expression was confirmed by western immunoblot analysis of various tissues sampled from GR2E rice (Oliva et al., 2016b).

The patterns of expression of *Zm*PSY1, CRTI, and PMI proteins in GR2E rice tissues were consistent with the activity of the endosperm-specific rice *GluA-2* promoter upstream of the *Zmpsy1* and *crtI* genes, and use of the constitutive maize polyubiquitin promoter for the *pmi* gene. Expression of *Zm*PSY1 and CRTI was detected only in milk, dough, and mature stage grain (Figure 15, lanes 3–5, panels B and D) and not in samples of bran, hulls, leaf, stem, or root tissue. In comparison, PMI expression was detected in all GR2E rice tissue types tested (Figure 15, lanes 3–10, panel F).

Based on the western blot analysis, the only routes of potential human or animal dietary exposure to *Zm*PSY1 and CRTI proteins would be via consumption of milled rice, or derived products. Potential dietary exposure to PMI protein could be via any rice-derived product.

In order to estimate potential human and animal dietary exposure to the *Zm*PSY1, CRTI, and PMI enzymes expressed in GR2E rice, the concentration of these proteins in plant tissues was determined by quantitative enzyme-linked immunosorbent assay (ELISA). Analyses were performed on samples of grain and straw, which represent the only potential pathways

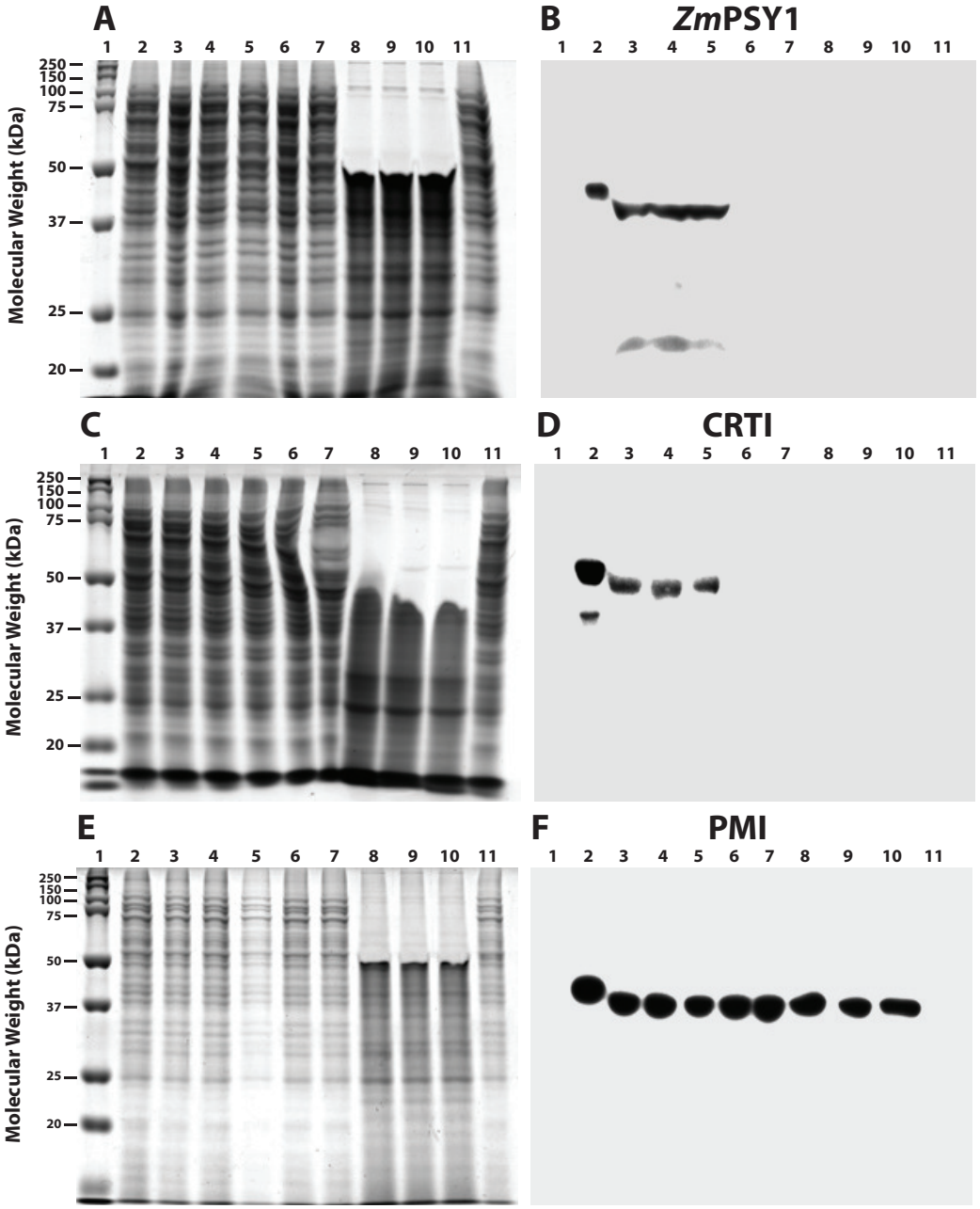


Figure 15. Total protein extracts were prepared from dough (BBCH 75; lane 3), milk (BBCH 85; lane 4), and mature (BBCH 90; lane 5) grain, bran (lane 6), hulls (lane 7), leaf (lane 8), stem (lane 9), and root tissue (lane 10) obtained from Kaybonnet rice containing event GR2E. Similar extracts were prepared from control non-transgenic Kaybonnet dough-stage grain (lane 11). For each analysis, the positive control sample (lane 2) consisted of control Kaybonnet extract spiked with 2.5 ng of purified *Zm*PSY1 protein (M20452-05; panels A and B), 25 ng of purified CRTI protein (M20454-02; panels C and D), or 6.25 ng purified PMI protein (21038G; panels E and F). Samples (*ca.* 40 µg total protein, panels A–D, or 7 µg total protein, panels E and F) were subjected to sodium dodecylsulfate (SDS)-polyacrylamide gel electrophoresis (PAGE). Gels were either stained for total protein with colloidal blue G250 (panels A, C, and E) or subjected to western immunoblot analysis (panels B, D, and F). For *Zm*PSY1 (panel B) and CRTI (panel D), the blots were labelled with monoclonal anti-*Zm*PSY1 (ELX1048; 2 µg/ml) and monoclonal anti-CRTI (ELX1043; 2 µg/ml), respectively, followed by alkaline phosphatase (AP)-conjugated goat anti-mouse IgG (1:10000). For PMI (panel F), the blot was labelled with horseradish peroxidase (HRP)-conjugated rabbit anti-PMI (1:100). For all blots, detection of bound antibody-conjugate was via chemiluminescent substrate development. Molecular weight standards are shown in lane 1 (panels A, C, and E)

of dietary exposure, that were collected from GR2E plants grown at four locations in the Philippines during the rainy season in 2015 (Oliva et al., 2016a) and again during the dry season in 2016 (Oliva et al., 2016c). Information on the planting and maintenance of the 2015 and 2016 field trials is provided in Appendix C, page 113. Materials and methods for ELISA testing are described in Appendix D, page 121.

Expression of the *Zm*PSY1 and CRTI proteins in event GR2E is driven by the endosperm-specific rice *GluA-2* promoter and measurable concentrations of both these proteins were found in all grain developmental stages but not in stem tissue (straw) (Table 8). For each protein, the highest concentrations were measured in samples of dough-stage grain (BBCH 85), ranging between *ca*. 308–359 ng/g and between *ca*. 54–68 ng/g for *Zm*PSY1 and CRTI, respectively, across both growing seasons. Across the four locations and two growing seasons, the highest concentrations of *Zm*PSY1 and CRTI measured in samples of mature grain were *ca*. 245 ng/g and 30 ng/g, respectively.

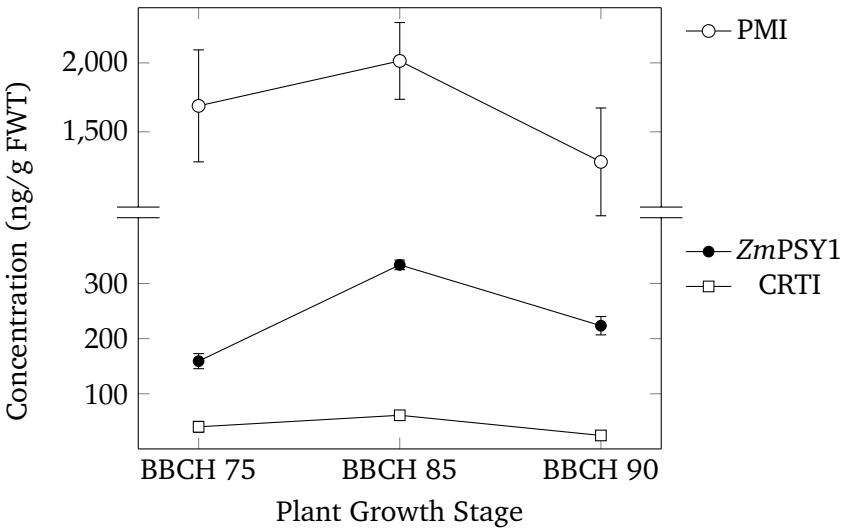


Figure 16. Samples of GR2E grain were collected at different developmental stages [BBCH 75 (milk stage), BBCH 85 (dough stage), and BBCH 90 (mature stage)] from four locations over two growing seasons in 2015-16 and the concentrations of *Zm*PSY1, CRTI, and PMI were determined by quantitative ELISA. The plotted values represent the mean concentration across locations and years for each protein, and the error bars represent the range of concentrations measured across locations over both growing seasons. In some cases, the size of the error bars was less than the symbol size used for plotting.

Concentrations of PMI protein were significantly higher than either *Zm*PSY1 or CRTI in samples from all grain growth stages (Figure 16), and were highest in dough-stage grain, averaging *ca*. 2015 ng/g across the four locations over both growing seasons. The mean PMI concentration in samples of mature GR2E rice grain was *ca*. 1282 ng/g across both growing seasons (Table 8). Since expression of the PMI protein was under control of the constitutive maize polyubiquitin promoter, it was also present in samples of straw tissue at concentrations ranging between 320–796 ng/g depending on location and growing season. The average concentration of PMI protein in GR2E straw tissue was *ca*. 482 ng/g fresh weight tissue (FWT).

As expected, the expression of *Zm*PSY1 and CRTI was limited to GR2E rice endosperm, with no detectable amounts present in samples of GR2E rice straw, while expression of PMI protein was readily quantifiable in samples of both grain and straw. For each of the proteins, the patterns and levels of expression were generally consistent between field trial locations, and across growing seasons. Concentrations of *Zm*PSY1, CRTI, and PMI were below the limit of quantification in all tissue samples collected from control non-transformed PSB Rc82 rice.

Table 8. Concentrations of *Zm*PSY1, CRTI, and PMI in samples of grain and straw collected from GR2E rice grown at four locations in the Philippines during two growing seasons in 2015-16

	Concentration (ng/g FWT)[†]					
	***Zm*PSY1**		**CRTI**		**PMI**	
Samples	**2015**	**2016**	**2015**	**2016**	**2015**	**2016**
Grain (milk stage)	156.6 (126.4–180.8)	161.7 (138.5–184.6)	43.6 (23.3–52.2)	36.6 (31.4–40.0)	1548.1 (1097–1798)	1828.4 (1352–2083)
Grain (dough stage)	338.6 (326.3–358.9)	328.9 (308.2–356.4)	61.8 (55.3–67.6)	60.3 (53.8–68.0)	1897.9 (1565–2197)	2130.7 (1912–2397)
Grain (mature stage)	226.6 (196.7–244.8)	220.2 (195.7–239.6)	26.1 (23.0–29.7)	22.8 (16.6–27.4)	1369.0 (915–1891)	1195.0 (765–1780)
Straw (mature stage)	< LOQ[a]	< LOQ	< LOQ	< LOQ	466.7 (320–617.5)	498.0 (339–795.5)

[†] Values represent the mean across four locations of three replicate samples collected from each location in the Philippines where event GR2E rice was grown during the rainy season in 2015 and again during the dry season in 2016 (n=12 for each season). The lowest and highest individual mean location values for each year are shown in parentheses. Concentrations are uncorrected for extraction efficiency and expressed in ng protein per gram fresh weight tissue (FWT).

[a] LOQ = Limit of quantification, which for the *Zm*PSY ELISA with straw (stem) samples was 66.4 ng/g FWT and for CRTI with straw samples was 6.2 ng/g FWT.

B.1.4. Estimated Maximum Human Dietary Exposure to Newly Expressed Proteins

The assessment of potential dietary exposure to new or novel proteins in the diet is an essential component of food and feed safety assessment. The assessment is conducted in a step-wise process where initially a conservative "worst-case" scenario is considered that reflects much higher levels of exposure than would actually be expected in order to account for any realistically conceivable exposure scenario (Garcia-Alonso, 2013). Risk assessment conducted under conservative worst-case scenarios helps ensure a margin of safety under actual exposure conditions.

Expression studies conducted to determine the concentrations of novel proteins in the genetically engineered plant provide data on the expected levels of the proteins in edible tissues, and food consumption databases provide estimates of expected dietary intake of the modified food.

Two approaches were followed to obtain estimates of daily rice consumption. In the first approach, historic rice utilization data for the highest rice-consuming countries in Asia, in comparison with the United States, were obtained from the USDA Production Supply and Distribution database and converted to *per capita* utilization estimates using the FAOSTAT population database. These values for 2011–2015 are presented in Table 9. Projected utilization values for the same countries were obtained from the International Rice Outlook:

International Rice Base Projections 2014–2024 publication of Wailes and Chavez (2015), also presented in Table 9 and Figure 17.

Table 9. Historic and projected per capita rice consumption in Asia and in the US

	Annual per capita Utilization (kg)[a]					Projected per capita Utilization (kg)[b]		
Country	**2011**	**2012**	**2013**	**2014**	**2015**	**2016/17**	**2020/21**	**2024/25**
Bangladesh	223.6	222.2	222.1	221.3	220.5	224.8	227.6	228.5
Cambodia	233.0	239.3	242.1	241.4	231.1	245.4	249.1	252.9
China	103.5	106.2	107.4	107.8	109.0	106.3	104.1	102.1
India	74.8	74.4	77.5	75.7	74.7	78.6	79.4	80.1
Indonesia	156.0	153.7	153.2	150.5	148.7	152.7	151.2	149.5
Malaysia	94.8	93.5	94.2	93.6	93.0	92.1	91.7	91.3
Philippines	136.1	133.8	131.7	133.1	131.6	130.0	130.0	130.5
Thailand	155.4	157.8	161.2	169.8	169.2	160.4	157.2	154.3
Vietnam	220.0	242.4	240.8	235.9	234.4	232.0	232.5	227.6
United States	11.2	12.0	12.5	13.0	12.0	13.5	13.5	13.3

[a] Consumption estimates were derived using data from the USDA PSD (Production, Supply, and Distribution) online database and the FAOSTAT population database.
[b] Projections excerpted from Table 68 in Wailes and Chavez (2015).

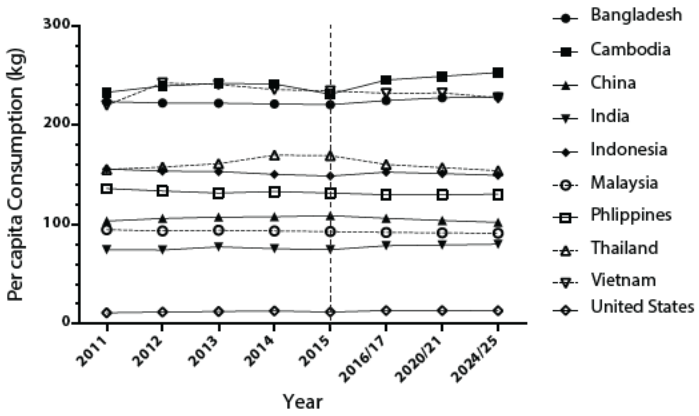


Figure 17. Per capita rice consumption in the highest-consuming countries in Asia compared with the United States. Values from 2011–2015 were derived using the USDA Production, Supply, and Distribution online database (USDA-FAS, 2016) and the FAOSTAT population database, while values after 2015 are projections taken from Wailes and Chavez (2015).

Using the highest projected *per capita* rice utilization in Cambodia of 253 kg/yr and an estimated average adult body weight of 57.7 kg in Asia (Walpole et al., 2012), the maximum daily rice intake was calculated as shown in equation 1.

$$\text{Daily Rice Intake} = \frac{253\ (\text{kg/yr})}{365 \times 57.7\ (\text{kg BW})} \times 1000\ (\text{g/kg}) = 12.0\ (\text{g/kg body weight}) \quad (1)$$

The second approach utilized data from the Food and Agriculture Organization of the United Nations (FAO)/World Health Organization (WHO) Chronic Individual Food Consumption Database summary statistics (CIFOCOss) currently containing summary statistics of 37 surveys from 26 countries. The CIFOCOss was initially developed to be used

by FAO/WHO scientific committees for dietary exposure assessment.[2] The CIFOCOss does not provide data for every country and population subgroup; however, available data for Asian countries are shown in Table 10. A further comparison of consumption data between countries in Asia and selected countries in Europe, Africa, and South America is shown in Figure 18.

Table 10. FAO/WHO CIFOCOss *per capita* daily rice consumption data for selected countries in Asia

Country	Population Subgroup	Number of Consumers Sampled	Mean Daily Consumption (g/kg body weight)[a]
Bangladesh	Adult women Children	474 545	9.84 12.32
China	General population Children	50,011 2,233	5.04 8.92
Japan	General population Children	2,711 71	6.15 10.5
Philippines	Children	1,114	5.72
Republic of Korea	General population Children	9,225 648	3.56 6.92
Thailand	General population	15,695	6.31

[a] Individual consumption data for rice (excl. wild rice; GC 0649) were obtained from the FAO/WHO Chronic Individual Food Consumption Database summary statistics (CIFOCOss) accessed on 16 May 2016. Available at http://www.who.int/foodsafety/databases/en/

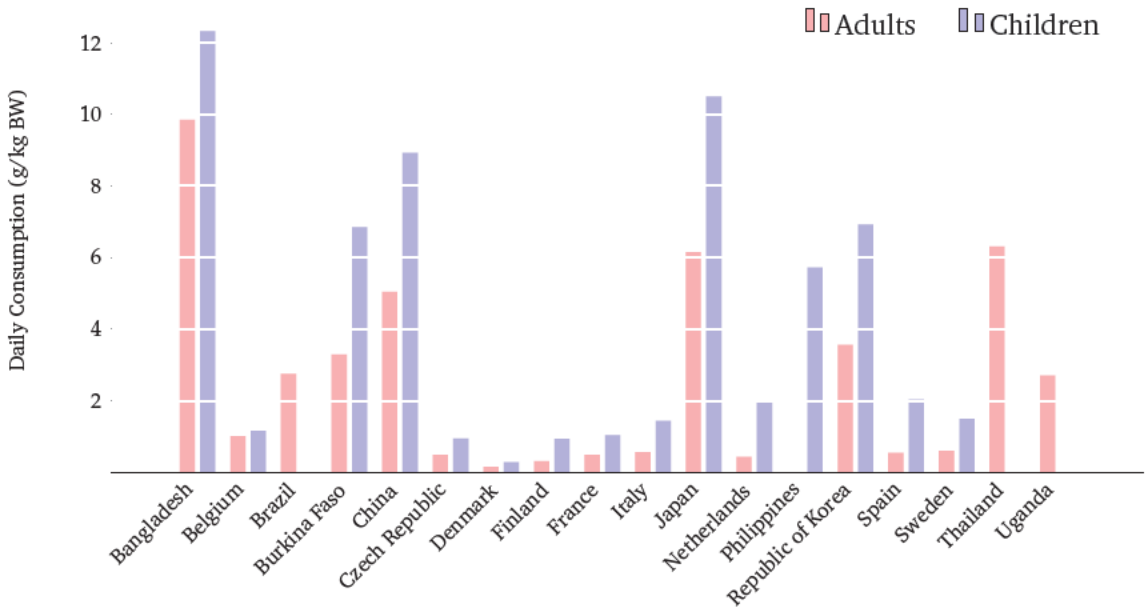


Figure 18. Daily individual rice consumption rates for adults and children in selected countries using data from the FAO/WHO Chronic Individual Food Consumption Database.

Based upon consideration of both approaches, a value of 12.5 g/kg body weight was chosen as the upper limit of mean daily dietary intake of rice. This value was judged as sufficient to account for consumption by all population subgroups, including children.

[2] The FAO/WHO Chronic Individual Food Consumption Database is available at: http://www.who.int/foodsafety/databases/en/

In deriving estimates of maximum potential daily dietary exposure to the *Zm*PSY1, CRTI, and PMI proteins expressed in GR2E rice, the following assumptions were used:

- Mean daily dietary rice consumption of 12.5 g/kg body weight.
- One hundred percent of the dietary rice intake is from GR2E rice.
- The grain concentrations of *Zm*PSY, CRTI, and PMI used for estimation are the highest values measured in samples of dough-stage grain collected from any individual trial site location in either 2015 or 2016, which were significantly higher than concentrations measured in mature grain at harvest.

Using these assumptions, the estimated maximum potential daily dietary exposure to each novel protein is shown in Table 11.

Table 11. Estimated maximum potential daily dietary exposure to *Zm*PSY1, CRTI, and PMI

Protein	Concentration (ng/g FWT)	Daily Dietary Exposure (µg/kg body weight)
*Zm*PSY1	359	4.49
CRTI	68	0.85
PMI	2397	29.96

[a] Daily dietary exposures (µg/kg body weight) to *Zm*PSY, CRTI, and PMI were calculated based on daily *per capita* rice consumption of 12.5 g/kg body weight.

The maximum potential human daily dietary exposures to *Zm*PSY1, CRTI, and PMI proteins from GR2E rice were estimated to be *ca*. 4.5, 0.85, and 30 μg/kg body weight, respectively, based on the highest concentrations of these proteins determined in dough-stage grain and a maximum daily rice intake of 12.5 g/kg body weight.

Rice consumption in Australia and New Zealand is significantly lower than in the high rice consuming nations of monsoon Asia. Based on results from the 2011-12 Australian Health Survey[3] average *per capita* daily rice consumption in Australia is 38.6 g for males and 25.8 g for females.

Using the higher consumption level for males, potential daily dietary exposure to *Zm*PSY1, CRTI, and PMI would be approximately 0.19, 0.035, and 1.25 μg/kg body weight, respectively, assuming an average body weight of 74.1 kg for persons in Oceania (Walpole et al., 2012).

More realistic estimates of exposure are likely to be significantly less considering that the maximum measured concentrations in mature GR2E rice grain at harvest were *ca*. 68, 44, and 79 percent of the maximum concentrations in dough-stage grain for *Zm*PSY1, CRTI, and PMI, respectively. Additionally, it is unlikely that 100 percent of daily rice consumption will be substituted with GR2E rice.

[3] Australian Bureau of Statistics, 43640DO005_20112012 Australian Health Survey: Nutrition First Results — Foods and Nutrients, 2011-12 — Australia. Available at: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0072011-12?OpenDocument (Accessed 02 September 2016)

B.1.5. Identity and Equivalence of ZmPSY1 Test Substance

Generally, the concentrations of newly expressed proteins present in genetically engineered plants are low, usually in the range of parts per million. In these cases, it is not feasible to utilize *in planta* expressed protein as a test substance for characterization studies that may require up to gram quantities of purified protein. This is certainly the situation for *Zm*PSY1 protein, where the levels of expressed protein in rice grain were significantly less than 0.5 $\mu g/g$ tissue, making the purification of any quantity of *Zm*PSY1 from GR2E rice highly impractical. The solution is to use a heterologous expression system (usually bacterial) to produce the needed quantities of purified test protein and then demonstrate functional equivalence to the plant-expressed form of the protein.

The *Zm*PSY1 protein was produced in an *Escherichia coli* over-expression system for use in pepsin digestibility and heat stability studies, and as a calibration standard for quantitative ELISA. Using the ChloroP Transit Peptide Predictor, the native *Zm*PSY1 protein is predicted to have a 62-residue N-terminal transit peptide that is processed to a 39.8 kDa (348 residues) mature plastid protein (Emanuelsson et al., 1999; Gallagher et al., 2004). The form of the protein expressed in recombinant *E. coli* consisted of an N-terminal truncated version of the *Zm*PSY1 protein corresponding to amino acids 50–410 of the precursor *Zm*PSY1 protein with a leader sequence containing a Factor Xa cleavage site and an N-terminal $(His)_6$-tag sequence. *Zm*PSY1 protein was solubilized from inclusion bodies using 6 M guanidine-HCl, refolded, and purified by batch affinity chromatography on Fractogel® EMD resin (Co^{2+}) followed by gel exclusion chromatography on Superdex G200. The purified protein was *ca*. 94 percent pure by reverse phase HPLC, migrated as a single species with an apparent molecular weight of *ca*. 42 kDa by SDS-PAGE, had an amino acid composition that corresponded closely with the predicted composition, displayed the expected enzymatic activity in catalyzing the production of 15-*cis*-phytoene from DMAPP and IPP, in the presence of active *A. thaliana* GGPP synthase, and was unambiguously identified as *Zm*PSY1 based on mass spectral analysis of peptide fragments following trypsin digestion (Figure 19).

The production and characterization of *Zm*PSY1 test protein is described in detail in MacKenzie (2016c).

The immunochemical cross-reactivity of plant- and microbial-expressed *Zm*PSY1 proteins was also demonstrated by western immunoblot analysis (Figure 15, panel B, page 45).

B.1.6. Identity and Equivalence of CRTI Test Substance

The CRTI protein was produced in an *Escherichia coli* over-expression system for use in pepsin digestibility, heat stability, and acute oral toxicity studies, and as a calibration standard for quantitative ELISA. The form of the protein expressed in recombinant *E. coli* consisted of a full-length version of the CRTI protein with a "GGS" linker sequence and a C-terminal $(His)_6$-tag. The protein was purified from clarified *E. coli* lysate by batch metal affinity chromatography on TALON® resin (Co^{2+}) followed by size exclusion chromatography on Superdex 200. The purified protein (Lot No. M20454-02) was *ca*. 92 percent pure by reverse phase HPLC, migrated as a single species with an apparent

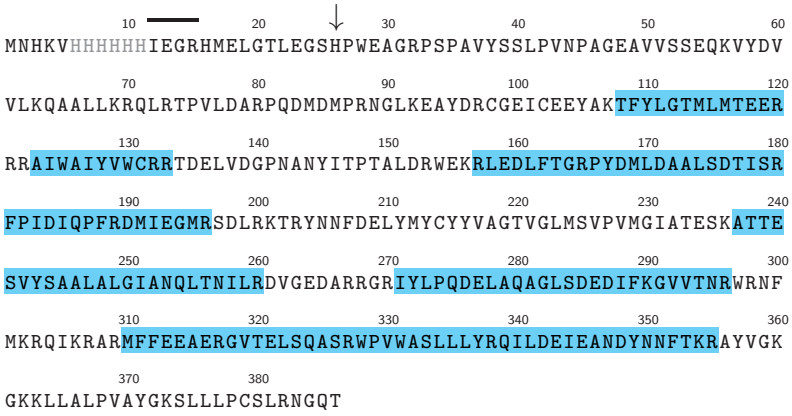


Figure 19. Amino acid sequence of the microbial-expressed *Zm*PSY1. The N-terminal $(His)_6$ tag and Factor Xa cleavage site (IEGR) are indicated within the 25-amino acid leader sequence. The start of the mature *Zm*PSY1 sequence is indicated with the downward arrow. Shaded amino acids correspond to sequences of peptides identified by mass spectral analysis following trypsin digestion of *Zm*PSY1 test protein.

molecular weight of *ca*. 50 kDa by SDS-PAGE, contained the predicted N-terminal sequence as confirmed by Edman N-terminal amino acid sequencing, had an amino acid composition that corresponded closely with the predicted composition, displayed the predicted enzymatic activity in catalyzing the conversion of liposome-incorporated phytoene to all-*trans*-lycopene, and was unambiguously identified as CRTI based on mass spectral analysis of peptide fragments following trypsin digestion (Figure 20).

The production and characterization of CRTI test protein is described in detail in MacKenzie (2016a).

The immunochemical cross-reactivity of plant- and microbial-expressed CRTI proteins was also demonstrated by western immunoblot analysis (Figure 15, panel D, page 45).

B.1.7. Bioinformatics Analyses of Newly Created ORFs

To investigate the possibility of creating new novel ORFs as a consequence of the T-DNA insertion in event GR2E rice, an open reading frame analysis was conducted to look for potential start-to-stop ORFs that spanned either the 5' or 3' junctional regions. This analysis examined each of three possible reading frames in both orientations (i.e., six possible reading frames in total) for potential ORFs capable of encoding sequences of 30 or more amino acids. Two ORFs were identified, one in the reverse (complementary) orientation that spanned the 5' T-DNA insert—genomic DNA border (Figure 9, ORF-1, 207 bp, 68 amino acids), and one in the forward orientation that spanned the 3' T-DNA insert—genomic DNA border (Figure 9, ORF-2, 240 bp, 79 amino acids).

To search for potential similarity to known toxins, the amino acid sequence of each ORF was queried against a toxin database using the FAST All sequence alignment tool (FASTA) to identify possible significant sequence similarity with known or potential toxins. The toxin database was created from a subset of sequences derived from the UniProt Knowledgebase,

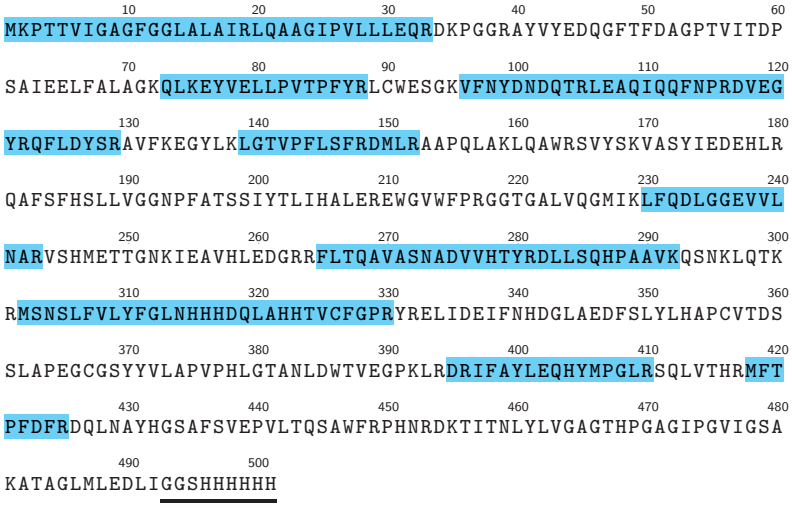


Figure 20. Amino acid sequence of the microbial-expressed CRTI. Shaded amino acids, representing 39 percent of the total sequence, correspond to sequences of 16 peptide fragments identified by mass spectral analysis following trypsin digestion of CRTI test protein. The C-terminal linker "GGS" sequence and $(His)_6$-tag is indicated with the bar.

Table 12. Location and identity of the ORFs spanning the 5' and 3' T-DNA–genomic DNA junctions in GR2E rice

ID	Nucleotide Location	Strand	Length (amino acids)	Deduced Amino Acid Sequence	No. of Allergen Hits	No. of FASTA Hits[a]
ORF-1	1830–2036	–	68 (7.9 kDa)	MIRLSFPAF SLNYQCLEK KRRFISLPV HLHQFCNCN SGCVYVSIH IHTHTLNFQ TILNRLSIL SPVLP	None	None
ORF-2	10870–11109	+	79 (8.8 kDa)	MLLDLLALQ EIYRCPGGQ HGRIRNVLL SCLSVNLFT PQFLGNGRC NQSMTTRLE SKQAQILDQ FIGGLSSLP LAFLLLQ	None	None

[a] Sequences were queried against a protein toxin database using FASTA with an *E*-value cutoff of 1×10^{-5}.

comprised of 550,552 manually annotated and reviewed sequences from Swiss-Prot and 60,971,489 automatically annotated, un-reviewed sequences from TrEMBL (The UniProt Consortium, 2014), that were selected using a keyword search on toxins (KW800). The collection contained a total of 25,572 sequences as of 20 February 2016, comprising 6,592 reviewed sequences from Swiss-Prot and 18,980 un-reviewed sequences from TrEMBL. The BLOSUM50 similarity scoring matrix was used for FASTA36 alignments (Pearson, 2013).

An *E*-score[4] acceptance criteria of 1×10^{-5} was used to identify sequences from the toxin database with potential for significant sequence similarity to the query sequences of each ORF.

The FASTA36 search resulted in no significant hits returned (Table 12, and Appendix B, page 110).

[4] The *E*-score is a parameter that describes the number of matches one can "expect" to see by chance when searching a database of a particular size. It decreases exponentially as the score of the match increases. The lower the *E*-score, or the closer it is to zero, the more "significant" the match. Typically, alignments between two sequences require an *E*-score of 1×10^{-5} or less to be considered to have sufficient sequence similarity to infer homology.

To assess the potential for allergenicity, the amino acid sequence of each ORF was compared to a peer-reviewed database of 1956 known and putative allergen and celiac protein sequences residing in the Food Allergy Research and Resource Program (FARRP) dataset at the University of Nebraska.[5] Potential identities between each of the query sequences and proteins in the allergen database were evaluated with the FASTA sequence alignment tool using an *E*-value threshold 1×10^{-5}. No significant alignments were returned. The standard greater than 35 percent identity threshold over any 80 amino acid length sequence alignment between a query sequence and an allergen was used to indicate the potential for cross-reactivity. No identity matches of greater than 35 percent over 80 residues were observed for either ORF-1 or ORF-2. Each query sequence was also evaluated for any eight contiguous identical amino acid matches to the allergens contained in the FARRP database. This was done using an algorithm that generates all possible eight-word peptides from both the query and dataset proteins and evaluates each query "word" against all dataset "words" for perfect matches. There were no eight contiguous identical amino acid matches observed for either ORF-1 or ORF-2 (Table 12 and Appendix B, page 111).

B.2. Characterization of Newly Expressed Proteins

B.2.1. Potential Toxicity of New Proteins

As a macronutrient, protein is an essential component of the human diet and, although individual proteins mediate a diverse range of biological functions, consumption of proteins as a class of dietary substances is not inherently associated with adverse effects (FAO/WHO, 1996). Only a small number of dietary proteins have the potential to exert anti-nutritional or toxic effects, or elicit allergic reaction in previously sensitized individuals.

Assessing the safety of newly-expressed proteins produced in the edible portions of a genetically engineered food crop is an integral component of the overall safety assessment. As there is currently no single criterion that is sufficiently predictive of potential toxicity or allergenicity, a "weight-of-evidence" approach is recommended for hazard identification that considers the history of use, amino acid sequence similarity to known toxins or allergens, function or mode of action, digestibility under standardized *in vitro* conditions, stability to heat or processing, and expression levels and potential dietary exposure (Codex, 2003; Delaney et al., 2008). Conventional toxicology studies are not considered necessary where the newly-expressed protein, or a closely related one, has been consumed safely in food at equivalent intakes or where the new substance is not present in the food (Codex, 2003; HC, 2006). Only when a potential for hazard has been identified, or when the previous assessment does not permit a determination of safety, is further hazard characterization warranted (Delaney et al., 2008; Hammond et al., 2013). This tier-2 characterization may include appropriate oral toxicity studies or other hypothesis-based toxicology studies when the protein's biochemical function suggests it may be potentially toxic to non-target organisms.

[5] The FARRP allergen protein database resides at AllergenOnline (http://www.allergenonline.org). Version 16 was released on January 27, 2016, and contains 1956 peer-reviewed sequences representing 778 taxonomic-protein groups.

A tiered "weight-of-evidence" approach was followed in assessing the safety of the *Zm*PSY1, CRTI, and PMI proteins expressed in GR2E rice, and is described in the following sections.

B.2.2. Bioinformatics Analyses

B.2.2.1 ZmPSY1 Protein

One component of the tier-1 safety assessment of proteins (Delaney et al., 2008) is a bioinformatic evaluation of the amino acid sequence similarity between the protein and known protein toxins (Codex, 2003), which is described in Oliva and MacKenzie (2016d) for the *Zm*PSY1 protein.

Potential structural similarities shared between the *Zm*PSY1 protein and sequences in a protein toxin database were evaluated using version 36 the FASTA (FASTA36) sequence alignment tool (Pearson and Lipman, 1988; Pearson, 1996, 2000). The FASTA36 program directly compares amino acid sequences (i.e., primary, linear protein structure) and the alignment data may be used to infer shared higher order structural similarities between two sequences (i.e., secondary and tertiary protein structures). Proteins that share a high degree of similarity throughout the entire sequence are often homologous.

The target sequence for bioinformatic analysis was the 410-amino acid *Zm*PSY1 sequence (Buckner et al., 1996), which is shown in Figure 21.

```
          10        20        30        40        50        60
MAIILVRAASPGLSAADSISHQGTLQCSTLLKTKRPAARRWMPCSLLGLHPWEAGRPSPA

          70        80        90       100       110       120
VYSSLPVNPAGEAVVSSEQKVYDVVLKQAALLKRQLRTPVLDARPQDMDMPRNGLKEAYD

         130       140       150       160       170       180
RCGEICEEYAKTFYLGTMLMTEERRRAIWAIYVWCRRTDELVDGPNANYITPTALDRWEK

         190       200       210       220       230       240
RLEDLFTGRPYDMLDAALSDTISRFPIDIQPFRDMIEGMRSDLRKTRYNNFDELYMYCYY

         250       260       270       280       290       300
VAGTVGLMSVPVMGIATESKATTESVYSAALALGIANQLTNILRDVGEDARRGRIYLPQD

         310       320       330       340       350       360
ELAQAGLSDEDIFKGVVTNRWRNFMKRQIKRARMFFEEAERGVTELSQASRWPVWASLLL

         370       380       390       400       410
YRQILDEIEANDYNNFTKRAYVGKGKKLLALPVAYGKSLLLPCSLRNGQT
```

Figure 21. Deduced amino acid sequence of the *Zm*PSY1 protein (EC 2.5.1.32; 15-*cis*-phytoene synthase, sp|P49085). The protein is 410 amino acids in length with a calculated molecular weight of *ca.* 46.5 kDa.

A FASTA36 bioinformatic alignment search using the *Zm*PSY1 amino acid sequence as the query sequence was performed against a toxin database to identify possible significant sequence similarity with known or potential toxins. The toxin database was created from a subset of sequences derived from the UniProt Knowledgebase, comprised of 550,116 manually annotated and reviewed sequences from Swiss-Prot and 55,270,679 automatically annotated, un-reviewed sequences from TrEMBL (The UniProt Consortium, 2014), that were selected using a keyword search on toxins (KW800). The collection contained a total of 24,098 sequences as of 21 January 2016, comprising 6,588 reviewed sequences from

Swiss-Prot and 17,510 un-reviewed sequences from TrEMBL. The BLOSUM50 similarity scoring matrix was used for FASTA36 alignments (Pearson, 2013).

An *E*-score acceptance criteria of 1×10^{-5} was used to identify sequences from the toxin database with potential for significant sequence similarity to the *Zm*PSY1 protein. The *E*-score is a parameter that describes the number of matches one can "expect" to see by chance when searching a database of a particular size. It decreases exponentially as the score of the match increases. The lower the *E*-score, or the closer it is to zero, the more "significant" the match. Typically, alignments between two sequences require an *E*-score of 1×10^{-5} or less to be considered to have sufficient sequence similarity to infer homology.

The toxin database search using the *Zm*PSY1 query sequence did not return any entries with an *E*-score less than 1×10^{-5} (Oliva and MacKenzie, 2016d).

Based on the results of this study, there were no sequence homology structural alerts for potential toxicity of the *Zm*PSY1 protein.

B.2.2.2 CRTI Protein

The bioinformatic evaluation of the amino acid sequence similarity between the CRTI protein and known protein toxins is described in Oliva and MacKenzie (2016b).

The target sequence for bioinformatic analysis was the mature 492-amino acid phytoene desaturase (CRTI) sequence (Misawa et al., 1990), which is shown in Figure 22.

```
          10         20         30         40         50         60
MKPTTVIGAGFGGLALAIRLQAAGIPVLLLEQRDKPGGRAYVYEDQGFTFDAGPTVITDP

          70         80         90        100        110        120
SAIEELFALAGKQLKEYVELLPVTPFYRLCWESGKVFNYDNDQTRLEAQIQQFNPRDVEG

         130        140        150        160        170        180
YRQFLDYSRAVFKEGYLKLGTVPFLSFRDMLRAAPQLAKLQAWRSVYSKVASYIEDEHLR

         190        200        210        220        230        240
QAFSFHSLLVGGNPFATSSIYTLIHALEREWGVWFPRGGTGALVQGMIKLFQDLGGEVVL

         250        260        270        280        290        300
NARVSHMETTGNKIEAVHLEDGRRFLTQAVASNADVVHTYRDLLSQHPAAVKQSNKLQTK

         310        320        330        340        350        360
RMSNSLFVLYFGLNHHHDQLAHHTVCFGPRYRELIDEIFNHDGLAEDFSLYLHAPCVTDS

         370        380        390        400        410        420
SLAPEGCGSYYVLAPVPHLGTANLDWTVEGPKLRDRIFAYLEQHYMPGLRSQLVTHRMFT

         430        440        450        460        470        480
PFDFRDQLNAYHGSAFSVEPVLTQSAWFRPHNRDKTITNLYLVGAGTHPGAGIPGVIGSA

         490
KATAGLMLEDLI
```

Figure 22. Deduced amino acid sequence of the mature CRTI phytoene desaturase protein derived from *Pantoea ananatis* (sp|P21685). The mature protein is 492 amino acids in length with a calculated molecular weight of 55.0 kDa.

A FASTA36 bioinformatic alignment search using the CRTI amino acid sequence as the query sequence was performed against a toxin database to identify possible significant sequence similarity with known or potential toxins. The toxin database was created from a subset of sequences derived from the UniProt Knowledgebase, comprised of 550,116 manually annotated and reviewed sequences from Swiss-Prot and 55,270,679 automatically

annotated, un-reviewed sequences from TrEMBL (The UniProt Consortium, 2014), that were selected using a keyword search on toxins (KW800). The collection contained a total of 24,098 sequences as of 16 January 2016, comprising 6,588 reviewed sequences from Swiss-Prot and 17,510 un-reviewed sequences from TrEMBL. The BLOSUM50 similarity scoring matrix was used for FASTA36 alignments (Pearson, 2013).

An *E*-score acceptance criteria of 1×10^{-5} was used to identify sequences from the toxin database with potential for significant sequence similarity to the CRTI protein. The *E*-score is a parameter that describes the number of matches one can "expect" to see by chance when searching a database of a particular size. It decreases exponentially as the score of the match increases. The lower the *E*-score, or the closer it is to zero, the more "significant" the match. Typically, alignments between two sequences require an *E*-score of 1×10^{-5} or less to be considered to have sufficient sequence similarity to infer homology.

A search using the CRTI query sequence returned three protein accessions from the toxin database with an *E*-score less than 1×10^{-5} (Table 13). The three sequence alignments (Figure 23) were to the N-terminal regions of L-amino acid oxidase (LAAO) enzymes from three species of venomous snakes: *Bungarus multicinctus* (many-banded krait, also known as the Taiwanese krait or the Chinese krait), *B. fasciatus* (banded krait), and *Daboia russelii* (Russell's viper).

Table 13. FASTA36 alignment results using CRTI query sequence against the toxin database

Entry	Description	Identity	*E*-Score[†]
sp|A8QL51	L-amino acid oxidase (*Bungarus multicinctus*)	37.0% identity in 81 aa overlap (2-79:53-133)	9.7×10^{-7}
sp|A8QL52	L-amino acid oxidase (*Bungarus fasciatus*)	35.8% identity in 81 aa overlap (2-79:53-133)	4.9×10^{-6}
sp|G8XQX1	L-amino acid oxidase (*Daboia russelii*)	32.1% identity in 81 aa overlap (2-79:52-132)	7.5×10^{-6}

[†] The FASTA36 alignment search was performed on 16 January 2016 using a protein toxin database containing 6,588 sequences from Swiss-Prot and 17,510 sequences from TrEMBL that had been selected using the keyword "toxin" (KW800). An *E*-score acceptance criteria of 1×10^{-5} was used to identify sequences from the toxin database with potential for significant sequence similarity to the CRTI protein.

LAAOs (EC 1.4.3.2) are widely distributed in many different species including insects, fungi, bacteria, and snakes, and are even found in plants where one of their catalytic products, ammonia, is used as a nitrogen source in cell metabolism. In snake venoms, LAAOs are found in high concentrations that vary according to each species of snake, which may contribute to the toxicity of ophidian envenomation. Snake venom LAAOs are non-toxic via the oral route of exposure. LAAOs exhibit catalytic specificity for long chain hydrophobic and aromatic amino acids and are active in a wide range of pHs and temperatures. Their structures, molecular masses, and isoelectric points are quite varied. They are able to induce changes in platelet function, which cause local effects on plasma clotting disorders among other things. LAAOs are capable of inducing apoptosis in various cell lines and show antimicrobial and anti-parasitic activity. The existence of LAAOs may be a means of protection against natural agents, parasites, and bacteria. For a recent review of LAAOs, see Izidoro et al. (2014).

LAAOs are flavoenzymes belonging to the class of oxidoreductases that catalyze the stereospecific oxidative deamination of L-amino acids. Snake venom LAAOs are usually homodimeric with cofactors FAD (flavin adenine dinucleotide) or FMN (flavin

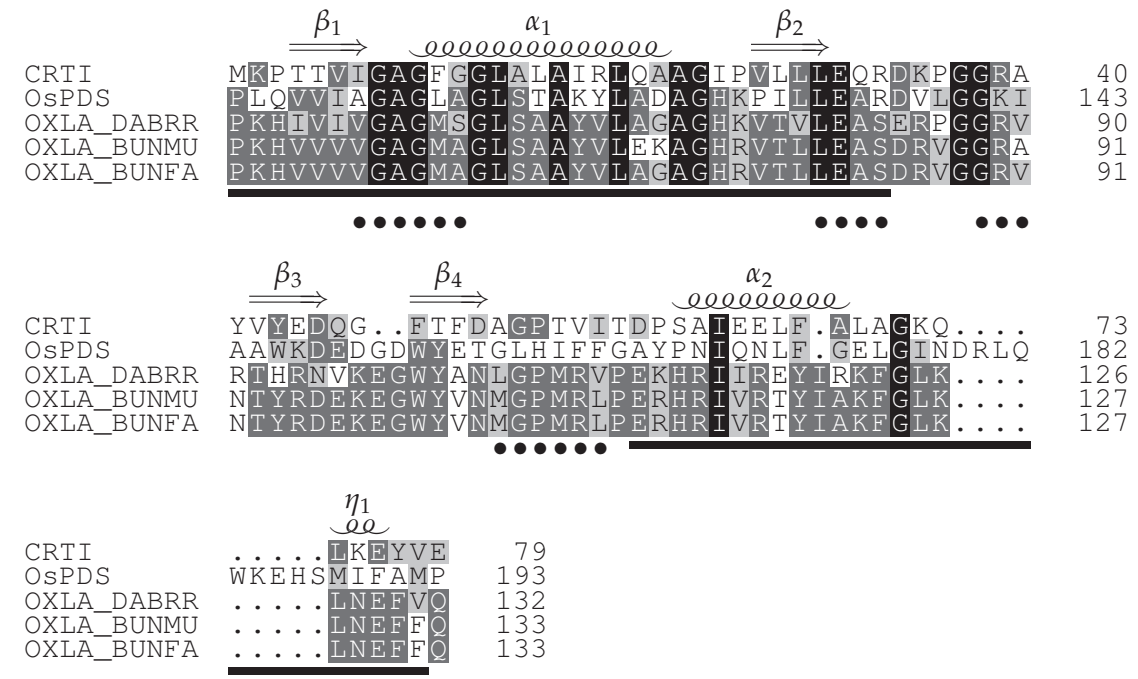


Figure 23. Alignment of the N-terminal region of the *P. ananatis* CRTI protein (sp|P21685, residues 1–79) with the corresponding regions of the rice (*O. sativa*) phytoene desaturase (OsPDS, sp|A2XDA1) and the three L-amino acid oxidase enzymes identified from the toxin database search. These were from *Bungarus multicinctus* (OXLA_BUNMU, sp|A8QL51), *B. fasciatus* (OXLA_BUNFA, sp|A8QL52), and *Daboia russelii* (OXLA_DABRR, sp|G8XQX1). The secondary structure elements of CRTI have been indicated above the alignment and the shaded bar underneath the alignment indicates the FAD-binding domains within the N-terminal region, with bullets indicating putative FAD binding residues, as described in Schaub et al. (2012). This figure was generated with TEXshade (Beitz, 2000).

mononucleotide) covalently linked to their chemical structure. As demonstrated by Ali et al. (1999), the presence of a highly conserved β-α-β-fold domain in the N-terminal region of an LAAO from *Eristicophis macmahoni* venom, is involved in binding the FAD cofactor.

As previously noted, while CRTI and OsPDS share limited overall amino acid sequence similarity, they are structurally related and share a homologous FAD-binding domain.

For comparison, a search using the 578-amino acid OsPDS query sequence (sp|A2XDA1) against the toxin database returned 24 matches to L-amino acid oxidases with *E*-scores ranging from 4.9×10^{-11} to 1×10^{-6} (Oliva and MacKenzie, 2016b). All of these matches reflected sequence similarities within the N-terminal motif involved in FAD cofactor binding.

The limited sequence similarity between the *P. ananatis* CRTI protein and the three L-amino acid oxidase accessions retrieved from the toxin database was due to homology between N-terminal motifs involved in FAD binding, and was not considered to be a structural alert for potential toxicity. As previously noted, snake venom LAAOs are non-toxic via the oral route and the N-terminal sequence similarity was also observed between the native rice phytoene desaturase and a range of L-amino acid oxidases.

B.2.2.3 PMI Protein

The BLASTP program was used to search the National Center for Biotechnology Information Entrez® Protein Database to determine whether the PMI amino acid sequence showed significant similarity to known or putative toxins (Harper, 2011a). The threshold value for determining significance of matches was based on searches conducted with randomly shuffled sequences of the amino acids comprising PMI. There were 1384 protein sequences identified as having significant sequence similarity to PMI amino acid sequence; however, none of these proteins were known or putative toxins. These results support the conclusion that there are no primary sequence structural alerts for potential toxicity associated with the PMI protein.

B.2.3. Stability to Proteolysis

Susceptibility to digestion has been routinely evaluated as part of the safety assessment of novel proteins introduced into genetically engineered plants. Proteins that are readily digestible are likely to behave like other dietary proteins upon ingestion and unlikely to result in allergic or toxic reactions. Astwood et al. (1996) were among the first to provide empirical evidence supporting a correlation between the pepsin resistance of a protein and its allergen status. Although a number of subsequent studies have indicated a weaker link between stability to digestion and allergenicity (Fu et al., 2002; Herman et al., 2007; Bøgh and Madsen, 2015), the resistance of a novel food protein to pepsin digestion under acidic conditions remains generally accepted as one factor to consider in a weight-of-evidence approach for assessing potential allergenicity and toxicity (Bannon et al., 2002; Codex, 2003; Goodman et al., 2005).

The pepsin digestibility assay is not meant to predict whether a given protein will always be digested in the stomach of the consumer, but it does provide a simple, standardized, *in vitro* correlative assay to evaluate relative protein digestibility (Thomas et al., 2004).

B.2.3.1 ZmPSY1 Protein

The *in vitro* pepsin resistance of *Zm*PSY1 protein was investigated by incubating purified *Zm*PSY1 protein for 0, 0.5, 1, 2, 5, 10, 20, 30, and 60 minutes at 37°C in the presence of simulated gastric fluid (SGF) pH 1.2 containing pepsin (Oliva, 2016). Control digestions with bovine serum albumin (BSA) and β-lactoglobulin were performed for 0, 1, and approximately 60 minutes under the same conditions. Samples were removed at stated time points and subjected to sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis.

Following exposure to SGF containing pepsin for 30 seconds, the earliest time point sampled during the digestion, no intact *Zm*PSY1 protein (*ca*. 42 kDa) was evident as assessed by either SDS-PAGE or western immunoblot analysis (Figure 24, lane 4 in panels A and B, respectively). Faint, low molecular weight degradation products were visible in samples removed up to two minutes of digestion (Figure 24, lanes 4–6, panel A), but not at later time points, and these were not detected in the western blot.

The results from positive and negative control digestions were as expected. The positive control protein, BSA, was rapidly digested with no detectable intact protein after one minute of digestion (Figure 24, panel B, lane 7), while intact β-lactoglobulin, which is known to be stable to pepsin, was present after 60 minutes of incubation in SGF containing pepsin (Figure 24, panel B, lane 4).

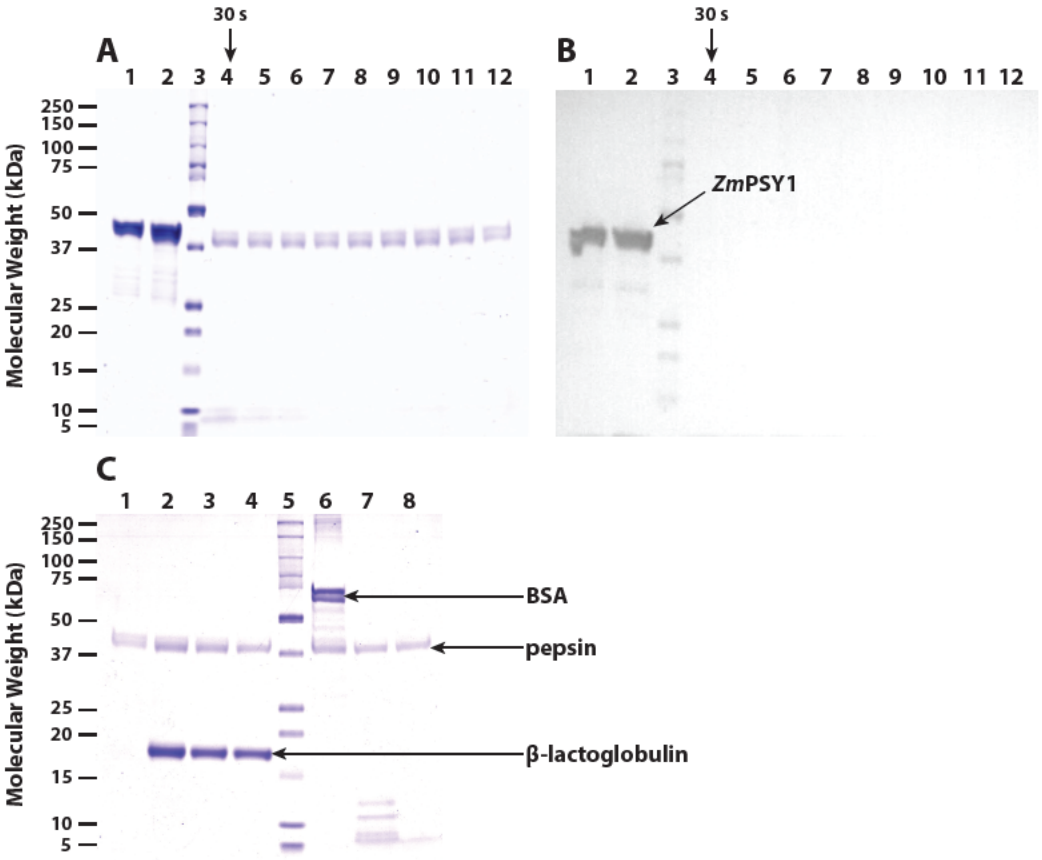


Figure 24. Panels A and B: Samples of *Zm*PSY1 protein purified from recombinant *E. coli* (M20452-05) were incubated in the presence of SGF pH 1.2 containing pepsin for 0 min (lane 2) and 0.5, 1, 2, 5, 10, 20, 30 or 60 min at 37°C (lanes 4–11) and then analyzed by SDS-PAGE. Gels were either stained for protein with colloidal blue G250 (panel A) or subjected to western immunoblot analysis (panel B) using rabbit anti-*Zm*PSY1 immunoglobulin and horseradish peroxidase-conjugated goat anti-rabbit IgG followed by precipitating substrate development. Control samples included *Zm*PSY1 protein diluted in gastric control fluid without pepsin (lane 1) and SGF solution containing pepsin (lane 12). Molecular weight standards are shown in lane 3. The apparent molecular weight of *Zm*PSY1 protein corresponds to *ca.* 42 kDa.

Panel C: In a separate set of control digestions, samples of β-lactoglobulin and BSA were incubated in the presence of SGF pH 1.2 containing pepsin for 0 (lanes 2 and 6, respectively), 1 (lanes 3 and 7, respectively), or 60 min (lanes 4 and 8, respectively) at 37°C and then analyzed by SDS-PAGE. Lane 1 contained SGF control. Molecular weight standards are shown in lane 5.

These data support the conclusion that *Zm*PSY1 protein will be readily digested as conventional dietary protein in a typical mammalian gastric environment. Based on the Codex (2003) guidelines, there would not be a concern of increased potential allergenicity or toxicity due to resistance of the *Zm*PSY1 protein to pepsin digestion.

B.2.3.2 CRTI Protein

The *in vitro* pepsin resistance of CRTI protein was investigated by incubating purified CRTI protein for 0, 0.5, 1, 2, 5, 10, 20, 30, and 60 minutes at 37°C in the presence of simulated gastric fluid (SGF) pH 1.2 containing pepsin (Oliva and Cueto, 2016). Control digestions with bovine serum albumin and β-lactoglobulin were performed for 0, 1, and approximately 60 minutes under the same conditions. Samples were removed at stated time points and subjected to SDS-PAGE analysis.

Following exposure to SGF containing pepsin for 30 seconds, the earliest time point sampled during the digestion, no intact CRTI protein was evident as assessed by either SDS-PAGE or western immunoblot analysis (Figure 25, lane 4 in panels A and B, respectively), and there was no evidence of stable lower molecular weight proteolytic fragments derived from CRTI.

The results from positive and negative control digestions were as expected. The positive control protein, BSA, was rapidly digested with no detectable intact protein after one minute of digestion (Figure 25, panel B, lane 7), while intact β-lactoglobulin, which is known to be stable to pepsin, remained undigested after 60 minutes of incubation in SGF containing pepsin (Figure 25, panel B, lane 4).

These data support the conclusion that CRTI protein will be readily digested as conventional dietary protein in a typical mammalian gastric environment. Based on the Codex (2003) guidelines, there would not be a concern of increased potential allergenicity or toxicity due to resistance of the CRTI protein to pepsin digestion.

B.2.3.3 PMI Protein

The susceptibility of PMI to proteolytic degradation in simulated mammalian gastric fluid (SGF) containing pepsin was evaluated using SDS-PAGE and western blot analysis (Nelson, 2008). The results from this study demonstrated that PMI was readily degraded with no intact protein (apparent molecular weight *ca*. 42.8 kDa) or degradation products detected following digestion for one minute.

These data support the conclusion that PMI protein will be readily digested as conventional dietary protein in a typical mammalian gastric environment. Based on the Codex (2003) guidelines, there would not be a concern of increased potential allergenicity or toxicity due to resistance of the PMI protein to pepsin digestion.

B.2.4. Acute Oral Toxicity

B.2.4.1 ZmPSY1 Protein

A tier-1 assessment of potential hazards associated with the *Zm*PSY1 protein, which considered the food crop source of the gene, lack of significant amino acid sequence similarity with known toxins and allergens, susceptibility to heat inactivation, and rapid digestibility concluded that further hazard characterization by animal toxicity testing was unnecessary.

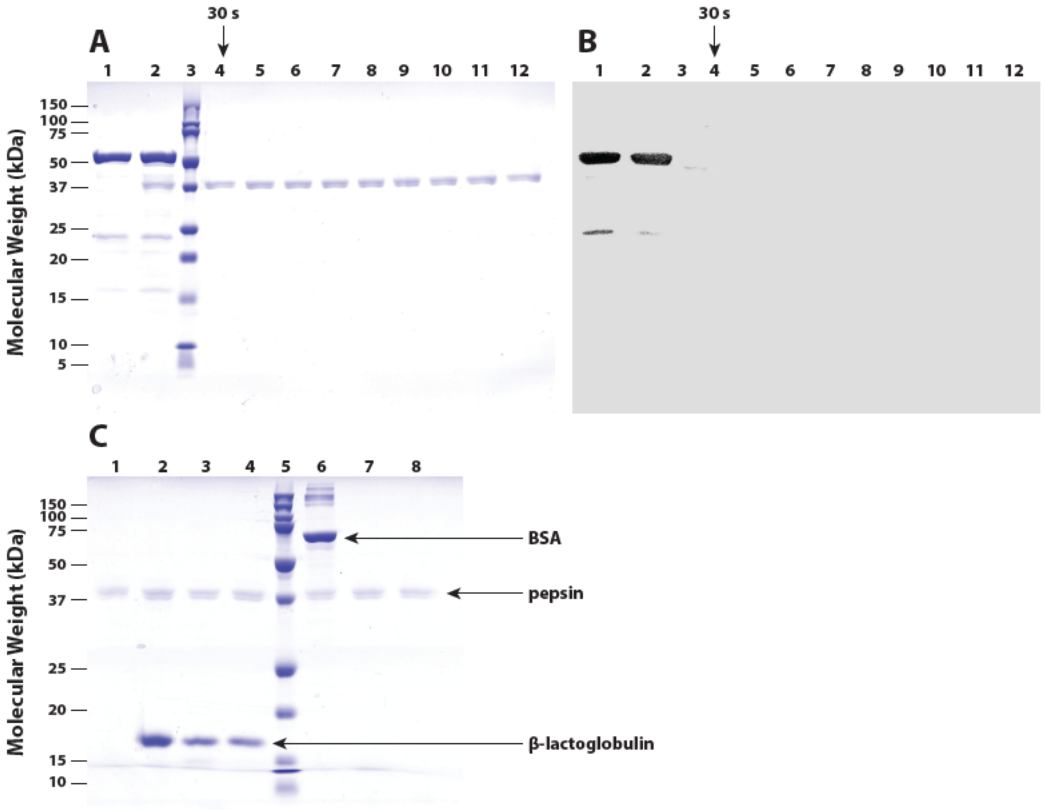


Figure 25. Panels A and B: Samples of CRTI protein purified from recombinant *E. coli* (M20454-02) were incubated in the presence of SGF pH 1.2 containing pepsin for 0 min (lane 2) and 0.5, 1, 2, 5, 10, 20, 30 or 60 min at 37°C (lanes 4–11) and then analyzed by SDS-PAGE. Gels were either stained for protein with colloidal blue G250 (panel A) or subjected to western immunoblot analysis (panel B) using rabbit anti-CRTI immunoglobulin (1:1000) and horseradish peroxidase-conjugated goat anti-rabbit IgG followed by precipitating substrate development. Control samples included CRTI protein diluted in gastric control fluid without pepsin (lane 1) and SGF solution containing pepsin (lane 12). Molecular weight standards are shown in lane 3.

Panel C: In a separate set of control digestions, samples of β-lactoglobulin and BSA were incubated in the presence of SGF pH 1.2 containing pepsin for 0 (lanes 2 and 6, respectively), 1 (lanes 3 and 7, respectively), or 60 min (lanes 4 and 8, respectively) at 37°C and then analyzed by SDS-PAGE. Lane 1 contained SGF control. Molecular weight standards are shown in lane 5.

B.2.4.2 CRTI Protein

An analysis of potential toxicological hazards associated with the CRTI protein that considered the lack of primary sequence structural alerts for toxicity, the defined mode of action of CRTI including its lack of thermal stability, and the rapid pepsin digestibility of CRTI, did not identify any concerns. Based on these considerations, it can be predicted that CRTI protein is unlikely to be acutely toxic by the oral route.

As a further assurance of safety, the potential for acute toxicity resulting from a single oral exposure to CRTI was investigated in mice (Mukerji, 2016). Groups of five male and five female Crl:CD1(ICR) mice were dosed orally by gavage with: formulation buffer (vehicle control; 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP)); bovine serum albumin (BSA; negative control; target dose 100 mg/kg body weight dissolved in vehicle control solution); or purified microbial-expressed CRTI (Lot No. M20603; 100 mg/kg body weight actual dose purified protein dissolved in vehicle control solution), at a volume of 13.45 ml/kg body weight, administered in two divided doses *ca*. 4

hours apart on test day 1. Body weights were evaluated on test days 1 (pre-fast and shortly prior to administration of the first dose), 2, 3, 5, 8, and 15. Clinical signs were evaluated seven times on test day 1 (distributed before and after each dose) and daily thereafter. On test day 15, all mice were euthanized and given a gross pathological examination.

All animals survived until the scheduled end of the study period on day 15 and there were no clinical signs of toxicity observed during the test period, nor were any gross lesions found in the mice at necropsy. There were no treatment-related effects on body weights for male or female mice over the duration of the study (Figure 26) and all mice experienced net weight gain by test day 15 compared with test day 1 (pre-fast).

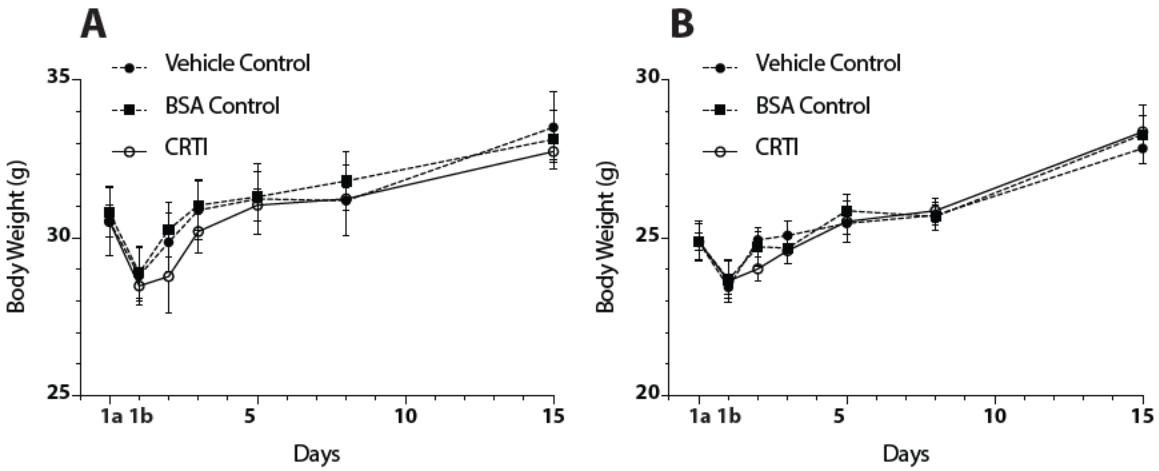


Figure 26. Groups of five fasted male (panel A) and five fasted female (panel B) Crl:CD1(ICR) mice were subjected to oral gavage with vehicle control buffer, or control BSA or CRTI test substance (Lot No. M20603, see page 140 for certificate of analysis (COA) and MacKenzie (2016b) for complete characterization) at a dosage of 100 mg/kg body weight. Individual animal body weights were measured on test day 1 before (1a) and after fasting (1b) and on days 2, 3, 5, 8, and 15. Mean body weights are plotted with standard errors (± SE).

Under the conditions of this testing, oral administration of CRTI test substance at a concentration of 100 mg/kg body weight produced no test substance-related clinical signs of toxicity, body weight losses, gross lesions, or mortality, further substantiating the predicted lack of acute oral toxicity of CRTI.

Based on very conservative assumptions in deriving a "worst-case" scenario for maximum potential daily dietary exposure to CRTI protein (0.85 μg/kg body weight; Section B.1.4, Table 11, page 50), the dose concentration used for acute oral toxicity testing of CRTI protein represents a margin of exposure of greater than 115,000–fold.

B.2.4.3 PMI Protein

Microbial-expressed PMI protein (89.5 percent purity) administered as a single oral gavage dose at 0 or 2000 mg/kg body weight followed by a 14-day observation period was well tolerated in male and female Crl:CD-1(ICR) mice (Korgaonkar, 2009). All mice survived without clinical signs of physical impairment or physiologic dysfunction. There were no test substance-related clinical observations. There were no test substance-related effects on body weight or weight gain, food consumption, or haematology parameters. There were no macroscopic or microscopic findings that were attributable to the test substance.

Collectively, clinical and anatomical pathology findings indicated that, under the conditions of this study, a specific target organ toxicity was not identified.

B.2.5. Potential Allergenicity of New Proteins

B.2.6. Bioinformatics Analyses

There is currently no definitive test that can be relied upon to predict allergic response in humans to a novel dietary protein, thus an integrated, stepwise, case-by-case approach is recommended. One component of this assessment includes an evaluation of whether significant amino acid sequence similarity exists between the new protein and known allergens. The purpose of a sequence homology comparison is to assess the extent to which a newly expressed protein is similar in structure to a known allergen, and this information may suggest whether there is any potential for allergic cross-reaction between the new protein and a known allergen.

As recommended by Codex (2003), IgE cross-reactivity between the newly expressed protein and a known allergen should be considered a possibility when there is more than 35 percent identity in a segment of 80 or more amino acids. The window size of 80 amino acids is meant to correspond with a typical domain size in a protein, and recognizes that a single domain may contain epitopes that mediate antibody binding.

It is also possible that proteins structurally unrelated to allergens, gliadins, and glutenins may contain smaller immunologically meaningful epitopes. A sliding window search was first recommended by Metcalfe et al. (1996) to determine if small regions of proteins that represented IgE-binding epitopes or T-cell epitopes shared identities with known allergens that would be missed in the global sequence alignment. Although considerable variation has been reported in the sizes of IgE binding epitopes, ranging from four amino acids in length to as much as 23 amino acids long, the vast majority consist of at least eight amino acids (Thomas et al., 2005; Bannon and Ogawa, 2006). Using a sliding window of less than eight amino acids can produce matches containing considerable uncertainty as to their significance depending on the length of the query sequence (Silvanovich et al., 2006) and are not useful to the allergy assessment process (Thomas et al., 2005). In fact, the use of any short-alignment criteria for predicting the allergenic potential of proteins has also been criticized (Goodman et al., 2008; Cressman and Ladics, 2009). An *in silico* study by Herman et al. (2009) found that the allergen-sequence pairs only sharing 8-amino-acid identity, but not >35 percent identity over 80 amino acids, were unlikely to be cross-reactive allergens.

A negative outcome of a sequence homology evaluation indicates that a newly expressed protein is not a known allergen and is unlikely to be cross-reactive to known allergens. A result indicating absence of significant sequence similarity should be considered along with the other data in assessing the allergenic potential of newly expressed proteins (Codex, 2003).

B.2.6.1 ZmPSY1 Protein

There is currently no definitive test that can be relied upon to predict allergic response in humans to a novel dietary protein, thus an integrated, stepwise, case-by-case approach

is recommended. One component of this assessment includes an evaluation of whether significant amino acid sequence similarity exists between the new protein and known allergens. The purpose of a sequence homology comparison is to assess the extent to which a newly expressed protein is similar in structure to a known allergen, and this information may suggest whether there is any potential for allergic cross-reaction between the new protein and a known allergen.

As recommended by Codex (2003), IgE cross-reactivity between the newly expressed protein and a known allergen should be considered a possibility when there is more than 35 percent identity in a segment of 80 or more amino acids. The window size of 80 amino acids is meant to correspond with a typical domain size in a protein, and recognizes that a single domain may contain epitopes that mediate antibody binding.

It is also possible that proteins structurally unrelated to allergens, gliadins, and glutenins may contain smaller immunologically meaningful epitopes. A sliding window search was first recommended by Metcalfe et al. (1996) to determine if small regions of proteins that represented IgE-binding epitopes or T-cell epitopes shared identities with known allergens that would be missed in the global sequence alignment. Although considerable variation has been reported in the sizes of IgE binding epitopes, ranging from four amino acids in length to as much as 23 amino acids long, the vast majority consist of at least eight amino acids (Thomas et al., 2005; Bannon and Ogawa, 2006). Using a sliding window of less than eight amino acids can produce matches containing considerable uncertainty as to their significance depending on the length of the query sequence (Silvanovich et al., 2006) and are not useful to the allergy assessment process (Thomas et al., 2005). In fact, the use of any short-alignment criteria for predicting the allergenic potential of proteins has also been criticized (Goodman et al., 2008; Cressman and Ladics, 2009). An *in silico* study by Herman et al. (2009) found that the allergen-sequence pairs only sharing 8-amino-acid identity, but not $>$35 percent identity over 80 amino acids, were unlikely to be cross-reactive allergens.

A negative outcome of a sequence homology evaluation indicates that a newly expressed protein is not a known allergen and is unlikely to be cross-reactive to known allergens. A result indicating absence of significant sequence similarity should be considered along with the other data in assessing the allergenic potential of newly expressed proteins (Codex, 2003).

As part of assessing the potential allergenicity of the *Zea mays* phytoene synthase, an evaluation of the significance of any amino acid sequence similarity with known allergens was conducted (Oliva and MacKenzie, 2016c).

To assess the potential for allergenic cross-reactivity, the 410-amino acid sequence encoded by the *Zmpsy1* gene was compared to a peer-reviewed database of 1956 known and putative allergen and celiac protein sequences residing in the FARRP16 dataset at the University of Nebraska.[6]

Potential identities between the *Zm*PSY1 query sequence and proteins in the allergen database were evaluated with the FASTA35 sequence alignment tool using the default

[6] The Food Allergy Research and Resource Program (FARRP) allergen protein database resides at AllergenOnline (http://www.allergenonline.org). Version 16 was released on January 27, 2016, and contains 1956 peer-reviewed sequences representing 778 taxonomic-protein groups.

parameters.[7] The recommended greater than 35 percent identity threshold over any 80-amino acid length sequence alignment between the query sequence and an allergen was used to indicate the potential for cross-reactivity. No identity matches of $>$ 35 percent over 80 residues were observed (Table 14).

The *Zm*PSY1 query sequence was also evaluated for any eight contiguous identical amino acid matches to the allergens contained in the FARRP database. This was done using an algorithm that generates all possible eight-word peptides from both the query and dataset proteins and evaluates each query "word" against all dataset "words" for perfect matches. There were no eight contiguous identical amino acid matches observed (Table 14).

Table 14. Search results using *Zm*PSY1 query sequence against the FARRP database

Database	AllergenOnline Database v16 (January 27, 2016)[a]
Input Query	>ZmPSY1 Zea mays phytoene synthase P49085 MAIILVRAA SPGLSAADS ISHQGTLQC STLLKTKRP AARRWMPCS LLGLHPWEA GRPSPAVYS SLPVNPAGE AVVSSEQKV YDVVLKQAA LLKRQLRTP VLDARPQDM DMPRNGLKE AYDRCGEIC EEYAKTFYL GTMLMTEER RRAIWAIYV WCRRTDELV DGPNANYIT PTALDRWEK RLEDLFTGR PYDMLDAAL SDTISRFPI DIQPFRDMI EGMRSDLRK TRYNNFDEL YMYCYYVAG TVGLMSVPV MGIATESKA TTESVYSAA LALGIANQL TNILRDVGE DARRGRIYL PQDELAQAG LSDEDIFKG VVTNRWRNF MKRQIKRAR MFFEEAERG VTELSQASR WPVWASLLL YRQILDEIE ANDYNNFTK RAYVGKGKK LLALPVAYG KSLLLPCSL RNGQT
Length	410
Number of 80mers	331
Number of Sequences with Hits	0
Number of 8mers	403
Number of Sequences with Exact 8mer Match	0

[a] Database search performed on 3 February 2016.

Based on the results of this study, the lack of potentially significant sequence similarity between the *Zm*PSY1 protein and known and putative allergens indicates that it is not a known allergen and is unlikely to be cross-reactive to known allergens.

B.2.6.2 CRTI Protein

To assess the potential for allergenic cross-reactivity, the 492-amino acid sequence encoded by the *crtI* gene was compared to a peer-reviewed database of 1956 known and putative allergen and celiac protein sequences residing in the FARRP16 dataset at the University of Nebraska (Oliva and MacKenzie, 2016c).[8]

Potential identities between the CRTI query sequence and proteins in the allergen database were evaluated with the FASTA35 sequence alignment tool using the default parameters.[9]

[7] The FASTA version 35.04 used on the AllergenOnline website uses the BLOSUM 50 scoring matrix (Henikoff and Henikoff, 1992), a "word size" of 2, and an expectation value score (*E*-score) cut-off of 1. The *E*-score is a parameter that describes the number of matches one can "expect" to see by chance when searching a database of a particular size. It decreases exponentially as the score of the match increases. The lower the *E*-score, or the closer it is to zero, the more "significant" the match.

[8] The Food Allergy Research and Resource Program (FARRP) allergen protein database resides at AllergenOnline (http://www.allergenonline.org). Version 16 was released on January 27, 2016, and contains 1956 peer-reviewed sequences representing 778 taxonomic-protein groups.

[9] The FASTA version 35.04 used on the AllergenOnline website uses the BLOSUM 50 scoring matrix (Henikoff and Henikoff, 1992), a "word size" of 2, and an expectation value score (*E*-score) cut-off of 1. The *E*-score is a parameter that describes the number of matches one can "expect" to see by chance when searching a database of a particular size. It decreases exponentially as the score of the match increases. The lower the *E*-score, or the closer it is to zero, the more "significant" the match.

The recommended greater than 35 percent identity threshold over any 80-amino acid length sequence alignment between the query sequence and an allergen was used to indicate the potential for cross-reactivity. No identity matches of > 35 percent over 80 residues were observed (Table 15).

The CRTI query sequence was also evaluated for any eight contiguous identical amino acid matches to the allergens contained in the FARRP database. This was done using an algorithm that generates all possible eight-word peptides from both the query and dataset proteins and evaluates each query "word" against all dataset "words" for perfect matches. There were no eight contiguous identical amino acid matches observed (Table 15).

Table 15. Search results using CRTI query sequence against the FARRP database

Database	AllergenOnline Database v16 (January 27, 2016)[a]
Input Query	>CRTI - phytoene desaturase, Pantoea ananatis MKPTTVIGA GFGGLALAI RLQAAGIPV LLLEQRDKP GGRAYVYED QGFTFDAGP TVITDPSAI EELFALAGK QLKEYVELL PVTPFYRLC WESGKVFNY DNDQTRLEA QIQQFNPRD VEGYRQFLD YSRAVFKEG YLKLGTVPF LSFRDMLRA APQLAKLQA WRSVYSKVA SYIEDEHLR QAFSFHSLL VGGNPFATS SIYTLIHAL EREWGVWFP RGGTGALVQ GMIKLFQDL GGEVVLNAR VSHMETTGN KIEAVHLED GRRFLTQAV ASNADVVHT YRDLLSQHP AAVKQSNKL QTKRMSNSL FVLYFGLNH HHDQLAHHT VCFGPRYRE LIDEIFNHD GLAEDFSLY LHAPCVTDS SLAPEGCGS YYVLAPVPH LGTANLDWT VEGPKLRDR IFAYLEQHY MPGLRSQLV THRMFTPFD FRDQLNAYH GSAFSVEPV LTQSAWFRP HNRDKTITN LYLVGAGTH PGAGIPGVI GSAKATAGL MLEDLI
Length	492
Number of 80mers	413
Number of Sequences with Hits	0
Number of 8mers	485
Number of Sequences with Exact 8mer Match	0

[a] Database search performed on 1 February 2016.

Based on the results of this study, the lack of potentially significant sequence similarity between the *Pantoea ananatis* CRTI protein and known and putative allergens indicates that it is not a known allergen and is unlikely to be cross-reactive to known allergens.

B.2.6.3 PMI Protein

To determine whether or not the PMI amino acid sequence showed biologically relevant amino acid sequence similarity to known or putative allergens, two different searches were performed against the FARRP AllergenOnline database, version 11.0, which contained 1,491 amino acid sequences of known and putative allergens (Harper, 2011b).

Potential identities between the PMI query sequence and proteins in the allergen database were evaluated with the FASTA35 sequence alignment tool using the default parameters. The recommended > 35 percent identity threshold over any 80-amino acid length sequence alignment between the query sequence and an allergen was used to indicate the potential for cross-reactivity. No identity matches of > 35 percent over 80 residues were observed.

The PMI query sequence was also evaluated for any eight contiguous identical amino acid matches to the allergens contained in the FARRP database. This was done using an algorithm that generates all possible eight-word peptides from both the query and dataset

proteins and evaluates each query "word" against all dataset "words" for perfect matches. There was one eight-amino acid identity match to a known allergen, α-parvalbumin from *Rana* species CH2001 (unidentified edible frog) (Hilger et al., 2002). This match was previously reported by Rabe (2004). Further investigation using sensitive serum screening methodology demonstrated no cross-reactivity between PMI and the serum from the single individual known to have demonstrated IgE-mediated allergy to this specific α-parvalbumin. The patient's serum did not recognize any portion of the PMI protein as an allergenic epitope. Therefore, the sequence identity between PMI and the α-parvalbumin from *Rana* species CH2001 is not biologically meaningful and has no implications for the potential allergenicity of PMI.

B.2.7. Heat Stability

For novel dietary proteins, the assessment of potential toxicity and allergenicity relies on a weight-of-evidence approach that includes an assessment of stability to heat or processing, in addition to other considerations such as amino acid sequence similarity to known protein toxins and allergens, and susceptibility to degradation in model digestion systems (Codex, 2003; FAO, 2001).

The potential for a protein to be denatured by heat treatment can be assessed *in vitro* by heating it in aqueous solution at temperatures ranging up to 100°C for 15–30 minutes, conditions commonly encountered during cooking or food processing (Hammond et al., 2013). Following heat treatment, the protein solution is returned to room temperature and the activity of the protein is measured. For enzymes, this involves a functional assay. Since heat denaturation does not necessarily result in protein degradation, heat stability studies are not generally applicable to structural proteins or proteins without known enzymatic or biological activities that can be tested (DBT, 2008).

B.2.7.1 ZmPSY1 Protein

The thermal stability of the *Zm*PSY1 protein was evaluated by measuring enzymatic activity using an high-pressure liquid chromatography (HPLC) method to monitor the production of 15-*cis*-phytoene from *in situ* produced geranylgeranyl diphosphate derived from the enzymatic conversion of dimethylallyl diphosphate and isopentenyl diphosphate in the presence of active geranylgeranyl diphosphate synthase (Welsch and Beyer, 2016). Samples of purified, microbial-expressed *Zm*PSY1 protein were subjected to heat treatment over a temperature incubation range of *ca*. 30–65°C for 15 minutes and then used in activity assays. Chloroform extracts of individual reaction mixtures were separated by reverse-phase HPLC and *Zm*PSY1 enzyme activity was calculated based on phytoene peak area measurements.

The microbial-expressed *Zm*PSY1 test substance was enzymatically active, catalyzing the production of 15-*cis*-phytoene from DMAPP and IPP, in the presence of active *A. thaliana* GGPP synthase, at the rate of *ca*. 28.4 pmol μg^{-1} min^{-1} under the assay conditions used in this study (Figure 27). Enzyme activity was irreversibly destroyed upon heat treatment, with 50 percent loss of activity following pre-incubation at *ca*. 42°C for 15 minutes and complete loss of activity following pre-incubation at 50°C for 15 minutes.

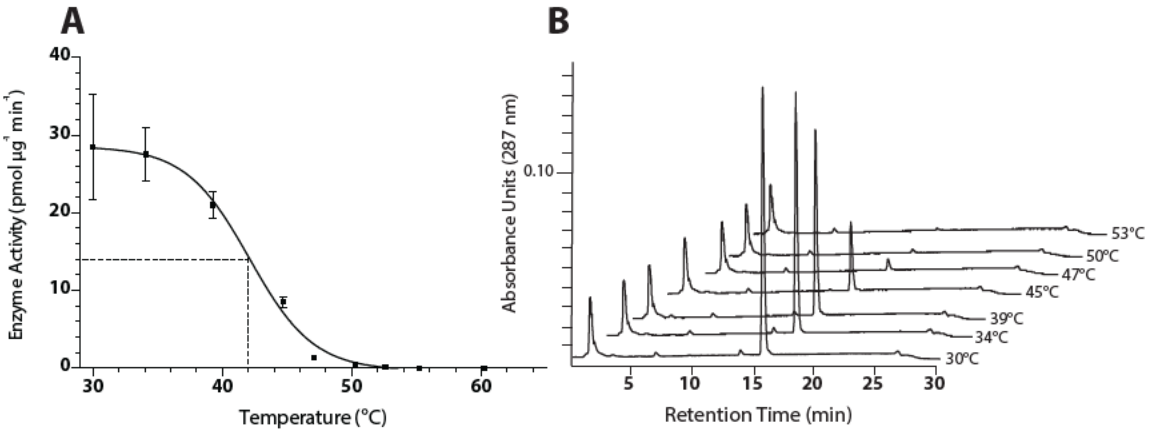


Figure 27. Individual samples of *Zm*PSY1 protein purified from recombinant *E. coli* (Lot No. M20452-05) were heated for 15 minutes at a designated temperature ranging from 30–65°C. Following this treatment, enzymatic production of 15-*cis*-phytoene was measured by HPLC. **Panel A** shows enzymatic activity (pmol μg^{-1} min^{-1}) versus pre-incubation temperature, where the values are means of two technical replicates and error bars represent the standard deviation around the mean value. **Panel B** shows HPLC chromatograms (287 nm) of chloroform:methanol extracts of selected *Zm*PSY1 activity assays. The area under the phytoene peak (retention time = 15.7 min) was used to calculate phytoene concentration. Pre-incubation temperatures are indicated to the right of each chromatogram trace.

The temperatures required to completely inactivate the *Zm*PSY1 enzyme were significantly lower than temperatures normally employed during cooking or processing, and it is therefore expected that dietary exposure to functional *Zm*PSY1 will be negligible.

B.2.7.2 CRTI Protein

The thermal stability of the phytoene desaturase (CRTI) protein was evaluated by measuring enzymatic activity using a spectrophotometric assay to monitor the conversion of liposome-incorporated 15-*cis*-phytoene to all-*trans*-lycopene (Schaub and Beyer, 2016). Samples of microbial-expressed CRTI protein (6 μg) were subjected to heat treatment over a temperature incubation range of *ca.* 30–60°C for 15 minutes, following which enzyme activity was measured at 37°C in the presence of 7 μM phytoene, 150 μM flavin adenine dinucleotide, 50 mM Tris-HCl pH 8.0, and 200 mM NaCl.

The microbial-expressed CRTI test substance was enzymatically active, catalyzing the conversion of liposome-incorporated phytoene to all-*trans*-lycopene at the rate of *ca.* 5.4 pmol all-*trans*-lycopene μg^{-1} min^{-1} under the assay conditions used in this study (Figure 28). Enzyme activity was irreversibly destroyed upon heat treatment, with 50 percent loss of activity following pre-incubation at *ca.* 51°C for 15 minutes and complete loss of activity following pre-incubation at 55°C for 15 minutes.

The temperatures required to completely inactivate the CRTI enzyme were significantly lower than temperatures normally employed during cooking or processing, and it is therefore expected that dietary exposure to functional CRTI will be negligible.

B.2.7.3 PMI Protein

The heat stability of PMI protein was investigated by measuring enzymatic activity following pre-incubation for 30 minutes at temperatures ranging from 25–95°C (Mims, 2009). PMI enzymatic activity was below that limits of quantification following pre-incubation at temperatures of 65°C and above.

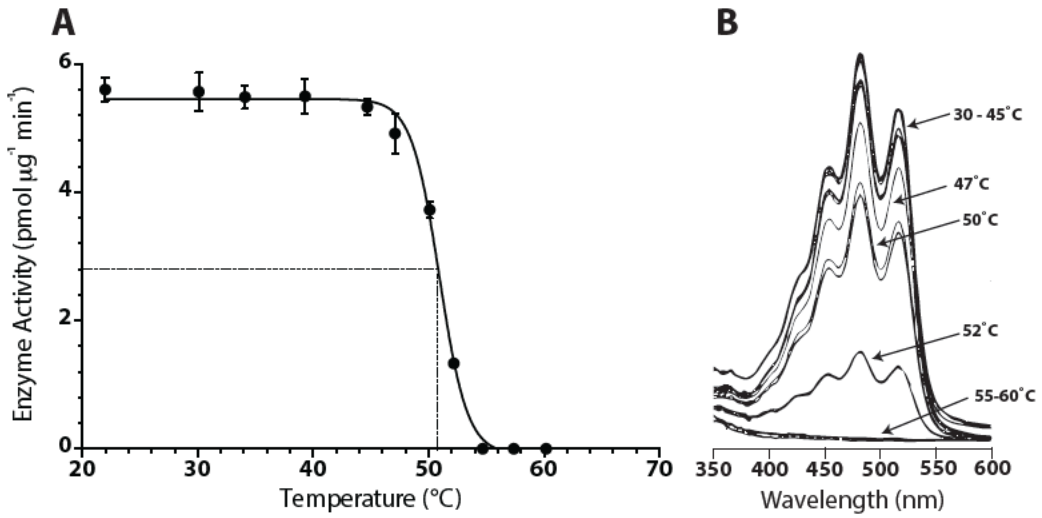


Figure 28. Individual samples of CRTI protein purified from recombinant *E. coli* (Lot No. M20454-02) were heated for 15 minutes at a designated temperature ranging from 30–60°C. Following this treatment, enzymatic conversion of 15-*cis*-phytoene to all-*trans*-lycopene was measured with a spectrophotometric assay. **Panel A** shows enzymatic activity (pmol μg^{-1} min^{-1}) versus pre-incubation temperature, where the values are means of two technical replicates and error bars represent the standard deviation around the mean value. **Panel B** illustrates the UV-visible spectral scans (350–600 nm) of the chloroform:methanol extracts of CRTI activity assays following different pre-incubation temperature regimes. The maximum absorbance at 483 nm was used to calculate the amount of all-*trans*-lycopene product.

The temperatures required to completely inactivate the PMI enzyme were significantly lower than temperatures normally employed during cooking or processing, and it is therefore expected that dietary exposure to functional PMI will be negligible.

B.2.8. Conclusions from Assessment of Potential Toxicity and Potential Allergenicity

B.2.8.1 ZmPSY1 Protein

The weight-of-evidence supporting the lack of identifiable hazards associated with *Zm*PSY1 included:

- The donor organism for the *Zm*PSY1-encoding gene was maize, a food crop with a long history of safe use.
- Based on the known expression of *Zm*PSY1 in yellow maize grain, there is a history of likely human and animal exposure to *Zm*PSY1 in the diet.
- The known function of *Zm*PSY1 in catalyzing the condensation of two molecules of geranylgeranyl diphosphate to yield 15-*cis*-phytoene is not similar to the activities of known protein toxins.
- Bioinformatics analyses showed that the *Zm*PSY1 protein did not display significant sequence similarity to known or putative protein toxins or allergens.
- The *Zm*PSY1 protein was rapidly degraded in simulated gastric fluid containing pepsin at pH 1.2. No intact *Zm*PSY1 or *Zm*PSY1-derived fragments were detected by western immunoblot labelling after 30 seconds of exposure to pepsin-containing SGF.

- The *Zm*PSY1 protein was rapidly and irreversibly inactivated following incubation for 15 minutes at 50°C, a temperature significantly lower than temperatures normally employed during cooking or processing of rice.

In combination, the scientific rationale and data described above support the conclusion that additional hazard characterization by animal toxicity testing is unnecessary and that the *Zm*PSY1 protein is unlikely to be toxic to humans or animals or result in allergic reaction.

B.2.8.2 CRTI Protein

The weight-of-evidence supporting the lack of identifiable hazards associated with CRTI included:

- The known function of CRTI in catalyzing the desaturation of phytoene to yield all-*trans*-lycopene is not similar to the activities of known protein toxins.
- Bioinformatics analyses showed that the CRTI protein did not display significant sequence similarity to known or putative protein toxins that would be considered a structural alert for potential toxicity, nor did it display any significant sequence similarity to known or putative allergens.
- The CRTI protein was rapidly degraded in simulated gastric fluid containing pepsin at pH 1.2. No intact CRTI or CRTI-derived fragments were detected after 30 seconds exposure to pepsin-containing SGF as assessed by Coomassie blue staining and western immunoblot labelling following SDS-PAGE analysis.
- The CRTI protein was rapidly and irreversibly inactivated following incubation for 15 minutes at 55°C, a temperature significantly lower than temperatures normally employed during cooking or processing of rice.

Additional hazard characterization through acute oral toxicity testing of CRTI protein in mice demonstrated a lack of any test substance-related adverse effects at a dosage of 100 mg/kg body weight, which represented a 115,000–fold margin of exposure relative to any conceivable human exposure (Table 11, page 50).

In combination, the scientific rationale and data described above support the conclusion that CRTI protein is unlikely to be toxic or allergenic to humans or animals.

B.2.8.3 PMI Protein

The weight-of-evidence supporting the lack of identifiable hazards associated with PMI included:

- The known function of PMI in catalyzing the reversible interconversion of mannose-6-phosphate and fructose-6-phosphate is ubiquitous in nature and not similar to the activities of known protein toxins.

- Bioinformatics analyses showed that the PMI protein did not display significant sequence similarity to known or putative protein toxins (Harper, 2011a), nor did it display any biologically meaningful sequence similarity to known or putative allergens (Harper, 2011b).
- The PMI protein was rapidly degraded in simulated gastric fluid containing pepsin at pH 1.2. No intact PMI or PMI-derived fragments were detected after one minute exposure to pepsin-containing SGF as assessed by western immunoblot labelling following SDS-PAGE analysis (Nelson, 2008).
- The PMI protein was rapidly and irreversibly inactivated following incubation for 30 minutes at 65°C (Mims, 2009), a temperature significantly lower than temperatures normally employed during cooking or processing of rice.

Additional hazard characterization through acute oral toxicity testing of PMI protein in mice demonstrated a lack of any test substance-related adverse effects at a dosage of 2000 mg/kg body weight (Korgaonkar, 2009). The tested dose represents a 69,000–fold margin of exposure relative to any conceivable human exposure due to consumption of GR2E rice (Table 11).

In combination, the scientific rationale and data described above, as previously reviewed by numerous regulatory authorities worldwide, support the conclusion that PMI protein is unlikely to be toxic or allergenic to humans or animals.

B.3. Other (Non-Protein) New Substances

B.3.1. Identity

Expression of the *Zm*PSY1 and CRTI enzymes in the endosperm tissue of GR2E rice results in the production of all-*trans*-lycopene from geranylgeranyl diphosphate (Figure 1). Lycopene is the substrate of two competing endogenous cyclases: ε-cyclase and β-cyclase, which acting together lead to the formation of α-carotene, whereas the action of β-cyclase alone forms β-carotene.

The principal carotenoid produced in GR2E rice endosperm is all-*trans*-β-carotene along with significantly lower amounts of other provitamin A carotenoids (e.g., all-trans-α-carotene, *cis* isomers of β-carotene, and β-cryptoxanthin) (Table 25).

The chemical name for β-carotene is (all-E)-1,1'-(3,7,12,16-tetramethyl-1,3,5,7,9,11,13,15,17-octadecanonaene-1,18-diyl)bis[2,6,6-trimethylcyclohexene]. The CAS Registry Number is 7235-40-7. The molecular formula is $C_{40}H_{56}$, its molecular weight is 536.88 g/mol and its structural formula is shown in Figure 29.

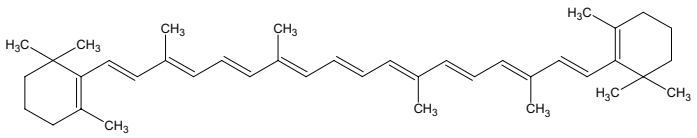


Figure 29. Chemical structure for β-carotene.

B.3.2. Sources and Dietary Intakes of β-Carotene

Beta-carotene is regarded as the major carotenoid present in the human diet (Johnson, 2002) and is found in significant quantities in a variety of green leafy and yellow-coloured vegetables and orange-coloured fruit (Table 16).

Table 16. β-Carotene content of selected foods

Food ID	Food Name	β-Carotene (mg/100 g)
13A12001	Basil, dried	27.13
13A12571	Sweet potato, chips, regular, purchased frozen, baked or roasted, with or without added fat	11.64
13A11835	Sweet potato, orange flesh, peeled or unpeeled, fresh or frozen, baked, roasted, fried, stir-fried, grilled or BBQ'd, no added fat	7.24
13A11688	Carrot, baby, peeled or unpeeled, fresh or frozen, raw	7.20
13A11794	Spinach, frozen, boiled, microwaved or steamed, drained	6.81
13A11712	Sweet potato, orange flesh, peeled or unpeeled, fresh or frozen, raw	6.60
13A12091	Mixed vegetables, for homemade mixed dishes, carrot, parsnip, radish, swede & turnip, raw	6.49
01B30372	Juice, carrot, home squeezed	6.07
13A12572	Sweet potato, orange flesh, mashed with fat not further defined	5.33
13A11901	Parsley, not further defined, raw	4.14
13A12084	Herbs, mixed, raw	3.21
13A11745	Tomato, sundried or semi-sundried	2.90
13A11788	Chives, raw	2.23
13A11795	Watercress, raw	1.98
13A11703	Spinach, fresh, raw	1.92
06D10478	Mango, peeled, raw	1.55
13A12623	Kale, raw	1.46
13A12095	Mixed leafy greens, for salad recipes, lettuce, spinach & rocket, raw	1.01
13A12127	Salad, garden, made from leafy greens, cucumber, onion & tomato, no added dressing	0.59
13A11683	Bok choy or choy sum, raw	0.55
10B10071	Olive, green or black, drained	0.28

† Data obtained from AUSNUT 2011-13 Food Nutrient Database. Available on the Internet at: http://www.foodstandards.gov.au/science/monitoringnutrients/ausnut/ausnutdatafiles/Pages/foodnutrient.aspxx (accessed 2 September 2016).

Based on a pooled analysis of 10 prospective cohort studies in North America and Europe that included more than 500,000 women, the mean daily intake of β-carotene is in the range of 2.7–6.4 mg (Koushik et al., 2006). Other data from the Australian Health Survey: Usual Nutrient Intakes, 2011-12, found that mean daily intakes of total provitamin A carotenoids range between 2.95–3.48 mg and 2.83–3.50 mg for adult men and women, respectively (Table 17).

Beta-carotene is an effective source of vitamin A in both conventional foods and vitamin supplements, and it is generally considered virtually nontoxic because humans tolerate high dietary dosages without apparent harm (Bendich, 1988; Hathcock et al., 1990). There are no reports of adverse effects arising from the consumption of naturally-occurring β-carotene in food. Standard toxicological tests, including teratogenic, mutagenic, and carcinogenic assays, have been performed on β-carotene without any evidence of harmful effects, as recently reviewed by EFSA (2012a). There is no evidence that conversion of β-carotene to vitamin A contributes to vitamin A toxicity, even when β-carotene is ingested in large amounts, and there is much circumstantial evidence that supplementary intakes of 15–50 mg/day are without side effects except for discolouration of the skin related to hypercarotenemia in some subjects at high intakes (Diplock, 1995). Dosages of β-carotene as high as 180 mg per day have been given to humans for several months during the treatment of light-sensitive skin diseases (e.g., protoporphyria) without observed adverse

Table 17. Daily dietary intakes of provitamin A carotenoids

		Age Group (years)[†]							
Percentiles		**2–3**	**4–8**	**9–13**	**14–18**	**19–30**	**31–50**	**51–70**	**> 70**
Males	**(mg/day)[a]**								
5		0.65	0.73	1.16	1.02	1.31	1.39	1.50	1.56
25		1.17	1.32	1.81	1.60	2.04	2.16	2.30	2.41
50		1.80	2.02	2.43	2.16	2.74	2.89	3.08	3.21
75		2.73	3.05	3.27	2.90	3.63	3.82	4.07	4.25
95		4.87	5.50	4.86	4.30	5.31	5.63	5.98	6.29
Mean Intake		2.16	2.42	2.65	2.35	2.95	3.12	3.32	3.48
Females	**(mg/day)**								
5		0.56	0.63	0.92	0.80	1.04	1.17	1.34	1.10
25		1.02	1.15	1.61	1.40	1.78	2.00	2.24	1.86
50		1.54	1.73	2.30	2.00	2.54	2.83	3.16	2.65
75		2.37	2.63	3.24	2.85	3.55	3.95	4.39	3.70
95		4.24	4.72	5.14	4.62	5.61	6.20	6.83	5.78
Mean Intake		1.87	2.10	2.58	2.27	2.83	3.15	3.50	2.94

[†] Data source: Australian Bureau of Statistics (43640DO003_20112012 Australian Health Survey: Usual Nutrient Intakes, 2011–12 — Australia). Available at http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0082011-12?OpenDocument

[a] Where information on levels of carotenes other than β-carotene in foods was available, this was included in the provitamin A total as β-carotene equivalents, according to the equation provitamin A = β-carotene + 0.5 × α-carotene + 0.5 × cryptoxanthin. This equation takes into account the differing biological activities of the different forms of provitamin A.

effects other than changes in skin colour (Mathews-Roth, 1986), which are reversible upon reduction of β-carotene intake.

B.3.3. Safety Considerations for β-Carotene Supplements

In 2012, the European Food Safety Authority (EFSA) Panel on Food Additives and Nutrient Sources added to Food (ANS) reviewed the possible link between the ingestion of high supplemental doses of β-carotene and cancer enhancement in heavy smokers. EFSA noted the findings of the ATBC (1994) (Alpha-Tocopherol Beta-Carotene Cancer Prevention) study and Carotenoid and Retinol Efficacy Trial (CARET) (Omenn et al., 1996) and also considered a meta-analysis of nine randomized controlled trials (RCTs) that demonstrated a lack of protection associated with β-carotene supplementation against cancer risk (Druesne-Pecollo et al., 2010). The meta-analysis, which included the ATBC and CARET trials, found an absence of any protective effect associated with β-carotene supplementation with regard to primary cancer risk and an increased risk of lung and stomach cancers in heavy smokers and asbestos workers supplemented with β-carotene at dose levels $\geq$ 20 mg/day. The authors, however, noted several significant caveats over the interpretation of their findings, in particular the importance of considering the influence of the dose and the population characteristics (e.g., nutritional deficiencies, exposure to carcinogens, gender, etc) when interpreting the results of RCTs.

In contrast, no increased lung cancer incidence was reported at supplemental dose levels of β-carotene varying from 6–15 mg/day for about five up to seven years (Druesne-Pecollo et al., 2010).

Other large studies did not show the same results as the ATBC and CARET studies. The Physicians' Health Study examined the effect of β-carotene supplementation (50 mg every other day) on cancer risk in more than 22,000 male physicians in the United States, of whom 11 percent were current smokers. In that population, β-carotene supplementation

for more than 12 years was not associated with an increased risk of lung cancer Hennekens et al. (1996).

Additionally, as reported by Baron et al. (2003), a clinical trial of the impacts of β-carotene (25 mg/day) and/or vitamins C and E (1,000 mg and 400 IU, respectively) indicated that among subjects who neither smoked nor drank alcohol, β-carotene strongly reduced the risk of recurrent colorectal adenomas; but among smokers and drinkers, β-carotene increased the risk. These data provide further evidence that β-carotene has different effects on smokers and nonsmokers.

The EFSA panel concluded that exposure to β-carotene from its use as a food additive and as a food supplement at a level below 15 mg per day does not give rise to concerns about adverse health effects in the general population. It also stated that no sensitive groups were identified from the available evidence at this exposure; therefore, the term *general population* encompasses all groups, including heavy smokers (EFSA, 2012c,a).

B.3.4. Potential Dietary Exposure to β-Carotene from GR2E Rice

Based on the highest rate of daily rice consumption reported for children in Bangladesh (12.5 g/kg body weight) used in the maximum dietary exposure estimate for *Zm*PSY1, CRTI, and PMI proteins (section B.1.4, page 47) and the highest concentration of β-carotene measured in samples of milled rice from field-grown plants of PSB Rc82 containing event GR2E (7.31 μg/g, Table 25, page 82), the maximum estimated daily dietary exposure to β-carotene is 91.4 μg/kg body weight. Using an average adult body weight of 57.7 kg in Asia (Walpole et al., 2012), this corresponds to a maximum daily intake of *ca*. 5.3 mg of β-carotene from GR2E rice.

In context, the maximum daily intake of β-carotene from GR2E rice based on the above scenario is equivalent to about 100 g of mashed, orange flesh, sweet potato (Table 16).

Based on results from the 2011-12 Australian Health Survey[10] average per capita daily rice consumption in Australia is 38.6 g for males and 25.8 g for females. Thus in the Australian context, the potential exposure to β-carotene from GR2E rice would be approximately 0.28 mg/day (using the higher consumption level for males), which is about the amount contained in a 100-g serving of green or black olives (Table 16), or about one-tenth the daily adult β-carotene consumption from all other food sources (Table 17).

More realistic dietary intakes are likely to be even lower as it is unlikely that all rice in the diet will be substituted with GR2E rice, and the known loss of β-carotene over time in all fruits and vegetables where it is found (Pénicaud et al., 2011; Mugode et al., 2014; Hidalgo and Brandolini, 2008; Oliveira and Carvalho, 2010), including GR2E rice, due to storage, processing, and cooking.

B.4. Toxicity of Novel Herbicide Metabolites

Not applicable.

[10] Australian Bureau of Statistics, 43640DO005_20112012 Australian Health Survey: Nutrition First Results — Foods and Nutrients, 2011-12 — Australia. Available at: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0072011-12?OpenDocument (Accessed 02 September 2016)

B.5. Compositional Analyses

For new varieties without purposefully altered nutritional properties, which includes the vast majority of currently authorized genetically modified crops, the compositional assessment is part of the weight-of-evidence approach for evaluating whether there were any unanticipated consequences of the genetic modification. The experience with genetically modified crop plants with introduced traits conferring insect-resistance and/or herbicide-tolerance has indicated that the incorporation of these traits has little biologically meaningful impact on the composition of key nutrients and anti-nutrients (Harrigan et al., 2010). As with products of conventional plant breeding, most compositional variation is due to environmental and agronomic factors, and the base genetics of the plant variety (Harrigan et al., 2007; Ricroch, 2013).

The compositional evaluation of crops intentionally modified to express altered nutritional properties is intended to determine whether the composition differs significantly from its traditional counterpart aside from the intended change in nutrient composition and to assess the safety of the intended change and any unintended changes. Nutritionally improved varieties may be expected to contribute significant new sources of dietary nutrients or other bioactive phytochemicals. To assess the safety and nutritional impact of these products, it is important to have knowledge of how much of these products will be consumed in the overall human diet or in animal feeds. The safety and nutritional quality of these products can only be assessed in the context of their proposed uses and consequent intake.

Rice grain (paddy) was chosen as the most appropriate test material for the compositional assessment of GR2E rice from a human food perspective as it the source of all rice food products used in commerce. Additionally, rice straw and processed bran were chosen as they are components of livestock animal feed. Analyses of key nutrients and anti-nutrients were performed on samples of paddy rice and straw collected from transgenic event GR2E and non-transgenic control rice grown over two growing seasons in 2015 and 2016 at four different locations representing typical rice growing conditions in the Philippines.

Compositional analyses were performed on samples of rice grain and straw obtained from event GR2E introgressed into PSB Rc82 (BC_5F_3 in 2015; BC_5F_4 in 2016; Figure 5, page 29) and near-isogenic control PSB Rc82 rice that were grown in side-by-side trials at four separate sites in the Philippines (Batac City, Los Baños, Muñoz, and San Mateo) during 2015 (Swamy et al., 2016a) and 2016 (Swamy et al., 2016b). At each site, planting and cultivation was done according to local agronomic practices. Three blocks (replicates) of each entry were established at each test site in a randomized complete block design. Grain and straw samples were collected from matured rice plants, the stage when typical grain harvest would occur. Grain samples were analyzed for key nutritional components, including proximates, fibre, sugars, fatty acids, amino acids, minerals, vitamins, and anti-nutrients. Samples of straw and bran were analyzed for proximates and minerals.

Information on the field trial locations and site management is provided in Appendix C (page 113) and further details on analytical procedures and statistical analyses are presented in Appendix E, page 127.

The first step in the compositional assessment was to test for differences in mean values between GR2E rice and the control PSB Rc82 rice. Where a statistically significant difference (p-value < 0.05) was identified in the multi-year combined-sites analysis, further context for interpreting the possible biological significance of the difference was gathered through comparisons with the range of values for each analyte reported in the published literature (OECD, 2016; NARO, 2011; Heuzé and Tran, 2015; Juliano and Bechtel, 1985) or available from the International Life Sciences Institute (ILSI) Crop Composition Database (ILSI, 2014). Analyte ranges for GR2E rice that fell within the combined literature range for that analyte were considered to be within the range of normal variability of conventional rice.

B.5.1. Proximates

The major constituents, or proximates, of rice straw and grain are carbohydrates, protein, fat and ash. Fibre is the predominant form of carbohydrate present in rice straw. Fibre is measured by the neutral detergent fibre (NDF) method, which measures the insoluble fibre: lignin, cellulose and hemicellulose. total dietary fibre (TDF) consists of the insoluble and soluble fibre (pectin). The acid detergent fibre (ADF) method solubilizes hemicellulose, measuring only cellulose and lignin.

There were no statistically significant differences in proximates and fibre between samples of straw obtained from GR2E and control PSB Rc82 rice (Table 18). With the exception of moisture content, which is dependent on the extent of drying of straw following harvest, the mean values for all proximates and fibre in rice straw were similar to the ranges reported in the literature.

Table 18. Proximate and fibre composition of straw harvested from event GR2E and control rice grown at four locations in the Philippines over two growing seasons in 2015 and 2016

	Proximates (% dry weight)[†]					Fibre (% dry weight)		
Samples	**Ash**	**Crude Fat**	**Crude Protein**	**CHO[a]**	**Moisture**	**ADF**	**NDF**	**Crude Fibre**
Event GR2E	25.73 (21.1–30.4)	2.58 (1.31–4.8)	6.13 (3.16–11.3)	66.12 (60.5–69.2)	12 (9.24–14)	52.6 (46.1–58)	62.53 (56.3–68.9)	30.1 (26.9–35.8)
Control PSBRc82	25.26 (20.6–30.2)	2.72 (0.995–6.15)	6.12 (2.85–10.2)	66.37 (62.1–70.5)	12.57 (8.12–22.1)	51.59 (44.7–59.3)	62.21 (50.2–69.7)	29.73 (26–34.9)
p-Value	0.4632	0.7668	0.974	0.7345	0.5932	0.5220	0.8184	0.6587
Literature Values								
ILSI (2014)	14.25 (10.8–18.9)	2.46 (1.92–3.52)	5.99 (4.02–8.33)	77.17 (71.0–81.6)	55.15 (41.7–73.7)	43.27 (36.1–55.3)	61.97 (51.9–70.3)	–[b]
Heuzé and Tran (2015)	18.1 (12.0–24.0)	1.4 (0.9–2.1)	4.2 (2.4–6.8)	–	7.2 (3.5–10.7)	42.4 (36.7–52.0)	69.1 (61.7–78.6)	35.1 (29.8–41.5)

[†] Values represent the least square mean of three replicate samples collected over two growing seasons from each of four locations in the Philippines (n=24 for each entry). For each analyte, the lowest and highest individual values across years and locations are shown in parentheses.
[a] CHO = carbohydrate by calculation; ADF = acid detergent fibre; NDF = neutral detergent fibre.
[b] Not reported.

In the multi-year combined-sites analysis, comparisons of proximates, fibre, and sugars in grain (paddy) samples derived from GR2E and control PSB Rc82 rice resulted in no statistically significant differences in any of the measured parameters (Table 19).

Table 19. Proximate, fibre, and sugar composition of grain harvested from event GR2E and control rice grown at four locations in the Philippines over two growing seasons in 2015 and 2016

	Proximates (% dry weight)†				
Samples	**Ash**	**Crude Fat**	**Crude Protein**	**CHO[a]**	**Moisture (%FW)**
Event GR2E	5.89 (4.95–7.17)	1.42 (0.838–2.16)	8.10 (6.07–11.2)	84.58 (81–86.9)	12.26 (11.1–13.8)
Control PSBRc82	6.02 (5–7.06)	1.34 (0.555–1.98)	8.26 (6.03–11.4)	84.38 (81.1–86.4)	12.32 (10.9–13.6)
p-Value	0.3739	0.7113	0.5445	0.5595	0.8023
Literature Values					
ILSI (2014)	4.77 (3.61–6.54)	2.76 (2.52–3.47)	8.55 (7.41–10.0)	83.91 (79.98–85.53)	16.85 (9.05–28.35)
Heuzé and Tran (2015)	5.9 (3.9–8.6)	2.1 (1.7–2.6)	8.3 (5.9–11.8)	–[b]	12.0 (7.6–16.4)

	Fibre (% dry weight)†					
Samples	**ADF**	**NDF**	**Crude Fibre**	**TDF**	**Amylose (%DW)**	**Starch (%DW)**
Event GR2E	18.53 (15.7–21.7)	22.1 (17.5–35.5)	11.96 (10.1–14.6)	16.96 (12.8–20.3)	12.87 (7.31–18.6)	59.52 (32.8–71.5)
Control PSBRc82	17.65 (15.6–18.8)	20.6 (16.2–32.8)	11.1 (10.1–12.3)	16.9 (11.4–21.4)	12.81 (6.76–18.6)	61.05 (28.1–73.9)
p-Value	0.3522	0.4774	0.2133	0.9549	0.9548	0.6888
Literature Values						
ILSI (2014)	15.06 (11.79–16.75)	18.49 (16.15–21.47)	14.51 (10.89–18.13)	19.15 (16.73–22.97)	–	–
Heuzé and Tran (2015)	13.3 (10.8–18.2)	21.5 (15.0–32.2)	11.1 (8.6–14.8)	–	–	64.2 (61.9–67.2)

† Values represent the least square mean of three replicate samples collected over two growing seasons from each of four locations in the Philippines (n=24 for each entry). For each analyte, the lowest and highest individual values across years and locations are shown in parentheses.
[a] CHO = carbohydrate by calculation; ADF = acid detergent fibre; NDF = neutral detergent fibre; TDF = total dietary fibre; DW = dry weight; FW = fresh weight.
[b] Not reported.

Table 20. Proximate, fibre, and mineral composition of bran produced from event GR2E and control rice

	Proximates (% dry weight)†				Minerals (g/kg DW)			Fibre (% dry weight)		
Samples	**Ash**	**Crude Fat**	**Crude Protein**	**CHO**	**Ca**	**P**	**Moisture**	**ADF[a]**	**NDF**	**Crude Fibre**
Event GR2E	10.3 (10-10.6)	24.1 (23.8-24.4)	15.1 (14.7-15.6)	50.4 (49.4-51.4)	0.618 (0.46-0.77)	25.9 (25.8-26)	11.5 (10.8-12.2)	16 (11.6-20.4)	23.7 (23.7-23.8)	8.97 (8.31-9.63)
Control PSBRc82	11.4 (10.7-12)	25.3 (24.6-26)	14.6 (14.1-15.2)	48.6 (47.7-49.5)	0.589 (0.55-0.63)	26.7 (25.5-27.8)	11.9 (10.9-12.9)	18 (13.8-22.1)	25.8 (24-27.6)	9.05 (9.01-9.08)
Literature Values										
Heuzé and Tran (2015)	9.4 (6.5–14)	17.2 (11.3–21.9)	14.8 (11.5–18.3)	48.7 (33–63.8)	0.7 (0.2–2.9)	17.0 (8.6–22.2)	9.9 (6.9–12.8)	11.2 (6.6–15.2)	25.2 (16.1–29.6)	8.6 (5.6–12.2)

† Values represent measured concentrations in bran produced from composite grain samples collected from four locations in the Philippines where event GR2E and control PSB Rc82 rice was grown during two growing seasons in 2015 and 2016.
[a] CHO = carbohydrate by calculation; ADF = acid detergent fibre; NDF = neutral detergent fibre; DW = dry weight; Ca = calcium; P = phosphorus.

Bulked samples of grain harvested from GR2E and control PSB Rc82 rice were processed into bran, which is a livestock feed ingredient derived from rice. Samples of processed bran were analyzed for proximates, fibre, and the minerals calcium and phosphorus (Table 20). The measured values for these analytes were within the respective ranges reported in the literature with the exception of crude fat and phosphorus, which were slightly higher in both GR2E and control PSB Rc82 rice.

B.5.2. Minerals

Several mineral ions are recognized as essential plant nutrients and are required by the plant in significant quantities. These macronutrients include calcium, phosphorous, and potassium. The micronutrient minerals, iron, copper, sodium, and zinc are incorporated in plant tissues in only trace amounts.

Table 21. Mineral composition of straw harvested from event GR2E and control rice grown at four locations in the Philippines over two growing seasons in 2015 and 2016

	Minerals (g/kg dry weight)†	
Samples	**Ca[a]**	**P**
Event GR2E	3.77 (1.75–7.06)	1.52 (0.82–3.61)
Control PSBRc82	3.64 (1.79–5.43)	1.47 (0.81–3.47)
p-Value	0.7033	0.6165
Literature Values		
Heuzé and Tran (2015)	2.9 (1.7–4.4)	0.9 (0.5–1.7)

† Values represent the least square mean of three replicate samples collected over two growing seasons from each of four locations in the Philippines (n=24 for each entry). For each analyte, the lowest and highest individual values across years and locations are shown in parentheses.
[a] Ca = calcium; P = phosphorus.

There were no statistically significant differences in concentrations of calcium and phosphorus measured in samples of GR2E and control PSB Rc82 rice straw across locations and growing seasons (Table 21). Mean values were within ranges for these two minerals reported in the literature.

Comparison of the mineral composition in samples of GR2E and control PSB Rc82 rice grain did not reveal any statistically significant differences in the concentrations of any measured analytes (Table 22). The mean concentrations of each of the minerals measured in samples from GR2E and control PSB Rc82 rice grain were within the ranges reported in the literature.

B.5.3. Amino Acids

A comparison of the amino acid composition of event GR2E and control PSB Rc82 rice grain is shown in Table 23. Across locations and growing seasons, there were no statistically

Table 22. Mineral composition of grain harvested from event GR2E and control rice grown at four locations in the Philippines over two growing seasons in 2015 and 2016

	Minerals (mg/100 g dry weight)[†]								
Samples	**Ca**[a]	**Cu**	**Fe**	**Mg**	**Mn**	**P**	**K**	**Na**	**Zn**
Event GR2E	22.5 (14.2–35)	0.391 (0.183–0.68)	3.96 (2.37–10.6)	130.5 (87.5–185)	6.61 (4.33–8.39)	327.1 (211–461)	345.9 (236–597)	1.50 (0.56–3.81)	2.31 (1.63–3.21)
Control PSBRc82	21.39 (15.3–29.8)	0.365 (0.215–0.506)	4.57 (2.58–9.08)	133.4 (102–157)	6.47 (4.85–7.78)	329.4 (241–383)	338.9 (222–472)	1.30 (0.543–3.07)	2.19 (1.73–2.78)
p-Value	0.5535	0.5441	0.5136	0.6312	0.6741	0.8698	0.6656	0.4668	0.5689
Literature Values									
ILSI (2014)	31.75 (25.1–42.7)	–[b]	5.64 (3.63–7.42)	–	–	289.1 (249–335)	–	–	–
Juliano and Bechtel (1985)	10–90	0.2–1.3	1.6–7.0	70–170	2.0–11	200–450	170–430	6.2–94	0.2–3.6
Heuzé and Tran (2015)	60 (20–150)	0.3 –	5.3 –	100 (30–140)	8.2 (4.6–11.7)	290 (190–470)	280 (190–350)	30.0 (0–100)	1.4 –

[†] Values represent the least square mean of three replicate samples collected over two growing seasons from each of four locations in the Philippines (n=24 for each entry). For each analyte, the lowest and highest individual values across years and locations are shown in parentheses.
[a] Ca = calcium; Cu = copper; Fe = iron; Mg = magnesium; Mn = manganese; P = phosphorus; K = potassium; Na = sodium; Zn = zinc.
[b] Not reported.

significant differences in the concentrations of any amino acids between samples of GR2E and control PSB Rc82 rice.

B.5.4. Vitamins

Samples of event GR2E and control PSB Rc82 rice grain were analyzed for concentrations of the water-soluble B vitamins (thiamine, riboflavin, niacin, pantothenic acid, pyridoxine, and folic acid), β-carotene, and α-tocopherol (Vitamin E). Alpha-tocopherol is the only form of vitamin E that is actively maintained in the human body and has the greatest nutritional significance (Linus Pauling Institute, 2004).

With the exception of β-carotene, which was intended to be elevated in GR2E rice, there were no statistically significant differences noted in the concentrations of any measured vitamins between GR2E and control PSB Rc82 rice in the combined-sites analysis over both growing seasons (Table 24).

Milled rice prepared from these grain (paddy) samples was also analyzed at IRRI for β-carotene and other carotenoids (Samia and Swamy, 2016). In these analyses, the concentrations of all-*trans*-β-carotene ranged from 1.96–7.31 $\mu g/g$ across locations and years (Table 25) and on average comprised *ca.* 59 percent of the total carotenoids as determined by HPLC.

The differences in β-carotene concentrations as determined at IRRI compared with values shown in Table 24 were attributed to a combination of differences in sample matrix (paddy *vs.* milled rice)[11] and differences in carotenoid extraction efficiencies between the Association of Official Analytical Chemists (AOAC) official method for β-carotene (941.15)

[11] Indica rice varieties have *ca.* 20 percent hull and 10 percent bran layers. For the study conducted at IRRI, typically, *ca.* 6.5 g of milled rice was produced from 10 g of rice paddy

Table 23. Amino acid composition of grain harvested from event GR2E and control rice grown at four locations in the Philippines over two growing seasons in 2015 and 2016

	Amino Acids (mg/100 g)†								
Samples	**Met**[a]	**Cys**	**Lys**	**Trp**	**Arg**	**Ile**	**His**	**Val**	**Leu**
Event GR2E	169.2 (124–228)	156.3 (117–214)	298.6 (216–443)	73.36 (44.2–107)	563.8 (409–737)	329.2 (241–446)	211.8 (162–276)	468.2 (341–627)	643.6 (464–886)
Control PSBRc82	166.2 (127–215)	155 (113–198)	294.4 (211–434)	74.93 (50.9–101)	564 (408–782)	331.9 (237–464)	214.5 (157–281)	474 (338–653)	650.5 (462–921)
p-Value	0.6019	0.7671	0.7111	0.7954	0.988	0.7871	0.6593	0.6987	0.743
Literature Values									
ILSI (2014)	190 (170–210)	180 (150–200)	290 (280–320)	100 (90–120)	570 (530–650)	300 (270–340)	220 (200–250)	430 (390–490)	620 (550–710)
Juliano and Bechtel (1985)	140–310	100–260	290–420	110–180	610–850	270–430	140–250	410–630	610–780
Samples	**Thr**	**Phe**	**Gly**	**Ala**	**Asp**	**Glu**	**Pro**	**Ser**	**Tyr**
Event GR2E	308.8 (235–400)	439.8 (324–593)	388.7 (293–495)	455.4 (329–625)	708.4 (493–1010)	1354 (942–1980)	376.4 (276–510)	400.6 (296–540)	213.5 (158–282)
Control PSBRc82	308.2 (217–409)	443.5 (317–622)	392.5 (292–511)	459.8 (331–628)	709.5 (497–994)	1360 (890–1990)	380.9 (278–521)	400.5 (231–556)	206.7 (133–291)
p-Value	0.9416	0.7992	0.7161	0.7404	0.9600	0.8975	0.6777	0.9975	0.5314
Literature Values									
ILSI (2014)	300 (270–330)	400 (360–440)	370 (340–420)	440 (380–500)	760 (680–850)	1240 (1100–1370)	350 (290–420)	400 (360–470)	140 (130–180)
Juliano and Bechtel (1985)	270–410	280–520	350–480	390–570	610–940	1310–1740	330–540	360–510	340–480

† Values represent the least square mean of three replicate samples collected over two growing seasons from each of four locations in the Philippines (n=24 for each entry). For each analyte, the lowest and highest individual values across years and locations are shown in parentheses.

[a] Met = methionine; Cys = cystine; Lys = lysine; Trp = tryptophan; Arg = arginine; Ile = isoleucine; His = histidine; Val = valine; Leu = leucine; Thr = threonine; Phe = phenylalanine; Gly = glycine; Ala = alanine; Asp = aspartic acid; Glu = glutamic acid; Pro = proline; Ser = serine; Tyr = tyrosine.

Table 24. Vitamin composition of grain harvested from event GR2E and control rice grown at four locations in the Philippines over two growing seasons in 2015 and 2016

	Vitamins (mg/kg dry weight)†						
Samples	**β-Carotene**[a]	**Thiamine** B1	**Niacin** B3	**Pantothenic Acid** B5	**Pyridoxine** B6	**Folic Acid** B9	**α-Tocopherol**
Event GR2E	1.26 (0.504–2.35)	3.13 (2.33–3.77)	36.9 (23.4–58.3)	9.15 (7.31–11.9)	2.72 (2.22–3.3)	0.912 (0.562–2.56)	2.98 (2.47–3.87)
Control PSBRc82	<LOQ [b] (<LOQ–0.07)	3.08 (2.35–3.81)	32.63 (20.2–48.8)	9.13 (7.22–11.4)	2.75 (2.1–5.42)	0.883 (0.393–1.54)	2.75 (2.1–3.5)
p-Value	NA[c]	0.6344	0.4171	0.9291	0.9185	0.881	0.3358
Literature Values							
ILSI (2014)	–[d]	5.61 (5.07–6.25)	54.26 (46.53–64.92)	–	–	–	–
Juliano and Bechtel (1985)	–	3.0–3.8	34–65	8–14	5–8	0.2–0.5	10–23

† Values represent the least square mean of three replicate samples collected over two growing seasons from each of four locations in the Philippines (n=24 for each entry). For each analyte, the lowest and highest individual values across years and locations are shown in parentheses.

[a] Concentrations of β-carotene are reported as all-*trans*-β-carotene.

[b] LOQ = Limit of quantification, which for β-carotene was 0.05 mg/kg dry weight. Values for riboflavin (vitamin B2) were below the LOQ of 0.9 mg/kg dry weight for all samples tested and are not included in this table.

[c] NA = not applicable. Statistical analysis was not performed as ≧ 80% of the values for an entry were below the LOQ.

[d] Not reported.

Table 25. Concentrations of carotenoids in milled rice from event GR2E and control rice grown at four locations in the Philippines over two growing seasons in 2015 and 2016

	Carotenoids (µg/g dry weight)[†]				
Samples	**β-Cryptoxanthin**	**all-*trans*-α-Carotene**	**all-*trans*-β-Carotene**	**9'-*cis*-β-Carotene**	**Total Carotenoids**
Event GR2E	0.312 (0.232–0.464)	0.713 (0.35–1.32)	3.57 (1.96–7.31)	0.762 (0.5–1.32)	5.88 (3.5–10.9)
Control PSBRc82	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ

[†] Values represent the mean of three replicate samples collected over two growing seasons from each of four locations in the Philippines (n=24 for each entry). For each analyte, the lowest and highest individual values across years and locations are shown in parentheses.
[a] LOQ = limit of quantification.

and the process employed at IRRI, which included a pre-incubation in 60°C water to swell the starch to improve extraction.

B.5.5. Fatty Acids

Three fatty acids account for *ca*. 93 percent of the total fatty acids in rice grain ILSI (2014). The most abundant fatty acids include oleic (C18:1 Δ9; *ca*. 40 percent) followed by linoleic (C18:2 Δ9,12; *ca*. 38 percent) and palmitic (C16:0; 15.4 percent). Less abundant, but occurring at measurable levels are stearic (C18:0; *ca*. 2 percent) and α-linolenic (C18:3 Δ9,12,15; *ca*. 1 percent).

The desaturation of oleic acid to form linoleic acid, and its subsequent desaturation to form α-linolenic acid, occurs only in plants, hence both linoleic and α-linolenic acids are essential fatty acids for mammals. For this reason, it was desirable to measure for any unintended changes in the levels of linoleic and α-linolenic acids, and their key precursors, palmitic, stearic and oleic acids, in grain from GR2E rice.

Other polyunsaturated and longer chain polyunsaturated fatty acids can all be synthesized by mammals from dietary sources of α-linolenic and linoleic acid. Additionally, the synthesis of palmitoleic (C16:1 Δ9) and saturated fatty acids with chain lengths greater than 18 (e.g., C20:0, C22:0, C24:0) can be accomplished in mammals through *de novo* fatty acid synthesis without dietary requirements for palmitic and stearic acids, respectively. Hence, small changes in the concentrations of any of these trace fatty acids in GR2E rice grain would have little or no biological significance to either humans or animals consuming GR2E rice products.

The complete fatty acid profile of grain from GR2E and control rice was determined and the results are summarized in Table 26. The concentrations of a number of fatty acids occurring in trace amounts in both GR2E and PSB Rc82 control rice grain samples were below the limit of quantification and are not reported (these are itemized in the footnotes to Table 26). In the combined-sites analysis over both growing seasons, the only statistically significant difference observed between GR2E and control PSB Rc82 rice samples was in the concentration stearic (C18:0) acid, which was approximately 6.5 percent higher for GR2E rice.

Table 26. Fatty acid composition of grain harvested from event GR2E and control rice grown at four locations in the Philippines over two growing seasons in 2015 and 2016

	Fatty Acids (% total fatty acids)†					
Samples	**Myristic** (C14:0)	**Palmitic** (C16:0)	**Palmitoleic** (C16:1)	**Stearic** (C18:0)	**Oleic** (C18:1)	**Linoleic** (C18:2)
Event GR2E	0.438 (0.346–0.518)	19.54 (18.8–20.4)	0.193 (0.16–0.218)	2.25 (1.95–2.78)	39.7 (38.4–41.3)	33.45 (32.4–34.4)
Control PSBRc82	0.393 (0.251–0.492)	18.45 (15.8–19.1)	0.191 (0.153–0.217)	2.11 (1.71–2.68)	40.23 (38.4–45.9)	34.03 (30.1–35.4)
p-Value	0.1983	0.2216	0.3614	***0.0487***	0.4866	0.3944
Literature Values						
ILSI (2014)	0.38 (0.32–0.48)	15.44 (14.9–16.94)	0.41 (0.26–0.93)	1.88 (1.68–2.09)	39.59 (37.49–40.49)	37.84 (37.51–38.49)
NARO (2011)[a]	0.70 (0.5–1.1)	21.90 (18.2–31.2)	0.20 (0.1–0.2)	2.00 (1.5–2.8)	36.90 (30.9–42.0)	34.70 (26.1–39.0)

Samples	**α-Linolenic** (C18:3)	**Arachidic** (C20:0)	**Eicosenoic** (C20:1)	**Behenic** (C22:0)	**Lignoceric** (C24:0)
Event GR2E	1.63 (1.34–2.31)	0.862 (0.739–0.989)	0.476 (0.405–0.538)	0.508 (0.407–0.585)	0.931 (0.687–1.25)
Control PSBRc82	1.64 (1.27–2.7)	0.887 (0.743–1.02)	0.515 (0.442–0.773)	0.543 (0.478–0.61)	1.00 (0.805–1.19)
p-Value	0.8871	0.1967	0.2857	0.142	0.1159
Literature Values					
ILSI (2014)	1.15 (1.12–1.21)	0.72 (0.66–0.79)	0.56 (0.54–0.58)	0.62 (0.48–0.82)	1.18 (1.06–1.34)
NARO (2011)	1.2 (0.9–1.6)	0.6 (0.4–0.7)	0.5 (0.4–0.6)	0.3 (0.2–0.6)	0.6 (0.4–0.9)

† Values represent the least square mean of three replicate samples collected over two growing seasons from each of four locations in the Philippines (n=24 for each entry). For each analyte, the lowest and highest individual values across years and locations are shown in parentheses. The concentrations of the following fatty acids were below the lower limit of quantification and are not reported: caprylic (C8:0); capric (C10:0); lauric (C12:0); pentadecanoic (C15:0); heptadecanoic (C17:0); eicosadienoic (C20:2); eicosatrienoic (C20:3); arachidonic (C20:4); erucic (C22:1); and nervonic (C24:1).

[a] Values from NARO (2011) are for brown rice.

B.5.6. Secondary Metabolites

Secondary metabolites are defined as those natural products that do not function directly in the primary biochemical activities that support growth, development, and reproduction of the organism in which they occur. Phytic acid (inositol hexakisphosphate; phytate when in salt form) is the main storage form of phosphorus in plant tissues, which is not in a bioavailable form for monogastric animals that lack the digestive enzyme phytase. Phytic acid also has strong binding affinity to important minerals such as calcium, magnesium, iron, and zinc, thus reducing the absorption of these minerals. Other than binding with minerals, phytic acid also binds to proteins, reducing their digestibility and thus amino acid bioavailability.

Trypsin inhibitors are proteins known to inhibit biologically active trypsin, interfere with digestion and ultimately absorption of food material, and act as anti-nutrients. Trypsin inhibitors are typical anti-nutritional components in soybeans, cereals, and potatoes. Proteinase inhibitors are of particular significance in animal nutrition causing growth depression and pancreatic hypertrophy.

Table 27. Secondary metabolites of grain harvested from event GR2E and control rice grown at four locations in the Philippines over two growing seasons in 2015 and 2016

Samples	Phytic Acid (% dry weight)	Trypsin Inhibitor (TIU/mg)
Event GR2E	0.861 (0.582–1.1)	0.924 (0.28–1.71)
Control PSBRc82	0.881 (0.605–1.23)	0.996 (0.032–4.17)
p-Value	0.6218	0.8282
Literature Values		
ILSI (2014)	0.76 (0.61–0.84)	–[a]

[†] Values represent the least square mean of three replicate samples collected over two growing seasons from each of four locations in the Philippines (n=24 for each entry). For each analyte, the lowest and highest individual values across years and locations are shown in parentheses.
[a] Not reported.

As recommended in the OECD consensus document on new rice varieties OECD (2016), grain samples from GR2E and control PSB Rc82 rice were analyzed for phytic acid. Although not required, samples were also analyzed for trypsin inhibitor activity. There were no statistically significant differences in the concentrations of phytic acid or in the levels of trypsin inhibitor between samples of GR2E and PSB Rc82 control rice (Table 27).

B.5.7. Conclusions from Compositional Analyses

Among the 69 compositional components that were assessed in samples of GR2E and control PSB Rc82 rice grain, and 10 components that were assessed in straw samples, the only statistically significant difference observed from the multi-year combined-site analysis was for stearic (C18:0) fatty acid measured in grain samples (not including the intended difference in β-carotene levels) (Table 28). With the exception of provitamin A carotenoids, the compositional parameters measured in samples of GR2E rice were within or similar to the range of natural variability of those components in conventional rice varieties with a history of safe consumption. Overall, no consistent patterns emerged to suggest that biologically meaningful changes in composition or nutritive value of the grain or straw had occurred as an unintended consequence of the genetic modification.

Table 28. Summary of statistically significant differences observed in the multi-year combined-site analysis of compositional parameters measured for GR2E and control PSB Rc82 rice

Analytical Component (units)	Event GR2E Mean[†]	PSB Rc82 Control Mean	Mean Difference (% of control)	Significance (*p*-Value)	Combined Literature Range
Grain Fatty Acids (% total fatty acids)					
Stearic (C18:0)	2.25	2.11	6.5%	0.0487	1.5–2.8
Grain Vitamins (mg/kg dry weight)					
β-Carotene	1.26	<LOQ[a]			

[†] Values represent the least square mean of three replicate samples collected over two growing seasons from each of four locations in the Philippines (n=24 for each entry).
[a] LOQ = Limit of quantification, which for β-carotene was 0.05 mg/kg dry weight.

C. Information Related to Nutritional Impact

C.1. Vitamin A Deficiency

Vitamin A refers to a group of fat-soluble retinoids, including retinol, retinal, and retinyl esters, and has long been understood to help regulate immune function and to reduce morbidity of infectious diseases (Green and Mellanby, 1928). Vitamin A is required for normal functioning of the visual system, maintenance of cell function for growth, epithelial integrity, production of red blood cells, immunity, and reproduction. Vitamin A is an essential nutrient in humans that cannot be synthesized *de novo* in the body, so it must be obtained through the diet (Bates, 1995). Different forms of vitamin A can be obtained from food, either preformed vitamin A in animal products, such as meat, eggs and dairy products, or provitamin A carotenoids, mainly β-carotene in plant products, such as green leafy and many types of yellow-coloured vegetables and orange-coloured fruit.

In Western societies, the provitamin A carotenoids derived from plants provide less than 30 percent of daily vitamin A intake, with preformed vitamin A derived from animal products providing the balance of daily vitamin A intake. The situation is reversed in many developing countries, where provitamin A carotenoids in vegetables and fruit are the principal sources of dietary vitamin A.

It has historically been assumed that provitamin A derived from various sources in a mixed diet generally yields an retinol activity equivalent (RAE) ratio of 12:1. When referring to specific carotenoids in the diet, it is assumed that the RAE ratio for all-*trans*-β-carotene is 12:1, whereas for β-cryptoxanthin and α-carotene, the ratio is 24:1. The provitamin A activity of 9-*cis* and 13-*cis* isomers of β-carotene is less than 10 percent of all-*trans*-β-carotene (Harrison, 2012). The bioavailability of provitamin A carotenoids in specific fruits and vegetables is highly variable, with an almost eight-fold difference in β-carotene conversion factors (on a weight basis), ranging from 3.6:1 to 28:1 for Golden Rice and leafy vegetables, respectively (Tang et al., 2009; Tang, 2010; Tang et al., 2012). The major factor that affects the vitamin A value of plant provitamin A carotenoids is the food matrix. However, it must be recognized that human subjects may have different abilities to convert provitamin A carotenoids to vitamin A, and that these differences may be due to the vitamin A status of the individual and also genetic variability in β-carotene metabolism.

It is difficult for children to fulfil their daily vitamin A requirements through plant foods alone, and consequently, vitamin A deficiency (VAD)[12] is common among children whose families cannot afford eggs and dairy products. The main underlying cause of VAD as a public health problem is a diet that is chronically insufficient in vitamin A that can lead to lower body stores that fail to meet physiologic needs.

In southeast Asian countries, the average per capita rice consumption is 130 kg/year or 360 g/day, providing nearly two-thirds of caloric intake. The high consumption of rice may reflect a lack of dietary diversity, which, when combined with the poor micronutrient

[12] Vitamin A deficiency exists when serum retinol levels are < 0.7 μmol/l. Mild deficiency is usually evident at serum retinol levels ranging between 0.7–$<$1.05 μmol/l, and normal status is regarded as $\geq$1.05 μmol/l serum retinol.

content of polished rice in general and lack of any provitamin A, specifically, is a risk factor for VAD (de Pee, 2014). Deficiency of sufficient duration or severity can lead to disorders that are common in vitamin A deficient populations, such as xerophthalmia (xeros = dryness; -ophthalmia = pertaining to the eye), the leading cause of preventable childhood blindness, anaemia, and weakened host resistance to infection, which can increase the severity of infectious diseases and risk of death (Scrimshaw and SanGiovanni, 1997; Christian et al., 1998). VAD increases vulnerability to a range of illnesses including diarrhoea, measles, and respiratory infections, which are the leading causes of mortality among children in low and middle income countries, where risk of infection and risk of mortality can be compounded by coexisting under-nutrition. The prevalence and severity of xerophthalmia, anaemia and the less-measurable "vicious cycle" between VAD and infection in vulnerable groups, notably young children and pregnant or lactating mothers, represent the most compelling consequences of VAD and underlie its significance as a public health problem around the world.

The WHO (2009) estimated that approximately 33 percent of the world's preschool age children are vitamin A deficient, (Table 29) and the situation is worse in Africa and southeast Asia where 44–50 percent of preschool children suffer from VAD.

Table 29. Prevalence of vitamin A deficiency by WHO region, 1995–2005

	Preschool Age Children[a]		Pregnant Women	
WHO Region	**Prevalence (%)[b]**	**People Affected (millions)**	**Prevalence (%)**	**People Affected (millions)**
Africa	44.4	56.4	13.5	4.18
Americas	15.6	8.68	2.0	0.23
South-East Asia	49.9	91.5	17.3	6.69
Europe	19.7	5.81	11.6	0.72
Eastern Mediterranean	20.4	13.2	16.1	2.42
Western Pacific	12.9	14.3	21.5	4.90
Global[c]	**33.3**	**190**	**15.3**	**19.1**

[a] Population subgroups: Preschool age children (< 5 years); Pregnant women (no age range defined).
[b] Excludes countries with a 2005 per capita Gross Domestic Product ≧ US$ 15,000.
[c] Source: (WHO, 2009).

Although the global prevalence has been declining over the past 20 years, VAD persists as a major public health nutrition problem, especially in sub-Saharan Africa and south Asia. In these two regions, VAD prevalence has changed little since 1991 and remains high at 48 percent and 44 percent, respectively (Stevens et al., 2015). This contrasts with significant reductions noted in other regions of the world, including east and southeast Asia and Oceania where VAD prevalence has fallen to approximately 6 percent (Stevens et al., 2015). Nevertheless, pockets of deficiency persist in many countries in this region, particularly in the context of decentralized health systems.

Since 2003 there have been significant declines in the prevalence of VAD in the Philippines, particularly among pregnant and lactating women where the most recently reported rates are nine and five percent, respectively. There have also been improvements in the vitamin A status of preschool age children; however, VAD remains a severe problem, ranging between 19.6–27.9 percent for 1–5-year and 6 to ≤ 12 month olds, respectively (FNRI, 2016).

In Bangladesh, a micronutrient status survey was conducted in 2011–2012 to provide representative national, rural, urban and urban slum status estimates (ICDDR, 2013). The prevalence of VAD among preschool and school age children suggest a serious public health problem, with respective rates found to be 20.5 percent and 20.9 percent, respectively. At the national level over half of the preschool age (56.3 percent) and school age children (53.3 percent) were found to exhibit at least a mild grade of VAD.[13] Significantly lower rates of VAD, 5.4 percent, were found for non-pregnant, non-lactating women although about one-third were affected by mild deficiency. The problem was more pronounced in urban slums areas rather than rural areas for all target groups, but was almost double in urban slums among children 6–59 months of age (38.1 percent) compared to rural areas (19.4 percent) (ICDDR, 2013).

C.2. Intended Nutritional Effect of GR2E Rice

Vitamin A intakes or requirements are generally expressed in terms of RAE, where one RAE is defined as the biological activity associated with 1 μg of all-*trans*-retinol. The ability of provitamin A compounds, such as β-carotene, to meet vitamin A requirements is limited by factors such as losses during storage, processing, and/or cooking, and, as already noted, the conversion efficiency of dietary β-carotene to RAEs.

The EARs[14] and RDAs[15] for vitamin A, expressed as RAEs and broken down by gender and age-group, are shown in Table 30.

The intended nutritional effect of GR2E rice is to complement existing VAD control efforts by supplying up to 30–50 percent of the EAR for vitamin A for preschool age children and pregnant or lactating mothers in high-risk countries, including Bangladesh, Indonesia, and the Philippines.

D. Other Information

D.1. Lack of Justification for Animal Feeding Studies

Compositional analysis is the cornerstone of the nutritional assessment of a food derived from a new plant variety. When compositional equivalence between the new food and its conventional counterpart has been established, the results of numerous published livestock feeding trials with modified varieties of maize, soybean, canola, cotton, and sugar beet, have confirmed that there are no significant differences in digestibility of nutrients, animal health, or animal performance (Flachowsky et al., 2005; Van Eenennaam and Young, 2014). Therefore, once compositional equivalence has been established, feeding studies with livestock species add little to the safety assessment (EFSA, 2006).

[13] Serum retinol $\geq$0.7 to $<$1.05 μmol/l.

[14] The EAR is the median daily intake value that is estimated to meet the requirement of half the healthy individuals in a life-stage and gender group. At this level of intake, the other half of the individuals in the specified group would not have their needs met. The EAR is used to calculate the recommended daily allowance (RDA). It is also used to assess the adequacy of nutrient intakes, and can be used to plan the intake of groups.

[15] The RDA is the average daily dietary intake level that is sufficient to meet the nutrient requirement of nearly all (97–98 percent) healthy individuals in a particular life-stage and gender group. The RDA is the goal for usual intake by an individual.

Table 30. Vitamin A dietary reference intakes

	µg per day (RAE)[a]			IU[b] per day (RAE)		
	EAR[c]	RDA[d]	UL[e]	EAR	RDA	UL
Infants[f]						
0–6 months	ND	400	600	ND	1333	2000
7–12 months	ND	500	600	ND	1667	2000
Children						
1–3 yr	210	300	600	700	1000	2000
4–8 yr	275	400	900	917	1333	3000
Males						
9–13 yr	445	600	1700	1483	2000	5667
14–18 yr	630	900	2800	2100	3000	9333
> 18 yr	625	900	3000	2083	3000	10000
Females						
9–13 yr	420	600	1700	1400	2000	5667
14–18 yr	485	700	2800	1617	2333	9333
> 18 yr	500	700	3000	1667	2333	10000
Pregnancy						
≦ 18 yr	530	750	2800	1767	2500	9333
> 18 yr	550	770	3000	1833	2567	10000
Lactation						
≦ 18 yr	885	1200	2800	2950	4000	9333
> 18 yr	900	1300	3000	3000	4333	10000

[a] RAE = retinol activity unit.
[b] IU = international unit
[c] The EAR is the median daily intake value that is estimated to meet the requirement of half the healthy individuals in a life-stage and gender group. At this level of intake, the other half of the individuals in the specified group would not have their needs met. The EAR is used to calculate the RDA. It is also used to assess the adequacy of nutrient intakes, and can be used to plan the intake of groups.
[d] The RDA is the average daily dietary intake level that is sufficient to meet the nutrient requirement of nearly all (97–98 percent) healthy individuals in a particular life-stage and gender group. The RDA is the goal for usual intake by an individual.
[e] UL = upper limit for pre-formed vitamin A only. There currently is no defined UL for β-carotene.
[f] Source: Health Canada dietary reference intakes for vitamins (http://www.hc-sc.gc.ca/fn-an/nutrition/reference/table/ref_vitam_tbl-eng.php) and the vitamin A nutrient reference intakes for Australia New Zealand (https://www.nrv.gov.au/nutrients/vitamin-a).

With the exception of elevated levels of provitamin A carotenoids (mainly β-carotene), GR2E rice is compositionally equivalent to conventional rice.

In addition to its natural occurrence in a range of foods, β-carotene is also approved as a food additive and colourant in amounts consistent with good manufacturing practice (i.e., without maximum permissible quantities). Beta-carotene is also used as an additive in animal nutrition as a source of provitamin A. The safety and efficacy of β-carotene supplementation of animal feed was recently reviewed by the EFSA Panel on Additives and Products or Substances used in Animal Feed (FEEDAP) who concluded that the use of β-carotene is safe for the target animals and that setting a maximum content in feed legislation was not considered necessary (EFSA, 2012b).

Considering the data from the compositional assessment of GR2E rice and the lack of any observed phenotypic characteristics indicative of unexpected, unintended, effects arising from the genetic modification process, and considering the established history of safe use of β-carotene, there are no plausible risk hypotheses that would indicate the need for animal feeding studies.

E. Conclusions

The purpose of this evaluation of GR2E rice was to determine whether the use of GR2E rice in food or feed could raise any new safety concerns relative to conventional rice, and was not intended to address questions related to the efficacy of GR2E rice in helping combat VAD in at-risk populations.

The assessment of GR2E rice has included a complete description of the genetic modification (e.g., gene sources, characterization of inserted DNA and site of integration within the host genome, stability, and inheritance), the safety of the newly expressed proteins (e.g., history of use in food, function, potential toxicity, potential allergenicity, and patterns and levels of expression), and a comprehensive nutrient compositional assessment to identify whether there were any unintended, unexpected, effects of the genetic modification.

Collectively, the data presented in this submission have not identified potential health and safety concerns, and support the conclusion that food and/or livestock animal feed derived from provitamin A enriched GR2E rice is as safe and nutritious as food or feed derived from conventional rice varieties.

F. References

Ali, S. A., Hamid, F., Abbasi, A., Zaidi, Z. H., and Shehnaz, D. (1999). Pharmacological effects of the leaf-nosed viper snake (Eristocophis macmahoni) venom and its HPLC fractions. *Toxicon*, 37(8):1095–1107.

Asfaw, A., Gurum, F., Alemayehu, F., and Rezene, Y. (2012). Analysis of multi-environment grain yield trials in mung bean Vigna radiate (L.) Wilczek based on GGE biplot in southern Ethiopia. *Journal of Agricultural Science and Technology*, 14(2):389–398.

Astwood, J. D., Leach, J. N., and Fuchs, R. L. (1996). Stability of food allergens to digestion in vitro. *Nature Biotechnology*, 14(10):1269–1273.

ATBC (1994). The effect of vitamin E and beta-carotene on the incidence of lung cancer and other cancers in male smokers. *New England Journal of Medicine*, 330:1029–1035.

Bannon, G., Goodman, R. E., Leach, J. N., Rice, E., Fuchs, R. L., and Astwood, J. D. (2002). Digestive stability in the context of assessing the potential allergenicity of food proteins. *Comments on Toxicology*, 8:271–285.

Bannon, G. A. and Ogawa, T. (2006). Evaluation of available IgE-binding epitope data and its utility in bioinformatics. *Molecular Nutrition and Food Research*, 50(7):638–644.

Baron, J. A., Cole, B. F., Mott, L., Haile, R., Grau, M., Church, T. R., Beck, G. J., and Greenberg, E. R. (2003). Neoplastic and Antineoplastic Effects of Âă-Carotene on Colorectal Adenoma Recurrence: Results of a Randomized Trial. *JNCI Journal of the National Cancer Institute*, 95(10):717–722.

Bates, C. (1995). Vitamin A. *The Lancet*, 345:31–35.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-6 (http://cran.r-project.org/package=lme4).

Beitz, E. (2000). TEXshade: shading and labeling of multiple sequence alignments using LATEX2 epsilon. *Bioinformatics (Oxford, England)*, 16(2):135–139.

Bendich, A. (1988). The safety of beta-carotene. *Nutrition and Cancer: An International Journal*, 11(4):207–214.

Besler, M. and Tanabe, S. (2001). Allergen Data Collection - Update: Rice. *Internet Symposium on Food Allergens*, 3:1–17.

Beyer, P., Al-Babili, S., Ye, X., Lucca, P., Schaub, P., Welsch, R., and Potrykus, I. (2002). Golden Rice: introducing the beta-carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat vitamin A deficiency. *The Journal of Nutrition*, 13(3):506S–510S.

Black, R. E., Allen, L. H., Bhutta, Z. A., Caulfield, L. E., de Onis, M., Ezzati, M., Mathers, C., and Rivera, J. (2008). Maternal and child undernutrition: global and regional exposures and health consequences. *The Lancet*, 371(9608):243–260.

Bøgh, K. L. and Madsen, C. B. (2015). Food allergens: is there a correlation between stability to digestion and allergenicity? *Critical Reviews in Food Science and Nutrition*.

Buckner, B., San Miguel, P., Janick-Buckner, D., and Bennetzent, J. L. (1996). The y1 gene of maize codes for phytoene synthase. *Genetics*, 143:479–488.

Burkhardt, P. K., Beyer, P., Wünn, J., Klöti, A., Armstrong, G. A., Schledz, M., Von Lintig, J., and Potrykus, I. (1997). Transgenic rice (Oryza sativa) endosperm expressing daffodil

(Narcissus pseudonarcissus) phytoene synthase accumulates phytoene, a key intermediate of provitamin A biosynthesis. *Plant Journal*, 11(5):1071–1078.

Cazzonelli, C. I. and Pogson, B. J. (2010). Source to sink: regulation of carotenoid biosynthesis in plants. *Trends in Plant Science*, 15(5):266–274.

CERA (2016). GM Crop Database.

Chen, B. Y., Heneen, W. K., and Simonsen, V. (1989). Comparative and genetic studies of isozymes in resynthesized and cultivated Brassica napus L., B. campestris L. and B. alboglabra Bailey. *Theoretical and Applied Genetics*, 77(5):673–9.

Christensen, A. H., Sharrock, R. A., and Quail, P. H. (1992). Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Molecular Biology*, 18(4):675–689.

Christian, P., Schulze, K., Stoltzfus, R. J., and West, K. P. (1998). Hyporetinolemia, illness symptoms, and acute phase protein response in pregnant women with and without night blindness. *The American Journal of Clinical Nutrition*, 67(6):1237–43.

Codex (2003). Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants CAC/GL 45-2003.

Coruzzi, G., Broglie, R., Edwards, C., and Chua, N.-H. (1984). Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1,5-bisphosphate carboxylase. *The EMBO Journal*, 3(8):1671–1679.

Coutinho, T. A. and Venter, S. N. (2009). Pantoea ananatis: an unconventional plant pathogen. *Molecular Plant Pathology*, 10(3):325–35.

Cressman, R. F. and Ladics, G. (2009). Further evaluation of the utility of "Sliding Window" FASTA in predicting cross-reactivity with allergenic proteins. *Regulatory Toxicology and Pharmacology*, 54(3 SUPPL.).

Cueto, M., Salcedo, M., Oliva, N., and Trijatmiko, K. (2016). Southern hybridization characterization of event IR-ØØGR2E-5 rice. Technical report, IR2015-07003 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

Dailey, T. A. and Dailey, H. A. (1998). Identification of an FAD superfamily containing protoporphyrinogen oxidases, monoamine oxidases, and phytoene desaturase. Expression and characterization of phytoene desaturase of Myxococcus xanthus. *Journal of Biological Chemistry*, 273(22):13658–13662.

DBT (2008). Protocols for food and feed safety assessment of GE crops. Technical report, Department of Biotechnology, MoST, Government of India.

de Lonlay, P. and Seta, N. (2009). The clinical spectrum of phosphomannose isomerase deficiency, with an evaluation of mannose treatment for CDG-Ib. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1792(9):841–843.

De Maayer, P., Chan, W. Y., Rubagotti, E., Venter, S. N., Toth, I. K., Birch, P. R., and Coutinho, T. A. (2014). Analysis of the Pantoea ananatis pan-genome reveals factors underlying its ability to colonize and interact with plant, insect and vertebrate hosts. *BMC Genomics*, 15:404.

De Moura, F. F., Palmer, A. C., Finkelstein, J. L., Haas, J. D., Murray-Kolb, L. E., Wenger, M. J., Birol, E., Boy, E., and Peña-Rosas, J. P. (2014). Are biofortified staple food crops

improving vitamin A and iron status in women and children? New evidence from efficacy trials. *Advances in Nutrition (Bethesda, Md.)*, 5(5):568–70.

de Pee, S. (2014). Proposing nutrients and nutrient levels for rice fortification. *Annals of the New York Academy of Sciences*, 1324(1):55–66.

Delaney, B., Astwood, J. D., Cunny, H., Conn, R. E., Herouet-Guicheney, C., MacIntosh, S., Meyer, L. S., Privalle, L., Gao, Y., Mattsson, J., and Levine, M. (2008). Evaluation of protein safety in the context of agricultural biotechnology. *Food and Chemical Toxicology*, 46(SUPPL. 2):872–897.

DellaPenna, D. and Pogson, B. J. (2006). Vitamin synthesis in plants: tocopherols and carotenoids. *Annual Review of Plant Biology*, 57(1):711–38.

Depicker, A., Stachel, S., Dhaese, P., Zambryski, P., and Goodman, H. M. (1982). Nopaline synthase: transcript mapping and DNA sequence. *Journal of Molecular and Applied Genetics*, 1(6):561–573.

Diplock, A. T. (1995). Safety of antioxidant vitamins and beta-carotene. *The American Journal of Clinical Nutrition*, 62(6 Suppl):1510S–1516S.

Doyle, J. J. and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19:11–15.

Druesne-Pecollo, N., Latino-Martel, P., Norat, T., Barrandon, E., Bertrais, S., Galan, P., and Hercberg, S. (2010). Beta-carotene supplementation and cancer risk: A systematic review and metaanalysis of randomized controlled trials. *International Journal of Cancer*, 127(1):172–184.

EFSA (2006). Guidance document of the scientific panel on genetically modified organisms for the risk assessment of genetically modified plants and derived food and feed. *EFSA Journal*, 99:1–100.

EFSA (2012a). Scientific Opinion on the re-evaluation of mixed carotenes (E 160a (i)) and beta-carotene (E 160a (ii)) as a food additive. *EFSA Journal*, 10(3):2593.

EFSA (2012b). Scientific opinion on the safety and efficacy of beta-carotene as a feed additive for all animal species and categories. *EFSA Journal*, 10(6):2737.

EFSA (2012c). Statement on the safety of β-carotene use in heavy smokers. *EFSA Journal*, 10(12):2953.

Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein science : A Publication of the Protein Society*, 8(5):978–984.

Ercolini, D., Russo, F., Torrieri, E., Masi, P., and Villani, F. (2006). Changes in the spoilage-related microbiota of beef during refrigerated storage under different packaging conditions. *Applied and Environmental Microbiology*, 72(7):4663–4671.

Ewing, B. and Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research*, 8(3):186–194.

FAO (2001). Evaluation of Allergenicity of Genetically Modified Foods. Technical Report January, Food and Agriculture Organization, United Nations.

FAO/WHO (1996). Strategies for assessing the safety of foods produced by biotechnology. Technical report, Joint FAO/WHO Expert Consultation. Food and Agriculture Organization, World Health Organization, Geneva, Switzerland.

FDA (2005). New protein consultation: phosphomannose isomerase (PMI). Technical report, US Food and Drug Administration.

Flachowsky, G., Chesson, A., and Aulrich, K. (2005). Animal nutrition with feeds from genetically modified plants. *Archives of Animal Nutrition*, 59:1–40.

FNRI (2016). Vitamin A deficiency (VAD) among Filipino preschool children, pregnant, and lactating women: 1993-2013. In *42nd FNRI Seminar Series, First 100 Days*, pages 1–12, Los Banos, Laguna, Philippines.

FSANZ (2006). Final assessment report: Application A564, food derived from insect-protetcted corn line MIR604. Technical report, Food Standards Australia New Zealand, Canberra, BC ACT 2610, Australia.

Fu, T. J., Abbott, U. R., and Hatzos, C. (2002). Digestibility of food allergens and nonallergenic proteins in simulated gastric fluid and simulated intestinal fluid - A comparative study. *Journal of Agricultural and Food Chemistry*, 50(24):7154–7160.

Gallagher, C. E., Matthews, P. D., Li, F., and Wurtzel, E. T. (2004). Gene duplication in the carotenoid biosynthetic pathway preceded evolution of the grasses. *Plant Physiology*, 135(3):1776–83.

Gannon, B., Kaliwile, C., Arscott, S. A., Schmaelzle, S., Chileshe, J., Kalungwana, N., Mosonda, M., Pixley, K., Masi, C., and Tanumihardjo, S. A. (2014). Biofortified orange maize is as efficacious as a vitamin A supplement in Zambian children even in the presence of high liver reserves of vitamin A: A community-based, randomized placebo-controlled trial. *American Journal of Clinical Nutrition*, 100(6):1541–1550.

Garcia-Alonso, M. (2013). Safety assessment of food and feed derived from GM crops: using problem formulation to ensure fit for purpose risk assessments. *Coll Biosafety Rev.*, 8:72–101.

Gemmecker, S., Schaub, P., Koschmieder, J., Brausemann, A., Drepper, F., Rodriguez-Franco, M., Ghisla, S., Warscheid, B., Einsle, O., and Beyer, P. (2015). Phytoene Desaturase from Oryza sativa: Oligomeric Assembly, Membrane Association and Preliminary 3D-Analysis. *PloS One*, 10(7):e0131717.

Goodman, R. E., Hefle, S. L., Taylor, S. L., and Van Ree, R. (2005). Assessing genetically modified crops to minimize the risk of increased food allergy: A review. *International Archives of Allergy and Immunology*, 137(2):153–166.

Goodman, R. E., Vieths, S., Sampson, H. a., Hill, D., Ebisawa, M., Taylor, S. L., and van Ree, R. (2008). Allergenicity assessment of genetically modified crops–what makes sense? *Nature Biotechnology*, 26(1):73–81.

Green, N. and Mellanby, E. (1928). Vitamin A as an anti-infective agent. *BMJ (Clinical research ed.)*, 2:691–696.

Hammond, B., Kough, J., Herouet-Guicheney, C., and Jez, J. M. (2013). Toxicological evaluation of proteins introduced into food crops. *Critical Reviews in Toxicology*, 43 Suppl 2:25–42.

Harper, B. (2011a). Phosphomannose isomerase (Entrez® accession number AAA24109): assessment of amino acid sequence similarity to known or putative toxins. Technical report, No. SSB-145-11 (unpublished), Syngenta Biotechnology, Inc., Research Triangle Park, NC,USA.

Harper, B. (2011b). Phosphomannose isomerase (Entrez® database accession number AAA24109): assessment of amino acid sequence similarity to known or putative allergens. Technical report, No. SSB-102-11 (unpublished), Syngenta Biotechnology, Inc., Research Triangle Park, NC,USA.

Harrigan, G. G., Lundry, D., Drury, S., Berman, K., Riordan, S. G., Nemeth, M. A., Ridley, W. P., and Glenn, K. C. (2010). Natural variation in crop composition and the impact of transgenesis. *Nature Biotechnology*, 28:402–404.

Harrigan, G. G., Stork, L. G., Riordan, S. G., Reynolds, T. L., Ridley, W. P., Masucci, J. D., MacIsaac, S., Halls, S. C., Orth, R., Smith, R. G., Wen, L., Brown, W. E., Welsch, M., Riley, R., McFarland, D., Pandravada, A., and Glenn, K. C. (2007). Impact of genetics and environment on nutritional and metabolite components of maize grain. *Journal of Agricultural and Food Chemistry*, 55(15):6177–6185.

Harrison, E. H. (2012). Mechanisms involved in the intestinal absorption of dietary vitamin A and provitamin A carotenoids. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, 1821(1):70–77.

Hathcock, J. N., Hattan, D. G., Jenkins, M. Y., McDonald, J. T., Sundaresan, P. R., and Wilkening, V. L. (1990). Evaluation of vitamin A toxicity. *The American Journal of Clinical Nutrition*, 52(2):183–202.

HC (2006). Guidelines for the safety assessment of novel foods. Technical report, Food Directorate, Health Canada, Ottawa, Ontario.

Henikoff, S. and Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences of the United States of America*, 89(22):10915–10919.

Hennekens, C. H., Buring, J. E., Manson, J. E., Stampfer, M., Rosner, B., Cook, N. R., Belanger, C., LaMotte, F., Gaziano, J. M., Ridker, P. M., Willett, W., and Peto, R. (1996). Lack of effect of long-term supplementation with beta carotene on the incidence of malignant neoplasms and cardiovascular disease. *The New England Journal of Medicine*, 334(18):1145–1149.

Herman, R., Song, P., and Thirumalaiswamysekhar, A. (2009). Value of eight-amino-acid matches in predicting the allergenicity status of proteins: an empirical bioinformatic investigation. *Clinical and Molecular Allergy*, 7:9.

Herman, R. A., Woolhiser, M. M., Ladics, G. S., Korjagin, V. A., Schafer, B. W., Storer, N. P., Green, S. B., and Kan, L. (2007). Stability of a set of allergens and non-allergens in simulated gastric fluid. *Int J Food Sci Nutr*, 58(2):125–141.

Heuzé, V. and Tran, G. (2015). Feedipedia, A Programme by INRA, CIRAD, AFZ, and FAO.

Hidalgo, A. and Brandolini, A. (2008). Kinetics of carotenoids degradation during the storage of einkorn (Triticum monococcum L. ssp. monococcum) and bread wheat (Triticum aestivum L. ssp. aestivum) flours. *Journal of Agricultural and Food Chemistry*, 56(23):11300–5.

Hilger, C., Grigioni, F., Thill, L., Mertens, L., and Hentges, F. (2002). Severe IgE-mediated anaphylaxis following consumption of fried frog legs: definition of alpha-parvalbumin as the allergen in cause. *Allergy*, 57(11):1053–8.

Hoffman, D. R. (1975). The specificities of human IgE antibodies combining with cereal grains. *Immunochemistry*, 12(6-7):535–8.

ICDDR (2013). National micronutrients status survey 2011-2012. Technical Report January, Centre for Nutrition and Food Security, icddr,b; UNICEF, Bangladesh; GAIN; IPHN, Dhaka, Bangladesh.

IFBC (1990). Biotechnologies and food: assuring the safety of foods produced by genetic modification. *Regulatory Toxicology and Pharmacology*, 12:S1–S196.

Ikezawa, Z., Miyakawa, K., Komatsu, H., Suga, C., Miyakawa, J., Sugiyama, A., Sasaki, T., Nakajima, H., Hirai, Y., and Suzuki, Y. (1992). A probable involvement of rice allergy in severe type of atopic dermatitis in Japan. *Acta Dermato-Venereologica. Supplementum*, 176:103–7.

ILSI (2014). Crop Composition Database, Version 5.0. International Life Sciences Institute, www.cropcomposition.org.

ISAAA (2016). GM Approval Database.

Izidoro, L. F. M., Sobrinho, J. C., Mendes, M. M., Costa, T. R., Grabner, A. N., Rodrigues, V. M., da Silva, S. L., Zanchi, F. B., Zuliani, J. P., Fernandes, C. F. C., Calderon, L. A., Stábeli, R. G., and Soares, A. M. (2014). Snake venom L-amino acid oxidases: trends in pharmacology and biochemistry. *BioMed Research International*, 2014:1–20.

Johnson, E. J. (2002). The role of carotenoids in human health. *Nutrition in Clinical Care*, 5(2):56–65.

Juliano, B. and Bechtel, D. (1985). The rice grain and its gross composition. In Juliano, B., editor, *Rice: Chemistry and Technology, 2nd ed.*, pages 17–57. American Association of Cereal Chemists, St. Paul, MN.

Keir, G., Winchester, B. G., and Clayton, P. (1999). Carbohydrate-deficient glycoprotein syndromes: inborn errors of protein glycosylation. *Annals of Clinical Biochemistry*, 36 (Pt 1):20–36.

Kenward, M. G. and Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 53(3):983–997.

Kim, S.-L., Veena, and Gelven, S. (2007). Genome-wide analysis of Agrobacterium T-DNA integration sites in the Arabidopsis genome generated under non-selective conditions. *Plant Journal*, 51:779–791.

Korgaonkar, C. K. (2009). PMI-0105 - Single-dose oral (gavage) toxicity study in mice with a 14-day observation period. Technical report, No. WIL-639011 (unpublished), Syngenta Crop Protection, Inc., Greensboro, NC 27409 USA.

Koushik, A., Hunter, D. J., Spiegelman, D., Anderson, K. E., Buring, J. E., Freudenheim, J. L., Goldbohm, R. A., Hankinson, S. E., Larsson, S. C., Leitzmann, M., Marshall, J. R., McCullough, M. L., Miller, A. B., Rodriguez, C., Rohan, T. E., Ross, J. A., Schatzkin, A., Schouten, L. J., Willett, W. C., Wolk, A., Zhang, S. M., and Smith-Warner, S. A. (2006). Intake of the major carotenoids and the risk of epithelial ovarian cancer in a pooled analysis of 10 cohort studies. *International Journal of Cancer*, 119(9):2148–2154.

Kumar, R., Srivastava, P., Kumari, D., Fakhr, H., Sridhara, S., Arora, N., Gaur, S. N., and Singh, B. P. (2007). Rice (Oryza sativa) allergy in rhinitis and asthma patients: a clinico-immunological study. *Immunobiology*, 212(2):141–7.

La Frano, M. R., Woodhouse, L. R., Burnett, D. J., and Burri, B. J. (2013). Biofortified cassava increases β-carotene and vitamin A concentrations in the TAG-rich plasma layer of American women. *The British Journal of Nutrition*, 110(2):310–20.

Lancashire, P. D., Bleiholder, H., van den Boom, T., Langeluddeke, P., Stauss, R., Weber, E., and Witzenberger, A. (1991). A uniform decimal code for growth stages of crops and weeds. *Annals of Applied Biology*, 119(3):561–601.

Lee, B. T. and Matheson, N. K. (1984). Phosphomannoisomerase and phosphoglucoisomerase in seeds of Cassia coluteoides and some other legumes that synthesize galactomannan. *Phytochemistry*, 23(5):983–987.

Leff, J. W. and Fierer, N. (2013). Bacterial Communities Associated with the Surfaces of Fresh Fruits and Vegetables. *PLoS ONE*, 8(3):1–9.

Lenth, R. (2013). lsmeans: Least-squares means. R package version 1.06-05.

Li, F., Vallabhaneni, R., Yu, J., Rocheford, T., and Wurtzel, E. T. (2008). The Maize Phytoene Synthase Gene Family: Overlapping Roles for Carotenogenesis in Endosperm, Photomorphogenesis, and Thermal Stress Tolerance. *Plant Physiology*, 147(3):1334–1346.

Linus Pauling Institute (2004). Micronutrient Information Center, Oregon State University. Available on the Internet: http://lpi.oregonstate.edu/infocenter/vitamins/vitaminE/.

MacKenzie, D. J. (2016a). Characterization of CRTI protein (Lot Number M20454-02) derived from a microbial expression system. Technical report, IR2016-02004 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

MacKenzie, D. J. (2016b). Characterization of CRTI protein (Lot Number M20603) derived from a microbial expression system. Technical report, IR2016-03002 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

MacKenzie, D. J. (2016c). Characterization of ZmPSY1 protein (Lot Number M20452-05) derived from a microbial expression system. Technical report, IR2016-02005 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines.

MacKenzie, D. J. (2016d). Estimated daily dietary exposure to ZmPSY1, CRTI, and PMI proteins expressed in rice event IR-ØØGR2E-5. Technical report, IR2016-02003 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

Malvolti, M. E., Paciucci, M., Cannata, F., and Fineschi, S. (1993). Genetic variation in Italian populations of Juglans regia L. In *Acta Horticulturae*, number 311, pages 86–94. International Society for Horticultural Science (ISHS), Leuven, Belgium.

Mathews-Roth, M. M. (1986). Beta-carotene therapy for erythropoietic protoporphyria and other photosensitivity diseases. *Biochimie*, 68(6):875–84.

Mehr, S. S., Kakakios, A. M., and Kemp, A. S. (2009). Rice: a common and severe cause of food protein-induced enterocolitis syndrome. *Archives of Disease in Childhood*, 94(3):220–3.

Metcalfe, D. D., Astwood, J. D., Townsend, R., Sampson, H., Taylor, S. L., and Fuchs, R. L. (1996). Assessment of the allergenic potential of foods derived from genetically engineered crop plants. *Critical Reviews in Food Science and Nutrition*, 36 Suppl:S165–86.

Miles, J. S. and Guest, J. R. (1984). Nucleotide sequence and transcriptional start point of the phosphomannose isomerase gene (manA) of Escherichia coli. *Gene*, 32(1-2):41–48.

Mims, G. (2009). Effect of temperature on phosphomannose isomerase as contained in test substance PMI-0105 as assessed by enzymatic activity. Technical report, No.

SSB-023-09 (unpublished), Syngenta Biotechnology, Inc., Research Triangle Park, NC,USA.

Misawa, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y., and Harashima, K. (1990). Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway by functional analysis of gene products expressed in Escherichia coli. *Journal of Bacteriology*, 172(12):6704–6712.

Mohanty, S. (2013). Trends in global rice consumption. *Rice Today*, 12(1):44–45.

Mugode, L., Ha, B., Kaunda, A., Sikombe, T., Mutale, R., Davis, C., Tanumihardjo, S., and De Moura, F. (2014). Carotenoid retention of bio fortified provitamin A maize (Zea mays L.) after Zambian traditional methods of milling, cooking and storage. *Journal of Agricultural and Food Chemistry*, 62:6317–6325.

Mukerji, P. (2016). CRTI: Acute oral toxicity study in mice. Technical report, IR2016-04001 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

Murray, M. G. and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research*, 8(19):4321–4326.

Muthayya, S., Sugimoto, J. D., Montgomery, S., and Maberly, G. F. (2014). An overview of global rice production, supply, trade, and consumption. *Annals of the New York Academy of Sciences*, 1324(1):7–14.

Muzhingi, T., Gadaga, T. H., Siwela, A. H., Grusak, M. A., Russell, R. M., and Tang, G. (2011). Yellow maize with high beta-carotene is an effective source of vitamin A in healthy Zimbabwean men. *American Journal of Clinical Nutrition*, 94(2):510–519.

NARO (2011). Food Composition Database for Safety Assessment of Genetically Modified Crops as Foods and Feeds.

Negrotto, D., Jolley, M., Beer, S., Wenck, A. R., and Hansen, G. (2000). The use of phosphomannose-isomerase as a selectable marker to recover transgenic maize plants (Zea mays L.) via Agrobacterium transformation. *Plant Cell Reports*, 19(8):798–803.

Nelson, A. (2008). In vitro digestibility of phosphomannose isomerase (PMI) as contained in test substance PMI-0105 under simulated mammalian gastric conditions. Technical report, No. SSB-034-07 A1 (unpublished), Syngenta Biotechnology, Inc., Research Triangle Park, NC,USA.

OECD (1999). Consensus Document on the Biology of Oryza sativa (Rice) ENV/JM/MONO(99)26. Technical Report 11, Organisation for Economic Cooperation and Development, Paris, France.

OECD (2004). Consensus document on compositional considerations for new varieties of rice (Oryza sativa): key food and feed nutrients and anti-nutrients. Technical Report ENV/JM/MONO(2004)15, Organisation for Economic Cooperation and Development.

OECD (2016). Revised consensus document on compositional considerations for new varieties of rice (Oryza sativa): key food and feed nutrients, anti-nutrients and other constituents. Technical report, Organisation for Economic Cooperation and Development, Paris, France.

OGTR (2005). The Biology and Ecology of Rice (Oryza sativa L .) in Australia. Technical Report February, Office of the Gene Technology Regulator, Australia, Canberra.

Oliva, N. (2016). Characterization of the in vitro pepsin digestibility of phytoene synthase protein (ZmPSY1). Technical report, IR2016-01002 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines.

Oliva, N. and Cueto, M. (2016). Characterization of the in vitro pepsin digestibility of phytoene desaturase protein (CRTI) using western blot analysis. Technical report, IR2016-07003 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

Oliva, N., Cueto, M., and Salcedo, M. (2016a). Concentrations of phytoene synthase (ZmPSY1), phytoene desaturase (CRTI), and phosphomannose isomerase (PMI) in grain and straw harvested from event IR-ØØGR2E-5 rice grown in the Philippines in 2015. Technical report, IR2015-08004 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines.

Oliva, N., Cueto, M., and Salcedo, M. (2016b). Western immunoblot analysis of ZmPSY1, CRTI, and PMI expression in different plant tissues from rice event IR-ØØGR2E-5. Technical report, IR2016-04003 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

Oliva, N., Cueto, M., and Salceo, M. (2016c). Concentrations of phytoene synthase (ZmPSY1), phytoene desaturase (CRTI), and phosphomannose isomerase (PMI) in grain and straw harvested from event IR-ØØGR2E-5 rice grown in the Philippines in 2016. Technical report, IR2016-04004 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines.

Oliva, N. and MacKenzie, D. J. (2016a). Amino acid sequence similarity search between Pantoea ananatis phytoene desaturase (CRTI) and known and putative protein allergens. Technical report, IR2016-02001 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

Oliva, N. and MacKenzie, D. J. (2016b). Amino acid sequence similarity search between Pantoea ananatis phytoene desaturase (CRTI) and known and putative protein toxins. Technical report, IR2016-01004 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines.

Oliva, N. and MacKenzie, D. J. (2016c). Amino acid sequence similarity search between Zea mays phytoene synthase (ZmPSY1) and known and putative protein allergens. Technical report, IR2016-02002 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines.

Oliva, N. and MacKenzie, D. J. (2016d). Amino acid sequence similarity search between Zea mays phytoene synthase (ZmPSY1) and known and putative protein toxins. Technical report, IR2016-01005 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines.

Oliveira, A. and Carvalho, L. (2010). Assessment and degradation study of total carotenoids and b-carotene in bitter yellow cassava (Manihot esculenta Crantz) varieties. *Afr J Food Chemistry*, 4(4):148–155.

Omenn, G. S., Goodman, G. E., Thornquist, M. D., Balmes, J., Cullen, M. R., Glass, A., Keogh, J. P., Meyskens, F. L., Valanis, B., Williams, J. H., Barnhart, S., and Hammar, S. (1996). Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. *N.Engl.J.Med.*, 334(18):1150–1155.

Paine, J., Shipton, C., Chaggar, S., Howells, R. M., Kennedy, M. J., Vernon, G., Wright, S. Y., Hinchliffe, E., Adams, J. L., Silverstone, A. L., and Drake, R. (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nature Biotechnology*, 23(4):482–487.

Pastorello, E. a., Farioli, L., Pravettoni, V., Ispano, M., Scibola, E., Trambaioli, C., Giuffrida, M. G., Ansaloni, R., Godovac-Zimmermann, J., Conti, A., Fortunato, D., and Ortolani, C. (2000). The maize major allergen, which is responsible for food-induced allergic reactions, is a lipid transfer protein. *The Journal of Allergy and Clinical Immunology*, 106(4):744–751.

Pearson, W. (2000). Flexible sequence similarity searching with the FASTA3 program package. *Methods in Molecular Biology*, 132:185–219.

Pearson, W. and Lipman, D. (1988). Improved tools for biological sequence comparison. *Proceedings of the National Academy of Sciences*, 85:2444–2448.

Pearson, W. R. (1996). *Computer Methods for Macromolecular Sequence Analysis*, volume 266 of *Methods in Enzymology*. Elsevier.

Pearson, W. R. (2013). Selecting the Right Similarity-Scoring Matrix. *Current Protocols in Bioinformatics*, 43:1–12.

Pénicaud, C., Achir, N., Dhuique-Mayer, C., Dornier, M., and Bohuon, P. (2011). Degradation of β-carotene during fruit and vegetable processing or storage: Reaction mechanisms and kinetic aspects: A review. *Fruits*, 66(6):417–440.

Pereira, A. and Aarts, M. G. (1998). Transposon tagging with the En-I system. *Methods in Molecular Biology*, 82:329–338.

Proudfoot, A. E., Turcatti, G., Wells, T. N., Payton, M. A., and Smith, D. J. (1994a). Purification, cDNA cloning and heterologous expression of human phosphomannose isomerase. *Eur J Biochem*, 219(1-2Glaxo Institute for Molecular Biology, Geneva, Switzerland. http://www.biomednet.com/db/medline/94139717):415–423.

Proudfoot, A. E. I., Payton, M. A., and Wells, T. N. C. (1994b). Purification and characterization of fungal and mammalian phosphomannose isomerases. *Journal of Protein Chemistry*, 13(7):619–627.

R Core Team (2015). R: A language and environment for statistical computing. Technical report, R Foundation for Statistical Computing, Vienna, Austria.

Rabe, S. (2004). Phosphomannose isomerase: assessment of amino acid sequence homology with known allergens. Technical report, No. SSB-116-04 (unpublished), Syngenta Biotechnology, Inc., Research Triangle Park, NC,USA.

Ranum, P., Pena-Rosas, J. P., and Garcia-Casal, M. N. (2014). Global maize production, utilization, and consumption. *Annals of the New York Academy of Sciences*, 1312(1):105–112.

Rao, S. and Rossmann, M. G. (1973). Comparison of super-secondary structures in proteins. *Journal of Molecular Biology*, 76(2):241–256.

Reed, J., Privalle, L., Powell, M. L., Meghji, M., Dawson, J., Dunder, E., Suttie, J., Wenck, A., Launis, K., Kramer, C., Chang, Y.-F. F., Hansen, G., Wright, M., and Sutthe, J. (2001). Phosphomannose isomerase: An efficient selectable marker for plant transformation. *In Vitro Cellular and Development Biology - Plant*, 37(2):127–132.

Ricroch, A. E. (2013). Assessment of GE food safety using '-omics' techniques and long-term animal feeding studies. *New Biotechnology*, 30(4):349–354.

Samia, M. and Swamy, M. (2016). Concentrations of beta-carotene and other carotenoids in grain samples from rice event IR-ØØGR2E-5. Technical report, IR2016-07004 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

Sanchez-Monge, R., Lopez-Torrejón, G., Pascual, C. Y., Varela, J., Martin-Esteban, M., and Salcedo, G. (2004). Vicilin and convicilin are potential major allergens from pea. *Clinical and Experimental Allergy : Journal of the British Society for Allergy and Clinical Immunology*, 34(11):1747–53.

Sandmann, G. (2009). Evolution of carotene desaturation: the complication of a simple pathway. *Archives of Biochemistry and Biophysics*, 483:169–174.

Schaub, P. and Beyer, P. (2016). Characterization of the heat stability of phytoene desaturase protein (CRTI). Technical report, IR2015-12001 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

Schaub, P., Yu, Q., Gemmecker, S., Poussin-Courmontagne, P., Mailliot, J., McEwen, A. G., Ghisla, S., Al-Babili, S., Cavarelli, J., and Beyer, P. (2012). On the structure and function of the phytoene desaturase CRTI from Pantoea ananatis, a membrane-peripheral and FAD-dependent oxidase/isomerase. *PLoS ONE*, 7(6).

Scrimshaw, N. S. and SanGiovanni, J. P. (1997). Synergism of nutrition, infection, and immunity: an overview. *The American Journal of Clinical Nutrition*, 66(2):464S–477S.

Sell, M., Steinhart, H., and Paschke, A. (2005). Influence of maturation on the alteration of allergenicity of green pea (Pisum sativum L.). *Journal of Agricultural and Food Chemistry*, 53(5):1717–1722.

Shure, M., Wessler, S., and Fedoroff, N. (1983). Molecular identification and isolation of the Waxy locus in maize. *Cell*, 35(1):225–233.

Silvanovich, A., Nemeth, M. A., Song, P., Herman, R., Tagliani, L., and Bannon, G. A. (2006). The value of short amino acid sequence matches for prediction of protein allergenicity. *Toxicological Sciences*, 90(1):252–258.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, 98:503–517.

Stein, A. J., Sachdev, H. P. S., and Qaim, M. (2006). Potential impacts of golden rice on public health in India. In *Contributed paper presented at the 26th Conference of the International Association of Agricultural Economists (IAAE), Gold Coast, Australia, August 12-18, 2006*, pages 1–15.

Stevens, G. A., Bennett, J. E., Hennocq, Q., Lu, Y., De-Regil, L. M., Rogers, L., Danaei, G., Li, G., White, R. A., Flaxman, S. R., Oehrle, S. P., Finucane, M. M., Guerrero, R., Bhutta, Z. A., Then-Paulino, A., Fawzi, W., Black, R. E., and Ezzati, M. (2015). Trends and mortality effects of vitamin A deficiency in children in 138 low-income and middle-income countries between 1991 and 2013: A pooled analysis of population-based surveys. *The Lancet Global Health*, 3(9):528–536.

Suarez, J. V., Banks, S., Thomas, P. G., and Day, A. (2014). A new F131V mutation in Chlamydomonas phytoene desaturase locates a cluster of norflurazon resistance mutations near the FAD-binding site in 3D protein models. *PloS ONE*, 9(6):e99894.

Suvarna, B. S. (2008). Rice , the allergen. *Nepal Journal of Science and Technology*, 9:195–199.

Swamy, M. and Samia, M. (2015). Segregation of the inserted DNA within multiple generations of rice event IR-ØØGR2E-5. Technical report, IR2015-08003 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

Swamy, M. and Samia, M. (2016). Stability of the elevated beta-carotene trait across multiple generations of rice event IR-ØØGR2E-5. Technical report, IR-2016-07001 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

Swamy, M., Samia, M., Boncodin, R., Rebong, D., Ordonio, R., and MacKenzie, D. J. (2016a). Nutrient composition of event IR-ØØGR2E-5 and non-transgenic control rice grown during the rainy season in 2015 in the Philippines. Technical report, IR2015-07001 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

Swamy, M., Samia, M., Boncodin, R., Rebong, D., Ordonio, R., Miranda, R., Rebong, A., Alibuyog, S., Marundan, S., Adeva, C., and MacKenzie, D. J. (2016b). Nutrient composition of event IR-ØØGR2E-5 and non-transgenic control rice grown during the dry season in 2016 in the Philippines. Technical report, IR2016-05001 (unpublished) International Rice Research Insstitute, Los Banos, Laguna, Philippines.

Takaiwa, F., Ebinuma, H., Kikuchi, S., and Oono, K. (1987). Nucleotide sequence of a rice glutelin gene. *FEBS Letters*, 221(1):43–47.

Tang, G. (2010). Bioconversion of dietary provitamin A carotenoids to vitamin A in humans. *American Journal of Clinical Nutrition*, 91(5):1468–1473.

Tang, G., Hu, Y., Yin, S. A., Wang, Y., Dallal, G. E., Grusak, M. A., and Russell, R. M. (2012). β-carotene in Golden Rice is as good as β-carotene in oil at providing vitamin A to children. *American Journal of Clinical Nutrition*, 96(3):658–664.

Tang, G., Jian, Q., Dolnikowski, G. G., Russell, R. M., and Grusak, M. A. (2009). Golden rice is an effective source of vitamin A. *American Journal of Clinical Nutrition*, 89(6):1776–1783.

The UniProt Consortium (2014). UniProt: a hub for protein information. *Nucleic Acids Research*, 43(D1):D204–212.

Thomas, K., Aalbers, M., Bannon, G., Bartels, M., Dearman, R. J., Esdaile, D. J., Fu, T. J., Glatt, C. M., Hadfield, N., Hatzos, C., Hefle, S. L., Heylings, J. R., Goodman, R. E., Henry, B., Herouet, C., Holsapple, M., Ladics, G. S., Landry, T. D., MacIntosh, S. C., Rice, E. A., Privalle, L. S., Steiner, H. Y., Teshima, R., Van Ree, R., Woolhiser, M., and Zawodny, J. (2004). A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. *Regulatory Toxicology and Pharmacology*, 39(2):87–98.

Thomas, K., Bannon, G., Hefle, S., Herouet, C., Holsapple, M., Ladics, G., MacIntosh, S., and Privalle, L. (2005). In silico methods for evaluating human allergenicity to novel proteins: International bioinformatics workshop meeting report, 23-24 February 2005. *Toxicological Sciences*, 88(2):307–310.

Todd, R. and Tague, B. W. (2001). Phosphomannose isomerase: A versatile selectable marker forArabidopsis thaliana germ-line transformation. *Plant Molecular Biology Reporter*, 19(4):307–319.

Trijatmiko, K., Arines, F., Oliva, N., Slamet-Loedin, I., and Kohli, A. (2016). Molecular analyses of transgenic plants. *Methods Mol Bio*, 1385:201–222.

Trijatmiko, K., Oliva, N., Alexandrov, N., Slamet-Loedin, I., and Chadha-Mohanty, P. (2015). Nucleotide sequence analysis of the inserted DNA and host genomic flanking regions in rice event IR-ØØGR2E-5. Technical report, IR2015-08001 (unpublished) International Rice Research Institute, Los Banos, Laguna, Philippines.

Tripathi, K., Warrier, R., Govila, O., and Ahuja, V. (2011). Biology of Oryza sativa L. (Rice). Technical report, Ministry of Environment and Forests, Government of India.

USDA-FAS (2015). Australia Grain and Feed Annual. Technical report, AS1506, USDA Foreign Agricultural Service, Canberra ACT, Australia.

USDA-FAS (2016). USDA PSD Online (https://apps.fas.usda.gov/psdonline/psdhome.aspx).

Van Eenennaam, L. and Young, E. (2014). Prevalence and impacts of genetically engineered feedstuffs on livestock populations. *Journal of Animal Science*, 92(10):4255–4278.

Wailes, E. J. and Chavez, E. C. (2015). International Rice Outlook: International Rice Baseline Projections 2014-2024. Technical report, University of Arkansas, Fayetteville.

Walpole, S., Prieto-Merino, D., Edwards, P., Cleland, J., Stevens, G., and Roberts, I. (2012). The weight of nations: an estimation of adult human biomass. *BMC Public Health*, 12(1):439.

Welsch, R., Arango, J., Bär, C., Salazar, B., Al-Babili, S., Beltrán, J., Chavarriaga, P., Ceballos, H., Tohme, J., and Beyer, P. (2010). Provitamin A accumulation in cassava (Manihot esculenta) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. *The Plant Cell*, 22(10):3348–3356.

Welsch, R. and Beyer, P. (2016). Characterization of the heat stability of phytoene synthase protein (ZmPSY1). Technical report, IR2015-12002 (unpublished), International Rice Research Institute, Los Banos, Laguna, Philippines.

WHO (2009). Global prevalence of vitamin A deficiency in populations at risk 1995-2005 : WHO global database on vitamin A deficiency. Technical report, World Health Organization, Geneva, Switzerland.

Yadav, N. S., Vanderleyden, J., Bennett, D. R., Barnes, W. M., and Chilton, M. D. (1982). Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proceedings of the National Academy of Sciences of the United States of America*, 79(20):6322–6326.

Ye, X., Al-Babili, S., Klöti, A., Zhang, J., Lucca, P., Beyer, P., and Potrykus, I. (2000). Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science (New York, N.Y.)*, 287(5451):303–305.

Zambryski, P., Depicker, A., Kruger, K., and Goodman, H. M. (1982). Tumor induction by Agrobacterium tumefaciens: analysis of the boundaries of T-DNA. *Journal of Molecular and Applied Genetics*, 1(4):361–70.

Zimmermann, R. and Ahmed, F. (2006). Rice biotechnology and its potential to combat vitamin A deficiency: a case study of golden rice in Bangladesh. Technical Report 1, No. 104, Zentrum für Entwicklungsforschung (ZEF), Bonn, Germany.

Zimmermann, R. and Qaim, M. (2004). Potential health benefits of Golden Rice: A Philippine case study. *Food Policy*, 29(2):147–168.

Appendices

A. Materials and Methods for Genetic Characterization of GR2E Rice

A.1. Southern Blot Characterization of GR2E Rice

Southern blot analyses were performed to investigate the number of sites of insertion of the pSYN12424 T-DNA in event GR2E rice, the integrity of introduced genetic elements, the absence of plasmid backbone sequences, and the multi-generational stability of the inserted DNA.

To determine the number of copies of the inserted DNA, samples of genomic DNA prepared from event GR2E in four different germplasm backgrounds were subject to restriction endonuclease digestion with either *Hin*dIII or *Sph*I enzymes followed by electrophoretic separation and transfer onto nylon membrane. Hybridizations were performed with DNA probes specific to the *Zmpsy1*, *pSSU-crtI*, and *pmi* genes.

The integrity of the inserted T-DNA was investigated by Southern hybridization of *Asc*I+*Xma*I-digested genomic DNA samples using the *Zmpsy1*, *pSSU-crtI*, and *pmi* probes.

The lack of integration of plasmid backbone sequences from pSYN12424 was confirmed using a mixture of five overlapping hybridization probes, which together corresponded to the entire plasmid backbone region.

The stability of the inserted DNA across multiple generations of GR2E rice was assessed by Southern blot analyses of genomic DNA samples prepared from the T_n generation (Kaybonnet) and the BC_3F_5, BC_4F_3, and BC_5F_3 generations of GR2E in BRRI *dhan* 29, IR64, and PSB Rc82 germplasm.

A.2. Genomic DNA Extraction

DNA isolation was performed as described by Murray and Thompson (1980). Pooled leaf tissue samples from three plants (*ca*. 2 g) were cooled in liquid nitrogen and ground to a fine powder using a mortar and pestle. Ground tissue was incubated with 20 ml pre-warmed 1.5 × cetyltrimethyl ammonium bromide (CTAB) buffer [75 mM Tris-HCl, pH 8.0, 1.05 M NaCl, 15 mM EDTA, 1.5% CTAB] for 20 min at 56°C followed by addition of 20 ml chloroform:isoamyl alcohol (24:1) and incubation at room temperature with moderate shaking for 20 min. Samples were centrifuged (3000 rpm × 20 min) and the supernatant was filtered using miracloth and transferred to a new tube and mixed (55 rpm, room temperature) with 2 ml of pre-warmed 10% CTAB in 0.7 M NaCl and 20 ml chloroform:isoamyl alcohol (24:1). Samples were centrifuged (3000 rpm × 30 min) and the uppermost layer was collected and transferred to a new tube. DNA was precipitated overnight with 30 ml of 1.0% CTAB precipitation buffer [50 mM Tris-HCl, pH 8.0, 10 mM ethylenediamine tetraacetic acid (EDTA), 1% CTAB] followed by centrifugation (3000 rpm × 20 min) at room temperature and the supernatant was discarded. The pellet was

dried for 10 min and was dissolved with 3 ml 1 M NaCl. RNAse was added to a final concentration of 13 μg/ml followed by incubation in a 56°C water bath for 1 hr. Two volumes of absolute ethanol were added and precipitated DNA was hooked using disposable loop and dip-washed in 0.5 ml 70% ethanol twice for 7 min and then in absolute ethanol for 5 min. Precipitated DNA was transferred to a new microcentrifuge tube, air dried, and resuspended with 600 μl 1/10 TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA).

A.3. Quantitation of Genomic DNA

Following extraction, the DNA samples were quantified on a spectrofluorometer using PicoGreen®reagent (Quant-iT™ PicoGreen dsDNA reagent and kits, Invitrogen) following a standard procedure. The DNA was also visualized on an agarose gel to confirm quantitation values from the PicoGreen analysis and to determine DNA quality.

A.4. Identification of GR2E Plants used for Southern Analysis

Leaf tissue samples from GR2E and control rice plants were analyzed by multiplex PCR zygosity testing to confirm the homozygous presence of the pSYN12424 T-DNA insert in samples derived from GR2E rice and absence of the insert in samples from control Kaybonnet rice. .

A.5. Digestion of DNA for Southern Blot Analyses

Genomic DNA samples extracted from selected GR2E and control rice plants were digested with restriction enzymes following a standard procedure. Approximately 7 μg of genomic DNA was digested using 30 units of restriction enzyme in a total volume of 150 μl, according to manufacturer's recommendations. The digestions were carried out at 37°C for 12–16 hours, followed by ethanol precipitation with 1/10 volume of 3 M sodium acetate (pH 5.2) and 2 volumes of 100% ethanol. After incubation at –20°C (0.5–2 hours) and centrifugation, the DNA was allowed to dry and then re-dissolved in 15 μl TE buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA) plus 3 μl of 6× Orange Ficoll loading dye (0.3% Orange G, 15% Ficoll 400, 2% 50× TAE buffer, 1% SDS). The reference plasmid, pSYN12424, was spiked into a control plant DNA sample in an amount equivalent to approximately one or 0.2 gene copies per rice genome and digested with *Sph*I enzyme to serve as a positive control for probe hybridization and to verify sizes of fragments internal to the plasmid on the Southern blot.

A.6. Electrophoretic Separation and Southern Transfer

Following restriction enzyme digestion, the resultant DNA fragments were electrophoretically separated by size through an agarose gel. DIG-labelled DNA Molecular Weight Marker VII (Roche), which is visible after DIG detection as described below, was used to determine the hybridizing fragment size on the Southern blots.

Agarose gels containing the separated DNA fragments were depurinated, denatured, and neutralized *in situ*, and transferred overnight via upward capillary transfer onto a nylon membrane in 20 × saline sodium citrate (SSC) buffer (3 M NaCl, 0.3 M sodium citrate, pH 7.0) using the method as described by Trijatmiko et al. (2016). The blot transfer sandwich

contained a bridge resting in a reservoir of 20 × SSC, soaked 3MM Whatman paper, the agarose gel and a nylon membrane same size as of the gel then covered with 3MM paper pre-wetted with 20 × SSC. To complete the assembly, a stack of paper towels was added followed by a glass plate and finally a 200–500 g weight. After the overnight transfer, the DNA bound to the membrane was immobilized by UV cross-linking (Stratalinker, Stratagene, La Jolla, CA).

A.7. DNA Probe Labelling for Southern Blot Hybridization

Probes for the *Zmpsy1*, *pSSU-crtI*, and *pmi* genes were used to detect genetic elements within the insertion. Probes covering the backbone region of plasmid pSYN12424 were used to verify absence of plasmid backbone DNA in GR2E rice (Figure 30). DNA fragments of the probe elements were generated by PCR from plasmid pSYN12424 using specific primers (Table 31). PCR fragments were electrophoretically separated on an agarose gel, excised and purified using a gel purification kit (Qiagen, Valencia, CA). DNA probes were generated from these fragments by PCR that incorporated a DIG-labelled nucleotide, [DIG-11]-dUTP, into the fragment. PCR labelling of isolated fragments was carried out according to the procedures supplied in the PCR DIG Probe Synthesis Kit (Roche).

Table 31. Location and sizes of DNA probes used for Southern hybridization

Probe	Primers (5' to 3')	Position on pSYN12424 (bp to bp)	Length (bp)
Zmpsy1	ATGGCCATCATACTCGTACGA CTAGGTCTGGCCATTTCTCAA	3893–3913 5105–5125	1233
pSSU-crtI	ATGGCTTCTATGATATCCTCTTCC TCAAATCAGATCCTCCAGCA	1087–1110 2717–2736	1650
pmi	ATGCAAAAACTCATTAACTCAGTGC TTACAGCTTGTTGTAAACACGC	7429–7453 8583–8604	1176
Backbone 1	CCACCAGCCAGCCAACAG CTTGTTCACAGCGCTTGAGA	9028–9045 10418–10437	1410
Backbone 2	GCCGCAATTCTGACGAACTG CGGTTCATGGATTCGGTTAG	10355–10374 11664–11683	1329
Backbone 3	GTGATCCGCTACGAGCTTCC CTCTCATCAACCGTGGCTCC	11559–11578 12821–12840	1282
Backbone 4	CTACGGCCAGGCAATCTACC TGGAGCGAACGACCTACACC	12681–12700 13751–13770	1090
Backbone 5	ATCTCAGTTCGGTGTAGGTC CTAAGAGAAAAGAGCGTTTATTAG	13741–13760 14316–14339	599

A.8. Probe Hybridization and Visualization

Labelled probes (Table 31) were hybridized to the target DNA on the nylon membranes at 42°C overnight and washed two times with 2 × SSC and 0.1% SDS wash buffer at room temperature and then two times at 68°C with 0.1 × SSC and 0.1% SDS wash buffer. The membrane was then equilibrated with maleic acid buffer (0.1 M maleic acid, 0.15 M Nacl, pH7.5) containing 0.3% Tween 20 and subjected to blocking solution (1 × maleic acid buffer, 1% blocking reagent, Roche, IN, USA) for 2 hrs. then onto antibody solution (1 × blocking solution containing 75 mU/ml anti-DIG AP, Roche, IN, USA) for 30 minutes at room temperature. Excess antibody solution was washed two times by incubating with 1

× maleic acid buffer with 0.3% Tween 20 and then equilibrated with 0.1 M Tris-HCl pH 9.5, 0.1 M NaCl. Blots were exposed to Kodak Biomax MS film (Eastman Kodak, Rochester, NY) for one or more time points to detect hybridizing fragments and to visualize molecular weight standards bound to the nylon membrane. Films were dried and photographs were taken using a white light transilluminator in a gel imaging system (Vilber Lourmat, France). The sizes of detected bands were documented for each digest and each probe.

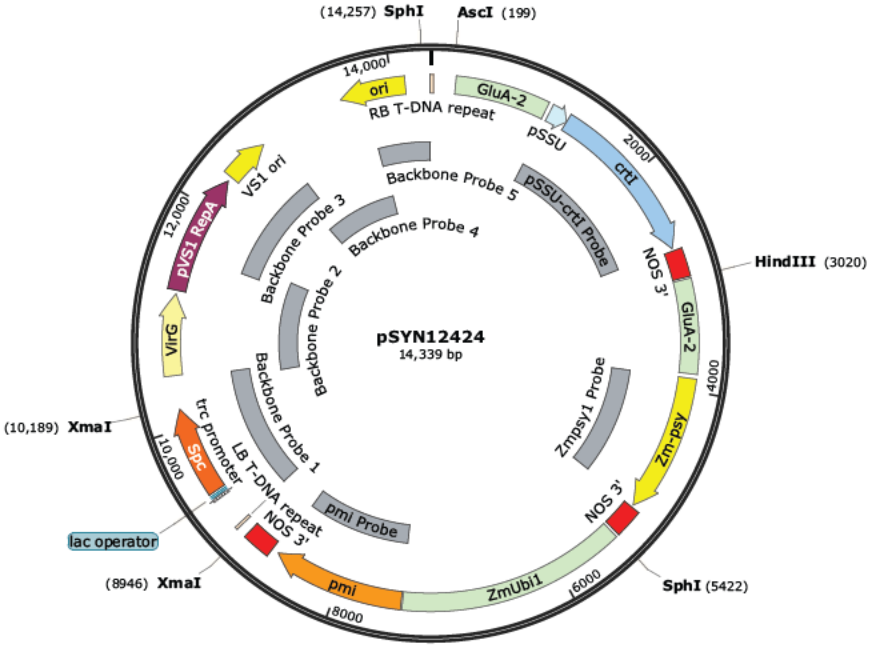


Figure 30. Schematic map of plasmid pSYN12424 showing the location of the DNA probes used for Southern hybridization characterization of GR2E rice.

A.9. Methods for Segregation Analysis

A.9.1. DNA Isolation

DNA isolation was performed essentially following the CTAB extraction method (Doyle and Doyle, 1987). Leaf tissue samples (2–3 inches in length) in 2-ml micro-centrifuge tubes were cooled in liquid nitrogen and pulverized to a fine powder using steel balls in a Geno Grinder (60 sec × 900 rpm, followed by cooling with liquid N_2 and re-grinding, if necessary). Ground tissue was incubated with 400 μl CTAB buffer (0.1 M Tris-HCl pH 8.0, 1.4 M NaCl, 2 mM EDTA, 2% CTAB) for 30 min at 65°C followed by extraction with 450 μl chloroform:isoamyl alcohol (24:1) for 20–30 min at room temperature. Samples were centrifuged (11000 rpm × 10 min) and the upper aqueous phase (200 μl) was mixed with 300 μl isopropanol and DNA was precipitated at −20°C (1 hour). Following centrifugation, pelleted DNA was washed twice by centrifugation with 70 percent ethanol and resuspended with 100 μl 10/1 TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) containing 10 μg/ml RNase. Isolated DNA was quantified spectrophotometrically by measuring absorbance at 260 nm and checked for quality by alkaline agarose gel electrophoresis.

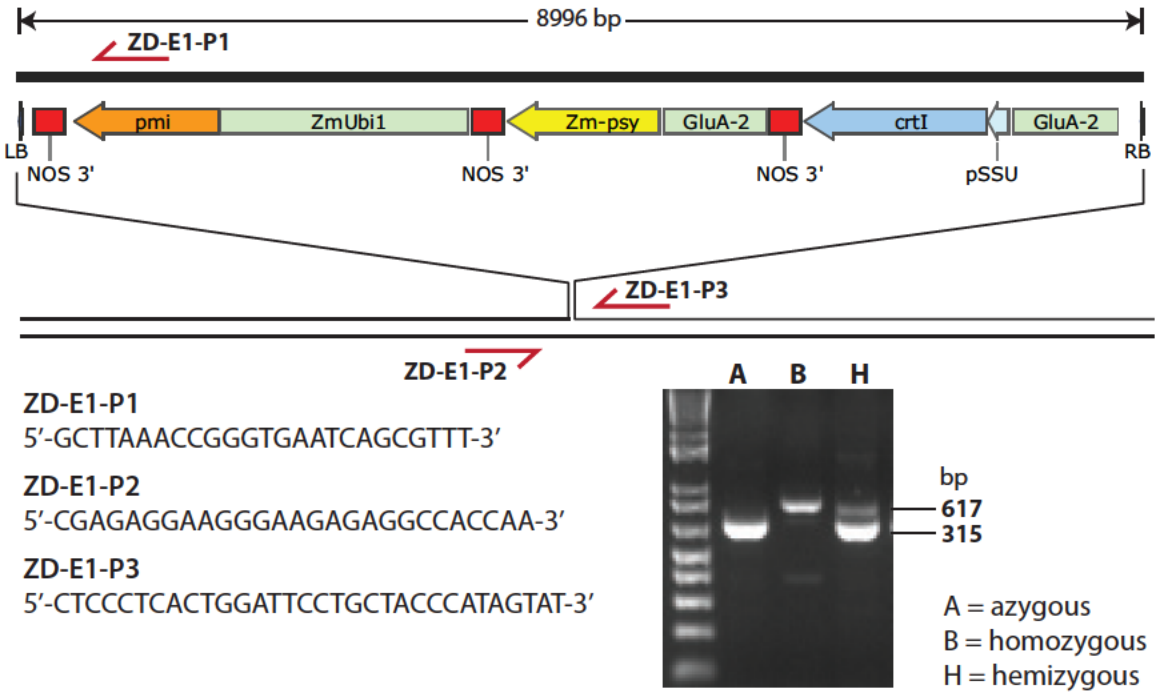


Figure 31. Genomic DNA isolated from individual plants was analyzed by multiplex PCR amplification employing three primers. Two primers, ZD-E1-P2 and ZD-E1-P3, were complementary to sequences in the left border (LB) and right border (RB) flanking host genomic regions, respectively, and primer ZD-E1-P1 was complementary to *pmi* gene sequences within the inserted T-DNA. Null segregants lacking the pSYN12424 T-DNA resulted in amplification of a single 315 bp fragment. Homozygous GR2E plants resulted in amplification of a single 617 bp fragment, and hemizygous plants resulted in the amplification of two fragments (315 bp and 617 bp)

A.9.2. Genotypic Analysis

DNA samples were analyzed by multiplex PCR amplification to determine zygosity of the pSYN12424 T-DNA insert as illustrated in Figure 31. Individual PCR reaction mixtures contained 300 ng genomic DNA, 0.5 mM of each primer (ZD-E1-P1, ZD-E1-P2, and Zd-E1-P3), 0.2 mM deoxynucleotide triphosphates, 1.5 mM $MgCl_2$, and 0.5 units KAPA Taq DNA polymerase (KAPA Biosystems) in 10 μl KAPA Taq buffer. After an initial 5 min denaturation at 95°C, PCR amplifications were performed for 35 cycles consisting of denaturation (95°C, 45 sec), annealing (55°C, 45 sec), and extension (72°C, 45 sec). A final extension step (72°C, 8 min) was included following the last PCR cycle, after which samples were held at 25°C.

Following PCR, reaction mixtures were analyzed by alkaline agarose gel electrophoresis in 1.5 percent agarose and amplified fragments were visualized with SYBR® Safe DNA Gel Stain (ThermoFisher Scientific).

A.9.3. Statistical Analysis

A Chi-square analysis was performed on the segregation results of each GR2E rice generation to compare the observed segregation ratio to the expected segregation ratio. This analysis tested the hypothesis that the introduced DNA was segregating in a Mendelian fashion. The critical value to reject the hypothesis at the 5 percent level is 3.84.

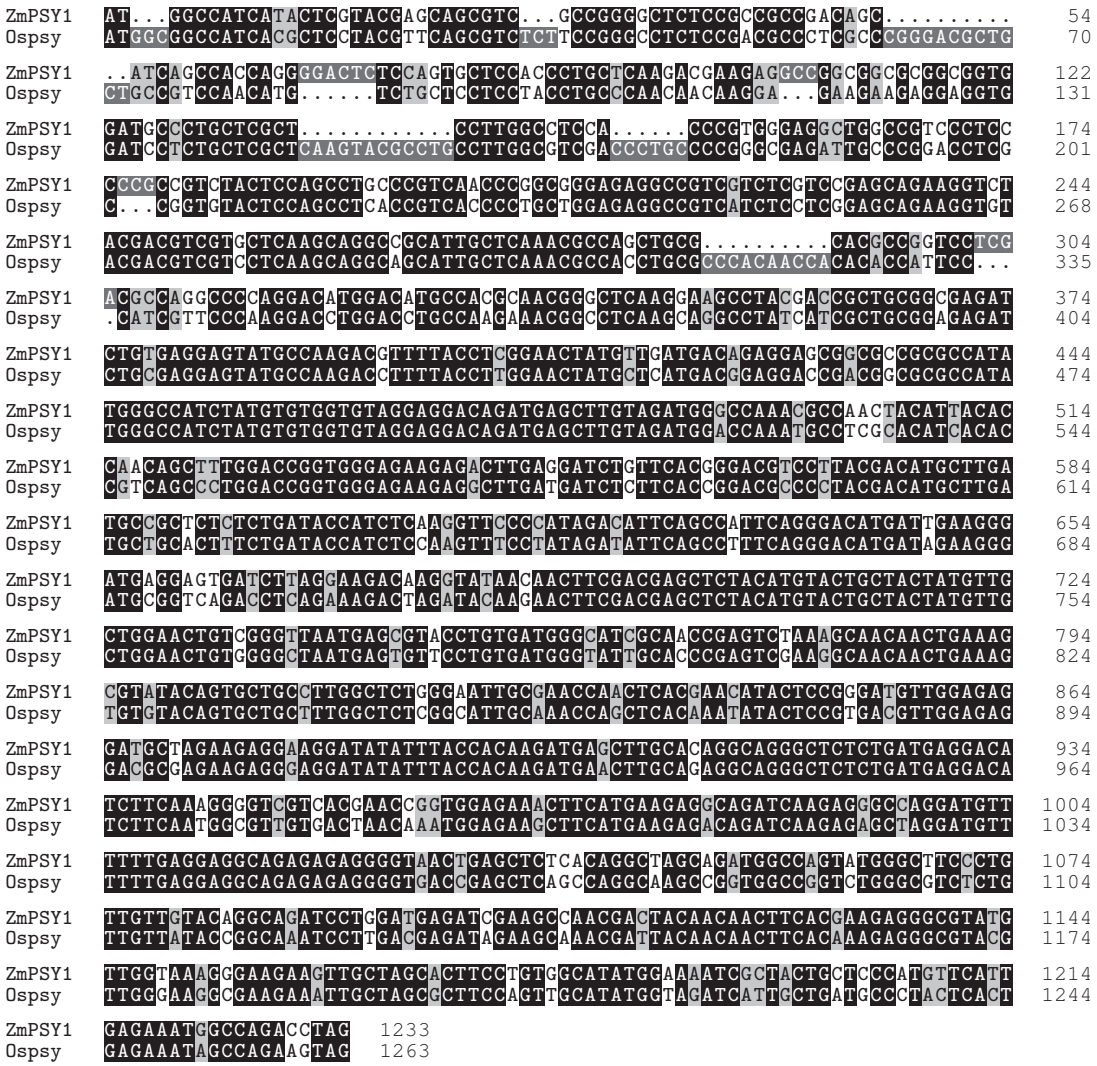


Figure 32. Alignment of the nucleotide sequences of the *Zea mays psy1* (*Zmpsy1*; GI:1098664) and *Oryza sativa psy* (*Ospsy*; GI:47678186) genes. These two genes share 86.9 percent similar residues and 82.8 percent identical residues. This figure was generated with TEXshade (Beitz, 2000).

B. Materials and Methods for Nucleotide Sequence Analysis of GR2E Rice

The objective of this study was to determine the nucleotide sequence of the inserted DNA within rice event GR2E (IR-ØØGR2E-5), including a portion of the 5' and 3' flanking host genomic sequences, in order to demonstrate overall integrity of the insert, contiguity of the functional elements, and to detect any individual base-pair changes. In addition, an open reading frame (ORF) analysis was conducted to investigate the possibility of creating any new start-to-stop ORFs spanning the 5' or 3' junctional regions that could potentially encode sequences homologous to known allergens or toxins.

B.1. DNA Extraction

Genomic DNA was extracted from leaf tissue obtained from a single GR2E rice progeny plant derived from the original transformation event in Kaybonnet germplasm using a modification of the urea-phenol method Pereira and Aarts (1998); Shure et al. (1983). Approximately 200 mg of leaf tissue was frozen in liquid nitrogen in a microcentrifuge tube and ground to a fine powder using a pestle. The frozen powder was resuspended in 250 μl of DNA extraction buffer (0.3 M NaCl, 50 mM Tris-HCl pH 7.5, 20 mM EDTA, 2% sarkosyl, 0.5% SDS, 5 M urea, 5% phenol) and ground once more. An additional 250 μl of extraction buffer was added and the mixture was extracted with an equal volume (500 μl) of phenol:chloroform (1:1). Following centrifugation (13000 rpm 10 min), the upper aqueous phase was removed and DNA was precipitated by the addition of 0.7 vol of isopropanol and incubation for 5 min at room temperature followed by centrifugation (13000 rpm 5 min). The DNA pellet was washed once with 70% ethanol, air-dried, and resuspended with 100 μl 10 mM Tris-HCl pH 8.0, 1 mM EDTA containing 10 μg/ml RNAse.

B.2. PCR Amplifications

The nucleotide sequence of the plasmid pSYN12424 T-DNA, together with preliminary sequence information from the 5' and 3' flanking genomic DNA (Syngenta), was used to design seven sets of oligonucleotide primers that were used to amplify the insert and flanking regions from GR2E rice genomic DNA as seven individual overlapping fragments (Table 32).

Polymerase chain reaction (PCR) amplifications were carried out using the KAPA HiFi HotStart PCR Kit (Kapa Biosystems, Wilmington, MA, USA; Catalogue No. KK2502) and a G-Storm GS1 thermocycler using the parameters in Table 33.

B.3. Nucleotide Sequencing

The PCR amplification fragments were cloned into pCR®4-TOPO® vector via blunt-end ligation and transformed into *E. coli* strain DH5α. Three colonies from each transformation were randomly selected and separately transferred into liquid culture and grown overnight at 37°C with shaking. Plasmid DNA prepared from each culture was submitted to Macrogen Incorporated (Seoul, Korea) for dye-terminator nucleotide sequencing using the ABI3730XL

Table 32. Primers used for nucleotide sequencing of the inserted DNA within GR2E rice

PCR Fragment	Primer	Primer Sequences (5′ – 3′)	Target Sequence Location (bp to bp)†	Amplified Fragment Size (bp)
A	GR2E F1	TCTGGAGGATCTACTGCTGC	1–20	2041
	GR2E R1	CGCTCATGATCAGATTGTCG	2022–2041	
B	GR2E-F2	CCCAGTGAGTAGTATTTTCAGTCGT	1642–1666	3137
	GR2E-R2	AAGACCGGCAACAGGATTC	4760–4778	
C	GR2E F3	TAGAGGACGGTCGCAGGTTC	3998–4017	1967
	GR2E R3	GGCCTCTTCGTCTTGAGCAG	5945–5964	
D	GR2E F4	TCCACCTTTCGTGTACCACA	5662–5681	1506
	GR2E R4	AAGACCGGCAACAGGATTC	7149–7168	
E	GR2E-F5	GGATGAGATCGAAGCCAACG	6952–6971	2018
	GR2E R5	CCACATCATCACAACCAAGC	8950–8969	
F	GR2E F6	GGATCTGTATGTGTGTGCCA	8820–8839	1868
	GR2E R6	AAGACCGGCAACAGGATTC	10668–10686	
G	GR2E F7	ACCACCATTAGCCAGCAGAG	10390–10409	2383
	GR2E R7	AGGAAGAATGCATCTGGAGC	12753–12772	

† Numbering is relative to the sequence of the insert and flanking regions from event GR2E rice Nucleotides corresponding to the 5′ flanking sequence (1 1988) T-DNA insert (1989 10984) and 3′ flanking sequence (10985 12772)

Table 33. Thermocycling parameters

Fragment(s)	Denaturation	Annealing	Extension	Cycles
A B G	3 min 95°C			1
	20 sec 98°C	15 sec 55°C	2 min 72°C	35
			7 min 72°C	1†
C	3 min 95°C			1
	20 sec 98°C	15 sec 55°C	2 min 72°C	35
			7 min 72°C	1
D	3 min 95°C			1
	20 sec 98°C	15 sec 55°C	90 sec 72°C	35
			7 min 72°C	1
E	3 min 95°C			1
	20 sec 98°C	15 sec 65°C	2 min 72°C	35
			7 min 72°C	1
F	3 min 95°C			1
	20 sec 98°C	15 sec 60°C	2 min 72°C	35
			7 min 72°C	1

† Following completion of the final extension incubation, samples were held at 4°C

analyzer with ABI BigDye® 3.1 terminator chemistry. Sequence analysis was performed using the Phred, Phrap, and Consed packages (University of Washington) to an error rate of less than 1 in 10,000 bases (Ewing and Green, 1998). Nucleotide sequences from each individual clone were aligned using AlignX™ (Vector NTI Advance™ v. 11.0) to obtain the final consensus sequence for each PCR amplification fragment.

B.4. BLASTP Searches using ORF-1 and ORF-2 as Query Sequences

B.4.1. ORF-1

```
# fasta36 -E 1e-5 orf-1.fasta toxin-database.fasta
FASTA searches a protein or DNA sequence data bank
 version 36.3.8c Dec, 2015(preload9)
Please cite:
```

```
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

Query: orf-1.fasta
  1>>>ORF-1 Right Border spanning open reading frame from rice event GR2E (Oryza sativa) - 68 aa
Library: toxin-database.fasta
  4273861 residues in 25572 sequences

Statistics:  Expectation_n fit: rho(ln(x))= 4.4975+/-0.000834; mu= 12.2290+/- 0.038
 mean_var=40.9307+/-12.620, 0's: 38 Z-trim(107.3): 43  B-trim: 0 in 0/49
 Lambda= 0.200470
 statistics sampled from 8389 (8389) to 8389 sequences
Algorithm: FASTA (3.8 Nov 2011) [optimized]
Parameters: BL50 matrix (15:-5), open/ext: -10/-2
 ktup: 2, E-join: 1 (0.755), E-opt: 0.2 (0.328), width:  16
 Scan time:  0.170
!! No sequences with E() < 1e-05

68 residues in 1 query   sequences
4273861 residues in 25572 library sequences
 Tcomplib [36.3.8c Dec, 2015(preload9)] (4 proc in memory [0G])
 start: Sat Feb 20 15:54:36 2016 done: Sat Feb 20 15:54:36 2016
 Total Scan time:  0.170 Total Display time:  0.000

Function used was FASTA [36.3.8c Dec, 2015(preload9)]
```

B.4.2. ORF-2

```
# fasta36 -E 1e-5 orf-2.fasta toxin-database.fasta
FASTA searches a protein or DNA sequence data bank
 version 36.3.8c Dec, 2015(preload9)
Please cite:
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

Query: orf-2.fasta
  1>>>ORF-2 Left Border spanning open reading frame from rice event GR2E (Oryza sativa) - 79 aa
Library: toxin-database.fasta
  4273861 residues in 25572 sequences

Statistics:  Expectation_n fit: rho(ln(x))= 4.9400+/-0.000628; mu= 10.6087+/- 0.029
 mean_var=36.8657+/-13.008, 0's: 50 Z-trim(107.2): 54  B-trim: 784 in 2/49
 Lambda= 0.211234
 statistics sampled from 8355 (8357) to 8355 sequences
Algorithm: FASTA (3.8 Nov 2011) [optimized]
Parameters: BL50 matrix (15:-5), open/ext: -10/-2
 ktup: 2, E-join: 1 (0.747), E-opt: 0.2 (0.327), width:  16
 Scan time:  0.190
!! No sequences with E() < 1e-05

79 residues in 1 query   sequences
4273861 residues in 25572 library sequences
 Tcomplib [36.3.8c Dec, 2015(preload9)] (4 proc in memory [0G])
 start: Sat Feb 20 15:55:38 2016 done: Sat Feb 20 15:55:38 2016
 Total Scan time:  0.190 Total Display time:  0.000

Function used was FASTA [36.3.8c Dec, 2015(preload9)]
```

B.5. Sequence Similarity Searches to Known Allergens using ORF-1 and ORF-2 Sequences as Queries

The deduced amino sequences of the two junction-spanning open reading frames (ORFs) identified in GR2E rice were used as query sequences to search for significant sequence similarities with 1956 known and putative allergen and celiac protein sequences residing in

the FARRP16 dataset at the University of Nebraska.[16] Search routines included a search for full-length alignments using FASTA to identify any significant sequence alignments (*E*-value threshold 1×10^{-5}) that would be indicative of possible cross-reactivity, as well as the sliding 80-mer search to look for >35 percent sequence identity over all possible 80 amino acid segments and a search for exact matches of 8 contiguous amino acids.

There were no significant alignments or significant amino acid sequence similarity between either ORF-1 (Table 34) or ORF-2 (Table 35) and allergens in the FARRP database.

Table 34. Search results using ORF-1 query sequence against the FARRP database

Database	AllergenOnline Database v16 (January 27, 2016)[a]
Input Query	>ORF-1 Right Border spanning open reading frame from rice event GR2E (Oryza sativa) MIRLSFPAF SLNYQCLEK KRRFISLPV HLHQFCNCN SGCVYVSIH IHTHTLNFQ TILNRLSIL SPVLP
Length	68
Full FASTA Search	No alignments (*E*-value: 1×10^{-5})
Number of 80mers	1
Number of Sequences with Hits	0
Number of 8mers	61
Number of Sequences with Exact 8mer Match	0

[a] Database search performed on 20 February 2016.

Table 35. Search results using ORF-2 query sequence against the FARRP database

Database	AllergenOnline Database v16 (January 27, 2016)[a]
Input Query	>ORF-2 Left Border spanning open reading frame from rice event GR2E (Oryza sativa) MLLDLLALQ EIYRCPGGQ HGRIRNVLL SCLSVNLFT PQFLGNGRC NQSMTTRLE SKQAQILDQ FIGGLSSLP LAFLLLQ
Length	79
Full FASTA Search	No alignments (*E*-value: 1×10^{-5})
Number of 80mers	1
Number of Sequences with Hits	0
Number of 8mers	72
Number of Sequences with Exact 8mer Match	0

[a] Database search performed on 20 February 2016.

[16] The FARRP allergen protein database resides at AllergenOnline (http://www.allergenonline.org). Version 16 was released on January 27, 2016, and contains 1956 peer-reviewed sequences representing 778 taxonomic-protein groups.

C. Description of 2015 and 2016 Field Trials

Event GR2E rice plants of the BC_5F_3 generation (2015 trials), the BC_5F_4 generation (2016 trials) and near-isogenic control PSB Rc82 rice plants (2015 and 2016 trials) were grown at four locations in the Philippines during the rainy season of 2015 and again in the dry season of 2016. The trials were used to generate rice paddy (grain) and straw samples for compositional analysis and for measuring concentrations of CRTI (phytoene desaturase), *Zm*PSY (*Zea mays* phytoene synthase), and PMI (phosphomannose isomerase). In addition, grain collected from these locations was pooled and used to prepare processed fractions (e.g., bran) for the analysis of proximates and minerals. This Appendix provides information relevant to the design and management of these trials.

C.1. Trial Locations and Key Dates

Trials were located at PhilRice research centers located near Batac City, Ilocos Norte (GPS coordinates: Lat. N 18.0547°, Long. E 120.5431°); Muñoz, Nueva Ecija (GPS coordinates: Lat. N 15.067°, Long. E 120.8903°); and San Mateo, Isabela (GPS coordinates: Lat. N 16.87658°, Long. E 121.5959°); and at the International Rice Research Institute, Los Baños (GPS coordinates: Lat. N 14.18333°, Long. E 121.25°) (Tables 36 and 37, Figure 33). The trial site locations match National Cooperative Test (NCT) locations with the exception of the San Mateo site, which replaces PhilRice-Murcia, Negros, Occidental.

Table 36. Locations, planting, and harvest dates for 2015 trials

	Dates				
Site Code	**Seed Sowing**	**Transplanting**	**Harvesting[a]**	**Permit No.**	**Location Description**
BC	5 Jun 2015	25 Jun 2015	6 Oct 2015	2015-0290	PhilRice-Batac Station, Brgy. Tabug, Batac City, Ilocos Norte
LB	29 May 2015	17 Jun 2015	1 Oct 2015	2015-0293	Robert S. Zeigler Experimental Station, Block F8-F9, IRRI, Los Baños, Laguna
MZ	5 Jun 2015	23 Jun 2015	1 Oct 2015	2015-0291	PhilRice-Central Experimental Station, Brgy. Maligaya, Muñoz, Nueva Ecija
SM	5 Jun 2015	25 Jun 2015	28 Sep 2015	2015-0292	PhilRice-Isabela Station, Brgy. Malasin, San Mateo, Isabela

[a] Indicates the start of harvesting, which at some locations extended over two days.

Table 37. Locations, planting, and harvest dates for 2016 trials

	Dates				
Site Code	**Seed Sowing**	**Transplanting**	**Harvesting**	**Permit No.**	**Location Description**
BC	15 Dec 2015	5 Jan 2016	26 Mar 2016	2015-0290	PhilRice-Batac Station, Brgy. Tabug, Batac City, Ilocos Norte
LB	17 Nov 2015	8 Dec 2015	10 Mar 2016	2015-0293	Robert S. Zeigler Experimental Station, Block F8-F9, IRRI, Los Baños, Laguna
MZ	26 Nov 2015	17 Dec 2015	15 Mar 2016	2015-0291	PhilRice-Central Experimental Station, Brgy. Maligaya, Muñoz, Nueva Ecija
SM	14 Jan 2016	4 Feb 2016	28–29 Apr 2016	2015-0292	PhilRice-Isabela Station, Brgy. Malasin, San Mateo, Isabela

Figure 33. Map showing field trial locations in the Philippines

C.2. Experimental Design

During each season, the field experiments were laid out in randomized complete block design with three replications. Each entry was planted in 10-row plots of 5 meters in length (20 cm × 20 cm spacing) for a total plot area of 10 m^2. There were 25 plants per row, totalling 250 plants per entry per plot. A portion of the buffer zone within the experimental site was used as seedbed.

C.3. Temperature and Precipitation Records

Daily mean, minimum and maximum temperatures (°C) and precipitation (mm) totals were recorded at each site from the date of planting until the date of final sample collection (Figures 34 and 35). Average monthly temperatures (minimum, maximum, and mean) and precipitation totals are provided in Tables 38 and 39 for 2015 and 2016 trials, respectively.

C.4. Maintenance Product Applications

Normal pest control and maintenance practices, consistent with rice production for the locale, were used to produce the crop, including border rows. Maintenance applications of fertilizers and pesticides at each site during 2015 are presented in Table 40 and for 2016 in Tables 41, 42, 43, and 44.

C.5. Control of Experimental Bias

Procedures employed to control the introduction of experimental bias in this study included random assignment of entries to planting blocks and uniform maintenance treatments across each plot area.

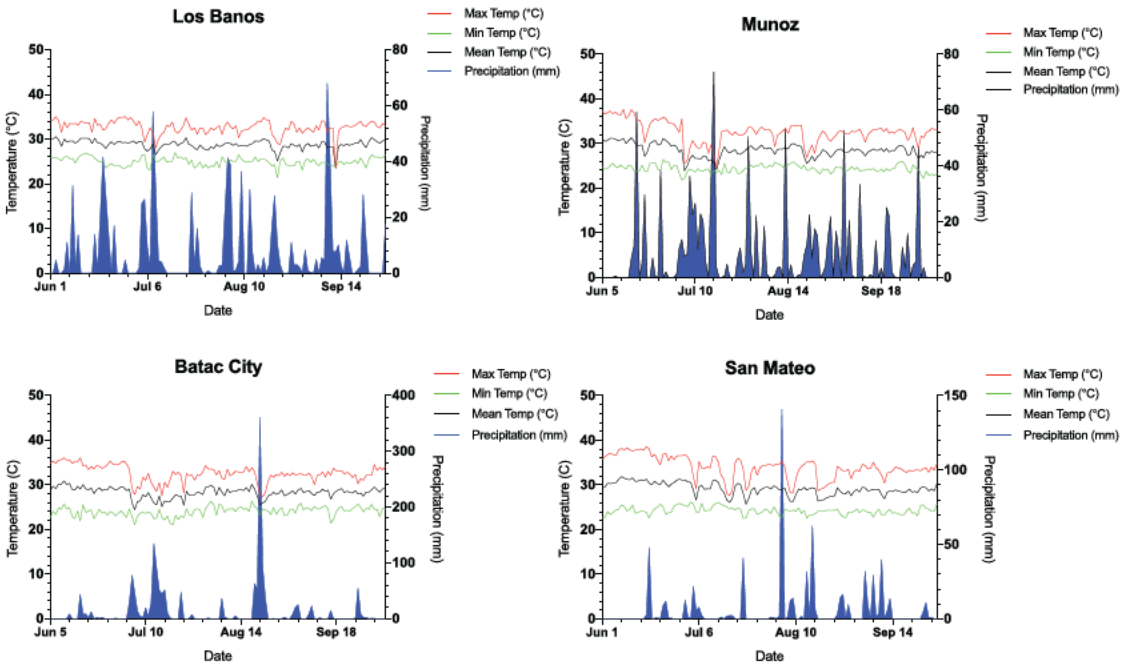


Figure 34. Temperature and precipitation profiles for each trial site location during the 2015 wet season

Table 38. Monthly temperature and precipitation data

Site Code	Month	Temperature (°C)			RH (%)[a]	WS (m/sec)	Total Rain (mm)
		Max	Min	Mean			
BC	June	34.8	23.8	29.3	76.3	–	88.7
	July	31.3	23.2	27.2	84.8	–	664.6
	August	32.1	24.5	28.3	82.9	–	653.2
	September	32.2	24.5	28.3	84.5	–	170.0
LB	June	33.7	25.2	29.4	83.2	1.2	162.7
	July	32.1	24.9	28.5	82.0	1.4	214.2
	August	32.3	24.8	28.5	84.4	1.3	254.7
	September	32.5	25.0	28.7	84.2	0.9	232.9
MZ	June	35.7	24.8	30.3	74.9	0.6	161.2
	July	30.9	23.9	27.4	79.9	7.7	385.4
	August	31.8	24.7	28.3	82.3	6.4	234.4
	September	32.5	24.4	28.5	82.3	5.6	238.1
SM	June	37.0	24.4	30.7	72.4	0.6	74.1
	July	33.3	24.7	29.0	73.1	0.6	100.9
	August	32.6	23.9	28.2	82.3	0.2	318.7
	September	33.4	24.1	28.8	81.9	0.1	153.3

[a] RH = average relative humidity (percent); WS = average wind speed (meters/second).

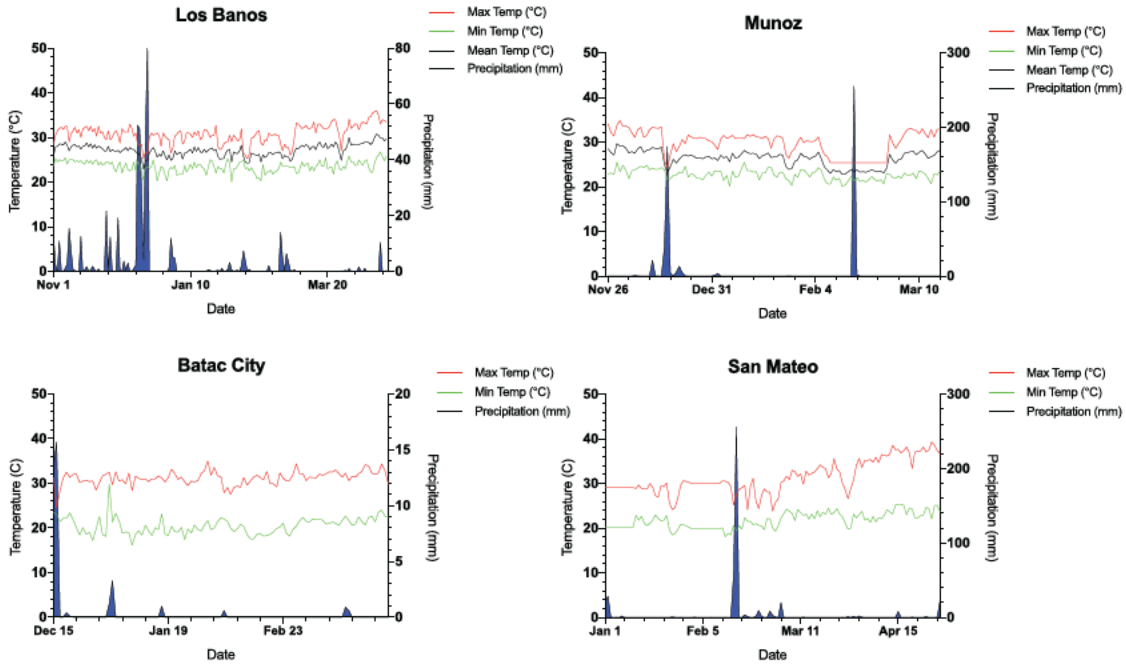


Figure 35. Temperature and precipitation profiles for each trial site location during the 2016 dry season

Table 39. Monthly temperature and precipitation data

Site Code	Month	Temperature (°C)			RH (%)[a]	WS (m/sec)	Total Rain (mm)
		Max	Min	Mean			
BC	December	30.5	20.6	25.5	77.0	0.51	29.6
	January	31.2	20.0	25.6	75.9	0.24	5.5
	February	30.9	19.8	25.3	68.0	0.40	0.6
	March	32.3	21.9	27.1	66.6	0.17	1.5
LB	November	31.2	24.5	27.9	82.7	1.6	107.2
	December	29.9	22.9	26.4	87.7	1.7	308.3
	January	30.1	23.1	26.6	82.7	1.8	17.2
	February	29.2	22.8	26.0	82.3	1.7	38.7
	March	31.8	23.5	27.7	81.4	1.3	2.8
	April[b]	33.7	24.5	29.1	79.5	1.4	13.7
MZ	November[c]	33.3	23.5	28.4	70.3	2.0	0.0
	December	31.0	23.3	27.1	76.4	1.6	262.7
	January	30.3	22.9	26.6	76.9	1.7	4.8
	February	26.7	21.9	24.3	78.8	1.9	257.0
	March[d]	31.7	22.5	27.1	81.9	1.4	0.0
SM	January	28.6	20.8	24.5	83.7	0.4	54.5
	February	29.0	20.3	23.6	84.6	0.7	340.2
	March	31.2	22.6	26.1	82.7	0.9	27.1
	April	36.6	23.4	29.4	69.4	1.0	35.8

[a] RH = average relative humidity (percent); WS = average wind speed (meters/second).
[b] Data up to April 19, 2016.
[c] Data from 26–30 November, 2015.
[d] Data from 1–17 March, 2016.

Table 40. Maintenance product applications at trial site locations in 2015

Site Code	Date	Product	Active Ingredient	Rate
BC	8 Jul 2015	Fertilizer	N-P-K	214.4 kg/ha
	14 Jul 2015	Zinc sulfate	Zinc sulfate	25 kg/ha
	28 Jul 2015	Fertilizer Prevathon Furadan	Urea (46-0-0) Chlorantraniliprole Carbofuran	65.2 kg/ha 450 mL/ha 25 kg/ha
	6 Aug 2015	Prevathon	Chlorantraniliprole	450 mL/ha
	10 Aug 2015	Mancozeb	Mancozeb	450 g/ha
	11 Aug 2015	Fertilizer	Urea (46-0-0)	65.2 kg/ha
	24 Aug 2015	Mancozeb Fertilizer	Mancozeb Urea (46-0-0)	450 g/ha 65.2 kg/ha
	3 Sep 2015	Prevathon Mancozeb	Chlorantraniliprole Mancozeb	450 mL/ha 450 g/ha
LB	6 Jun 2015	Roundup 2,4-D Herbicide	Glyphosate 2,4-D	2.67 L/ha 2.67 L/ha
	11 Jun 2015	Fertilizer	N-P-K	214.4 kg/ha
	18 Jun 2015	Bayluscide	Niclosamide (ethanolamine salt, 20.3%)	2.67 L/ha
	1 Jul 2015	Fertilizer	Urea (46-0-0)	59.94 kg/ha
	7 Jul 2015	Cartap	Cartap hydrochloride	1 kg/ha
	13 Jul 2015	Fertilizer	Urea (46-0-0)	59.94 kg/ha
	15 Jul 2015	Bug Buster	Pyrethrins	2.67 L/ha
	24 Jul 2015	Hopcin	2-sec-butylphenyl-N-methylcarbamate	2.67 L/ha
	5 Aug 2015	Prevathon	Chlorantraniliprole	1.34 L/ha
	16 Aug 2015	Hopcin	2-sec-butylphenyl-N-methylcarbamate	5.35 L/ha
	24 Aug 2015	Cartap	Cartap hydrochloride	534 g/ha
	12 Sep 2015	Cymbush 5EC	Cypermethrin	2.67 L/ha
MZ	23 Jun 2015	Bayluscide	Niclosamide, methyl isobutyl ketone	1.05 L/ha
	7 Jul 2015	Fertilizer Furadan	N-P-K Carbofuran	60 kg/ha 33.4 kg/ha
	16 Jul 2015	Prevathon	Chlorantraniliprole	400 mL/ha
	7 Aug 2015	Fertilizer Furadan	Urea (46-0-0) Carbofuran	30 kg/ha 33.4 kg/ha
	1 Sep 2015	Furadan	Carbofuran	33.4 kg/ha
	25 Sep 2015	Cymbush 5EC Actara	Cypermethrin Thiamethoxam	450 mL/ha 950 g/ha
SM	6 Jul 2015	Fertilizer	N-P-K	214.4 kg/ha
	22 Jul 2015	Actara	Thiamethoxam	950 g/ha
	24 Jul 2015	Fertilizer Zinc phosphide	Urea (46-0-0) Zinc phosphide	65.2 kg/ha 950 g/ha
	27 Jul 2015	Actara	Thiamethoxam	950 g/ha
	30 Jul 2015	Actara	Thiamethoxam	950 g/ha
	3 Aug 2015	Fertilizer	N-P-K	214.4 kg/ha
	5 Aug 2015	Fertilizer Kocide	Urea (46-0-0) Copper hydroxide	65.2 kg/ha 1.9 kg/ha
	11 Aug 2015	Kocide	Copper hydroxide	1.9 kg/ha
	17 Aug 2015	Zinc phosphide	Zinc phosphide	950 g/ha
	19 Aug 2015	Kocide	Copper hydroxide	1.9 kg/ha
	3 Sep 2015	Zinc phosphide	Zinc phosphide	950 g/ha

Table 41. Maintenance product applications at Batac City trial site in 2016

Date	Product	Active Ingredient	Rate
4 Jan 2016	Vermicompost		5 tons/ha
12 Jan 2016	Fertilizer	N-P-K (14-14-14)	30 kg/ha
19 Jan 2016	Prevathon Mancozeb	Rynaxypyr Fungicide (Dithane M-45)	1.48 L/ha 500 g/ha
25 Jan 2016	Fertilizer Zinc sulfate	Urea (46-0-0) Zinc sulfate	30 kg/ha 25 kg/ha
28 Jan 2016	Furadan	Carbofuran	25 kg/ha
9 Feb 2016	Fertilizer	Urea (46-0-0)	30 kg/ha
11 Feb 2016	Funguran	Copper hydroxide	1.27 kg/ha
16 Feb 2016	Fertilizer	Urea (46-0-0)	30 kg/ha
2 Mar 2016	Funguran	Copper hydroxide	1.27 kg/ha

Table 42. Maintenance product applications at Los Baños trial site in 2016

Date	Product	Active Ingredient	Rate
3 Dec 2015	Organic Kuhol Buster	Metaldehyde	157 kg/ha
6 Dec 2015	Fertilizer	NPK (14-14-14)	214.4 kg/ha
8 Dec 2015	Bayluscide Sofit 300 EC	Niclosamide (ethanolamine salt) 20.3% Pretilachlor	2.19 L/ha 0.94 L/ha
20 Dec 2015	Cymbush 5EC Cartap	Cypermethrin Cartap Hydrochloride	2.5 L/ha 0.62 kg/ha
21 Dec 2015	Fertilizer	Urea (46-0-0)	65.24 kg/ha
28 Dec 2015	Fertilizer	Urea (46-0-0)	65.24 kg/ha
5 Jan 2016	Cymbush 5EC Regent	Cypermethrin Fipronil	2.5 L/ha 62.8 kg/ha
16 Feb 2016	Ascend Score	Fipronil 200 g/L Difenoconazole	3.14 L/ha 0.12 L/ha

Table 43. Maintenance product applications at Muñoz trial site in 2016

Date	Product	Active Ingredient	Rate
17 Dec 2015	Bayluscide	Niclosamide, methyl isobutyl ketone	1.05 L/ha
28 Dec 2015	Fertilizer Furadan	N-P-K (14-14-14) Carbofuran	60 kg/ha 33.4 kg/ha
28 Jan 2016	Fertilizer Furadan	Urea (46-0-0) Carbofuran	60 kg/ha 33.4 kg/ha

Table 44. Maintenance product applications at San Mateo trial site in 2016

Date	Product	Active Ingredient	Rate
15 Feb 2016	Fertilizer Fertilizer	N-P-K (14-14-14) Urea (46-0-0)	429 kg/ha 130.4 kg/ha
19 Feb 2016	Actara Chix	Thiamethoxam beta-Cypermethrin	950 g/ha 950 g/ha
24 Feb 2016	Actara Chix	Thiamethoxam beta-Cypermethrin	950 g/ha 950 g/ha
25 Feb 2016	Fertilizer	Urea (46-0-0)	130.4 kg/ha
29 Feb 2016	Actara Chix	Thiamethoxam beta-Cypermethrin	950 g/ha 950 g/ha
3 Mar 2016	Actara Chix	Thiamethoxam beta-Cypermethrin	950 g/ha 950 g/ha
7 Mar 2016	Fertilizer Fertilizer	N-P-K (14-14-14) (21-0-0)	429 kg/ha 289 kg/ha
13 Mar 2016	Kocide	Copper hydroxide	1.9 kg/ha
15 Mar 2016	Actara	Thiamethoxam	950 g/ha
18 Mar 2016	Prevathon	Chlorantraniliprole	900 g/ha
22 Mar 2016	Kocide Chix	Copper hydroxide beta-Cypermethrin	1.9 kg/ha 950 g/ha
29 Mar 2016	Actara Kocide	Thiamethoxam Copper hydroxide	950 g/ha 1.9 kg/ha
31 Mar 2016	Furadan Prevathon	Carbofuran Chlorantraniliprole	20 kg/ha 900 g/ha
4 Apr 2016	Chix	beta-Cypermethrin	950 g/ha
7 Apr 2016	Furadan Prevathon	Carbofuran Chlorantraniliprole	20 kg/ha 900 g/ha
11 Apr 2016	Chix	beta-Cypermethrin	950 g/ha
13 Apr 2016	Prevathon	Chlorantraniliprole	900 g/ha
19 Apr 2016	Prevathon	Chlorantraniliprole	900 g/ha
26 Apr 2016	Prevathon	Chlorantraniliprole	900 g/ha

C.6. Rice Growth Stages

Table 45. Phenological growth stages and BBCH identification keys for rice

Principal Growth Stage		Sub-Stage	Descriptions[†]
0	Germination	00	Dry seed (caryopsis)
		03	Seed imbibition complete
		05	Radicle emerged from caryopsis
		07	Coleoptile emerged from caryopsis
1	Leaf development[a]	10	Imperfect leaf unrolled, tip of first true leaf visible
		11...19	First leaf unfolded, subsequent leaves unfolded up to nine or more leaves unfolded (11...19)
2	Tillering[b]	21	Beginning of tillering: first tiller detectable
		22	Two tillers detectable
		23...29	Stages continuous until maximum number of tillers detectable
3	Stem elongation[c]	30	Panicle initiation or green ring stage: chlorophyll accumulates in the stem tissue, forming a green ring
		32	Panicle formation: panicle 1–2 mm in length
		34	Internode elongation or jointing stage: internodes begin to elongate, panicle more than 2 mm long (variety-dependent)
		39	Flag leaf stage: flag leaf unfolded, collar regions (auricle and ligule) of flag leaf and penultimate leaf aligned (pre-boot stage)
4	Booting	41	Early boot stage: upper part of stem slightly thickened, sheath of flag leaf about 5 cm out of penultimate leaf sheath
		43	Mid-boot stage: sheath of flag leaf 5–10 cm out of the penultimate leaf sheath
		45	Late boot stage: flag leaf sheath swollen, sheath of flag leaf more than 10 cm out of penultimate leaf sheath
		49	Flag leaf sheath open
5	Inflorescence emergence, heading	51	Beginning of panicle emergence: tip of inflorescence emerged from sheath
		52...58	20%, 30%...80% of panicle emerged (stages 52...58, respectively)
		59	End of panicle emergence: neck node level with the flag leaf auricle, anthers not yet visible
6	Flowering, anthesis	61	Beginning of flowering: anthers visible at top of panicle
		65	Full flowering: anthers visible on most spikelets
		69	End of flowering: all spikelets have completed flowering but some dehydrated anthers may remain
7	Development of fruit	71	Watery ripe: first grains have reached half their final size
		73	Early milk
		75	Medium milk: grain content milky
		77	Late milk
8	Ripening	83	Early dough
		85	Soft dough: grain content soft but dry, fingernail impression not held, grains and glumes still green
		87	Hard dough: grain content solid, fingernail impression held
		89	Fully ripe: grain hard, difficult to divide with thumbnail
9	Senescence	92	Over-ripe: grain very hard, cannot be dented by thumbnail
		97	Plant dead and collapsing
		99	Harvested product

[†] As described in Lancashire et al. (1991).
[a] A leaf is unfolded when its ligule is visible or the tip of the next leaf is visible.
[b] Tillering or stem elongation may occur earlier than stage 13; in this case continue with stages 21 or 30.
[c] If stem elongation begins before the end of tillering continue with stage 30.

D. Materials and Methods for Protein Expression Analysis

In order to estimate potential human and animal dietary exposure to the *Zm*PSY1, CRTI, and PMI enzymes expressed in GR2E rice, the concentration of these proteins in plant tissues was determined by quantitative ELISA. Analyses were performed on samples of grain and straw, which represent the only potential pathways of dietary exposure, that were collected from GR2E plants grown at four locations in the Philippines during the rainy season in 2015 (Oliva et al., 2016a) and again during the dry season in 2016 (Oliva et al., 2016c). This appendix provides information on materials and methods used for ELISA testing for *Zm*PSY1, CRTI, and PMI proteins.

D.1. ELISA Calibration Standards

D.1.1. ZmPSY1 Protein

Lot Number	M20452-05
Date of Manufacture	14 November 2015
Concentration	2.39 mg/ml (amino acid analysis)
Formulation Buffer	50 mM Tris-HCl pH 7.6, 600 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP), 0.035% lauryldimethylamine-N-oxide (LDAO)
Purity by SDS-PAGE	>95%
Storage	–80°C
Manufacturer	Aldevron, 5602 Research Park Blvd, Madison, WI 53719

The production and characterization of this lot of *Zm*PSY1 protein was described in MacKenzie (2016c) and a certificate of analysis is provided on page 138.

D.1.2. CRTI Protein

Lot Number	M20454-02
Date of Manufacture	9 October 2015
Concentration	1.10 mg/ml (amino acid analysis)
Formulation Buffer	50 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP)
Purity by SDS-PAGE	97.8%
Storage	–80°C
Manufacturer	Aldevron, 5602 Research Park Blvd, Madison, WI 53719

The production and characterization of this lot of CRTI protein is described in MacKenzie (2016a) and a certificate of analysis is included on page 139.

D.1.3. PMI Protein

Lot Number	21038G-PMI
Date of Manufacture	18 November 2010
Concentration	11.9 mg/ml; Bradford method
Formulation Buffer	50 mM potassium phosphate pH 7.62
Purity by SDS-PAGE	>90%
Storage	–20°C
Manufacturer	Center for Biocatalysis and Bioprocessing, University of Iowa, Coralville, IA 52241

A certificate of analysis for this lot of PMI protein is included on page 141.

D.2. Test and Control Substances

For each year, the test substance (entry 01) consisted of rice seed containing event GR2E (Table 46) from the BC_5F_3 or BC_5F_4 generations following introgression and repeated back-crossing with PSB Rc82 as the recurrent parent (Figure 36). The control substance (entry 02) consisted of non-transgenic unmodified PSB Rc82 rice.

Table 46. Samples for novel protein concentration analysis

Entry	Identity	Sample Type	Field Locations	Replications	No. of Samples
01	GR2E	Grain (milk)	4	3	12
		Grain (dough)	4	3	12
		Grain (mature)	4	3	12
		Straw (mature)	4	3	12
02	PSB Rc82	Grain (milk)	4	3	12
		Grain (dough)	4	3	12
		Grain (mature)	4	3	12
		Straw (mature)	4	3	12

† Grain developmental stages of milk, dough, and mature corresponded to BBCH 75, BBCH 85, and BBCH 90, respectively, as described in Lancashire et al. (1991).

D.3. Field Phase

The field phase of this study was conducted at four sites located in the rice-growing regions of the Philippines during 2015 and 2016. Planting was done according to local agronomic practices and all procedures and planting activities were documented in the Field Trial Notebook for each trial site. Normal pest control and maintenance practices, consistent with rice production for the locale, were used to produce the crop, including border rows. A description of maintenance practices (fertilization, herbicide, and pesticide applications) and local meteorological data are included in Appendix C, page 113.

D.4. Tissue Collection

Grain and straw samples for protein concentration analysis were collected from impartially selected, healthy, representative plants from each entry. For each tissue type, one sample was collected from each of three replicated blocks of GR2E and control PSB Rc82 rice grown at each trial site location during each growing season (i.e., three samples per tissue and sampling time per site).

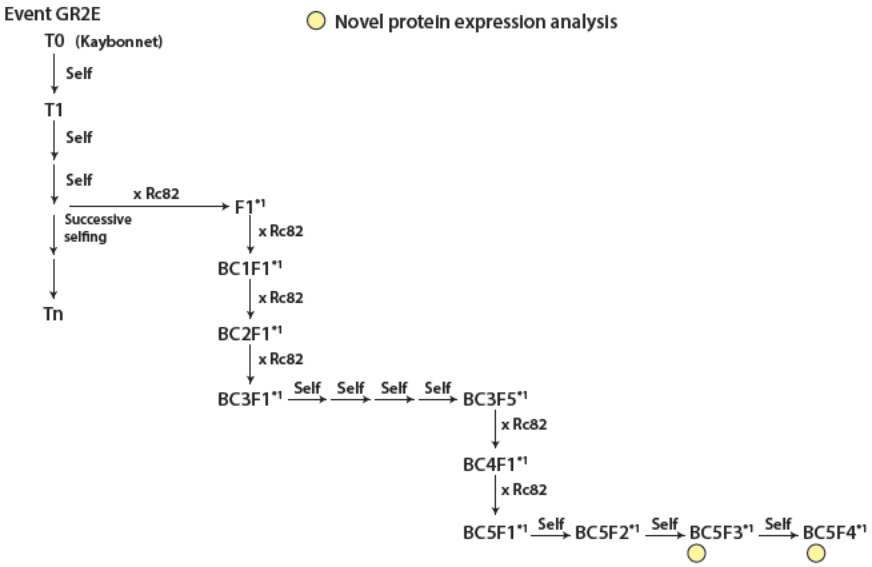


Figure 36. Breeding pedigree of event GR2E used for novel protein expression analyses. Plants from the BC_5F_3 or BC_5F_4 generations in PSB Rc82 background were planted at four locations in the Philippines during 2015 and 2016, respectively.

D.4.1. Grain Samples

Samples of rice grain were collected at the milk (*ca*. BBCH 75), dough (*ca*. BBCH 85), and mature (*ca*. BBCH 90) stages of development (Lancashire et al., 1991) (see also Table 45 in Appendix C.6 for descriptions of rice growth stages). Each individual sample was a composite of material obtained from at least five representative plants within each block. Approximately 3–4 panicles were collected from each representative plant and placed in a pre-labelled net bag, one per block. Following collection of all plants per block, samples were placed on dry ice and transported to IRRI where grains for each sample and growth stage were removed from the panicles, mixed, and *ca*. 100 g placed in pre-labelled 50 ml screw-top Falcon tubes and placed in –80°C storage until processed for analysis.

D.4.2. Straw Samples

Samples of rice straw were collected at harvest (*ca*. BBCH 90). Each individual sample was a composite of material obtained from at least five representative plants within each block. Following collection, straw was removed from the collection sack, chopped into small pieces *ca*. 12 cm in length, mixed, and *ca*. 100 g placed in a pre-labelled sample bag for shipment to IRRI on dry ice.

D.5. Quantitative ELISA

Concentrations of the *Zm*PSY1, CRTI, and PMI proteins were determined using specific quantitative ELISA methods.

D.5.1. Sample Extraction

For grain samples, hulls were removed from individual grains prior to grinding and extraction. All samples of rice grain and straw were ground to a fine powder in liquid nitrogen using a mortar and pestle and 100 mg aliquots of each sample were used for extraction. Samples to be tested for *Zm*PSY1 were resuspended with 10 volumes (*ca*. 1 ml) of 30 mM sodium phosphate pH 7.1, 0.15 M NaCl, 1% (v/v) Triton X-100, vortexed for 1 min, and incubated on ice for 30 min. Samples to be tested for CRTI were similarly extracted; however, the extraction buffer was composed of 25 mM Tris-HCl pH 7.3 containing 0.5% (w/v) SB3-10 Zwittergent (caprylyl sulfobetaine). Samples for PMI testing were resuspended in 500 volumes (*ca*. 50 ml) of 10 mM sodium phosphate pH 7.4 containing 0.05% (v/v) Tween-20, vortexed for 1 min, and placed on ice for 30 min.

D.5.2. ZmPSY1 ELISA

The *Zm*PSY1 ELISA method utilized the QuantiPlate™ Kit for PSY (EnviroLogix, Portland, ME, USA; AP097-V100). Aliquots (100 μl) of standards (analyzed in triplicate wells) and diluted samples (analyzed in triplicate) were incubated in EIA microtitre plate wells, previously coated with ELX1036 anti-*Zm*PSY1 monoclonal antibody (MAb), for two hours at room temperature with gentle agitation (rotary plate shaker, 200 rpm). Wells were washed six times with 300 μl volumes of phosphate buffered saline (PBS)-Tween (10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 2.7 mM KCl, 137 mM NaCl, pH 7.4 containing 0.05% v/v Tween-20) using a Labexim LMW4 plate washer and then incubated with 100 μl of HRP-conjugated ELX1035 anti-*Zm*PSY1 MAb for two hours at room temperature with gentle agitation (as above). Wells were washed as before and bound HRP-antibody complex was detected by adding 100 μl of HRP chromogenic substrate, 3,3',5,5'–tetramethylbenzidine (TMB), and allowing colour development for 30 min prior to addition of 100 μl stop reagent (1 N HCl). The optical density (OD) at 450 nm was measured using a microplate reader, with OD_{650nm} background subtraction.

D.5.3. CRTI ELISA

The CRTI ELISA method utilized the QuantiPlate™ Kit for CRTI (EnviroLogix, AP096-V100) and followed the same protocol as described for the *Zm*PSY1 ELISA with the following exceptions: microtitre plate wells were coated with ELX1024 anti-CRTI MAb and HRP-conjugated ELX1043 MAb was used for detection.

D.5.4. PMI ELISA

The quantification of PMI protein utilized the AgraQuant™ PMI Plate (Romer Labs® Union, MO, USA; Part No. 7020106) ELISA and was performed according to the manufacturer's instructions. Protein standards and test samples were analyzed in triplicate, incubation times for samples and enzyme conjugate were 20 min each at room temperature without agitation. Plate washing after each incubation step was as described above for the *Zm*PSY1 ELISA. Following HRP chromogenic substrate colour development, the OD_{450nm} was measured using a microplate reader.

D.5.5. Calculations for Determining Protein Concentration

A standard curve was included on each ELISA plate and data were fitted to the four-parameter logistic function shown in equation (1).

$$y = d + \frac{a - d}{1 + (x/c)^b} \tag{2}$$

Where y is the response value (i.e., optical density) and x is the known standard concentration. Interpolation of the sample concentration (ng/ml) was accomplished by solving for x in equation (2) using values for a, b, c, and d determined from the standard curve.

$$x \text{ (ng/ml)} = c\left(\frac{a - d}{y - d} - 1\right)^{\frac{1}{b}} \tag{3}$$

Sample concentrations (ng/ml) were converted to amounts per weight of tissue homogenized for each sample type and analyte, and expressed as ng protein per gram fresh weight tissue.

$$\text{Amount (ng/g FWT)} = \frac{x \text{ (ng/ml)} \times \text{Dilution Factor} \times \text{Extraction Volume (ml)}}{\text{Sample Weight (g)}} \tag{4}$$

D.6. Confirmation of Trait Presence

Samples of GR2E and control rice grain (milk, dough, and mature growth stages) from one replicate block at each trial site location were analyzed by multiplex PCR zygosity testing to confirm the homozygous presence of the pSYN12424 T-DNA insert in samples derived from GR2E rice and the absence of the insert in samples from control PSB Rc82 rice. The principle of the assay is illustrated in Figure 31 (Appendix A, page 107).

DNA isolation was performed essentially following the CTAB extraction method Doyle and Doyle (1987). Individual PCR reaction mixtures contained 300 ng genomic DNA, 0.5 mM of each primer (ZD-E1-P1, ZD-E1-P2, and ZD-E1-P3), 0.2 mM deoxynucleotide triphosphates, 1.5 mM $MgCl_2$, and 0.5 units KAPA Taq DNA polymerase (KAPA Biosystems) in 10 μl KAPA Taq buffer. After an initial 5 min denaturation at 95°C, PCR amplifications were performed for 35 cycles consisting of denaturation (95°C, 45 sec), annealing (55°C, 45 sec), and extension (72°C, 45 sec). A final extension step (72°C, 8 min) was included following the last PCR cycle, after which samples were held at 25°C.

Following PCR, reaction mixtures were analyzed by agarose gel electrophoresis in 1.5 percent agarose and amplified fragments were visualized with SYBR® Safe DNA Gel Stain (ThermoFisher Scientific).

The homozygous presence of the pSYN12424 T-DNA was confirmed for all GR2E rice samples tested and absence of the insert was confirmed for control PSB Rc82 rice samples (Figures 37 and 38).

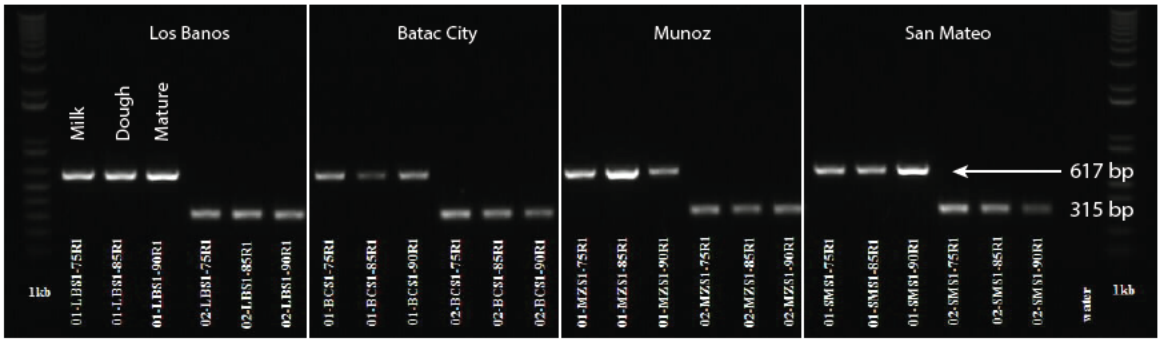


Figure 37. Genomic DNA was prepared from samples of GR2E and control rice grain (milk, dough, mature stages) collected from one replicate block at the Batac City, Los Baños, Muñoz, and San Mateo locations in 2015. Multiplex PCR was performed to confirm the homozygous presence of the pSYN12424 T-DNA insert in samples derived from GR2E plants (i.e., amplification of 617 bp fragment) and the absence of the T-DNA insert samples from control PSB Rc82 plants (i.e., amplification of 315 bp fragment).

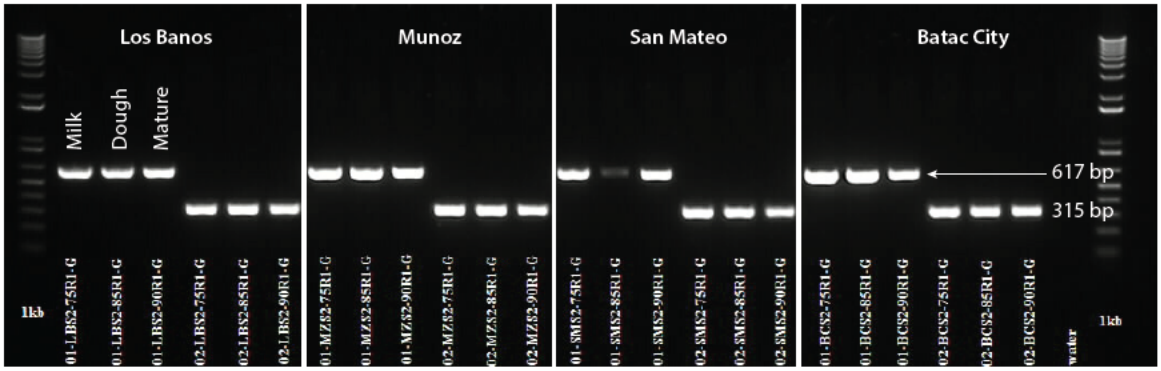


Figure 38. Genomic DNA was prepared from samples of GR2E and control rice grain (milk, dough, mature stages) collected from one replicate block at the Batac City, Los Baños, Muñoz, and San Mateo locations in 2016. Multiplex PCR was performed to confirm the homozygous presence of the pSYN12424 T-DNA insert in samples derived from GR2E plants (i.e., amplification of 617 bp fragment) and the absence of the T-DNA insert samples from control PSB Rc82 plants (i.e., amplification of 315 bp fragment).

E. Materials and Methods for Composition Analysis

Compositional analyses were performed on samples of rice grain and straw obtained from event GR2E introgressed into PSB Rc82 (BC_5F_3 in 2015; BC_5F_4 in 2016; Figure 5, page 29) and near-isogenic control PSB Rc82 rice that were grown in side-by-side trials at four separate sites in the Philippines (Batac City, Los Baños, Muñoz, and San Mateo) during 2015 and 2016. This appendix briefly describes the experimental design, procedures, and statistical analysis conducted on the composition data derived from multi-year trials of GR2E rice in the Philippines and further details can be found in the reports of Swamy et al. (2016a) and Swamy et al. (2016b).

E.1. Experimental Design

The field phase of this study was conducted during the rainy season in 2015 and the dry season in 2016 at four sites located in the rice-growing regions of the Philippines (Table 47). Planting was done according to local agronomic practices. Three blocks (replicates) of each entry (event GR2E in PSB RC82 background and near-isogenic non-transgenic control PSB RC82) were established at each test site in a randomized complete block design. Each entry was planted in 10-row plots of 5 meters in length (20 cm × 20 cm spacing) for a total plot area of 10 m^2. There were 25 plants per row, totalling 250 plants per entry per plot.

Table 47. Trial locations for grain and straw production

Trial Site Code	DOST-BC Permit No.	Approval Date	Trial Size (m^2)	Location Description
BC	2015-0290	15 May 2015	200	PhilRice-Batac Station, Brgy. Tabug, Batac City, Ilocos Norte.
LB	2015-0293	7 May 2015	200	Robert S. Zeigler Experimental Station, Block F8-F9, Los Baños, Laguna.
MZ	2015-0291	8 May 2015	200	PhilRice Central Experimental Station, Brgy. Maligaya, Muñoz, Nueva Ecija.
SM	2015-0292	14 May 2015	200	PhilRice Isabela Station, Brgy. Malasin, San Mateo, Isabela.

E.2. Maintenance of Field Plots

Normal pest control and maintenance practices, consistent with rice production for the locale, were used to produce the crop, including border rows. For more details on trial site management, see Appendix C, page 113.

E.3. Sample Collection and Shipment

Grain and straw samples were collected from matured rice plants, the stage when typical grain harvest would occur. Grain was collected from at least 150 plants per plot, excluding the outer rows, and pooled into a single sample per plot. Straw was collected from at least 8 plants per plot, randomly selected and excluding the outer rows, and pooled into a single sample per plot.

Following collection, samples were shipped frozen to EPL Bio-Analytical Services (Niantic, IL) for processing and analysis.

E.4. Compositional Parameters

Consistent with guidance contained within the OECD consensus document on compositional considerations for new rice varieties (OECD, 2016), the following parameters were analyzed in samples of paddy rice and straw:

E.4.1. Paddy Rice

Proximates and Fibre: Moisture, crude protein, crude fat, ash, ADF, NDF, crude fibre, TDF, and carbohydrate.

Sugars: Total starch and amylose.

Fatty Acids: Caprylic (C8:0), capric (C10:0), lauric (C12:0), myristic (C14:0), pentadecanoic (C15:0), palmitic (C16:0), palmitoleic (C16:1 Δ9); heptadecanoic (C17:0), stearic (C18:0), oleic (C18:1 Δ9), linoleic (C18:2 Δ9,12), α-linolenic (C18:3 Δ9,12,15), arachidic (C20:0), eicosenoic (C20:1), eicosadienoic (C20:2 Δ11,14), eicosatrienoic (C20:3 Δ11,14,17), arachidonic (C20:4 Δ5,8,11,14), behenic (C22:0), erucic (C22:1 Δ13), lignoceric (C24:0), and nervonic (C24:1 Δ15).

Amino Acids: Lysine, arginine, glycine, histidine, isoleucine, leucine, phenylalanine, threonine, valine, alanine, aspartic acid, glutamic acid, proline, serine, tyrosine, cystine, methionine, and tryptophan.

Minerals: Calcium, phosphorus, magnesium, potassium, zinc, manganese, copper, iron, and sodium.

Vitamins: Thiamine (B1), riboflavin (B2), niacin (B3), pantothenic acid (B5), pyridoxine (B6), folic acid (B9), α-tocopherol (E), and β-carotene.

Anti-Nutrients: Phytic acid and trypsin inhibitor.

E.4.2. Rice Straw and Bran

Proximates and Fibre: Moisture, crude protein, crude fat, ash, ADF, NDF, crude fibre, and carbohydrate.

Minerals: Calcium and phosphorus.

The methods used for analysis of rice grain compositional parameters are itemized in Table 48 and the methods used for the analysis of rice straw and bran are listed in Table 49.

Table 48. Rice grain analytical methods

Compositional Parameter	SOP No.	Method Reference	Reporting Units†
Proximates and Fibre			
Moisture	4	AOAC 925.09	%
Crude Protein	20	Foss-Tecator (Kjeldahl)	% DB
Crude Fat	230	AOCS Method AM 5-04; Ankom Technology	% DB
Ash	2	AOAC 923.03	% DB
Carbohydrate	494	USDA Agricultural Handbook	% DB
Acid Detergent Fibre	3	Ankom Technology	% DB
Neutral Detergent Fibre	9	Ankom Technology	% DB
Crude Fibre	5	Ankom Technology	% DB
Total Dietary Fibre	359	Ankom Technology; AOAC 991.43; Foss-Tecator	% DB
Sugars			
Total starch	462	AOAC 996.11	% DB
Amylose	NC-549	Megazyme	% FW
Minerals			
Listing below: (Calcium, phosphorus, magnesium, potassium, zinc, manganese, copper, iron, and sodium)	60	AOAC 999.11; CEM Corporation Operation Manual (ICP-OES)	% DB
Amino Acids			
Cystine and Methionine	279	AOAC 994.12	% DB
Tryptophan	22	*J. Micronutr. Anal.* 7:27–25 (1990)	% DB
Remaining 15 amino acids (Lysine, arginine, glycine, histitdine, isoleucine, leucine, phenylalanine, threonine, valine, alanine, aspartic acid, glutamic acid, proline, serine, and tyrosine)	58	*J. Chromatogr. A.* 670:59–66 (1994); Waters Corporation	% DB
Fatty Acids			
Complete fatty acid profile	319	AOAC 939.05; AOCS Method Ce 1e-91; AOCS Method Ce 2-66	% Total Fatty Acids
Vitamins			
Thiamine (B1) and riboflavin (B2)	281	AACC Method 86-80	mg/kg DB
Niacin (B3)	28	AACC Method 86-51	mg/kg DB
Pantothenic Acid (B5)	41	AOAC 945.74	mg/kg DB
Pyridoxine (B6)	220	AACC Method 86-31	mg/kg DB
Folic acid (B9)	26	AACC Method 86-47	mg/kg DB
α-Tocopherol (vitamin E)	346	*J. Am. Oil. Chem. Soc.* 61:1231–1234 (1984); *Anal. Sci.* 21:1545–1548 (2005)	mg/kg DB
β-Carotene	341	AOAC 941.15; *Journal of AOAC Int.* 87(5):1070–1082 (2004)	mg/kg DB
Anti-Nutrients			
Phytic acid	235	AOAC 986.11	% DB
Trypsin inhibitor	12	AOCS Method Ba 12-75	TIU/mg DW

† FW = Fresh Weight; DB = Dry Basis; DW = Dry Weight; TIU = Trypsin Inhibitor Units

E.5. Statistical Analysis

All statistical analyses were conducted in *R* (R Core Team, 2015), including the "lmerTest" (Bates et al., 2014) package.

E.5.1. Mixed Model for Multi-Year Combined-Sites Analysis

For a given compositional analyte, data were analyzed using the following linear mixed model (Asfaw et al., 2012):

$$y_{ijk} = \mu_i + l_j + s_k + (\mu l)_{ij} + (\mu s)_{ik} + (ls)_{jk} + (\mu ls)_i + \varepsilon_{ijk}$$

Table 49. Rice straw and bran analytical methods

Compositional Parameter	SOP No.	Method Reference	Reporting Units[†]
Proximates and Fibre			
Moisture	25	AOAC 930.15	%
Crude Protein	20	Foss-Tecator (Kjeldahl)	% DB
Crude Fat	141	AOCS Method AM 5-04; Ankom Technology	% DB
Ash	2	AOAC 923.03	% DB
Carbohydrate	494	USDA Agricultural Handbook	% DB
Acid Detergent Fibre	3	Ankom Technology	% DB
Neutral Detergent Fibre	9	Ankom Technology	% DB
Crude Fibre	5	Ankom Technology	% DB
Minerals			
Calcium and phosphorus	60	AOAC 999.11; CEM Corporation Operation Manual (ICP-OES)	% DB

[†] FW = Fresh Weight; DB = Dry Basis; DW = Dry Weight

where μ_i denotes the mean of the i^{th} entry (fixed effect), l_j denotes the effect of the j^{th} site (random effect), s_k denotes the effect of the s^{th} season (random effect), $(\mu l)_{ij}$ denotes the interaction between the entries and sites (random effect), $(\mu s)_{ik}$ denotes the interaction between the entries and growing seasons (random effect), $(ls)_{jk}$ denotes the interaction between sites and growing seasons (random effect), $(\mu ls)_i$ denotes the interaction between entries, locations, and growing seasons (random effect), and ε_{ijk} denotes the residual error (random effect).

E.5.2. R Statistical Procedures

The “lmer” procedure from the “lmerTest” package was used to fit the linear mixed model and to generate estimates of variance components and p-values. For each compositional analyte, the least squares (LS)-mean value across years and sites was estimated from the corresponding statistical model for GR2E rice and the control PSB Rc82 rice using the "lsmeans" package (Lenth, 2013). The 95 percent confidence interval (CI) for each of the entry means was formed by:

$$\text{LS-mean} \pm t_{0.975,v} \times \text{standard error of LS-mean}$$

where $t_{0.975,v}$ denotes the upper 0.025 percentage point of a t-distribution with v degrees of freedom. The degrees of freedom were determined by Kenward-Roger method (Kenward and Roger, 1997).

E.5.3. Statistical Comparisons and Interpretations

The first step in the evaluation was to test for differences in LS-mean values between GR2E rice and the control PSB Rc82 rice. Where a statistically significant difference (p-value < 0.05) was identified in the multi-year combined-sites analysis, further context for interpreting the possible biological significance of the difference was gathered through comparisons with the range of values for each analyte reported in the published literature (e.g., (OECD, 2016; NARO, 2011; Heuzé and Tran, 2015; Juliano and Bechtel, 1985)) or available from the ILSI Crop Composition Database (ILSI, 2014). Analyte ranges for GR2E rice that fell within the combined literature range for that analyte were considered to be within the range of normal variability of conventional rice.

E.6. Multi-Year Combined-Sites Statistical Tables

Compositional data were fitted to a linear mixed model using the "lmer" procedure of the R “lmerTest” package (Bates et al., 2014). For any given component, the form of the model was:

$$param.mod <- lmer(param \sim geno\ (1|loc) + (1|year) + (1|geno:loc) + (1|geno:year) + (1|loc:year) + (1|geno:loc:year), data=dataframe)$$

Where: param = a specific compositional component; geno = genotype (event GR2E or PSB Rc82 control); loc = trial site location (Batac City, Los Baños, Muñoz, San Mateo); and year = growing season, 2015 or 2016.

Least square means and 95 percent confidence intervals were derived from the fitted model using the “lsmeans” package using the Kenward-Roger method for estimation of degrees of freedom (Kenward and Roger, 1997).

The remainder of this page is intentionally left blank.

Table 50. Multi-year combined-sites analysis for straw proximates, fibre, and minerals

Analyte	Statistic	GR2E	RC82
moisture (% FW)	LSMean	12	12 57
	95% C	9 55<->14 45	10 12<->15 02
	Range	(9 24-14)	(8 12-22 1)
	N	24	24
	Pr(>F)	0 5932	
	Conditional R2	0 336	
	LSMeansDi	-0 571	
	LSMeansDi C	-3 73<->2 58	
	LSMeansPctDi	-4 54	
	LSMeansPctDi C	-29 6<->20 6	
crude at (% DB)	LSMean	2 58	2 716
	95% C	0 9377<->4 221	1 074<->4 357
	Range	(1 31-4 8)	(0 995-6 15)
	N	24	24
	Pr(>F)	0 7668	
	Conditional R2	0 912	
	LSMeansDi	-0 136	
	LSMeansDi C	-1 47<->1 2	
	LSMeansPctDi	-5 01	
	LSMeansPctDi C	-54<->44	
carbohydrates (% DB)	LSMean	66 12	66 37
	95% C	64 32<->67 92	64 57<->68 17
	Range	(60 5-69 2)	(62 1-70 5)
	N	24	24
	Pr(>F)	0 7345	
	Conditional R2	0 39	
	LSMeansDi	-0 249	
	LSMeansDi C	-1 92<->1 42	
	LSMeansPctDi	-0 375	
	LSMeansPctDi C	-2 89<->2 14	
crude ibre (% DB)	LSMean	30 1	29 73
	95% C	26 32<->33 88	25 95<->33 51
	Range	(26 9-35 8)	(26-34 9)
	N	24	24
	Pr(>F)	0 6587	
	Conditional R2	0 844	
	LSMeansDi	0 373	
	LSMeansDi C	-2 08<->2 83	
	LSMeansPctDi	1 25	
	LSMeansPctDi C	-7 01<->9 52	
calcium (% DB)	LSMean	0 3765	0 3643
	95% C	0 182<->0 5709	0 1698<->0 5587
	Range	(0 175-0 706)	(0 179-0 543)
	N	24	24
	Pr(>F)	0 7033	
	Conditional R2	0 956	
	LSMeansDi	0 0122	
	LSMeansDi C	-0 0363<->0 0608	
	LSMeansPctDi	3 35	
	LSMeansPctDi C	-9 96<->16 7	
crude protein (% DB)	LSMean	6 129	6 118
	95% C	2 904<->9 354	2 893<->9 343
	Range	(3 16-11 3)	(2 85-10 2)
	N	24	24
	Pr(>F)	0 974	
	Conditional R2	0 931	
	LSMeansDi	0 0114	
	LSMeansDi C	-0 966<->0 989	
	LSMeansPctDi	0 186	
	LSMeansPctDi C	-15 8<->16 2	
ash (% DB)	LSMean	25 73	25 26
	95% C	22 29<->29 17	21 82<->28 7
	Range	(21 1-30 4)	(20 6-30 2)
	N	24	24
	Pr(>F)	0 4632	
	Conditional R2	0 831	
	LSMeansDi	0 475	
	LSMeansDi C	-1 03<->1 98	
	LSMeansPctDi	1 88	
	LSMeansPctDi C	-4 06<->7 82	
acid detergent ibre (% DB)	LSMean	52 6	51 59
	95% C	44 45<->60 75	43 44<->59 74
	Range	(46 1-58)	(44 7-59 3)
	N	24	24
	Pr(>F)	0 522	
	Conditional R2	0 836	
	LSMeansDi	1 01	
	LSMeansDi C	-3 04<->5 05	
	LSMeansPctDi	1 95	
	LSMeansPctDi C	-5 89<->9 8	
neutral detergent ibre (% DB)	LSMean	62 53	62 21
	95% C	56 04<->69 01	55 73<->68 69
	Range	(56 3-68 9)	(50 2-69 7)
	N	24	24
	Pr(>F)	0 8184	
	Conditional R2	0 735	
	LSMeansDi	0 316	
	LSMeansDi C	-2 96<->3 59	
	LSMeansPctDi	0 507	
	LSMeansPctDi C	-4 75<->5 77	
phosphorus (% DB)	LSMean	0 1521	0 1469
	95% C	0 0404<->0 2638	0 0352<->0 2586
	Range	(0 082-0 361)	(0 0811-0 347)
	N	24	24
	Pr(>F)	0 6165	
	Conditional R2	0 933	
	LSMeansDi	0 0052	
	LSMeansDi C	-0 0072<->0 0176	
	LSMeansPctDi	3 54	
	LSMeansPctDi C	-4 9<->12	

Table 51. Multi-year combined-sites analysis for grain proximates and fibre

Analyte	Statistic	GR2E	RC82
moisture (% FW)	LSMean	12 26	12 32
	95% C	5 609<->18 9	5 669<->18 96
	Range	(11 1-13 8)	(10 9-13 6)
	N	24	24
	Pr(>F)	0 8023	
	Conditional R2	0 903	
	LSMeansDiff	-0 06	
	LSMeansDiffC	-1 01<->0 895	
	LSMeansPctDiff	-0 487	
	LSMeansPctDiffC	-8 24<->7 27	
crude fat (% DB)	LSMean	1 419	1 34
	95% C	0 4929<->2 346	0 4131<->2 266
	Range	(0 838-2 16)	(0 555-1 98)
	N	24	24
	Pr(>F)	0 7113	
	Conditional R2	0 502	
	LSMeansDiff	0 0798	
	LSMeansDiffC	-1 94<->2 1	
	LSMeansPctDiff	5 96	
	LSMeansPctDiffC	-145<->157	
carbohydrates (% DB)	LSMean	84 58	84 38
	95% C	73 73<->95 44	73 53<->95 24
	Range	(81-86 9)	(81 1-86 4)
	N	24	24
	Pr(>F)	0 5595	
	Conditional R2	0 936	
	LSMeansDiff	0 198	
	LSMeansDiffC	-2 69<->3 08	
	LSMeansPctDiff	0 235	
	LSMeansPctDiffC	-3 18<->3 65	
starch (% DB)	LSMean	59 52	61 05
	95% C	11 24<->107 8	12 76<->109 3
	Range	(32 8-71 5)	(28 1-73 9)
	N	24	24
	Pr(>F)	0 6888	
	Conditional R2	0 538	
	LSMeansDiff	-1 52	
	LSMeansDiffC	-8 15<->5 1	
	LSMeansPctDiff	-2 49	
	LSMeansPctDiffC	-13 3<->8 36	
crude fibre (% DB)	LSMean	11 96	11 1
	95% C	10 68<->13 25	9 812<->12 39
	Range	(10 1-14 6)	(10 1-12 3)
	N	24	24
	Pr(>F)	0 2133	
	Conditional R2	0 569	
	LSMeansDiff	0 863	
	LSMeansDiffC	-2 3<->4 03	
	LSMeansPctDiff	7 77	
	LSMeansPctDiffC	-20 7<->36 3	
total dietary fibre (% DB)	LSMean	16 96	16 9
	95% C	9 736<->24 18	9 683<->24 12
	Range	(12 8-20 3)	(11 4-21 4)
	N	24	24
	Pr(>F)	0 9549	
	Conditional R2	0 398	
	LSMeansDiff	0 053	
	LSMeansDiffC	-9 03<->9 13	
	LSMeansPctDiff	0 314	
	LSMeansPctDiffC	-53 4<->54	
crude protein (% DB)	LSMean	8 104	8 256
	95% C	4 433<->11 78	4 584<->11 93
	Range	(6 07-11 2)	(6 03-11 4)
	N	24	24
	Pr(>F)	0 5445	
	Conditional R2	0 916	
	LSMeansDiff	-0 152	
	LSMeansDiffC	-0 425<->0 122	
	LSMeansPctDiff	-1 84	
	LSMeansPctDiffC	-5 15<->1 48	
ash (% DB)	LSMean	5 894	6 021
	95% C	1 537<->10 25	1 663<->10 38
	Range	(4 95-7 17)	(5-7 06)
	N	24	24
	Pr(>F)	0 3739	
	Conditional R2	0 869	
	LSMeansDiff	-0 126	
	LSMeansDiffC	-0 423<->0 17	
	LSMeansPctDiff	-2 1	
	LSMeansPctDiffC	-7 02<->2 83	
amylose (% DB)	LSMean	12 87	12 81
	95% C	-8 455<->34 19	-8 506<->34 14
	Range	(7 31-18 6)	(6 76-18 6)
	N	24	24
	Pr(>F)	0 9548	
	Conditional R2	0 784	
	LSMeansDiff	0 051	
	LSMeansDiffC	-2<->2 1	
	LSMeansPctDiff	0 398	
	LSMeansPctDiffC	-15 6<->16 4	
acid detergent fibre (% DB)	LSMean	18 53	17 65
	95% C	16 28<->20 78	15 4<->19 89
	Range	(15 7-21 7)	(15 6-18 8)
	N	24	24
	Pr(>F)	0 3522	
	Conditional R2	0 611	
	LSMeansDiff	0 882	
	LSMeansDiffC	-4 74<->6 5	
	LSMeansPctDiff	5	
	LSMeansPctDiffC	-26 8<->36 8	
neutral detergent fibre (% DB)	LSMean	22 1	20 6
	95% C	-4 307<->48 51	-5 811<->47
	Range	(17 5-35 5)	(16 2-32 8)
	N	24	24
	Pr(>F)	0 4774	
	Conditional R2	0 615	
	LSMeansDiff	1 5	
	LSMeansDiffC	-1 67<->4 68	
	LSMeansPctDiff	7 3	
	LSMeansPctDiffC	-8 12<->22 7	

Table 52. Multi-year combined-sites analysis for grain minerals

Analyte	Statistic	GR2E	RC82
calcium (mg/100 g DB)	LSMean	22 5	21 39
	95% C	4 522<->40 49	3 41<->39 37
	Range	(14 2-35)	(15 3-29 8)
	N	24	24
	Pr(>F)	0 5535	
	Conditional R2	0 832	
	LSMeansDi	1 11	
	LSMeansDi C	-2 38<->4 6	
	LSMeansPctDi	5 2	
	LSMeansPctDi C	-11 1<->21 5	
iron (mg/100 g DB)	LSMean	3 958	4 565
	95% C	-1 65<->9 566	-1 043<->10 17
	Range	(2 37-10 6)	(2 58-9 08)
	N	24	24
	Pr(>F)	0 5136	
	Conditional R2	0 313	
	LSMeansDi	-0 607	
	LSMeansDi C	-1 64<->0 43	
	LSMeansPctDi	-13 3	
	LSMeansPctDi C	-36<->9 43	
manganese (mg/100 g DB)	LSMean	6 606	6 472
	95% C	5 421<->7 791	5 287<->7 657
	Range	(4 33-8 39)	(4 85-7 78)
	N	24	24
	Pr(>F)	0 6741	
	Conditional R2	0 569	
	LSMeansDi	0 134	
	LSMeansDi C	-0 336<->0 603	
	LSMeansPctDi	2 07	
	LSMeansPctDi C	-5 19<->9 32	
potassium (mg/100 g DB)	LSMean	345 9	338 9
	95% C	-40 97<->732 7	-47 88<->725 8
	Range	(236-597)	(222-472)
	N	24	24
	Pr(>F)	0 6656	
	Conditional R2	0 867	
	LSMeansDi	6 91	
	LSMeansDi C	-32 9<->46 7	
	LSMeansPctDi	2 04	
	LSMeansPctDi C	-9 71<->13 8	
zinc (mg/100 g DB)	LSMean	2 306	2 19
	95% C	1 866<->2 745	1 751<->2 63
	Range	(1 63-3 21)	(1 73-2 78)
	N	24	24
	Pr(>F)	0 5689	
	Conditional R2	0 637	
	LSMeansDi	0 115	
	LSMeansDi C	-0 706<->0 937	
	LSMeansPctDi	5 26	
	LSMeansPctDi C	-32 2<->42 8	
copper (mg/100 g DB)	LSMean	0 3913	0 3648
	95% C	0 227<->0 5557	0 2005<->0 5292
	Range	(0 183-0 68)	(0 215-0 506)
	N	24	24
	Pr(>F)	0 5441	
	Conditional R2	0 471	
	LSMeansDi	0 0265	
	LSMeansDi C	-0 0212<->0 0742	
	LSMeansPctDi	7 26	
	LSMeansPctDi C	-5 81<->20 3	
magnesium (mg/100 g DB)	LSMean	130 5	133 4
	95% C	44 11<->216 8	47 07<->219 8
	Range	(87 5-185)	(102-157)
	N	24	24
	Pr(>F)	0 6312	
	Conditional R2	0 595	
	LSMeansDi	-2 97	
	LSMeansDi C	-15 4<->9 49	
	LSMeansPctDi	-2 22	
	LSMeansPctDi C	-11 6<->7 11	
phosphorus (mg/100 g DB)	LSMean	327 1	329 4
	95% C	187 4<->466 8	189 6<->469 1
	Range	(211-461)	(241-383)
	N	24	24
	Pr(>F)	0 8698	
	Conditional R2	0 589	
	LSMeansDi	-2 25	
	LSMeansDi C	-36 5<->32	
	LSMeansPctDi	-0 683	
	LSMeansPctDi C	-11 1<->9 72	
sodium (mg/100 g DB)	LSMean	1 502	1 304
	95% C	-2 815<->5 82	-3 014<->5 621
	Range	(0 563-3 81)	(0 543-3 07)
	N	24	24
	Pr(>F)	0 4668	
	Conditional R2	0 911	
	LSMeansDi	0 199	
	LSMeansDi C	-1 22<->1 61	
	LSMeansPctDi	15 2	
	LSMeansPctDi C	-93 3<->124	

Table 53. Multi-year combined-sites analysis for grain fatty acids

Analyte	Statistic	GR2E	RC82
caprylic (% otal FA)	Mean	NaN	NaN
	Range	(n -- n)	(n -- n)
	N	0	0
lauric (% otal FA)	Mean	NaN	NaN
	Range	(n -- n)	(n -- n)
	N	0	0
pentadecanoic (% otal FA)	Mean	NaN	NaN
	Range	(n -- n)	(n -- n)
	N	0	0
palmitic (% otal FA)	LSMean	19 54	18 45
	95% C	19 06<->20 02	17 97<->18 93
	Range	(18 8-20 4)	(15 8-19 1)
	N	24	24
	Pr(>F)	0 2216	
	Conditional R2	0 711	
	LSMeansDi	1 09	
	LSMeansDi C	0 631<->1 54	
	LSMeansPctDi	5 89	
	LSMeansPctDi C	3 42<->8 36	
heptadecanoic (% otal FA)	Mean	NaN	NaN
	Range	(n -- n)	(n -- n)
	N	0	0
stearic (% otal FA)	LSMean	2 246	2 109
	95% C	2 056<->2 436	1 92<->2 299
	Range	(1 95-2 78)	(1 71-2 68)
	N	24	24
	Pr(>F)	0 04872	
	Conditional R2	0 956	
	LSMeansDi	0 137	
	LSMeansDi C	0 003<->0 27	
	LSMeansPctDi	6 48	
	LSMeansPctDi C	0 142<->12 8	
linoleic (% otal FA)	LSMean	33 45	34 03
	95% C	30 2<->36 7	30 78<->37 28
	Range	(32 4-34 4)	(30 1-35 4)
	N	24	24
	Pr(>F)	0 3944	
	Conditional R2	0 704	
	LSMeansDi	-0 586	
	LSMeansDi C	-1 25<->0 0786	
	LSMeansPctDi	-1 72	
	LSMeansPctDi C	-3 67<->0 231	
arachidic (% otal FA)	LSMean	0 8623	0 8873
	95% C	0 7898<->0 9349	0 8147<->0 9599
	Range	(0 739-0 989)	(0 743-1 02)
	N	24	24
	Pr(>F)	0 1967	
	Conditional R2	0 93	
	LSMeansDi	-0 025	
	LSMeansDi C	-0 0422<->-0 0078	
	LSMeansPctDi	-2 82	
	LSMeansPctDi C	-4 76<->-0 879	
eicosadienoic (% otal FA)	Mean	0 126	NaN
	Range	(0 126-0 126)	(n -- n)
	N	1	0
arachidonic (% otal FA)	Mean	NaN	NaN
	Range	(n -- n)	(n -- n)
	N	0	0
erucic (% otal FA)	Mean	NaN	NaN
	Range	(n -- n)	(n -- n)
	N	0	0
lignoceric (% otal FA)	LSMean	0 9311	1 002
	95% C	-0 4023<->2 265	-0 3317<->2 335
	Range	(0 687-1 25)	(0 805-1 19)
	N	24	24
	Pr(>F)	0 1159	
	Conditional R2	0 921	
	LSMeansDi	-0 0707	
	LSMeansDi C	-0 139<->-0 0027	
	LSMeansPctDi	-7 06	
	LSMeansPctDi C	-13 8<->-0 27	

Analyte	Statistic	GR2E	RC82
capric (% otal FA)	Mean	NaN	NaN
	Range	(n -- n)	(n -- n)
	N	0	0
myristic (% otal FA)	LSMean	0 438	0 3927
	95% C	0 0906<->0 7854	0 0453<->0 7401
	Range	(0 346-0 518)	(0 251-0 492)
	N	24	24
	Pr(>F)	0 1983	
	Conditional R2	0 948	
	LSMeansDi	0 0453	
	LSMeansDi C	-0 122<->0 213	
	LSMeansPctDi	11 5	
	LSMeansPctDi C	-31 1<->54 2	
palmitoleic (% otal FA)	LSMean	0 1934	0 1908
	95% C	0 0319<->0 355	0 0293<->0 3523
	Range	(0 16-0 218)	(0 153-0 217)
	N	24	24
	Pr(>F)	0 3614	
	Conditional R2	0 958	
	LSMeansDi	0 0026	
	LSMeansDi C	-0 0027<->0 0079	
	LSMeansPctDi	1 36	
	LSMeansPctDi C	-1 42<->4 14	
oleic (% otal FA)	LSMean	39 7	40 23
	95% C	38 69<->40 72	39 22<->41 25
	Range	(38 4-41 3)	(38 4-45 9)
	N	24	24
	Pr(>F)	0 4866	
	Conditional R2	0 648	
	LSMeansDi	-0 529	
	LSMeansDi C	-1 42<->0 36	
	LSMeansPctDi	-1 31	
	LSMeansPctDi C	-3 52<->0 895	
linolenic (% otal FA)	LSMean	1 626	1 641
	95% C	-1 259<->4 51	-1 244<->4 525
	Range	(1 34-2 31)	(1 27-2 7)
	N	24	24
	Pr(>F)	0 8871	
	Conditional R2	0 853	
	LSMeansDi	-0 0149	
	LSMeansDi C	-0 984<->0 954	
	LSMeansPctDi	-0 908	
	LSMeansPctDi C	-60<->58 1	
eicosenoic (% otal FA)	LSMean	0 4755	0 5148
	95% C	0 3909<->0 56	0 4302<->0 5994
	Range	(0 405-0 538)	(0 442-0 773)
	N	24	24
	Pr(>F)	0 2857	
	Conditional R2	0 484	
	LSMeansDi	-0 0393	
	LSMeansDi C	-0 105<->0 0268	
	LSMeansPctDi	-7 63	
	LSMeansPctDi C	-20 5<->5 21	
eicosatrienoic (% otal FA)	Mean	NaN	NaN
	Range	(n -- n)	(n -- n)
	N	0	0
behenic (% otal FA)	LSMean	0 5083	0 5427
	95% C	0 1223<->0 8942	0 1567<->0 9287
	Range	(0 407-0 585)	(0 478-0 61)
	N	24	24
	Pr(>F)	0 142	
	Conditional R2	0 837	
	LSMeansDi	-0 0344	
	LSMeansDi C	-0 0573<->-0 0116	
	LSMeansPctDi	-6 34	
	LSMeansPctDi C	-10 6<->-2 14	
nervonic (% otal FA)	Mean	0 271	NaN
	Range	(0 14-0 403)	(n -- n)
	N	2	0

Table 54. Multi-year combined-sites analysis for grain amino acids

Analyte	Statistic	GR2E	RC82
methionine (mg/100 g DB)	LSMean	169.2	166.2
	95% CI	120<->218.5	117<->215.5
	Range	(124-228)	(127-215)
	N	24	24
	Pr(>F)	0.6019	
	Conditional R2	0.835	
	LSMeansDiff	3	
	LSMeansDiffCI	-4.91<->10.9	
	LSMeansPctDiff	1.8	
	LSMeansPctDiffCI	-2.95<->6.56	
lysine (mg/100 g DB)	LSMean	298.6	294.4
	95% CI	62.91<->534.3	59.32<->529.4
	Range	(216-443)	(211-434)
	N	24	23
	Pr(>F)	0.7111	
	Conditional R2	0.871	
	LSMeansDiff	4.25	
	LSMeansDiffCI	-10.7<->19.2	
	LSMeansPctDiff	1.44	
	LSMeansPctDiffCI	-3.62<->6.51	
arginine (mg/100 g DB)	LSMean	563.8	564
	95% CI	392.9<->734.7	393.1<->734.9
	Range	(409-737)	(408-782)
	N	24	24
	Pr(>F)	0.988	
	Conditional R2	0.86	
	LSMeansDiff	-0.25	
	LSMeansDiffCI	-41.8<->41.3	
	LSMeansPctDiff	-0.0443	
	LSMeansPctDiffCI	-7.41<->7.32	
histidine (mg/100 g DB)	LSMean	211.8	214.5
	95% CI	155.4<->268.1	158.2<->270.9
	Range	(162-276)	(157-281)
	N	24	24
	Pr(>F)	0.6593	
	Conditional R2	0.841	
	LSMeansDiff	-2.79	
	LSMeansDiffCI	-10.7<->5.11	
	LSMeansPctDiff	-1.3	
	LSMeansPctDiffCI	-4.98<->2.38	
leucine (mg/100 g DB)	LSMean	643.6	650.5
	95% CI	348.2<->939	355.1<->945.9
	Range	(464-886)	(462-921)
	N	24	24
	Pr(>F)	0.743	
	Conditional R2	0.876	
	LSMeansDiff	-6.88	
	LSMeansDiffCI	-34.8<->21	
	LSMeansPctDiff	-1.06	
	LSMeansPctDiffCI	-5.34<->3.23	
phenylalanine (mg/100 g DB)	LSMean	439.8	443.5
	95% CI	300.8<->578.8	304.5<->582.4
	Range	(324-593)	(317-622)
	N	24	24
	Pr(>F)	0.7992	
	Conditional R2	0.835	
	LSMeansDiff	-3.67	
	LSMeansDiffCI	-34.9<->27.6	
	LSMeansPctDiff	-0.827	
	LSMeansPctDiffCI	-7.87<->6.22	
alanine (mg/100 g DB)	LSMean	455.4	459.8
	95% CI	201.9<->708.8	206.4<->713.2
	Range	(329-625)	(331-628)
	N	24	24
	Pr(>F)	0.7404	
	Conditional R2	0.902	
	LSMeansDiff	-4.42	
	LSMeansDiffCI	-22.1<->13.3	
	LSMeansPctDiff	-0.961	
	LSMeansPctDiffCI	-4.81<->2.89	
glutamic acid (mg/100 g DB)	LSMean	1354	1360
	95% CI	426.2<->2282	432.4<->2288
	Range	(942-1980)	(890-1990)
	N	24	24
	Pr(>F)	0.8975	
	Conditional R2	0.887	
	LSMeansDiff	-6.17	
	LSMeansDiffCI	-74.2<->61.8	
	LSMeansPctDiff	-0.453	
	LSMeansPctDiffCI	-5.45<->4.55	
serine (mg/100 g DB)	LSMean	400.6	400.5
	95% CI	222.6<->578.6	222.5<->578.6
	Range	(296-540)	(231-556)
	N	24	24
	Pr(>F)	0.9975	
	Conditional R2	0.839	
	LSMeansDiff	0.0417	
	LSMeansDiffCI	-19.1<->19.2	
	LSMeansPctDiff	0.0104	
	LSMeansPctDiffCI	-4.76<->4.78	
cystine (mg/100 g DB)	LSMean	156.3	155
	95% CI	71.46<->241.1	70.17<->239.8
	Range	(117-214)	(113-198)
	N	24	24
	Pr(>F)	0.7671	
	Conditional R2	0.886	
	LSMeansDiff	1.29	
	LSMeansDiffCI	-4.59<->7.18	
	LSMeansPctDiff	0.833	
	LSMeansPctDiffCI	-2.96<->4.63	
tryptophan (mg/100 g DB)	LSMean	73.36	74.93
	95% CI	58.09<->88.62	59.66<->90.19
	Range	(44.2-107)	(50.9-101)
	N	24	24
	Pr(>F)	0.7954	
	Conditional R2	0.536	
	LSMeansDiff	-1.57	
	LSMeansDiffCI	-35.9<->32.8	
	LSMeansPctDiff	-2.1	
	LSMeansPctDiffCI	-48<->43.8	
isoleucine (mg/100 g DB)	LSMean	329.2	331.9
	95% CI	180.9<->477.4	183.7<->480.1
	Range	(241-446)	(237-464)
	N	24	24
	Pr(>F)	0.7871	
	Conditional R2	0.877	
	LSMeansDiff	-2.75	
	LSMeansDiffCI	-16.6<->11.1	
	LSMeansPctDiff	-0.829	
	LSMeansPctDiffCI	-5.01<->3.36	
valine (mg/100 g DB)	LSMean	468.2	474
	95% CI	285.9<->650.6	291.6<->656.4
	Range	(341-627)	(338-653)
	N	24	24
	Pr(>F)	0.6987	
	Conditional R2	0.872	
	LSMeansDiff	-5.75	
	LSMeansDiffCI	-24.9<->13.4	
	LSMeansPctDiff	-1.21	
	LSMeansPctDiffCI	-5.25<->2.83	
threonine (mg/100 g DB)	LSMean	308.8	308.2
	95% CI	192.1<->425.5	191.5<->424.9
	Range	(235-400)	(217-409)
	N	24	24
	Pr(>F)	0.9416	
	Conditional R2	0.869	
	LSMeansDiff	0.583	
	LSMeansDiffCI	-10.8<->12	
	LSMeansPctDiff	0.189	
	LSMeansPctDiffCI	-3.51<->3.89	
glycine (mg/100 g DB)	LSMean	388.7	392.5
	95% CI	278.9<->498.5	282.7<->502.2
	Range	(293-495)	(292-511)
	N	24	24
	Pr(>F)	0.7161	
	Conditional R2	0.859	
	LSMeansDiff	-3.79	
	LSMeansDiffCI	-17.4<->9.82	
	LSMeansPctDiff	-0.966	
	LSMeansPctDiffCI	-4.44<->2.5	
aspartic acid (mg/100 g DB)	LSMean	708.4	709.5
	95% CI	173.2<->1244	174.4<->1245
	Range	(493-1010)	(497-994)
	N	24	24
	Pr(>F)	0.96	
	Conditional R2	0.902	
	LSMeansDiff	-1.12	
	LSMeansDiffCI	-33.3<->31	
	LSMeansPctDiff	-0.159	
	LSMeansPctDiffCI	-4.69<->4.37	
proline (mg/100 g DB)	LSMean	376.4	380.9
	95% CI	211.8<->541	216.3<->545.5
	Range	(276-510)	(278-521)
	N	24	24
	Pr(>F)	0.6777	
	Conditional R2	0.9	
	LSMeansDiff	-4.46	
	LSMeansDiffCI	-18.1<->9.14	
	LSMeansPctDiff	-1.17	
	LSMeansPctDiffCI	-4.74<->2.4	
tyrosine (mg/100 g DB)	LSMean	213.5	206.7
	95% CI	136.9<->290	130.2<->283.2
	Range	(158-282)	(133-291)
	N	24	24
	Pr(>F)	0.5314	
	Conditional R2	0.812	
	LSMeansDiff	6.75	
	LSMeansDiffCI	-21.4<->34.9	
	LSMeansPctDiff	3.27	
	LSMeansPctDiffCI	-10.3<->16.9	

Table 55. Multi-year combined-sites analysis for grain vitamins

Analyte	Statistic	GR2E	RC82
beta-carotene (mg/kg DB)	Mean	1 26	0 0651
	Range	(0 504-2 35)	(0 0585-0 0701)
	N	24	5
riboflavin (mg/kg DB)	Mean	NaN	NaN
	Range	(nf-- nf)	(nf-- nf)
	N	0	0
folic acid (mg/kg DB)	LSMean	0 9116	0 8827
	95% C	0 5007<->1 323	0 4718<->1 294
	Range	(0 562-2 56)	(0 393-1 54)
	N	24	24
	Pr(>F)	0 881	
	Conditional R2	0 249	
	LSMeansDiff	0 0289	
	LSMeansDiffC	-0 691<->0 748	
	LSMeansPctDiff	3 27	
	LSMeansPctDiffC	-78 2<->84 8	
pantothenic acid (mg/kg DB)	LSMean	9 152	9 133
	95% C	7 312<->10 99	7 293<->10 97
	Range	(7 31-11 9)	(7 22-11 4)
	N	24	24
	Pr(>F)	0 9291	
	Conditional R2	0 905	
	LSMeansDiff	0 0192	
	LSMeansDiffC	-0 564<->0 603	
	LSMeansPctDiff	0 21	
	LSMeansPctDiffC	-6 18<->6 6	
alpha-tocopherol (mg/kg DB)	LSMean	2 975	2 746
	95% C	0 1472<->5 803	-0 0815<->5 574
	Range	(2 47-3 87)	(2 1-3 5)
	N	24	24
	Pr(>F)	0 3358	
	Conditional R2	0 643	
	LSMeansDiff	0 229	
	LSMeansDiffC	0 0619<->0 396	
	LSMeansPctDiff	8 33	
	LSMeansPctDiffC	2 25<->14 4	
thiamine (mg/kg DB)	LSMean	3 127	3 081
	95% C	2 607<->3 647	2 561<->3 601
	Range	(2 33-3 77)	(2 35-3 81)
	N	24	24
	Pr(>F)	0 6344	
	Conditional R2	0 82	
	LSMeansDiff	0 0465	
	LSMeansDiffC	-0 0927<->0 186	
	LSMeansPctDiff	1 51	
	LSMeansPctDiffC	-3 01<->6 02	
niacin (mg/kg DB)	LSMean	36 9	32 63
	95% C	-55 54<->129 3	-59 81<->125 1
	Range	(23 4-58 3)	(20 2-48 8)
	N	24	24
	Pr(>F)	0 4171	
	Conditional R2	0 959	
	LSMeansDiff	4 27	
	LSMeansDiffC	-0 78<->9 32	
	LSMeansPctDiff	13 1	
	LSMeansPctDiffC	-2 39<->28 5	
pyridoxine (mg/kg DB)	LSMean	2 722	2 75
	95% C	1 994<->3 45	2 022<->3 477
	Range	(2 22-3 3)	(2 1-5 42)
	N	24	24
	Pr(>F)	0 9185	
	Conditional R2	0 497	
	LSMeansDiff	-0 0279	
	LSMeansDiffC	-2 63<->2 57	
	LSMeansPctDiff	-1 01	
	LSMeansPctDiffC	-95 6<->93 5	

Table 56. Multi-year combined-sites analysis for grain secondary metabolites

Analyte	Statistic	GR2E	RC82
phytic acid (% DB)	LSMean	0 8609	0 8814
	95% C	0 7459<->0 9759	0 7664<->0 9963
	Range	(0 582-1 1)	(0 605-1 23)
	N	24	24
	Pr(>F)	0 6218	
	Conditional R2	0 706	
	LSMeansDiff	-0 0204	
	LSMeansDiffC	-0 0703<->0 0294	
	LSMeansPctDiff	-2 31	
	LSMeansPctDiffC	-7 98<->3 34	
trypsin inhibitor (T U/mg DB)	LSMean	0 9239	0 9962
	95% C	-4 793<->6 64	-4 72<->6 713
	Range	(0 28-1 71)	(0 0322-4 17)
	N	24	24
	Pr(>F)	0 8282	
	Conditional R2	0 826	
	LSMeansDiff	-0 0723	
	LSMeansDiffC	-4 1<->3 95	
	LSMeansPctDiff	-7 26	
	LSMeansPctDiffC	-411<->397	

F. Test Substance Certificates of Analysis

F.1. ZmPSY1 Protein

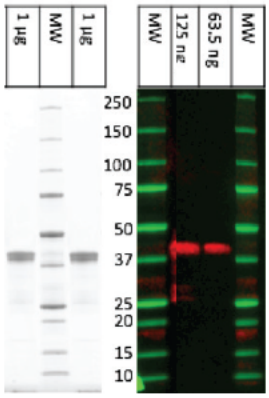
Product Information

Name	*Zm*PSY1 (*Zea mays* phytoene synthase)
Lot Number	M20452-05
Date of Manufacture	14 November 2015
Concentration	2.39 mg/ml (Amino Acid Analysis)
Formulation Buffer	50 mM Tris-HCl pH 7.6, 600 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP), 0.035% lauryldimethylamine-N-oxide (LDAO)
Storage	–80°C
Source	Rosetta™-2 DE3 (*Escherichia coli*) transformed with plasmid pCOLD1-*Zm*PSY1
Manufacturer	Aldevron, 5602 Research Park Blvd, Madison, WI 53719

Test Results

Characteristic	Method	Analysis Date	Result	Testing Facility
Concentration	Bradford	7 December 2015	2.97 mg/ml	Aldevron
Purity	SDS-PAGE	7 December 2015	>95%	Aldevron
Purity	RP-HPLC	7 December 2015	93.7%	Aldevron
Identity/ Concentration	Amino Acid Analysis	25 November 2015	Good agreement between theoretical and actual amino acid composition; concentration 2.39 mg/ml	Alphalyse Inc.
Identity/ Molecular Weight	SDS-PAGE	7 December 2015	*ca.* 42 kDa	Aldevron
Identity/ Sequence	MALDI-TOF Tryptic peptide mapping	1 December 2015	PASS; Sequence coverage 48%	Alphalyse Inc.
Identity/ Sequence	N-terminal sequencing	26 November 2015	More than one N-terminal sequence was observed in the sample	Alphalyse Inc.
Identity/ Activity	HPLC determination of phytoene production	22 December 2015	*ca.* 28.4 pmol μg^{-1} min^{-1} of 15-*cis*-phytoene from DMAPP and IPP, in the presence of active *A. thaliana* GGPP synthase	Albert-Ludwigs Universität Freiburg

A



B

```
  1 MNHKVHHHHH HIEGRHMELG TLEGSHPWEA GRPSPAVYSS LPVNPAGEAV
 51 VSSEQKVYDV VLKQAALLKR QLRTPVLDAR PQDMDMPRNG LKEAYDRCGE
101 ICEEYAKTFY LGTMLMTEER RRAIWAIYVW CRRTDELVDG PNANYITPTA
151 LDRWEKRLED LFTGRPYDML DAALSDTISR FPIDIQPFRD MIEGMRSDLR
201 KTRYNNFDEL YMYCYYVAGT VGLMSVPVMG IATESKATTE SVYSAALALG
251 IANQLTNILR DVGEDARRGR IYLPQDELAQ AGLSDEDIFK GVVTNRWRNF
301 MKRQIKRARM FFEEAERGVT ELSQASRWPV WASLLLYRQI LDEIEANDYN
351 NFTKRAYVGK GKKLLALPVA YGKSLLLPCS LRNGQT
```

Panel A: SDS PAGE and anti-His Tag western blot analysis of ZmPSY1 final product. SDS PAGE conditions: 4-15% BioRad TGX gel, BioRad Precision Plus Protein Markers, 300 v for 18 min, SYPRO-Orange stain. Western blot: Rockland anti-6xHis 1°, Rockland antimouse-HRP 2°, SuperSignal West Pico Chemiluminescent Substrate. Purity is >95%

Panel B: Amino acid sequence of the microbial-expressed ZmPSY1 protein. The N-terminal leader sequence containing the 6xHis-tag and the factor Xa cleavage site (IEGR) is indicated with the bar. The protein is 386 amino acids in length with a calculated molecular weight of 44.2 kDa.

Original or exact copies of all raw data and pertinent information are archived at the International Rice Research Institute, Pili Drive, UPLB, Los Baños, 4031 Laguna, Philippines.

[redacted]

January 5, 2016

Date

F.2. CRTI Protein

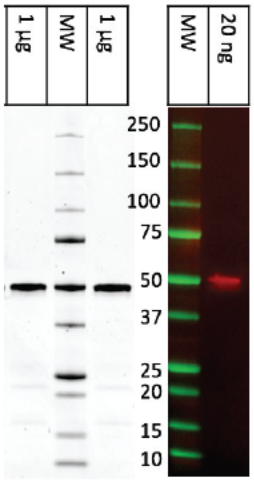
Product Information

Name	CRTI (*Pantoea ananatis* phytoene desaturase)
Lot Number	M20454-02
Date of Manufacture	9 October 2015
Concentration	1.10 mg/ml (Amino Acid Analysis)
Formulation Buffer	50 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP)
Storage	–80°C
Source	JM109 (*Escherichia coli*) transformed with plasmid pCRTI-His (pSYN13740)
Manufacturer	Aldevron, 5602 Research Park Blvd, Madison, WI 53719

Test Results

Characteristic	Method	Analysis Date	Result	Testing Facility
Concentration	Bradford	18 November 2015	3.81 mg/ml	Aldevron
Purity	SDS-PAGE	18 November 2015	97.8%	Aldevron
Purity	RP-HPLC	18 November 2015	91.6%	Aldevron
Identity/ Concentration	Amino Acid Analysis	6 November 2015	Good agreement between theoretical and actual amino acid composition; concentration 1.10 mg/ml	Alphalyse Inc.
Identity/ Molecular Weight	SDS-PAGE	18 November 2015	Apparent MW of *ca.* 50 kDa	Aldevron
Identity/ Sequence	MALDI-TOF Tryptic peptide mapping	6 November 2015	Sequence coverage 39% allowed unambiguous identification as CRTI_PANAN (UniProt entry P21685)	Alphalyse Inc.
Identity/ Sequence	N-terminal sequencing	4 November 2015	The predominant amino terminal sequence in the sample was MKPTTVIGAG, matching the CRTI N-terminus	Alphalyse Inc.
Identity/ Activity	Spectrophotometric assay	21 December 2015	*ca.* 5.40 pmol μg^{-1} min^{-1} of all-*trans*-lycopene produced from liposome-incorporated phytoene	Albert-Ludwigs Universität Freiburg

A



B

```
1    MKPTTVIGAG FGGLALAIRL QAAGIPVLLL EQRDKPGGRA YVYEDQGFTF
51   DAGPTVITDP SAIEELFALA GKQLKEYVEL LPVTPFYRLC WESGKVFNYD
101  NDQTRLEAQI QQFNPRDVEG YRQFLDYSRA VFKEGYLKLG TVPFLSFRDM
151  LRAAPQLAKL QAWRSVYSKV ASYIEDEHLR QAFSFHSLLV GGNPFATSSI
201  YTLIHALERE WGVWFPRGGT GALVQGMIKL FQDLGGEVVL NARVSHMETT
251  GNKIEAVHLE DGRRFLTQAV ASNADVVHTY RDLLSQHPAA VKQSNKLQTK
301  RMSNSLFVLY FGLNHHHDQL AHHTVCFGPR YRELIDEIFN HDGLAEDFSL
351  YLHAPCVTDS SLAPEGCGSY YVLAPVPHLG TANLDWTVEG PKLRDRIFAY
401  LEQHYMPGLR SQLVTHRMFT PFDFRDQLNA YHGSAFSVEP VLTQSAWFRP
451  HNRDKTITNL YLVGAGTHPG AGIPGVIGSA KATAGLMLED LIHHHHHH
```

Panel A: SDS PAGE and anti-His Tag western blot analysis of CRTI final product. SDS PAGE conditions: 4-15% BioRad TGX gel, BioRad Precision Plus Protein Markers, 300 v for 18 min, SYPRO-Orange stain. Western blot: Rockland anti-6xHis 1°, Rockland antimouse-HRP 2°, SuperSignal West Pico Chemiluminescent Substrate. Purity is approximately 98%

Panel B: Amino acid sequence of the microbial-expressed CRTI protein. The C-terminal 6xHis-tag is indicated with the bar. The protein is 501 amino acids in length with a calculated molecular weight of 56.0 kDa.

Original or exact copies of all raw data and pertinent information are archived at the International Rice Research Institute, Pili Drive, UPLB, Los Baños, 4031 Laguna, Philippines.

[redacted]

January 8, 2016

Date

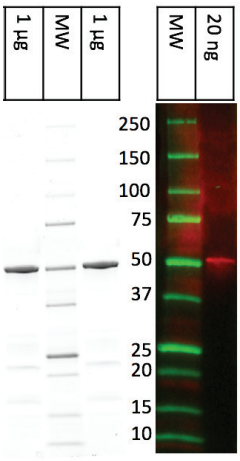
Product Information

Name	CRTI (*Pantoea ananatis* phytoene desaturase)
Lot Number	M20603
Date of Manufacture	25 February 2016
Concentration	4.38 mg/ml (amino acid analysis)
Formulation Buffer	50 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP)
Storage	–80°C
Source	JM109 (*Escherichia coli*) transformed with plasmid pCRTI-His (pSYN13740)
Manufacturer	Aldevron, 5602 Research Park Blvd, Madison, WI 53719

Test Results

Characteristic	Method	Analysis Date	Result	Testing Facility
Concentration	Absorbance 280 nm ($\varepsilon^{0.1\%}_{280nm} = 1.147$)	25 February 2016	4.23 mg/ml	Aldevron
Purity	SDS-PAGE	25 February 2016	> 90%	Aldevron
Purity	RP-HPLC	25 February 2016	84.9%	Aldevron
Identity/ Concentration	Amino Acid Analysis	14 March 2016	Good agreement between theoretical and actual amino acid composition; concentration 4.38 mg/ml	Alphalyse Inc.
Identity/ Molecular Weight	SDS-PAGE	25 February 2016	Apparent MW of *ca.* 50 kDa	Aldevron
Identity/ Sequence	MALDI-TOF Tryptic peptide mapping	9 March 2016	Sequence coverage 55% allowed unambiguous identification as CRTI_PANAN (UniProt entry P21685)	Alphalyse Inc.
Identity/ Sequence	N-terminal sequencing	9 March 2016	The predominant amino terminal sequence in the sample was MKPTTVIGAG, matching the CRTI N-terminus	Alphalyse Inc.
Identity/ Activity	Spectrophotometric assay	22 March 2016	*ca.* 3.47 pmol μg^{-1} min^{-1} of all-*trans*-lycopene produced from liposome-incorporated phytoene	Albert-Ludwigs Universität Freiburg

A



B

```
1     MKPTTVIGAG FGGLALAIRL QAAGIPVLLL EQRDKPGGRA YVYEDQGFTF
51    DAGPTVITDP SAIEELFALA GKQLKEYVEL LPVTPFYRLC WESGKVFNYD
101   NDQTRLEAQI QQFNPRDVEG YRQFLDYSRA VFKEGYLKLG TVPFLSFRDM
151   LRAAPQLAKL QAWRSVYSKV ASYIEDEHLR QAFSFHSLLV GGNPFATSSI
201   YTLIHALERE WGVWFPRGGT GALVQGMIKL FQDLGGEVVL NARVSHMETT
251   GNKIEAVHLE DGRRFLTQAV ASNADVVHTY RDLLSQHPAA VKQSNKLQTK
301   RMSNSLFVLY FGLNHHHDQL AHHTVCFGPR YRELIDEIFN HDGLAEDFSL
351   YLHAPCVTDS SLAPEGCGSY YVLAPVPHLG TANLDWTVEG PKLRDRIFAY
401   LEQHYMPGLR SQLVTHRMFT PFDFRDQLNA YHGSAFSVEP VLTQSAWFRP
451   HNRDKTITNL YLVGAGTHPG AGIPGVIGSA KATAGLMLED LIHHHHHH
```

Panel A: SDS PAGE and anti-His Tag western blot analysis of CRTI final product. SDS PAGE conditions: 4-15% BioRad TGX gel, BioRad Precision Plus Protein Markers, 300 v for 18 min, SYPRO-Orange stain. Western blot: Rockland anti-6xHis 1°, Rockland antimouse-HRP 2°, SuperSignal West Pico Chemiluminescent Substrate. Purity is > 90% by SDS-PAGE.

Panel B: Amino acid sequence of the microbial-expressed CRTI protein. The C-terminal 6xHis-tag is indicated with the bar. The protein is 501 amino acids in length with a calculated molecular weight of 56.0 kDa.

Original or exact copies of all raw data and pertinent information are archived at the International Rice Research Institute, Pili Drive, UPLB, Los Baños, 4031 Laguna, Philippines.

[redacted]

March 23, 2016

Date

IRRI

THE UNIVERSITY OF IOWA

CENTER FOR
BIOCATALYSIS AND BIOPROCESSING

University of Iowa Research Park
2501 Crosspark Road, Suite C100 MTF
Coralville, Iowa 52241-3486
319-335-4900 *Fax* 319-335-4901
www.uiowa.edu/~biocat

Certificate of Analysis

Product: Phosphomannose Isomerase (PMI)

Client: Donald Danforth Plant Science Center
975 N. Warson Road
St. Louis, MO 63132

Test	Test Method	Result
Appearance	Visual Inspection	Pale green solution, free of any particulates
pH	pH at 25°C	7.62 (Buffer: 50 mM Potassium Phosphate, 100mM NaCl)
Product Purity	SDS-PAGE	Major band at 44 kDa with an apparent purity of ≥90% purity by visual examination
Protein Concentration	Bradford	11.9 mg/mL

Storage Condition: Store at -20°C or below.

[redacted]

Date: 20 JUNE 2011

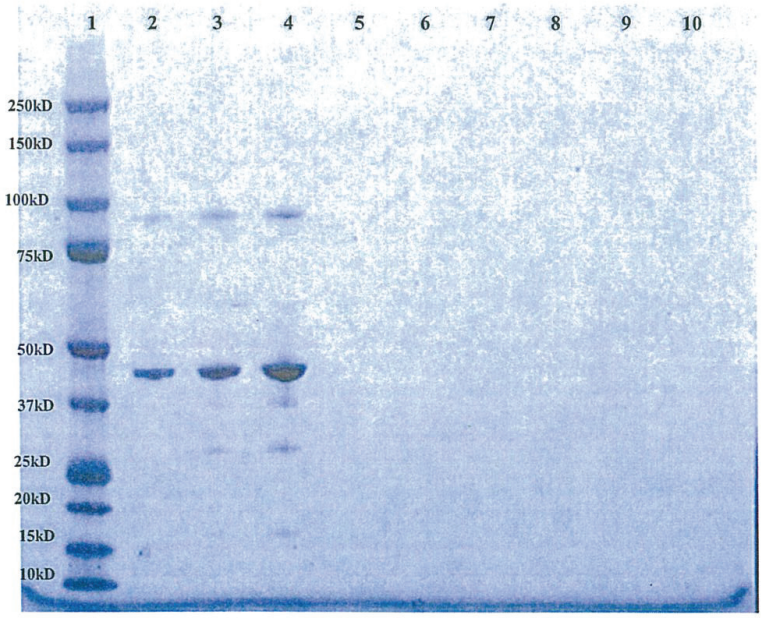
OFFICE OF THE VICE PRESIDENT
FOR RESEARCH

SDS PAGE REPORT

PROJECT: 21038G-PMI Date: 18NOV10
User: AK
Gel Description: PMI final sample
Gel Type: NuPAGE® 4-12% Bis-Tris Gel 1.5mm X 10 well
Gel # 102

Lane	µl	Comments
1	10	Precision Plus Protein Standard (Bio-Rad)
2	5	PMI final sample (0.5 µg)
3	10	PMI final sample (1.0 µg)
4	20	PMI final sample (2.0 µg)
5	---	Empty
6	---	Empty
7	---	Empty
8	---	Empty
9	---	Empty
10	---	Empty

STAIN PROCEDURE: Simply Blue Safe Stain Microwave Method
SAMPLE PREPARATION: 25 µL of sample (4.2 µL of PMI final sample and 20.8 µL of water) was mixed with 25 µL of Laemmli sample buffer containing 5% (v/v) β-mercaptoethanol. The sample was vortexed and boiled for 6 minutes.



[redacted]