Rapid and targeted introgression of *fgr* gene through marker-assisted backcrossing in rice (*Oryza sativa* L.)

Journal:	<i>Genome</i>
Manuscript ID	gen-2017-0100.R1
Manuscript Type:	Article
Date Submitted by the Author:	29-Jul-2017
Complete List of Authors:	Cheng, Acga; University of Malaya, Institute of Biological Sciences; Universiti Kebangsaan Malaysia, School of Biosciences and Biotechnology Ismail, Ismanizan; Universiti Kebangsaan Malaysia, School of Biosciences and Biotechnology; Universiti Kebangsaan Malaysia, Institute of Systems Biology (INBIOSIS) Osman, Mohamad; University Putra Malaysia, Department of Crop Science Hashim, Habibuddin ; Institut Penyelidikan and Kemajuan Pertanian Mohd Zainual, Nur Samahah; Institut Penyelidikan and Kemajuan Pertanian, Agrobiodiversity and Environment Research Centre; Universiti Kebangsaan Malaysia, School of Biosciences and Biotechnology
Is the invited manuscript for consideration in a Special Issue? :	This submission is not invited
Keyword:	<i>fgr</i> gene, fragrant, marker-assisted backcrossing, recurrent parent genome, rice

SCHOLARONE™
Manuscripts

**Rapid and targeted introgression of *fgr* gene through marker-assisted
backcrossing in rice (*Oryza sativa* L.)**

Acga Cheng^{1,2*}, Ismanizan Ismail^{2,3}, Mohamad Osman⁴, Habibuddin Hashim⁵ and Nur Samahah
Mohd Zainual^{2,6}

¹Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala
Lumpur, Malaysia. Email: acgacheng@um.edu.my

²School of Biosciences and Biotechnology, Faculty of Science, Universiti Kebangsaan
Malaysia, 43600 Bangi, Selangor Darul Ehsan, Malaysia.

³Institute of Systems Biology (INBIOSIS), Universiti Kebangsaan Malaysia, 43600 Bangi,
Selangor Darul Ehsan, Malaysia. Email: maniz@ukm.my

⁴Department of Crop Science, Faculty of Agriculture, Universiti Putra Malaysia, 43300
Serdang, Malaysia. Email: mohamad_osm@upm.edu.my

⁵Malaysia Agricultural Research and Development Institute (MARDI), 13200 Kepala Batas,
Pulau Pinang, Malaysia. Email: habibuddinhashim@gmail.com

⁶Agrobiodiversity and Environment Research Centre, MARDI Headquarters, 43300 Serdang,
Selangor Darul Ehsan, Malaysia. Email: nur.zanuar@gmail.com

*Corresponding author: Acga Cheng (email: acgacheng@um.edu.my)

Abstract

While it is crucial for developing countries like Malaysia to achieve self-sufficiency in rice (*Oryza sativa* L.), it is equally critical to be able to produce high-quality rice, specifically fragrant rice, which demands are often met through importation. The present study was aimed at developing high-yielding fragrant rice, in a timely and cost-effective manner. A marker-assisted backcross (MABC) approach was optimised to introgress the fragrance gene (*fgr*) into two high-yielding Malaysian varieties; MR84 and MR219, within two years utilising less than fifty molecular markers. Coupled with phenotypic screening, one single foreground marker (*fgr*-SNP) and forty-eight background markers were selected and utilised; revealing recovery of at least 90% of recurrent parent genome (RPG) in merely two backcross generations. Collectively, the yield potential of the developed BC₂F₂ lines (BLs) was higher ($P>0.05$) than the donor parent; MRQ74, and similar ($P<0.05$) to both the recurrent parents; MR84 and MR219. In addition, some of the developed BLs showed good grain quality, such as having long grain. We believe that this is the first report comprising the validation and utilisation of the single functional marker system (*fgr*-SNP) in introgressing the *fgr* gene into different rice varieties.

Keywords: *fgr* gene; fragrant; marker-assisted backcrossing; recurrent parent genome; rice

Introduction

In a world of rising population and booming economy, there is a great importance in developing sufficient key food staples with desirable qualities, and such is the case of rice (*Oryza sativa* L.). Being one of the three biggest cereal crops in the world, along with maize (*Zea mays*) and wheat (*Triticum aestivum*), the global rice production has kept pace with the demands of an increasing population over the past half-century by virtue of the Green Revolution (Biswajit et al. 2013; Muthayya et al. 2014). The current annual yield enhancement rate of rice, however, is showing signs of slowing down due mainly to the sweeping climate change (Khoury et al. 2014; Massawe et al. 2016). With two billion more people to feed by the mid-century, the increase in rice yield potential is vital not only to meet the immediate demands but also for the sustainability of the world food security (Massawe et al. 2016). Furthermore, a concurrent strong economic growth in many developing countries, such as India and Malaysia, has boosted demand for high-quality rice (Biswajit et al. 2013; Khush 2001). To such a degree, the need to develop rice variety with high-yielding and superior quality has become more urgent, and this will require a concerted and committed effort from rice breeders all over the world.

Fragrance, or aroma, is one of the most highly-valued grain quality traits that increased the popularity of the Basmati and Jasmine rice. The advent of molecular markers in the 1990s, coupled with the completion of the rice genome sequence in the mid-2000s, has facilitated the discovery of the functional genes associated with grain fragrance in rice (Hashemi et al. 2015). These include the major fragrance gene, *fgr*, which is a single recessive gene located on rice chromosome 8. This gene was reported to be responsible for the production of 2-acetyl-1-pyrroline (2-AP); the key fragrance constituent in cooked rice. The accumulation of 2-AP was reported to result from an eight base-pair deletion and three single nucleotide polymorphisms

(SNPs) in the seventh exon of betaine aldehyde dehydrogenase 2 (*badh2*) gene, which could be the *fgr* gene (Bergman et al. 2002; Bradbury et al. 2005; Cheng et al. 2014). A perfect marker system for fragrance genotyping; the *fgr*-SNP, has proved to be functional in several studies in discriminating between fragrant and non-fragrant individuals in segregating rice populations (Yeap et al. 2013; Lau et al. 2017). Nevertheless, to the best of our knowledge, a full-scale validation and utilisation of the single *fgr*-SNP system in introgressing the *fgr* gene into different high-yielding rice varieties has yet to be done in any breeding programme in Malaysia, and possibly beyond.

Phenotypic selection for fragrance in rice, commonly through sensory evaluation, can be challenging and complicated as the trait has a relatively low heritability (Yeap et al. 2013). The integration of effective molecular markers, along with phenotypic analysis, can increase the effectiveness of selective breeding (Hospital 2005; Lau et al. 2017). Marker-assisted backcrossing (MABC), the simplest form of marker-assisted selection (MAS), has enormous potential to introgress the *fgr* gene into diverse rice varieties (Cheng et al. 2015; Yeap et al. 2013). The past decade has seen numerous successful MABC programmes in rice, particularly in enhancing the resistance of the crop to certain diseases and pests, such as plant brown hopper resistance (Jairin et al. 2009) and blast resistance (Ragimekula et al. 2013; Tanweer et al. 2015). Due to the rapid evolution of pathogens, breeding strategies for durable resistance focus largely on increasing crop gene or genotype diversity to slow down the evolutionary changes (McDonald 2014). Rice breeding efforts in Malaysia over the past half-decade have been focused on improving its yield and resistance to biotic and abiotic stresses, and its quality enhancement has often been left out until recently (Cheng et al. 2015). In the present study, we described results of an optimised MABC breeding scheme in introgressing the *fgr* gene from a Malaysian

fragrant rice variety MRQ74 to two high-yielding Malaysian rice varieties; MR84 and MR219. Our results showed that more than 90% of recurrent parent genome (RPG) was generally recovered in two backcross generations, indicating that the *fgr*-SNP can be applied to a large-scale study in MABC to produce high-yielding fragrant rice varieties.

Materials and methods

Plant materials

All plant materials, including the parental varieties, F₁, and backcross lines, were obtained at the Malaysian Agricultural Research and Development Institute (MARDI), Seberang Perai, Malaysia. Fig. 1 shows the crossing scheme used in the present study to develop high-yielding fragrant BC₂F₂ lines (hereinafter referred as “BLs”). Two sets of BLs were derived from the crossing of one donor parent; MRQ74, with two recurrent parents; MR84 (hereinafter referred as “Cross-1”) and MR219 (hereinafter referred as “Cross-2”). MRQ74 is a long-grain fragrant variety, while MR84 and MR219 are short-grain and medium-grain non-fragrant varieties, respectively (Asfaliza et al. 2012; Shamsudin et al. 2016).

The selection process in this study involved both phenotypic and molecular tools (Fig. 1). Immediately upon confirming the hybridity of plants in the two developed F₁ populations, true hybrid heterozygous plants (Aa) were backcrossed with each of the recurrent parents (AA), producing seeds of the BC₁F₁ generation. Foreground selection was performed using the tightly linked marker *fgr*-SNP, consisting of four allele-specific primers including four allele-specific primers: external antisense primer (EAP); external sense primer (ESP); internal fragrant antisense primer (IFAP); and Internal non-fragrant sense primer (INSP) (Bradbury et al. 2005; Cheng et al. 2014). Five heterogeneous individual plants with the desired allele (Aa) were then backcrossed again; subsequently producing the seeds of the BC₂F₁ generation. The similar

selection steps were followed for the BC₂F₁ generation.

With an additional step of background selection, the best five to eight heterogeneous BC₂F₁ plants (Aa) were selected to be selfed; producing the seeds of the BC₂F₂ generation. The background selection to assess the recovery of recurrent parent genome (RPG) was performed based on the utilisation of 48 polymorphic simple sequence repeats (SSRs) selected from Cheng *et al.* (2014). The primer sequences are provided in Supplementary Table 1. At the final stage, homozygous BC₂F₂ individuals carrying the target allele (aa) and other desired traits such as appropriate plant height (approximately 70 to 90 cm) and medium- or long-grain were analysed for their agronomic performance (IRRI 2002). The best ten fragrant BLs developed from each of the crosses were reported here.

DNA extraction and PCR amplification

Leaves of 4-week-old rice seedlings were collected from the parental varieties and individuals from each generation for DNA isolation. Genomic DNA was extracted using genomic DNA extraction kit according to the manufacturer's protocol (Qiagen, USA). The quality of DNA samples was examined using a 1% agarose gel prepared in 1x Tris-Acetate-EDTA acid (TAE).

Polymerase chain reaction (PCR) amplification was carried out in 25 µl reaction mixtures, following the protocol described by Bradbury *et al.* (2005). Each PCR reaction contained 2.0 µl of genomic DNA, 5.0 µl of 1x Green GoTaq Flexi buffer, 1.5 µl of 25 mM MgCl₂, 0.5 µl of dNTP mix, and 0.25 µl of GoTaq Flexi DNA. The targeted fragments were amplified using a Mastercycler Gradient (Eppendorf, Germany) with the following protocol: 95 °C for 5 min, followed by 30 cycles of 91 °C for 1 min, 55 °C for 1.5 min and 72 °C for 2 min, and 5 min at 72 °C for the final extension. PCR products from *fgr*-SNP for foreground selection

were separated on standard 1% agarose gel at 100 V for 1 h, while products from the SSR markers for background selection were separated on high resolution 3% MetaPhor agarose gel at 80 V for 3.5 h. The 3% MetaPhor agarose gel was prepared following the protocol described by Cheng *et al.* (2012). The resolved PCR bands were detected by staining the agarose or MetaPhor agarose gels for 30 s with ethidium bromide (EtBr), followed by destaining for 30 min in distilled water and visualisation with a UV gel imager (Alpha Innotech, USA).

Phenotypic screening of parental varieties and BC₂ individuals

For phenotypic analysis, the presence or absence of fragrance was determined by sensory evaluation of rice leaves and grains according to methods of Sood (1978) and Golam *et al.* (2010), respectively, with minor modifications. Leaf aromatic test (hereinafter referred as “LAT”) was conducted by cutting about 0.2 g leaf samples into pieces and placing into 10 ml of 1.7% potassium hydroxide (KOH) for 10 min at room temperature. Subsequently, the samples were smelled and rated for fragrance by a panel of three experts on a scale of 1 to 3; where 1 represented absence of fragrance, 2 represented presence of mild fragrance, and 3 represented presence of strong fragrance. Separately, grain aromatic test (hereinafter referred as “GAT”) was performed by soaking fifty milled rice grains in 10 ml of 1.7% KOH for at least one hour. Like the LAT, the samples for GAT were also scored by three experts on a rating scale of 1 to 3. Both sensory tests were performed in triplicate.

Agronomic performance of the most promising fragrant BC₂F₂ lines

The developed fragrant BLs having a maximum recovery of RPG along with phenotypic similarity with the recurrent parent were used to evaluate the agronomic traits. Several

parameters related to yield and grain quality were recorded; including plant height (PLHT), days to 50% flowering (DFFL), days to maturity (DMT), grain length (GRLG), grain width (GRWH), grain shape (GRSP), thousand-grain weight (GRWT), yield per plant (YPL), biomass yield (BY), and harvest index (HI). These parameters were recorded from the ten best-selected BLs, along with the parents; MRQ74, MR84 and MR219.

Statistical Analysis

An analysis for the goodness of fit to the expected ratio of 1:2:1 was calculated for each BC₂F₂ populations using the chi-square test. The recorded BC₂F₂ segregation data were subjected to descriptive statistics and analysis of variance (ANOVA). The mean difference for the selected best BLs and the recurrent parents MR84 and MR219 was analyzed using *t*-test. All analysis was performed using Minitab 16 (Minitab Inc., USA).

Results

Foreground and background selection

The *fgr*-SNP marker used in the present study showed polymorphism between the donor parent MRQ74 and the two recurrent parents, MR84 and MR219. Fig. 2 shows an example of amplified *fgr*-SNP products from MRQ74, MR219, and 28 BLs derived from the cross between these varieties. Product size of ~580 bp observed for all the genotypes representing the positive control from the *fgr*-SNP primers, amplified by the two external primers EAP and ESP (Bradbury et al. 2005; Sathivel et al. 2009). In both BC₁F₁ and BC₂F₁ generation, five best plants were selected to be further crossed. Among the 150 plants in BC₂F₂ population from Cross-1, 39 of them had homozygous *fgr* alleles (aa) similar to MRQ74, 79 plants were heterozygotes (Aa), and 32 plants

had homozygous *fgr* alleles (AA) similar to MR84. The ratio of 39:79:32 agreed with the expected 1:2:1 segregation ratio according to the chi-square test ($\chi^2=1.080 < \chi_{0.052}=3.841$). In Cross-2, 31, 77, and 42 BC₂F₂ plants had aa, Aa, and AA alleles, respectively. The ratio of 31:77:42 was also in agreement with the expected 1:2:1 segregation ratio ($\chi^2=1.720 < \chi_{0.052}=3.841$).

For background selection, a total of 48 SSRs were selected from our mapping study (Cheng et al. 2014), which aimed to identify quantitative trait loci (QTLs) for fragrance, along with the other two major quality traits namely amylose content and cooked grain elongation. In the mapping study, 96 out of 212 selected markers were found to be polymorphic and distributed over the twelve rice chromosomes, covering 2086.8 cM of the genome. Half of the 96 polymorphic markers, which were unlinked to the *fgr* gene, were utilised in the present study to determine the recovery of the RPG. As recommended in Neeraja *et al.* (2007), at least three markers per rice chromosome were selected. The average distance between adjacent markers ranged between 4.9 cM and 39.6 cM. All the 48 selected markers were polymorphic between the donor and each of the recurrent parents. The average recovery of the genome of MR84 and MR219 in the selected fragrant BLs was 91.9% and 90.2%, respectively. The genomic proportions of the parents in these lines are shown in Table 1.

Phenotypic screening of parental varieties and BLs

In both LAT and GAT analyses, the donor parent MRQ74 having the *fgr* gene, expressed a strong fragrance with a score of 3, while both the recurrent parents did not express fragrance with a score of 1. The each ten selected BLs from each of the crosses carrying the *fgr* gene expressed a strong fragrance with a score of 3 (Table 1). However, a couple of other non-selected

BLs from each of the crosses carrying the *fgr* gene were found to have only mild-fragrance in both LAT and GAT, with a sensory score of 2 (Fig. 3).

Agronomic performance of the developed fragrant BLs

The agronomic performance of the parental varieties and the most promising fragrant BLs from each of the crosses is presented in Table 2. Collectively, all traits, with the exception of BY, showed significant differences ($P < 0.05$) among the populations. The fragrant BLs from Cross-1 had similar yield-related traits; including DMT, GRWT, and YPL, with the recurrent parent, MR84. However, the BLs showed significantly better GRGL, GRWH, and GRSP than MR84; and these numbers were similar ($P < 0.05$) to the donor parent MRQ74. Separately, fragrant BLs from Cross-2 did not have similar ($P > 0.05$) grain quality traits with the donor parent MRQ74 but had similar GRSP ($P < 0.05$) with the recurrent parent MR219. Other similar traits between these lines with MR219 were DFFL, DMT, and YPL. Nevertheless, significant differences ($P > 0.05$) were found for PLHT, GRGL, and GRWH, between the BLs and MR219. Both developed BC_2F_2 populations had significantly different PLHT than the respective parents within their cross.

Discussion

As the economies of developing countries are on the rise, the demand for staple food with improved quality is projected to grow rapidly in the next several decades (Floros et al. 2010). For countries like Malaysia which aim to be self-sufficient in rice, a concerted effort to develop rice varieties with good yield and quality through rapid and efficient breeding strategies is, therefore, crucial. To the best of our knowledge, this is the first report of a full-scale validation and utilisation of a single functional marker system, *fgr*-SNP, in introgressing the *fgr* gene into different rice varieties. The selection of parents is often one of the critical steps in breeding

programmes to develop superior lines (Shamsudin et al. 2016). It should be noted that the lack of desirable recombinants from many crosses involving Basmati varieties and modern varieties is one of the major barriers in breeding high-yielding fragrant rice (Nematzadeh et al. 2004). Two crosses were established in the present study, one of which was a controlled cross (Cross-1). A highly fragrant Malaysian rice variety MRQ74, also known as Maswangi, was selected as the sole donor parent in this study. This variety possesses several desirable grain quality traits, notably for its long and slender grain, but has relatively low yield potential (Shamsudin et al. 2016). The MR84 was selected as the recurrent parent in the control cross due mainly to its rapid grain filling ability (Teo et al. 2011). The recurrent parent in the other cross was MR219, an elite Malaysian rice variety with superior yield and good grain quality (Alias et al. 2001). Our ultimate aim in the present study was to develop fragrant BLs with the yield potential of MR219, and our results demonstrated that it could be achieved within a two-year time span.

The foreground marker used in this study; *fgr*-SNP, has been validated both in our mapping (Cheng et al. 2014) and multiplexing (Cheng et al. 2015) studies. In our QTL analysis, we figured that the most effective QTL for the fragrance trait was located on chromosome 8 between markers RM223 and AROE, which is consistent with most of the literature (Amarawathi et al. 2008; Jain et al. 2006; Sakthivel et al. 2009). Based on the chi-square analysis on the BC₂F₂ generation data, the segregation ratios obtained from both Cross-1 and Cross-2 were in good agreement with the theoretical ratio of 1:2:1. This is what was expected, given that *fgr*-SNP used in the present studies was a functional marker for fragrance trait that has been validated in numerous studies (Cheng et al. 2015; Yeap et al. 2013). Nevertheless, foreground marker segregation distortion was observed in some previous studies, and this may arise due to environmental effects (Jin et al. 2010; Lau et al. 2017). The segregation ratios in the present

study were in close agreement with the expected ratio perhaps because all the backcross plants were grown in Seberang Perai; one of the few major rice cultivation areas in Malaysia where the environment is particularly suited for growing rice.

In general, RPG recovery can be accelerated by using markers for background selection (Servin and Hospital 2002). A couple of well-placed markers, about three to four markers on a chromosome within 100 cM, can provide adequate coverage of the genome in a backcross programme (Neeraja et al. 2007; Servin and Hospital 2002). Hinged on the background selection analysis in the present study, the average recovery of RPG for both parents in selected fragrant BLs was higher than the theoretical average value (i.e. 87.5%) after two generations of backcrossing. Our results were in agreement with many other MAS studies in rice (Ellur et al. 2016; Rajpurohit et al. 2011) demonstrating that continued selection in a self-pollinated BC₂ generation with the aid of molecular markers would lead to a higher recovery in RPG. Nonetheless, some studies have reported otherwise (Jairin et al. 2009), and this could be because the background screening was not done in the early backcross generations.

In the phenotypic analysis, the recurrent parents MR84 and MR219 were scored 1 for both LAT and GAT, representing non-fragrant; and the donor parent MRQ74 scored 3 as fragrant. All of the selected fragrant BLs from both crosses were scored 3 (Table 1). Only a handful of non-selected BLs was found to have mild-fragrance with a score of 2 in both LAT and GAT. This suggests the reliability of the *fgr*-SNP marker in determining the fragrant and non-fragrant rice, which supported the findings from previous studies (Cheng et al. 2015; Yeap et al. 2013). Nevertheless, the results are in contradiction with a small number of the previous studies which reported that some backcross lines that carried the homozygous *fgr* alleles were scored as non-fragrant, whereas some lines without the alleles were scored as mildly fragrant (Rajpurohit

et al. 2011).

The MABC approach in the present study has clearly demonstrated an ability to accelerate the breeding process of high-yielding fragrant lines. Within two backcross generations, a considerable increase in YPL was observed among the selected fragrant BLs derived from both Cross-1 and Cross-2. As compared to the donor parent MRQ74, the average increase of YPL among the two BC₂F₂ populations was 32.5%. In addition, the selected fragrant BLs showed significantly higher HI than MRQ74 ($P < 0.05$). Interestingly, the PLHT of these improved lines were significantly different than the parents ($P < 0.05$), whereby their height was higher than the donor parent but lower than the recurrent parent (MR84 or MR219). A similar finding was reported in Sabu *et al.* (2006), whereby the BC₂F₂ families in their study showed different mean values of culm height than both parents; *Oryza rufipogon* ($P < 0.05$) and MR219 ($P < 0.01$). The optimal plant height for rice ranges between 80 cm and 90 cm (Sabu et al. 2006), and the developed BLs in the present study fell within this optimal range. The DFFL and DMT of the selected fragrant BLs from both crosses were similar ($P < 0.05$) to their respective recurrent parent, and significantly lower than the donor parent ($P > 0.05$). These results are in line with a recent backcross breeding programme in Malaysia, where the DFFL and DMT of the developed backcross lines closely resembled the recurrent parent MR219, with a genome recovery of more than 95% (Tanweer et al. 2015).

The grain quality traits of the selected fragrant BLs derived from Cross-1, including GRGL, GRWH, and GRSP, were significantly different than MR84 ($P < 0.05$), and they might inherit these traits from MRQ74. Nevertheless, these lines possessed the same GRWT as MR84, which is known to have a rapid grain filling ability (Teo et al. 2011). On the other hand, selected fragrant BLs derived from Cross-2 had similar GRSP ($P > 0.05$) but higher GRGL and GRWH

($P < 0.05$) than MR219. In contrast to those from Cross-1, they did not show similarity to MRQ74 on any of the grain quality traits. It is worth noting that these lines had significantly higher GRWT ($P > 0.05$) than both parents; demonstrating that these lines have a great potential to be further developed.

Conclusions

The results of the present study, in general, provides evidence of accuracy and reliability for the *fgr*-SNP marker to be applied directly to large-scale MABC programmes for the development of high-yielding fragrant varieties. Within only two backcross generations, at least 90% of the recurrent parents' genomes were recovered, and fragrant BLs were developed; demonstrating that introgression of *fgr* gene with MABC breeding is much faster than that of conventional breeding. For the most part, the developed fragrant BLs showed better yield-related traits than the donor parent MRQ74; possessing a similar yield potential with the recurrent parent MR219, which has been commercially grown by local farmers since its establishment in 2001. These BLs also showed better grain quality than MR219. The present study has overall, provided a clear, fast, and yet affordable route to introgressing *fgr* gene into rice genotypes, and this would benefit researchers especially those with limited resources.

Acknowledgements

The authors would like to thank the Ministry of Agriculture and Agro-based Industry Malaysia for their research fund support, and Malaysia Agricultural Research and Development Institute (MARDI) for providing the seeds and field facilities support.

326 **References**

- 327 Alias, I., Mohammad, H., Othman, O., Saad, A., Habibuddin, H., Azlan, S., et al. 2001.
328 Pembentukan dan prestasi varietas padi baru MR219. MARDI Report. 196.
- 329 Amarawathi, Y., Singh, R., Singh, A.K., Singh, V.P., Mohapatra, T., Sharma, T.R., et al. 2008.
330 Mapping of quantitative trait loci for basmati quality traits in rice (*Oryza sativa* L.). Mol Breed.
331 21: 49-65.
- 332 Asfaliza, R., Rafii, M.Y., Saleh, G., Omar, O., Puteh, A. 2012. Combining ability and heritability
333 of selected rice varieties for grain quality traits. Aus J Crop Sci. 6(12): 1718-1723.
- 334 Bergman, C.J., McClung, A.M., Pinson, S.R., Fjellstrom, R.G. 2002. Development of PCR-
335 based markers associated with cooked rice kernel elongation and aroma. Proc RTWG, Arkansas.
- 336 Bishwajit, G., Sarker, S., Kpoghomou, M., Gao, H., Jun, L., Yin, D., et al. 2013. Self-sufficiency
337 in rice and food security: a South Asian perspective. Agric. Food Secur. 2:10.
- 338 Bradbury, L.M.T., Henry, R.J., Jin, Q., Reinke, R.F., Waters, D.L.E. 2005. A perfect marker for
339 fragrance genotyping in rice. Mol. Breed. 16: 279-283.
- 340 Cheng, A., Ismanizan, I., Mohamad, O., Habibuddin, H. 2012. Simple and rapid molecular
341 techniques for identification of amylose levels in rice varieties. Int. J. Mol. Sci. 13: 6156-6166.
- 342 Cheng, A., Ismanizan, I., Mohamad, O., Habibuddin, H. 2014. Mapping of quantitative trait loci
343 for aroma, amylose content and cooked grain elongation traits in rice. Plant Omics J. 7(3):152-
344 157.

- 345 Cheng, A., Ismanizan, I., Mohamad, O., Habibuddin, H. 2015. High resolution agarose-based
346 system for single-tube genotyping of *fgr* and *Waxy* genes in rice: MAGE to displace PAGE?.
347 Plant Omics J. 8(4): 348-352.
- 348 Ellur, R.K., Khanna, A., Yadav, A., Pathania, S., Rajashekara, H., Singh, V.K., et al. 2016.
349 Improvement of Basmati rice varieties for resistance to blast and bacterial blight diseases using
350 marker assisted backcross breeding. Plant Sci. 242: 330-341.
- 351 Floros, J.D., Newsome, R., Fisher, W., Barbosa-Canovas, G.V., Chen, H., Dunne, C.P., et al.
352 2010. Feeding the world today and tomorrow: The importance of food science and technology.
353 An IFT scientific review. Compr. Rev. Food Sci. F. 9: 572-599.
- 354 Golam, F., NorZulaani, K., Jennifer, A.H., Subhal, B., Zulqarnain, M., Osman, M., et al. 2010.
355 Evaluation of kernel elongation ratio and aroma association in global popular aromatic rice
356 cultivars in tropical environment. Afr. J. Agric. Res. 5(12): 1515-1522.
- 357 Hashemi, G.F.S., Rafii, M.Y., Ismail, M.R., Mohamed, M.T.M., Rahim, H.A., Latif, M.A., et al.
358 2015. Application of an effective statistical technique for an accurate and powerful mining of
359 quantitative trait loci for rice aroma trait. PLOS ONE. 10(6): e0129069.
- 360 Hospital, F. 2005. Selection in backcross programmes. Philos. Trans. R. Soc. Lond. B. Biol. Sci.
361 360(1459): 1503-1511.
- 362 IRRI. 2002. Standard evaluation system for rice. International Rice Research Institute, Los
363 Banos, Philippines.

- 364 Jain, N., Jain, S., Saini, N., Jain, R.K. 2006. SSR analysis of chromosome 8 regions associated
365 with aroma and cooked kernel elongation in Basmati rice. *Euphytica*. 152: 259-273.
- 366 Jairin, J., Teangdeerith, S., Leelagud, P., Kothcharerk, J., Sansen, K., Yi, M., Vanavichit, A., et
367 al. 2009. Development of rice introgression lines with brown planthopper resistance and
368 KDML105 grain quality characteristics through marker-assisted selection. *F. Crop Res.* 110:
369 263-271.
- 370 Jin, L., Lu, Y., Xiao, P., Sun, M., Corke, H., Bao, J.S. 2010. Genetic diversity and population
371 structure of a diverse set of rice germplasm for association mapping. *Theor. Appl. Genet.* 121:
372 475-487.
- 373 Khoury, C.K., Bjorkman, A.D., Dempewolf, H., Ramirez-villegas, J., Guarino, L., Jarvis A., et
374 al. 2014. Increasing homogeneity in global food supplies and the implications for food security.
375 *PNAS*. 111: 4001-4006.
- 376 Khush, G.S. 2001. Green revolution: the way forward. *Nat. Rev. Genet.* 2: 815-822.
- 377 Lau, W.C.P., Rafii, M.Y., Ismail, M.R., Puteh, A., Latif, M.A., Asfaliza, R., et al. 2017.
378 Development of advanced fragrant rice lines from MR269 $\times$ Basmati 370 through marker-
379 assisted backcrossing. *Euphytica*. 213:11.
- 380 Massawe, F., Mayes, S., Cheng, A. 2016. Crop diversity: an unexploited treasure trove for food
381 security. *Trends Plant Sci.* 21(5): 365-368.
- 382 McDonald, B.A. 2014. Using dynamic diversity to achieve durable disease resistance in
383 agricultural ecosystems. *Trop. Plant Pathol.* 39(3): 191-196.

- 384 Muthayya, S., Sugimoto, J.D., Montgomery, S., Maberly, G.F. 2014. An overview of global rice
385 production, supply, trade, and consumption. *Ann. N. Y. Acad. Sci.* 1324: 7-14.
- 386 Neeraja, C.N., Maghirang-Rodriguez, R., Pamplona, A., Heuer, S., Collard, B.C.Y.,
387 Septiningsih, E.M., et al. 2007. A marker-assisted backcross approach for developing
388 submergence-tolerant rice cultivars. *Theor Appl Genet.* 115: 767-76.
- 389 Nematzadeh, G.A., Huang, N., Khush, G.S. 2004. Mapping the gene for aroma in rice (*Oryza*
390 *sativa* L.) by bulk segregant analysis via RAPD markers. *J. Agric. Sci. Technol.* 6: 129-137.
- 391 Ragimekula, N., Varadarajula, N.N., Mallapuram, S.P., Gangimeni, G., Reddy, R.K.,
392 Kondreddy, H.R. 2013. Marker assisted selection in disease resistance breeding. *J Plant Breed*
393 *Genet.* 1: 90-109.
- 394 Sabu, K.K., Abdullah, M.Z., Lim, L.S., Wickneswari, R. 2006. Development and evaluation of
395 advanced backcross families of rice for agronomically important traits. *Com. Biom. Crop Sci.* 1:
396 111-123.
- 397 Rajpurohit, D., Kumar, R., Kumar, M., Paul, P., Awasthi, A., Basha, P., et al. 2011. Pyramiding
398 of two bacterial blight resistance and a semidwarfing gene in Type 3 Basmati using marker-
399 assisted selection. *Euphytica.* 178: 111-126.
- 400 Sakthivel, K., Sundaram, R., Shobha, R.N., Balachandran, S., Neeraja, C. 2009. Genetic and
401 molecular basis of fragrance in rice. *Biotechnol. Adv.* 27: 468-473.
- 402 Servin, B., Hospital, F. 2002. Optimal positioning of markers to control genetic background in
403 marker-assisted backcrossing. *J Hered.* 93: 214-217.

404 Shamsudin, N.A.A., Swamy, B.P.M., Ratnam, W., Cruz, M.T.S., Raman, A., Kumar, A. 2016.
405 Marker assisted pyramiding of drought yield QTLs into a popular Malaysian rice cultivar,
406 MR219. BMC Genet. 17:30.

407 Sood, B.C. 1978. A rapid technique for aroma determination in rice. Indian J. Genet. Plant
408 Breed. 38: 268-271.

409 Tanweer, F.A., Rafii, M.Y., Sijam, K., Rahim, H.A., Ahmed, F., Ashkani, S., et al. 2015.
410 Introgression of blast resistance genes (putative *Pi-b* and *Pi-kh*) into elite rice cultivar MR219
411 through marker-assisted selection. Front. Plant Sci. 6:1002.

412 Teo, C.J., Thuc, L.V., Yeoh, K.A., Chan, W.S., Chivukula, S.V., Namasivayam, P., et al. 2011.
413 Profiling the differentially expressed genes in two rice varieties during rapid grain-filling stages.
414 Acta Physiol. Plant. 33(6) :2259-2268.

415 Yeap, H.Y., Faruq, G., Zakaria, H.P., Harikrishna, J.A. 2013. The efficacy of molecular markers
416 analysis with integration of sensory methods in detection of aroma in rice. Sci. W. J.
417 doi.10.1155/2013/569268.

418

419

420

421

422

List of captions

Fig 1. Crossing scheme for the development of fragrant BC₂F₂ lines

Fig 2. Examples of amplified *fgr*-SNP products separated using 1% agarose gel electrophoresis at 100 V for 1 h. L: 100 bp ladder; Samples 1, 6, 7, 11, and 12: BC₂F₂ individuals with homozygous *fgr* alleles similar to MRQ74. Samples 2-4, 8, 9, 13-17, 19, 20, and 23-28: BC₂F₂ individuals with heterozygous *fgr* alleles. Samples 5, 10, 18, 21, and 22: homozygous *fgr* alleles similar to MR219. P1: MR219; and P2: MRQ74.

Fig 3. Distributions of the sensory score for (A) LAT and (B) GAT in BC₂F₂ lines. The BC₂F₂ lines derived from Cross-1 and Cross-2 are represented in yellow and green, respectively.

Tables

Table 1. Genetic background analysis and phenotypic screening of selected fragrant BC₂F₂ lines

Cross	Selected individual	Genomic proportion (%)			Sensory score	
		*R	[†] D	[‡] H	LAT	GAT
Cross-1	RU13002-1-6	91.7	4.2	4.1	3	3
	RU13002-2-8	93.8	2.1	4.1	3	3
	RU13002-2-12	89.6	4.2	6.6	3	3
	RU13002-3-3	93.8	2.1	4.1	3	3
	RU13002-5-10	91.7	4.2	4.1	3	3
	RU13002-5-12	93.8	4.2	2.0	3	3
	RU13002-7-5	89.6	4.2	6.2	3	3
	RU13002-8-1	91.7	2.1	6.2	3	3
	RU13002-9-3	93.8	4.2	2.0	3	3
	RU13002-9-9	89.6	6.3	4.1	3	3
Cross-2	RU13005-2-3	87.5	4.2	8.3	3	3
	RU13005-2-11	89.6	6.3	4.1	3	3
	RU13005-3-3	87.5	6.3	6.2	3	3
	RU13005-4-7	91.7	4.2	4.1	3	3
	RU13005-4-10	91.7	2.1	6.2	3	3
	RU13005-5-9	89.6	4.2	6.6	3	3
	RU13005-7-12	87.5	8.3	4.2	3	3
	RU13005-7-13	93.8	4.2	2.0	3	3
	RU13005-8-7	89.6	4.2	6.6	3	3
	RU13005-9-5	91.7	4.2	4.1	3	3

*R: Recurrent parent; [†]D: Donor parent; [‡]H: Heterozygous

Table 2. Performance of major agronomic traits of selected BC₂F₂ lines carrying *fgr* gene

Traits	MRQ74	MR84	Fragrant BLs	MRQ74	MR219	Fragrant BLs
PLHT (cm)	68.1a	99.9b	95.5c	68.1a	89.7b	85.9c
DFFL (day)	98a	93b	95c	98a	87b	88b
DMT (day)	123a	118b	119b	123a	115b	117b
GRGL (mm)	10.05a	9.84b	9.96a	10.05a	9.72b	9.85c
GRWH (mm)	1.96a	1.97b	1.96a	1.96a	1.98b	2.00c
GRSP	5.13a	4.99b	5.08a	5.13a	4.90b	4.90b
GRWT (g)	24.3a	24.9b	25.6b	24.3a	24.3a	27.7b
YPL (g)	31.5a	37.3b	44.9b	31.5a	40.4b	38.6b
BY	92.7a	98.2a	92.8a	92.7a	91.8a	83.1a
HI	0.265a	0.382b	0.483c	0.319ab	0.267a	0.365b

*Significance at 5% level with independent t-test

461

Draft

Figure 1.

Figure 2.

Draft

Figure 3.

Draft

Supplementary Table 1. Primer sequences used for foreground and background selection

No	Primers name	Sequences (5'-3')
1	EAP	AGTGCTTTACAAAGTCCCGC
2	ESP	TTGTTTGGAGCTTGCTGATG
3	IFAP	CATAGGAGCAGCTGAAATATATACC
4	INSP	CTGGTAAAAAGATTATGGCTTCA
5	RM237-F	CAAATCCCGACTGCTGTCC
6	RM237-R	TGGGAAGAGAGCACTACAGC
7	RM212-F	CCACTTTCAGCTACTACCAG
8	RM212-R	CACCCATTTGTCTCTCATTATG
9	RM431-F	TCCTGCGAACTGAAGAGTTG
10	RM431-R	AGAGCAAAACCCTGGTTCAC
11	RM104-F	GGAAGAGGAGAGAAAGATGTGTGTCG
12	RM104-R	TCAACAGACACACCGCCACCGC
13	RM279-F	GCGGGAGAGGGATCTCCT
14	RM279-R	GGCTAGGAGTTAACCTCGCG
15	RM555-F	TTGGATCAGCCAAAGGAGAC
16	RM555-R	CAGCATTGTGGCATGGATAC
17	RM53-F	ACGTCTCGACGCATCAATGG
18	RM53-R	CACAAGAACTTCCTCGGTAC
19	RM174-F	AGCGACGCCAAGACAAGTCGGG
20	RM174-R	TCCACGTCGATCGACACGACGG
21	RM6-F	GTCCCCTCCACCCAATTC
22	RM6-R	TCGTCTACTGTTGGCTGCAC
23	RM240-F	CCTTAATGGGTAGTGTGCAC
24	RM240-R	TGTAACCATTCCTTCCATCC
25	RM208-F	TCTGCAAGCCTTGTCTGATG
26	RM208-R	TAAGTCGATCATTGTGTGGACC
27	RM207-F	CCATTTCGTGAGAAGATCTGA
28	RM207-R	CACCTCATCTCGTAACGCC
29	RM36-F	CAACTATGCACCATTTGTCGC
30	RM36-6	GTACTCCACAAGACCGTACC
31	RM7-F	TTCGCCATGAAGTCTCTCG
32	RM7-R	CCTCCCATCATTTGTTGTT
33	RM251-F	GAATGGCAATGGCGCTAG
34	RM251-R	ATGCGGTTCAAGATTCGATC
35	RM273-F	GAAGCCGTCGTGAAGTTACC
36	RM273-R	GTTTCCTACCTGATCGCGAC
37	RM252-F	TTCGCTGACGTGATAGGTTG
38	RM252-R	ATGACTTGATCCCGAGAACG
39	RM241-F	GAGCCAAATAAGATCGCTGA
40	RM241-R	TGCAAGCAGCAGATTTAGTG
41	RM348-F	CCGCTACTAATAGCAGAGAG
42	RM348-R	GGAGCTTTGTTCTTGCGAAC
43	RM163-F	ATCCATGTGCGCCTTTATGAGGA
44	RM163-R	CGCTACCTCCTTCACTTACTAGT
45	RM164-F	TCTTGCCCGTCACTGCAGATATCC
46	RM164-R	GCAGCCCTAATGCTACAATTCTTC
47	RM440-F	CATGCAACAACGTACACCTTC
48	RM440-R	ATGGTTGGTAGGCACCAAAG
49	RM421-F	AGCTCAGGTGAAACATCCAC
50	RM421-R	ATCCAGAATCCATTGACCCC
51	RM3-F	ACACTGTAGCGGCCACTG
52	RM3-R	CCTCCACTGCTCCACATCTT
53	RM3628-F	AATCATGCCTAGAGCATCGG
54	RM3628-R	GTTCAACATGGGTGCAGATG

55	RM30-F	GGTTAGGCATCGTCACGG
56	RM30-R	TCACCTCACCACACGACACG
57	RM340-F	GGTAAATGGACAATCCTATGGC
58	RM340-R	GACAAATATAAGGGCAGTGTGC
59	RM18-F	TTCCCTCTCATGAGCTCCAT
60	RM18-R	GAGTGCCTGGCGCTGTAC
61	RM234-F	ACAGTATCCAAGGCCCTGG
62	RM234-R	CACGTGAGACAAAGACGGAG
63	RM70-F	GTGGACTTCATTTCAACTCG
64	RM70-R	GATGTATAAGATAGTCCC
65	RM560-F	GCAGGAGGAACAGAATCAGC
66	RM560-R	AGCCCGTGATACGGTGATAG
67	RM152-F	GAAACCACCACACCTCACCG
68	RM152-R	CCGTAGACCTTCTTGAAGTAG
69	RM38-F	ACGAGCTCTCGATCAGCCTA
70	RM38-R	TCGGTCTCCATGTCCAC
71	RM515-F	TAGGACGACCAAAGGGTGAG
72	RM515-R	TGGCCTGCTCTCTCTCTC
73	RM210-F	TCACATTCGGTGGCATTG
74	RM210-R	CGAGGATGGTGTGTTCACTTG
75	RM242-F	GGCCAACGTGTGTATGTCTC
76	RM242-R	TATATGCCAAGACGGATGGG
77	RM201-F	CTCGTTTATTACCTACAGTACC
78	RM201-R	CTACCTCCTTTCTAGACCGATA
79	RM215-F	CAAAATGGAGCAGCAAGAGC
80	RM215-R	TGAGCACCTCCTTCTCTGTAG
81	RM258-F	TGCTGTATGTAGCTCGCACC
82	RM258-R	TGGCCTTTAAAGCTGTCGC
83	RM228-F	CTGCCATTAGTCCTTGG
84	RM228-R	GCTTGCGGCTCTGCTTAC
85	RM333-F	GTACGACTACGAGTGTACCAA
86	RM333-R	GTCTTCGCGATCACTCGC
87	RM202-F	CAGATTGGAGATGAAGTCCTCC
88	RM202-R	CCAGCAAGCATGTCAATGTA
89	RM287-F	TTCCCTGTAAAGAGAGAAATC
90	RM287-R	GTGTATTTGGTGAAAGCAAC
91	RM206-F	ACTCCACTATGACCCAGAG
92	RM206-R	GAACAATCCCTTCTACGATCG
93	RM144-F	TGCCCTGGCGCAAATTTGATCC
94	RM144-R	GCTAGAGGAGATCAGATGGTAGTGCATG
95	RM20-F	ATCTTGTCCCTGCAGGTCAT
96	RM20-R	GAAACAGAGGCACATTTTCATTG
97	RM4-F	TTGACGAGGTCAGCACTGAC
98	RM4-R	AGGGTGTATCCGACTCATCG
99	RM19-F	CAAAAACAGAGCAGATGAC
100	RM19-R	CTCAAGATGGACGCCAAGA