Genotyping USDA Rice Mini-Core Collection With Functional Markers for Important Agronomic Traits

Kehu Li (li_kehu@yahoo.com)  
Guizhou University  https://orcid.org/0000-0002-4851-919X

Yongyi Ge
Guizhou University

Lily Yan Wang
The University of Hong Kong

Research Article

Keywords: Rice, USDA mini-core, Functional markers

Posted Date: August 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-788277/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

The USDA rice mini-core collection was established to capture the diversity of an entire collection of over 18,700 accessions of global origins for efficient germplasm evaluation and exploration. Previous studies have investigated its genetic diversity and population structure using genome-wide SSR markers. Many important agronomic traits that are fundamental to rice breeding programs, however, remain to be explored. Functional markers can be developed based on polymorphic sites within genes affecting phenotypic variation in, e.g., starch physicochemical properties, nutritional qualities and biotic resistance. These markers can be used for genotyping and hence differentiating phenotypes among rice accessions. In this study, we employed 12 pairs of functional markers (SNP and Indel) to genotype all 217 accessions constituting the USDA rice mini-core. These markers are highly associated with starch physicochemical properties (intron 1 G/C SNP, 23bp duplication in exon 2, exon 6 C/A SNP, exon 10 C/T SNP of Waxy gene, GC/TT SNPs of SSIIa gene, G/C SNP of SBE3 gene), glutelin content (3.5 kb deletion in Lgc1 gene), grain length (C/A SNP in GS3 gene), brown planthopper resistance (InDel in Bph 14 gene) and rice blast resistance (InDel in Pi54 and Pit gene). Using these functional markers, all the 217 accessions of the mini-core are characterized for aforementioned agronomic traits associated alleles/genes. The results of this study will help breeders select parental materials with desirable allele/gene combinations and phenotypes among mini-core accessions for rice breeding programs.

1. Introduction

Rice is the major staple food for nearly half of the world's population. Diverse rice germplasm providing raw materials for parents selection is required by various rice breeding programs aiming at various breeding goals, e.g. yield lifting, quality improvement, biotic and abiotic resistance enhancement. The USDA rice mini-core collection which is composed of 217 accessions was selected from the whole USDA rice gene bank. As a set of germplasm, potentially valuable for breeding, its diversity had been evaluated based on phenotypic and genotypic markers in previous studies (Agrama et al. 2009; Li et al. 2010), the results confirmed that this mini-core collection is a good representative of over 18,700 genotypes from 115 countries contained in USDA rice gene bank. Some successful studies on its application in gene-trait association analysis are available (Li et al. 2011; Bryant et al. 2013). However, in the respect of breeding practice, many fundamental yet important information on agronomic traits remained to be explored.

The knowledge on phenotypic and genotypic traits is a prerequisite for parent and offspring selection in breeding. DNA markers that can identify desired genes/alleles in rice germplasm are valuable because they can reveal the existence of the genes without laborious phenotype investigation. However, many markers are distant from targeted genes, occasional uncoupling of the marker from trait might happen in cycles of meiosis, this could lead to errors in selection of desired traits (Perumalsamy et al. 2010). Functional molecular markers (FMs) are derived from functional polymorphisms within or around genes that causally affect phenotypic variations (Anderson and Lubberstedt 2003). FMs have absolute advantages over other DNA markers since they are fully diagnostic of the target trait allele. As information on cloned genes becomes available, FMs development are enabled and FMs genotyping
should serve as a reliable alternative to cumbersome phenotype investigation since they always show consistency between functional traits and sequence polymorphisms in alleles.

1.1 Functional markers for rice starch physicochemical properties

As the major component in the rice grain, starch accounts for over 90% of endosperm weight, its physicochemical properties strongly influence rice eating and cooking qualities. Two types of starch polymers exist in rice endosperm: amylose and amylopectin. Amylose is linear with few branches while amylopectin is highly branched (Tester et al. 2004, Hizukuri et al. 2006). It has been established that amylose content and the fine structure of amylopectin are the main determinants of rice eating and cooking qualities (Juliano 1985, Fujita et al. 2003, Zhang et al. 2011, Syahariza et al. 2013).

Apparent amylose content (AAC) is one of key parameters for evaluation of rice eating and cooking qualities. Rice cultivars can be grouped into five classes according to AAC: glutinous (AAC 1–2%), very low (2–9%), low (10–20%), intermediate (20–25%) and high (> 25%) (Kumar and Khush, 1987). Genetic studies have revealed that Waxy (Wx) gene, which encodes granule bound starch synthase (GBSS) was the primary gene responsible for amylose biosynthesis in rice endosperm. GBSS activity is affected by the naturally occurred or human generated sequence variations in Wx gene, on transcriptional or post-transcriptional levels, then AAC phenotypic variations occurred subsequently. Three functional alleles at the Waxy locus, namely Wxa, Wxb and wx, were first revealed by Sano in 1984 (Sano 1984). Glutinous rice has the recessive wx allele which harbors a 23 bp duplication in the exon 2 of Wx gene. Due to this insertion, gene expression was disrupted, a prematurely terminated translation occurred, then normal GBSS synthesis failed (Wanchana et al. 2003). Rice cultivars with Wxb allele have low AAC while those with Wxa allele have intermediate or high AAC (Sano 1984; Sano et al. 1985). The difference between Wxa and Wxb alleles is the G/T substitution in putative 5’ splice site in intron 1 of Wx gene. This G to T mutation in Wxb could reduce the efficiency of GBSS pre-mRNA processing, thus resulted in a lower level of spliced mature mRNA and a subsequent lower AAC (Bligh et al. 1998; Cai et al. 1998; Hirato et al. 1998; Isshiki et al. 1998). Another two additional functional Wx alleles, namely Wxop and Wxin were found by Mikami et al. (2008). Wxop and Wxin are associated with very low (< 10%) and intermediate (20–25%) AAC respectively. The A to G substitution occured in the exon 4 and the A to C substitution in the exon 6 in Wx gene were the characters of Wxop and Wxin alleles respectively. In a study involving 70 rice cultivars (Tian et al. 2009), Wx gene were classified into Wx-I, Wx-II and Wx-III haplotypes and they were associated with glutinous, intermediate and high AAC phenotypes respectively. Finally, in a most recent and advanced study (Teng et al. 2012), a total of five functional Wx alleles have been found in single-segment substitution lines involving 17 parental rice cultivars. Among the five functional Wx alleles reported in this paper, wx, Wxt and Wxg1 were identical to the previously reported wx, Wxb and Wxin alleles (Sano 1984; Mikami et al. 2008). Compared with Wxg1, Wxg3 had a C to A substitution in the exon 6, Wxg2 had an extra
C to T substitution in the exon 10. Importantly, these five Wx alleles, namely Wx$^{a3}$, Wx$^{a2}$, Wx$^{a1}$, Wx$^t$ and wx were perfectly associated with five AAC classes: High II, High I, Intermediate, Low, and Glutinous.

Amylopectin is the other type of starch in rice grain, its content and fine structure are also crucial to rice eating, cooking and processing qualities. Genetically, ADP-glucose pyrophosphorylase (AGPase), soluble starch synthase (SS), starch branching enzyme (SBE) and starch de-branching enzyme (DBE) work together to synthesize amylopectin (Nakamura 2002, Nakamura et al. 2010). Each group of enzyme has multiple subunits or isoforms with specific roles. Among them, SSIIa has been reported as the major gene responsible for gelatinization temperature (Umemoto et al. 2004). Two consecutive GC/TT SNPs could differentiate rice cultivars with low GT from those having intermediate or high GT. PCR-based markers targeting this polymorphism were designed (Bao et al. 2006). Additionally, the C/G SNP downstream of the SBE3 gene was significantly associated with amylose content and viscosity properties (Lu et al. 2012).

1.2 Functional markers for blast resistance, bacteria blight resistance and brown planthopper resistance

Rice blast is one of the most devastating diseases affecting the rice crop worldwide. To date, more than 50 blast resistance genes have been characterized in rice. Among them, one of the major blast resistance genes Pi54 (once named as Pik$^r$) exhibited universal resistance to predominant races of the pathogen (Sharma et al. 2002). In an allele mining of Pi54 gene, Ramkumar et al. (2011) revealed a 144 bp insertion/deletion (Indel) in the exonic region of the gene. Pi54 MAS, a PCR-based co-dominant molecular marker targeting this Indel was designed. In a following validation of Pi54 MAS, 105 rice genotypes were investigated, the conclusion was that Pi54 MAS could more accurately distinguish nearly all the resistant/susceptible cultivars than any early reported marker for Pi54 gene, hence this functional marker is suitable for application in MAS.

Pit is another rice blast resistance gene conferring rice cultivars broad-spectrum resistance (Kiyosawa 1972). The resistance function was mainly conferred by up-regulated promoter activity through the insertion of a long terminal repeat (LTR) retrotransposon upstream of Pit. tK 59 is a PCR-based DNA marker derived from LTR, with this marker, five resistant cultivars were successfully identified from 68 diverse rice genotypes (Hayashi et al. 2010).

Bacteria leaf blight is also a serious disease throughout the rice growing area. To date, over 20 resistance genes have been identified (Wang et al. 2007, Ram et al. 2010). A recessive gene xa5 provides resistance to several races of pathogens (Garris et al. 2003). A Simple PCR-based functional marker targeting two consecutive nucleotide polymorphisms in xa5 gene that can distinguish resistance from susceptibility had been developed and verified using rice cultivars phenotyped for bacterial blight resistance (Iyer-Pascuzzi et al. 2007).

The Brown planthopper (Nilaparvata lugens) is a severe constraint to rice production (Way and Heong 1994, Normile 2008). To date, over 20 resistance genes have been identified in rice cultivars and wild rice
species (Jairin et al. 2010, Jena and Kim 2010, Rahman et al. 2009, Yara et al. 2010). Among them, Bph14 transferred from *Oryza officinalis*, is one of genes conferring rice brown planthopper resistance (Huang et al. 2001). With recent cloning of *Bph14* gene (Du et al. 2009), Zhou et al. (2013) developed PCR-based codominant functional markers amenable for MAS.

1.3 Other functional markers

High yield had been not the only goal in recent rice breeding programs. More and more interesting was attracted by rice eating and cooking quality improvement and functional food development. With knowledge generated in agronomically important genes’ isolation, more information are available for developing FMs directly used in germplasm evaluation, gene/allele identification and offspring selection in breeding.

Low glutelin content rice is characterized by less mature gultelin and more prolamine than those of normal varieties. Since large amount of glutelin absorption can lead to proteometabolism disturbance and exacerbation, low glutelin rice is viewed as functional food for patients with kidney failure. Genes *Lgc 1* (*glu-1*), *glu-2* and *glu-3* were revealed responsible for low glutelin content (Miyahara et al. 1996, Iida et al. 1997, Qu et al. 2002). Chen et al. (2010) designed a PCR-based InDel marker targeting a 3.5-Kb deletion in *Lgc 1*, with which rice cultivars with low glutelin content could be identified by genotyping.

*GS3* is the most important grain length QTL, a single nucleotide polymorphism C/A in its exon 2 principally differentiates rice grain length (Fan et al. 2006). SF28 is a cleaved amplified polymorphic sequence marker targeting C/A SNP in the *GS3* gene, in an association analysis, 38 out of 180 rice varieties with *GS3*-A allele had grain length ranging from 8.8 to 10.7 mm, while the rest 142 ones with *GS3*-C allele had grain length ranging from 6.4 to 8.8 mm (Fan et al. 2009), this suggested this marker could be used in MAS with great reliability.

The objective of the present study was to genotype USDA rice mini-core collection with twelve functional markers associated with eight important agronomic traits, for the purpose of providing guidelines for selecting parent materials with desired gene/alleles for rice breeding. Besides, the polymorphism detected in *Waxy* gene in this study could be applied in a following association analysis with AAC variations, in order to achieve a better understanding of rice amylose biosynthesis and find marker combinations that could explain more AAC variations.

2. Materials And Methods

2.1 Plant materials

The USDA rice mini-core collection used in this study was provided by USDA-ARS.

2.2 DNA isolation
Rice seeds were germinated, whole genomic DNA was extracted from five seedlings of each accession using CTAB method (Doyle 1991).

2.3 Functional markers for \( Wx \)-intron 1, \( Wx \)-exon 2, \( Wx \)-exon 6, \( Wx \)-exon 10, \( SSIIa \) and \( SBE3 \)

Four polymorphisms reported in the \( Waxy \) gene, G/T SNP in the intron 1 (\( Wx \)-Intron 1), 23 bp duplication in the exon 2 (\( Wx \)-Exon 2), A/C SNP in the exon 6 (\( Wx \)-Exon 6) and C/T SNP in the exon 10 (\( Wx \)-Exon 10), were chosen for genotyping the USDA rice mini-core collection.

Two primers oligo 484 and oligo W2R, restriction endonuclease \( AccI \) (BRL) (Ayers et al. 1997), were used to genotype the intron 1 G/T SNP. Two fragments (129 and 128 bp) designated a G SNP, whereas no digestion suggested a T SNP.

For 23 bp InDel in the exon 2, primers Glu-23F and Glu-23R were used for detection (Wanchana et al. 2003). A 196 bp fragment indicated the duplication while 173 bp fragment meant no duplication.

Three primers oligo Ex6-F-A M13, Ex6-F-C and Ex6-R were included in one PCR tube to genotype exon 6 A/C SNP (Caffagni et al. 2013), A 217 bp fragment indicated A SNP while 198bp designated C SNP.

For the exon 10 C/T SNP, two primers designed in this study10F-10R and restriction endonuclease \( ApaI \) were used in genotyping, no digestion meant T SNP and two fragments (146 and 196 bp) suggested C SNP.

For GC/TT consecutive SNPs in \( SSIIa \) gene, two-pairs confronting primers, oligo F7, F22, R1 and R21 were applied in one PCR tube for genotyping (Bao et al. 2006), a 832 bp fragment correspond to TT SNP while 543 bp fragment to GC SNP. Oligo 492 and oligo 493, along with restriction endonuclease \( SpeI \) were used to determine the C/G SNP in \( SBE3 \) (Han et al. 2004), Two fragments (215 and 295 bp) indicated a G SNP while no digestion suggested a C SNP.

2.4 Functional markers for \( Pi54 \), \( Pit \) and \( Bph14 \)

Primers \( Pi54 \) MAS were used in determination of the 144 bp InDel polymorphism in the exonic region of \( Pi54 \) gene (Ramkumar et al. 2011), a 216 bp product was for resistant allele and 359 bp fragment for susceptible allele.

Three primers, oligo tK59, tRn1 and tRn2 were applied in one PCR tube to detect resistant \( Pit \) allele, which produced two fragments (733 and 530 bp) (Hayashi et al. 2010). Primers \( xa5F \), \( xa5R \) and restriction endonuclease \( BsrI \) were used to identify TC/AG SNPs in \( xa5 \) gene, two fragments (551 and 400 bp) suggested TC SNP (susceptible allele) and no digestion designated AG SNP (Iyer-Pascuzzi et al. 2007).

For \( Bph14 \), a codominant marker system Bph14P/N, comprising four primers were used to search for resistant allele, a 566 bp fragment produced in resistant allele (Zhou et al. 2013).

2.5 Functional marker for \( Lgc1 \) and \( GS3 \)
Primers InDel-Lgc1-2 were used to identify the 3.5-kb deletion in low glutelin content gene 1 (Lgc1), those accession without this deletion, produced a 509 bp fragment (Chen et al. 2010).

Primers SF28 and restriction endonuclease PstI were used to determine the C/A SNP in GS3 gene (Fan et al. 2009), two fragments (110 and 26 bp) designated C SNP, whereas no digestion indicated A SNP.

2.6 DNA marker analysis

The PCR reaction was performed in a 10 µl reaction mixture containing 20 ng of template DNA, 1X PCR buffer, 2mM MgCl₂, 0.2 mM dNTPs, 0.2 µM of each primer and 1 unit of Taq DNA polymerase. All amplifications were performed on a PTC-100 thermal cycler (MJ Research, Inc.) under the following conditions: 5 min at 94°C, followed by 45 s at 94°C, 60 s at Tₐ (Annealing temperature, Table 1), 60 s at 72°C for 35 cycles, and 7 min at 72°C for a final extension.

The PCR products of SSIIa, Pit, Pi54, Bph14 and Lgc1 gene were resolved on a 2.0% agarose gel containing 0.05 µl/mL gel red in 1X TBE buffer, then visualized using a gel documentation system. PCR products of Wx-exon 2, Wx-exon 6 were separated on 8% polyacrylamide denaturing gels and visualized by silver staining (Bassam et al. 1991).

5µl of PCR products of Wx-intron 1, Wx-exon 10, SBE3, xa5, and GS3 were digested by restriction endonucleases (Biolabs) (Table 1), according to manufactures. The products of restriction digestion of Wx-exon 10, SBE3, xa5 were separated by gel electrophoresis (2.0% agarose gel) and visualized by gel documentation system. For Wx-intron 1 and GS3, the digested products were separated on 8% polyacrylamide gel and detected by silver staining.

3. Results

All 217 accessions in mini-core collection were genotyped using twelve functional markers for eight agronomic traits (Table 1, Supplementary Table 1).

3.1 Allelic frequencies of Wx gene, SSIIa gene and SBE3 gene

Four Wx gene markers have been chosen in genotyping the collection: G/T SNP in the 5’ splice site of intron 1 (Wx-Intron 1), 23 bp duplication in exon 2 (Wx-Exon 2), A/C SNP in exon 6 (Wx-Exon 6) and C/T SNP in exon 10 (Wx-Exon 10).

Concerning Wx-Intron 1 G/T SNP, 154 rice accessions carried G allele (71.0%) while only 63 (29.0%) for T allele (Table 2). T allele was reported associated with low AAC while G allele with intermediate and high AAC (Cai et al. 1998). For exon 6 A/C SNP, the genotyping results showed A allele (163 accessions, 75.1%) was predominant in mini-core (Table 2), this polymorphism work as a complement to Wx-Intron 1, could differentiate rice varieties with intermediate and high AAC (Chen et al. 2008). Only 15 accessions (6.9%) carried the 23 bp duplication in Wx gene exon 2 (Table 2), this duplication had been only found in glutinous rices (Wanchana et al. 2003). As far as the Wx-exon 10 C/T SNP concerned, 178 accessions...
(82.0%) carried C allele while only 39 accessions (18.0%) for T allele (Table 2), as complement to Wx-Intron 1 G/T polymorphism and Wx-exon 6 A/C polymorphism, Wx-exon 10 could further divide rice cultivars with high AAC to high I and high II subclasses (Teng et al. 2012).

Generally, in mini-core, for each of the four Wx gene polymorphic sites, always one allele was much more frequent than its counterpart (Table 2). Based on all four Wx marker loci, a total of 8 haplotypes could be found in USDA rice mini-core collection (Table 3). A total of 200 accessions in the USDA rice mini-core collection could be grouped into the previously reported five Wx haplotypes (Table 3), proposed by Teng et al. (2012). Wx\textsuperscript{g3} (G-1-A-C) haplotype, associated with the high II AAC class and characterized with a G SNP in intron 1, no duplication in exon 2, A SNP in exon 6 and C SNP in exon 10, was the most frequent haplotype in mini-core, harbored in 75 accessions (34.5 %, Table 2). The second largest group is Wx\textsuperscript{g1} (G-1-C-C) carried by 43 accessions (19.8%), which is associated with high I AAC. Wx\textsuperscript{g2} (G-1-A-T) and wx\textsuperscript{t} (T-1-A-C), associated with intermediate and low AAC, have 29 (13.4%) and 38 (17.5%) accessions respectively (Fig 1). What is noteworthy, seventeen accessions could not be assigned into previously reported Wx haplotypes. Among them, eight accessions, namely (plant identification, Supplementary Table 1) “Onu B”, “WC 3532”, “Grassy”, “Buphopa”, “Bombon”, “Egyptian Wild Type”, “Jyanak”, “IR 64” belonged to “T-1-C-C” haplotype, another eight accessions, namely “Hsin Hsing Pai Ku”, “10340”, “CM1, HAIPONG”, “P35”, “Bogarigbeli”, “Magoti”, “M202”, “Nipponbare” fell into “T-1-A-T” group and one accession “Lua Chua Chan” was “T-1-C-T” haplotype.

Two pairs of confronting markers targeting GC/TT polymorphism in SSIIa gene were chosen to genotype the mini-core. The GC allele was predominant, harbored by 183 accessions (83.9%). Contrastingly, only 35 accessions (16.1%) had TT allele (Table 2). TT allele was reported associated with low gelatinization temperature (GT) while GC allele with intermediate and high GT. Pretty same pattern observed in SBE3 C/G SNP, as 145 accessions (66.8%) carried G allele while 72 accessions (33.2%) had C allele (Table 2).

In general, for all the six functional markers employed in present study targeting starch physicochemical properties, genotyping results showed in USDA mini-core collection, for each one of them, one allele was always more frequent than its counterpart. Relatively, the most equally distributed alleles occurred in SBE3 C/G polymorphism, but still 66.8% SBE3-G allele to 33.2% SBE3-C allele. The most unequally distribution observed in Wx-Exon 2, as 23 bp duplication was only observed in 6.9% cases, by the way, this duplication was reported responsible for glutinous AAC.

### 3.2 Allelic frequencies of Pi54 gene, Pit gene and Bph14 gene

Pi54 and Pit were both of genes conferring rice blast resistance. The Pi54 resistant allele has a 144 bp deletion in exon2 region of Pi54 and the Pit resistant allele has an insertion of a long terminal repeat retrotransposon upstream of Pit. In mini-core, 99 (45.6%) accessions carried Pi54 resistant allele, but Pit resistant allele was only found in 4 (1.8%) accessions (Table 2). Noticeably, two accessions “SHIMIZU MOCHI” and “Padi Pohon Batu” were found heterogenous for the Pi54 allele (Fig 2a). Furthermore, three
accessions “Warrangal Culture 1252”, “Traneup Beykher” and “4484” had both resistant alleles of \textit{Pi54} and \textit{Pit} (Supplementary Table 1).

Resistant allele of \textit{Bph14} gene was not found in USDA rice mini-core collection, nevertheless resistant allele was neither found in Chinese rice mini-core collection, which comprises 200 varieties, selected from more than 60000 accessions of Chinese cultivated rice (Zhou et al. 2013).

3.3 Allelic frequencies of \textit{Lgc 1} gene and \textit{GS3} gene

For \textit{GS3} C/A SNP, C allele is more frequently found in mini-core, in contrast to 151 accessions (69.6%) carrying C allele, only 66 accessions (30.4%) had A allele (Table 2), C allele was also found more frequent in another study involving 180 rice genotypes (Fan et al. 2009), 142 (79%) of them have the C SNP.

In USDA rice mini-core collection, there were 58 accessions (26.7%, Table 2) having the 3.5 kb deletion in \textit{Lgc1} gene which caused low glutelin content phenotype.

4. Discussion

Germplasm resource is essentially important for plant breeding in respect of parents selection. Phenotype and genotype investigation is a prerequisite for parents selection in any breeding program. Although the USDA rice mini-core collection is a valuable set of germplasm, little is known on its phenotypic and genotypic characterization of many important agronomic traits, such like disease resistance, eating and cooking qualities. On the other hand, the investigation of phenotype is usually laborious. For example, to identify parental materials with pathogen resistance, usually inoculation test and long-term observation is needed.

Molecular marker assisted selection (MAS) is helpful in parents and offspring selection to save cost and labor in breeding. However, many markers are distant from responsible genes, when undesired recombination between marker and candidate genes happened, the marker will lose its accuracy and reliability in selection of interested traits (Rafalski and Tingey, 1993). Functional markers are derived from within or around genes causally affect phenotypic variations thus have been directly employed with great efficiency and reliability to identify desirable alleles in many breeding programs (Ellis et al. 2002, Peter et al. 2004, Louis et al. 2005, Iyer-Pascuzzi and McCouch 2007, Caffagni et al. 2013).

Since starch is the major component in rice grain, accounting for nearly 90% of endosperm weight, the nucleotide sequence variations in starch biosynthesis related genes and their association with starch physicochemical properties has long been a research focus in rice breeding. \textit{Waxy} gene is primarily responsible for amylose biosynthesis. Sequence variations in five polymorphic sites of \textit{Wx} gene can explain most of amylose content variations. Firstly, due to a 23 bp duplication in exon 2 (wx allele), gene translation is prematurely terminated, so amylose can not be synthesized because of the absence of normal GBSS in glutinous rice (Wanchana et al. 2003). In USDA rice mini-core collection, 15 accessions harbored this 23bp duplication, so these rices are supposed to be glutinous/waxy rice with AAC less than
2%. For those rice without this duplication, if they have a T SNP in *Wx* gene intron 1 pupative 5” splice site, because of the lowered efficiency of the first intron splicing out, the amount of mature mRNA was reduced and of the GBSS enzyme. In short, rice with *Wx*-intron 1-T allele tend to show low amylose content (Sano 1984, Sano et al. 1985). In mini-core, *Wx*-intron 1-G allele was much more frequent (71.0%), these rice should fall to category of low AAC class. As complement to Intron 1-G allele, the A/C SNP in exon 6 has the ability to distinguish intermediate AAC from low and high AAC. In detail, C SNP is associated with intermediate AAC class while A SNP is associated with both low and high AAC class (Larkin and Park, 2003). Furthermore, the C/T SNP in exon 10 could distinguish rice with high I AAC from high II AAC: G-1-A-C allele associated with higher AAC (high II) than G-1-A-T (high I) (Teng et al. 2012), However, So far, no alleles had been reported distinguishing very low AAC from low AAC.

Based on these four polymorphic sites, all 217 accessions in USDA rice mini-core collection have been assigned to eight waxy gene haplotypes. G-1-A-C with 75 accessions (35.6%) was the most predominant *Wx* gene haplotype in mini-core, it was associated with high II AAC class (Teng et al. 2012). Apart from the previously reported five haplotypes, three new haplotypes have been found in 17 accessions, T-1-C-C, T-1-A-T and T-1-C-T. According to previous knowledge, T SNP in waxy gene intron 1 leaded to low AAC phenotype, hence these 17 accessions may demonstrate low AAC phenotype. Nevertheless, besides *wx* gene, other genes also play their roles in AAC determination. Among them, a C to G transversion in *SBE3* gene was reported leading to decreased AAC and increased RVA profile (Lu et al. 2012). This substitution occurred at the 63rd nucleotide downstream of *OsBEIIb* gene termination codon. Instead of directly causing amino acid substitution, it may affect gene translation activity. In mini-core, 145 accessions were revealed carrying *SBE3*-G allele.

The question is, what is the interaction between *Wx* alleles and the alleles of other starch biosynthesis related genes? For instance, with C/G allele in *SBE3* gene? On the other hand, conventionally, low AAC class is further divided to low and very low AAC subclasses (Kumar and Khush, 1987). Do the new *wx* alleles found in mini-core have the ability to distinguish rice genotypes with very low AAC (2–9 %) from that with low AAC (10–20 %)? In the following study, AAC will be measured as it is also a fundamental information about mini-core and on the other hand, the data will allow an association analysis which will shed light on the aforementioned questions raised in the present study.

Gelatinization temperature (GT) is one of important indicators for rice eating, cooking and processing quality. Genetically, GT is determined by *starch synthase Ila* (*SSIIa*) gene (Umemoto et al. 2004). Two consecutive SNPs can differentiate rice with low GT from those with intermediate or high GT with accuracy at 90%. PCR-based markers targeting GC/TT were developed (Bao et al. 2006). This important gene marker could be used to predict GT of cultivars and applied in marker-assisted selection to improve rice grain quality (Lu et al. 2010). In mini-core, GC allele is more frequent (182 accessions, 83.9%), with GC allele, rice show high or intermediate GT in 90% cases. GC allele was also frequently found in another survey involving 334 rice breeding lines and 172 rice landraces, totally, 346 rice genotypes (68.38%) carried GC allele and noticeably, only two had TT in 172 rice landraces.
Pi 54 and Pit are both of important genes being manipulated in rice breeding programs, the functional markers employed in present study were designed targeting functional polymorphisms of Pi54 and Pit gene, their reliabilities have been validated. 98 accessions (45.6%) have been found carrying functional allele of Pi54 gene. Comparably, in another set of diverse rice varieties, 38 out of 105 (36%) were found carrying functional alleles, suggesting a close result with our study in terms of allele frequency (Ramkumar et al. 2011). For Pit gene, only 4 accessions were revealed having resistant allele, this low frequency is not a surprise as in another study, functional Pit allele was only found in 5 accessions from 68 cultivars of the NIAS (National Institute of Agrobiological Sciences, Japan) global rice core collection (Hayashi et al. 2010). What is noteworthy, among these four accessions harboring functional Pit allele, three of them, namely “Warrangal Culture 1252”, “Tranoeup Beykher”, and “4484” also carried functional Pi54 allele. With functional markers, individuals with disease resistant alleles could be easily identified and through gene pyramiding, cultivar with combined alleles will show wider spectrum resistance.

Resistant allele of Bph14 gene was not found in USDA rice mini-core collection, furthermore, resistant allele was neither found in Chinese rice mini-core collection, which comprises 200 varieties, selected from more than 60000 accessions of Chinese cultivated rice (Zhou et al. 2013). This indicates Bph 14 may be not present in O. sativa.

Grain length plays its role in determining grain appearance, milling quality and affecting grain yield. It is therefore an important agronomic trait in rice breeding. For GS3 C/A SNP, C allele was more frequently found in mini-core, in contrast to 151 accessions (69.6%) carrying C allele, only 66 accessions (30.4%) had A allele. C allele was also found more frequent in another study involving 180 rice genotypes, 142 (79%) of them carrying GS3-C allele, with C allele, rice grain showed shorter grain length (Fan et al. 2009).

As revealed by Kusaba et al. (2003), a 3.5 kb deletion occurred in Lgc 1 gene resulted in remarkable suppression of glutelin accumulation in rice. For breeding low glutelin rice, conventionally, low glutelin content mutant generated by physical or chemical mutagenesis is an important genetic resource (Iida et al. 1993, 1997; Qu et al. 2002). In this present study, a total of 58 accessions (26.7%) having the 3.5 kb deletion in Lgc1 gene were identified in USDA rice mini-core collection. It is conceivable that new and suitable parent material could be found in these 58 accessions in respect of breeding low glutelin rice variety.

5. Conclusions

In conclusion, all 217 accessions of USDA rice mini-core collection had been genotyped using twelve functional markers for eight agronomic traits. We hope the results of this study could help breeders select parental materials with desirable allele/gene combinations and phenotypes among mini-core accessions for rice breeding programs, also the Wx gene polymorphisms identified in this study could be included in an association analysis with amylose content variations, in order to develop powerful marker combination that can differentiate rice cultivars belong to different AAC classes.
Declarations

Author Contributions: Conceptualization, Kehu Li; methodology, Yongyi Ge and Lily Yan Wang; software, Kehu Li; validation, Yongyi Ge and Lily Yan Wang; formal analysis, Kehu Li; investigation, Kehu Li; resources, Kehu Li; data curation, Kehu Li; writing—original draft preparation, Kehu Li; writing—review and editing, Yongyi Ge and Lily Yan Wang; visualization, Kehu Li; supervision, Kehu Li; project administration, Kehu Li; funding acquisition, Kehu Li. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Agrama HA, Yan WG, Lee F, Fjellstrom R, Chen MH, Jia M, McClung A (2009) Genetic assessment of a mini-core subset developed from the USDA rice genebank. *Crop Science, 49*:1336-1346.
2. Anderson JR, Lubberstedt T (2003) Functional markers in plants. *Trends in Plant Science, 8*:554-560.
3. Ayers NM, McClung AM, Larkin PD, Bligh HFJ, Jones CA, Park WD (1997) Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germ plasm. *Theoretical and Applied Genetics, 94*:773-781.
4. Bao JS, Corke H, Sun M (2006) Nucleotide diversity in starch synthase IIa and validation of single nucleotide polymorphisms in relation to starch gelatinization temperature and other physicochemical properties in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics, 113*:1171-1183.
5. Bassam BJ, Anolles GC, Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry, 196*:80-83.
6. Bligh HFJ, Larkin PD, Roach PS, Jones CA, Fu H, Park WD (1998) Use of alternate splice sites in granule-bound starch synthase mRNA from low-amylose rice varieties. *Plant Molecular Biology, 38*:407-415.
7. Bryant RJ, Jackson AK, Yeaster KM, Yan WG, McClung AM, Fjellstrom RG (2013) Genetic variation and association mapping of protein concentration in brown rice using a diverse rice germplasm collection. *Cereal Chemistry, 90 (5)*:445-452.
8. Caffagni A, Albertazzi G, Gavina G, Ravaglia S, Gianinetti A, Pecchioni N, Milc J (2013) Characterization of an Italian rice germplasm collection with genetic markers useful for breeding to improve eating and cooking quality. *Euphytica, 194*:383-399.
9. Cai XL, Wang ZY, Xing YY, Zhang JL, Hong MM (1998) Aberrant splicing of intron 1 leads to the heterogeneous 5’ UTR and decreased expression of waxy gene in rice cultivars of intermediate amylose content. *Plant Journal, 14 (4)*:459-465.
10. Chen MH, Bergman CJ, Pinson SRM, Fjellstrom RG (2008) Waxy gene haplotypes: Associations with pasting properties in an international rice germplasm collection. *Journal of Cereal Science, 48 (3)*:781-788.
11. Chen T, Tian MX, Zhang YD, Zhu Z, Zhao L, Zhao Q Y, Lin J, Zhou L H, Wang C L (2010) Development of simple functional markers for low glutelin content gene 1 (Lgc1) in rice (Oryza sativa). *Rice Science*, **17**(3):173-178.

12. Doyle JJ (1991) DNA protocols for plants-CTAB total DNA isolation. In: Hewitt GM (ed) Molecular techniques in taxonomy. Springer, Berlin Heidelberg New York, pp:283-293.

13. Du B, Zhang WL, Liu BF, Hu J, Wei Z, Shi Z Y, He RF, Zhu LL, Chen RZ, Han B, et al. (2009) Identification and characterization of Bph14, a gene conferring resistance to brown planthopper in rice. *Proceedings of the National Academy of Sciences of the United States of America*, 52:22163-22168.

14. Ellis MH, Spielmeyer W, Gale KR, Rebetzke GJ, Richards RA (2002) “Perfect” markers for the Rht-B1b and Rht-D1b dwarfing genes in wheat. *Theoretical and Applied Genetics*, 105:1038-1042.

15. Fan CC, Xing YZ, Mao HL, Lu TT, Han B, Xu CG, Li XH, Zhang QF (2006) GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theoretical and Applied Genetics*, 112:1164-1171.

16. Fan CC, Yu SB, Wang CG, Xing YZ (2009) A causal C-A mutation in the second exon of GS3 highly associated with rice grain length and validated as a functional marker. *Theoretical and Applied Genetics*, 118:465-472.

17. Fujita N, Kubo A, Suh DS, Wong KS, Jane JL, Ozawa K, Takaiwa F, Inaba Y, Nakamura Y (2003) Antisense inhibition of isoamylase alters structure of amylopectin and the physicochemical properties of starch in rice endosperm. *Plant and Cell Physiology*, 44 (6):607-618.

18. Garris AJ, McCouch SR, Kresovich S (2003) Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding xa5 locus of rice (Oryza sativa L.). *Genetics*, 165:759-769.

19. Han YP, Xu ML, Liu XY, Yan CJ, Korban SS, Chen XL, Gu MH (2004) Genes coding for starch branching enzymes are major contributors to starch viscosity characteristics in waxy rice (Oryza sativa L.). *Plant Science*, 166:357-364.

20. Hayashi K, Yasuda N, Fujita Y, Koizumi S, Yoshida H (2010) Identification of the blast resistance gene Pit in rice cultivars using functional markers. *Theoretical and Applied Genetics*, 121:1357-1367.

21. Hirano HY, Eiguchi M, Sano Y (1998) A single base change altered the regulation of the Waxy gene at the post-transcriptional level during the domestication of rice. *Molecular Biology and Evolution*, 15:978-987.

22. Hizukuri S, Abe J I, Hanashiro I (2006) Starch: analytical aspects. In Carbohydrates in Food, 2nd edn. Edited by Eliasson AC, CRC Press, Boca Raton, FL. pp:305-390.

23. Huang Z, He G C, Shu L H, Li X H, Zhang Q F (2001) Identification and mapping of two brown planthopper resistance genes in rice. *Theoretical and Applied Genetics*, 102:929-934.

24. Iida S, Amano E, Nishio T (1993) A rice (Oryza sativa L.) mutant having a low content of glutelin and a high content of prolamine. *Theoretical and Applied Genetics*, 87 (3):374-378.
25. Iida S, Kusaba M, Nishio T (1997) Mutants lacking glutelin subunits in rice: Mapping and combination of mutated glutelin genes. *Theoretical and Applied Genetics*, 94 (2):177-183.

26. Isshiki M, Morino K, Nakajima M, Okagaki RJ, Wessler SR, Izawa T, Shimamoto K (1998) A naturally occurring functional allele for the waxy locus has a GT to TT mutation at the 5' splice of the first intron. *Plant Journal*, 15:133-138.

27. Iyer-Pascuzzi AS, McCouch SR (2007) Functional markers for *xa5*-mediated resistance in rice (*Oryza sativa* L.). *Molecular Breeding*, 19: 291-296.

28. Jairin J, Sansen K, Wongbong W, Kothcharerj K (2010) Detection of a brown planthopper resistance gene *bhp4* at the same chromosomal position for *Bph3* using two different genetic backgrounds of rice. *Breeding Science*, 60:71-75.

29. Jena KK, Kim SM (2010) Current status of brown planthopper (BPH) resistance and genetics. *Rice*, 3: 161-171.

30. Juliano BO (1985) Criteria and tests for rice grain quality In: B. O. Juliano (Ed.) Rice chemistry and technology, American Association of Cereal Chemists, Incorporated, Saint Paul, MN, USA pp:443-513.

31. Kiyosawa S (1972) The inheritance of blast resistance transferred from some *indica* varieties in rice. *Bulletin of the National Institute of Agricultural Sciences Series D*, 23:69-96.

32. Kumar I, Khush GS (1986) Genetic analysis of different amylose levels in rice. *Crop Science*, 27:1167-1172.

33. Kusaba M, Miyahara K, Iida S, Fukuoka H, Takano T, Sassa H, Nishimura M, Nishio T (2003) *Low glutelin content 1*: A dominant mutation that suppresses the glutelin multigene family via RNA silencing in rice. *Plant Cell*, 15 (6):1455-1467.

34. Larkin PD, Park WD (2003) Association of *waxy* gene single nucleotide polymorphisms with starch characteristics in rice (*Oryza sativa* L.). *Molecular Breeding*, 12:335-339.

35. Li XB, Yan WG, Agrama H, Hu BL, Jia LM, Jia M, Jackson A, Moldenhauer K, McClung A, Wu DX (2010) Genotypic and phenotypic characterization of genetic differentiation and diversity in the USDA rice mini-core collection. *Genetica*, 138:1221-1230.

36. Li XB, Yan WG, Agrama H, Jia LM, Shen XH, Jackson A, Moldenhauer K, Yeater K, McClung A, Wu DX (2011) Mapping QTLs for improving grain yield using the USDA rice mini-core collection. *Planta*, 234:347-361.

37. Louis MTB, Robert JH, Jin QS, Russel FR, Daniel LEW (2005) A perfect marker for fragrance genotyping in rice. *Molecular Breeding*, 16:279-283.

38. Lu FH, Park YJ (2012) An SNP downstream of the *OsBEllb* gene is significantly associated with amylose content and viscosity properties in rice (*Oryza sativa* L.). *Journal of Cereal Science*, 56:706-712.

39. Mikami I, Uwatoko N, Ikeda Y, Yamaguchi J, Hirano HY, Suzuki Y, Sano Y (2008) Allelic diversification at the *wx* locus in landraces of Asian rice. *Theoretical and Applied Genetics*, 116 (7):979-989.
40. Miyahara K, Kusaba M, Sassa E, Iida S, Takano T, Nishio T, Nishio T (1996) Analysis of glutelin gene in rice low glutelin line “LGC-1”. *Breeding Science*, 46 (Suppl 1):42.

41. Nakamura Y (2002) Towards a better understanding of the metabolic system for amylopectin biosynthesis in plants: rice endosperm as a model tissue. *Plant Cell Physiol* 43(7):718-725

42. Nakamura Y, Utsumi Y, Sawada T, Aihara S, Utsumi C, Yoshida M, Kitamura S (2010) Characterization of the reactions of starch branching enzymes from rice endosperm. *Plant and Cell Physiology*, 51 (5):776-794.

43. Normile D. 2008. Reinventing rice to feed the world. *Science*, 321, 330-333.

44. Perumalsamy S, Bharani M, Sudha M, Nagarajan P, Arul L, Saraswathi R, Balasubramanian P, Ramalingam J (2010) Functional marker-assisted selection for bacterial leaf blight resistance genes in rice (*Oryza sativa* L.). *Plant Breeding*, 129:400-406.

45. Peter E, Brian R, Graham S (2004) Allele-specific markers within the barley stem rust resistance gene (*Rpg 1*). *Barley Genetics Newsletter*, 33: 7.

46. Qu LQ, Wei XL, Satoh H, Kumamaru T, Ogawa M, Takaiwa F (2002) Inheritance of alleles for glutelin α-2 subunit genes in rice and identification of their corresponding cDNA clone. *Theoretical and Applied Genetics*, 105 (8):1099-1108.

47. Rafalski JA, Tingey SV (1993) Genetic diagnostics in plant breeding: PAPDs, microsatellites and machines. *Trends in Genetics*, 96:957-963.

48. Rahman ML, Jiang WZ, Chu SH, Qiao YL, Ham TH, Woo MO, Lee J, Khanam MS, Chin JH, Jeung JU, *et al.* (2009) High resolution mapping of two rice brown planthopper resistance genes, *Bph(t)* and *Bph21(t)*, originating from *Oryza minuta*. *Theoretical and Applied Genetics*, 119:1237-1246.

49. Ram T, Laha GS, Gautam SK, Deen R, Madhav MS, Brar DS (2010) Identification of new gene introgressed from *Oryza brachyantha* with broad-spectrum resistance to bacterial blight of rice in India. *Rice Genetics Newsletter*, 25: 57.

50. Ramkumar G, Srinivasarao K, Madhan MK, Sudarshan I, Sivarajani AKP, Gopaakrishna K, Neeraja CN, Syahariza ZA, Sar S, Hasjim J, Tizzotti MJ, Gilbert RG (2013) The importance of amylose and amylopectin fine structure for starch digestibility in cooked rice grains. *Food Chemistry*, 136:742-749.

51. Sano Y 1984. Differential regulation of *waxy* gene expression in rice endosperm. *Theoretical and Applied Genetics*, 68:467-473.

52. Sano Y, Katsumata M, Amano E (1985) Correlations between the amounts of amylose and *Wx* protein in rice endosperm. *SABRAO Journal*, 17:121-127.

53. Sharma TR, Chauhan RS, Singh BM, Paul R, Sagar V, Rathore R (2002) RAPD and pathotype analysis of *Magnaporthe grisea* population from North-western Himalayan region of India. *Journal of Phytopathology*, 150:649-656.

54. Teng B, Zeng RZ, Wang YC, Liu ZQ, Zhang ZM, Zhu HT, Ding XH, Li WT, Zhang GQ (2012) Detection of allelic variation at the *Wx* locus with single-segment substitution lines in rice (*Oryza sativa* L.). *Molecular Breeding*, 30:583-595.
55. Tester RF, Karkalas J, Qi X (2004) Starch-composition, fine structure and architecture. *Journal of Cereal Science*, 39:151-165.

56. Tian ZX, Qian Q, Liu QQ, Yan MX, Liu XF, Yan CJ, Liu GF, Gao ZY, Tang SZ, Zeng DL, Wang YH, Umemoto T., Aoki N, Lin H, Nakamura Y, Maruyama S (2004) Natural variation in rice starch synthase iiia affects enzyme and starch properties. *Functional Plant Biology*, 31 (7):671-684.

57. Wanchana S, Toojinda T, Tragoonrung S, Vanavichit A (2003) Duplicated coding sequence in the waxy allele of tropical glutinous rice (*Oryza sativa* L.). *Plant Science*, 165:1193-1199.

58. Wang J, Wan X, Li H, Pfeiffer WH, Crouch J, Wan J (2007) Application of identified QTL-marker associations in rice quality improvement through a design-breeding approach. *Theoretical and Applied Genetics*, 115:87-100.

59. Way MJ, Heong KL (1994) The role of biodiversity in the dynamics and management of insect pests of tropical irrigated rice-a review. *Bulletin of Entomological Research*, 84:567-587.

60. Yara A, Phi CN, Matsumura M, Yoshimura A, Yasui H (2010) Development of near-isogenic lines for *BPH25(t)* and *BPH26(t)*, which confer resistance to the brown planthopper, *Nilaparvata lugens* (Stal.) in indica rice “ADR52”. *Breeding Science*, 60:639-647.

61. Yu JM, Gu MH, Li JY (2009) Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. *Proceedings of the National Academy of Sciences of the United States of America*, 106:21760-21765.

62. Zhang G, Cheng Z, Zhang X, Guo X, Su N, Jiang L, Mao L, Wan JM (2011) Double repression of soluble starch synthase genes SSIIa and SSIIIa in rice (*Oryza sativa* L.) uncovers interactive effects on physic-chemical properties of starch. *Genome*, 54: 448-459.

63. Zhou L, Chen ZJ, Lang XY, Du B, Liu K, Yang GC, Hu G, Li SH, He GC, You AQ (2013) Development and validation of a PCR-based functional marker system for the brown planthopper resistance gene *Bph14* in rice. *Breeding Science*, 63:347-352

**Tables**

Due to technical limitations, table 1-3 is only available as a download in the Supplemental Files section.

**Figures**
Figure 1

Wx haplotype frequencies in USDA rice mini-core collection. Wx gene haplotypes were based on the G/T SNP in intron 1, the 23 bp duplication in exon 2, the A/C SNP in exon 6 and C/T SNP in exon 10. Wx: T-2-A-C; Wxt: T-1-A-C; Wxg1: G-1-C-C; Wxg2: G-1-A-T; Wxg3: G-1-A-C; Wxt1: T-1-C-C; Wxt2: T-1-A-T; Wxt3: T-1-C-T.
Figure 2

a. Functional marker for the detection of 144 bp deletion in Pi54 gene (2% agarose gel). Lanes 1-14 are rice accessions GSOR 311697, NSGC 5935, A100943-R, W 1193, NSGC 5953, Mayang Khang, E B Gopher, RD 218, C 5560, LEAH, Quinimpol, Ao Chiu 2 Hao, 81B/25, Karang Serang respectively. Lane M: 100 bp DNA Ladder (Takara).

b. Functional marker for the detection of TC/AG SNP in Xa5 gene (2% agarose gel). Lanes 1-14 are rice accessions GSOR 311697, NSGC 5935, A100943-R, W 1193, NSGC 5953, Mayang
Khang, E B Gopher, RD 218, C 5560, LEAH, Quinimpol, SHIMIZU MOCHI, Padi Pohon Batu, Gazan respectively. Lane M: 100 bp DNA Ladder (Takara).

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- Suppl.Table1.docx
- Tables.pdf