Grain quality-based selection and QTL mapping in a Ciherang-sub1 backcross population

Estria F Pramudyawardani1,2, H Aswidinnoor1*, B S Purwoko1, W B Suwarno1, H Verdeprado3, M R Islam3, B C Y Collard3,4
1 Department of Agronomy and Horticulture, Faculty of Agriculture, Bogor Agricultural University, Indonesia
2 Indonesian Center for Rice Research, Sukamandi, West Java, Indonesia
3 Plant Breeding Platform, International Rice Research Institute, Los Banos, Philippines
4 Yanco Agricultural Institute, NSW Department of Primary Industries, Yanco, Australia

E-mail: hajrial@apps.ipb.ac.id

Abstract. Ciherang has been a dominant variety in Indonesia for almost 2 decades and is considered to be a premium rice in Indonesia. Despite the importance of this variety, limited information is available regarding the genetic control of grain quality traits. A backcross population (Ciherang-Sub1*2/IRRI148) was developed for mapping a range of traits. This study was performed to select promising families with similar rice grain quality as Ciherang-Sub1 and to identify QTLs associated with grain quality from Ciherang-Sub1.

Nine quality traits (chalkiness; seed length, width, and shape; brown rice, milled rice percentage, head rice recovery percentage, amylose content) and 3 complementary data (gelatinization temperature, gel consistency, milling potential degree) were measured at the IRRI Grain quality and nutrition centre. Selection using two clustering methods identified 32 promising families which had similar grain quality as Ciherang-Sub1. Using the ICIM QTL mapping method, 12 grain quality QTLs were identified. Eight of those were newly identified QTLs (qSL2.1, qSL8.1, qSWd1.1, qBrP1.1, qBrP3.1, qBrP9.1, qMP2.1 and qGT6.2). Large effect QTLs were identified for the seed length (R² = 63.61%), seed shape (R² = 41.29%), and gelatinization temperature (R² = 41.52%). Both parents contributed alleles that affected quality, where 2 QTLs from seed length and width increased seed shape. The promising families and QTLs identified in this study could be useful for enhancing rice grain quality using molecular marker assisted breeding and developing new rice varieties with premium quality in the future.

1. Introduction

Rice grain quality is a critical consideration for the rice breeders [1,2]. The acceptance of a new rice variety depends on the farmers, millers and consumer’s preferences [3,4] and mostly led by grain quality and yield traits. The preferred physical quality and or eating quality would increase the economic value of a new rice variety. Studies on the connection of price with grain quality in Philippines showed that most of the middle to high class consumers would choose better eating quality rice [5]. This fact encourages rice researchers to develop new high yielding rice varieties with premium quality. Combining the grain quality QTLs pyramiding with marker assisted selection (MAS) method might be an effective approach for the development of high quality breeding lines in...
rice [6].

Ciherang is one of the most popular rice HYV in Indonesia [7], released in 2000 [8]. Preferred rice quality characteristic of Ciherang is one from many reasons for its wide adoption in Indonesia [7,9,10]. Ciherang-Sub1 was developed using marker-assisted backcrossing and has 87.5% similarity to Ciherang [11]. This variety has also been released in Bangladesh as BINAdhan 11 in 2013 and becoming popular due to its submergence tolerance ability and shorter growth duration (110 days) in Aman (rainfed) season [12]. So it is a good choice to select as a parent in breeding program or for QTL mapping. Selection of backcross families derived from Ciherang-Sub1 and the identification of QTLs for quality from this parent were the two main objectives of this study. This work extends a previous QTL mapping study targeting agronomic traits [13].

2. Materials and methods

2.1. Plant material

Rice grain quality measurements were performed on 147 BC1F1:2 families, derived from a Ciherang-Sub1*2/IRRI148 population (hereafter named “#109”). Both parents were included as checks. Full details regarding the development of this BC1F1 population were reported in Pramudyawardani et al. [13].

2.2. Quality data

The grain quality analysis was performed by the IRRI Grain quality and nutrition center (GQNC). The observed characters included physical traits (chalkiness, seed length, seed width, seed shape, percent brown rice, percent milled rice, percentage head rice recovery), milling potential degree, and cooking traits (amylose content, gelatinization temperature, gel consistency). The measurements were based on the IRRI grain quality standard laboratory procedure. (http://www.knowledgebank.irri.org/training/fact-sheets/postharvest-management/rice-quality-fact-sheet-category) and the scoring based on standard evaluation systems for rice [14]. Ciherang-Sub1 was used for the quality reference.

2.3. Selection based on cluster analysis

The family selection based on the cluster analysis, with Ciherang-Sub1 as parent reference. Ciherang-Sub1 was used as reference because of it superiority for the grain quality characteristics, compared to IRRI148. Cluster analysis is a statistical analysis that aims to separate objects into several groups that have similar or different characteristics between groups with one another (Juliano, 1991). The Gower dissimilarity method was used based on the diversity of the collected data (nominal and ordinal data type). The partition around medoids (PAM) method and Neighbor-Joining (NJ) method were used to classify the similar quality group and the closest family to Ciherang-Sub1. The appropriate numbers of grouping determined by the PAM silhouette. Families with similar quality group and had closest relatedness with Ciherang-Sub1 were selected in the breeding program.

2.4. QTL mapping

The previously published map of this population [13] was used, incorporating quality data as the phenotypic data. Full details of map construction and QTL analysis were reported in Pramudyawardani et al. [13]. Nine of 11 observed traits were included for the QTL mapping analysis, which were: chalkiness, seed length, seed width, seed shape, brown rice percentage, milled rice percentage, head rice recovery percentage, amylose content, and the gelatinization temperature. While gel consistency and milling potential degree were not be able to used due to the non-compliance of the normality data. The gelatinization temperature data was changed to ordinal data, where low temperature (L) classified as 0, intermediate temperature (I) classified as 2, and the I/L category classified as 1. The QTL nomenclature followed McCouch and CGSNL [14]. QTLs detected from this study were compared to the previous studies, based on the specific trait and map information, using the Gramene. (http://archive.gramene.org/qlt/) and QTARO databases (http://qtaro.abr.affrc.go.jp/qtab/table).
2.5. Data analysis
Simple descriptive statistics were used to illustrate the population heterogeneity. Histogram graphs were applied to inspect the distribution of main trait data and the normal Q-Q plot was applied to determine the normality of the data. The descriptive analysis calculation using SPSS 16.0. The clustering analysis calculation using PBSTAT-CL statistic software (freely available from www.pbstat.com). The QTL analysis was running by QTL ICIMapping V3.2 software [16], using the ICIM method (inclusive composite interval mapping). Significance levels for QTL analysis were determined using 5000 permutations.

3. Results and Discussion
Developing new rice varieties with preferred quality is a challenge for rice breeders, due to the broad characteristics of the “premium” rice classification [5,1]. Large number of new varieties have been released, but only few of them are widely-adopted by farmers; quality is one of the most important factors regarding adoption [3,4].

Rice quality is dependent upon the variety, environment, production, harvesting conditions and post-harvest management, milling, and marketing technique [17]. The original Ciherang variety is a well-known variety for its grain quality, broad adaptation and high yield potential, and Ciherang-Sub1 retained all of these characteristics with improved submergence tolerance [11]. IRRI148 was used as donor parent for enhancing Ciherang-Sub1 for drought tolerance, but combining them might reduce the quality of their offspring, because IRRI148 has lower quality class compared to Ciherang-Sub1. Selection of the backcross population should consider maintaining the Ciherang-Sub1 quality in their offspring, along with selection for other critical traits.

The early generation (BC1F1:2) was used in this study, because selection for quality traits can be very effective at early generations (Concepcion et al, 2015). Selecting preferable rice quality is a great challenge, especially for the eating preferences, because there are differences categories of “good taste” preference [2]. Various studies have been done to match the quality-characters demanded by consumers, but the suitable methods have not yet developed to reduce the gap [4]. In this study we used the commonly-measured traits for evaluating quality.

3.1. Descriptive statistics of rice quality.
Rice quality data of BC1F1:2 population and parents are shown in Table 1. The descriptive statistics illustrate the differences of quality between parents, where Ciherang-Sub1 has better seed shape and milling quality for Indonesia compared to IRRI148. The seed material was taken from the same field experiment. Both parents were classified as intermediate for the amylose content. Other data did not show much differences between parents, except for gelatinization temperature, where Ciherang-Sub1 was classified having intermediate gelatinization temperature while IRRI148 classified as low. Based on the mean value, most of the observed traits of the BC1F1:2 population have equal quality as Ciherang-Sub1, except for chalkiness (4.68%) and head rice recovery (13.04%) which had lower quality than Ciherang-Sub1 (chalkiness = 1.7% and head rice recovery = 61.85%). The mean BC1F1:2 percentage of milled rice mean (72.12%) was higher than Ciherang-Sub1 (70.75%).

3.2. Selection based on clustering analysis.
“Cluster analysis is the art of finding groups in data” [18]. The aim for clustering was to reduce the number of families for the selection necessities. This study used two different clustering methods, which is PAM and NJ method. The PAM method was aimed to sorted number of families into groups, while the NJ method was aimed classified the most similar families to Ciherang-Sub1. The head rice recovery percentage data was not included in the clustering analysis due to the low moisture content of the seeds, which leads to the lower head rice recovery percentage. From the PAM average silhouette width vs. number of clusters graph, indicated that the number of clustering group is 4, which
determined from the highest peak of the output. Ciherang-Sub1 was included in the first group and IRRI148 in the second group. Thirty-two families grouped with Ciherang-Sub1 and 38 families grouped with IRRI148 (Fig. 1). The neighbor-joining clustering method (NJ) was selected for the clustering method based on the correlations among dissimilarity and cophenetic distances, where NJ has the highest significant value among other methods (0.67). Clustering group resulted from the NJ method formed two major groups, where both parents were classified into the same group.

**Table 1.** Rice quality performances of BC1F1:2 population and parents

| Traits         | Parent means | BC1F1:2 |          |          |          |          |
|----------------|--------------|---------|----------|----------|----------|----------|
|                | Cih-Sub1     | IRRI148 | Mean     | SD       | Range    |          |
| Chalkiness (%) | 1.70         | 2.00    | 4.68     | ± 3.15   | 0.70 - 20.00 |          |
| Seed length (mm) | 6.82     | 6.11    | 6.48     | ± 0.22   | 6.04 - 6.93 |          |
| Seed width (mm) | 2.25      | 2.40    | 2.29     | ± 0.05   | 2.16 - 2.39 |          |
| Seed shape     | 3.04         | 2.55    | 2.83     | ± 0.13   | 2.54 - 3.19 |          |
| Brown rice (%) | 79.30        | 77.20   | 79.23    | ± 0.53   | 77.80 - 80.30 |          |
| Milled rice (%)| 70.75        | 69.05   | 72.12    | ± 0.69   | 69.90 - 74.00 |          |
| Head rice      | 61.85        | 51.10   | 17.84    | ± 13.04  | 1.00 - 64.60 |          |
| Amylose        | 22.10        | 23.75   | 23.11    | ± 1.09   | 17.10 - 25.60 |          |

**Figure 1.** Dendrogram of the BC1F1:2 population, based on the NJ clustering method

Selection was based on the both clustering methods used, where families which grouped on the same as Ciherang-Sub1 (from PAM method) and has closest distance (from NJ method) with Ciherang-Sub1 were selected. From this combined method resulted 32 families which has preferred rice quality categories, where 9 among them have 0.83 to 0.90 similarity value (Table 2). The selected families based on both clustering methods will be continued for the research material for the rice quality enhancing breeding program. This data will be combined with other agronomic data to refine selection for the breeding program.

3.3. Rice quality QTL mapping

Ten QTLs were identified for physical quality traits, and 2 QTLs were identified for cooking quality (Table 3). QTLs with large effect came from 4 seed length QTLs ($R^2 = 63.61\%$), 1 seed shape QTL ($R^2 = 41.29\%$), and 2 gelatinization temperature QTLs ($R^2 = 41.52\%$). Eight out of 12 QTLs were classified as newly identified in rice genome, based on the Gramene and QTARO rice QTL database.
Table 2. Nine closest families with similar quality as Ciherang-Sub1 (NJ method)

| Family | NJ value | Observed rice quality data |
|--------|----------|----------------------------|
|        |          | Chlk | SL | SWd | SS | Amy | BrP | MP | GT | GC | MPt |
| ChlSub1 | 1.7 | 6.8 | 2.2 | 3.0 | 22.1 | 79.3 | 70.8 | 2 | 79 | 1 |
| G103 | 0.90 | 1.6 | 6.7 | 2.3 | 3.0 | 22.5 | 79.1 | 71.7 | 2 | 100 | 1 |
| G87 | 0.88 | 3.4 | 6.7 | 2.3 | 3.0 | 22.4 | 78.9 | 71.0 | 2 | 100 | 1 |
| G25 | 0.88 | 3.1 | 6.7 | 2.3 | 3.0 | 22.5 | 79.4 | 71.4 | 2 | 100 | 1 |
| G140 | 0.86 | 3.2 | 6.8 | 2.2 | 3.1 | 24.0 | 79.4 | 72.0 | 2 | 100 | 1 |
| G149 | 0.86 | 2.5 | 6.5 | 2.3 | 2.9 | 22.1 | 79.1 | 72.0 | 2 | 100 | 1 |
| G146 | 0.85 | 4.3 | 6.8 | 2.2 | 3.0 | 23.5 | 79.4 | 72.0 | 1 | 75 | 1 |
| G136 | 0.84 | 4.1 | 6.2 | 2.3 | 2.9 | 22.1 | 79.1 | 71.7 | 2 | 100 | 2 |
| G53 | 0.83 | 3.1 | 6.8 | 2.2 | 3.1 | 22.9 | 79.4 | 72.0 | 1 | 75 | 1 |
| G75 | 0.83 | 3.8 | 6.7 | 2.3 | 2.9 | 23.6 | 79.0 | 70.9 | 2 | 100 | 1 |
| IRRI148 | 0.54 | 2.0 | 6.1 | 2.4 | 2.5 | 23.8 | 77.2 | 69.1 | 0 | 77 | 1 |

NJ value = similarity value came from NJ method; Chlk = chalkiness; SL = seed length; SWd = seed width; SS = seed shape; Amy = amylose content; BrP = brown rice percentage; MP = milled rice percentage; GT = gelatinization temperature; GC = gel consistency; MPt = milling potential degree.

Table 3. Rice quality QTL identified from Ciherang-Sub1 BC population

| Trait   | Chr | QTL name | Position (Mbp) | Flanking markers | LOD | PVE (%) | Add | Similar QTL |
|---------|-----|----------|----------------|------------------|-----|---------|-----|-------------|
|         |     |          |                | Left             | Right |       |      |             |
| Physical quality | | | | | | | | |
| SL      | 2   | qSL2.1   | 24.81          | id2008866        | id2010969 | 3.99 | 4.93 | -0.10 New |
|         | 3   | qSL3.1   | 13.98          | id3006510        | id3008333 | 3.53 | 7.46 | 0.13 gl3 |
|         | 3   | qSL3.2   | 21.98          | id3008333        | id3010345 | 26.16 | 47.30 | 0.13 qLWR |
|         | 8   | qSL8.1   | 23.20          | id8006485        | id8006789 | 3.25 | 21.32 | 0.30 New |
| SWd     | 1   | qSWd1.1  | 18.16          | id1010609        | id1011535 | 2.86 | 8.58 | -0.03 New |
| SS      | 3   | qLWR3.1  | 21.98          | id3008333        | id3010345 | 26.16 | 47.30 | 0.13 qLWR |
| BrP     | 1   | qBrP1.1  | 36.16          | id1022408        | id1023892 | 3.22 | 6.58 | 0.29 New |
|         | 3   | qBrP3.1  | 12.98          | id3006510        | id3008333 | 7.80 | 18.50 | 0.49 New |
|         | 9   | qBrP9.1  | 21.32          | id9007161        | id9007763 | 2.94 | 5.96 | 0.27 New |
| MP      | 2   | qMP2.1   | 8.81           | wd200409         | id2004774 | 2.80 | 6.91 | 0.36 New |
|         |     |          |                |                  |       |        |      |             |
| Cooking quality | | | | | | | | |
| GT      | 6   | asv6b    | 7.68           | id6003649        | id6006147 | 9.22 | 27.58 | 0.54 asv6b |
|         | 6   | qGT6.2   | 15.68          | id600870         | id6009683 | 4.89 | 13.94 | 0.38 New |

Chr: chromosome; LOD: logarithm of the odds ratio; PVE: phenotypic variance explained; SL = seed length; SWd = seed width; SS = seed shape; BrP = brown rice percentage; MP = milled rice percentage; GT = gelatinization temperature. Add (additive effect) = (x̅AA - x̅aa)/2.

Two novel QTLs were detected for seed length on chromosome 2 (qSL2.1 at 24.81 Mbp), and chromosome 8 (qSL8.1 at 23.20 Mbp); 1 novel QTL was detected for seed width on chromosome 1 (qSWd1.1 at 18.64 Mbp); 3 novel QTLs were detected for brown rice percentage at chromosome 1 (qBrP1.1 at 16.16 Mbp), chromosome 3 (qBrP3.1 at 12.98 Mbp), and chromosome 9 (qBrP9.1 at 21.32 Mbp); 1 novel QTL was detected for milled rice percentage at chromosome 2 (qMP2.1 at 8.81 Mbp); and 1 novel QTL was detected for gelatinization temperature on chromosome 6 (qGT6.2 at
15.68 Mbp). Two QTLs for grain shape (\textit{qSL2.1} and \textit{qSL8.1}) and one QTL for grain width were able to increase grain size.

Quality is a complex trait, including the physical, cooking, and sensory preferences [2]. The discovery of QTLs and selection using marker assisted selection (MAS) for quality will provide breeders with efficient tools to develop new premium rice varieties [19,20,4,21]. More extensive studies of gene/QTLs controlling the quality traits are needed to verify the QTLs detected in this study.

4. Conclusion

Two clustering methods were applied for the selection purposes and identified 32 selected families which had similar quality as Ciherang-Sub1. Nine among them were highly related to Ciherang-Sub1, namely: G103, G87, G25, G140, G149, G146, G136, G53, and G75. QTL mapping analysis identified 11 QTLs of 6 rice chromosomes where 8 QTLs were novel. While the other 4 QTLs have the same region position with the previous studies, namely: \textit{qSL3.1}, \textit{qSL3.2}, \textit{qLWR3.1}, and \textit{asv6b}. Both novel and previously identified rice quality QTLs are valuable for rice breeders as supporting data to develop high yielding varieties with preferred grain quality. Selected families from this study will be continued in breeding programs for enhancing rice grain quality using MAS. Rice grain quality-trait QTLs identified in our study could be useful developing the premium rice, especially for Indonesia and other South-east Asian consumers.

Acknowledgements

This study is a part of final assignment of IPB’s Doctoral degree funded by IAARD scholarship. The experiments and travels were funded by the Lee Foundation Scholarship.

References

[1] CALINGACION M, LABORTE A, NELSON A, et al. (2014) Diversity of global rice markets and the science required for consumer-targeted rice breeding. \textit{PLoS One}, 9 (1): 1–12.
[2] CONCEPCION J C T, OUK M, ZHAO D, et al. (2015) The need for new tools and investment to improve the accuracy of selecting for grain quality in rice. \textit{Field Crops Research}, 182: 60–67.
[3] MANZANILLA D O, PARIS T R, VERGARA G V, et al. (2011) Submergence risks and farmers’ preferences: implications for breeding Sub1 rice in Southeast Asia. \textit{Agricultural Systems}, 104(4): 335–347.
[4] ANACLETO R, CUEVAS R P, JIMENEZ R, et al. (2015) Prospects of breeding high-quality rice using post-genomic tools. \textit{Theor Appl Genet}, 128(8):1449-1466.
[5] CUEVAS R P, PEDE V O, MCKINLEY J, et al. (2016) Rice grain quality and consumer preferences: a case study of two rural towns in the Philippines. \textit{PloS One}, 11(3): 1–17.
[6] ZHAO X, ZHOU L, PONCE K, et al. (2015) The usefulness of known genes / QTLs for grain quality traits in an indica population of diverse breeding lines tested using association analysis. \textit{Rice}, 8 (29): 1-13.
[7] RUSKANDAR A. (2007) Penyebaran padi unggul baru di Jawa Barat. \textit{Warta Penelitian dan Pengembangan Pertanian}, 29 (3): 5–6. (in Bahasa)
[8] ICRR (Indonesian Center for Rice Research). (2010) Deskripsi Varietas Padi. Sukamandi: ICRR: 1-117. (in Bahasa)
[9] RUSKANDAR A. (2009) Varietas Ciherang makin mendominasi. \textit{Warta Penelitian dan Pengembangan Pertanian}, 31 (6): 11–13. (in Bahasa)
[10] INDRASARI S D. (2011) Mutu gizi dan mutu rasa beras varietas unggul Ciherang. \textit{Warta Penelitian Dan Pengembangan Pertanian}, 33 (2): 8–10. (in Bahasa)
[11] NUGRAHA Y, HIDAYATUN N, YULIANI D. (2017) Phenotypic performance of Ciherang-Sub1 near isogenic line as an adaptive variety for flooding conditions. Indonesian Journal of \textit{Agricultural Science}, 18 (1): 7–16.
[12] ISLAM M, BEGUM S, HASSAN A, et al. (2015) BINA dhan11 and BINA dhan12, two flash flood tolerant Aman rice varieties. Leaflet. (in Bengali)

[13] PRAMUDYAWARDANI E F, ASWIDINNOOR H, PURWOKO B S, et al (2018) Genetic analysis and QTL mapping for agronomic and yield-related traits in Ciherang-Sub1 rice backcross populations. *Plant Breed Biotech*, 6 (3): 177-192.

[14] IRRI (International Rice Research Institute). (2013) Standard Evaluation System (SES) for Rice. Los Banos Philippines: *IRRI*: 1-65.

[15] MCCOUCH S R, CGSNL (Committee on Gene Symbolization, Nomenclature and Linkage, Rice Genetics Cooperative). (2008) *Gene nomenclature system for rice*. Rice, 1:72–84.

[16] MENG L, LI H, ZHANG L, et al. (2015) QTL IciMapping: integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop Journal*, 3 (3): 269–283.

[17] BIRLA D S, MALIK K, SAINGER M, et al. (2017) Progress and challenges in improving the nutritional quality of rice (*Oryza sativa* L.). Critical Reviews in Food Science and Nutrition, 57(11): 2455–2481.

[18] KAUFMAN L, ROUSSEUW P J. 1990. Finding groups in data: an introduction to cluster analysis. *Biometrics*, New Jersey USA: A John Wiley and Sons, Inc.

[19] STAMP P, VISSER R. (2012) The twenty-first century, the century of plant breeding. *Euphytica*, 186(3): 585–591.

[20] ZHAO W G, JONG C W, KWON S W, et al. (2012) Association analysis of physicochemical traits on eating quality in rice (*Oryza sativa* L.). *Euphytica*. doi: 10.1007/s10681-012-0820-z.

[21] CALINGACION M, FANG L, QUIATHHON-BAEZA L, et al. (2015) Delving deeper into technological innovations to understand differences in rice quality. *Rice*, 8(6): 2–10.