Full Length Article

Genetic Characteristics of Amylose Content in Medium-Grain Rice (*Oryza sativa*)

Bui Phuoc Tam[1,4][†], Pham Thi Be Tu[2][*] and Nguyen Thi Pha[3][†]
[1]*PhD student, Can Tho University, Can Tho city 94000, Vietnam*
[2]*College of Agriculture, Can Tho University, Can Tho city 94000, Vietnam*
[3]*Institute of Food and Biotechnology, Can Tho University, Can Tho city 94000, Vietnam*
[4]*Loc Troi Agricultural Research Institute, An Giang Province 90000, Vietnam*
[*]For correspondence: ptbtu@ctu.edu.vn; phuoctam1987@gmail.com
[†]Contributed equally to this work and are co-first authors

Abstract

Amylose content (AC) is one of the key traits related to rice grain quality. It decided the hard or soft grain rice. Consumers generally prefer cooked rice with a soft and sticky texture. AC is a single-gene trait and is controlled by the *Waxy* gene located on chromosome 6. Genetic studies related to AC in medium-grain rice varieties underpin the identification of genes and linked molecular markers that few previous studies mentioned. The study's purpose is to investigate the genetic characteristics of AC in medium-grain rice on 12 chromosomes. A rice diversity panel of 114 genotypes was applied to evaluate AC and its correlation with the agronomic and quality characteristics. Genome-Wide Association Study (GWAS) was utilized to identify significant SNPs. The results showed that the AC of varieties had a large range from 10.83 to 30.12%, in which, the group of varieties with low AC (less than 20%) accounted for 21.9%. AC was strongly correlated with gel consistency (GC) but not with gelatinization temperature (GT) and agronomic traits. The GWAS analysis of AC identified 9 significant SNPs on chromosomes 1, 2, 4, 8, 9, 11 and 12. The frequency of alleles widely ranged from 0.94 to 56.60%. On chromosome 6, 74 significant SNPs had allele frequencies ranging from 3.8 to 67.0%. In which, two SNPs fd7 (1764762 bp) and fd8 (176707 bp) are located in the *Waxy* gene region. These findings provided the basic genetic foundation for high-quality medium-grain rice breeding programs. © 2024 Friends Science Publishers

Introduction

Rice (*Oryza sativa* L.) grain consists of approximately 90% starch and AC is widely recognized as the most important factor affecting eating and cooking quality (Tian *et al.* 2009; Cai *et al.* 2015; Zhou *et al.* 2018). AC in rice is classified into high (>25%), intermediate (20–25%), low (12–20%), very low (5–12%), or waxy (0–2%) (Juliano *et al.* 1981; Yu *et al.* 2012). Rice with a high AC tends to cook hard and dry, whereas rice with a intermediate AC tends to be softer and stickier, and rice with a low AC is generally quite soft and sticky. The major gene for amylose synthesis in rice is the *Waxy* (*Wx*) gene located on chromosome 6, which encodes the granule bound starch synthase (Septiningsih *et al.* 2003; Fan *et al.* 2005; Tian *et al.* 2009). Other minor QTLs were also detected on the 12 chromosomes of rice in previous studies.

AC has various tend in different rice species. According to Feng *et al.* (2017), evaluating 635 rice varieties (including pure rice and hybrid rice), the AC of japonica rice (short grain) was significantly lower (P < 0.05) than indica rice (long grain). This result was also recorded by IRRI (2013). Sano (1984) showed that the *Wx* locus has three different alleles, *Wx^a*, *Wx^b* and *wx*, in the rice species of indica (or long grain), japonica (or short grain), and sticky rice, respectively.

GWAS mapping makes it possible to simultaneously screen a very large number of accessions for genetic variation underlying diverse complex traits. One great advantage of the GWAS design for rice is the homozygous nature of most rice varieties, which makes it possible to employ a "genotype or sequence once and phenotype many times over" strategy, whereby once the lines are genomically characterized, the genetic data can be reused many times across different phenotypes and environments (Zhao *et al.* 2011). In rice, GWAS is gaining widespread use and there are reports (Huang *et al.* 2010; Zhao *et al.* 2011; Vanniarajan *et al.* 2012; Courtois *et al.* 2013; Ueda *et al.* 2014; Kumar *et al.* 2015).

The objectives of this study were to: (1) evaluate the

agronomic traits and analyze the amylose content, gelatinization temperature and gel consistency, (2) identify the correlation of amylose content with agronomic traits and other quality indexes, and (3) investigate single nucleotide polymorphisms (SNPs) in the whole of rice chromosomes and the target chromosome.

Materials and Methods

Place and time

The experiments were conducted at the experimental station of CanTho University, CanTho city, Vietnam in the Summer-Fall season of 2020.

Materials

The group of medium-grain rice varieties included 114 varieties selected from the 44k-SNP rice diversity panel of the International Rice Research Institute (IRRI). The data set of 44,100 SNP were referenced from Zhao *et al.* (2011).

Evaluation of the agronomic characteristics

The rice plants were grown on the alluvial soil in the 2020 Summer-Fall season. The experiment was performed as a randomized complete block design with three replications, each replication was 3 m^2. Each rice plant individual was separately transplanted in a 20 × 20 cm distance. Rice farming was applied the alternate wetting and drying method with the fertizier formula of 80 N: 60 P$_2$O$_5$: 60 K$_2$O (kg/ha). Agronomic characteristics were evaluated including duration (days), plant height (cm), panicle length (cm), number of panicles per square meter (panicles), number of filled grain per panicle (grains), percentage of unfilled grains, 1000 grain weight (gram), and yield (tons/ha).

Analysis of amylose content (AC)

The AC of the milled rice samples was determined according to the method of Juliano (1971); Graham (2002). Milled rice flour (100 mg) was soaked in 1 mL of 95% ethanol and 9.0 mL of 1 N NaOH in a 50-mL glass test tube and allowed to stand undisturbed overnight for 16 h. Afterwards, distilled water (90 mL) was added to bring the solution up to 100 mL, and 0.5 mL aliquot was transferred into a 20-mL test tube containing 5 mL distilled water. Then, 0.1 mL of 1 M CH$_3$COOH was added and the solution was mixed thoroughly using a vortex mixer followed by addition of 0.2 mL of iodine solution (0.15% I$_2$ in 1.5% KI). The solution was then diluted to 10 mL using 4.2 mL distilled water. To develop the calibration curves for the determination of amylose content in a rice sample, 40 mg Avebe potato amylose (standard amylose) put in a 50 mL test tube and proceeded as described above. Then, 0.1, 0.2, 0.3, 0.4, 0.5 mL of the standard amylose sample

solution were transferred into 20 mL test tubes and proceeded in the same way as for the test sample. Construction of the calibration curve was carried by converting from the spectral reading to the percentage of amylose content according to the followed formula (y = ax + b, where y is the absorbance OD, and x is the amount of amylose in the measured sample (mg/L) (Graham 2002).

Analysis of gelatinization temperature (GT)

The GT was determined using the alkali digestion test (IRRI 2014). A duplicate set of six whole-milled kernels without cracks was selected and placed in a petri dish (8.0 cm in diameter). Ten mL of 1.7% KOH solution was added. The samples were arranged to provide enough space between kernels to allow for spreading. The dishes were covered and incubated for 23 h at 30°C. The starchy endosperm was rated visually based on a seven-point numerical spreading scale as a standard evaluation system for rice: high (1–2), high or intermediate (3), intermediate (4–5), and low (6–7) (IRRI 2014).

Analysis of gel consistency (GC)

Analysis of gel consistency according to the method of Tang *et al.* (1991). Milled rice flour (100 mg) was put into a glass test tube (13 x 100 mm). Then, 0.2 mL of 95% ethanol containing 0.03% green thymol was placed into the test tube. 2 mL of 0.2 N KOH was further added to the test tube and shaken well on a vortex machine. The test tube was covered and placed in a pot of boiling water (100°C) for 8 min. Test tubes were cooled to room temperature for 5 min and placed in an ice bath for 20 min. Test tubes were removed and placed horizontally for 1 h. The gel consistency is the length of the moved gel measured from the bottom of the test tube to the end of the gel. The classification of gel consistency is applied according to the standard evaluation system for rice of IRRI (2014): soft (61–80 mm), medium (41–60 mm), and hard (< 40 mm).

Investigation of single nucleotide polymorphisms (SNPs)

A 44,100 SNP genotyping data (Zhao *et al.* 2011) were used for association of SNP variants with the phenotypes. GWAS by GAPIT (Genomic Association and Prediction Integrated Tool) was done using R-studio 3.2.4 software (Zhiwu Zhang Laboratory 2020). The GAPIT tool optimized enriched compressed mixed linear model approach (eCMLM). To combine association results, the significance threshold was set to P < 0.0001. Information on SNPs was collected using the Rice SNP-Seek database (https://snp-seek.irri.org/). The QTL mapping of related SNPs was made using the Ritchie Lab Visualization database (http://visualization.ritchielab.org/phenograms/plot). The steps of GWAS analysis were based on the GAPIT tool, which is depicted in Fig. 1.

Statistical methods

Data were collected and stored, and graphs were designed using the Microsoft Office Excel 2013 program. Descriptive statistics of traits were carried using STAR 2.0.1 software (IRRI) (http://bbi.irri.org/products). GWAS by GAPIT was analyzed by R-studio 3.4.1 software (https://rstudio.com/products/rstudio/download). A chart of a correlation matrix was displayed by R-studio 3.4.1 software with the 'PerformanceAnalytics' package.

Results

Evaluation of the agronomic characteristics

The agronomic characteristics of the medium-grain rice varieties are showed in Table 1. Rice varieties had short to average growth duration and were not affected by photoperiod. The agronomic characteristics had a wide variance and a small standard error.

Determination of AC, GT and GC

Amylose content has a large difference among rice varieties, ranging from 10.83 to 30.12% (Table 2). Cooked rice kernels with high AC (>25%) occupying 36.0% of the total of rice varieties were dry, separate, less tender, and become hard upon cooling, whereas those with low (≤ 20%) with 21.9% were glossy, soft, and sticky. Intermediate AC (20–25%) rice made up 42.1% was soft and flaky (Fig. 2). The experiment recorded that GT in the medium-grain rice varied from score 3 to score 7 (Table 2). GT in the high group (≤ score 3) accounted for about 4.4%, in the intermediate group (score 4–5) accounted for 23.7% and the rest was the low group (score 6–7) made up the majority, about 71.9%. For GC, most varieties were in the medium classification (41–60 mm), accounting for 51.8%. The GC in the soft group was 34.2% while this in the hard group was about 14.0%.

Correlation between AC and agronomic traits and quality characteristics in medium grain rice

The correlation between agronomic and quality characteristics was showed at Fig. 3. For agronomic characteristics, AC was positively correlated on average (+0.44) with number of panicles/m^2, and negatively correlated with plant height, panicle length and yield. However, AC was not significantly or not correlated with duration, number of filled grains per panicle, percentage of unfilled grain, and 1000 grain weight. AC was not strongly correlated with the degree of GT and this correlation tended to be negative, the correlation coefficient was (-0.19). AC was negatively associated with GC. The correlation coefficient recorded between these two indicators was (-0.85).
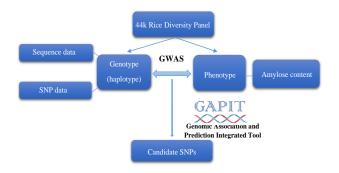

Fig. 1: Investigation of single nucleotide polymorphism (SNP) for rice genotypes with different amylose content on chromosomes of rice
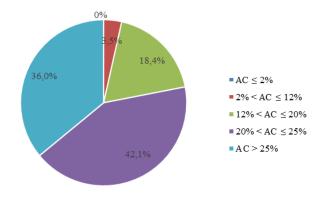

Fig. 2: Groups of AC in medium grain rice varieties
Notes: AC: Amylose content

Fig. 3: Correlation between AC and agronomic and quality characteristics of medium-grain rice varieties
Notes: Statistics at 95% significance level; TGST: Duration; Caocay: Plant height; Daibong: Panicle length; Sobong: Number of panicles per m^2; Hatchac: Number of filled grains per panicle; Tllep: Percentage of unfilled grain; W1000: 1000 grain weight; AC: Amylose content; GT: Gelatinization temperature; GC: Gel consistency

Investigation of single nucleotide polymorphism (SNP) for varieties with different AC on chromosomes of rice

Through analysis of AC in the medium-grain rice accessions by GWAS, nine significant SNPs (LOD score ~ 3) were identified on chromosomes 1, 2, 4, 8, 9, 11 and 12 (Fig. 4, 5). These SNPs significantly linked the target gene

Table 1: Descriptive statistics of agronomic characteristics in the medium-grain rice

Characteristics	Min	Max	Mean	Range	Variance	SD	SE
Duration (*days*)	73.00	132.00	98.40	59.00	104.36	10.22	0.96
Plant height (*cm*)	75.10	194.30	121.00	119.20	425.65	20.63	1.93
Panicle length (*cm*)	16.30	32.50	24.30	16.20	11.02	3.32	0.31
No. of panicle/m^2 (*panicles*)	165.00	544.50	240.40	379.50	4316.56	65.70	6.15
No. of filled grain/panicle (*grains*)	42.00	182.90	94.90	140.90	506.43	22.50	2.11
Percentage of uhfilled grain (*%*)	4.80	59.40	30.60	54.60	172.59	13.14	1.23
1000 grain weight (*gr*)	15.80	30.90	21.90	15.10	8.30	2.88	0.27
Yield (*tons/ha*)	1.15	6.05	3.04	4.90	0.95	0.97	0.09

Notes: MIN: minimum; MAX: maximum; MEAN: average; SD: standard deviation; SE: standard error

Table 2: Descriptive statistics of AC, GT, and GC in the medium-grain rice

Characteristics	Min	Max	Mean	Range	Variance	SD	SE
AC (%)	10.83	30.12	22.80	19.29	4.56	20.83	0.43
GT (score)	3.00	7.00	5.86	4.00	0.95	0.91	0.09
GC (mm)	31.00	96.00	57.26	65.00	15.33	235.13	1.44

Notes: AC: Amylose Content; GT: Gelatinization Temperature; GC: Gel Consistency; MIN: minimum; MAX: maximum; MEAN: average; SD: standard deviation; SE: standard error

Table 3: Significant SNPs related to AC located on different chromosomes in medium-grain rice varieties

SNP ID	Chr.	Position (bp)	P-value	LOD-value	O-allele	R-allele	F (%)	MAF	R^2
id2000259	2	349130	2.66×10^{-4}	3.6	T*	A*	10.13	0.038	15.2
id2000249	2	346742	5.02×10^{-4}	3.3	T*	C*	11.19	0.057	13.9
id9001853	9	6668351	5.63×10^{-4}	3.2	A	T	18.87	0.252	13.7
id12001229	12	2966521	6.48×10^{-4}	3.2	A*	C*	34.00	0.490	13.4
id11005587	11	16106487	9.38×10^{-4}	3.0	A	G	56.60	0.376	12.6
id8004295	8	16118582	11.24×10^{-4}	2.9	T	G	0.94	0.024	12.2
id4001458	4	3525063	12.04×10^{-4}	2.9	G	A	24.53	0.457	12.1
id2015370	2	34318478	13.08×10^{-4}	2.9	G	A	29.25	0.386	11.9
id1028358	1	42975052	13.68×10^{-4}	2.9	C	T	9.43	0.095	11.8

Notes: Chr.: Chromosome; LOD: Logarithm of the odds ratio; O-allele: Original allele; R-allele: Replaced allele; F: Allele frequency; MAF: Minor Allele Frequency; R^2: The ratio of expression of each QTL to the total expression of the corresponding characteristic; *: information of alleles were based on the Rice SNP-Seek database (IRRI)
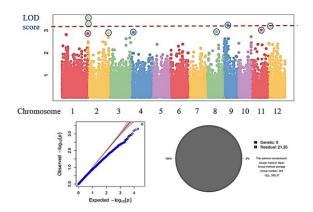
at P ≤ 0.001 and their allele frequencies ranged from 10.13 to 56.60% (Table 3). This shows that there was a difference in amylose content among medium grain rice varieties. Based on the Rice SNP-Seek database (IRRI), among nine identified SNPs, two (id2000259 and id2000249) on chromosome 2 and one SNP (id12001229) on chromosome 12 were selected. SNP id2000259 at position 349130 bp had the substitution of the T allele by the A allele with the frequency of 10.13% and explained 15.2% of the variations in AC of the medium-grain rice varieties. Similarly, SNP id2000249 has a substitution of the T allele by the C allele at position 346742 bp on chromosome 2. This allele had the substitution frequency of 11.99% and explained 13.9% of the variations. On chromosome 12, SNP id12001229 had a substitution of A by C at position 2966521 bp, the frequency of substitution allele was 34.00% and the SNP explained 13.4% of the variations in phenotype. Among 3265 SNPs located on chromosome 6, 74 SNP positions were identified with LOD score ≥ 3 at the significance level p ≤ 0.001 (Fig. 6). The frequency of occurrence of alternative alleles of SNPs varied widely, from 3.8 to 67.0% (Table 4).

When investigation of SNPs focusing on chromosome 6, among 3265 discovered SNPs, 74 SNP positions were identified with LOD score ≥ 3 at the significance level p ≤ 0.001 (Fig. 6). The frequency of occurrence of alternative alleles of SNPs varied widely, from 3.8 to 67.0% (Table 4).

For the level of significance of the SNPs, a total of 74 SNPs were selected with LOD values ranging from 3.0 to 11.0. In which, 27 SNPs were with a LOD value of 3.0 ≤ LOD < 4.0 (accounting for 36.5%), 25 SNPs with 4.0 ≤ LOD < 5.0 (accounting for 33.8%), and 22 SNPs with LOD value ≥ 5.0 (accounting for 29.7%). Especially, the study showed that two SNPs *fd7* (1764762 bp) and *fd8* (1770007 bp) were located on the *Waxy* gene region (Fig. 7). The SNP fd7 had the highest LOD value among the SNPs (LOD = 10.96) that had the replacement of the G allele with the T allele at the frequency of 17.9% and explained 11.63% of the variations in AC. The SNP fd8 with a LOD value of 5.65 had a replacement of allele A with allele C at the frequency of 14.2% and explained 9.26% of phenotypic variations.

Discussion

The agronomic characteristics of these rice varieties had a big range and a large phenotypic variance. This is a rich source of variation that allowed the exploitation of desirable traits. Based on the morphological and agronomic characteristics, the breeder can predict and orient the hybridization between rice varieties to achieve the highest breeding efficiency. Anwar *et al.* (2009) and Oladosu *et al.* (2018) suggested that path analysis for rice yield can be calculated directly through directly or indirectly measurement by yield components. For rice varieties, short-

Fig. 4: Analysis results of single nucleotide polymorphism (SNP) based on GWAS for AC in medium-grain rice varieties on 12 chromosomes in rice
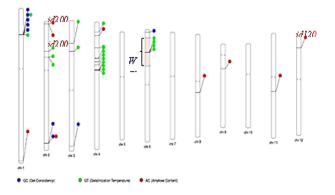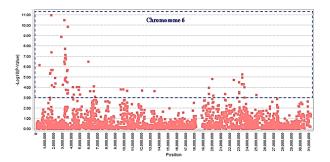

Fig. 5: QTL mapping of candidate SNPs related to amylose content, gel consistency, and gelatinization temperature in medium-grain rice varieties over 12 chromosomes
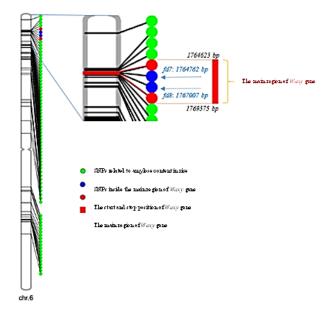
Notes: chr.: chromosome; AC: Amylose Content; GT: Gelatinization Temperature; GC: Gel Consistency; The region coverred by pink color was the expected region of Waxy gene

Fig. 6: Analysis results of single nucleotide polymorphism (SNP) based on GWAS for amylose content in medium-grain rice varieties on the chromosome 6 in rice

Fig. 7: QTL map of candidate SNPs related to AC in medium-grain rice varieties on chromosome 6

duration help to increase the number of crops per year, moderate plant height to prevent falling, and yield components are required such as long panicle, good tillering, high number of filled seeds, and big seed weight. The most respected characteristic is the actual yield, high yield helps to increase production efficiency and increase profits for farmers.

Besides the agronomic characteristics, cooking and eating quality is one of the important factors determining the value of rice grains. AC, GT and GC are three important quality parameters closely related to the softness of rice grains when cooked as well as cooled. AC among rice varieties was diversified and this result is similar to many previous studies (Manners 1979; Juliano 1992; Patindol *et al.* 2015). Through analysis of AC, it was found that in medium grain rice varieties, the group of varieties with low AC accounted for a lower percentage than other groups of amylose. This as a potential amylose group for breeding new rice varieties with low AC. Similarly, the value of GT and GC in the medium-grain rice varieties was

a large range. These results inferred that there are diversified quality characteristics in the medium rice to permit exploitation for rice breeding programs. A rice variety is considered a good cooking quality when converging factors such as low AC ($\leq$ 20%), low GT (score 6 to 7), soft GC (> 60 mm), aroma, and other valuable nutritious properties. This was also mentioned by Custodio *et al.* (2019).

In the correlation between AC and other traits, AC did not have a strong correlation with agronomic characteristics. This correlation was similarly noted in previous studies (Graham-Acquaah *et al.* 2018; Sharifi 2018).

AC was lowly correlated with GT and this result was similar to the conclusions of Jennings *et al.* (1979), Hossain *et al.* (2009); and Pang *et al.* (2016). AC was strongly associated with GC. This implied that the lower the AC, the greater the GC, the more flexible the rice, and vice versa. This result was similarly recorded in previous studies (Lapitan *et al.* 2009; Ritika *et al.* 2010; Zhang *et al.* 2020).

Table 4: Significant SNPs related to amylose content located on the chromosomes 6 in medium-grain rice varieties

No.	SNP ID	Position(bp)	P-value	LOD-value	O- allele	R- allele	F (%)	R^2
1	**fd7**	**1764762**	**1.09×10^{-11}**	**10.96**	**G**	**T**	**17.9**	**11.63**
2	id6002622	3251812	3.27×10^{-11}	10.49	C	T	18.9	9.57
3	id6002910	3638340	1.40×10^{-10}	9.85	G	A	20.8	9.08
4	id6002257	2905046	1.33×10^{-9}	8.88	T	C	17.9	8.23
5	id6002745	3330294	1.89×10^{-8}	7.72	G	A	22.6	7.47
6	id6001346	1793697	4.12×10^{-8}	7.39	C	T	8.5	7.10
7	id6002767	3364148	7.54×10^{-8}	7.12	G	T	24.5	6.79
8	id6002778	3377594	7.54×10^{-8}	7.12	A	G	24.5	6.79
9	id6002798	3408987	1.36×10^{-7}	6.87	A	G	25.5	6.86
10	id6002753	3332532	1.55×10^{-7}	6.81	G	A	22.6	6.51
11	id6002804	3414296	1.94×10^{-7}	6.71	A	G	24.5	6.40
12	id6002613	3244706	3.32×10^{-7}	6.48	T	A	23.6	6.21
13	id6003812	5976978	3.37×10^{-7}	6.47	T	A	12.3	6.20
14	id6002690	3289852	5.42×10^{-7}	6.27	C	A	23.6	6.28
15	id6000293	418306	7.15×10^{-7}	6.15	G	A	16.0	5.97
16	id6002888	3599544	2.08×10^{-6}	5.68	G	A	12.3	5.47
17	id6001434	1892145	2.09×10^{-6}	5.68	G	A	8.5	10.19
18	**fd8**	**1767007**	**2.22×10^{-6}**	**5.65**	**A**	**C**	**14.2**	**9.26**
19	wd6000096	3564084	2.59×10^{-6}	5.59	A	G	25.5	5.48
20	id6001256	1658758	4.10×10^{-6}	5.39	C	G	27.4	5.31
21	id6001209	1612742	4.76×10^{-6}	5.32	G	T	27.4	5.12
22	id6012152	23441321	5.59×10^{-6}	5.25	C	A	16.0	5.45
23	id6012159	23442920	1.11×10^{-5}	4.95	C	G	65.1	4.77
24	id6002810	3427968	1.13×10^{-5}	4.95	G	A	27.4	4.76
25	id6001645	2194965	1.23×10^{-5}	4.91	G	A	14.2	5.10
26	id6002750	3330720	1.43×10^{-5}	4.84	C	T	23.6	7.73
27	id6010574	20037449	1.58×10^{-5}	4.80	C	T	40.6	4.62
28	id6011692	22493522	1.93×10^{-5}	4.71	T	A	28.3	4.58
29	id6002701	3312612	2.03×10^{-5}	4.69	G	A	24.5	4.51
30	id6002807	3427237	2.55×10^{-5}	4.59	T	A	25.5	5.02
31	id6002874	3589970	3.24×10^{-5}	4.49	T	C	21.7	4.47
32	id6011886	23086388	3.28×10^{-5}	4.48	T	C	26.4	4.91
33	id6001672	2234192	4.27×10^{-5}	4.37	G	A	22.6	4.37
34	id6012170	23444308	4.89×10^{-5}	4.31	C	T	3.8	4.13
35	id6012182	23464841	4.89×10^{-5}	4.31	T	C	3.8	4.13
36	wd6003011	23402917	4.92×10^{-5}	4.31	T	C	3.8	4.17
37	id6012207	23610475	5.17×10^{-5}	4.29	G	A	3.8	4.30
38	id6002447	3078075	5.64×10^{-5}	4.25	C	G	11.3	4.07
39	id6001376	1832609	6.39×10^{-5}	4.19	A	G	13.2	4.02
40	ud6000315	6602235	8.16×10^{-5}	4.09	T	C	31.1	3.93
41	id6004221	6601428	8.33×10^{-5}	4.08	A	T	31.1	4.09
42	id6001567	2091584	9.19×10^{-5}	4.04	G	A	10.4	3.86
43	id6003135	4211403	9.63×10^{-5}	4.02	A	G	17.0	3.84
44	wd6000289	4683858	9.63×10^{-5}	4.02	T	C	17.0	3.84
45	id6003349	4732170	9.63×10^{-5}	4.02	C	T	17.0	3.84
46	id6003422	4929359	9.63×10^{-5}	4.02	G	A	17.0	3.84
47	id6012242	23696829	9.76×10^{-5}	4.01	T	G	42.5	3.85
48	id6010440	19662774	1.09×10^{-4}	3.96	A	G	29.2	3.82
49	id6006125	9652883	1.57×10^{-4}	3.80	G	A	13.2	3.62
50	id6006227	9931739	1.57×10^{-4}	3.80	C	A	36.8	3.94
51	id6002818	3430848	1.70×10^{-4}	3.77	A	G	26.4	3.75
52	id6003402	4869557	1.78×10^{-4}	3.75	A	T	16.0	3.85
53	id6002912	3638528	1.94×10^{-4}	3.71	G	T	59.4	3.50
54	ud6000539	12067053	2.01×10^{-4}	3.70	A	G	65.1	3.60
55	id6006423	10381968	2.12×10^{-4}	3.67	C	G	44.3	3.59
56	id6008098	13486344	2.29×10^{-4}	3.64	T	G	67.0	3.43
57	id6004211	6595944	2.32×10^{-4}	3.63	G	C	30.2	3.43
58	id6003901	6176014	2.33×10^{-4}	3.63	T	C	29.2	3.87
59	id6011935	23107424	2.87×10^{-4}	3.54	T	A	64.2	3.36
60	id6011933	23106969	3.73×10^{-4}	3.43	T	C	60.4	3.51
61	id6002930	3648304	3.75×10^{-4}	3.43	C	T	53.8	3.80
62	id6002885	3598148	3.92×10^{-4}	3.41	A	G	24.5	3.41
63	id6010438	19661511	4.30×10^{-4}	3.37	C	T	30.2	3.17
64	id6011277	21501803	4.46×10^{-4}	3.35	A	T	29.2	3.90
65	id6003463	5035086	4.90×10^{-4}	3.31	G	C	15.1	3.35
66	id6003170	4317102	5.16×10^{-4}	3.29	C	T	17.0	3.01
67	id6013459	25054910	5.56×10^{-4}	3.25	A	T	43.5	3.44
68	ud6000850	20159426	6.27×10^{-4}	3.20	C	T	22.6	3.07
69	id6004306	6724842	8.29×10^{-4}	3.08	G	C	59.4	2.95
70	id6004317	6732352	8.44×10^{-4}	3.07	T	C	60.4	2.91
71	id6003109	4162872	9.39×10^{-4}	3.03	C	A	21.7	3.00
72	id6012253	23700701	9.72×10^{-4}	3.01	C	T	54.7	3.18
73	id6001216	1632684	9.80×10^{-4}	3.01	T	C	16.0	2.88
74	id6011826	22864363	10.2×10^{-4}	2.99	G	A	53.8	3.66

Notes: Chr.: Chromosome; LOD: Logarithm of the odds ratio; O-allele: Original allele; R-allele: Replaced allele; F: Allele frequency; R^2: The ratio of expression of each QTL to the total expression of the corresponding characteristic

Therefore, when determining the AC or GC, the breeders can predict the range of the remaining parameter. Moreover, within the same AC group, rice varieties with softer GC were preferred (Morgante and Olivieri 1993; Hirano *et al.* 1998).

In studying the associations between phenotypic and genotypic variances, LOD score (-$\log_{10}$P) was applied to select the significant SNPs. LOD scores of 3 or more were often indicative of a QTL (quantitative trait loci), or a genetic region involved in the expression of the phenotype (Nurnberger *et al.* 2008). On 12 chromosomes in rice, GWAS detected nine candidate SNPs, of which, two (id2000259 and id2000249) on chromosome 2 and one SNP (id12001229) on chromosome 12 had the presence of information about a specific base substitution in the Rice SNP-Seek database. These SNPs related to AC on chromosomes 2 and 12 were identified in previous studies. QTL SNPs associated with AC on chromosome 2 had been recognized by the studies of Tan *et al*. (1999); Lee *et al.* (2003); Yun *et al*. (2016). Additionally, on chromosome 12, many studies also recorded significant SNPs related to AC including Lee *et al*. (2003); Wan *et al*. (2004); Wada *et al*. (2006). Therefore, there SNPs (id2000259, id2000249, and id12001229) have a high possibility of controlling AC in medium-grain rice.

Focusing on chromosome 6, significant SNPs were widely distributed on the target chromosome and concentrated near the *Waxy* target gene region from position 1,764,623 to 1,769,575 bp (http://shigen.nig.ac.jp). Among them, two SNP fd7 and fd8 were located right in the *Waxy* gene region. This indicated that the *Waxy* gene also contribute to control AC in the medium-grain rice.

The number of SNPs and their significance identified on 12 chromosomes were usually lower than those on only one target chromosome (chromosome 6). This can be explained by the number of chromosomes, the more chromosomes are considered, the fewer candidate SNPs are identified, and the lower the LOD score value of the SNPs.

Conclusion

The AC in the medium-grain rice varieties varied widely from 10.83 to 30.12%. In which, the group of varieties with low AC of less than 20% accounted for a low percentage. AC in this study was strongly correlated with GC but not with GT. The genome-wide association study on 12 chromosomes for AC in medium-grain rice varieties identified three SNPs, id2000259, id2000249 on chromosome 2, and id12001229 on chromosome 12. On chromosome 6, 74 SNPs were associated with differences in AC of rice varieties. In which, two SNPs, fd7 (1764762 bp) and fd8 (176707 bp), were located in the *Waxy* gene region. The candidate SNPs are the base to find the candidate genes and molecular markers related to AC in the medium-grain rice varieties.

Author Contributions

Bui Phuoc Tam carried out the experimental work, participated in the sequence alignment, and drafted the manuscript. Pham Thi Be Tu and Nguyen Thi Pha conceived of the study and participated in its design and coordination. The authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

All authors declare no conflict of interest.

Data Availability

Data presented in this study will be available on a fair request to the corresponding author.

Ethics Approval

Not applicable to this paper.

References

Anwar J, MA Ali, M Hussain, W Sabir, MA Khan, M Zulkiffal, M Abdullah (2009). Assessment of yield criteria in bread wheat through correlation and path analysis. *J Anim Plant Sci* 19:185–188

Cai HY, DZ Xu, L Zhou, JP Cheng, ZH Zhang, JP Wu, AQ You (2015). Development of PCR-based SNP marker of rice *Waxy* gene with confronting two-pair primers. *Rus J Gen* 51:787–791

Courtois B, A Audebert, A Dardou, S Roques, T Ghneim-Herrera, G Droc, J Frouin, L Rouan, E Goze, A Kilian, N Ahmadi, M Dingkuhn (2013). Genome-wide association mapping of root traits in a japonica rice panel. *PLoS One* 8:e78037

Custodio MC, RP Cuevas, J Ynion, AG Laborte, ML Velasco, M Demont (2019). Rice quality: How is it defined by consumers, industry, food scientists, and geneticists? *Trends Food Sci Technol* 92:122–137

Fan CC, XA Yu, YZ Xing, CG Xu, LJ Luo, Q Zhang (2005). The main effects, epistatic effects and environmental interactions of QTLs on the cooking and eating quality of rice in a doubled haploid line population. *Theor Appl Genet* 110:1445–1452

Feng F, Y Li, X Qin, Y Liao, HM Siddique (2017). Changes in rice grain quality of indica and japonica type varieties released in China from 2000 to 2014. *Front Plant Sci* 8:1863

Graham RD (2002). *A Proposal for IRRI to establish a grain quality and nutrition research center.* IRRI discussion paper series, No. 44, International Rice Research Institute, Los Banos, The Philippines

Graham-Acquaah S, K Saito, K Traore, I Dieng, A Alognon, S Bah, A Sow, JT Manful (2018). Variations in agronomic and grain quality traits of rice grown under irrigated lowland conditions in West Africa. *Food Sci. Nutr* 6:970–982

Hirano HY, M Eiguchi, Y Sano (1998). A single base change altered the regulation of the *Waxy* gene at the post-transcriptional level during the domestication of rice. *Mol Biol Evol* 15:978–987

Hossain MS, AK Singh, Fasih-uz-Zaman (2009). Cooking and eating characteristics of some newly identified inter sub-specific (indica/japonica) rice hybrids. *Sci Asia* 35:320–325

Huang X, X Wei, T Sang, Q Zhao, Q Feng, Y Zhao, C Li, C Zhu, T Lu, Z Zhang (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet* 42:961–967

IRRI (International Rice Research Institute) (2014). *Standard Evaluation System for Rice*, 5th edn. International Rice Research Institute, Los Banos, The Philippines

IRRI (International Rice Research Institute) (2013). Rice science for a better world, IRRI Annual Report [e-book]. Available at: http://books.irri.org/AR2013_content.pdf. (Accessed: 29 January 2020)

Jennings PR, WR Coman, HE Kauman (1979). *Rice Improvement,* 1st edn. International Rice Research Institute, Los Banos, Philippines

Juliano BO (1971). A simplified assay for milled-rice amylose. *Cereal Sci Today* 16:334–340

Juliano BO (1992). Rice starch properties and grain quality. *Denpun Kagaku* 39:11–21

Juliano BO, CM Perez, AB Blakeney, DT Castillo, N Kongseree, B Laignelet, ET Lapis, VVS Murty, CM Paule, BD Webb (1981). International cooperative testing on the amylose content of milled rice. *Starch/Starke* 33:157–162

Kumar V, A Singh, SVA Mithra, SL Krishnamurthy, SK Parida, S Jain, KK Tiwari, P Kumar, AR Rao, SK Sharma, JP Khurana, NK Singh, T Mohapatra (2015). Genome-wide association mapping of salinity tolerance in rice (*Oryza sativa*). *DNA Res* 22:133–145

Lapitan VC, ED Redoña, T Abe, DS Brar (2009). Mapping of quantitative trait loci using a doubled-haploid population from the cross of indica and japonica cultivars of rice. *Crop Sci* 49:1620–1628

Lee JS, WG Ha, JK Chang, KL Ryu, JH Cho, Y-C Song, O-K Kwon, SJ Yang, HY Kim, HS Suh (2003). *QTL Analysis for Grain Quality Properties in a Japonica Rice Combination*. Available at: http://www.regional.org.au/au/asa/2004/poster/3/4/1/865_leejs.htm. [Accessed 29 January 2020]

Manners DJ (1979). The enzymic degradation of starches. In: *Polysaccharides in Food*, pp:75–91. JMV Blanshard, JR Mitchell (Eds.). Butterworths, London

Morgante M, AM Olivieri (1993). PCR – amplified microsatellites as markers in plant genetics. *Plant J* 1:175–182

Nurnberger JI, W Berrettini, AB Niculescu (2008). Genetics of psychiatric disorders. In: *The Medical Basis of Psychiatry,* 3rd edn, pp:487–518. Fatemi SH, PJ Clayton, N Sartorius (Eds.). Springer Nature, Berlin, Germany

Oladosu Y, MY Rafii, U Magaji, N Abdullah, G Miah, SC Chukwu, G Hussin, A Ramli, I Kareem (2018). Genotypic and phenotypic relationship among yield components in rice under tropical conditions. *BioMed Res Intl* 00:8936767

Pang Y, J Ali, X Wang, NJ Franje, JE Revilleza, J Xu, Z Li (2016). Relationship of rice grain amylose, gelatinization temperature and pasting properties for breeding better eating and cooking quality of rice varieties. *PLoS One* 11:e0168483

Patindol JA, TJ Siebenmorgen, YJ Wang (2015). Impact of environmental factors on rice starch structure: A review. *Starch/Starke* 67:42–54

Ritika BY, BS Khatkar, BS Yadav (2010). Physicochemical, morphological, thermal and pasting properties of starches isolated from rice cultivars grown in India. *Intl J Food Props* 13:1339–1354

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

Septiningsih EM, KR Trijatmiko, S Moeljopawiro, SR McCouch (2003). Identifcation of quantitative trait loci for grain quality in an advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative O. *rufipogon*. *Theor Appl Genet* 107:1433–1441

Sharifi P (2018). Sequential path analysis for determination of relationships between yield-related characters and yield and amylose content in rice. *Phil J Crop Sci* 43:73–79

Tan YF, JX Li, SB Yu, YZ Xing, CG Xu, Q Zhang (1999). The three important traits for cooking and eating quality of rice grains are controlled by a single locus in an elite rice hybrid, Shanyou 63. *Theor Appl Genet* 99:642–648

Tang SX, GS Khush, BO Juliano (1991). Genetics of gel consistency in rice (*Oryza sativa* L.). *J Genet* 70:69–78

Tian Z, Q Qian, Q Liu, M Yan, X Liu, C Yan, G Liu, Z Gao, S Tang, D Zeng, Y Wang, J Yu, M Gu, J Li (2009). Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. *Proc Nat Academy Sci* 106:21760–21765

Ueda Y, F Frimpong, Y Qi, E Matthus, L Wu, S Holler, T Kraska, M Frei (2014). Genetic dissection of ozone tolerance in rice (*Oryza sativa* L.) by a genome-wide association study. *J Exp Bot* 66:293–306

Vanniarajan C, KK Vinod, A Pereira (2012). Molecular evaluation of genetic diversity and association studies in rice (*Oryza sativa* L.). *J Genet* 91:9–19

Wada T, Y Uchimura, T Ogata, M Tsubone, Y Matsue (2006). Mapping of QTLs for physicochemical properties in japonica rice. *Breed Sci* 56:253–260

Wan XY, JM Wan, CC Su, CM Wang, WB Shen, JM Li, HL Wang, L Jiang, SJ Liu, LM Chen, H Yasui, A Yoshimura (2004). QTL detection for eating quality of cooked rice in a population of chromosome segment substitution lines. *Theor Appl Genet* 110:71–79

Yu S, Y Ma, L Menager, DW Sun (2012). Physicochemical properties of starch and flour from different rice cultivars. *Food Bioprocess Technol* 5:626–637

Yun Y-T, C-T Chung, Y-J Lee, H-J Na, J-C Lee, S-G Lee, K-W Lee, Y-H Yoon, J-W Kang, H-S Lee, J-Y Lee, S-N Ahn (2016). QTL mapping of grain quality traits using introgression lines carrying oryza rufipogon chromosome segments in japonica rice. *Rice* 9:62

Zhang A, Y Gao, Y Li, B Ruan, S Yang, C Liu, B Zhang, H Jiang, G Fang, S Ding, N Jahan, L Xie, G Dong, Z Xu, Z Gao, L Guo, Q Qian (2020). Genetic analysis for cooking and eating quality of super rice and fine mapping of a novel locus *QGC10* for gel consistency. *Front Plant Sci* 11:342

Zhao K, CW Tung, GC Eizenga, MH Wright, ML Ali, AH Price, GJ Norton, MR Islam, A Reynolds, J Mezey (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat Commun* 2:467

Zhiwu Zhang Laboratory (2020). *User manual for GAPIT – Genomic Association and Prediction Integrated Tool, Version 3*. Washington State University, Pullman, Washington, USA

Zhou L, S Chen, G Yang, W Zha, H Cai, S Li, Z Chen, K Liu, H Xu, A You (2018). A perfect functonal marker for the gene of intermediate amylose content Wx-In in rice (*Oryza sativa* L.). *Crop Breed Appl Biotechnol* 18:103–109