QUANTITATIVE TRAIT LOCI ANALYSIS FOR ANTIOXIDANT TRAITS IN RICE *Oryza sativa* L.

Miss Saranyu Thaworn

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Biotechnology
Common Course
FACULTY OF SCIENCE
Chulalongkorn University
Academic Year 2020

การวิเคราะห์ตำแหน่งยีนของลักษณะเชิงปริมาณสำหรับลักษณะต้านออกซิเดชันในข้าว *Oryza sativa* L.

น.ส.ศรัณยู ถาวร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาเทคโนโลยีชีวภาพ ไม่สังกัดภาควิชา/เทียบเท่า
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2563

Thesis Title	QUANTITATIVE TRAIT LOCI ANALYSIS FOR ANTIOXIDANT TRAITS IN RICE *Oryza sativa* L.
By	Miss Saranyu Thaworn
Field of Study	Biotechnology
Thesis Advisor	WARALUK KASETTRANAN, Ph.D.
Thesis Co Advisor	Assistant Professor CHANITA PALIYAVUTH, Ph.D. Assistant Professor MONNAT PONGPANICH, Ph.D.

Accepted by the FACULTY OF SCIENCE, Chulalongkorn University in Partial Fulfillment of the Requirement for the Doctor of Philosophy

..... Dean of the FACULTY OF SCIENCE
(Professor POLKIT SANGVANICH, Ph.D.)

DISSERTATION COMMITTEE

..... Chairman
(Associate Professor PATTAMAWADEE YANATATSANEEJIT, Ph.D.)

..... Thesis Advisor
(WARALUK KASETTRANAN, Ph.D.)

..... Thesis Co-Advisor
(Assistant Professor CHANITA PALIYAVUTH, Ph.D.)

..... Thesis Co-Advisor
(Assistant Professor MONNAT PONGPANICH, Ph.D.)

..... Examiner
(Assistant Professor KANOGWAN SERAYPHEAP, Ph.D.)

..... Examiner
(YUPYN CHINTAPAKORN, Ph.D.)

..... External Examiner
(Supaporn Junbuathong, Ph.D.)

ศรัณยู ถาวร : การวิเคราะห์ตำแหน่งยีนของลักษณะเชิงปริมาณสำหรับลักษณะต้านออกซิเดชันในข้าว *Oryza sativa* L.. (QUANTITATIVE TRAIT LOCI ANALYSIS FOR ANTIOXIDANT TRAITS IN RICE *Oryza sativa* L.) อ.ที่ปรึกษาหลัก : อ. ดร. วราลักษณ์ เกษตรานันท์, อ.ที่ปรึกษาร่วม : ผศ. ดร.ชนิตา ปาลิยะวุฒิ,ผศ. ดร.มนนัทธ์ พงษ์พานิช

การวิจัยครั้งนี้ดำเนินการเพื่อศึกษาลักษณะต้านออกซิเดชัน ซึ่งประกอบด้วยปริมาณสารฟีนอลิคทั้งหมด (TPC) สารฟลาโวนอยด์ทั้งหมด (TFC) และฤทธิ์ต้านออกซิเดชัน (AC) ด้วยวิธีการทดสอบ ABTS ในเมล็ดข้าวพันธุ์ไทย 174 พันธุ์ ประกอบด้วยข้าวที่มีเยื่อหุ้มเมล็ดสีขาว 152 พันธุ์ สีแดง 10 พันธุ์ และสีม่วง 12 พันธุ์ ประเมินการวิเคราะห์กลุ่มแบบขั้นตอนและพันธุ์ข้าวทั้งหมดสามารถจำแนกได้เป็นสี่กลุ่มตามการต้านออกซิเดชัน กลุ่มที่ 1 ประกอบด้วยพันธุ์ที่มีเยื่อหุ้มเมล็ดสีขาว 72 พันธุ์ และมีค่า TPC TFC และ AC ต่ำที่สุด กลุ่มที่ 2 ประกอบด้วยพันธุ์ที่มีเยื่อหุ้มเมล็ดสีขาว 80 พันธุ์ และแสดงค่า TFC สูงกว่ากลุ่มที่ 1 กลุ่มที่ 3 ประกอบด้วยพันธุ์ที่มีเยื่อหุ้มเมล็ดสีแดง 6 พันธุ์ และพันธุ์ที่มีเยื่อหุ้มเมล็ดสีม่วง 10 พันธุ์ ค่า TPC TFC และ AC ของกลุ่มนี้สูงกว่ากลุ่มที่ 1 และ 2 สุดท้ายคือกลุ่มที่ 4 ประกอบด้วยพันธุ์ที่มีเยื่อสีแดง 4 พันธุ์ และพันธุ์ที่มีเยื่อหุ้มเมล็ดสีม่วง 2 พันธุ์ โดยมีค่า TPC TFC และ AC สูงที่สุด

การศึกษาความสัมพันธ์เชื่อมโยงในจีโนมดำเนินการด้วยซอฟต์แวร์ GEMMA และ MLM โดยใช้เครื่องหมาย SNP 209594 เครื่องหมายกับลักษณะต้านออกซิเดชันของข้าวไทย 159 พันธุ์ QTLs 8 ตำแหน่งถูกระบุว่าอยู่บนโครโมโซมที่ 1 4 5 6 7 และ 8 ในบรรดา QTLs เหล่านี้ มี 1 QTL ที่ยืนยันว่าพบยีน *Rc* และ *OsCHS2* ซึ่งเป็นยีนต้านออกซิเดชัน ส่วนอีก 7 QTLs มี 12 candidate genes (MYB family transcription factors 7 ยีน helix-loop-helix family of transcriptional regulatory proteins 3 ยีน และ WD-repeat domain protein families 2 ยีน) ของการต้านออกซิเดชัน นอกจากนี้ QTLs บนโครโมโซมที่ 6 ยังวางตัวอยู่ใกล้กับยีน *OsC1* ดังนั้น เครื่องหมาย SNP 17 เครื่องหมาย ใน 16 loci บนโครโมโซมที่ 1 5 6 7 และ 8 ได้ถูกคัดเลือกจากเครื่องหมาย SNP 106 เครื่องหมาย ใน 38 loci ของ QTLs ทั้งหมดเพื่อตรวจสอบโปรไฟล์ของเครื่องหมาย SNP ในข้าวพันธุ์กข41 และไรซ์เบอร์รี่ (พ่อแม่) ด้วย sanger sequencing ผลการทดลองพบว่าเครื่องหมาย SNPs 3 เครื่องหมายแสดงโพลิมอร์ฟิซึมระหว่างข้าวทั้งสองพันธุ์ เครื่องหมาย SNP ทั้ง 3 เครื่องหมายได้รับการประเมินด้วยการวิเคราะห์ HRM และพบว่าเฉพาะเครื่องหมาย SNP ที่ตำแหน่ง 4219208 สามารถแสดงความแตกต่างของโปรไฟล์พันธุกรรมระหว่างพ่อแม่ เครื่องหมาย SNP นี้ใช้สำหรับลักษณะต้านออกซิเดชัน ได้ถูกตรวจสอบในประชากร F_2 ที่คัดเลือกไว้ ผลการทดลองแสดงว่า TPC TFC และ AC มีค่าสัมประสิทธิ์ของการกำหนดต่ำ คือ 8.62×10^{-3}% 3.51×10^{-2}% และ 7.10×10^{-3}% ตามลำดับ ผลการทดลองนี้เสนอให้เห็นว่าเครื่องหมาย SNP ที่ตำแหน่ง 4219208 ไม่สามารถใช้เป็นเครื่องหมายโมเลกุลเพื่อการคัดเลือกสารต้านออกซิเดชันสูงในข้าวได้ อย่างไรก็ตามยังมี candidate genes อื่นในการศึกษาครั้งนี้ที่ยังไม่ได้รับคัดเลือกเพื่อตรวจหาเครื่องหมายโมเลกุล ยีนเหล่านี้เป็นไปได้ที่จะแสดงโพลิมอร์ฟิซึมและอาจจะใช้เป็นเครื่องหมายโมเลกุลใหม่สำหรับการต้านออกซิเดชันในเมล็ดข้าวได้

การศึกษานี้ช่วยให้เข้าใจเกี่ยวกับพันธุศาสตร์ของลักษณะต้านออกซิเดชันในเมล็ดข้าวได้ชัดเจนยิ่งขึ้น ลักษณะต้านออกซิเดชันในประชากรนี้แสดงออกด้วยองค์ประกอบทางพันธุกรรมโดยปฏิกิริยาแบบผลบวก นอกจากนี้การพบ candidate genes ใหม่ที่มีความสัมพันธ์กับลักษณะต้านออกซิเดชันเป็นส่วนสำคัญในการพัฒนาเครื่องหมายโมเลกุลสำหรับการปรับปรุงพันธุ์พืช

สาขาวิชา เทคโนโลยีชีวภาพ ลายมือชื่อนิสิต .....
ปีการศึกษา 2563 ลายมือชื่อ อ.ที่ปรึกษาหลัก .....
ลายมือชื่อ อ.ที่ปรึกษาร่วม .....
ลายมือชื่อ อ.ที่ปรึกษาร่วม .....

5772869023 : MAJOR BIOTECHNOLOGY

KEYWORD: Rice, Antioxidant, GWAS, SNPs

Saranyu Thaworn : QUANTITATIVE TRAIT LOCI ANALYSIS FOR ANTIOXIDANT TRAITS IN RICE *Oryza sativa* L.. Advisor: WARALUK KASETTRANAN, Ph.D. Co-advisor: Asst. Prof. CHANITA PALIYAVUTH, Ph.D.,Asst. Prof. MONNAT PONGPANICH, Ph.D.

This research was conducted to study the antioxidant traits that contained the total phenolic content (TPC), total flavonoid content (TFC), and the antioxidant capacity (AC), by ABTS assay in the seed of 174 Thai rice cultivars, comprising 152 white, 10 red, and 12 purple pericarp cultivars. The hierarchical cluster analysis was evaluated, and the rice cultivars were classified into four clusters based on the antioxidant. Cluster I comprised of 72 white pericarp cultivars and had the lowest TPC, TFC, and AC. Cluster II contained 80 white cultivars and showed higher TFC values than cluster I. Cluster III consisted of 6 red and 10 purple cultivars. The TPC, TFC, and AC values of this cluster were higher than cluster I and II. Finally, cluster IV comprised of 4 red and 2 purple cultivars.

The Genome-wide association study was performed by GEMMA software and MLM using 209594 SNPs and the antioxidant traits of 159 Thai rice cultivars. The eight QTLs were identified on chromosomes 1, 4, 5, 6, 7, and 8. Among these, one QTL was confirmed *Rc* and *OsCHS2* genes which were the antioxidant genes. The seven QTLs were 12 candidate genes (7 MYB family transcription factors, 3 helix-loop-helix family of transcriptional regulatory proteins, and 2 WD-repeat domain protein families) of antioxidants. In addition, a QTL on chromosome 6 was closely located with the *OsC1* gene. Thus, 17 SNPs in 16 loci on chromosomes 1, 5, 6, 7, and 8 were selected from 106 SNPs in 38 loci of all QTLs for examining SNP profiles in RD41 and Riceberry (parents) by sanger sequencing. The result revealed that the three SNPs exhibited polymorphism between two rice cultivars. The three SNPs were evaluated by HRM analysis and found that only the SNP at position 4219208 could express the difference of genotyping profile between parents. This SNP for antioxidant traits was validated in the selected F_2 population. The results showed that the low values of coefficient of determination (R^2) for TPC, TFC, and AC were 8.62×10^{-3}%, 3.51×10^{-2}%, and 7.10×10^{-3}%, respectively. This result suggested that the SNP at position 4219208 was unable to be the molecular marker for high antioxidant selection in rice. However, there were other candidate genes of this study that were not selected for detecting molecular markers. These genes could be possible exhibit polymorphism and may able be a new molecular marker for the antioxidant in the rice seed.

This study could help in a clearer understanding of the genetics of antioxidant traits in the rice seed. The antioxidant traits in this population were expressed by genetic components with the additive genes action. Moreover, the discovery of new candidate genes associated with antioxidant traits should be an important part of the development of molecular markers for plant breeding.

Field of Study:	Biotechnology	Student's Signature .....
Academic Year:	2020	Advisor's Signature .....
		Co-advisor's Signature .....
		Co-advisor's Signature .....

ACKNOWLEDGEMENTS

The researcher would like to express sincere thanks and appreciation to her dissertation advisor, Dr. Waraluk Kasettranan, and two co-advisor, Asst. Prof. Chanita paliyavuth and Asst. Prof. Monnat Pongpanich for their kind assistance and continuous encouragement throughout conducting her dissertation. Their teachings and advice were valuable and beneficial to her academic work. Her sincere thanks are also for Prof. Dr. Supachitra Chadchawan and Assist. Prof. Dr. Jittra Piapukiew for valuable suggestions and endurable supports related to this research work.

Many grateful thanks are to Dr. Supaporn Junbuathong at Pathumthani Rice Research Center, Thanyaburi, Pathum Thani, for her provision of knowledge and suggestion on rice hybridization and permission of using the rice planting area as a part of the experiment. Her thanks and gratitude are extended to the dissertation committees, Asst. Prof. Dr. Kanogwan Seraypheap and Dr. Yupyn Chintapakorn, for their excellent dissertation verification, corrections and suggestions.

In addition, her special and grateful thanks are for Assoc. Prof. Dr. Sunthorn Kohtbantau, who attentively and thoroughly read the whole text for absolute perfection of the proper English usages.

Last but not least, her heart-felt and fondest thanks are for her beloved family, Mom and Dad. Without their support, encouragement, and understanding, this highest educational goal of hers could not have been attained.

This thesis was supported by The 90th Anniversary of Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund), the Agricultural Research Development Agency (a public organization) (PRP6105020640), the Program in Biotechnology, the Faculty of Science, Chulalongkorn University, Thailand, and the Department of Botany, Faculty of Science, Chulalongkorn University, Thailand.

Saranyu Thaworn

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

Introduction

Rice (Oryza sativa L.) is cultivated and consumed by a large number of people in Asia (Bhattacharjee et al., 2002). A whole rice seed contains macronutrients and micronutrients including carbohydrates, proteins, fatty acids, vitamins, and minerals (Lee et al., 2019). In rice bran, several phytochemicals are antioxidants, and the major antioxidant contains a phenolic compound (Gunaratne et al., 2013). Phenolic compounds can provide health benefits associated with reducing the risk of some chronic diseases, such as coronary heart disease, type-2 diabetes, and cancers (Kazemzadeh et al., 2014; Shao and Bao, 2015; Sun et al., 2010). Therefore, rice consumption is provided healthily and can eliminate pathogenic free radicals.

Coloured rice or pigmented rice is mainly red and purple from the pericarp layers. It is rich in the antioxidant compound and has more antioxidant capacity than white pericarp rice or non-pigmented rice (Goffman and Bergman, 2004; Shen et al., 2009). Therefore, the consumption of pigmented rice is increasing. The previous research revealed that the phenolic and flavonoid contents correlated to the rice antioxidant activities (Pramai and Jiamyangyuen, 2016; Shen et al., 2009). Therefore, an understanding of the genetic antioxidant activities and related antioxidant genotypes could be the basis for improving rice with higher antioxidants by the breeding program. The previous study of Sanghamitra et al. (2018) estimated genetic parameters and heritability of total phenolic content (TPC), total flavonoid content (TFC) and antioxidant capacity (AC) in 11 India pigmented rice cultivars. The results found that phenotypic coefficient of variation (PCV), genotypic coefficient of variation (GCV), and broad-sense heritability (H^2) were high for the antioxidant traits, but the differences between PCV and GCV were low. These results may be due to additive genes and transmissible potentiality to offspring, which were pre-requisite for breeders. Thus, the improvement of antioxidants in pigmented rice is possible.

GWAS is based on SNP markers and becomes a powerful tool for the detection of the candidate genes and new molecular markers for marker-assisted selection (MAS). Zhao et al. (2018) used GWAS to find cadmium (Cd) accumulation candidate genes from 312 diverse rice accessions. The 14 QTLs for Cd accumulation were identified, and QTL namely qCd3-2 was found in the *OsNRAMP2* gene, encoding Cd transportation protein in rice. This gene was used to study Cd accumulation in the rice seeds by the sequence analysis of 15 low and 15 high Cd accumulation rice accession. The different sequence of *OsNRAMP2* gene in low and high Cd accumulation rice was validated and expressed in 4 low and 4 high Cd accumulation in rice shoot by qRT-PCR. The results showed that the *OsNRAMP2* gene expression was increased in high Cd accumulation rice, but did not change significantly in low Cd accumulation rice. Therefore, *OsNRAMP2* can be used as the molecular marker for low Cd accumulation in rice. Moreover, Li et al. (2019) used GWAS to find the alkalinity tolerance gene in 295 japonica rice varieties. Eight QTLs associated with alkalinity tolerance traits including alkalinity tolerance (SAT), the concentrations of Na+ and K+ in the shoots (SNC and SKC, respectively). One QTL was associated with the SAT, SNC, and SNK on chromosome 3 with R^2 13.36-13.64%. The qRT-PCR analysis was used to identify the expression of *OsIRO3* locating in QTL on chromosome 3. The sequence analysis showed one 7-bp indel for the alternate allele in *OsIRO3*. Then, the 7-bp indel was used as the molecular marker for genotyping 60 tolerant and 66 sensitive rice varieties. The result indicated that 85.00% of the tolerant varieties showed the tolerant allele, and 83.33% of the sensitive varieties showed the sensitive allele. The 7-bp indel of *OsIRO3* was used for alkalinity tolerance via marker-assisted selection (MAS). Therefore, GWAS can be a good tool for detection of the candidate antioxidant genes in rice for new molecular markers.

This study aimed to identify quantitative trait loci (QTL) associated with TPC, TFC, or AC in seeds of Thai rice. Normally, high-antioxidant rice was selected by the

colour of the rice pericarps, but the relationship between the degree of antioxidant and colour was unknown. The results of this study can provide the candidate genes for new molecular markers to select high-antioxidant rice, which is likely to be more accurate than using the colour of the rice pericarp.

CHAPTER II

Literature Review

2.1 Rice

Rice (*Oryza sativa*. L) is the food crop of Asia that is classified into three subspecies including japonica, javanica, and indica. Japonica subspecies was originated in East Asia, such as China, Japan, and Korea. Javanica subspecies was originated in Java Islands, which were cultivated and consumed inland. Indica subspecies was originated in East or South East Asia. A large amount of consumed rice is indica and japonica subspecies. Rice seed (Figure 1) is divided into four parts which are 25% hull, 8% bran, 2% embryo, and 65% endosperm (Muthayya et al., 2014). One cup of cooked rice comprises 44.8 g carbohydrate, 5.03 g protein, 1.76 g fat, 11.28 mg vitamins, and 21.55 mg mineral (Lee et al., 2019).

The layers of rice caryopsis, from the outside to the inside, are hull, bran, endosperm, and embryo (Figure 1). The antioxidant compounds of rice seed are mostly distributed in the bran (pericarp, seed coat, nucellus, and aleurone layer) (Figure 1). Moreover, the phytochemicals in rice bran contain antioxidant capacity which has the potential to protect against several chronic diseases such as coronary heart disease, type-2 diabetes, and cancers (Kazemzadeh et al., 2014; Shao and Bao, 2015; Sun et al., 2010). The major compound of phytochemicals in rice is consists of phenolics, flavonoids, anthocyanins, proanthocyanins, vitamins, amino acids, phytosterols, γ-oryzanol, and others (Ravichanthiran et al., 2018).

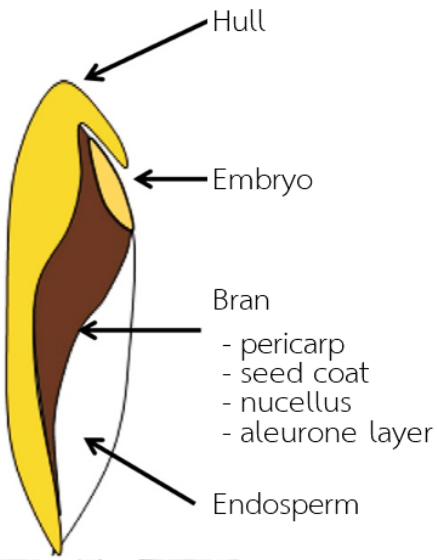


Figure 1 The components of the rice seed including hull, bran, endosperm, and embryo (Peanparkdee and Iwamoto, 2019).

Unpolished rice is called brown rice that mentions the caryopsis after the rice hull or husk is removed utilizing the mortar and pestle or rubber rolls. So, the rice seeds have pericarp, seed coat, nucellus, and aleurone layer, and they show the different colours of the pericarp such as white, red, and purple (Figure 2) (Babu et al., 2009; Ravichanthiran et al., 2018). Therefore, rice seeds can be classified as pigmented rice (red and purple pericarp) and non-pigmented rice (white pericarp) (Deng et al., 2013).

RD41 was generated by Rice Department, Ministry of Agriculture and Cooperative, from the cross of F_1 (“CNT85059-27-1-3-2” and “Supanburi60”) and “RP217-635-8”. This rice cultivar is photoperiod insensitivity, has a higher yield (722 kg/rai), and a 105-day harvesting time. The pericarp and seed coat colour of this cultivar are white. Non-pigmented rice contains low levels of phytochemicals, which have low antioxidant activities (Bureau of Rice Research and Development, 2020).

Riceberry was generated by the Rice Science Center of Kasetsart University, Kamphaeng Saen campus, from the cross of “Hom nil” and “Khao dawk mali 105”.

This rice cultivar is photoperiod insensitivity, has a lower yield (300-500 kg/rai), and has a 130-day harvesting time. Its pericarp colour is purple and has high antioxidant activities, which is one of the important determinants of rice nutritional quality (Poosri et al., 2019).

Figure 2 The pigmented rice seeds. (A) red, (B) purple, and (C) white pericarps.

2.2 Antioxidants

2.2.1 Antioxidant compound

Antioxidant compounds are identified as organic molecules that are beneficial to human health. They have the potential to protect the human cell-damaging from being destroyed by free radicals and reactive oxygen species that might otherwise exert harmful metabolic effects. The antioxidant compounds in rice seed are phytochemical, and the most dependent on the contents of phenolic compounds including phenolic acids, flavonoid, anthocyanins, and proanthocyanins (Goufo and Trindade, 2014).

Phenolic acids are secondary metabolites that are largely distributed in fruits, vegetables, and cereal grains. There are one or more phenolic rings with one or more organic carboxylic acid functions (Figure 3). Twelve phenolic acids are detected in rice, ranging from 7.3 to 8.7 mg/100 g in the endosperm, 177.6 to 319.8 mg/100 g in the bran, and 477.6 mg/100 in the husk, depending on the rice pericarp colour. The most phenolic acids found in rice are ferulic acid (56–77% of total phenolic acids), followed by p-coumaric acid (8–24%), sinapic acid (2–12%), gallic acid (1–6%), protocatechuic acid (1–4%), p-hydroxybenzoic acid (1–2%), vanillic acid (1%), and syringic acid (1%). Minor phenolic acids found in rice are caffeic, chlorogenic,

cinnamic, and ellagic acids and each compound reported for less than 1% of total phenolic acids (Goufo and Trindade, 2014; Ravichanthiran et al., 2018).

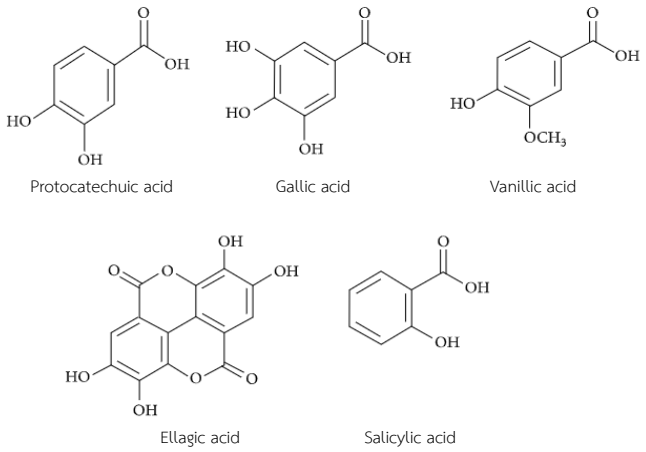


Figure 3 The structure of phenolic acid (Saibabu et al., 2015).

The flavonoids consist of a 15-carbon skeleton. Two aromatic rings (A-and B-rings) are linked by a three-carbon chain (structure C6-C3-C6) (Figure 4). Flavonoids can donate electrons for stopping chain reactions by the hydroxyl groups in the three-carbon chain (Goufo and Trindade, 2014). Flavonoids are one group of polyphenols that are classified as chalcones, flavones, flavonols, flavandiols, anthocyanidins, and proanthocyanidins (Figure 4) (Falcone Ferreyra et al., 2012). Flavonoids are generated by the general phenylpropanoid pathway, following by the flavonoid pathway. Malonyl CoA and p-coumaroyl CoA are substrates of the flavonoid pathway (Shih et al., 2008). Flavonoids content in rice seeds is generally much higher than in other parts, leaf sheet, leaf blade and it imparts purple and red colouration. Tricin (77% of all flavonoids) is the most abundant flavonoid that mainly existed in whole seed (Goufo and Trindade, 2014).

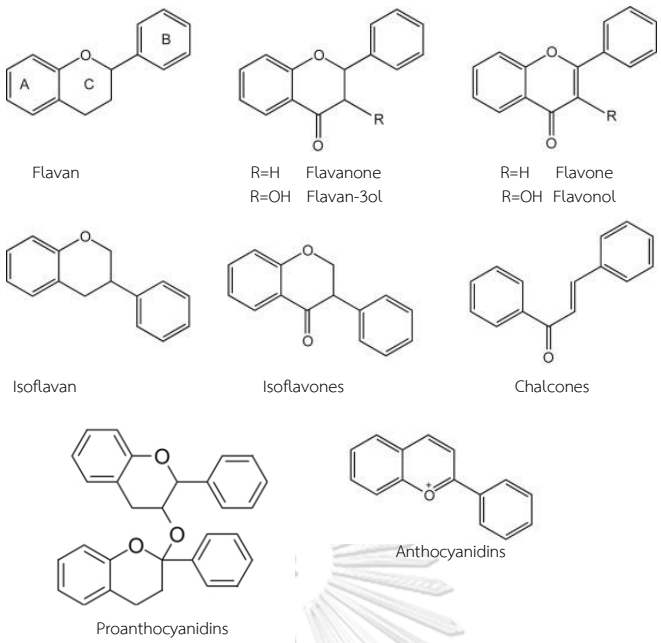


Figure 4 The structure of flavonoid (Hernández-Rodríguez et al., 2019).

Anthocyanidins and proanthocyanidins are a class of flavonoids and exhibited purple and red in rice pericarp, respectively. Several anthocyanidins have been found and separated from the coloured rice, including peonidin 3-glucoside, pelargonidin 3,5-diglucoside, cyanidin 3-glucoside, cyanidin 3-galactoside, cyanidin 3-rutinoside, cyanidin 3,5-diglucoside, and malvidin 3-galactoside (Deng et al., 2013). Moreover, anthocyanidins are intermediate in the biosynthesis of proanthocyanidins which are composed of catechin and epicatechin monomers (Goufo and Trindade, 2014).

2.2.2 The 2,2'-azino-bis-3-ethylbenzthiazoline-6-sulphonic acid (ABTS)

The method for determining the antioxidant activity is described as a decolorization assay appropriate to both lipophilic and hydrophilic antioxidants, including flavonoids, hydroxycinnamates, carotenoids, and plasma antioxidants. The ABTS active form, $ABTS^{\bullet+}$, is the hydrogen donated to the antioxidants and established by the oxidation of ABTS with potassium persulfate (Re et al., 1999)

Awika et al. (2003) reported the advantages of ABTS assay. This method is suitable for the analysis of antioxidant capacity because of its simplicity, diversity, reproducibility, and flexible usage in multiple media to verify both hydrophilic and lipophilic antioxidant capacity of food extracts and physiological fluids.

2.2.3 The antioxidant literature in rice

Studies on the antioxidant compounds (TPC, and TFC) and antioxidant capacity (AC) were the major studies for explaining the level and activity of antioxidants in rice. Shen et al. (2009) studied TPC, TFC, and AC in 481 rice cultivars including 423 white, 52 red, and 6 purple pericarps. They found that red and purple pericarp rice showed higher TPC, TFC, and AC than white pericarp rice. The antioxidant compounds were significantly positively correlated with the antioxidant activity, and the five components explained 83.7% of the total variance of the observed variables. Moreover, Pramai and Jiamyangyuen (2016) studied the colour parameters and antioxidant properties of Thai rice cultivars with purple, red, and white pericarps which were collected from many locations in the North of Thailand. The results suggested the high positive correlation between the antioxidant compounds and the antioxidant activity, and the rice cultivars could be classified into four groups based on colour parameters and antioxidant properties.

2.3 Single-nucleotide polymorphism (SNP)

SNP is a DNA sequence variation arising when a single nucleotide adenine (A), thymine (T), cytosine (C), or guanine (G) in one position has been replaced by another nucleotide. Different individuals have different variations. For example, two sequenced DNA fragments from two individuals were GGATTAT and GGATGAT; these contained one SNP where the alleles were T and G. Almost common SNPs have two alleles. Within a population, SNPs were the rare allele frequency ($>$ 1%) to observe in a particular population and were found once every several hundred base pairs in the genome (Brody, 2016). The SNPs in the human genome were used to understand gene functions, estimate the likelihood of certain diseases, and approximate

response to medicines (Shastry, 2007 and 2009). Moreover, the SNPs in plants were used for finding molecular markers in genes and the QTL discovery (Mammadov et al., 2012).

Nigro et al. (2017) detected the QTL of phenolic acid using 112 tetraploid cultivars of wheat. The QTLs were detected by GWAS of phenolic acid and used 13639 SNPs. The results indicated the two QTLs where *p*-coumaric acid was associated with the phenylalanine ammonia-lyase (*PAL2*) and p-coumarate 3-hydroxylase (*C3H*) genes on chromosome arms 2AL and 1AL, respectively.

2.4 Genome-wide association study (GWAS)

GWAS is an approach used in genetics research to associate specific genetic variations with specific phenotypes. The method involved scanning the genomes from many different individuals and looking for genetic markers that could be used to predict the presence of a phenotype. Once such genetic markers were identified, they could be used to understand how genes contribute to the phenotype (Tam et al., 2019).

GWAS is based on SNP markers and becomes a powerful tool for detecting the candidate genes and new molecular markers for marker-assisted selection (MAS). The SNP markers can be high-throughput sequence-based markers or array-based markers and distributed throughout the genome of an organism (Mammadov et al., 2012). Thus, this method was used to study organisms including humans and plants.

In the previous study, the potential roles of SNPs in the genetic mechanism of bipolar disorder were found using GWAS based on SNP markers (Qi et al., 2020). Moreover, Tao et al. (2018) used GWAS to find candidate genes of aluminum (Al) toxicity tolerance from 211 indica rice accessions with 395553 SNP markers. 21 QTLs for Al toxicity tolerance traits were identified at the seedling stage. Especially, QTL in chromosome 2 was found *Nrat1* gene, encoding metal ion transporter protein in rice. This gene was used to fine-mapping study by bulked-segregant analysis combined with whole-genome resequencing (BSA-seq) and linkage analysis in an F_2 population derived from the cross of Al tolerant accession; CC105 and super susceptible

accession; CC180. A total of 228 SNPs in the coding sequence and promoter region were observed in the *Nrat1* gene. However, two SNP markers in the *Nrat1* promoter and coding regions both play important role in regulating Al toxicity tolerance in 327 diversity rice accessions.

2.5 Genetic parameters

Genetic parameters were estimated by a statistical approach from the experiment data to understand phenotypes. The phenotypes of the population were controlled by a combination of genetic and environmental effects. Thus, the genetic parameters, including phenotypic coefficient of variation, genotypic coefficient of variation, and heritability were determined to predict the phenotype stability in the next generation (Mehboob et al., 2016; Sanghamitra et al., 2018).

Roy and Shil (2020) determined genetic parameters and heritability (H^2) of plant height, flag leaf length, flag leaf width, panicle length, grain per panicle, grain length, grain breadth, 1000 grain weight, active tillering, heading date, and maturity time from F_5 lines (Tulaipanji × IR64), F_3 lines (Tulaipanji × IR64 × PB1460), and F_3 lines (Badshabhog × Swarna sub1). The results found that PCV, GCV, and H^2 were high for 1000 grain weight and maturity time. However, the difference between PCV and GCV was low. These results may be due to additive genes and transmissible potentiality to the offspring.

2.6 High resolution melting (HRM) analysis

HRM analysis is the technique to classify the genotype of a genetic marker by the different melting profiles of each genotype. First, the amplification of DNA templates is performed, and the fluorescent dye is attached to PCR products. This fluorescent dye is bounded to double-stranded DNA to fluoresce brightly. Then, the PCR products were separated into single-stranded by increasing the temperature (65-95 °C). Finally, the melting profile of PCR products was conducted and analyzed to identify the genotype by the software (Farrar and Wittwer, 2017).

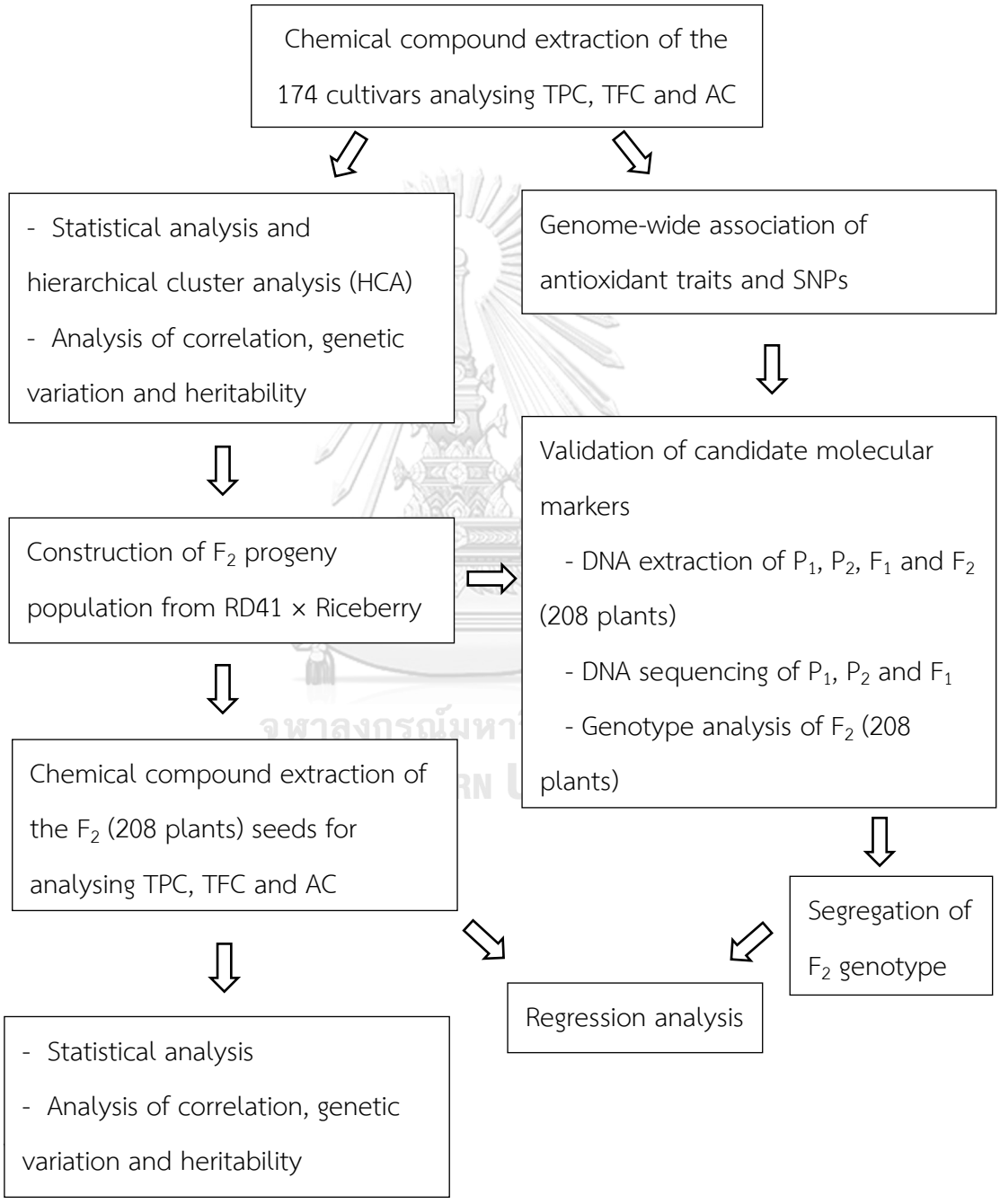
Garritano et al. (2009) used HRM for genotyping and mutant screening to find genetic variations in the tumor suppressor gene *TP53*. The results showed that the

HRM technique revealed high sensitivity and specificity for mutation screening. Therefore, HRM was suitable for mutation scanning of numerous amplicons. Moreover, Ceri et al. (2019) used HRM for genotyping to validate the percent agreement of SNPs. The results showed that the percent agreement of the c.1918G > T and c.2266A > C SNPs were 100% and 90%, respectively. Thus, the validation confirmed by PCR-HRM is a precise, accurate, and specific method for SNP genotyping.

CHAPTER III

Methodology

3.1 Flowchart of methodology



3.2 Plant materials

The seeds of 174 selected indica rice cultivars (Appendix A) were obtained from the Center of Excellence in Environment and Plant Physiology (Chulalongkorn University), Pathumthani Rice Research Center, Chum Phae Rice Research Center, Phatthalung Rice Research Center, and Maejo University, Thailand. These comprised 152 white, 10 red, and 12 purple pericarp cultivars. After that, the seeds were extracted and analyzed for TPC, TFC, and AC.

3.3 Antioxidant traits

3.3.1 Chemical preparation (Appendix B)

3.3.2 Sample preparation

Rice seeds were dehulled. A 0.5 g of sample from each cultivar was powdered by a mixer mill (MM 400, Restsch, Germany). The rice flour of each sample was mixed with 10 mL of extraction solvent [1% (v/v) hydrochloric acid in methanol] and extracted by the rotator shaker at room temperature for 24 h (Zhang et al., 2010a). The mixture was centrifuged at 4,000 g at 4 °C for 15 min and the supernatant was collected. The residue was re-extracted in the same procedure and the two supernatants were combined and filtrated by Whatman® Grade 542 syringe filter. All extracts were stored at 4 °C until the analyses were performed. A sample from each cultivar was extracted in three replications. The extracts were used for the measurements of the TPC, TFC, and AC in terms of the ABTS radical scavenging assay.

3.3.3 Total phenolic content (TPC)

The TPC was determined using the Folin-Ciocalteu colorimetric method proposed by Zhang et al. (2006), with slight modifications. In each well of a microplate, 20 µL samples were extracted, then 100 µL of 10% (v/v) Folin-Ciocalteu reagent and 100 µL deionized water were added; the mixture was incubated at room temperature for 1 min. Next, 80 µL of 7.5% (w/v) sodium carbonate solution was added to the mixture and then incubated for 30 min in the dark at room

temperature. The absorbance was measured at 765 nm using a UV-visible absorbance microplate reader (SpectraMax® M3, Molecular Devices, USA).

The gallic acid stock solution was diluted to 0.313, 0.156, 0.078, 0.039, and 0.020 mg/mL, which were used as the standard reference by examining TPC with the same method of sample extract. The standard curve was plotted by varying the concentration of gallic acid and the absorbance at 765 nm (Figure 20). The TPC was expressed as mg of gallic acid equivalents per g of rice (mg GAE/g rice).

3.3.4 Total flavonoid content (TFC)

The TFC was determined according to Herald et al. (2012) and Shen et al. (2009) with slight modifications. In each well of a microplate, 200 µL of 50% (v/v) ethanol and 7.5 µL of 5% (w/v) sodium nitrite were combined into the reaction mixture, followed by 10 µL of the respective extracted sample, and the reaction mixture was incubated for 5 min before adding 15 µL of 10% (w/v) aluminum chloride to the mixture and incubated for 5 min. Then, 50 µL of 1 M sodium hydroxide was added and the mixture was incubated for 15 min before reading the absorbance of the final solution at 400 nm using the microplate reader (SpectraMax® M3, Molecular Devices, USA).

The rutin was used as a standard reference by examining TFC with the same method of sample extract. The standard curve was plotted by varying the concentration of rutin from 0.05 to 0.80 mg/mL with their absorbance at 400 nm (Figure 21). The TFC was calculated as mg of rutin equivalent per g of rice (mg RE/g rice).

3.3.5 Antioxidant capacity (AC)

The AC was determined using the ABTS radical scavenging assay. First, the working solution of ABTS ($ABTS^{\cdot+}$ cation radical) was prepared overnight by the reaction of 7 mM of ABTS solution and 2.45 mM of potassium persulfate solution in the dark at room temperature (Re et al., 1999). Then, the $ABTS^{\cdot+}$ cation radical

solution was diluted with deionized water to adjust the absorbance at 734 nm (A_{734}) to 0.700 ± 0.02. Next, 200 µL of diluted $ABTS^{\cdot+}$ cation radical solution was transferred to 96-well plates, followed by 20 µL of the sample extract was added to each well and incubated for 30 min in the dark. The A_{734} values were converted to a percentage inhibition using Eq. (1),

$$\text{Percentage inhibition} = \frac{\text{A734 of control} - \text{A734 of sample}}{\text{A734 of control}} \times 100 \quad (1)$$

The control solution was prepared by mixing 20 µL solvent extraction and 200 µL of diluted $ABTS^{\cdot+}$ cation radical solution. This mixture was incubated for 30 min in the dark before reading the absorbance of the final solution at 734 nm.

The ranged of 5 to 100 ug/mL and percentage inhibition of ascorbic acid was used to plot the standard curve (Figure 22). The AC of all the samples was measured in terms of µg ascorbic acid equivalent antioxidant capacity per g of rice (µg AEAC/g rice).

3.3.6 Statistical analysis and hierarchical cluster analysis (HCA)

TPC, TFC, and AC were examined for normal distribution using the Kolmogorov-Smirnov test, and the level of significance was set at $p < 0.05$. In addition, one-way analysis of variance (ANOVA) and Duncan's multiple range test (DMRT) were used to test the difference of TPC, TFC, and AC between pericarp colours ($p < 0.05$). The statistical analyses were performed using the IBM SPSS Statistics version 22 software (International Business Machines, 2013).

The agglomerative HCA based on the three variables, TPC, TFC, and AC for 174 rice cultivars was performed using Ward's method (Milligan, 1980) implemented in the JMP9 software (Statistical Analysis Software Institute Inc, 2010). The result obtained from HCA was the dendrogram showing the relationships among 174 rice cultivars.

3.3.7 Analysis of correlation, genetic variation, and heritability

3.3.7.1 Correlation analysis

Pearson correlation was used to determine the correlation coefficients among the different parameters in the white, red, and purple pericarp cultivars, and the p-values of the correlation were determined at the $p < 0.05$ and 0.01 levels by IBM SPSS Statistics version 22 software (International Business Machines, 2013).

3.3.7.2 Genetic variation and heritability in 174 rice cultivars

The one-way ANOVA was performed using the IBM SPSS Statistics version 22 software (International Business Machines, 2013). ANOVA was used to calculate phenotypic (σ_p^2) and genotypic (σ_g^2) variances. Both parameters were used to estimated phenotypic (PCV) and genotypic (GCV) coefficient of variation by the method of Miller et al. (1958). Moreover, Broad-sense heritability (H^2) was calculated using Allard (1999) method by σ_p^2 and σ_g^2. These parameters were calculated by equation (2)-(6) (Appendix F).

$$\sigma_g^2 = \left(\frac{MSG - MSE}{r}\right) \quad (2)$$

$$\sigma_p^2 = \sigma_g^2 + \sigma_e^2 \quad (3)$$

$$GCV = \frac{\sqrt{\sigma_g^2}}{\bar{x}} \times 100 \quad (4)$$

$$PCV = \frac{\sqrt{\sigma_p^2}}{\bar{x}} \times 100 \quad (5)$$

$$H^2 = \frac{\sigma_g^2}{\sigma_p^2} \quad (6)$$

Where:

σ_e^2 = environmental variance.

MSG = mean square of genotypes (mean square between groups from the ANOVA table in Appendix F).

MSE = error mean square (mean square within groups from the ANOVA table in Appendix F).

r = number of replications.

$\bar{x}$ = mean of a trait.

3.4 Genome-wide association of antioxidant traits and SNPs

Genotype data from an SNP database of 159 selected rice cultivars with three pericarp colours: 150 white, 4 red, and 5 purple were obtained from the Center of Excellence in Environment and Plant Physiology, Chulalongkorn University. The markers data was prepared by following the study of Lekklar et al. (2019). In brief, SNP markers with minor allele frequency less than 0.05 or missing proportions greater than 0.4 were removed. A total number of 209,594 SNPs in promoter and exome regions of 12 chromosomes was subsequently used for the GWAS analysis described below.

The association analysis between these SNPs markers and antioxidant traits (TPC, TFC, and AC) were performed using GEMMA software (Zhou and Stephens, 2012 and 2014) which was based on a mixed linear model (Zhang et al., 2010b). Bonferroni correction was performed to adjust for multiple comparisons. The significance cut-off values were set at 5.18×10^{-7} and 4.42×10^{-7} for promoter and exome regions, respectively. The Manhattan and quantile-quantile (Q-Q) plots were generated using the R package qqman (Turner, 2014).

To define significant QTLs, SNPs that were significant in two or more antioxidant traits, with the coefficient of determination (R^2) more than 0.1 and had a physical distance within 200 kb of each other were grouped into the same QTL. However, if a significant SNP had no other significant SNP nearby, it was considered false positive. Then, we searched for genes that reside on the QTL region and 200 kb upstream and downstream of the QTL region. We regarded these genes as candidate genes.

3.5 Validation of candidate molecular markers

3.5.1 Construction of F_2 progeny population from RD41 × Riceberry

"RD41" (P_1) is white pericarp rice and has low antioxidants. "Riceberry" (P_2) is with the purple pericarp and has a higher antioxidant than the RD41. The harvesting period of the RD41 is 105 days, which is faster than Riceberry (130 days). Both cultivars are non-photoperiod sensitivity rice (Bureau of Rice Research and Development, 2020; Poosri et al., 2019). The F_1 was derived from RD41 × Riceberry by the hot air emasculation method. The female panicles of the RD41 were treated with water vapor at 45 °C for 10 minutes. The pollen grains lost viability while the stigma remained function. Then, the process of pollination could be done by bagging flowering male panicles of Riceberry (RB) above the emasculated female panicle in the glassine cross bag. The cross seed (F_1 seed) was mature and ready to be harvested in 28-30 days. The F_1 was cultivated to F_2 for at least 200 seeds.

Rice seeds were cultivated in an experimental field in the growing season with the same growing condition at Pathumthani Rice Research Center. The cultivated rice included the following: 20 plants of each parent, 15 plants of F_1, and 208 plants of F_2. After harvest, the seeds were analyzed for TPC, TFC, and AC.

3.5.2 DNA Extraction

The young leaves of P_1, P_2, F_1, and F_2 (208 plants) were collected for DNA extraction. The 25 mg dried leaves of each plant were homogenized with the stainless steel bead by the mixer mill (MM 400, Restsch, Germany) in a 1.5 mL microcentrifuge tube. The genomic DNA was extracted using E.Z.N.A.® Plant DNA Kit (Omega Bio-Tek, USA). Briefly, the 800 µL of P1 buffer (extraction buffer) was added to the tubes of the powdered leaf before incubating at 65 °C for 10 min. A lysate was mixed with 140 µL of P2 buffer (acetate buffer) and centrifuged at 10,000 g for 10 min. The 500 µL of clear lysate was transferred to a new 1.5 mL microcentrifuge tube, and 350 µL of isopropanol was added before centrifuging at 14,000 g for 2 min to pellet DNA. The supernatants were discarded, and the pellets were resuspended

by adding 300 µL of the heated sterile deionized water. The 4 µL of RNase, 150 µL of P3 buffer (binding buffer), and 300 µL of absolute ethanol were mixed into the resuspended pellets. The mixtures were transferred to the HiBind® DNA Mini Column before centrifuging at 10,000 g for 1 min. The HiBind® DNA Mini Columns were washed two times by 650 µL of DNA wash buffer before centrifuging at 10,000g for 1 min. The HiBind® DNA Mini Columns were dried by centrifuging at 10,000 g for 2 min. The 100 µL elution buffer was added to the HiBind® DNA Mini Columns in the 1.5 mL microcentrifuge tubes before incubating at room temperature for 5 min and centrifuging at 10,000 g for 1 min. The genomic DNA was determined for the quality by 0.8% agarose gel electrophoresis, and the concentration and purity were assessed by determining the ratio of absorbance at 260 and 280 nm. The genomic DNA was kept at -20 °C before genotyping by high-resolution melt (HRM) analysis.

3.5.3 DNA sequencing

The primers of 17 loci were designed (Appendix G). These primers were used to amplify the genomic DNA of P_1, P_2, and F_1 for analyzing the polymorphism of 17 SNPs. The PCR reaction was conducted in 30 µL of amplification volume including 1X PCR buffer (Invitrogen, USA), 1.5 mM $MgCl_2$ (Invitrogen, USA), 0.2 mM dNTP (Biotech rabbit, Germany), 0.5 unit of DNA polymerase (Invitrogen, USA), 0.5 µM primers (Bionics, South Korea) and 2 ng of genomic DNA. The reactions were carried out by GenePro Thermal Cycler (Bioer Technology, China) using the PCR condition at 94 °C for 3 min, followed by 35 cycles at 94 °C for 45 sec, 57-62 °C depending on primer (Appendix G) for 30 sec, 72 °C for 90 sec and final extension at 72 °C for 10 min. The PCR products were examined through 2% agarose gel electrophoresis in 1X TBE buffer and sequenced by Bionics Co., Ltd. (Seoul, South Korea).

3.5.4 Genotyping analysis

Three polymorphic SNPs markers were used to classify the genotype patterns of 208 F_2 plants by high-resolution melt (HRM) analysis. The PCR reaction was performed in 20 µL of amplification volume including 1X SsoFast™ EvaGreen®

Supermix (Bio-Rad, Canada), 0.25 μM primers (Bionics, South Korea), and 2 ng of genomic DNA. The amplification protocol was denaturation at 98 ℃ for 30 sec followed by a 35 cycle of 98 ℃ for 2 sec, or 52-62 ℃ depending on primer (Appendix G) for 10 sec and the plate was read. Then, the PCR mixtures were heated to 98 ℃ for 30 sec and cooled to 65 ℃ for 30 sec. The melt curve was performed by increasing the temperature from 65 ℃ to 90 ℃ (increasing the temperature by 0.2 ℃ at a time), 10 sec dwell times. The melt profiles of each F_2 were analyzed by Melt Analysis™ software (Bio-Rad, Canada).

3.5.5 Segregation and regression analysis

The Chi-square test was conducted to determine whether the genotypic ratio in F_2 progeny was 1:2:1 as expected.

The regression analysis was used to determine the relationship between the SNP genotype and the antioxidant traits (TPC, TFC, and AC) using the program R (R Core Team, 2020).

3.5.6 Genetic variation and heritability in 208 F_2 plants.

The phenotypic ($\boldsymbol{\sigma}_p^2$) and genotypic ($\boldsymbol{\sigma}_g^2$) variances were determined from 20 plants of each parent, 15 plants of F_1, and 208 plants of F_2. Both parameters were used to calculated GCV, PCV, H^2 (Globerson et al., 1987), and percentage of genetic advance (%GA) using equations (4), (5), (6), and (7) (Johnson et al., 1955).

$$\%GA = \frac{K \times \sqrt{V_P} \times H^2}{\bar{x}} \times 100 \qquad (7)$$

Where: K = the selection differential (K = 2.06 at 5% selection intensity)

CHAPTER IV

Results and Discussions

4.1 Antioxidant traits in 174 rice cultivars

4.1.1 TPC

The TPC results of the 174 rice cultivars were summarized in Figure 5A. The TPC of white pericarp cultivars ranged from 0.86–1.93 mg GAE/g with an average of 1.24 mg GAE/g. The red pericarp cultivars had the highest TPC levels, ranging from 3.85–5.79 mg GAE/g with an average of 4.70 mg GAE/g, while the TPC of purple pericarp cultivars were slightly lower than red pericarp, ranging from 3.01–4.74 mg GAE/g with an average of 4.07 mg GAE/g. The lowest and highest TPC of the white pericarp cultivar was RD19 and Ai tai, the red pericarp cultivar was Sang yod and Lai mahk and the purple pericarp cultivar was Hawm dam and Gam nahng payah, respectively. The average of the red pericarp cultivar was about 3.8-fold higher than the white pericarp cultivar group, while the average of the purple pericarp cultivar was about 3.3-fold higher than the white pericarp cultivar. However, in a previous study, the differences were greater than in this study. TPC in red pericarp cultivars was 4.6-fold higher than white pericarp cultivars (Xu et al., 2016).

Significant differences ($p < 0.05$) in the TPC were observed between the three groups of rice. The red pericarp cultivars had the highest average TPC level followed by the purple pericarp cultivars, and the white pericarp cultivars (Appendix D). This result is in agreement with Goffman and Bergman (2004), who studied the phenolic content in the seeds of 133 rice cultivars, where the red pericarp cultivars exhibited the highest TPCs. Shen et al. (2009) also reported that the TPC was positively correlated with the colour parameters of rice seeds including *L**, *a**, *b**, *C,* and *H°*. Whereas L*, a*, b*, C, and H° is an international standard for colour measurements, adopted by the Commission Internationale d'Eclairage (CIE) in 1976. Jun et al. (2012) analyzed the phenolic acids of white, red, and purple rice cultivars by HPLC. Their results indicated that protocatechuic acid, p-hydroxybenzoic acid, gallic acid, vanillic

acid, syringic acid, p-coumaric acid, ferulic acid, and sinapic acid were the major phenolic compounds in the red and purple rice cultivars and were presented at higher levels than in the white rice cultivars.

A previous study (Pramai and Jiamyangyuen, 2016) analyzed unpolished rice seeds of 20 Thai white and colour pericarp cultivars, and found that the highest TPC was in the purple cultivar, Niaw dam pleuk khow (6.34 mg GAE/g), while the lowest TPC was found in the white pericarp cultivar, Phitsanulok 2 (0.31 mg GAE/g). In our study, the purple cultivar Gam nahng payah (4.74 mg GAE/g) and the red cultivar Lai mahk (5.79 mg GAE/g) had the highest TPC in each group. The lowest TPC in the white pericarp rice was found in RD19 (0.86 mg GAE/g), while the Ai tai cultivar (1.93 mg GAE/g) was the white pericarp rice with the highest TPC.

4.1.2 TFC

The TFC showed values ranging from 2.03–5.07, 6.49–11.13, and 6.57–10.25 mg RE/g for the white, red, and purple pericarp rice groups, respectively, (Figure 5B). The lowest and highest TFC of the white pericarp cultivar group was Khao luang and Ai tai, the red pericarp cultivar group was Loi hah ruang and Chaw pli khao, and the purple pericarp cultivar group was Gam liaw and Gam nahng payah, respectively. The average TFC of the white, red, and purple pericarp cultivar groups were 2.98 mg RE/g, 8.62 mg RE/g, and 8.66 mg RE/g, respectively. A significant difference ($p < 0.05$) in the TFC was observed between the white pericarp cultivar group and the two groups of coloured pericarp cultivars, but no significant difference in the TFC was observed between the red and purple pericarp cultivar group. The red and purple pericarp cultivar groups had a higher average TFC level than the white pericarp cultivar group (Appendix D). Both averages of the red and purple pericarp cultivar groups were about 2.9-fold higher than the white pericarp cultivar group. These results suggested that pigmented rice had a higher TFC, compared to white rice.

Previous studies had indicated that the predominant flavonoids were proanthocyanins in the red pericarp cultivars and anthocyanins in the purple pericarp cultivars (Hosoda et al., 2018; Thitipramote et al., 2016). Moreover, proanthocyanins

and anthocyanins are present in many coloured plants, including orange, red, pink, mauve, purple, and blue. These colours in plants depended on the number of hydroxyl groups on the B-ring and glycosylation of the anthocyanidins (Tanaka et al., 2008). In a study on the colours of rice pericarp, Ghasemzadeh et al. (2018) suggested that quercetin, apigenin, catechin, luteolin, myricetin, and anthocyanin were the flavonoid compounds in white pericarp and pigmented rice, but that red and purple rice showed higher levels than white rice. Therefore, it should be noted here that the white rice cultivars had a lower TFC than the red and purple cultivars (Shen et al., 2009).

4.1.3 AC

The AC was measured using ABTS radical scavenging assay and was expressed as equivalents of ascorbic acid in Figure 5C. The AC of white rice was significantly lower than the other two groups of pigmented rice. Among the white rice cultivars, the average AC was 186.78 μg AEAC/g and ranged from 78.32–349.33 μg AEAC/g, whereas, among the red rice cultivars, the average AC was 1,641.90 μg AEAC/g, and ranged from 1,363.67–2,301.92 μg AEAC/g. The purple rice showed slightly lower AC values than the red rice, ranging from 727.24–1,880.62 μg AEAC/g with an average of 1,232.55 μg AEAC/g. The lowest and highest AC of the white pericarp cultivar group was Ma yom and Gwian hak, the red pericarp cultivar group was Sang yod and Lai mahk, and the purple pericarp cultivar group was Hawm nin and Gam feuang. The AC levels in each rice colour group (Figure 5C) were significant differences ($p<0.05$) between an individual colour group. The red pericarp cultivar group had the highest average AC level followed by the purple pericarp cultivar group, and the white pericarp cultivar group (Appendix D). The average of the red pericarp cultivar group was about 8.8-fold higher than the white pericarp cultivar group, while the average of the purple pericarp cultivar group was about 6.6-fold higher than the white pericarp cultivar group. The difference in the AC between the coloured pericarp rice and the white pericarp rice has been attributed to the presence of phenolic and flavonoid contents in the red and purple pericarps, which have strong antioxidant activities (Goffman and Bergman, 2004; Shao et al., 2018; Shen et al., 2009; Xu et al., 2016)

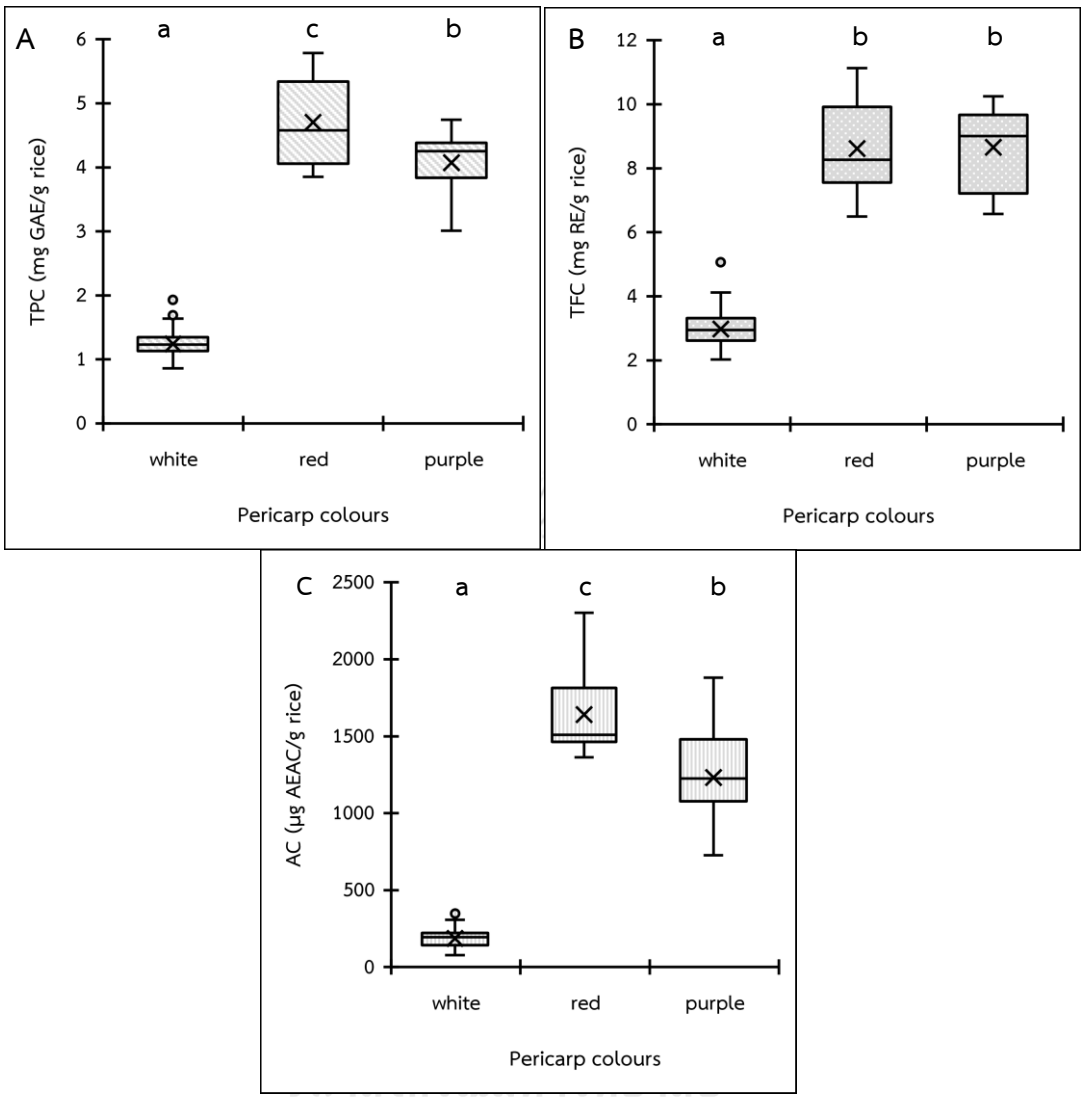


Figure 5 Boxplots of the (A) TPC, (B) TFC, and (C) AC of the white rice (152), red rice (10), and purple rice (12) cultivars.

Boxplots show the median (horizontal line in the box), standard deviation (bottom and top box lines) and average (mark in the box), and the range. For each parameter, means with a different letter are significantly different ($p < 0.05$; DMRT).

4.2 Distribution of TPC, TFC, and AC of 174 rice cultivars

Results from the Kolmogorov-Smirnov test showed that our data were not normal distribution (Figure 6). The data were separated into two groups according to pericarp colour in all three traits. White pericarp had a lower value of antioxidant

traits than coloured pericarp rice. In TPC, white pericarp had TPC values ranging from 0.86-1.93 mg GAE/g, while red and purple pericarp had TPC values ranging from 3.01-5.79 mg GAE/g, respectively (Figure 6A). In TFC, the range was from 2.03 to 5.07 mg RE/g for white pericarp and 6.49 to 11.13 mg RE/g for red and purple pericarp (Figure 6B). In AC, the range was from 78.32 to 349.33 μg AEAC/g for white pericarp and 727.24 to 2301.92 μg AEAC/g for red and purple pericarp, respectively (Figure 6C).

The previous studies found the relationship between the pericarp colours and antioxidant value (Goffman and Bergman, 2004; Shao et al., 2018; Shen et al., 2009; Xu et al., 2016). Antioxidant traits measured from the part of rice seed including husk, bran, and endosperm were analyzed. The results indicated that the antioxidant traits of bran were higher than husk and endosperm. The pericarp was the part of bran (Butsat and Siriamornpun, 2010). TPC, TFC, and AC varied widely not only according to colour but also cultivars. Our result was in agreement with the previous research (Shao et al., 2018).

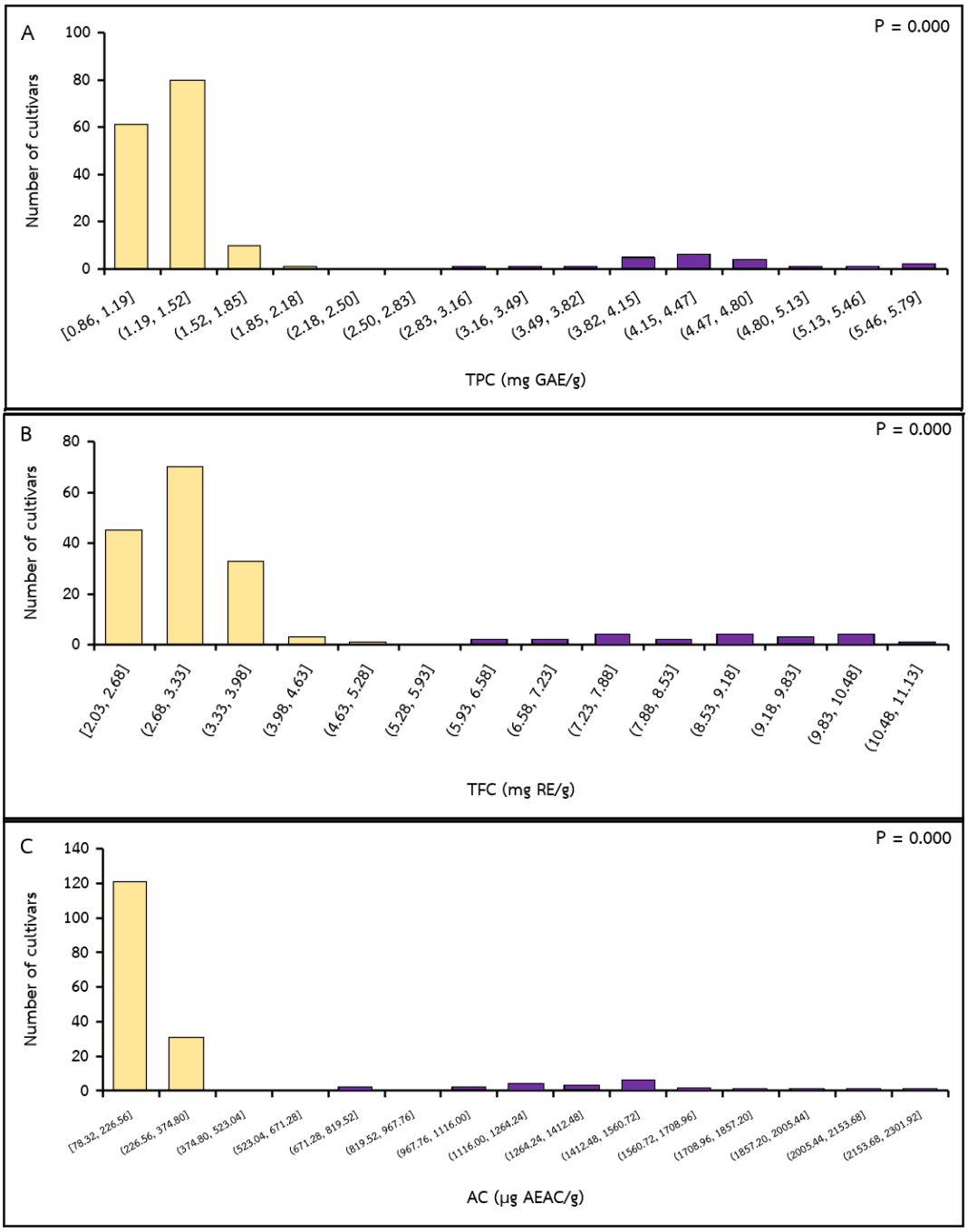


Figure 6 Frequency distribution of the antioxidant content and capacity in 174 Thai rice cultivars ($p < 0.05$).

(A) TPC, (B) TFC and (C) AC of the white (□) and coloured (■) pericarp rice groups.

4.3 HCA of 174 rice cultivars

The dendrogram of 174 rice cultivars based on these three antioxidant parameters (TPC, TFC, and AC) obtained from HCA was presented in Figure 7. The distance value of HCA ranged from zero (low difference) to 22 (high difference). The results showed that the 174 rice cultivars were grouped into four clusters at a distance value of 2.60. Cluster I comprised 72 white rice cultivars and had the lowest antioxidant contents and AC. In cluster II, the largest cluster, comprised of 80 white rice cultivars and had about the same range of TPC and AC values as found in cluster I. However, the TFC value of the white rice in this group was higher than the white rice in cluster I. Cluster III comprised 16 cultivars, six red rice cultivars, and ten purple rice cultivars. TPC, TFC, and AC values in this cluster were higher than in clusters I and II but were lower than in cluster IV. Finally, cluster IV had six cultivars, four red and two purple cultivars. This cluster exhibited the highest values of antioxidant contents and AC. Among the coloured pericarp cultivars, Hawm nin had the lowest TFC and AC, and Hawm dam expressed the lowest TPC. The red rice Lai mahk and the purple rice Gam feuang in cluster IV showed the highest antioxidant content and AC, respectively (Figure 7 and Table 1).

A previous evaluation of 20 rice cultivars from the North of Thailand classified them by the HCA method into four clusters according to their colours and antioxidant properties, i.e., total anthocyanin content, TPC, TFC, α-tocopherol, γ-oryzanol, and AC (Pramai and Jiamyangyuen, 2016). Cluster I comprised one purple and four red rice cultivars, cluster II had eight purple rice cultivars, cluster III had four purple rice cultivars with the highest TPC, TFC, and AC, and cluster IV had three white rice cultivars. Their cluster analysis suggested that the white rice cultivars clearly showed different TPC, TFC, and AC values (lower) compared to the red and purple rice cultivars. Most red and purple rice cultivars demonstrated similar TPC, TFC, and AC levels in clusters I and II, although some purple cultivars had higher

values of antioxidant content and AC than other red and purple rice cultivars in cluster III. Our results were broadly following their HCA clustering results.

A significant difference ($p < 0.05$) in the TPC and TFC was observed between the four groups (Table 2). Cluster IV showed the highest value of TPC, TFC, and AC, whereas cluster I showed the lowest value of these antioxidant content and capacity.

High TPC, TFC, and AC values were found in six pigmented rice cultivars in cluster IV (Table 1) including five Thai landraces (Gam feuang, Lai mahk, Chaw pli khao, Gam nahng payah, and Hawn daeng) and one Sri Lankan rice cultivar (Rathu heenati). Previously, Rathu heenati was reported to exhibit a moderate level of resistance against the brown planthopper. It is an important genetic resource, but there was no report on its TPC, TFC, and AC level (Jairin et al., 2007). Rice breeders could provide the cultivars that could lead to improved nutritional quality in rice via conventional and molecular breeding programs.

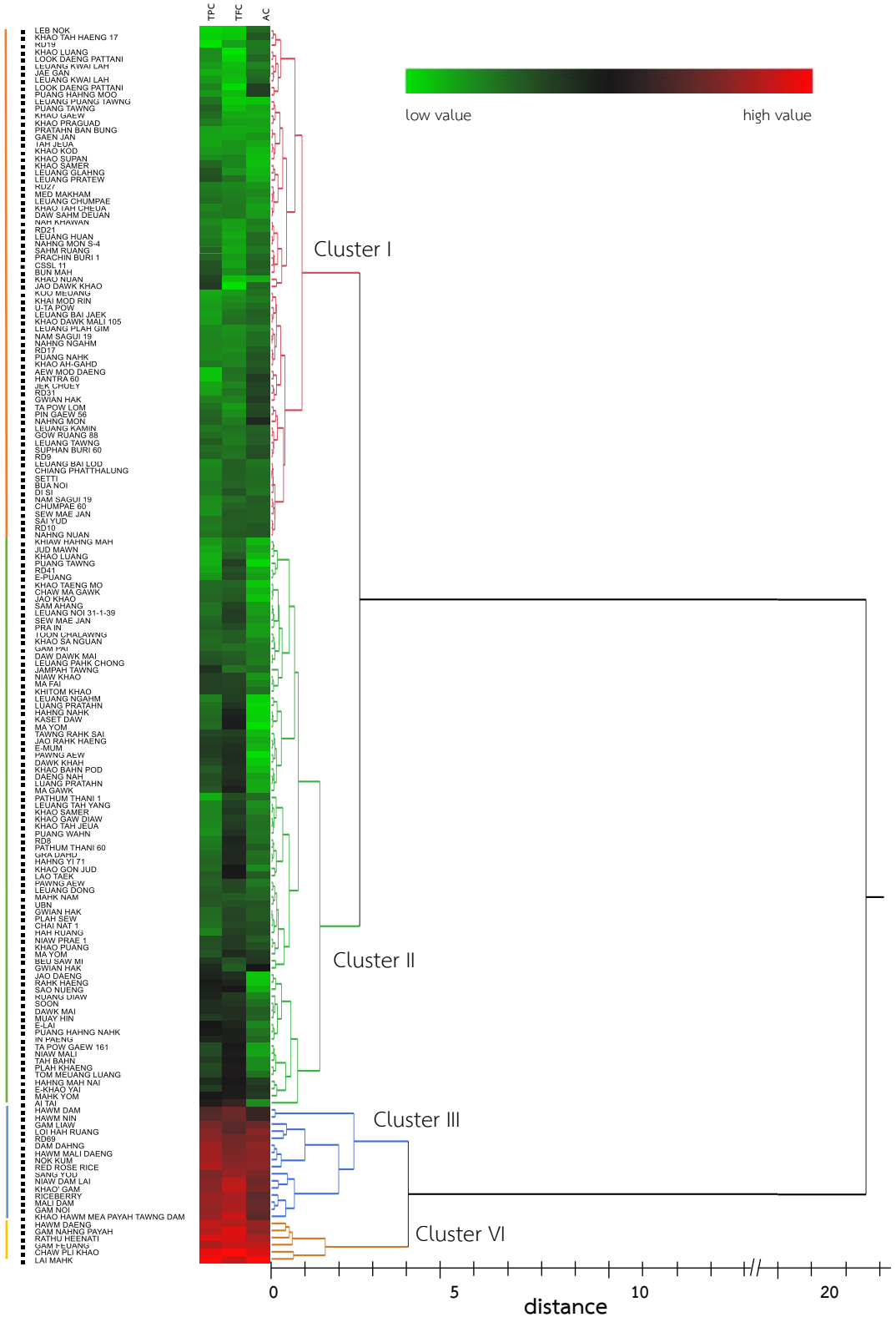


Figure 7 Dendrogram from HCA showing the relationship between the TPC, TFC, and AC (using ABTS assay) in 174 Thai rice cultivars.

Table 1 Clusters of rice cultivars according to HCA based on the TPC, TFC, and AC (ABTS assay).

Cluster number	Number of cultivars	Cultivars
Cluster I	72	Aew mod daeng, Bua noi, Bun mah, Chiang phatthalung, Chumpae 60, CSSL 11, Daw sahm deuan, Di si, Gaen jan, Gow ruang 88, Gwian hak, Hantra 60, jae gan, Jao dawk khao, Jek chuey, Khai mod rin, Khao ah-gahd, Khao dawk mali 105, Khao gaew, Khao kod, Khao luang, Khao nuan, Khao praguad, Khao samer, Khao supan, Khao tah cheua, Khao tah haeng 17, Koo meuang, Leb nok, Leuang bai jaek, Leuang bai lod, Leuang chumpae, Leuang glahng, Leuang huan, Leuang kamin, Leuang kwai lah, Leuang kwai lah, Leuang plah gim, Leuang pratew, Leuang puang tawng, Leuang tawng, Look daeng pattani, Look daeng pattani, Med makham, Nah khawan, Nahng mon, Nahng mon s-4, Nahng ngahm, Nahng nuan, Nam sagui 19, Nam sagui 19, Pin gaew 56, Prachin buri 1, Pratahn ban bung, Puang hahng moo, Puang nahk, Puang tawng, RD10, RD17, RD19, RD21, RD27, RD31, RD9, Sahm ruang, Sai yud, Setti, Sew mae jan, Suphan buri 60, Ta pow lom, Tah jeua, U-ta pow

Table 1 (Continued)

Cluster number	Number of cultivars	Cultivars
Cluster II	80	Ai tai, Beu saw mi, Chai nat 1, Chaw ma gawk, Daeng nah, Daw dawk mai, Dawk khah, Dawk mai, E-khao yai, E-lai, E-mum, E-puang, Gam pai, Gra dahd, Gwian hak, Gwian hak, Hah ruang, Hahng mah nai, Hahng nahk, Hahng yi 71, In paeng, Jampah tawng, Jao daeng, Jao khao, Jao rahk haeng, Jud mawn, Kaset daw, Khao bahn pod, Khao gaw diaw, Khao gon jud, Khao luang, Khao puang, Khao sa nguan, Khao samer, Khao taeng mo, Khao tah jeua, Khiaw hahng mah, Khitom khao, Lao taek, Leuang dong, Leuang ngahm, Leuang noi 31-1-39, Leuang pahk chong, Leuang tah yang, Luang pratahn, Luang pratahn, Ma fai, Ma gawk, Ma yom, Ma yom, Mahk nam, Mahk yom, Muay hin, Niaw khao, Niaw mali, Niaw prae 1, Pathum thani 1, Pathum thani 60, Pawng aew, Pawng aew, Plah khaeng, Plah sew, Pra in, Puang hahng nahk, Puang tawng, Puang wahn, Rahk haeng, RD41, RD8, Ruang diaw, Sam ahang, Sao nueng, Sew mae jan, Soon, Ta pow gaew 161, Tah bahn, Tawng rahk sai, Tom meuang luang, Toon chalawng, UBN
Cluster III	16	Dam dahng, Gam liaw, Gam noi, Hawm dam, Hawm mali daeng, Hawm nin, Khao' gam, Khao hawm mea payah tawng dam, Loi hah ruang, Mali dam, Niaw dam lai, Nok kum, RD69, Red rose rice, Riceberry, Sang yod
Cluster IV	6	Chaw pli khao, Gam feuang, Gam nahng payah, Hawm daeng, Lai mahk, Rathu heenati

Table 2 Comparison of the TPC, TFC, and AC in four clusters (Mean (SD)).

Character	Cluster I	Cluster II	Cluster III	Cluster IV
TPC	1.16 (0.11)[a]	1.33 (0.17)[b]	4.07 (0.45)[c]	5.13 (0.46)[d]
TFC	2.57 (0.25)[a]	3.35 (0.37)[b]	8.12 (1.09)[c]	10.03 (0.67)[d]
AC	174.62 (45.18)[a]	200.29 (55.70)[a]	1258.82(254.15)[b]	1844.77 (294.77)[c]

Cluster is from the HCA, see Figure 7. Means in the same row with a different letter are significantly different ($p < 0.05$; DMRT).

4.4 Analysis of correlation, genetic variation, and heritability of 174 rice cultivars

4.4.1 Correlation analysis

The pairwise correlation between TPC, TFC, and AC in 174 rice cultivars was shown in Figure 8. High positive significant correlations (r>0.90) were detected, ranging from 0.925 to 0.971. The significant correlation was detected between the TPC and TFC (r = 0.966, p<0.01), TPC and AC (r = 0.971, p<0.01) and TFC and AC (r = 0.925, p<0.01). The results were in agreement with Pramai and Jiamyangyuen (2016) and Shen et al. (2009), who studied the correlation between the TPC, TFC, and AC in white, red, and purple pericarp rice. The results of this and previous studies suggested that the antioxidant compounds are closely related to antioxidant activity. Thus, the related genes of TPC, TFC, and AC may be at the same or close location in the genome (Shen et al., 2009).

4.4.2 Genetic variation and heritability

The genetic variation parameters of 174 rice cultivars were calculated from one-way ANOVA (Appendix F). First, phenotypic (σ_p^2), genotypic (σ_g^2), and environmental (σ_e^2) variances were estimated. The three parameters were used to calculate PCV, GCV, and H^2 (Table 3). These parameters explained the net effect of genetic and environmental factors on the traits (Mulugeta et al., 2012). Broad-sense heritability values of TPC, TFC and AC were 0.83, 0.70 and 0.88, respectively. The low heritability was under 0.30, the moderate heritability was ranging from 0.30 to 0.60,

and the high heritability is over 0.60 (Robinson et al., 1949). The broad-sense heritability of antioxidant traits evaluated in this study was very high, indicating that the phenotypic variations of measured antioxidant traits were mainly affected by genetic factors. Therefore, the additive gene action was involved in all traits.

The PCV of TPC, TFC, and AC was 69.60, 60.34, and 131.73, respectively, and the GCV of TPC, TFC, and AC were 63.43, 50.41, and 123.79, respectively. The differences between the PCV and GCV of TPC, TFC, and AC were 6.17, 9.93, and 7.94, respectively. These scores were under 10, indicating a small difference. The difference between the PCV and GCV implied the influence of the environmental effect (Mehboob et al., 2016). A high difference reflected high environmental influence, whereas a small difference revealed high genetic influence. The results of phenotypic data analysis suggested the TPC, TFC, and AC were under high genetic influence. All traits could be used as selection criteria for rice antioxidant quality improvement programs. These results were supported by Sanghamitra et al. (2018) study, which reported the high broad-sense heritability value and the low difference between the PCV and GCV of TPC, TFC, and AC in 11 rice cultivars.

4.5 Genome-wide association study (GWAS).

One hundred and fifty-nine cultivars were used to study the genetic association with TPC, TFC, and AC. GWAS was conducted on 96,480 SNPs in promoter and 113,114 SNPs in the exome, 209,594 SNPs in total. Manhattan and quantile-quantile (Q-Q) plots showed the significant SNPs on promoter and exome regions (Figure 9 and 10). 158 significant SNPs were found on all chromosomes except chromosomes 10 and 11.

4.5.1 TPC

TPC was associated with 120 significant SNPs on chromosomes 1, 2, 3, 4, 5, 6, 7, 8, and 9 (Figure 9A1 and 10A1).

Chromosome 1 had five significant SNPs from position 18062779 to 18384219 bp. There was one significant SNP, which was located near LOC_Os01g33000.1, in the promoter region. Four significant SNPs were in the exome region. Two significant SNPs were in each LOC_Os01g32890.1 and another LOC_Os01g33090.1. The three loci encoded expressed protein.

Chromosome 2 had one significant SNP in the exome region at position 32908792 bp., The significant SNP was in LOC_Os02g53740.2 encoding myosin XI-I.

Chromosome 3 had one significant SNP in the promoter region at position 8753058 bp. The significant SNP was located close to LOC_Os03g15860.1 encoding mitochondrial carrier protein.

Chromosome 4 had 55 significant SNPs from position 23374014 to 23749030 bp and 33528752 to 33551098 bp. The promoter region had 27 significant SNPs, which were located close to LOC_Os04g39489.1, LOC_Os04g39510.1 LOC_Os04g39600.1, LOC_Os04g39610.1, LOC_Os04g56240.1, and LOC_Os04g56280.1. These genes encoded, for example, amino acid transporter, fasciclin domain-containing protein, glycerophosphoryl diester phosphodiesterase family protein, and lipase. The exome region had 28 significant SNPs, which were located in LOC_Os04g39280.1 LOC_Os04g39460.1, LOC_Os04g39510.1, LOC_Os04g39540.1 LOC_Os04g39560.1, LOC_Os04g39570.1, LOC_Os04g39629.1, LOC_Os04g39650.1 LOC_Os04g39830.1, and LOC_Os04g39840.1. These genes encoded, for example, ATPOT1, NBS-LRR type disease resistance protein, WRKY35, RNA recognition motif-containing protein, receptor protein kinase CLAVATA1 precursor, and Os4bglu10 - beta-glucosidase.

Chromosome 5 had 29 significant SNPs in the promoter region between position 25670289 to 29326298 bp. These significant SNPs were located close to LOC_Os05g44190.1, LOC_Os05g51100.1, LOC_Os05g51119.1, and LOC_Os05g50920.1

encoding protein such as OsFBX179 - F-box domain-containing protein and transmembrane amino acid transporter protein.

Chromosome 6 had seven significant SNPs in the promoter region from position 4203030 to 4234533 bp. These significant SNPs were located close to LOC_Os06g08510.1, LOC_Os06g08530.1, and LOC_Os06g08550.1. These genes encoded, for example, ubiquitin carboxyl-terminal hydrolase domain-containing protein and BTBN14 - Bric-a-Brac, Tramtrack, Broad Complex BTB domain with non-phototropic hypocotyl 3 NPH3 domain.

Chromosome 7 had 12 significant SNPs from position 6067391 to 6143116 bp and 26553409 to 26553573 bp. The promoter region had seven significant SNPs, which were located close to LOC_Os07g11100.1, LOC_Os07g11120.1, and LOC_Os07g44440.1 encoding hypothetical protein, hydrolase NUDIX family domain-containing protein, and peroxiredoxin, respectively. The exome region had five significant SNPs, which were located in LOC_Os07g11020.1, LOC_Os07g11060.1, LOC_Os07g11070.1, and LOC_Os07g11100.1 encoding, for example, RC - bHLH transcription factor, DUF630/DUF632 domains containing protein, and a hypothetical protein.

Chromosome 8 had eight significant SNPs from position 20675493 to 20856464 bp. The promoter region had one significant SNP, which was located close to LOC_Os08g33200.1 encoding 1-phosphatidylinositol-4-phosphate 5-kinase/ zinc ion binding protein. The exome region had seven significant SNPs, which were located in LOC_Os08g33320.1, LOC_Os08g33340.1, LOC_Os08g33370.2, LOC_Os08g33390.1, and LOC_Os08g33430.1. These genes encoded, for example, RAD23 DNA repair protein, 14-3-3 protein, regulator of chromosome condensation domain-containing protein, and formin.

Chromosome 9 had two significant SNPs in the promoter region from positions 16425550 and 16425614 bp. These significant SNPs were in the LOC_Os09g27020.1 locus, encoding retrotransposon protein.

4.5.2 TFC

TFC was associated with 70 significant SNPs on chromosome 1, 2, 3, 4, 5, 6, 7, 8, and 12 (Figure 9B1 and 10B1) following,

Chromosome 1 had six significant SNPs from position 18062779 to 18387417 bp. The promoter region had two significant SNPs, which were located close to LOC_Os01g33000.1 and LOC_Os01g33400.1 encoding expressed protein and helix-loop-helix DNA-binding domain-containing protein, respectively. The exome region had four significant SNPs. These significant SNPs were located in two loci, LOC_Os01g32890.1 and LOC_Os01g33090.1. Both loci encoded expressed protein.

Chromosome 2 had three significant SNPs at positions 23131824 to 32908792 bp. The promoter region had two significant SNPs, which were located closely to LOC_Os02g38250.1 encoding EMB1745. Another significant SNP was located in LOC_Os02g53740.1 in the exome region encoding myosin XI-I.

Chromosome 3 had one significant SNP in the promoter region at position 8753058 bp. The significant SNP was located close to LOC_Os03g15860.1 encoding mitochondrial carrier protein.

Chromosome 4 found 18 significant SNPs from position 23508984 to 33530279 bp. There was one significant SNP in the promoter region, which was located close to LOC_Os04g56240.1 encoding lipase. The others, 17 significant SNPs in the exome region which were located in LOC_Os04g39460.1 encoding NBS-LRR type disease resistance protein, and the others were in LOC_Os04g39510.1 and LOC_Os04g39560.1. These genes expressed proteins.

Chromosome 5 had five significant SNPs in the promoter region from position 25670289 to 29326298 bp. The significant SNPs were located close to two loci, LOC_Os05g44190.1 and LOC_Os05g51119.1. Both loci encoded expressed protein.

Chromosome 6 had seven significant SNPs in the promoter region from position 4203030 to 4234533 bp. These significant SNPs were located close to LOC_Os06g08510.1, LOC_Os06g08530.1, and LOC_Os06g08550.1. These genes encoded, for example, ubiquitin carboxyl-terminal hydrolase domain-containing protein and BTBN14 - Bric-a-Brac, Tramtrack, Broad Complex BTB domain with non-phototropic hypocotyl 3 NPH3 domain.

Chromosome 7 had 21 significant SNPs from position 6067391 to 6244746 bp. The promoter region had five significant SNPs, which were located close to two loci, LOC_Os07g11100.1 and LOC_Os07g11120.1 encoding hypothetical protein and hydrolase NUDIX family domain-containing protein, respectively. The exome region had 16 significant SNPs, which were located in LOC_Os07g11020.1, LOC_Os07g11060.1, LOC_Os07g11070.1, LOC_Os07g11100.1, LOC_Os07g11250.1, LOC_Os07g11280.1, LOC_Os07g11290.1 LOC_Os07g11320.1, and LOC_Os07g11330.1. These genes encoded, for example, RC - bHLH transcription factor, DUF630/DUF632 domains containing protein, hypothetical protein, pentatricopeptide containing protein, RAL1 - seed allergenic protein, RA5/RA14/RA17 precursor, and RAL2 - seed allergenic protein RA5/RA14/RA17 precursor.

Chromosome 8 had eight significant SNPs from position 20675493 to 20856464 bp. The promoter region had one significant SNP, which was located close to LOC_Os08g33200.1 encoding 1-phosphatidylinositol-4-phosphate 5-kinase/ zinc ion binding protein. The exome region had seven significant SNPs, which were located in LOC_Os08g33320.1, LOC_Os08g33340.1, LOC_Os08g33370.1, LOC_Os08g33390.1, and LOC_Os08g33430.1. These genes encoded, for example, RAD23 DNA repair protein,

14-3-3 protein, regulator of chromosome condensation domain-containing protein, and formin.

Chromosome 12 had one significant SNP in the promoter region at position 26924121 bp. The significant SNP was located close to LOC_Os12g43400.1, encoding expressed protein.

4.5.3 AC

AC was associated with 125 significant SNPs on chromosomes 1, 2, 3, 4, 5, 6, 7, and 8 (Figure 9C1 and 10C1) following,

Chromosome 1 had four significant SNPs in the exome region from position 18062779 to 18384219 bp. The significant SNPs were in two loci, LOC_Os01g32890.1 and LOC_Os01g33090.1. Both loci encoded expressed protein.

Chromosome 2 had three significant SNPs in the exome region from position 32866357 to 32908792 bp. The significant SNPs were in LOC_Os02g53700.1, LOC_Os02g53710.1, and LOC_Os02g53740.1 loci encoding DENN domain-containing protein, expressed protein, and myosin XI-I, respectively.

Chromosome 3 had one significant SNP in the promoter region at position 8753058 bp. The significant SNP was located close to LOC_Os03g15860.1 encoding mitochondrial carrier protein.

Chromosome 4 had 59 significant SNPs at positions 20769381 to 28537266. There were 17 significant SNPs in the promoter region, which were located close to LOC_Os04g39489.1, LOC_Os04g39510.1, LOC_Os04g39600.1, and LOC_Os04g39610.1. These genes encoded, for example, amino acid transporter, fasciclin domain-containing protein, and glycerophosphoryl diester phosphodiesterase family protein. The exome region had 42 significant SNPs, which were located in LOC_Os04g34290.1, LOC_Os04g39260.1, LOC_Os04g39280.1, LOC_Os04g39290.1 LOC_Os04g39460.1, LOC_Os04g39510.1, LOC_Os04g39540.1, LOC_Os04g39560.1 LOC_Os04g39570.1,

LOC_Os04g39629.1, LOC_Os04g39650.1, LOC_Os04g39830.1 LOC_Os04g39840.1, and LOC_Os04g47990.1. These loci encoded, for example, protein kinase, CAF1 family ribonuclease containing protein, ATPOT1, heavy metal transport/detoxification protein, NBS-LRR type disease resistance protein, WRKY35, RNA recognition motif-containing protein, receptor protein kinase CLAVATA1 precursor, Os4bglu10 - beta-glucosidase, and of zinc finger domain-containing protein.

Chromosome 5 had 25 significant SNPs in the promoter region from position 25676784 to 29326298 bp. The significant SNPs were located close to LOC_Os05g44190.1, LOC_Os05g50920.1, LOC_Os05g51100.1, and LOC_Os05g51119.1. These genes encoded, for example, transmembrane amino acid transporter protein and OsFBX179 - F-box domain-containing protein.

Chromosome 6 had eight significant SNPs from position 4203030 to 4234533 bp. The promoter region had seven significant SNPs, which were located close to three loci, LOC_Os06g08510.1, LOC_Os06g08530.1, and LOC_Os06g08550.1. LOC_Os06g08530.1 locus encoded ubiquitin carboxyl-terminal hydrolase domain-containing protein. While, LOC_Os06g08550.1 locus encoded BTBN14 - bric-a-brac, tramtrack, broad complex BTB domain with non-phototropic hypocotyl 3 NPH3 domain. The exome region had one significant SNP at position 26776007 bp. This significant SNP was located in LOC_Os06g44350.1, encoding the hypothetical protein.

Chromosome 7 had 17 significant SNPs from position 6067391 to 6143116 bp and 26552674 to 26553574 bp. The promoter region had 12 significant SNPs, which were located close to three loci, LOC_Os07g11100.1, LOC_Os07g11120.1, and LOC_Os07g44440.1 encoding hypothetical protein, hydrolase NUDIX family domain-containing protein, and peroxiredoxin, respectively. The exome region had five significant SNPs, which were located in LOC_Os07g11020.1, LOC_Os07g11060.1, LOC_Os07g11070.1, and LOC_Os07g11100.1 loci. These loci encoded, for example,

RC - bHLH transcription factor, DUF630/DUF632 domains containing protein and a hypothetical protein.

Chromosome 8 had eight significant SNPs from position 20675493 to 20856464 bp. The promoter region had one significant SNP, which was located close to LOC_Os08g33200.1 locus encoding 1-phosphatidylinositol-4-phosphate 5-kinase/ zinc ion binding protein. The exome region had seven significant SNPs, which were located in LOC_Os08g33320.1, LOC_Os08g33340.1, LOC_Os08g33370.1, LOC_Os08g33390.1, and LOC_Os08g33430.1. These genes encoded, for example, RAD23 DNA repair protein, 14-3-3 protein, regulator of chromosome condensation domain-containing protein, and formin.

4.5.4 QTLs and candidate genes

There were 106 significant SNPs in 38 loci on chromosomes 1 to chromosome 8 that were associated with two or more antioxidant traits and based on the coefficient of determination (R^2) more than 0.1. A total of eight QTLs were identified for antioxidant content and capacity (Figure 11).

The first QTL region was on chromosome 1, named QTAN1_1. This QTL had five significant SNPs, which were located in three loci, LOC_Os01g32890.1, LOC_Os01g33000.1, and LOC_Os01g33090.1 loci. One locus, LOC_Os01g33000.1 was associated with TPC and TFC. Meanwhile, LOC_Os01g32890.1 and LOC_Os01g33090.1 loci were related to TPC, TFC, and AC (Figure 11 and Table 15). The two SNPs in LOC_Os01g32890.1 were explained based on the highest R^2 and accounted for 31.21%, 33.23%, and 25.40% of phenotypic variation in TPC, TFC, and AC, respectively. QTAN1_1 was located at 324 kb upstream from LOC_Os01g33400.1 locus, which encoded the helix-loop-helix DNA-binding domain (Table 4) and 6.99 Mb upstream from *Rd* gene. The Helix-loop-helix DNA-binding domain is a gene regulator, which was a gene involved in controlling the expression of other genes (Massari and Murre, 2000). Although the association between the antioxidant traits and the helix-

loop-helix DNA-binding domain has not been verified nor has the protein from this locus had not been studied for controlling flavonoid structural genes. But, there was reported that the *Rd* gene expresses dihydroflavonol 4-reductase which is an essential enzyme controlling anthocyanin and proanthocyanin biosynthesis (Furukawa et al., 2007).

There was no QTL detected on chromosomes 2 and chromosome 3 because there was only one significant SNP (Chr. 2: LOC_Os02g53740.2 and Chr. 3: LOC_Os03g15860.1) on each chromosome was associated with three traits (Figure 11 and Table 15).

The second QTL region, QTAN4_1 (Figure 11) had 45 significant SNPs in 14 loci on chromosome 4, which were associated with two or three traits of TPC, TFC, and AC. The one SNP near LOC_Os04g39600.1 was shown the highest R^2 and accounted for 11.41% and 13.70% of phenotypic variation in TPC and AC, respectively. The QTL, QTAN4_1 was co-located with LOC_Os04g39470.1, encoding the MYB family transcription factor (Table 4). MYB family transcription factor is one of three gene regulators associated with enhanced antioxidant contents through the transcriptional pathway (Xu et al., 2015). However, the association between the antioxidant traits and this MYB family transcription factor had not been verified, and the protein from this locus had not been studied for controlling flavonoid structural genes (Zheng et al., 2019). In addition, QTAN4_1 was located at 4.2 Mb upstream from *OSB1* (*Ra*) and *OSB2* genes, which had reported that the *OSB1* and *OSB2* genes encoded anthocyanin regulatory Lc protein and Myc-type basic-helix-loop-helix (bHLH) transcription factor of anthocyanin, respectively (Sakamoto et al., 2001).

The third and fourth QTL regions, QTAN5_1 and QTAN5_2 were on chromosome 5 (Figure 11). Two QTLs for the antioxidant traits on chromosome 5 that were consisted of at least two traits of TPC, TFC, and AC, which were represented by four loci (LOC_Os05g44190.1, LOC_Os05g50920.1, LOC_Os05g51100.1,

and LOC_Os05g51119.1). Among these loci, QTAN5_1 in LOC_Os05g44190.1 locus and QTAN5_2 in LOC_Os05g50920.1, LOC_Os05g51100.1, and LOC_Os05g51119.1 loci have been detected (Figure 11). Significant QTL, QTAN5_1 has associated at least two traits of TPC, TFC, and AC. The two SNPs near LOC_Os05g44190.1 were shown the highest R^2, which was explained by 12.25% and 13.03% of the total phenotypic variances for TPC and TFC, respectively. This QTL region was located at 116 kb upstream from LOC_Os05g44320.1 locus, encoding the WD domain, G-beta repeat domain-containing protein, which had reported that is one of three important regulatory factor families related to the biosynthetic pathway (Quattrocchio et al., 2006). QTAN5_2 was located downstream from QTAN5_1 and associated with two or three traits of TPC TFC and AC. The one SNP near LOC_Os05g50920.1 was exhibited the highest R^2, which was explained 28.27% and 29.80% of the total phenotypic variances for TPC and AC, respectively. This region was located at 115 kb upstream and 21 kb downstream from LOC_Os05g51160.1 locus (Myb transcription factor) and LOC_Os05g50900.1 locus (helix-loop-helix DNA-binding protein), respectively (Table 4). Both regulatory factor proteins were from different families but were recruited to start transcription of the antioxidant genes (Feller et al., 2011).

The fifth QTL region, QTAN6_1 (Figure 11) had seven significant SNPs in three loci on chromosome 6 (LOC_Os06g08510.1, LOC_Os06g08530.1, and LOC_Os06g08550.1) were associated with TPC, TFC, and AC (Figure 11). The four SNPs near LOC_Os06g08550.1 were explained the highest R^2 and accounted for 11.93%, 13.79%, and 10.06% of phenotypic variation in TPC, TFC, and AC, respectively. The QTAN6_1 was located at 217 kb and 38 kb downstream from LOC_Os06g08290.1 locus (MYB family transcription factor) and LOC_Os06g08500.1 locus (bHLH transcription factor), respectively (Table 4). Moreover, the *OsC1* gene (LOC_Os06g10350.1) was located at 1.08 Mb downstream from QTAN6_1. This gene encoded an MYB transcription factor that was reported to regulate the flavonoid biosynthesis pathway in the rice leaf. The MYB transcription factor combined with

bHLH and WDR factors to the MYB-bHLH-WDR complex. This complex is the regulating flavonoid biosynthesis protein in rice leaf, which regulated the flavonoid genes including *OsF3H1*, *OsDFR*, and *OsANS1* (Zheng et al., 2019).

The sixth and seventh QTL regions, QTAN7_1 and QTAN7_2 (Figure 11 and Table 4) had 12 significant SNPs on chromosome 7 residing on six loci (LOC_Os07g11020.1, LOC_Os07g11060.1, LOC_Os07g11070.1, LOC_Os07g11100.1, LOC_Os07g11120.1, and LOC_Os07g44440.1). Among these loci, QTAN7_1 in LOC_Os07g11020.1, LOC_Os07g11060.1, LOC_Os07g11070.1, LOC_Os07g11100.1, and LOC_Os07g11120.1, and QTAN7_2 in LOC_Os07g44440.1 have been defined (Figure 11). QTAN7_1 was associated with TPC, TFC, and AC. The two SNPs near LOC_Os07g11100.1 were revealed the highest R^2, which was explained 23.34%, 22.19%, and 24.61% of the total phenotypic variances for TPC, TFC, and AC, respectively. The result indicated that the *Rc* gene (LOC_Os07g11020.1) was located in QTAN7_1. Furukawa et al. (2007) reported that the *Rc* gene encoded bHLH transcription factor regulating proanthocyanidin for controlling proanthocyanidin biosynthesis. Moreover, this QTL was located at 162 kb upstream from the *OsCHS2* gene, which had reported that the *OsCHS2* gene encoded chalcone synthase, the first enzyme of the flavonoid biosynthesis pathway (Shih et al., 2008). QTAN7_2 was located downstream from QTAN7_1 and associated with TPC and AC. The one SNPs near LOC_Os07g44440.1 was shown the highest R^2, which was explained 20.86% and 23.48% of the total phenotypic variances for TPC and AC (Table 4), respectively. This region was located at 195 kb and 8 kb downstream from LOC_Os07g44090.1 locus (Myb-related protein Hv33) and LOC_Os07g44410.1 locus (WD40-like Beta Propeller Repeat family protein), respectively.

Finally, the eighth QTL region, QTAN8_1 had eight significant SNPs in six loci (LOC_Os08g33200.1, LOC_Os08g33320.1, LOC_Os08g33340.1, LOC_Os08g33370.2, LOC_Os08g33390.1, and LOC_Os08g33430.1) and were associated with TPC, TFC, and AC (Figure 11 and Table 4). The one SNP near LOC_Os08g33200.1 was explained the

highest R^2 and accounted for 34.47%, 35.81%, and 32.17% of phenotypic variation in TPC, TFC, and AC (Table 4), respectively. The QTAN8_1 contained two loci (LOC_Os08g33050.1 and LOC_Os08g33150.1) encoding MYB family transcription factors. This QTL was located at 347 kb upstream from LOC_Os08g33660.1 locus (MYB family transcription factor).

There were numerous GWAS studies related to antioxidant traits in rice. Shao et al. (2011) extracted DNA and antioxidants from leaves and seed in 461 rice accessions including 361 white, 50 red, and 6 purple pericarp rice accessions. They used the Q model and Q+K model to study the association between phenotype (TPC, TFC, AC, and colour parameters) and genotype (100 SSR, *frg*, *Rc,* and *Ra* markers). *Rc* and *Ra* (*OSB1*) markers were associated with TPC, TFC, AC, and some colour parameters. This study confirmed that *Rc* and *Ra* genes were major effect loci for antioxidant traits. Another study by Xu et al. (2016) used GWAS to study the association between SNP markers and TPC, TFC, and AC of red and white pericarp rice. The results indicated that 30 QTLs including *Rc* and LOC_Os11g39020 were found with low P-value and high R^2 on chromosomes 1, 6, 7, 9, and 11. Moreover, LOC_Os11g39020 was a new potential gene that related to ferulic acid in whole and white-grain panels.

The antioxidant candidate genes from this study were reported in Table 4. The candidate gene positions were close to the flavonoid synthesis gene. There were genes in the regulation class similar to MYB family transcription factors (7 genes), helix-loop-helix family of transcriptional regulatory proteins (3 genes), and WD-repeat domain proteins family (2 genes). These proteins together form a protein complex that had been studied to be involved in flavonoid biosynthesis. Moreover, *Rc, OsC1, and OsCHS2* were the candidate genes in this study. The *Rc* and *OsC1* genes are the regulatory genes, while *OsCHS2* is a structural gene in the flavonoid biosynthesis pathway (Shih et al., 2008). In the previous studies, the expression of *Rc* and *OsC1* genes was identified. Furukawa et al. (2007) studied the expression of the *Rc* gene

involved in proanthocyanidin synthesis in red and white pericarp rice by using mutation and transgenic method. The results suggested that the dominant allele (*Rc*) in the red pericarp rice was responsible for the accumulation of proanthocyanidin, but the recessive allele (*rc*) in the white pericarp rice was not. In addition, Zheng et al. (2019) studied the *OsC1* gene involved in the anthocyanin biosynthesis pathway of rice leaves. The results showed that *OsC1* regulated the flavonoid biosynthesis pathway together with the bHLH and WDR genes. The protein products of three genes combined to form the MYB-bHLH-WD complexes, which regulated the transcription of structural genes to encoding enzymes for the flavonoid biosynthesis pathway. QTAN7_1 and QTAN6_1 were located closely to *Rc* and *OsC1* genes on chromosomes 7 and 6, respectively. Therefore, it might be possible to use SNP polymorphisms in these QTLs to validate antioxidant activities in rice.

The prior studies of the association between antioxidant traits and markers were reported by Shao et al. (2011) and Xu et al. (2016). Both studies found significant markers on chromosomes 1, 4, 6, 7, 8, 9, 10, and 11. Thus, the candidate genes of antioxidants were included *Ra*, *Rc*, and LOC_Os11g39020. In addition, our study showed significant QTLs on chromosomes 1, 4, 5, 6, 7, and 8. The antioxidant candidate genes were found *Rc*, *OsC1*, *OsCHS2*, MYB family transcription factors genes (7 genes), helix-loop-helix family of transcriptional regulatory proteins (3 genes), and WD-repeat domain proteins family (2 genes). *Rc* gene was the same candidate gene as Shao et al. (2011) and Xu et al. (2016). While, *OsC1, OsCHS2*, MYB family transcription factors genes (7 genes), helix-loop-helix family of transcriptional regulatory proteins (3 genes), and WD-repeat domain proteins family (2 genes) were newly found candidate genes. These results added it more likely to discover antioxidant genes and enhanced the understanding of antioxidant synthesis. For this reason, it can beneficial for future plant breeding.

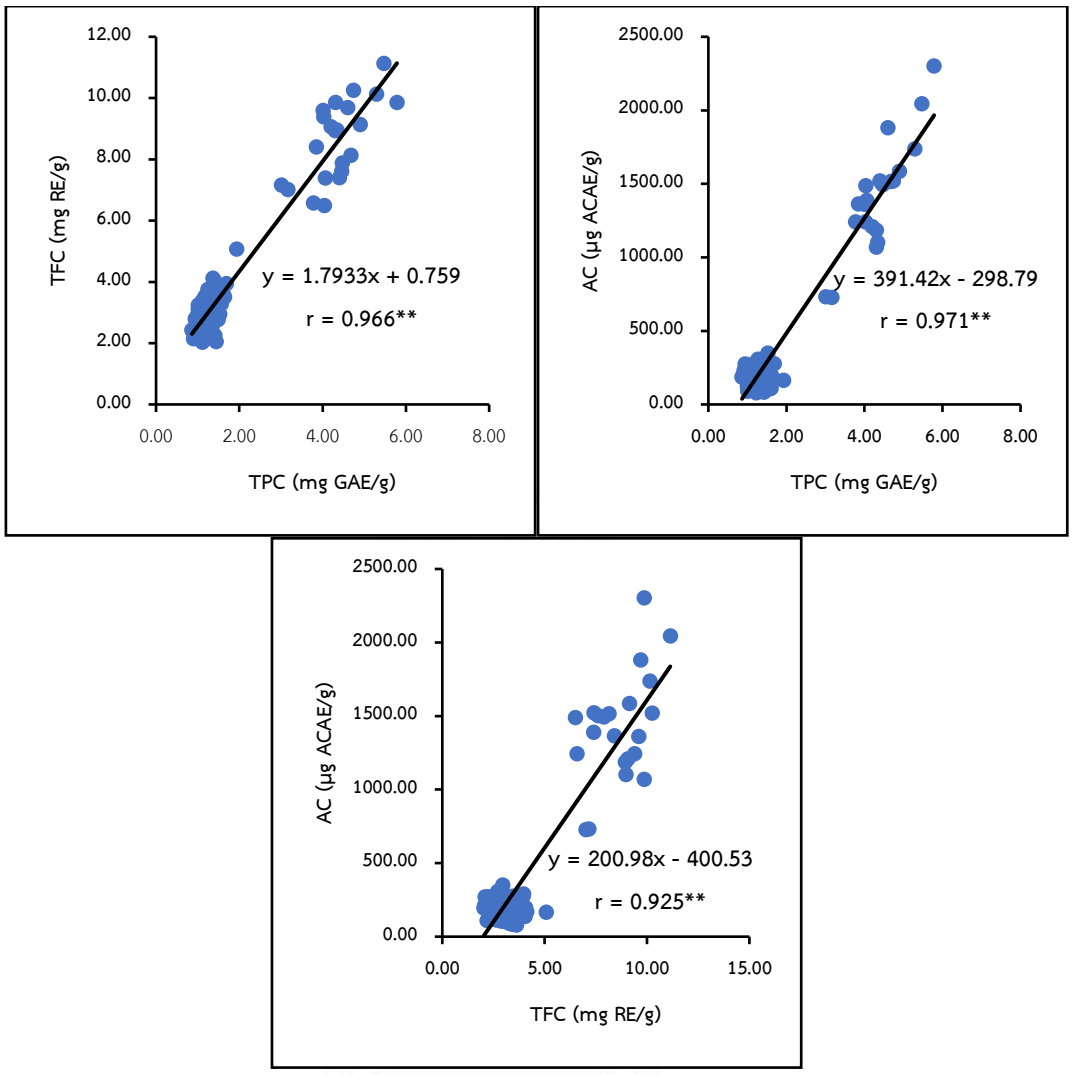


Figure 8 Correlation analysis of antioxidant traits TPC, TFC, and AC in 174 Thai rice cultivars.

Table 3 Genetic variability parameters values for antioxidant traits (TPC, TFC, and AC) in 174 Thai rice cultivars.

Traits	PCV	GCV	H^2
TPC	69.60	63.43	0.83
TFC	60.34	50.41	0.70
AC	131.73	123.79	0.88

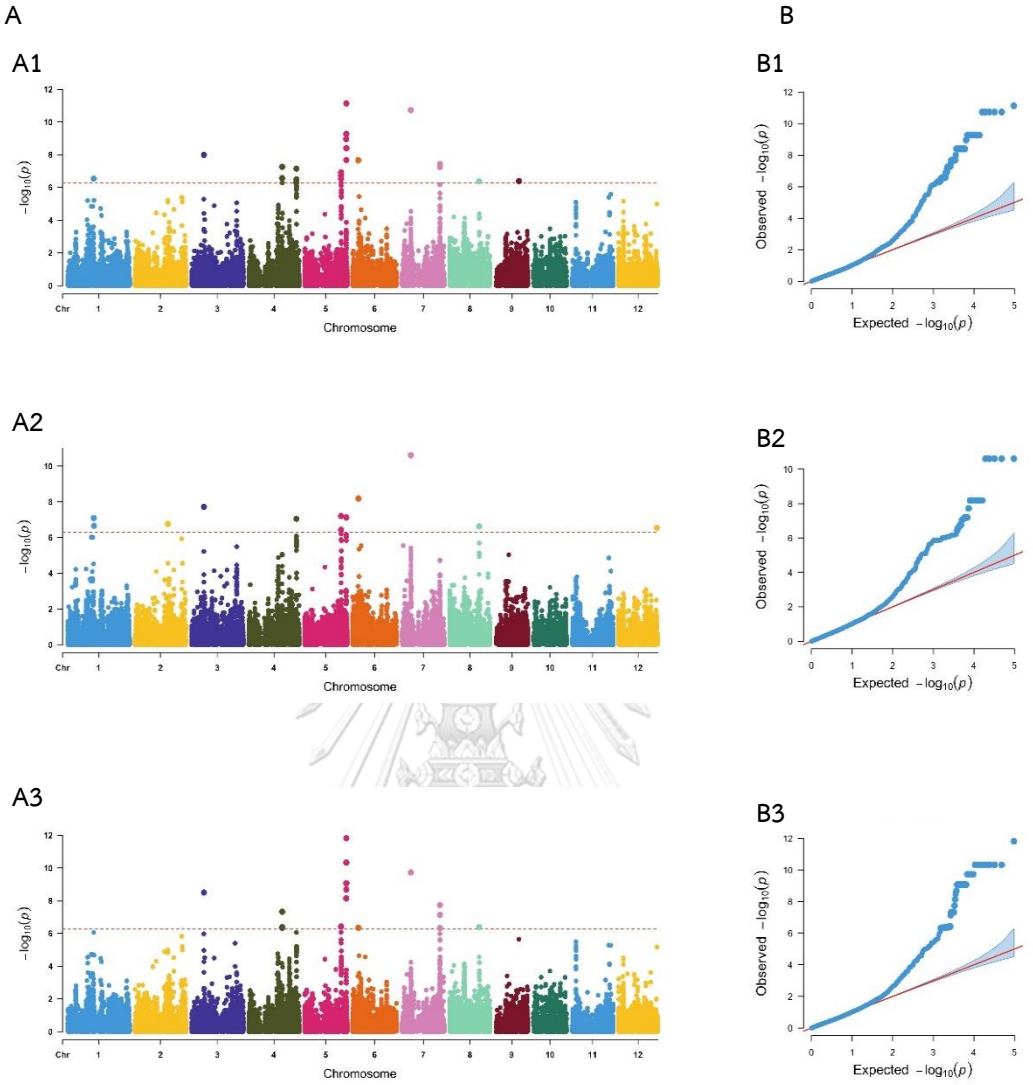


Figure 9 GWAS results for SNPs loci associated with TPC, TFC, and AC in promoter regions.

(A) Manhattan plots from association mapping using a mixed linear model. The x-axis shows the SNP along the 12 Oryza sativa chromosomes. The y-axis shows the negative logarithm (base 10) of the P-value (-log 10(P)). The dash horizontal line represents the significant threshold. A1-A3 showed a plot for TPC, TFC, and AC, respectively. (B) Quantile-Quantile (Q-Q) plot of observed against expected P-value from GWAS analysis. B1-B3 showed plots for TPC, TFC, and AC, respectively.

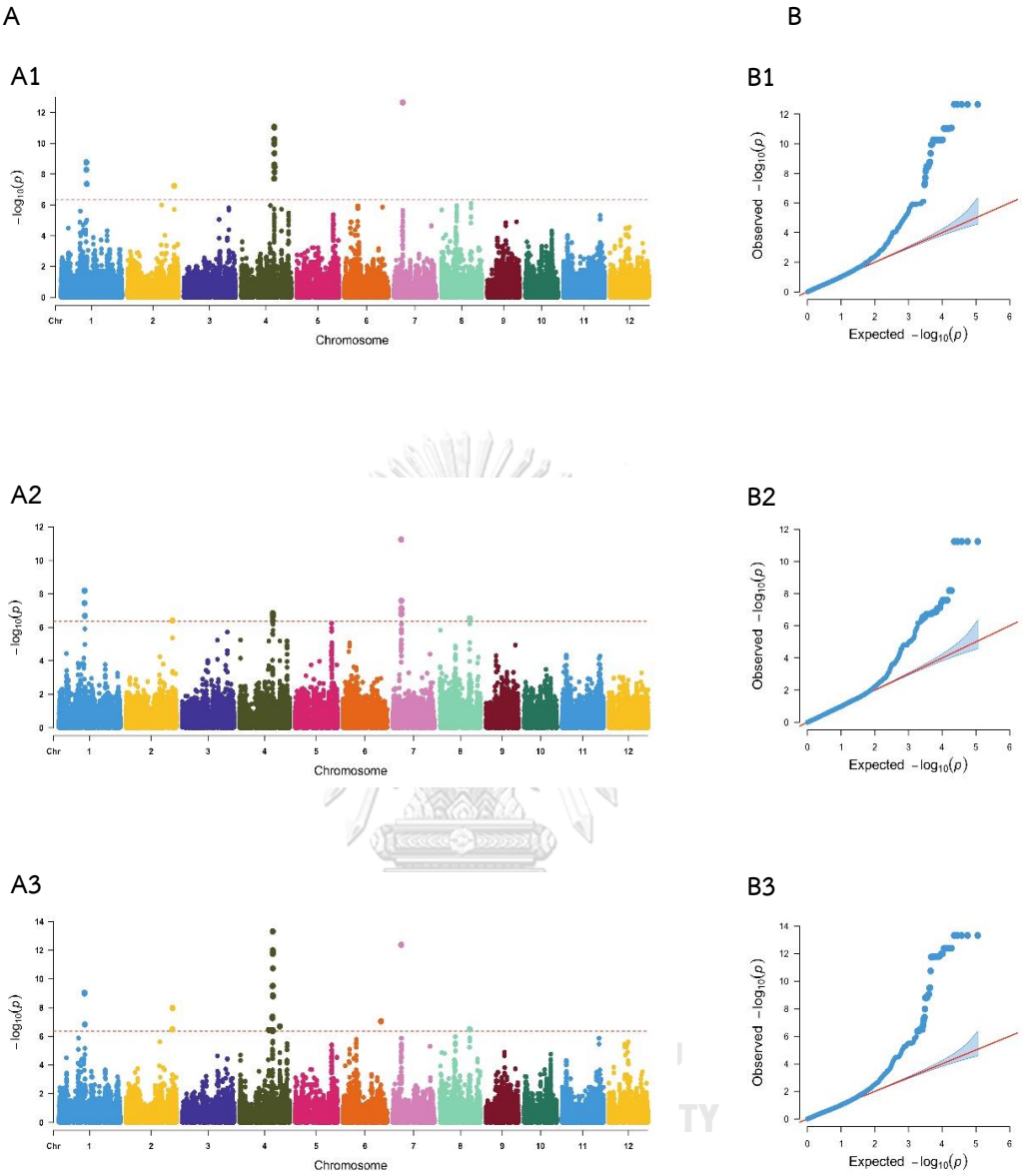


Figure 10 GWAS results for SNPs loci associated with TPC, TFC, and AC in exome regions.

(A) Manhattan plots from association mapping using a mixed linear model. The x-axis shows the SNP along the 12 *Oryza sativa* chromosomes. The y-axis shows the negative logarithm (base 10) of the P-value (-log 10(P)). The dash horizontal line represents the significant threshold. A1-A3 showed a plot for TPC, TFC, and AC, respectively. (B) Quantile-Quantile (Q-Q) plot of observed against expected P-value from GWAS analysis. B1-B3 showed plots for TPC, TFC, and AC, respectively.

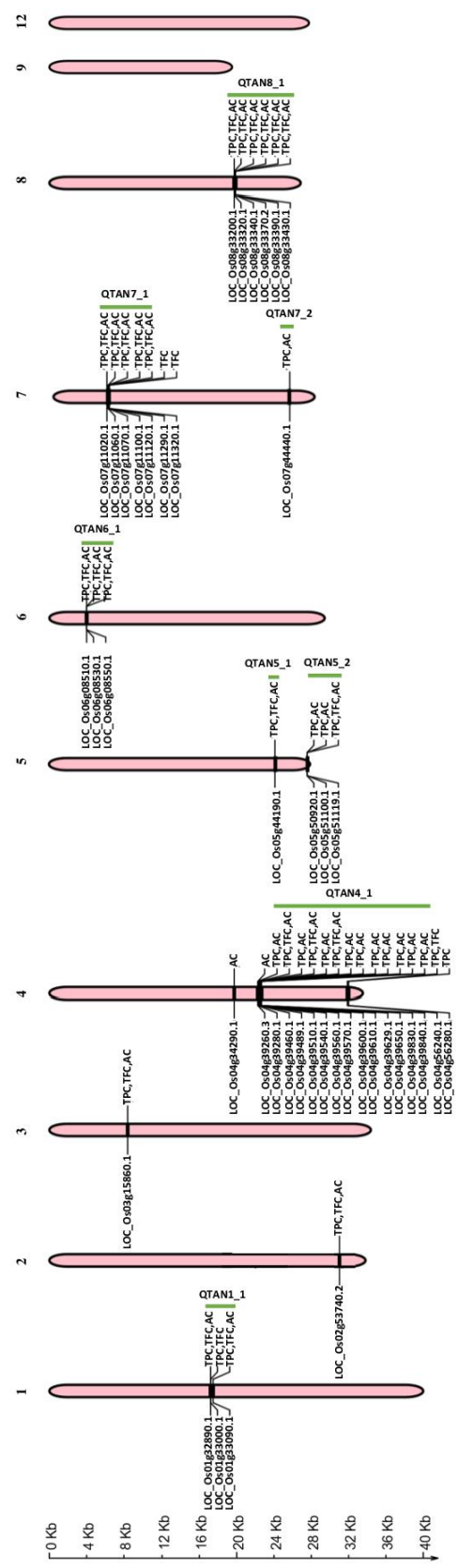


Figure 11 Distribution of 38 loci and 8 QTL positions on chromosome 1, 2, 3, 4, 5, 6, 7, and 8.

Table 4 QTLs and candidate genes by GWAS for TPC, TFC, and AC.

QTLs	Chr.	SNP	Locus*	Trait	P-value	R^2	Candidate genes
QTAN1_1	1	5	LOC_Os01g32890.1	TPC	1.73×10^{-9}	31.21	helix-loop-helix DNA-binding domain (LOC_Os01g33400.1)
				TFC	6.50×10^{-9}	33.23	
				AC	9.31×10^{-10}	25.40	
QTAN4_1	4	45	LOC_Os04g39600.1	TPC	5.45×10^{-8}	11.41	MYB family transcription factor (LOC_Os04g39470.1)
				AC	4.73×10^{-8}	13.70	
QTAN5_1	5	7	LOC_Os05g44190.1	TPC	1.92×10^{-7}	12.25	WD domain, G-beta repeat domain containing protein
				TFC	3.72×10^{-7}	13.03	(LOC_Os05g44320.1)
QTAN5_2	5	22	LOC_Os05g50920.1	TPC	1.11×10^{-9}	28.87	helix-loop-helix DNA-binding protein (LOC_Os05g50900.1)
				AC	7.17×10^{-9}	29.80	Myb transcription factor (LOC_Os05g51160.1)
QTAN6_1	6	7	LOC_Os06g08550.1	TPC	2.17×10^{-8}	11.93	MYB family transcription factor (LOC_Os06g08290.1)
				TFC	6.62×10^{-9}	13.79	BHLH transcription factor (LOC_Os06g08500.1)
				AC	4.48×10^{-7}	10.06	
QTAN7_1	7	10	LOC_Os07g11100.1	TPC	1.87×10^{-11}	23.34	*Rc* (LOC_Os07g11020.1)
				TFC	2.45×10^{-11}	22.19	*OsCHS2* chalcone synthase (LOC_Os07g11440.1)
				AC	1.88×10^{-10}	24.61	

Table 4 (Continued)

QTLs	Chr.	SNP	Locus*	Trait	P-value	R^2	Candidate genes
QTAN7_2	7	2	LOC_Os07g44440.1	TPC	5.63×10^{-8}	20.86	myb-related protein Hv33 (LOC_Os07g44090.1)
				AC	7.24×10^{-8}	23.48	WD40-like Beta Propeller Repeat family protein (LOC_Os07g44410.1)
QTAN8_1	8	8	LOC_Os08g33200.1	TPC	4.23×10^{-7}	34.47	MYB family transcription factor (LOC_Os08g33050.1)
				TFC	2.37×10^{-7}	35.81	MYB family transcription factor (LOC_Os08g33150.1)
				AC	4.07×10^{-7}	32.17	MYB family transcription factor (LOC_Os08g33660.1)

Remark: * The locus had at least one significant SNP that was exhibited the highest R^2 of antioxidant traits.

4.6 Validation of candidate molecular markers for TPC, TFC, and AC.

4.6.1 F_2 population and antioxidant traits.

4.6.1.1 F_1 and F_2 progenies

F_1 progenies were cross-pollinated between RD41 (white pericarp colour) and RB (purple pericarp colour) at Pathumthani Rice Research Center. The 70 plants of F_1 were cultivated to collect F_2 seed. The result showed that 43 F_1 plants had F_2 seeds with light purple colours as shown in Figure 12C. Because they were made from cross-pollination (RD41 × RB), 37 plants had white pericarp seeds due to self-pollination on RD41 plants. The flowering date and plant height (Table 5) of RD41 were ranging from 69-75 days and 79-92 cm, respectively. While, RB were ranging from 98-106 days and 105-118 cm, respectively. Forty-three plants of F_1 were ranging from 61-82 days and 77-109 cm, respectively. F_1 expressed intermediate values of the flowering date and plant height between parents. This result was agreeable with the previous research (Kamara, 2015; Rahman et al., 2013)

The 20 plants of the individual parent, 15 F_1 and 208 F_2 were cultivated to collect seeds, separately for each plant. These seeds were dehulled and analyzed for TPC, TFC, and AC.

Figure 12 Characteristic rice kernels of P_1 (A), P_2 (B), and F_2 (C) with white, purple, and light purple pericarp colours, respectively.

Table 5 Flowering date and plant height comparison of parents and F_1 plant.

Traits	RD41 (P_1)	RB (P_2)	F_1
flowering date	69-75	98-106	61-82
plant height	79-92	105-118	77-109

4.6.1.2 Measurement of the TPC, TFC and AC

The distribution of TPC in 208 F_2 from the cross of RD41 × RB was shown in Figure 13A. The TPC levels of F_2 ranged from 0.68-6.67 mg GAE/g with an average of 2.33±1.23 mg GAE/g. Moreover, the average TPC of P_1, P_2, and F_1 were 1.05±0.13, 4.77±0.43, and 2.26±0.28 GAE/g, respectively.

The distribution of TFC in 208 F_2 was shown in Figure 13B. The TFC of F_2 ranged from 1.92-14.20 mg RE/g with an average of 5.48±2.43 mg RE/g. Moreover, the average TFC of P_1, P_2, and F_1 were 3.32±0.72, 9.83±1.42, and 5.94±0.86 RE/g, respectively.

The distribution of AC in 208 F_2 was shown in Figure 13C. The AC levels of F_2 ranged from 20.88-2,324.52 mg μg AEAC/g with an average of 588.65±453.26 μg AEAC/g. Moreover, the average AC of P_1, P_2, and F_1 were 197.85±48.95, 1,727.37±246.30, and 516.59±126.94 μg AEAC/g, respectively.

The TPC, TFC, and AC traits presented continuous variation in the F_2 population. Moreover, transgressive segregation was found, that is some F_2 plants had higher antioxidant contents and capacity than the parents (Figure 13). The results indicated that these antioxidant traits were controlled by quantitative inheritance. This inheritance was caused by polygenes. These results agreed with total anthocyanin content (TAC) in the report of Mapan et al. (2014). The distribution of TAC was studied in F_2 and F_3 from the cross between Kumdoisaket (purple pericarp rice cultivar) and Pathumthani 1 (white pericarp rice cultivar). The results found that the TAC values of the 298 F_2 population skewed continuously toward Pathumthani 1, and some F_2 plants showed TAC values higher than the parents. Thus, there is a quantitative inheritance with the polygenes. The anthocyanin content is one class of flavonoids in the seed of rice and showed features of high antioxidant capacity in purple pericarp rice. The quantitative inheritance of this trait was confirmed by many studies of anthocyanin content. Therefore, this study confirmed that the antioxidant traits were controlled by quantitative inheritance in TPC, TFC, and AC.

The higher antioxidant content and capacity in the F_2 population than the parents were examined and revealed that it was dues to transgressive segregation which forms new genotypes in the offspring from the recombined parental allele. These traits had additive main effects (Poehlman and Sleper, 1995). This result agreed with Jin et al. (2009) study. TPC, TFC, and AC of 127 doubled haploid lines from the cross of indica rice and japonica rice had transgressive segregation that three traits in the F_2 population exhibited higher and lower value than the parents.

However, some antioxidant genes were reported in many previous studies (Furukawa et al., 2007; Jin et al., 2009; Park et al., 2016; Poehlman and Sleper, 1995; Sakamoto et al., 2001; Shih et al., 2008). The antioxidant biosynthesis was controlled by many structural and regulatory genes. The structural genes produced enzymes for the synthesis of antioxidants. While the regulatory genes produced proteins for controlling the transcription of structural genes (Park et al., 2016; Shih et al., 2008). The *Rd* gene is a structural gene and encodes DFR protein, which is the first enzyme in the pathway of proanthocyanidin and anthocyanin biosynthesis (Furukawa et al., 2007). However, *Rc*, *OSB1, and OSB2* genes are regulatory genes of proanthocyanidin and anthocyanin in a biosynthesis pathway. The *Rc* genes encode the protein bHLH transcription factor for the structural genes of proanthocyanidin biosynthesis in red pericarp rice (Furukawa et al., 2007). In addition, other genes, *OSB1* and *OSB2* have been shown to encode Lc protein and Myc-type basic-helix-loop-helix (bHLH) regulating the anthocyanin level in rice (Sakamoto et al., 2001).

4.6.1.3 Correlation

The pairwise correlations among TPC, TFC, and AC in 208 F_2 were shown in Figure 14. High positive significant correlations ($r>0.900$) were detected among these antioxidant traits. Pearson correlation coefficients ranged from 0.910 to 0.953. The results suggested that TPC and TFC in this population displayed high antioxidant capacity. This agreed with Umnajkitikorn and Faiyue (2016) study. The relationship of TPC and AC was studied in geminated rice of purple pericarp (Kum Doi Saket). The

result showed that TPC and AC had a significantly strong positive correlation (r=0.844, p<0.01). Jin et al. (2009) studied antioxidant compounds and antioxidant activity in double haploid lines from the cross between “Zhai-Ye-Qing 8” and “Jing-Xi”. The results showed that the correlation coefficients between TPC–TFC and TPC-AC were 0.581 and 0.446, respectively. Both relationships were moderate. The highly positive correlation among these antioxidant traits implied that high TPC appeared with high TFC and AC in rice (Shen et al., 2009).

4.6.1.4 PCV, GCV, H^2, and %GA

The genetic variation parameters of 208 F_2 were calculated from variance equations using the variance for antioxidant traits of P_1, P_2, F_1, and F_2 populations (Appendix F). First, phenotypic ($\boldsymbol{\sigma}_p^2$), genotypic ($\boldsymbol{\sigma}_g^2$), and environmental ($\boldsymbol{\sigma}_e^2$) variances were determined. Three parameters were used to calculate PCV, GCV, and H^2 (Table 6). The PCV of TPC, TFC, and AC was 52.77, 44.13, and 76.81, respectively, and the GCV of TPC, TFC, and AC were 51.20, 40.08, and 71.98, respectively. The results showed a very low difference between the PCV and GCV for all traits. The difference between the PCV and GCV implied the environmental effects. The high Broad-sense heritabilities of TPC, TFC, and AC in this study were 0.94, 0.82, and 0.88, respectively, and the high %GA of TPC, TFC, and AC were 102.57, 75.16, and 127.38, respectively. These results suggested that these antioxidant traits were under a high genetic influence with the additive gene action (Mehboob et al., 2016; Roy and Shil, 2020). Roy and Shil (2020) performed the cross-pollination between many rice cultivars. The plant height, 1000 seed weight, and harvested time of F_3 and F_5 offsprings were measured. High H^2 and %GA were found and explained that all traits in this rice population had high genetic effects with additive action.

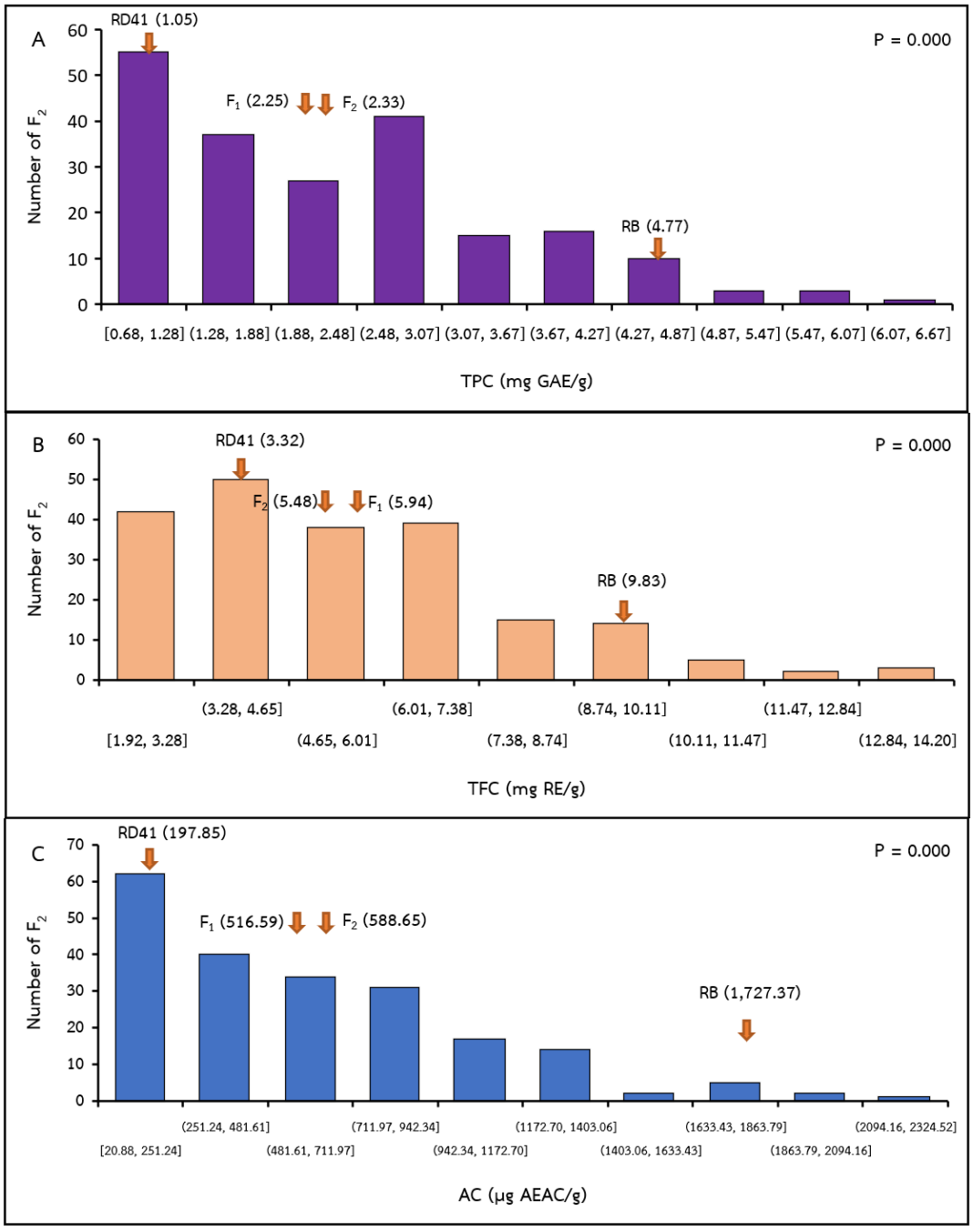


Figure 13 Frequency distribution of (A) TPC, (B) TFC, and (C) AC of 208 F_2 from the cross of RD41 × RB.

The arrows indicate the average antioxidant traits of P_1 (RD41), P_2 (RB), F_1, and F_2 populations.

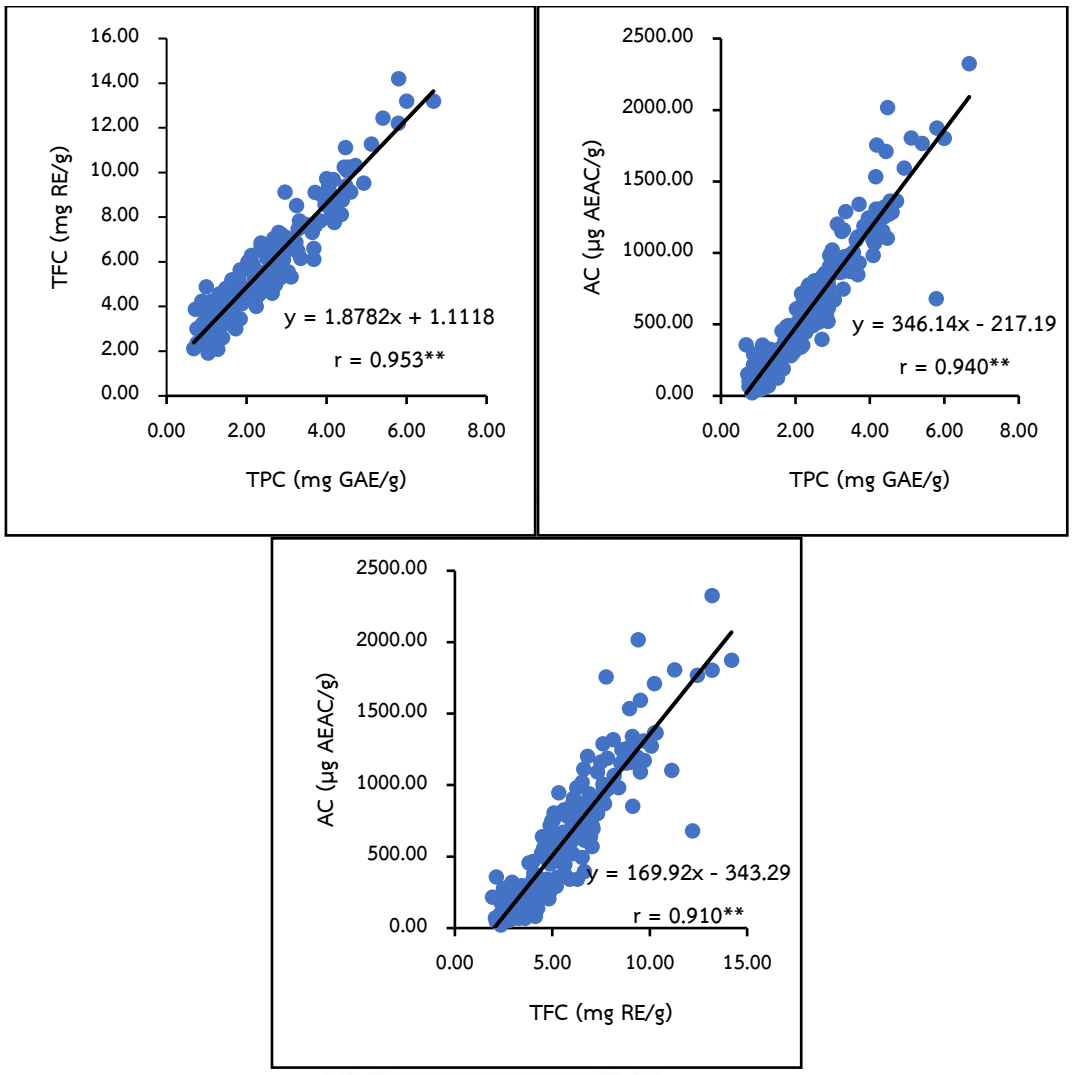


Figure 14 Correlation analysis among antioxidant traits, TPC, TFC, and AC of 208 F_2 from the cross between RD41 × RB.

Table 6 Genetic variability parameters for the antioxidant traits (TPC, TFC, and AC) of 208 F_2 from the cross between RD41 × RB.

Traits	PCV	GCV	H^2	%GA
TPC	52.77	51.20	0.94	102.57
TFC	44.13	40.08	0.82	75.16
AC	76.81	71.98	0.88	127.38

4.6.2 DNA sequencing

The polymorphisms of 17 significant SNPs in 16 loci on rice chromosomes 1, 4, 5, 6, 7, and 8 were selected from 106 significant SNPs in 38 loci. The 17 primers of the SNP border region were designed to detect the SNP profiles (Table 13). The PCR products of the parental DNA (P_1 and P_2) from these primers were examined by sager sequencing. The DNA sequences from each primer were compared between the parents. The results indicated that 14 SNPs showed the same DNA base. Thus, these SNPs were expressed as monomorphic DNA profiles. However, the three significant SNPs showed off polymorphic profiles between P_1 and P_2 DNA bases (Figure 15). The first polymorphic SNP genotypes, CC (P_1), and TT (P_2) are located in LOC_Os01g33090.1 at position 18384219 on chromosome 1 (Figure 15A). The second polymorphic SNP genotypes, AA (P_1), and GG (P_2) are located in LOC_Os06g08550.1 at position 4234533 on chromosome 6 (Figure 15B). The third polymorphic SNP genotypes, TT (P_1), and GG (P_2) are located on LOC_Os06g08530.1 at position 4219208 on chromosome 6 (Figure 15C). These polymorphic SNPs were used to genotype the F_2 population by HRM analysis in the next step.

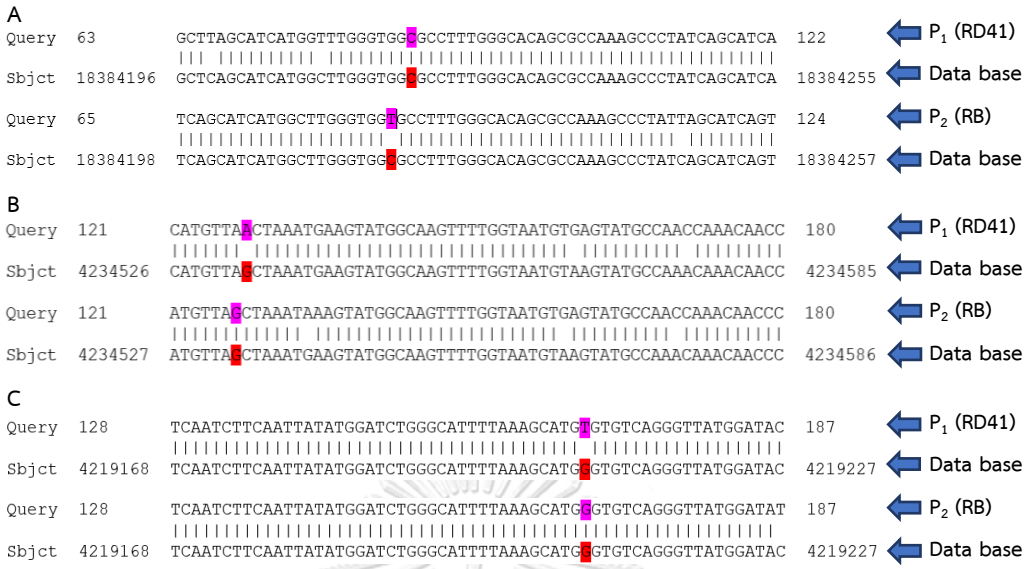


Figure 15 Sequence alignment of three polymorphic SNPs comparison between RD41 (P_1) and RB (P_2) DNA.

(A) SNP profile at position 18384219, (B) SNP profile at position 4234533, and (C) SNP profile at position 4219208.

4.6.3 Genotyping analysis

The polymorphic SNPs of RD41 and RB were found at positions 18384219, 4234533, and 4219208. The three primer sets of these positions were designed and shown in Table 14. Genotyping polymorphisms of the F_2 population were performed by HRM analysis. The amplifications were produced using a CFX96 real-time PCR system with SsoFast™ EvaGreen® Supermix. The melt curve shape and melting temperature of different genotypes were identified by Precision Melt Analysis ™ software.

The SNP at position 18384219 was determined for polymorphisms among P_1, P_2, and F_1 by sanger sequencing showing genotype CC, TT, and CT for P_1, P_2, and F_1, respectively. HRM analysis was used to amplify DNA and provide the melting curve and melting temperature to explore the genotype of P_1, P_2, and F_1. The results showed that the difference curve for this SNP was not grouped and was shown the

error and unstable of an imperfect reannealing of the amplicons (Figure 16A). These graphs exhibited the different patterns of reannealing profiles for the same amplicons. This phenomenon was caused by the PCR products localizing on the AT or GC-rich regions or short repeats (Figure 16B). Thus, 208 F_2 could not be analyzed for SNP genotype at this position.

The SNP at position 4234533 was determined for polymorphisms among P_1, P_2, and F_1 by sanger sequencing showing genotype AA, GG, and AG for P_1, P_2, and F_1, respectively. The PCR products of this SNP in P_1, P_2, and F_1 were amplified and provided melting curves and melting temperatures by HRM analysis. These PCR products were 98 base pairs. The results were shown in Figure 17. The amplicons clustered into one group and could not be identified for the difference among AA, GG, and AG (Figure 17A). As a result of there were two SNPs in the PCR products (A/G and reversed G/A). Thus, 208 F_2 could not be analyzed for SNP genotype at this position.

The SNP at position 4219208 was determined for polymorphisms among P_1, P_2, and F_1 by sanger sequencing showing genotype TT, GG, and TG for P_1, P_2, and F_1, respectively. The PCR products of this SNP in P_1, P_2, F_1, and 208 F_2 were amplified and provided melting curves and melting temperatures, which were normalized (Figure 18A). The different curves of genotyping data from HRM analysis were shown in Figure 18B. The red line group was represented by genotype TT of P_1 (RD41). In addition, the green and blue lines were symbolized by genotype TG and GG of F_1, and P_2 (RB), respectively. The red lines were the baseline of differential thermal analysis for classifying other curves of different genotypes, and the different curves in Figure 18B were established three clusters of genotypes. Then, the DNA of 208 F_2 plants was genotyped by the same HRM condition of P_1, P_2, and F_1. The F_2 amplicons were clustered similarly melted profiles into three groups by the auto-grouping function in Figure 18C. The three groups of the F_2 were shown genotype TT, TG, and GG in similarity groups of RD41, F_1, and RB, respectively. The proportion of genotype

TT: TG: GG was 54: 99: 55. This result was analyzed by Chi-square test and regression analysis in the next step.

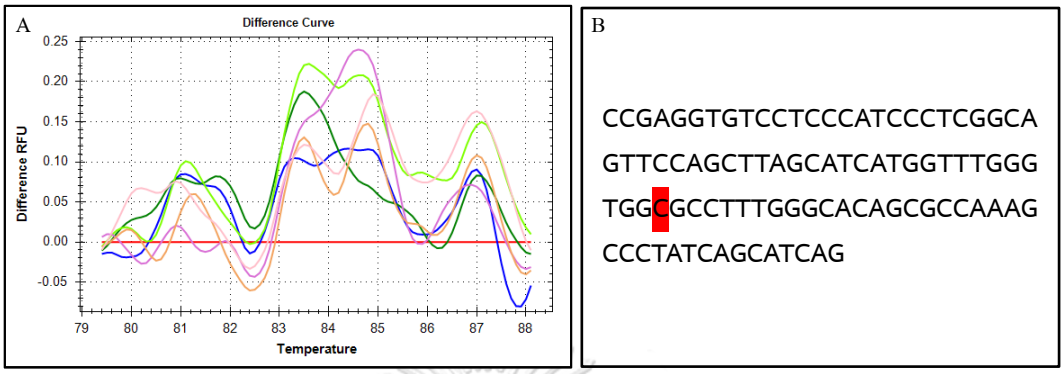


Figure 16 HRM analysis results for the genotyping of SNP at position 18384219. A) difference curve and B) DNA sequence.
The red highlight was the target SNP.

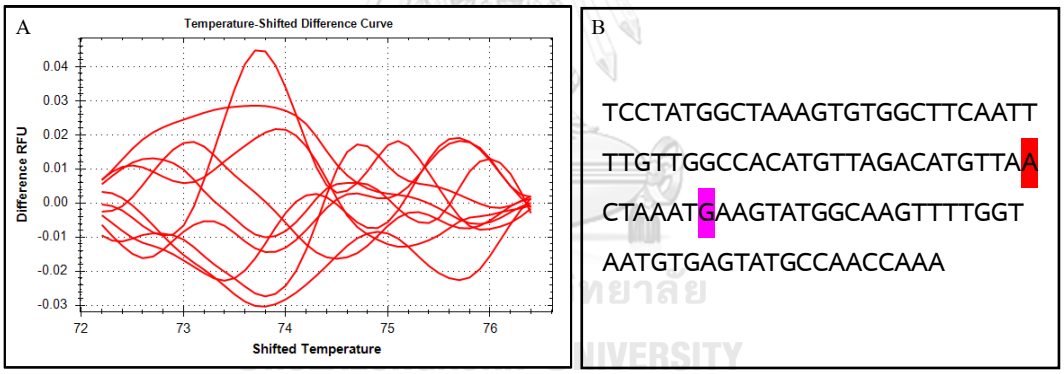
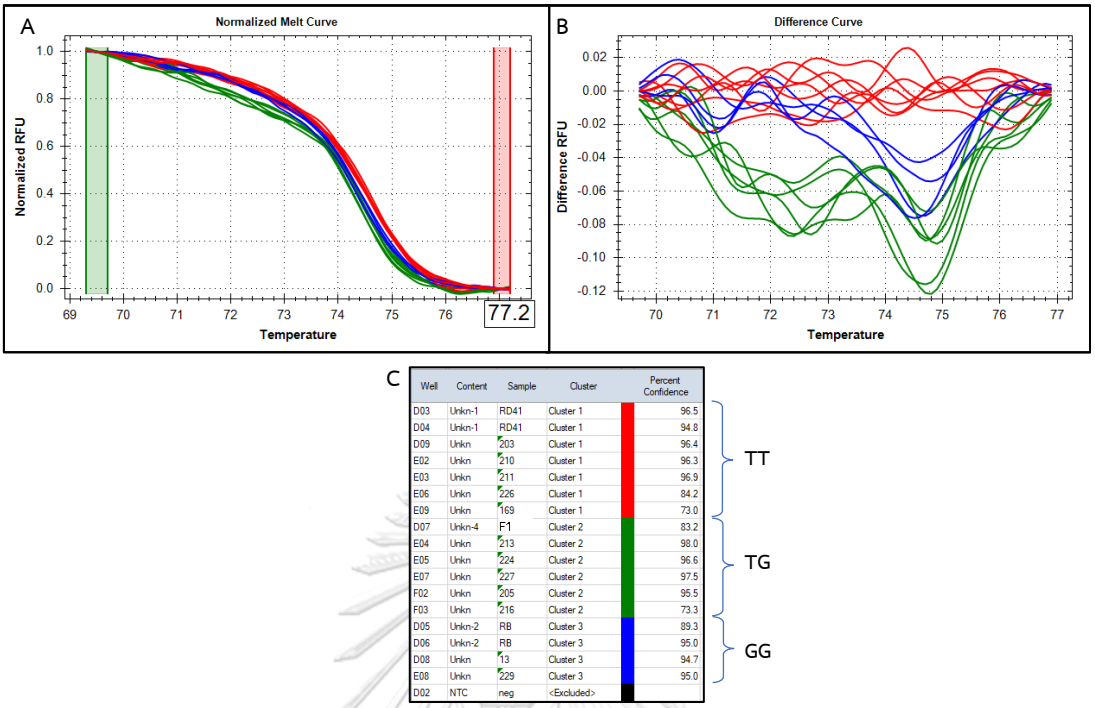


Figure 17 HRM analysis results for the genotyping of SNP at position 4234533. A) difference curve and B) DNA sequence.
The red highlight was the target SNP, and the purple highlight was another SNP in the amplicon.



Well	Content	Sample	Cluster	Percent Confidence
D03	Unkn-1	RD41	Cluster 1	96.5
D04	Unkn-1	RD41	Cluster 1	94.8
D09	Unkn	203	Cluster 1	96.4
E02	Unkn	210	Cluster 1	96.3
E03	Unkn	211	Cluster 1	96.9
E06	Unkn	226	Cluster 1	84.2
E09	Unkn	169	Cluster 1	73.0
D07	Unkn-4	F1	Cluster 2	83.2
E04	Unkn	213	Cluster 2	98.0
E05	Unkn	224	Cluster 2	96.6
E07	Unkn	227	Cluster 2	97.5
F02	Unkn	205	Cluster 2	95.5
F03	Unkn	216	Cluster 2	73.3
D05	Unkn-2	RB	Cluster 3	89.3
D06	Unkn-2	RB	Cluster 3	95.0
D08	Unkn	13	Cluster 3	94.7
E08	Unkn	229	Cluster 3	95.0
D02	NTC	neg	<Excluded>	

Figure 18 HRM analysis results for the genotyping of SNP at position 4219208 using Precision Melt Analysis software.

(A) melt profile of the TT genotype (red lines), TG genotype (green lines), and GG (blue lines) after normalization; (B) difference curve of genotyping data and (C) discrimination of sequence variation by clustering similar melt profiles into the same group.

4.6.4 Segregation and regression analysis

The Chi-square test was performed for SNP at position 4219208 to determine the segregation ratio whether was under the null hypothesis of 1: 2: 1 ratio in the F_2 population. The proportion of TT: TG: GG was 54: 99: 55, showing a Chi-square value of approximately 0.490 (p>0.5). This result was failed to reject the null hypothesis (H_0). Thus, the ratio of SNP genotypes at position 4219208 in 208 F_2 plants was exhibited 1: 2: 1. Accordingly, HRM analysis was a sensitive and specific method for the molecular diagnosis of an SNP genotype (Ceri et al., 2019).

The association between SNP at position 4219208 and the antioxidant traits (TPC, TFC, and AC) was analyzed using simple linear regression analysis in the R program (R Core Team, 2020). The results showed no relationship between the SNP genotypes at position 4219208 and each antioxidant trait from TPC, TFC, and AC traits. The coefficient of determination (R^2) was explained as low values 8.62×10^{-3}%, 3.51×10^{-2}%, and 7.10×10^{-3}% for TPC, TFC, and AC, respectively. Moreover, the results from these associations were shown no significant difference of average TPC, TFC, and AC between the SNP genotypes (TT, TG, and GG) by DMRT (Figure 19). Therefore, no association between the SNP at this position and the antioxidant traits. This SNP marker could not be used as a molecular marker to select for antioxidant content in rice.

In this study, there was no association between the SNP genotypes and the traits, so the R^2 exhibited very low values. In addition, Schrag et al. (2009) study the association between AFLP markers and the yield content traits (seed yield and seed dry matter content) in hybrid maize. The results showed the relationship between AFLP markers and these two traits. Regression analysis confirmed the high significance R^2, 46%, and 70% supported for seed yield and seed dry matter content, respectively. These molecular markers showed high efficiency predicted for the yield content in hybrid maize. This could help to improve greatly the efficiency of the

commercial hybrid breeding program. Thus, the markers for marker-assisted selection should be demonstrated the high-value R^2.

In this study, no relationships between the SNP genotypes and the antioxidant traits were found, which may be affected by the excessive distance between the genetic positions of the SNP marker and one of the antioxidant genes. Both genetic positions were not linked, and each genetic position was shown independent inheritance and assortment (Snustad and Simmons, 2006). However, the relationship between the SNP genotypes and antioxidant traits was detected by using another position of retrieving SNP from GWAS or studying the relationship in the other site of genomic DNA.

However, few molecular antioxidant markers were reported in previous studies. The first marker, the *Ra* marker was reported by Lim and Ha (2013). This marker was located in *Ra* (*OSB1*) genes and used to identify the purple and white pericarp rice. This polymorphism was shown in the digestion of PCR products with the *Bam*HI site in purple pericarp rice. The second maker was studied by Furukawa et al. (2007). This marker was located in the *Rc* gene at the seventh exon and used to classified the difference PCR amplicons of red and white pericarp rice with 14-bp deletion. The third maker was in *Rd* (*OsDFR*) gene, which was exhibited the two different DNA bands of coloured and white pericarp rice (Lim and Ha, 2013). These genes were transcribed and translated to enzymes and proteins for antioxidant biosynthesis in the flavonoid pathway (Shih et al., 2008), and the antioxidant traits and colour of pericarp rice seed were shown the significant relationship (Pramai and Jiamyangyuen, 2016; Shen et al., 2009). These molecular markers were utilized to identify antioxidant content and capacity among coloured (red and purple) and white pericarp rice. However, the GWAS results in this study found many candidate antioxidant genes. These candidate genes should examine the polymorphic sequences of DNA and find the relationship with antioxidant traits. The accurate antioxidant markers will be generated from that DNA polymorphic sequences

associated with antioxidant traits and will use to select these traits in plant breeding programs. The antioxidant contents and capacities in the seed of rice were regulated by polygenes. As a result of the antioxidant genes accumulation, the rice seed was indicated high antioxidation. Thus, the selection for high antioxidants in rice seed was completed by pyramiding antioxidant genes. The discovery of the other relationship between markers and antioxidant genes may increase the likelihood of successful selection.

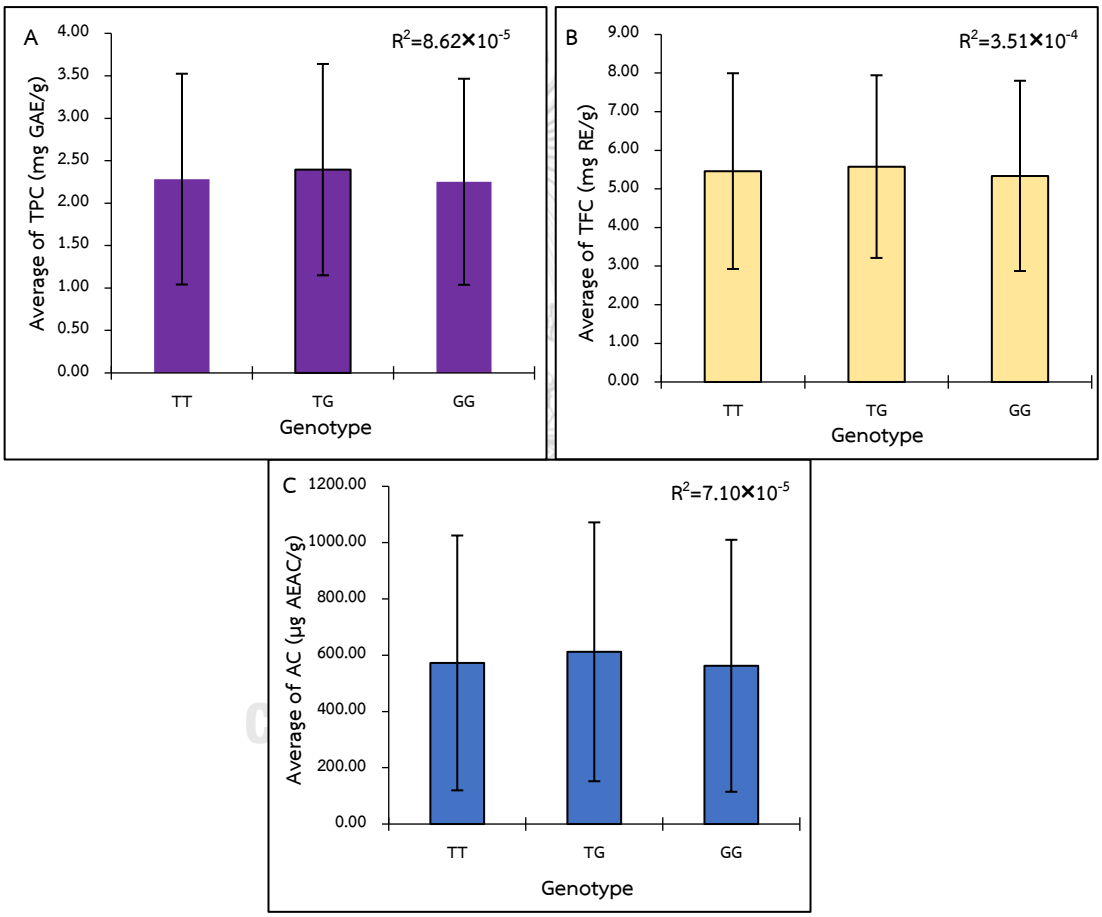


Figure 19 Average of antioxidant traits (A) TPC, (B) TFC, and (C) AC and the coefficient of determination (R^2)* of 208 F_2 population

Remark: *The R^2 was analyzed by regression analysis between antioxidant traits and the genotype at position 4219208.

CHAPTER V

Conclusion

The study on the quantitative trait loci analysis for antioxidant traits in rice *Oryza sativa* L. had four purposes: 1) to analyze the total phenolic content (TPC), total flavonoid content (TFC), and the antioxidant capacity (AC), 2) to conduct the genome-wide association study (GWAS) on antioxidant traits of Thai rice cultivars, 3) to estimate the genetic parameters (PCV, GCV, H^2, and %GA) of TPC, TFC, and AC in 208 F_2 from the cross between RD41 × RB, and 4) to study the relationship between single nucleotide polymorphisms (SNPs) and TPC, TFC, and AC in 208 F_2 from the cross of RD41 × RB. The findings were concluded as follows.

1) One hundred and seventy-four Thai rice cultivars, comprising 152 white, 10 red, and 12 purple pericarp cultivars were studied and the results revealed that the white pericarp cultivars exhibited the lower TPC, TFC, and AC values than the other coloured cultivars. All antioxidant traits had a highly positive significant correlation ($r>0.925$, $p<0.01$). The hierarchical cluster analysis (HCA) classified the rice cultivars into four clusters. Cluster I and II comprised white rice while cluster III and IV comprised coloured rice. Cluster I and II comprised of 72 and 80 white pericarp cultivars. respectively. Cluster I had the lowest antioxidant contents and capacity, but cluster II showed higher TFC values than cluster I. Cluster III comprised of six red and ten purple pericarp cultivars. The TPC, TFC, and AC values of this cluster were higher than cluster I and II, but were lower than cluster IV. Finally, cluster IV comprised four red and two purple pericarp cultivars, which had a high TPC, TFC, and AC.

2) The genome-wide association study (GWAS) was performed by genome-wide efficient mixed-model association (GEMMA) software which employed the mixed linear model (MLM) using 209594 SNPs and the antioxidant traits of 159 Thai

rice cultivars. GWAS results were identified eight QTLs on the chromosome 1, 4, 5, 6, 7, and 8. There were 2 genes (*Rc* and *OsCHS2*) and 12 candidate genes (7 MYB family transcription factors, 3 helix-loop-helix family of transcriptional regulatory proteins and 2 WD-repeat domain protein families). In addition, a QTL on chromosome 6 was closely located to *OsC1*.

3) In the 208 F_2 population from the cross between RD4 × RB, the histogram of TPC, TFC and AC were continuous but had a non-normal distribution. Moreover, some of the F_2 had higher or lower antioxidants than the parents because of transgressive segregation. The antioxidant traits were quantitative inheritance controlled by polygenes. High positive correlations between TPC, TFC, and AC were detected. The PCV, GCV, H^2, and %GA had a high value. The difference between PCV and GCV was low. The antioxidant traits in this population were expressed by genetic components with the additive gene action.

4) The SNP at position 4219208 in LOC_Os06g08530.1 locus had three genotypes (TT, TG, and GG) with the normal segregation ratio of 1: 2: 1 (TT: TG: GG), but the associations between genotype profiles and three antioxidant traits in 208 plants of F_2 were low. Accordingly, LOC_Os06g08530.1 encoding ubiquitin carboxyl-terminal hydrolase domain-containing protein was not related to the antioxidant gene; this SNP position could not become a molecular marker for antioxidant selection.

Two other approaches might be successful in finding the molecular markers for antioxidant traits. First, the candidate genes of low and high antioxidant rice groups should study the expression of genes by genetic engineering, and differently expressed candidate genes will be found in the different DNA sequences. These sequences can use for suitable molecular markers. Second, the eighty-nine significant SNPs should examine the polymorphic profiles between low and high antioxidant rice cultivars, and the polymorphic SNPs will be determined the association of

antioxidants in mapping populations for identifying the molecular markers to select high antioxidant traits in the breeding program.

REFERENCES

Allard, R. W. (1999). *Principles of plant breeding* (2nd ed.). New York: J. Wiley.

Awika, J. M., Rooney, L. W., Wu, X., Prior, R. L. and Cisneros-Zevallos, L. (2003). Screening methods to measure antioxidant activity of sorghum (*Sorghum bicolor*) and sorghum products. *Journal of Agricultural and Food Chemistry.* 51: 6657-6662. doi:10.1021/jf034790i

Babu, P. D., Subhasree, R. S., Bhakyaraj, R. and Vidhyalakshmi, R. (2009). Brown rice-beyond the color reviving a lost health food- a review. *American-Eurasian Journal of Agronomy.* 2: 67-72.

Bhattacharjee, P., Singhal, R. S. and Kulkarni, P. R. (2002). Basmati rice: a review. *International Journal of Food Science & Technology.* 37: 1-12. doi:10.1046/j.1365-2621.2002.00541.x

Brody, T. (2016). Chapter 19 - Biomarkers. In T. Brody (Ed.), *Clinical Trials (Second Edition)* (pp. 377-419). Boston: Academic Press.

Bureau of Rice Research and Development. (2020). Rice Knowledge Bank: RD41. Rice Department, Ministry of Agriculture and Cooperatives, Bangkok, Thailand. Retrieved from http://www.ricethailand.go.th/rkb3/title-index.php-file=content.php&id=121.htm (30/11/2020)

Butsat, S. and Siriamornpun, S. (2010). Antioxidant capacities and phenolic compounds of the husk, bran and endosperm of Thai rice. *Food Chemistry.* 119: 606-613. doi:10.1016/j.foodchem.2009.07.001

Ceri, A., Pavic, M., Horvat, I., Radic Antolic, M. and Zadro, R. (2019). Development and validation of a rapid method for genotyping three P-selectin gene polymorphisms based on high resolution melting analysis. *Journal of Clinical Laboratory Analysis.* 33: 1-8. doi:10.1002/jcla.22698

Deng, G. F., Xu, X. R., Zhang, Y., Li, D., Gan, R. Y. and Li, H. B. (2013). Phenolic compounds and bioactivities of pigmented rice. *Critical reviews in food science and nutrition.* 53: 296-306. doi:10.1080/10408398.2010.529624

Falcone Ferreyra, M. L., Rius, S. P. and Casati, P. (2012). Flavonoids: biosynthesis,

biological functions, and biotechnological applications. *Frontiers in plant science.* 3: 1-12. doi:10.3389/fpls.2012.00222

Farrar, J. S. and Wittwer, C. T. (2017). Chapter 6 - High-resolution melting curve analysis for molecular diagnostics. In G. P. Patrinos (Ed.), *Molecular Diagnostics (Third Edition)* (pp. 79-102): Academic Press.

Feller, A., Machemer, K., Braun, E. L. and Grotewold, E. (2011). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *The Plant journal.* 66: 94-116. doi:10.1111/j.1365-313X.2010.04459.x

Furukawa, T., Maekawa, M., Oki, T., Suda, I., Iida, S., Shimada, H., Takamure, I. and Kadowaki, K.-I. (2007). The *Rc* and *Rd* genes are involved in proanthocyanidin synthesis in rice pericarp. *The Plant Journal.* 49: 91-102. doi:10.1111/j.1365-313x.2006.02958.x

Garritano, S., Gemignani, F., Voegele, C., Nguyen-Dumont, T., Le Calvez-Kelm, F., De Silva, D., Lesueur, F., Landi, S. and Tavtigian, S. V. (2009). Determining the effectiveness of high resolution melting analysis for SNP genotyping and mutation scanning at the *TP53* locus. *BMC Genetics.* 10: 1-12. doi:10.1186/1471-2156-10-5

Ghasemzadeh, A., Karbalaii, M. T., Jaafar, H. Z. E. and Rahmat, A. (2018). Phytochemical constituents, antioxidant activity, and antiproliferative properties of black, red, and brown rice bran. *Chemistry Central Journal.* 12: 1-13. doi:10.1186/s13065-018-0382-9

Globerson, D., Genizi, A. and Staub, J. E. (1987). Inheritance of seed weight in *Cucumis sativus* (L.) var. *sativus* and var. *hardwickii* (Royle) Kitamura. *Theoretical and Applied Genetics.* 74: 522-526. doi:10.1007/BF00289832

Goffman, F. D. and Bergman, C. J. (2004). Rice kernel phenolic content and its relationship with antiradical efficiency. *Journal of the Science of Food and Agriculture.* 84: 1235-1240. doi:10.1002/jsfa.1780

Goufo, P. and Trindade, H. (2014). Rice antioxidants: phenolic acids, flavonoids, anthocyanins, proanthocyanidins, tocopherols, tocotrienols, γ-oryzanol, and phytic acid. *Food Science & Nutrition.* 2: 75-104. doi:10.1002/fsn3.86

Gunaratne, A., Wu, K., Li, D., Bentota, A., Corke, H. and Cai, Y. Z. (2013). Antioxidant activity and nutritional quality of traditional red-grained rice varieties containing proanthocyanidins. *Food Chemistry.* 138: 1153-1161. doi:10.1016/j.foodchem.2012.11.129

Herald, T. J., Gadgil, P. and Tilley, M. (2012). High-throughput micro plate assays for screening flavonoid content and DPPH-scavenging activity in sorghum bran and flour. *Journal of the science of food and agriculture.* 92: 2326-2331. doi:10.1002/jsfa.5633

Hernández-Rodríguez, P., Baquero, L. P. and Larrota, H. R. (2019). Chapter 14 - Flavonoids: potential therapeutic agents by their antioxidant capacity. In M. R. S. Campos (Ed.), *Bioactive Compounds* (pp. 265-288): Woodhead Publishing.

Hosoda, K., Sasahara, H., Matsushita, K., Tamura, Y., Miyaji, M. and Matsuyama, H. (2018). Anthocyanin and proanthocyanidin contents, antioxidant activity, and in situ degradability of black and red rice grains. *Asian-Australasian Journal of Animal Sciences.* 31: 1213-1220. doi:10.5713/ajas.17.0655

International Business Machines. (2013). *IBM SPSS Statistics for Windows, Version 22.0*. Armonk, New York: IBM Corp.

Jairin, J., Phengrat, K., Teangdeerith, S., Vanavichit, A. and Toojinda, T. (2007). Mapping of a broad-spectrum brown planthopper resistance gene, *Bph3*, on rice chromosome 6. *Molecular Breeding.* 19: 35-44. doi:10.1007/s11032-006-9040-3

Jin, L., Xiao, P., Lu, Y., Shao, Y., Shen, Y. and Bao, J. (2009). Quantitative trait loci for brown rice color, phenolics, flavonoid contents, and antioxidant capacity in rice grain. *Cereal Chemistry Journal.* 86: 609-615. doi:10.1094/CCHEM-86-6-0609

Johnson, H. W., Robinson, H. F. and Comstock, R. E. (1955). Estimates of genetic and environmental variability in soybeans[1]. *Agronomy Journal.* 47: 314-318. doi:10.2134/agronj1955.00021962004700070009x

Jun, H.-I., Song, G.-S., Yang, E.-I., Youn, Y. and Kim, Y.-S. (2012). Antioxidant activities and phenolic compounds of pigmented rice bran extracts. *Journal of Food Science.* 77: 759-764. doi:10.1111/j.1750-3841.2012.02763.x

Kamara, N. (2015). *Genetic analysis of agronomic traits in Oryza sativa X O. sativa Cross.* Master's Thesis, Crop and soil sciences, Graduate school, Kwame

Nkrumah University of Science and Technology.

Kazemzadeh, M., Safavi, S. M., Nematollahi, S. and Nourieh, Z. (2014). Effect of brown rice consumption on inflammatory marker and cardiovascular risk factors among overweight and obese non-menopausal female adults. *International Journal of Preventive Medicine.* 5: 478-488.

Lee, J.-S., Sreenivasulu, N., Hamilton, R. S. and Kohli, A. (2019). Brown rice, a diet rich in health promoting properties. *Journal of Nutritional Science and Vitaminology.* 65: 26-28. doi:10.3177/jnsv.65.S26

Lekklar, C., Pongpanich, M., Suriya-arunroj, D., Chinpongpanich, A., Tsai, H., Comai, L., Chadchawan, S. and Buaboocha, T. (2019). Genome-wide association study for salinity tolerance at the flowering stage in a panel of rice accessions from Thailand. *BMC Genomics.* 20: 1-18. doi:10.1186/s12864-018-5317-2

Li, N., Zheng, H., Cui, J., Wang, J., Liu, H., Sun, J., Liu, T., Zhao, H., Lai, Y. and Zou, D. (2019). Genome-wide association study and candidate gene analysis of alkalinity tolerance in japonica rice germplasm at the seedling stage. *Rice.* 12: 1-12. doi:10.1186/s12284-019-0285-y

Lim, S.-H. and Ha, S.-H. (2013). Marker development for the identification of rice seed color. *Plant Biotechnology Reports.* 7: 391-398. doi:10.1007/s11816-013-0276-1

Mammadov, J., Aggarwal, R., Buyyarapu, R. and Kumpatla, S. (2012). SNP markers and their impact on plant breeding. *International Journal of Plant Genomics.* 2012: 1-11. doi:10.1155/2012/728398

Mapan, P., Thipkan, S., Thebault Prom-u-thai, C., Kaladee, D. and Jamjod, S. (2014). Early generation selection for high anthocyanin and photoperiod insensitivity in F_2 population between Kumdoisaket and Pathumthani 1. *Naresuan Phayao Journal.* 7: 160-171.

Massari, M. E. and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Molecular and cellular biology.* 20: 429-440. doi:10.1128/mcb.20.2.429-440.2000

Mehboob, A., Mazhar, I., Zishan, G., Bilal Ahmad, K., Aneela, K., Maria, S. and Nayyer iqbal, K. (2016). Genetic analysis of F_2 population of tomato for quantitative traits in the cross Bushbeef X Nagina. *Advances in Plants & Agriculture Research.*

4: 422-424. doi:10.15406/apar.2016.04.00162

Miller, P. A., Williams, J. C., Robinson, H. F. and Comstock, R. E. (1958). Estimates of genotypic and environmental variances and covariances in upland cotton and their implications in selection[1]. *Agronomy Journal.* 50: 126-131. doi:10.2134/agronj1958.00021962005000030004x

Milligan, G. W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika.* 45: 325-342. doi:10.1007/BF02293907

Mulugeta, S., Sentayehu, A. and Kassahun, B. (2012). Genetic variability, heritability, correlation coefficient and path Analysis for yield and yield related traits in upland rice (*Oryza sativa* L.). *Journal of Plant Sciences.* 7: 13-22. doi:10.3923/jps.2012.13.22

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: 7-14. doi:10.1111/nyas.12540

Nigro, D., Laddomada, B., Mita, G., Blanco, E., Colasuonno, P., Simeone, R., Gadaleta, A., Pasqualone, A. and Blanco, A. (2017). Genome-wide association mapping of phenolic acids in tetraploid wheats. *Journal of Cereal Science.* 75: 25-34. doi:10.1016/j.jcs.2017.01.022

Park, S., Choi, M. J., Lee, J. Y., Kim, J. K., Ha, S. H. and Lim, S. H. (2016). Molecular and biochemical analysis of two rice flavonoid 3'-hydroxylase to evaluate their roles in flavonoid biosynthesis in rice grain. *International Journal of Molecular Sciences.* 17: 1-14. doi:10.3390/ijms17091549

Peanparkdee, M. and Iwamoto, S. (2019). Bioactive compounds from by-products of rice cultivation and rice processing: Extraction and application in the food and pharmaceutical industries. *Trends in Food Science & Technology.* 86: 109-117. doi:doi.org/10.1016/j.tifs.2019.02.041

Poehlman, J. M. and Sleper, D. A. (1995). *Breeding field crops* (4th ed.). Iowa: Iowa State University Press.

Poosri, S., Thilavech, T., Pasukamonset, P., Suparpprom, C. and Adisakwattana, S. (2019). Studies on Riceberry rice (*Oryza sativa* L.) extract on the key steps related to

carbohydrate and lipid digestion and absorption: A new source of natural bioactive substances. *NFS Journal.* 17: 17-23. doi:10.1016/j.nfs.2019.10.002

Pramai, P. and Jiamyangyuen, S. (2016). Chemometric classification of pigmented rice varieties based on antioxidative properties in relation to color. *Songklanakarin Journal of Science & Technology.* 38: 463-472. doi:10.14456/sjst-psu.2016.62

Qi, X., Wen, Y., Li, P., Liang, C., Cheng, B., Ma, M., Cheng, S., Zhang, L., Liu, L., Kafle, O. P. and Zhang, F. (2020). An integrative analysis of genome-wide association study and regulatory SNP annotation datasets identified candidate genes for bipolar disorder. *International Journal of Bipolar Disorders.* 8: 1-7. doi:10.1186/s40345-019-0170-z

Quattrocchio, F., Baudry, A., Lepiniec, L. and Grotewold, E. (2006). The regulation of flavonoid biosynthesis. In E. Grotewold (Ed.), *The Science of Flavonoids* (pp. 97-122). New York: Springer.

R Core Team. (2020). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.r-project.org/index.html (30/11/2020)

Rahman, M. M., Lee, K. E., Lee, E. S., Matin, M. N., Lee, D. S., Yun, J. S., Kim, J. B. and Kang, S. G. (2013). The genetic constitutions of complementary genes *Pp* and *Pb* determine the purple color variation in pericarps with cyanidin-3-O-glucoside depositions in black rice. *Journal of Plant Biology.* 56: 24-31. doi:10.1007/s12374-012-0043-9

Ravichanthiran, K., Ma, Z. F., Zhang, H., Cao, Y., Wang, C. W., Muhammad, S., Aglago, E. K., Zhang, Y., Jin, Y. and Pan, B. (2018). Phytochemical profile of brown rice and its nutrigenomic implications. *Antioxidants (Basel).* 7: 1-16. doi:10.3390/antiox7060071

Re, R., Pellegrini, N., Proteggente, A., Pannala, A., Yang, M. and Rice-Evans, C. (1999). Antioxidant activity applying an improved ABTS radical cation decolorization assay. *Free Radical Biology and Medicine.* 26: 1231-1237. doi:10.1016/S0891-5849(98)00315-3

Robinson, H. F., Comstock, R. E. and Harvey, P. H. (1949). Estimates of heritability and the degree of dominance in corn[1]. *Agronomy Journal.* 41: 353-359.

doi:10.2134/agronj1949.00021962004100080005x

Roy, S. C. and Shil, P. (2020). Assessment of genetic heritability in rice breeding lines based on morphological traits and caryopsis ultrastructure. *Scientific Reports.* 10: 1-17. doi:10.1038/s41598-020-63976-8

Saibabu, V., Fatima, Z., Khan, L. A. and Hameed, S. (2015). Therapeutic potential of dietary phenolic acids. *Advances in Pharmacological Sciences.* 2015: 1-10. doi:10.1155/2015/823539

Sakamoto, W., Ohmori, T., Kageyama, K., Miyazaki, C., Saito, A., Murata, M., Noda, K. and Maekawa, M. (2001). The *purple leaf* (*Pl*) locus of rice: the Pl^{w} allele has a complex organization and includes two genes encoding basic helix-loop-helix proteins involved in anthocyanin biosynthesis. *Plant and Cell Physiology.* 42: 982-991. doi:10.1093/pcp/pce128

Sanghamitra, P., Sah, R. P., Bagchi, T. B., Sharma, S. G., Kumar, A., Munda, S. and Sahu, R. K. (2018). Evaluation of variability and environmental stability of grain quality and agronomic parameters of pigmented rice (*O. sativa* L.). *Journal of Food Science and Technology.* 55: 879-890. doi:10.1007/s13197-017-2978-9

Schrag, T. A., Möhring, J., Maurer, H. P., Dhillon, B. S., Melchinger, A. E., Piepho, H.-P., Sørensen, A. P. and Frisch, M. (2009). Molecular marker-based prediction of hybrid performance in maize using unbalanced data from multiple experiments with factorial crosses. *Theoretical & Applied Genetics.* 118: 741-751. doi:10.1007/s00122-008-0934-9

Shao, Y. and Bao, J. (2015). Polyphenols in whole rice grain: genetic diversity and health benefits. *Food Chemistry.* 180: 86-97. doi:10.1016/j.foodchem.2015.02.027

Shao, Y., Hu, Z., Yu, Y., Mou, R., Zhu, Z. and Beta, T. (2018). Phenolic acids, anthocyanins, proanthocyanidins, antioxidant activity, minerals and their correlations in non-pigmented, red, and black rice. *Food Chemistry.* 239: 733-741. doi:10.1016/j.foodchem.2017.07.009

Shao, Y., Jin, L., Zhang, G., Lu, Y., Shen, Y. and Bao, J. (2011). Association mapping of grain color, phenolic content, flavonoid content and antioxidant capacity in dehulled rice. *Theoretical and Applied Genetics.* 122: 1005-1016. doi:10.1007/s00122-010-1505-4

Shastry, B. S. (2007). SNPs in disease gene mapping, medicinal drug development and evolution. *Journal of Human Genetics.* 52: 871-880. doi:10.1007/s10038-007-0200-z

Shastry, B. S. (2009). SNPs: impact on gene function and phenotype. In A. A. Komar (Ed.), *Single Nucleotide Polymorphisms: Methods and Protocols* (pp. 3-22). Totowa, New Jersey: Humana Press.

Shen, Y., Jin, L., Xiao, P., Lu, Y. and Bao, J. (2009). Total phenolics, flavonoids, antioxidant capacity in rice grain and their relations to grain color, size and weight. *Journal of Cereal Science.* 49: 106-111. doi:10.1016/j.jcs.2008.07.010

Shih, C. H., Chu, H., Tang, L. K., Sakamoto, W., Maekawa, M., Chu, I. K., Wang, M. and Lo, C. (2008). Functional characterization of key structural genes in rice flavonoid biosynthesis. *Planta.* 228: 1043-1054. doi:10.1007/s00425-008-0806-1

Snustad, D. P. and Simmons, M. J. (2006). *Principles of genetics* (4th ed.): John Wiley & Sons.

Statistical Analysis Software Institute Inc. (2010). *Using JMP 9* (3rd ed.). North Carolina: SAS®.

Sun, Q., Spiegelman, D., van Dam, R. M., Holmes, M. D., Malik, V. S., Willett, W. C. and Hu, F. B. (2010). White rice, brown rice, and risk of type 2 diabetes in US men and women. *Archives of internal medicine.* 170: 961-969. doi:10.1001/archinternmed.2010.109

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G. and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics.* 20: 467-484. doi:10.1038/s41576-019-0127-1

Tanaka, Y., Sasaki, N. and Ohmiya, A. (2008). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *The Plant journal.* 54: 733-749. doi:10.1111/j.1365-313X.2008.03447.x

Tao, Y., Niu, Y., Wang, Y., Chen, T., Naveed, S. A., Zhang, J., Xu, J. and Li, Z. (2018). Genome-wide association mapping of aluminum toxicity tolerance and fine mapping of a candidate gene for *Nrat1* in rice. *PLOS ONE.* 13: e0198589. doi:10.1371/journal.pone.0198589

Thitipramote, N., Pradmeeteekul, P., Nimkamnerd, J., Chaiwut, P., Pintathong, P. and

Thitilerdecha, N. (2016). Bioactive compounds and antioxidant activities of red (Brown red jasmine) and black (Kam leum pua) native pigmented rice. *International Food Research Journal.* 23: 410-414.

Turner, S. D. (2014). Qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv.* 1-2. doi:10.1101/005165

Umnajkitikorn, K. and Faiyue, B. (2016). Enhancing antioxidant properties of germinated Thai rice (*Oryza sativa* L.) cv. Kum doi saket with salinity. *Rice Research: Open Access.* 1: 103-110. doi:10.4172/jrr.1000103

Xu, F., Bao, J., Kim, T.-S. and Park, Y.-J. (2016). Genome-wide association mapping of polyphenol contents and antioxidant capacity in whole-grain rice. *Journal of Agricultural and Food Chemistry.* 64: 4695-4703. doi:10.1021/acs.jafc.6b01289

Xu, W., Dubos, C. and Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB–bHLH–WDR complexes. *Trends in Plant Science.* 20: 176-185. doi:10.1016/j.tplants.2014.12.001

Zhang, M. W., Zhang, R. F., Zhang, F. X. and Liu, R. H. (2010a). Phenolic profiles and antioxidant activity of black rice bran of different commercially available varieties. *Journal of Agricultural and Food Chemistry.* 58: 7580-7587. doi:10.1021/jf1007665

Zhang, Q., Zhang, J., Shen, J., Silva, A., Dennis, D. A. and Barrow, C. J. (2006). A simple 96-well microplate method for estimation of total polyphenol content in seaweeds. *Journal of Applied Phycology.* 18: 445-450. doi:10.1007/s10811-006-9048-4

Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., Bradbury, P. J., Yu, J., Arnett, D. K., Ordovas, J. M. and Buckler, E. S. (2010b). Mixed linear model approach adapted for genome-wide association studies. *Nature Genetics.* 42: 355-360. doi:10.1038/ng.546

Zhao, J., Yang, W., Zhang, S., Yang, T., Liu, Q., Dong, J., Fu, H., Mao, X. and Liu, B. (2018). Genome-wide association study and candidate gene analysis of rice cadmium accumulation in grain in a diverse rice collection. *Rice.* 11: 61-75. doi:10.1186/s12284-018-0254-x

Zheng, J., Wu, H., Zhu, H., Huang, C., Liu, C., Chang, Y., Kong, Z., Zhou, Z., Wang, G., Lin,

Y. and Chen, H. (2019). Determining factors, regulation system, and domestication of anthocyanin biosynthesis in rice leaves. *New Phytologist.* 223: 705-721. doi:10.1111/nph.15807

Zhou, X. and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics.* 44: 821-824. doi:10.1038/ng.2310

Zhou, X. and Stephens, M. (2014). Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods.* 11: 407-409. doi:10.1038/nmeth.2848

Appendix A

Table 7 Lists 174 rice cultivars according to the colours of the rice pericarp.

Colours of rice pericarp	Number of cultivars	Cultivars
white	152	Aew mod daeng, Ai tai, Beu saw mi, Bua noi, Bun mah, Chai nat 1, Chaw ma gawk, Chiang phatthalung, Chumpae 60, CSSL 11, Daeng nah, Daw dawk mai, Daw sahm deuan, Dawk khah, Dawk mai, Di si, E-khao yai, E-lai, E-mum, E-puang, Gaen jan, Gam pai, Gow ruang 88, Gra dahd, Gwian hak, Gwian hak, Gwian hak, Hah ruang, Hahng mah nai, Hahng nahk, Hahng yi 71, Hantra 60, In paeng, jae gan, Jampah tawng, Jao daeng, Jao dawk khao, Jao khao, Jao rahk haeng, Jek chuey, Jud mawn, Kaset daw, Khai mod rin, Khao ah-gahd, Khao bahn pod, Khao dawk mali 105, Khao gaew, Khao gaw diaw, Khao gon jud, Khao kod, Khao luang, Khao luang, Khao nuan, Khao praguad, Khao puang, Khao sa nguan, Khao samer, Khao samer, Khao supan, Khao taeng mo, Khao tah cheua, Khao tah haeng 17, Khao tah jeua, Khiaw hahng mah, Khitom khao, Koo meuang, Lao taek, Leb nok, Leuang bai jaek, Leuang bai lod, Leuang chumpae, Leuang dong, Leuang glahng, Leuang huan, Leuang kamin, Leuang kwai lah, Leuang kwai lah, Leuang ngahm, Leuang noi 31-1-39, Leuang pahk chong, Leuang plah gim, Leuang pratew, Leuang puang tawng, Leuang tah yang, Leuang tawng, Look daeng pattani, Look daeng pattani, Luang pratahn, Luang pratahn, Ma fai, Ma gawk, Ma yom, Ma yom, Mahk nam, Mahk yom, Med makham, Muay hin,

Table 7 (Continued)

Colours of rice pericarp	Number of cultivars	Cultivars
		Nah khawan, Nahng mon, Nahng mon s-4, Nahng ngahm, Nahng nuan, Nam sagui 19, Nam sagui 19, Niaw khao, Niaw mali, Niaw prae 1, Pathum thani 1, Pathum thani 60, Pawng aew, Pawng aew, Pin gaew 56, Plah khaeng, Plah sew, Pra in, Prachin buri 1, Pratahn ban bung, Puang hahng moo, Puang hahng nahk, Puang nahk, Puang tawng, Puang tawng, Puang wahn, Rahk haeng, RD10, RD17, RD19, RD21, RD27, RD31, RD41, RD8, RD9, Ruang diaw, Sahm ruang, Sai yud, Sam ahang, Sao nueng, Setti, Sew mae jan, Sew mae jan, Soon, Suphan buri 60, Ta pow gaew 161, Ta pow lom, Tah bahn, Tah jeua, Tawng rahk sai, Tom meuang luang, Toon chalawng, UBN, U-ta pow
red	10	Chaw pli khao, Hawm daeng, Hawm mali daeng, Lai mahk, Loi hah ruang, Nok kum, Rathu heenati, RD69, Red rose rice, Sang yod
purple	12	Dam dahng, Gam feuang, Gam liaw, Gam nahng payah, Gam noi, Hawm dam, Hawm nin, Khao' gam, Khao hawm mea payah tawng dam, Mali dam, Niaw dam lai, Riceberry

Appendix B

Chemical preparation

1. 1% (v/v) hydrochloric acid in methanol

990 mL of methanol (Honeywell, USA) was mixed with 10 mL hydrochloric acid (Merck, Germany).

2. 10% (v/v) Folin-Ciocalteu reagent

10 mL of Folin-Ciocalteu reagent (Merck, Germany) was diluted in 900 mL of deionized water.

3. 7.5% (w/v) sodium carbonate

7.5150 g of sodium carbonate (Ajax Finechem, Australia) was dissolved and adjusted the volume to 100 mL with deionized water in a volumetric flask.

4. Stock standard (12.5 mg/mL) gallic acid

0.3125 g of gallic acid (Bio Basic, Canada) was dissolved and made up the volume to 25 mL with methanol in a volumetric flask.

5. 50% methanol

75 mL of methanol was mixed with 75 mL deionized water.

6. 10% aluminum chloride

10.2040 g of aluminum chloride (Sigma-Aldrich, Germany) was dissolved and made up the volume to 100 mL with 50% methanol in a volumetric flask.

7. 5% (w/v) sodium nitrate

2.5773 g of sodium nitrate (Sigma-Aldrich, Germany) was dissolved and made up the volume to 50 mL with deionized water in a volumetric flask.

8. 1 M NaOH

4.0 g of sodium hydroxide (Ajax Finechem, Australia) was dissolved and made up the volume to 100 mL with deionized water in a volumetric flask.

9. 50% ethanol

25 mL of ethanol (Merck, Germany) was mixed with 25 mL deionized water.

10. 80% methanol

80 mL of methanol was mixed with 20 mL deionized water.

11. Stock standard (1.0 mg/mL) rutin

0.025 g of rutin (Sigma-Aldrich, Germany) was dissolved and made up the volume to 25 mL with 80% methanol in a volumetric flask.

12. 7.12 mM ABTS

0.01950 g of 2,2′-Azino-bis (3-ethylbenzothiazoline-6-sulfonic acid) diammonium salt (Amresco, USA) was dissolved and made up to 5 mL with deionized water.

13. 140 mM potassium persulfate

0.9461 g of potassium persulfate (Sigma-Aldrich, Germany) was dissolved and made up the volume to 25 mL with deionized water in a volumetric flask.

14. Working solution of ABTS ($ABTS^{\cdot+}$ cation radical)

35 µL of 7.12 mM ABTS was mixed with 1.965 mL of 140 mM potassium persulfate. This mixture was incubated in the dark at room temperature overnight.

15. Stock standard (400 ugs/mL) ascorbic acid

0.0100 g of ascorbic acid (Bio Basic, Canada) was dissolved and made up the volume to 25 mL with deionized water in a volumetric flask.

Appendix C

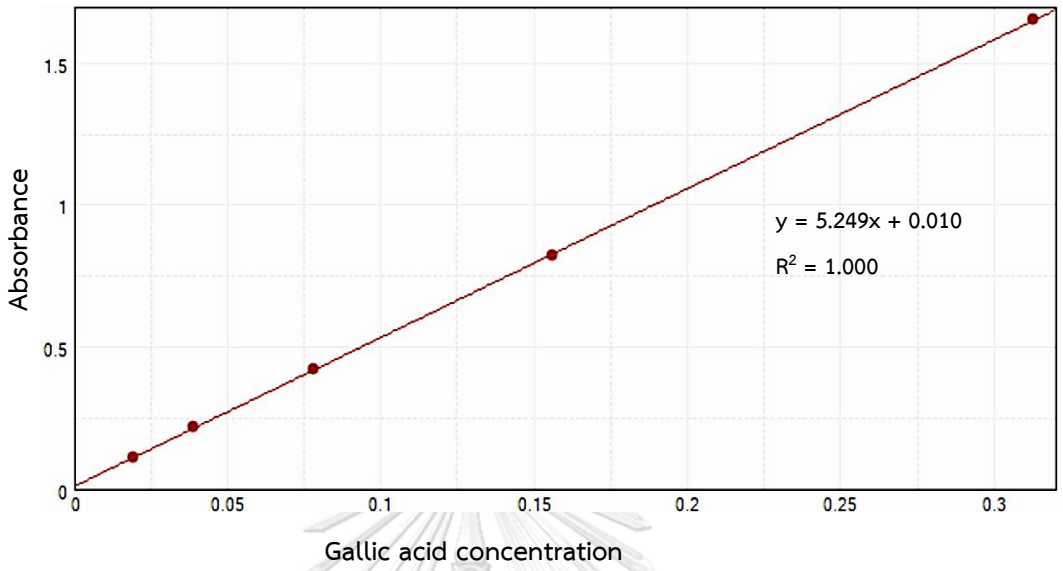


Figure 20 Gallic acid standard curve for TPC.

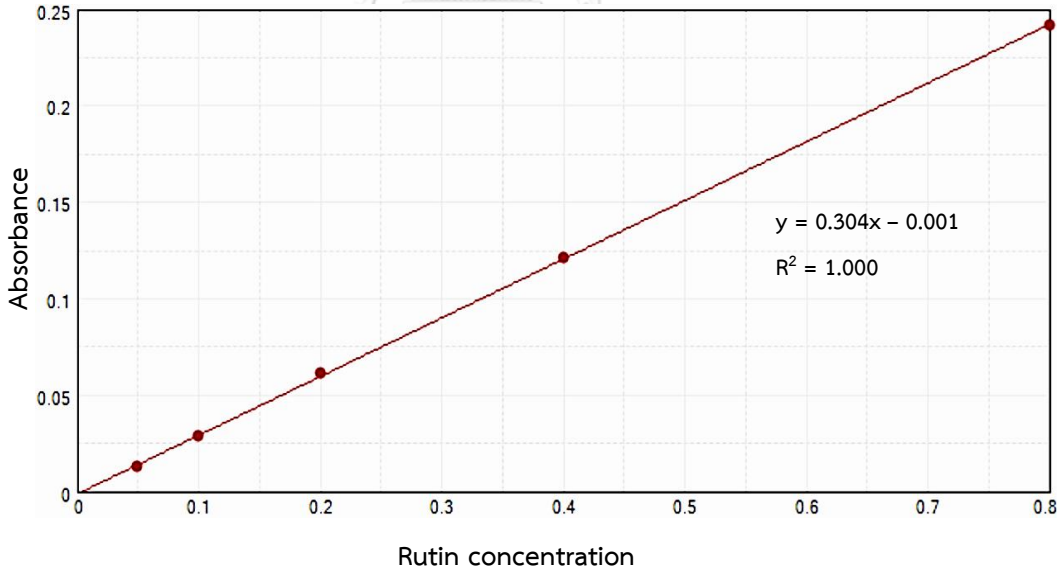


Figure 21 Rutin standard curve for TFC.

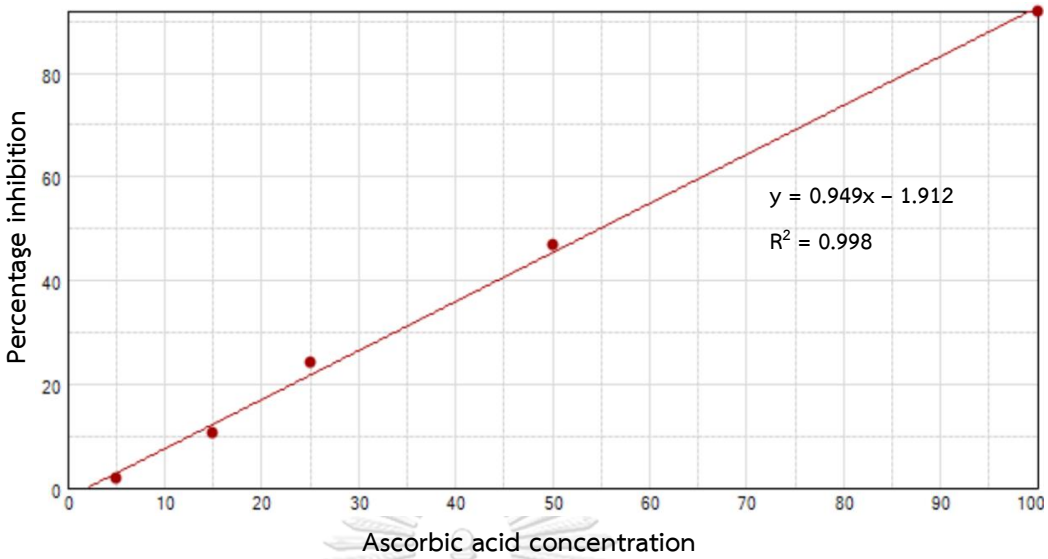


Figure 22 Ascorbic acid standard curve for AC.

Example of TPC

The linear equation of gallic acid standard curve

$$y = 5.249x + 0.010$$

Where: y = absorbance at 765 nm; x = gallic acid concentration (mg/mL); V = total volume of extract (20 mL); W = weight of rice (0.5 g); C = concentration from the calibration curve (mg/ml)

$$0.098 = 5.249x + 0.010$$

$$x = (0.098 - 0.010) \div 5.249$$

$$x = 0.0168 \text{ mg/mL}$$

$$TPC = C \times V \div W$$

$$0.67 \text{ mg GAE/g} = 0.0168 \text{ mg/mL} \times 20 \text{ mL} \div 0.5 \text{ g}$$

Appendix D

Table 8 One-way ANOVA results for TPC, TFC, and AC testing difference between groups of pericarp colour (white, red, and purple).

Traits		Sum of Squares	df	Mean Square	F	Sig.
TPC	Between Groups	188.768	2	94.384	1439.742	.000
	Within Groups	11.210	171	.066		
	Total	199.979	173			
TFC	Between Groups	614.896	2	307.448	712.874	.000
	Within Groups	73.749	171	.431		
	Total	688.645	173			
AC	Between Groups	30076764.875	2	15038382.437	1070.358	.000
	Within Groups	2402526.178	171	14049.861		
	Total	32479291.053	173			

Table 9 The Duncan's multiple range tests (DMRT) for TPC, TFC, and AC determining the significant differences between groups of pericarp colours (white, red, and purple).

Traits		Pericarp colour	N	Subset for alpha = 0.05		
				1	2	3
TPC	Duncan[a,b]	white	152	1.244		
		purple	12		4.075	
		red	10			4.703
		Sig.		1.000	1.000	1.000
TFC	Duncan[a,b]	white	152	2.982		
		red	10		8.616	
		purple	12		8.657	
		Sig.		1.000	.861	
AC	Duncan[a,b]	white	152	186.778		
		purple	12		1232.555	
		red	10			1641.904
		Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 15.797.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Appendix E

Table 10 The average of antioxidant traits (TPC, TPC, and AC) and pericarp colour in 174 rice cultivars.

G.S.No.	Cultivar	TPC (mg GAE/g)	TFC (mg RE/g)	AC (μg AEAC/g)	Pericarp colour
10049	Aew mod daeng	0.95	2.80	241.48	white
8100	Ai tai	1.93	5.07	164.51	white
23595	Beu saw mi	1.45	3.05	282.51	white
1990	Bua noi	1.20	2.89	203.95	white
3031	Bun mah	1.35	2.57	214.51	white
20712	Chai nat 1	1.22	3.13	247.19	white
2042	Chaw ma gawk	1.24	2.95	124.24	white
9742	Chaw pli khao	5.48	11.13	2043.73	red
21964	Chiang phatthalung	1.16	2.93	201.12	white
16235	Chumpae 60	1.12	2.90	220.94	white
-	CSSL 11	1.33	2.43	225.37	white
22850	Daeng nah	1.36	3.39	144.60	white
3271	Dam dahng	4.40	7.40	1520.64	purple
22817	Daw dawk mai	1.32	2.93	187.95	white
21792	Daw sahm deuan	1.18	2.73	152.57	white
12160	Dawk khah	1.38	3.44	103.45	white
15127	Dawk mai	1.52	3.41	224.48	white
22367	Di si	1.18	3.03	196.07	white
23114	E-khao yai	1.43	3.96	288.51	white
14082	E-lai	1.64	3.51	174.21	white
5639	E-mum	1.44	3.40	123.13	white
23239	E-puang	1.09	3.14	145.49	white
22653	Gaen jan	1.04	2.41	156.16	white
4490	Gam feuang	4.60	9.69	1880.63	purple
3321	Gam liaw	3.78	6.57	1242.63	purple

Table 10 (Continued)

G.S.No.	Cultivar	TPC (mg GAE/g)	TFC (mg RE/g)	AC (μg AEAC/g)	Pericarp colour
-	Gam nahng payah	4.74	10.25	1520.00	purple
-	Gam noi	4.20	9.07	1207.99	purple
23191	Gam pai	1.25	2.78	186.36	white
1193	Gow ruang 88	1.23	2.73	230.85	white
15905	Gra dahd	1.24	3.49	198.19	white
9064	Gwian hak	1.16	2.65	275.72	white
13201	Gwian hak	1.24	3.16	229.89	white
9064	Gwian hak	1.53	2.95	349.33	white
22395	Hah ruang	1.16	3.22	251.80	white
21707	Hahng mah nai	1.53	3.80	272.47	white
5583	Hahng nahk	1.23	3.52	93.05	white
1710	Hahng yi 71	1.26	3.47	204.68	white
16232	Hantra 60	0.95	2.77	274.45	white
-	Hawm daeng	4.90	9.13	1584.68	red
-	Hawm dam	3.01	7.16	732.01	purple
-	Hawm mali daeng	4.46	7.61	1502.27	red
-	Hawm nin	3.17	7.01	727.24	purple
13981	In paeng	1.55	3.57	209.60	white
12507	Jae gan	1.00	2.29	201.25	white
5211	Jampah tawng	1.49	2.77	208.00	white
3256	Jao daeng	1.56	3.40	104.89	white
3285	Jao dawk khao	1.44	2.05	209.79	white
3330	Jao khao	1.26	3.00	107.16	white
12492	Jao rahk haeng	1.42	3.32	134.32	white
2387	Jek chuey	1.04	2.63	262.18	white
1693	Jud mawn	1.04	2.84	142.95	white
23189	Kaset daw	1.26	3.59	98.64	white

Table 10 (Continued)

G.S.No.	Cultivar	TPC (mg GAE/g)	TFC (mg RE/g)	AC (μg AEAC/g)	Pericarp colour
1961	Khai mod rin	1.05	2.57	185.96	white
22358	Khao ah-gahd	1.19	2.64	228.84	white
19877	Khao bahn pod	1.30	3.39	115.77	white
-	Khao dawk mali 105	1.07	2.81	215.97	white
6152	Khao gaew	1.23	2.37	129.13	white
23113	Khao' gam	4.01	9.59	1361.27	purple
2534	Khao gaw diaw	1.15	3.21	197.55	white
3226	Khao gon jud	1.24	3.76	167.73	white
-	Khao hawm mea payah tawng dam	4.31	9.85	1069.13	purple
12503	Khao kod	1.08	2.52	120.51	white
7282	Khao luang	1.00	2.94	123.47	white
5533	Khao luang	1.12	2.03	194.48	white
6425	Khao nuan	1.41	2.25	129.57	white
3677	Khao praguad	1.17	2.45	144.04	white
9362	Khao puang	1.35	3.28	244.80	white
1697	Khao sa nguan	1.24	2.96	176.39	white
7200	Khao samer	1.14	3.31	176.31	white
3770	Khao samer	1.26	2.63	114.04	white
2296	Khao supan	1.12	2.64	113.43	white
3810	Khao taeng mo	1.24	2.88	104.72	white
12270	Khao tah cheua	1.15	2.74	141.81	white
828	Khao tah haeng 17	0.92	2.20	233.69	white
3808	Khao tah jeua	1.13	3.29	201.01	white
2063	Khiaw hahng mah	1.08	2.80	118.67	white
21708	Khitom khao	1.42	3.21	199.61	white
20795	Koo meuang	1.03	2.49	197.64	white
7025	Lai mahk	5.79	9.85	2301.92	red

Table 10 (Continued)

G.S.No.	Cultivar	TPC (mg GAE/g)	TFC (mg RE/g)	AC (μg AEAC/g)	Pericarp colour
22836	Lao taek	1.30	3.67	224.11	white
3979	Leb nok	0.90	2.15	205.33	white
14155	Leuang bai jaek	1.06	2.76	223.04	white
14154	Leuang bai lod	1.13	2.94	192.56	white
21782	Leuang chumpae	1.24	2.71	158.03	white
22046	Leuang dong	1.31	3.15	209.64	white
12498	Leuang glahng	1.33	2.52	120.67	white
7303	Leuang huan	1.17	2.37	206.95	white
21790	Leuang kamin	1.19	2.79	218.25	white
5551	Leuang kwai lah	1.07	2.27	176.74	white
7285	Leuang kwai lah	1.08	2.24	214.35	white
12616	Leuang ngahm	1.16	3.23	98.80	white
588	Leuang noi 31-1-39	1.20	3.26	148.18	white
21775	Leuang pahk chong	1.35	3.00	183.23	white
7293	Leuang plah gim	1.13	2.55	182.83	white
5681	Leuang pratew	1.35	2.79	130.03	white
7214	Leuang puang tawng	1.21	2.17	107.93	white
14685	Leuang tah yang	1.14	3.23	165.32	white
3318	Leuang tawng	1.29	2.71	224.48	white
6230	Loi hah ruang	4.04	6.49	1488.49	red
21963	Look daeng pattani	1.15	2.08	269.51	white
23303	Look daeng pattani	1.12	2.12	212.40	white
2975	Luang pratahn	1.36	3.45	135.72	white
2975	Luang pratahn	1.22	3.35	87.08	white
5650	Ma fai	1.42	3.17	177.83	white
13972	Ma gawk	1.32	3.57	133.95	white
13988	Ma yom	1.23	3.61	78.32	white

Table 10 (Continued)

G.S.No.	Cultivar	TPC (mg GAE/g)	TFC (mg RE/g)	AC (μg AEAC/g)	Pericarp colour
22802	Ma yom	1.32	3.45	275.00	white
22835	Mahk nam	1.33	2.97	215.28	white
21706	Mahk yom	1.69	3.95	276.79	white
-	Mali dam	4.31	8.93	1185.89	purple
22391	Med makham	1.20	2.69	172.48	white
5580	Muay hin	1.44	3.40	217.67	white
22379	Nah khawan	1.13	2.45	187.29	white
3638	Nahng mon	1.28	2.69	306.72	white
1204	Nahng mon s-4	1.20	2.44	206.23	white
2066	Nahng ngahm	1.15	2.61	201.12	white
3151	Nahng nuan	1.20	3.00	238.57	white
3023	Nam sagui 19	1.14	2.57	204.82	white
15833	Nam sagui 19	1.12	2.81	229.19	white
21240	Niaw dam lai	4.03	9.39	1242.07	purple
607	Niaw khao	1.42	3.17	154.96	white
22875	Niaw mali	1.38	4.04	137.24	white
23405	Niaw prae 1	1.36	3.27	220.99	white
-	Nok kum	4.48	7.88	1493.48	red
23898	Pathum thani 1	1.01	3.09	203.72	white
17770	Pathum thani 60	1.20	3.47	225.36	white
12159	Pawng aew	1.43	3.44	83.28	white
21577	Pawng aew	1.27	3.16	188.61	white
1581	Pin gaew 56	1.26	2.57	260.53	white
3241	Plah khaeng	1.37	4.12	169.19	white
22206	Plah sew	1.24	3.19	228.84	white
13967	Pra in	1.29	3.11	145.55	white
23406	Prachin buri 1	1.26	2.48	215.68	white

Table 10 (Continued)

G.S.No.	Cultivar	TPC (mg GAE/g)	TFC (mg RE/g)	AC (μg AEAC/g)	Pericarp colour
20375	Pratahn ban bung	1.06	2.37	138.21	white
2089	Puang hahng moo	1.08	2.27	269.29	white
23233	Puang hahng nahk	1.62	3.59	195.41	white
12266	Puang nahk	1.13	2.64	237.92	white
18442	Puang tawng	1.01	3.24	90.96	white
574	Puang tawng	1.28	2.25	137.36	white
2069	Puang wahn	1.11	3.39	197.19	white
12177	Rahk haeng	1.61	3.53	109.57	white
-	Rathuheti	5.29	10.13	1737.65	red
4790	RD10	1.17	2.93	235.30	white
3999	RD17	1.15	2.57	224.14	white
4000	RD19	0.86	2.43	186.80	white
4791	RD21	1.17	2.44	177.81	white
7125	RD27	1.18	2.61	155.04	white
24533	RD31	1.06	2.77	256.46	white
-	RD41	1.05	3.05	135.49	white
-	RD69	4.07	7.39	1388.07	red
3091	RD8	1.18	3.52	201.43	white
2601	RD9	1.27	2.76	247.24	white
-	Red rose rice	4.68	8.13	1515.08	red
-	Riceberry	4.34	8.96	1101.16	purple
22834	Ruang diaw	1.56	3.28	174.05	white
7288	Sahm ruang	1.27	2.38	190.91	white
22366	Sai yud	1.19	2.95	220.73	white
13987	Sam ahang	1.22	3.17	145.56	white
15101	Sang yod	3.85	8.40	1363.67	red
5699	Sao nueng	1.58	3.75	130.92	white

Table 10 (Continued)

G.S.No.	Cultivar	TPC (mg GAE/g)	TFC (mg RE/g)	AC (μg AEAC/g)	Pericarp colour
5678	Setti	1.16	2.92	206.25	white
3116	Sew mae jan	1.10	2.99	220.79	white
4001	Sew mae jan	1.27	3.23	160.63	white
22492	Soon	1.52	3.33	200.74	white
16240	Suphan buri 60	1.24	2.80	243.94	white
1578	Ta pow gaew 161	1.36	3.87	142.11	white
22357	Ta pow lom	1.22	2.45	250.08	white
21695	Tah bahn	1.44	3.88	167.96	white
5545	Tah jeua	1.04	2.51	134.40	white
10917	Tawng rahk sai	1.39	3.21	117.79	white
22784	Tom meuang luang	1.40	4.05	198.79	white
3637	Toon chalawng	1.25	2.95	148.72	white
-	UBN	1.30	3.05	239.61	white
2067	U-ta pow	1.08	2.65	207.17	White

Appendix F

1. PCV, GCV, and H^2 of TPC, TFC, and AC in 174 rice cultivars

Table 11 One-way ANOVA results for TPC, TFC, and AC testing difference between 174 rice cultivars.

Traits		Sum of Squares	df	Mean Square	F	Sig.
TPC	Between Groups	599.936	173	3.468	15.693	0.000
	Within Groups	76.899	348	0.221		
	Total	676.835	521			
TFC	Between Groups	2065.935	173	11.942	7.931	0.000
	Within Groups	523.978	348	1.506		
	Total	2589.913	521			
AC	Between Groups	97437873.156	173	563224.700	23.654	0.000
	Within Groups	8286319.451	348	23811.263		
	Total	105724192.607	521			

Example of TPC

$$\sigma_g^2 = \left(\frac{MSG - MSE}{r}\right)$$

$$\sigma_g^2 = \left(\frac{3.468 - 0.221}{3}\right)$$

$$\sigma_g^2 = \left(\frac{3.247}{3}\right)$$

$$\sigma_g^2 = 1.082$$

$$\sigma_e^2 = MSE$$

$$\sigma_e^2 = 0.221$$

$$\sigma_p^2 = \sigma_g^2 + \sigma_e^2$$

$$\sigma_p^2 = 1.082 + 0.221 = 1.303$$

$$PCV = \frac{\sqrt{\sigma_p^2}}{\bar{x}} \times 100 = \frac{\sqrt{1.303}}{1.64} \times 100 = 69.60$$

$$GCV = \frac{\sqrt{\sigma_g^2}}{\bar{x}} \times 100 = \frac{\sqrt{1.082}}{1.64} \times 100 = 63.43$$

$$H^2 = \frac{\sigma_g^2}{\sigma_p^2} \times 100 = \frac{1.082}{1.303} = 0.8304$$

2. PCV, GCV, H^2 and %GA of TPC, TFC, and AC in 208 F_2

A variance of P_1, P_2, F_1, and F_2 were calculated from $V = \frac{\Sigma(x-\bar{x})}{n}$. The results were shown in Table 12.

Table 12 The variance of P_1, P_2, F_1 and F_2 of TPC, TFC, and AC

Trait	V_{P1}	V_{P2}	V_{F1}	V_{F2}
TPC	1741.71	166.56	752.64	15091.47
TFC	19221.63	4659.75	6925.31	58591.51
AC	46.32	1.83	12.17	164.77

Example of TPC

$\sigma_p^2 = V_{F2}$

$= 1,5091.47$

$\sigma_e^2 = \frac{V_{P1}+V_{P2}+V_{F1}}{3} = \frac{1,741.71+166.56+752.64}{3}$

$= 886.97$

$\sigma_g^2 = \sigma_p^2 - \sigma_e^2$

$= 1,5091.47-886.97$

$= 1,4204.50$

$$PVC = \frac{\sqrt{\sigma_p^2}}{\bar{x}} \times 100 = \frac{\sqrt{1,5091.47}}{232.80} \times 100 = 52.77$$

$$GVC = \frac{\sqrt{\sigma_g^2}}{\bar{x}} \times 100 = \frac{\sqrt{1,4204.50}}{232.80} \times 100 = 51.20$$

$$H^2 = \frac{\sigma_g^2}{\sigma_p^2} = \frac{1,4204.50}{1,5091.47} = 0.9412$$

$$\%GA = \frac{K \times \sqrt{\sigma_p^2 \times H^2}}{\bar{x}} \times 100 = \frac{2.06 \times \sqrt{1,5091.47} \times 0.9412}{232.80}$$

$$= 102.32$$

Appendix G

Table 13 The 16 primers were used to amplify and collected genome sequencing data of RD41 (P_1), RB (P_2), and F_1 ($P_1 \times P_2$).

Locus	No.	Primer names	Primer sequences (5' ⟶ 3')	T_A^*
LOC_Os01g32890	1	F1-18063339seq	CTCGTCATTGCCATTCCACT	58
		R1-18063339seq	CAACATGCTGCTGCTCAAGA	
LOC_Os01g33090	2	F1-18207624seq	ATCAAGCACCAGAAGCACTG	57
		R1-18207624seq	TCTGCGACATGCCTCATAGA	
	3	F1-18384219seq	CACCTGTGCCTATGTGGAAC	58
		R1-18384219seq	ACCTGCTTCAAGGATGGCTA	
LOC_Os03g15860	4	F3-8753058seq	AGGCTATCTCTGTTGCTCGT	57
		R3-8753058seq	AGAAGGAGGCTTAGTTTCACCA	
LOC_Os05g50920	5	F5-29221090seq	CCTGTACTTTACCCACAGTAGAT	60
		R5-29221090seq	TCATCAAGCCAACTAACCGC	
LOC_Os05g51100	6	F5-29314992seq	GAGCACGGATTCAGATTGGC	57
		R5-29314992seq	AGAGTAAAACTGAGCAGACGG	
LOC_Os05g51119	7	F5-29326298seq	CGAGGTGGCGAATTTGAGTT	58
		R5-29326298seq	TGAAACAGCAGCACCAGAAC	
LOC_Os06g08510	8	F6-4203030seq	GATTTGGTTGGTGAGGTGGG	61
		R6-4203030seq	GCAGAGAATTTTGCTGTGACA	
LOC_Os06g08530	9	F6-4219208seq	GGGAGTATTCGTCTTCATTCTGG	61
		R6-4219208seq	CCTGGTGGGTGTATGTTGAT	
LOC_Os06g08550	10	F6-4234533seq	AGTTAGGCATTCGTTTGGTTAGT	58
		R6-4234533seq	TGGTGATGACAGGTTGACAGT	
LOC_Os07g11020	11	F7-6067391seq	CAGCACTCAACATCCAACCC	57
		R7-6067391seq	AATGCCAAGAGTGACAAGGC	
LOC_Os07g11060	12	F7-6110492seq	ACAGCTTGTGACTCCTCATCT	62
		R7-6110492seq	TCTTGTTTGGATGCGAAGGG	
LOC_Os07g11070	13	F7-6119155seq	ATGGTGTTGCTCGCTTGTTT	58
		R7-6119155seq	TGAAAGGACAGTTGCTTCCA	

Table 13 (Continued)

Locus	No.	Primer names	Primer sequences (5' ⟶ 3')	T_A*
LOC_Os07g11100	14	F7-6133379seq	TGTCATTCGCACCTACCTGT	62
		R7-6133379seq	GTGCTGAGGATGCTGATGGT	
LOC_Os07g11120	15	F7-6141196seq	GGAGAGAGAGGGGTAGGGTT	60
		R7-6141196seq	TTTGTCGGCGTCTTTCAACG	
LOC_Os07g44440	16	F7-26553573seq	GTCCCACACTCCTCTCCATC	60
		R7-26553573seq	CCATTGGTTGCCTCACTGTC	
LOC_Os08g33200		F8-20675493seq	AAACACCACCATCAAAGTGT	58
		R8-20675493seq	ACTTCCTCCGGCCTTAAGTA	

Remark: * T_A is annealing temperature.

Table 14 Three polymorphic primers were used to genotype in 208 F_2 (RD41 × RB) by HRM analysis.

Locus	Primer Name	Primer Sequence (5' ⟶ 3')	Annealing temperature
LOC_Os01g33090	HRM-F-18384219	CCGAGGTGTCCTCCCATCC	62
	HRM-R-18384219	CTGATGCTRATAGGGCTTTGGCG	
LOC_Os06g08550	HRM-F-4234533	CCTATGGCTAAAGTGTGGMTTC	58
	HRM-R-4234533	TGGTTGGCATACTCACATTACC	
LOC_Os06g08530	HRM-F-4219208	CTTCAATTATATGGATCTGGGC	52
	HRM-R-4219208	CGAGATTTAGTCAACTACTATTAGG	

Table 15 Summary of significant SNPs in promoter and exome regions were associated with two or more antioxidant traits.

Chr.	Traits	Position	SNP	*p*-value	R^2	Part	Locus	
1	TPC	18123656	A/C	2.90×10^{-7}	0.3003	promoter	LOC_Os01g33000.1	expressed protein
	TFC			8.09×10^{-8}	0.3669			
	TPC	18062779	C/T	5.25×10^{-9}	0.3121	exome	LOC_Os01g32890.1	expressed protein
	TFC			3.59×10^{-8}	0.3323			
	AC			1.02×10^{-9}	0.2540			
	TPC	18063339	A/T	1.73×10^{-9}	0.3121			
	TFC			6.50×10^{-9}	0.3323			
	AC			9.31×10^{-10}	0.2540			
	TPC	18207624	C/T	1.73×10^{-9}	0.3121	exome	LOC_Os01g33090.1	expressed protein
	TFC			6.50×10^{-9}	0.3323			
	AC			9.31×10^{-10}	0.2540			
	TPC	18384219	C/T	4.38×10^{-8}	0.1531			
	TFC			2.06×10^{-7}	0.1831			
	AC			1.49×10^{-7}	0.1191			

Table 15 (Continued)

Chr.	Traits	Position	SNP	*p*-value	R^2	Part	Locus	
2	TPC	32908792	C/T	5.83×10^{-8}	0.0953	exome	LOC_Os02g53740.2	XI-I, putative
	TFC			4.01×10^{-7}	0.0934			
	AC			1.07×10^{-8}	0.0837			
3	TPC	8753058	G/T	1.02×10^{-8}	0.2571	promoter	LOC_Os03g15860.1	mitochondrial carrier protein
	TFC			1.90×10^{-8}	0.2119			
	AC			3.19×10^{-9}	0.2855			

Table 15 (Continued)

Chr.	Traits	Position	SNP	*p*-value	R^2	Part	Locus	
4	TPC	23523608	C/T	2.72×10^{-7}	0.0730	promoter	LOC_Os04g39489.1	amino acid transporter
	AC			4.14×10^{-7}	0.0779			
	TPC	23523791	T/C	5.45×10^{-8}	0.0887			
	AC			4.73×10^{-8}	0.1046			
	TPC	23524227	C/T	2.72×10^{-7}	0.0730	promoter	LOC_Os04g39510.1	expressed protein
	AC			4.14×10^{-7}	0.0779			
	TPC	23524291	G/A	2.72×10^{-7}	0.0730			
	AC			4.14×10^{-7}	0.0779			
	TPC	23524452	C/T	2.72×10^{-7}	0.0730			
	AC			4.14×10^{-7}	0.0779			
	TPC	23524579	C/T	5.45×10^{-8}	0.0887			
	AC			4.73×10^{-8}	0.1046			
	TPC	23524837	C/T	2.72×10^{-7}	0.0730			
	AC			4.14×10^{-7}	0.0779			
	TPC	23524847	T/C	2.72×10^{-7}	0.0730			
	AC			4.14×10^{-7}	0.0779			

Table 15 (Continued)

Chr.	Traits	Position	SNP	*p*-value	R^2	Part	Locus	
	TPC	23524996	A/C	5.45×10^{-8}	0.0887			
	AC			4.73×10^{-8}	0.1046			
	TPC	23525054	G/T	5.45×10^{-8}	0.0887			
	AC			4.73×10^{-8}	0.1046			
	TPC	23525169	G/A	2.72×10^{-7}	0.0730			
	AC			4.14×10^{-7}	0.0779			
	TPC	23525262	C/T	2.72×10^{-7}	0.0730			
	AC			4.14×10^{-7}	0.0779			
	TPC	23525278	G/A	2.67×10^{-7}	0.1002			
	AC			4.18×10^{-7}	0.1113			
	TPC	23600805	C/A	5.45×10^{-8}	0.1141	promoter	LOC_Os04g39600.1	fasciclin domain containing protein
	AC			4.73×10^{-8}	0.1370			
	TPC	23604978	G/A	5.45×10^{-8}	0.0887	promoter	LOC_Os04g39610.1	glycerophosphoryl diester
	AC			4.73×10^{-8}	0.1046			phosphodiesterase family protein
	TPC	23606423	T/C	2.72×10^{-7}	0.0730			
	AC			4.14×10^{-7}	0.0779			

Table 15 (Continued)

Chr.	Trait	Position	SNP	*p*-value	R^2	Part	Locus	
	TPC	33530279	C/T	7.04×10^{-8}	0.3027	promoter	LOC_Os04g56240.1	lipase
	TFC		C/T	8.96×10^{-8}	0.2915			
	TPC	23374014	C/T	1.96×10^{-8}	0.0396	exome	LOC_Os04g39280.1	ATPOT1
	AC			3.19×10^{-10}	0.0580			
	TPC	23508984	C/T	9.77×10^{-11}	0.0561	exome	LOC_Os04g39460.1	NBS-LRR type disease resistance protein
	TFC			3.88×10^{-7}	0.0461			
	AC			1.32×10^{-12}	0.0575			
	TPC	23509020	A/T	8.56×10^{-12}	0.0725			
	TFC			1.61×10^{-7}	0.0452			
	AC			4.87×10^{-14}	0.0982			
	TPC	23509168	G/A	9.27×10^{-12}	0.0503			
	TFC			1.43×10^{-7}	0.0304			
	AC			4.87×10^{-14}	0.0651			
	TPC	23509181	C/G	9.27×10^{-12}	0.0725			
	TFC			1.43×10^{-7}	0.0452			
	AC			4.87×10^{-14}	0.0982			

Table 15 (Continued)

Chr.	Traits	Position	SNP	*p*-value	R^2	Part	Locus	
	TPC	23509581	G/A	5.60×10^{-11}	0.0585			
	TFC			2.24×10^{-7}	0.0500			
	AC			1.80×10^{-12}	0.0599			
	TPC	23524227	C/T	5.54×10^{-11}	0.0439	exome	LOC_Os04g39510.1	expressed protein
	TFC			1.88×10^{-7}	0.0371			
	AC			1.66×10^{-12}	0.0423			
	TPC	23524291	G/A	5.54×10^{-11}	0.0439			
	TFC			1.88×10^{-7}	0.0371			
	AC			1.66×10^{-12}	0.0423			
	TPC	23524452	C/T	5.54×10^{-11}	0.0439			
	TFC			1.88×10^{-7}	0.0371			
	AC			1.66×10^{-12}	0.0423			
	TPC	23524579	C/T	1.13×10^{-10}	0.0646			
	AC			1.02×10^{-12}	0.0859			
	TPC	23524581	A/G	1.13×10^{-10}	0.0646			
	AC			1.02×10^{-12}	0.0859			

Table 15 (Continued)

Chr.	Traits	Position	SNP	*p*-value	R^2	Part	Locus
	TPC	23524837	C/T	5.54×10^{-11}	0.0439		
	TFC			1.88×10^{-7}	0.0371		
	AC			1.66×10^{-12}	0.0423		
	TPC	23524847	T/C	5.54×10^{-11}	0.0439		
	TFC			1.88×10^{-7}	0.0371		
	AC			1.66×10^{-12}	0.0423		
	TPC	23524996	A/C	9.27×10^{-12}	0.0503		
	TFC			1.43×10^{-7}	0.0304		
	AC			4.87×10^{-14}	0.0651		
	TPC	23525054	G/T	9.27×10^{-12}	0.0503		
	TFC			1.43×10^{-7}	0.0304		
	AC			4.87×10^{-14}	0.0651		
	TPC	23525169	G/A	5.54×10^{-11}	0.0439		
	TFC			1.88×10^{-7}	0.0371		
	AC			1.66×10^{-12}	0.0423		

Table 15 (Continued)

Chr.	Traits	Position	SNP	*p*-value	R^2	Part	Locus	
	TPC	23525262	C/T	5.54×10^{-11}	0.0439			
	TFC			1.88×10^{-7}	0.0371			
	AC			1.66×10^{-12}	0.0423			
	TPC	23544685	C/T	4.45×10^{-10}	0.0535	exome	LOC_Os04g39540.1	expressed protein
	AC			1.83×10^{-11}	0.0532			
	TPC	23555753	G/A	5.60×10^{-11}	0.0439	exome	LOC_Os04g39560.1	expressed protein
	TFC			2.24×10^{-7}	0.0371			
	AC			1.80×10^{-12}	0.0423			
	TPC	23556184	T/A	5.54×10^{-11}	0.0439			
	TFC			1.88×10^{-7}	0.0371			
	AC			1.66×10^{-12}	0.0423			
	TPC	23558044	C/T	5.54×10^{-11}	0.0439			
	TFC			1.88×10^{-7}	0.0371			
	AC			1.66×10^{-12}	0.0423			

Table 15 (Continued)

Chr.	Traits	Position	SNP	*p*-value	R^2	Part	Locus	
	TPC	23582229	T/C	2.40×10^{-9}	0.0452	exome	LOC_Os04g39570.1	WRKY35
	AC			3.05×10^{-10}	0.0413			
	TPC	23617785	A/G	3.38×10^{-9}	0.1112	exome	LOC_Os04g39629.1	RNA recognition motif containing protein
	AC			1.61×10^{-9}	0.1072			
	TPC	23617878	C/T	3.38×10^{-9}	0.1112			
	AC			1.61×10^{-9}	0.1072			
	TPC	23628731	C/G	7.44×10^{-9}	0.1062	exome	LOC_Os04g39650.1	receptor protein kinase CLAVATA1 precursor
	AC			1.40×10^{-9}	0.1026			
	TPC	23727012	C/T	3.38×10^{-9}	0.1112	exome	LOC_Os04g39830.1	expressed protein
	AC			1.61×10^{-9}	0.1072			
	TPC	23730975	C/T	3.38×10^{-9}	0.1261	exome	LOC_Os04g39840.1	Os4bglu10 - beta-glucosidase homologue,
	AC			1.61×10^{-9}	0.1278			similar to Os4Bglu12 exoglucanase/
	TPC	23749030	C/T	3.38×10^{-9}	0.1112			beta-glucosidase
	AC			1.61×10^{-9}	0.1072			

Table 15 (Continued)

Chr.	Traits	Position	SNP	*p*-value	R^2	Part	Locus	
5	TPC	25670289	A/C	1.92×10^{-7}	0.1225	promoter	LOC_Os05g44190.1	expressed protein
	TFC			3.72×10^{-7}	0.1303			
	TPC	25675758	C/T	2.94×10^{-7}	0.0924			
	TFC			6.31×10^{-8}	0.0926			
	TPC	25675805	A/C	2.94×10^{-7}	0.1225			
	TFC			6.31×10^{-8}	0.1303			
	TPC	25675824	T/C	2.94×10^{-7}	0.0924			
	TFC			6.31×10^{-8}	0.0926			
	TPC	25676784	G/A	1.22×10^{-7}	0.0691			
	AC			3.77×10^{-7}	0.0738			
	TPC	25676839	A/C	1.22×10^{-7}	0.0843			
	AC			3.77×10^{-7}	0.0884			
	TPC	25676856	G/A	1.22×10^{-7}	0.0856			
	AC			3.77×10^{-7}	0.0911			
	TPC	29221090	C/A	1.11×10^{-9}	0.2887	promoter	LOC_Os05g50920.1	transmembrane amino acid transporter
	AC			7.17×10^{-9}	0.2980			protein

Table 15 (Continued)

Chr.	Traits	Position	SNP	*p*-value	R^2	Part	Locus	
	TPC	29314992	C/T	5.43×10^{-10}	0.1975	promoter	LOC_Os05g51100.1	OsFBX179 - F-box domain containing
	AC			4.74×10^{-11}	0.2509			protein
	TPC	29315254	T/C	5.43×10^{-10}	0.1975			
	AC		T/C	4.74×10^{-11}	0.2509			
	TPC	29315387	A/G	5.43×10^{-10}	0.1975			
	AC			4.74×10^{-11}	0.2509			
	TPC	29315515	C/G	4.03×10^{-9}	0.1410			
	AC			8.48×10^{-10}	0.1638			
	TPC	29315568	C/T	4.03×10^{-9}	0.1410			
	AC			8.48×10^{-10}	0.1638			
	TPC	29315698	G/A	4.03×10^{-9}	0.1410			
	AC			8.48×10^{-10}	0.1638			
	TPC	29315815	C/T	5.43×10^{-10}	0.1975			
	AC			4.74×10^{-11}	0.2509			
	TPC	29315858	T/C	4.03×10^{-9}	0.1410			
	AC			8.48×10^{-10}	0.1638			

Table 15 (Continued)

Chr.	Traits	Position	SNP	p-value	R^2	Part	Locus
	TPC	29316010	T/C	4.03×10^{-9}	0.1410		
	AC			8.48×10^{-10}	0.1638		
	TPC	29316074	C/G	2.08×10^{-8}	0.1034		
	AC			2.06×10^{-9}	0.1246		
	TPC	29316096	G/T	5.43×10^{-10}	0.1975		
	AC			4.74×10^{-11}	0.2509		
	TPC	29316097	A/T	5.43×10^{-10}	0.1975		
	AC			4.74×10^{-11}	0.2509		
	TPC	29316314	A/T	4.03×10^{-9}	0.1410		
	AC			8.48×10^{-10}	0.1638		
	TPC	29316350	C/T	4.03×10^{-9}	0.1410		
	AC			8.48×10^{-10}	0.1638		
	TPC	29316791	C/T	5.43×10^{-10}	0.1975		
	AC			4.74×10^{-11}	0.2509		
	TPC	29316931	T/G	5.43×10^{-10}	0.1975		
	AC			4.74×10^{-11}	0.2509		

Table 15 (Continued)

Chr.	Traits	Position	SNP	p-value	R^2	Part	Locus	
	TPC	29315724	A/C	4.03×10^{-9}	0.1410			
	AC				0.1638			
	TPC	29315726	A/C	4.03×10^{-9}	0.1410			
	AC				0.1638			
	TPC	29315727	A/C	4.03×10^{-9}	0.1410			
	AC				0.1638			
	TPC	29316060	A/C	4.03×10^{-9}	0.1410			
	AC				0.1638			
	TPC	29326298	A/G	7.49×10^{-12}	0.1975	promoter	LOC_Os05g51119.1	expressed protein
	TFC			7.47×10^{-8}	0.1409			
	AC			1.53×10^{-12}	0.2509			

Table 15 (Continued)

Chr.	Traits	Position	SNP	p-value	R^2	Part	Locus	
6	TPC	4203030	A/C	2.17×10^{-8}	0.1193	promoter	LOC_Os06g08510.1	expressed protein
	TFC			6.62×10^{-9}	0.1379			
	AC			4.48×10^{-7}	0.1006			
	TPC	4219208	G/T	2.17×10^{-8}	0.1193	promoter	LOC_Os06g08530.1	ubiquitin carboxyl-terminal hydrolase domain
	TFC			6.62×10^{-9}	0.1379			containing protein
	AC			4.48×10^{-7}	0.1006			
	TPC	4220525	C/A	2.17×10^{-8}	0.1193			
	TFC			6.62×10^{-9}	0.1379			
	AC			4.48×10^{-7}	0.1006			
	TPC	4233259	A/C	2.17×10^{-8}	0.1193	promoter	LOC_Os06g08550.1	BTBN14 - Bric-a-BrAC, Tramtrack, Broad
	TFC			6.62×10^{-9}	0.1379			Complex BTB domain with non-phototropic
	AC			4.48×10^{-7}	0.1006			hypocotyl 3 NPH3 domain
	TPC	4233682	A/C	2.17×10^{-8}	0.1193			
	TFC			6.62×10^{-9}	0.1379			
	AC			4.48×10^{-7}	0.1006			

Table 15 (Continued)

Chr.	Traits	Position	SNP	p-value	R^2	Part	Locus
	TPC	4234321	G/A	2.17×10^{-8}	0.1193		
	TFC			6.62×10^{-9}	0.1379		
	AC			4.48×10^{-7}	0.1006		
	TPC	4234533	G/A	2.17×10^{-8}	0.1193		
	TFC			6.62×10^{-9}	0.1379		
	AC			4.48×10^{-7}	0.1006		

Table 15 (Continued)

Chr.	Traits	Position	SNP	p-value	R^2	Part	Locus	
7	TPC	6132796	G/T	1.87×10^{-11}	0.2334	promoter	LOC_Os07g11100.1	hypothetical protein
	TFC			2.45×10^{-11}	0.2219			
	AC			1.88×10^{-10}	0.2461			
	TPC	6133379	A/G	1.87×10^{-11}	0.2334			
	TFC			2.45×10^{-11}	0.2219			
	AC			1.88×10^{-10}	0.2461			
	TPC	6141196	T/C	1.87×10^{-11}	0.2334	promoter	LOC_Os07g11120.1	hydrolase, NUDIX family, domain
	TFC			2.45×10^{-11}	0.2219			containing protein
	AC			1.88×10^{-10}	0.2461			
	TPC	6141248	G/A	1.87×10^{-11}	0.2334			
	TFC			2.45×10^{-11}	0.2219			
	AC			1.88×10^{-10}	0.2461			
	TPC	6143116	A/G	1.87×10^{-11}	0.2334			
	TFC			2.45×10^{-11}	0.2219			
	AC			1.88×10^{-10}	0.2461			

Table 15 (Continued)

Chr.	Traits	Position	SNP	p-value	R^2	Part	Locus	
	TPC	26553409	T/G	5.63×10^{-8}	0.2086	promoter	LOC_Os07g44440.1	peroxiredoxin
	AC			7.24×10^{-8}	0.2348			
	TPC	26553573	T/C	3.72×10^{-8}	0.1307			
	AC			1.82×10^{-8}	0.1654			
	TPC	6067391	G/A	2.20×10^{-13}	0.1979	exome	LOC_Os07g11020.1	rc - bHLH transcription factor regulating
	TFC			5.73×10^{-12}	0.1644			proanthocyanidin production in seeds
	AC			4.16×10^{-13}	0.1988			
	TPC	6110492	T/C	2.20×10^{-13}	0.1979	exome	LOC_Os07g11060.1	expressed protein
	TFC			5.73×10^{-12}	0.1644			
	AC			4.16×10^{-13}	0.1988			
	TPC	6118390	C/G	2.20×10^{-13}	0.1979	exome	LOC_Os07g11070.1	DUF630/DUF632 domains containing
	TFC			5.73×10^{-12}	0.1644			protein
	AC			4.16×10^{-13}	0.1988			
	TPC	6119155	C/T	2.20×10^{-13}	0.1979			
	TFC			5.73×10^{-12}	0.1644			
	AC			4.16×10^{-13}	0.1988			

Table 15 (Continued)

Chr.	Traits	Position	SNP	p-value	R^2	Part	Locus	
	TPC	6132796	G/T	2.20×10^{-13}	0.1979	exome	LOC_Os07g11100.1	hypothetical protein
	TFC			5.73×10^{-12}	0.1644			
	AC			4.16×10^{-13}	0.1988			

Table 15 (Continued)

Chr.	Traits	Position	SNP	p-value	R^2	Part	Locus	
8	TPC	20675493	A/C	4.23×10^{-7}	0.3447	promoter	LOC_Os08g33200.1	1-phosphatidylinositol-4-phosphate 5-
	TFC			2.37×10^{-7}	0.3581			kinase/ zinc ion binding protein
	AC			4.07×10^{-7}	0.3217			
	TPC	20776463	A/G	7.92×10^{-7}	0.2230	exome	LOC_Os08g33320.1	expressed protein
	TFC			2.95×10^{-7}	0.2621			
	AC			3.24×10^{-7}	0.1749			
	TPC	20777221	C/T	7.92×10^{-7}	0.2230			
	TFC			2.95×10^{-7}	0.2621			
	AC			3.24×10^{-7}	0.1749			
	TPC	20800570	C/T	7.92×10^{-7}	0.2230	exome	LOC_Os08g33340.1	RAD23 DNA repair protein
	TFC			2.95×10^{-7}	0.2621			
	AC			3.24×10^{-7}	0.1749			
	TPC	20812864	G/A	7.92×10^{-7}	0.2230	exome	LOC_Os08g33370.2	14-3-3 protein
	TFC			2.95×10^{-7}	0.2621			
	AC			3.24×10^{-7}	0.1749			

Table 15 (Continued)

Chr.	Traits	Position	SNP	p-value	R^2	Part	Locus	
	TPC	20823925	C/G	7.92×10^{-7}	0.2230	exome	LOC_Os08g33390.1	Regulator of chromosome condensation
	TFC			2.95×10^{-7}	0.2621			domain containing protein
	AC			3.24×10^{-7}	0.1749			
	TPC	20855542	A/G	7.92×10^{-7}	0.2230	exome	LOC_Os08g33430.1	formin
	TFC			2.95×10^{-7}	0.2621			
	AC			3.24×10^{-7}	0.1749			
	TPC	20856464	C/T	7.92×10^{-7}	0.2230			
	TFC			2.95×10^{-7}	0.2621			
	AC			3.24×10^{-7}	0.1749			

VITA

NAME	Miss Saranyu Thaworn
DATE OF BIRTH	July 14, 1983
PLACE OF BIRTH	Srisaket
INSTITUTIONS ATTENDED	Bachelor of Scienec (Biotecnology), Silapakorn University, Thailand. Master of Science (Genetics), Chulalongkorn University, Thailand. (GENETIC ANALYSIS OF AGRONOMIC CHARACTERS RELATED TO YIELD IN VEGETABLE SOYBEAN Glycine max (L.) Merrill)
HOME ADDRESS	9/171, Bang Pakok, Rat Burana, Bangkok 10140
PUBLICATION	1. Thaworn, S., Chadchawan, S., Paliyavuth, C. and Kasettranan, W. (2021). Clustering of white, red and purple rice cultivars according to their total phenolic content, total flavonoid content and antioxidant capacity in their grains. Agriculture and Natural Resources. 55: 89-97. doi:10.34044/j.anres.2021.55.1.12 2. ศรัณยู ถาวร ชนิตา ปาลิยะวุฒิ และ วราลักษณ์ เกษตรานันท์ (2021). การถ่ายทอดลักษณะต้านออกซิเดชันของข้าวกล้องในประชากรข้าวรุ่น F2 จากคู่ผสม “RD41” × “Riceberry”.แก่นเกษตร. (accepted)