Thesis Abstract

Molecular Genetic Studies on the Regulation Mechanism of Phytic Acid Content in Rice Grains

Name: Ishara Perera

Student ID: 4910161001

Introduction

Phytic acid (PA, *myo*-inositol hexa*kis*phosphate) is the most abundant storage form of phosphorus (P) in cereal and legume seeds. PA strongly chelates with metal cations such as calcium (Ca), magnesium (Mg), potassium (K), iron (Fe) and zinc (Zn) and form insoluble salts referred as phytate. PA is usually considered an anti-nutrient as it reduces the bioavailability of important micronutrients inside the human intestine. As the non-ruminant animals lack the digestive enzyme phytase that hydrolyzes P from the PA molecules, they cannot digest PA in food. To overcome P deficiency in non-ruminants, phytase is usually given as a supplement which then results in excess P excretion leading to environmental problems, such as eutrophication. Therefore, reducing the PA content of crops have arisen as a new strategy for improving micronutrient bioavailability. Though several studies have been carried out in the recent past to reduce low PA content in rice, still there is limited understanding on PA metabolism and regulation mechanism of PA accumulation in rice. To date, no previous research has investigated the association of PA content and genotypes using genome-wide mapping and further genetic basis underlying the PA accumulation. I first carried out a literature review to understand the existing information on physiological, molecular and genetic factors and the importance of PA content in plants, animals and the environment. Then the experiments were conducted with the objective of identifying low PA rice from the natural variation of rice and understanding the molecular genetic basis of regulation mechanisms of PA accumulation in rice grains.

Literature Review

Micronutrient deficiency is a serious health issue affecting one-third of the global population. Amongst, Zn deficiency, causes impaired growth, increased susceptibility to infections and increased mortality and affects around one-third of the world's population. Moreover, populations in developing countries are at a high risk of Zn deficiency due to high intake of vegetarian diets. Compared to other cereals, rice is (*Oryza sativa* L.) a poor source of essential micronutrients to

fulfill daily human nutritional requirements, however, rice is the major nutrient source for more than one-half of the global population and the most important crop in Asia. Therefore, improvement in the nutrient content of rice would benefit a large population around the world, especially in developing countries. Thus, Zn biofortification is considered a key solution that seems to be the most sustainable and cost-effective approach to combat global nutritional issue.

Among the several approaches to reduce PA content in rice, classical plant breeding techniques have been used to generate low phytic acid (*lpa*) crops by impairing the synthesis or transport of PA in seed. Several researches have been proceeded to reduce the level of PA in grains by developing *lpa* mutants through γ-irradiation or chemically induced mutagenesis in cereals including rice. However, these mutants observed poor agronomic and yield characteristics which limited their use in plant breeding. Then another approach using transgenic techniques have been used and some transgenic rice lines developed with low PA content, did not show any yield losses. However, the progress in developing *lpa* rice is modest compared to the success made for other cereals.

Several genes responsible for PA accumulation in rice grain have been identified, including eight genes that are expressed in developing rice grains. Efforts to improve micronutrient bioavailability in rice has seemed challenging until the present. Identifying elite accessions through screening the existing germplasm would be important in developing rice using both conventional and marker-assisted breeding programs. Genome-wide association study (GWAS) is an efficient tool that has been widely used to investigate the genetic basis of complex diseases and traits in plants. GWA mapping has been successfully used to identify causal relationship between genetic polymorphism within a species and the phenotypic differences observed between individuals. In rice, GWAS have been used to identify numerous Quantitative Trait Loci (QTL) responsible for morphological and grain elemental components. However, only two QTL have been identified in rice for PA content so far.

I have focused on the key determinants for PA concentration in rice grain, mainly the genetic and environmental factors including the soil and climatic conditions affect the PA content according to the published literature and discussed the possible molecular methods and approaches for manipulating the PA content to increase micronutrient bioavailability in rice grain. In general, I explained the possible strategies toward manipulating grain PA contents, such as the use of *lpa*

mutants advanced molecular genetic techniques such as transgenics and DNA markers. Further, described the possibilities concerning two key approaches, such as manipulating the distribution of grain PA and micronutrient and the PA/micronutrient ratio in the grain.

Identification of Low Phytic Acid and High Zn Bioavailable Rice from 69 accessions of the World Rice Core Collection

First, I identified the natural variation of PA and Zn contents in rice and evaluated the impact of PA on Zn bioavailability. Further, the micronutrient and heavy metal contents of accessions identified as low and high PA rice were determined to document effects of the PA content on other elemental contents owing to their importance in human nutrition and health. Then, a GWAS was conducted to identify the significant genetic polymorphisms controlling PA content in the world rice core collection (WRC).

I have used 69 accessions of WRC, obtained from the National Institute of Agrobiological Sciences (NIAS) Genebank in Tsukuba, Japan (Genebank Project, NARO). The accessions were transplanted at the experimental field in Itakura, Gunma, Japan or in the glasshouse at Toyo University, Iatkura during May 2017. Harvested seeds dried to 13% moisture and stored at room temperature were used in further experiments. These WRC accessions represented three major sub-type groups as *japonica, indica* and *aus.*

First, I determined the PA content of brown rice samples in the WRC using a Phytic Acid Assay Kit (Megazyme International, Ireland) and Zn content was determined by inductively coupled plasma mass spectrometry (ICPMS) method. Then the total daily absorption of Zinc (TAZ) was calculated from a mathematical model using the measured PA and Zn contents. I have observed higher variation in PA and Zn contents, while WRC 5 and WRC 6 recorded the lowest and highest PA contents recording 8.24 mg/g and 17.41 mg/g respectively. The highest Zn content was observed in WRC 28 (18.95 µg/g), whereas the lowest value was observed in WRC 66 (47.56 µg/g). Furthermore, no significant correlation (p>0.05) was observed between the PA and Zn contents among WRC accessions. The calculated mean Zn intake from the WRC accessions ranged from 0.602 – 1.222 mg Zn/day. A negative correlation (p<0.05) was observed between the PA content and TAZ values, whereas the correlation was positive (p<0.05) between the Zn content and TAZ values. Further considering the sub-groups, there was no any significant differences

among the *japonica*, *indica* and *aus* groups or among the geographical regions (p>0.05). The results suggest PA content is more important for improving Zn bioavailability in the rice grains.

Further, correlations among the phenotypic traits (number of days to heading, culm length, panicle number, panicle length, grain length, grain width, seed length, seed width, seed weight, Zn content, PA content and amylose content) of WRC were analyzed to identify other traits affecting the PA content. The PA content was negatively correlated (p<0.05) only with the amylose content, and no significant correlations were observed with any other phenotypic traits.

Mineral and heavy metal contents of low (WRC 5, WRC 12, WRC 30) and high PA (WRC 6, WRC 22, WRC 44) rice accessions were determined by ICPMS method. There were no significant differences (p>0.05) in the contents of Ca, Manganese (Mn), Fe and Copper (Cu) among the accessions. Levels of Ca, Mn, Fe and Cu in low PA rice ranged from 105.0 – 115.7 µg/g, 18.2 – 29.1 µg/g, 7.9 – 9.5 µg/g and 2.5 – 3.9 µg/g, whereas in the high PA accessions ranged from 109.4 – 126.3 µg/g, 24.5 – 38.8 µg/g, 10.2 – 12.9 µg/g and 2.9 – 3.9 µg/g, respectively. There were no considerable differences in the heavy metal content among low and high PA accessions. These results suggest that PA and mineral accumulation are independent from each other so that the improving PA trait could be possible without affecting mineral contents.

I have conducted GWA mapping for contents of PA, Zn and TAZ using PLINK software. The SNPs for PA did not co-localize with the reported PA biosynthesis genes. No common SNPs were observed among PA and Zn traits, also suggest that PA and Zn accumulation is regulated by independent genes. Further, Zn bioavailability in rice is strongly affected by the PA content of the grain. So that reducing the concentration of PA in rice could be a possible approach for improving Zn bioavailability since PA content and other agronomic traits are independently regulated.

Genome-Wide Association and Seed Proteomic Studies Identify Important Genes for Phytic Acid Accumulation in Rice

To understand the underlying regulation mechanism of PA accumulation, comparison of gene-expression profiles and comparative proteomics analysis was carried out at the critical PA accumulation stage of the developing grain between low and high PA accessions.

First, I have conducted GWAS analysis using 700,000 significant nucleotide polymorphism (SNP) from a high-density rice array (HDRA) database. 62 accessions where genotypic data available were used to calculate principal components (PCs) and generate kinship matrix (K) with the centered identity-by-state. The genotype data were filtered for SNPs with a minimum frequency of 0.05 for the minor allele and 90% of call rate of the accessions. Then the analysis was done in the Trait Analysis by Association Evolution and Linkage (TASSEL) version 5.2.5 using different models, including General Linear Model (GLM) with Q (population structure), GLM with K (kinship), Mix Linear Model (MLM) with Q and MLM with Q+K. The significant threshold was set at $p < 4.5 \times 10^{-4}$ and all the annotated genes within 200 kbp of each significant SNP region were obtained from Rice Annotation Project Database (RAP-DB). This resulted 10 significant SNPs associated with PA content identified by the best fit model, GLM as observed in Quantile-Quantile plot while MLM showed an overcorrection of data. GWA mapping identified 194 candidate genes within a 200-kb genomic region in the RAP-DB, however the significant SNPs identified do not co-localize with the already known PA biosynthetic genes. Interestingly, *myo*-inositol-3-phosphate synthase (*INO1)* gene located close to the significant SNP in Chromosome 3. This suggested that *INO1* might influence the PA content in rice grain. Heatmap was developed from the identified candidate genes using the expression data available in the rice microarray database. I have identified 17 genes with high expression (with >2.0 normalized signal intensity) focusing the developing rice embryo during both 10 and 14 days after flowering (DAF). Among them, Os05g0525900 (zinc finger CCCH domain-containing protein 37), a transcription factor was identified within the peak region of SNP 7 in Chromosome 5 which may have potential effect on regulation of PA content in rice.

Then gene sequence analysis was conducted for *INO1* gene to understand any nucleotide polymorphism for the difference of PA contents in WRC 5 and WRC 6. The results revealed there was no any difference in the *INO1* gene sequence between two accessions, suggesting a different mechanism other than nucleotide difference might affected the variation in PA content.

Another experiment was conducted to determine the critical PA accumulation stage in the developing rice grains using the selected low and high PA WRC accessions. The seed samples for determining the PA content and gene expression analysis were collected from 5, 10, 15, 20, 25 and 30 DAF of the two accessions. At 5 DAF, PA content was very low in both the accessions and

started higher accumulation during 10 DAF. The rate of PA accumulation was significantly higher than WRC 5 during 10 DAF. Next, I have examined the expression of *INO1* gene, at this significant PA accumulation stage. Results of the differential expression of *INO1* gene in WRC 5 and WRC 6 showed a significant difference between two accessions at 10 DAF whereas WRC 6 showed a higher expression of the gene with a 2-fold increase than the low PA accession.

Further to understand the regulation mechanism and to identify important metabolic proteins expressing at 10 DAF, WRC 5 and WRC 6 seeds were subjected to proteome analysis using liquid chromatography-mass spectrometry. A total of 940 proteins were identified while 96% of the proteins were common to both WRC 5 and WRC 6 at 90% threshold level. Proteins were considered as being differentially expressed when proteins have both a fold- change more than 20% and a *p*-value below 0.05. Among the characterized proteins, *INO1*, which catalyzing the rate limiting first step of PA biosynthesis pathway was found having significant difference in expression ($p<0.05$) among the low and high PA rice in this study, while WRC 6 having 1.8-fold change increase than in WRC 5. Further, 19 and 26 unique proteins were observed only in WRC 5 and WRC 6 separately at 10 DAF in rice grains.

To better understand the gene expression level differences between the two accessions, I have sequenced the promoter region (1000 bp upstream of the start codon) of *INO1* gene to identify any different *cis*-regulatory elements among the accessions. However, I did not observe any difference in the promoter region among the two accessions. The difference between the accessions might be due to the transcription factors affecting the PA genes. However, further investigations are needed to confirm the possible causes for this variation.

Conclusion

The results suggest that not only the gene itself but the expression regulation of INO1 at early developmental stage is important in PA accumulation in rice grains. These research findings would advance the understanding on the regulation mechanism of PA accumulation and provide insights for research approaches toward improving micronutrient bioavailability in rice. However, further investigations, particularly on physiological, molecular ang genetic mechanisms of P uptake, translocation and accumulation along with the environmental factors would be important.