Rapid identification of a new gene influencing low amylose content in rice landraces (*Oryza sativa* L.) using genome-wide association study with specific-locus amplified fragment sequencing

Journal:	<i>Genome</i>
Manuscript ID	gen-2016-0104.R4
Manuscript Type:	Article
Date Submitted by the Author:	28-Nov-2016
Complete List of Authors:	Yang, Xinghai; Guangxi Academy of Agricultural Sciences, Rice Research Institute. Nong, Baoxuan; Guangxi Academy of Agricultural Sciences, Rice Research Institute. Xia, Xiuzhong; Guangxi Academy of Agricultural Sciences, Rice Research Institute. Zhang, Zongqiong; Guangxi Academy of Agricultural Sciences, Rice Research Institute. Zeng, Yu; Guangxi Academy of Agricultural Sciences, Rice Research Institute. Liu, Kaiqiang; Guangxi Academy of Agricultural Sciences, Rice Research Institute. Deng, Guofu; Guangxi Academy of Agricultural Sciences, Rice Research Institute. Li, Danting; Guangxi Academy of Agricultural Sciences, Rice Research Institute.
Keyword:	Rice landraces, Very low amylose content, SLAF-seq, GWAS, Candidate gene

SCHOLARONETM
Manuscripts

**Rapid identification of a new gene influencing low amylose content in rice
landraces (*Oryza sativa* L.) using genome-wide association study with
specific-locus amplified fragment sequencing**

Xinghai Yang¹, Baoxuan Nong¹, Xiuzhong Xia¹, Zongqiong Zhang¹, Yu Zeng¹,
Kaiqiang Liu¹, Guofu Deng¹, Danting Li^{1*}

¹Rice Research Institute, Guangxi Academy of Agricultural Sciences, Nanning
530007, China

*Corresponding author: (Telephone number: 86-771-3244040; E-mail:
13978672750@163.com, ricegl@163.com.)

Address: 174 East Daxue Road, Nanning, Guangxi 530007, China

Abstract: Starch is the major components of milled rice, and amylose content (AC) affects eating quality. In this study, genome-wide association study (GWAS) with specific-locus amplified fragment sequencing (SLAF-seq) data was performed for AC in a panel of 419 core collection for rice landraces. Using the compressed mixed linear method based on the Q+K model, we identified a new gene *LAC6* (Chr6: 5.65 Mb ~ 6.04 Mb) associated with AC in the low amylose content panel. The candidate genes of *LAC6* was detected through qRT-PCR in rice panicles. The results indicated that *LOCs06g11130* was up-regulated, and *LOC_Os06g11340* showed significantly down-regulated, which was most likely a key candidate gene of *LAC6*. In conclusion, the findings provide a certain theoretical basis of molecular biology for genetic improvement of AC in rice and rice quality variety breeding.

Key words: Rice landraces, Very low amylose content, SLAF-seq, GWAS, Candidate gene

Introduction

As one of the most important crops, rice (*Oryza sativa* L.) has the largest planting area in the world, and it is a staple food for more than half of the world population (Ashikari et al. 2005). Rice landraces have evolved from their wild progenitors under human and natural selection, leading to the maintenance of high genetic distribution, which is the basis in rice breeding and germplasm innovation.

Starch consisting of amylose and amylopectin is the main storage substance in rice, and its content is about 90% of milled rice weight. AC is one of the most important factors of eating quality of cooked rice (Preiss 1991; Smith et al. 1997; He et al. 1998). Generally, high AC of rice varieties leads to cooking, dilatant, sticky rice and poor softness, and the rice is easy to harden, resulting in poor flavor. Low AC causes sticky, elastic, soft, easy to digest. Medium AC represents good viscosity and softness, not sticky after cold, soft, good palatability. Therefore, understanding of the genetic mechanisms underlying AC is the most effective way to improve the rice eating quality (Sano 1984; Sano et al. 1986; Wang et al. 1990; Mikami et al. 1999; Sato et al. 2002; Mikami et al. 2008). AC is controlled by a major gene and multiple minor genes. The *Waxy* (*Wx*) gene encodes a granule-bound starch synthase (GBSS) that plays a major role in amylose synthesis of rice (Preiss 1991; Wang et al. 1995; Smith et al. 1997). In rice, waxy and non-waxy rice are mainly caused by the difference between *Wx* and *wx*, a pair of alleles in this locus. At present, the allele of *Wx* gene has been identified in rice, such as *Wx^a*, *Wx^b*, *Wx^{mq}*, *Wx^{mp}*, *Wxⁱⁿ*, *wx* and so on (Sato et al. 2002; Mikami et al. 2008; Liu et al. 2009; Yang et al. 2013). Previous studies

have reported genetic and molecular analyses of the dull endosperm in rice. *Dull* genes of EM47, 2120 and 2035 are assigned to chromosomes 6, 9 and 6, respectively (Kaushik 1991). The *du-1* and *du-2* genes of rice encode SR protein regulate splicing factors that are involved in alternative splicing of pre-mRNA in rice (Isshiki et al. 2000). The *du-1* gene is located on chromosome 7, and it is not allelic to *du-2* (Yano et al. 1988). *Du-1* encodes a novel Prp1 protein (Zeng et al. 2007), and it is located on chromosome 10. *Du-3* is an mRNA cap-binding protein gene regulating AC in Japonica rice seeds, and it is located on chromosome 2. Takemoto-Kuno et al. mapped a QTL (*qAC2*) on chromosome 2, which is able to interact with *Wx* and control the low AC of rice (Takemoto-kuno et al. 2015).

As an efficient strategy, genome-wide association study (GWAS) is often used to establish the relationships between the genome-wide high-density molecular markers and phenotypic traits by scanning data of molecular markers in target populations (Aranzana et al. 2005). Based on the level of molecular markers, high-density genetic markers are required in GWAS, such as single nucleotide polymorphisms (SNPs), copy number variations (CNVs), simple sequence repeats (SSRs) and so on. It can directly assess DNA mutations at the genome-wide level, and assumptions of candidate genes are no longer required. GWAS is regarded as a powerful approach, and it is widely used to indentify underlying genes with complex diseases at an unprecedented speed (Hansen et al. 2001; Klein et al. 2005). Subsequently, a series of GWAS investigations have been published (Helgadottir et al. 2007; Wang et al. 2010; Jia et al. 2013; Xu et al. 2012). At present, GWAS has been successfully applied to

the dissection of yield, resistance and other agronomic traits in rice ([Huang et al. 2010](#); [Huang et al. 2012](#); [Wang et al. 2014](#); [Si et al. 2016](#); [Yano et al. 2016](#)) .

Guangxi is located in the south of China, which is probably the first place of development of cultivated rice ([Huang et al. 2012](#)). In the present study, we aimed to identify a substantial number of loci related to AC. We investigated AC via GWAS based on genotyping 1,090,071 SNPs variants across 419 rice landraces from Guangxi Zhuang Autonomous Region. We mapped a new gene *LAC6* in the very low AC varieties via GWAS. The findings will be helpful to understand the genetic mechanisms underlying AC and Provides theoretical basis for improvement of rice eating quality.

Materials and methods

Plant materials and determination of AC

The natural population consisted of 419 experimental materials, which were landraces as the core collections of rice germplasm derived from Guangxi. These materials were germinated and planted in the experimental field (in Nanning, China at N 22.85°, E 108.26°) from July 2014 to November 2014. Seeds were harvested at maturity and used for AC measurement.

AC was determined as previously described ([Juliano 1985](#)). Milled rice flour (100 $\pm$ 0.5 mg) was digested with 1 mL of 95% (v/v) ethanol and 9.0 mL of 1.00 mol/L NaOH overnight. Then the solution was mixed with 1.5 mL KI solution containing 0.2% I₂ and diluted with 1.0 mL of 1 mol/L acetic acid. The amylose-iodine color was measured at a wavelength of 620 nm using a spectrophotometer (TU-1810, China).

DNA extraction

Young quadrifoliate leaves from natural individuals were collected, frozen in liquid nitrogen, and then transferred to a -70°C freezer. Total genomic DNA was extracted from each leaf sample following the cetyltrimethylammonium bromide (CTAB) method (Porebski et al. 1997) with some modifications to the components of the CTAB buffer (8.19 g sodium chloride and 2 g CTAB in 100 mL of 20 mM EDTA, 100 mM Tris, pH 8.0). DNA concentration and integrity were assessed with a Bio-spec Nano spectrophotometer (Shimadzu, Kyoto, Japan) and electrophoresis (0.8% agarose gel) with a lambda DNA standard.

Specific-locus amplified fragment sequencing (SLAF-seq)

All DNA samples were used to construct modified SLAF library as previously described (Sun et al. 2013). Briefly, genome DNA was completely digested with HaeIII (New England Biolabs, NEB), a single-nucleotide A overhang was added to the digested fragments using Klenow Fragment (3' $\rightarrow$ 5' exo-) (NEB) at 37°C , and then the Duplex Tag-labeled sequencing adapters (PAGE purified, Life Technologies, USA) were ligated to the A-tailed DNA using T4 DNA ligase. The PCR reaction was performed using diluted restriction-ligation samples, dNTPs, Q5[®] High-Fidelity DNA Polymerase and PCR primers: 5'-AATGATACGGCGACCACCGA-3' and 5'-CAAGCAGAAGACGGCATACG-3' (PAGE purified, Life Technologies, USA). The PCR products were purified using Agencourt AMPure XP beads (Beckman Coulter, High Wycombe, UK) and then pooled. The pooled samples were separated by electrophoresis on a 2% agarose gel. Fragments of 400~450 bp (with indexes and

adaptors) in size were excised and purified using QIAquick Gel Extraction Kit (QIAGEN, USA). The gel-purified product was sequenced on the Illumina HiSeq 2500 system (Illumina, Inc.; San Diego, CA, USA) according to the manufacturer's instructions.

Genome-wide association study

The Short Oligonucleotide Alignment Program 2 (SOAP2) was used to map raw paired-end reads onto the reference genome (Li et al. 2009). The SLAF groups were generated by reads that were mapped to the same position. If an accession was only partly digested by the restriction enzymes, some reads mapped to the reference genome should have overlaps with two SLAF tags. In SNP calling, minor allele frequency threshold was set at 0.05, and all samples were genotyped by GATK package (Mckenna et al. 2010). If depth of minor allele was larger than 1/3 of the sample total depth, the genotype was regarded as heterozygous.

The mean, range, variance, standard deviation and skew of AC were calculated. GWAS was carried out using a compressed mixed linear model (Zhang et al. 2010) as follows:

$$Y = X\alpha + Q\beta + K\mu + e$$

Where Y is the phenotype value vector, X is the genotype value vector, and α is the effect of genotype and the population structure matrix calculated by the ADMIXTURE program. The proportion of each group was fitted as covariates: β is the effect of population structure vector of each group, K is the genotype matrix, μ is the weight vector of each marker, and e is the random error. Finally, each SNP locus

had an assigned value associated with related traits.

To identify candidate genes and predict their putative functions in the associated loci for the corresponding strain, we used gene annotation information from the Rice Genome Annotation Project (<http://rice.plantbiology.msu.edu/>).

Plant total RNA extraction, reverse transcription and qRT-PCR analysis

Panicles from rice plants in CK, group I and group II were collected on day 12 after flowering, frozen in liquid nitrogen, and used for RNA extraction. Total RNA was isolated from leaves using CWBio Ultrapure RNA Kit (Beijing, China). RNA concentration and quality were estimated using a BioPhotometer Plus spectrophotometer (Eppendorf, Hamburg, Germany). Reverse transcription was performed using TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix (Beijing, China). Primers for amplification of 14 genes were designed using Primer Premier 5.0 and are listed in [Supplementary Table 1](#). qRT-PCR was performed using a AceQ_qPCR_SYBR Green_Master_Mix(Nanjing, China). The rice actin gene was used as an internal standard (forward: 5'-GAACTGGTATGGTCAAGGCTG-3', reverse: 5'-ACACGGAGCTCGTTGTAGAAG-3'). Reactions were performed in 96-well plates. Briefly, after an initial denaturation step at 95°C for 30s, the amplification was carried out with 45 cycles at a melting temperature of 95°C for 10 s and an annealing temperature of 55°C for 60 s. qRT-PCR analysis was performed on a qTOWER2.2 (AnalytikJena, Germany). The relative expression levels of the target genes were determined using the $2^{-\Delta\Delta Ct}$ method.

Results

AC of 419 rice landraces

We investigated the AC of 419 rice landraces, and [Supplementary Table 2](#) summarizes the results. The classification of AC was identified as waxy (0-2%) , very low (2-12%), low (12-20%), intermediate (20-25%), and high (>25%) ([Xu et al. 2013](#); [Biselli et al. 2014](#))([Fig. 1](#)).

Analysis of SLAF-seq data and SLAF markers

A total of 33.5 Gb raw data were generated from Illumina sequencing and SLAF library construction, which contained 261,385,070 pair-end reads with a length of 80 bp. The GC (guanine-cytosine) content was 44.06%, and Q20 percentage (a quality score of 20) was 88.39%. Sample groups with an average sequencing depth greater than 2 were selected to define SLAF markers. Finally, 67,665 SLAF markers distributed throughout the genomes, and 56,768 SLAF were polymorphic markers. The distribution of all the SNPs in the genomes of 419 samples was determined according to the number of SNPs per 100 kb in the genome ([Fig. 2](#)). Finally, 1,090,071 SNPs were eligible for inclusion in the GWAS.

Group structure and clustering analysis

Based on SNPS, we calculated the sample group structure by admixture software, and the K-value that produced the best dataset was 6 ([Fig. 3a](#)). Proportions of the seven groups were fitted as covariates and substituted into the general linear model to calculate a p value. The clustering results were shown in [Fig. 3b](#).

Validation of known genes using GWAS

To investigate the genotypic variation underlying AC to rice, GWAS was carried out to identify the associated loci in landraces. Both the general linear model (GLM) (Fig. 4a and Fig. 4b) and compressed mixed linear model (MLM) (Fig. 4c and Fig. 4d) were used to identify association signals. Comparison of the mixed linear model, and the general linear model revealed 83 significant association signal differences (Supplementary Table 3 and Supplementary Table 4). The compressed mixed linear model, which took genome-wide patterns of genetic relatedness into account, greatly reduced false positives (Huang et al. 2010; Zhang et al. 2010). Therefore we used mixed linear model to analyze the 419 rice landraces.

In this study, a total of 116 significant association signals were identified as strong association signals with $P < 4.78 \times 10^{-8}$ (false discovery rate < 0.01) from the compressed mixed linear model (Fig. 4a). We focused on three of these loci, and two association signals for AC were located close to known genes that have been previously identified.

We analyzed a peak on chromosome 6, which was mapped close to *Wx*. We estimated a candidate region from 1.67 Mb to 1.92 Mb (250 kb) according to the LD decay. The peak SNP, Chr06_1797551 ($P = 3.97 \times 10^{-35}$), explaining up to 46.3% of the phenotypic variance (Supplementary Table 5) was located 26.9 kb downstream of the *Wx* gene (*LOC_Os06g04200*), a rice GBSS gene, and it was the major gene controlling AC in rice (Sano 1984; Hirano and Sano 2000).

Similarly, we analyzed another peak on chromosome 6, which was mapped close to *SSIIa*. We estimated a candidate region from 6.60 Mb to 6.87 Mb (272 kb) according

to the LD decay. The peak SNP, Chr06_ 6733351 ($P=9.67\times 10^{-15}$), explaining up to 32.0% of the phenotypic variance ([Supplementary Table 3](#)), *LOC_Os06g12450* was located near this locus and described as soluble starch synthase II. This finding suggested that it was *SSIIa*, which has significant correlation with AC ([Varavinit et al. 2003](#); [Su et al. 2014](#)). The identification of *Wx* and *SSIIa* indicated that GWAS with SLAF-seq could be used to efficiently identify genes associated with agronomic traits. We discussed the third highest peak signal on chromosome 6 below.

Identification of new gene

We used the same strategy on loci associated with AC, which have not been reported previously. We identified a major locus Chr06_ 5843510 ($P = 3.77\times 10^{-13}$) ([Supplementary Table 5](#)), which was responsible for very low AC in this rice landraces. We designated this locus as *LAC6* (low AC on chromosome 6). Further analysis showed that the *LAC6* explained about 18.9% of the variance in very low AC panels ([Supplementary Table 5](#)).

We estimated a candidate region from 5.65 Mb to 6.04 Mb (380 kb) according to the LD decay. All the polymorphisms in the candidate region were classified into four groups ([Supplementary Table 6](#)). Group I included polymorphisms that were significantly associated with trait variation in the GWAS ($P = 4.78\times 10^{-8}$), and they were predicted to occur in exon region. Group II included polymorphisms significantly associated with trait, and they were located at the promoter region (≤ 2 kb from the 5' flanking sequences of genes). Group III included polymorphisms significantly associated with trait, and they were located at the 5'UTR or 3'UTR

region. Group IV included polymorphisms significantly associated with trait, and they were located at the intron region. Given the four groups of the significant polymorphism-associated loci, we carefully analyzed the 13 predicted genes within the 380-kb region and *waxy* gene (*LOC_Os06g04200*). We selected Kasalath, five lines each of *waxy* and very low AC varieties to measure the expression levels of these 14 genes in panicles after removing hulls via quantitative RT-PCR (qRT-PCR) (Fig. 5). Of the 14 genes, two genes shown significant difference in *waxy* and very low AC varieties as compared with Kasalath. The expression levels of *LOC_Os06g11130*, encoding gibberellin receptor GID1L2, showed about 2-fold higher expression in *waxy* and very low AC varieties than in Kasalath. Conversely, the expression levels of two genes, encoding retrotransposon protein (*LOC_Os06g11340*), starch synthase (*LOC_Os06g04200*), were significantly lower in *waxy* and very low AC varieties than in Kasalath, and those in very low AC varieties were significantly higher than *waxy* varieties.

Discussion

The combination of SLAF-seq technology, and GWAS provides an efficient method to identify genomic regions related to traits. SLAF-seq technology produced millions of candidate SNPs, which were widely distributed in the whole genome (Sun et al. 2013). Using high-density genome-wide SNPs in GWAS has been proved to be an efficient tool for gene discovery.

Both GLM and the compressed MLM were used in this GWAS. We detected 199 significant association loci using the GLM, and 116 significant association loci were

detected using the compressed MLM. Compared with the GLM, the compressed MLM had less loci. Because the compressed MLM controls for population structure and relatedness withing GWAS, leading to greatly reduced false positives ([Huang et al. 2010](#); [Zhang et al. 2010](#)).

In this study, we detected genes for AC using GWAS baed on SLAF-seq. As we expected, a major AC gene *Wx*, was detected in the whole panel.

Meanwhile, the gene *SSIIa*, which plays primarily functions for starch gelatinization temperature ([Gao et al. 2003](#)) and roles during rice grain starch synthesis([Zhang et al. 2011](#)), was detected in the whole panel. This finding indicated that the *SSIIa* genes likely have a new function for AC, This result is in agreement with the previous finding that AC were very significant correlation with gelatinization temperature in rice ([Varavinit et al. 2003](#); [Sun et al. 2005](#); [Su et al. 2014](#)). A new gene mapping was generated from 5.65 Mb to 6.04 Mb on chromosome 6. We defined the candidate regions with significant signals based on LD, extracted the causal genes based on the function of polymorphisms, and classified four groups including 21 polymorphisms in t13 genes. qRT-PCR shown that the gene expression levels, of two genes, *LOCOs06g11130* and *LOC_Os06g11340* , were significantly higher in the very low AC than in Kasalath. The gene *LOCOs06g11130*, encoding gibberellin receptor GID1L2, was up-regulated in panicles in waxy and very low AC varieties. Previous study hasreported that the aleurone tissue of immature wheat grains can also be induced to respond to GA₃ ([Cornford et al. 1986](#)). The grains naturally become responsive to gibberellins when grains are older than 45 days post anthesis of wheat

(Armstrong et al. 1982). The expression levels of the gene *LOC_Os06g11340* was similar to *LOC_Os06g04200* (*waxy* gene), which encodes retrotransposon protein, was down-regulated in waxy and very low AC varieties. These results indicate that they could play an important role in AC in rice.

Conclusions

In summary, our results further confirmed that GWAS based on SLAF-seq is a powerful tool for the identification of genes associated with agronomic traits in rice. A total of 116 significant association loci were identified and validated two known genes *Wx* and *SSIa*. Meanwhile, we detected a new gene *LAC6* which could be the important region of influencing the amylose content in rice. The expression levels of *LAC6* in rice panicle were characterized. Thus, the systematic analysis of AC results may provide valuable information for rice quality improvement.

Conflict of interest

The authors declare that there is no conflict of interest associated with this work.

Acknowledgements

This study was financially supported by Guangxi Natural Science Foundation of China (2015GXNSFBA139054), Guangxi's Ministry of Science and Technology (14121001-1-6) and Guangxi Academy of Agricultural Sciences (2015YT15).

References

- Aranzana, M.J., Kim, S., Zhao K., Bakker, E., Horton, M., Jakob, K., et al. 2005. Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. *PLoS Genet*, **1**(5),e60. doi:[10.1371/journal.pgen.0010060](https://doi.org/10.1371/journal.pgen.0010060).
- Armstrong, C., Black, M., Chapman, J.M., Norman, H.A., Angold, R.1982. The induction of sensitivity to gibberellin in aleurone tissue of developing wheat grains. *Planta*, **154**(6),573. doi:[10.1007/BF00403003](https://doi.org/10.1007/BF00403003).
- Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takashi, T., Nishimura, A., et al. 2005. Cytokinin oxidase regulates rice grain production. *Science*, **309**(5735):741-5. doi:[10.1126/science.1113373](https://doi.org/10.1126/science.1113373). PMID:[15976269](https://pubmed.ncbi.nlm.nih.gov/15976269/).
- Biselli, C., Cavalluzzo, D., Perrini, R., Gianinetti, A., Bagnaresi, P., Urso, S., et al. 2014. Improvement of marker-based predictability of Apparent Amylose Content in japonica rice through GBSSI allele mining. *Rice*, **7**(1), 1. doi:[10.1186/1939-8433-7-1](https://doi.org/10.1186/1939-8433-7-1). PMID: [24383761](https://pubmed.ncbi.nlm.nih.gov/24383761/).
- Cornford, C.A., Black, M., Chapman, J.M., Baulcombe, D.C. 1986. Expression of α -amylase and other gibberellin-regulated genes in aleurone tissue of developing wheat grains. *Planta*, **169**(3),420. doi:[10.1007/BF00392140](https://doi.org/10.1007/BF00392140). PMID: [24232656](https://pubmed.ncbi.nlm.nih.gov/24232656/).
- Gao, Z., Zeng, D., Cui, X., Zhou, Y., Yan, M., Huang, D., et al. 2003. Map-based cloning of the *ALK* gene, which controls the gelatinization temperature of rice. *Science in China (Series C)*, **46**(6), 661-668. doi:[10.3969/j.issn.1674-7305.2003.06.011](https://doi.org/10.3969/j.issn.1674-7305.2003.06.011).

- Hansen, M., Kraft, T., Ganestam, S., Säll, T., Nilsson, N.O. 2001. Linkage disequilibrium mapping of the bolting gene in sea beet using AFLP markers. *Genet RES*, **77**(1),61-66. PMID:[11279831](#).
- He, P., Shen, L., Lu, C., Chen, Y., Zhu, L. 1998. Analysis of quantitative trait loci which contribute to anther culturability in rice. *MOL Breeding*, **4**,165-172. doi:[10.1023/A:1009692221152](#).
- Helgadóttira, T.T., Jonasdóttir, J.G., Gudbjartsson, M.V., Matthíasdóttir, J.S., Gylfason, V.D., Shah, Q.A., et al. 2007. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*, **316**(5380),1491-1493. PMID: [17478679](#).
- Hirano, H.Y., SANO, Y., 2000. Comparison of Waxy gene regulation in the endosperm and pollen in *Oryza sativa* L. *Genes Genet Syst*, **75**, 245-249. PMID: [11245217](#).
- Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., et al. 2010. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet*, **42**(11),961-7. doi:[10.1038/ng.695](#). PMID: [20972439](#).
- Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., et al. 2011. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet*, **44**(1),32-39. doi: [10.1038/ng.1018](#). PMID: [22138690](#).
- Huang, X., Kurata, N., Wei, X., Wang, Z.X., Wang, A., Zhao, Q., et al. 2012. A map of rice genome variation reveals the origin of cultivated rice. *Nature*,

- 490**(7421),497-501. doi:[10.1038/nature11532](https://doi.org/10.1038/nature11532). PMID:[23034647](https://pubmed.ncbi.nlm.nih.gov/23034647/).
- Isshiki, M., Morino, K., Nakajima, M., Okagaki, R.J., Wessler, S.R., Izawa, T., et al. 1998. A naturally occurring functional allele of the rice *waxy* locus has a GT to TT mutation at the 5' splice site of the first intron. *Plant J*, **15**(1),133-138. doi:[10.1046/j.1365-3113X.1998.00189.x](https://doi.org/10.1046/j.1365-3113X.1998.00189.x). PMID: [9744101](https://pubmed.ncbi.nlm.nih.gov/9744101/).
- Jia, G., Huang, X., Zhi, H., Zhao, Y., Zhao, Q., Li, W., et al. 2013. A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). *Nat Genet*, **45**(8),957-961. doi: [10.1038/ng.2673](https://doi.org/10.1038/ng.2673). PMID: [23793027](https://pubmed.ncbi.nlm.nih.gov/23793027/).
- Juliano, B. 1985. Rice chemistry and technology. American Association of Cereal Chemists, Sanit Pual Minnesota, USA. pp. 443-513.
- Kaushik, R.P., Khush, G.S. 1991. Genetic analysis of endosperm mutants in rice *Oryza sativa* L. *Theor Appl Genet*. **83**(2),146-152. doi:[10.1007/BF00226243](https://doi.org/10.1007/BF00226243). PMID:[24202350](https://pubmed.ncbi.nlm.nih.gov/24202350/).
- Klein, R.J., Zeiss, C., Chew, E.Y., Tsai, J.T., Sackler, R.S., Haynes, C., et al. 2005. Complement factor H polymorphism in age-related macular degeneration. *Science*, **308**(5720),385-389. doi:[10.1126/science.1109557](https://doi.org/10.1126/science.1109557). PMID:[15761122](https://pubmed.ncbi.nlm.nih.gov/15761122/).
- Li, R., Yu, C., Li, Y., Lam, T.W., Yiu, S.M., Kristiansen, K., et al. 2009. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics*, **25**(15),1966-1967. doi:[10.1093/bioinformatics/btp336](https://doi.org/10.1093/bioinformatics/btp336). PMID:[19497933](https://pubmed.ncbi.nlm.nih.gov/19497933/).
- Liu, L., Ma, X., Liu, S., Zhu, C., Jiang, L., Wang, Y., et al. 2009. Identification and characterization of a novel *Waxy* allele from a Yunnan rice landrace. *Plant Mol*

- Biol, **71**(6),609-626. doi:[10.1007/s11103-009-9544-4](https://doi.org/10.1007/s11103-009-9544-4). PMID:[19760367](https://pubmed.ncbi.nlm.nih.gov/19760367/).
- Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytksy, A., et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome RES*, **20**(9),1297-1303. doi:[10.1101/gr.107524](https://doi.org/10.1101/gr.107524). PMID: [20644199](https://pubmed.ncbi.nlm.nih.gov/20644199/).
- Mikami, I., Aikawa, M., Hiro-yuki, H., Sano, Y. 1999. Altered tissue-specific expression at the *Wx* gene of the opaque mutants in rice. *Euphytica*, **105**(2),91-97. doi:[10.1023/A:1003457209225](https://doi.org/10.1023/A:1003457209225).
- Mikami, I., Uwatoko, N., Ikeda, Y., Yamaguchi, J., Hirano, H.Y., Suzuki, Y., et al. 2008. Allelic diversification at the *wx* locus in landraces of Asian rice. *Theor Appl Genet*, **116**(7),979-989. doi:[10.1007/s00122-008-0729-z](https://doi.org/10.1007/s00122-008-0729-z). PMID:[18305920](https://pubmed.ncbi.nlm.nih.gov/18305920/).
- Mikami, I., Aikawa, M., Hiro-yuki, H., Sano, Y.1999. Altered tissue-specific expression at the *Wx* gene of the opaque mutants in rice. *Euphytica*, **105**(2),91-97. doi:[10.1023/A:1003457209225](https://doi.org/10.1023/A:1003457209225).
- Porebski, S., Bailey, L.G., Baum, B. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol Biol Rep*, **15**(1),8-15. doi:[10.1007/BF02772108](https://doi.org/10.1007/BF02772108).
- Preiss, J. 1991. Biology and molecular biology of starch synthesis and its regulation. *Oxf Surv Plant Mol Cell Biol*, **7**,59-114. PMID:[92020506](https://pubmed.ncbi.nlm.nih.gov/92020506/).
- Sano, Y. 1984. Differential regulation of waxy gene expression in rice endosperm. *Theor Appl Genet*, **68**(5),467-473. doi:[10.1007/BF00254822](https://doi.org/10.1007/BF00254822). PMID:[24257739](https://pubmed.ncbi.nlm.nih.gov/24257739/).
- Sano, y., Katsumata , M., Okuno, K. 1986. Genetic studies of speciation in cultivated

- rice. 5 Inter- and intra-specific differentiation in the waxy gene expression in rice. *Euphytica*, **35**(1),1-9. doi:[10.1007/BF00028534](https://doi.org/10.1007/BF00028534).
- Sato, H., Suzuki, Y., Sakai, M., Imbe, T. 2002. Molecular Characterization of *Wx^{-mq}*, a Novel Mutant Gene for Low-amylose Content in Endosperm of Rice (*Oryza sativa* L.). *Breeding Science*, **52**(2),131-135.doi:[10.1270/jsbbs.52.131](https://doi.org/10.1270/jsbbs.52.131).
- Si, L., Chen, J., Huang, X., Gong, H., Luo, J., Hou, Q., et al. 2016. *OsSPL13* controls grain size in cultivated rice. *Nat Genet*, **48**(4),447-456. doi:[10.1038/ng.3518](https://doi.org/10.1038/ng.3518). PMID:[26950093](https://pubmed.ncbi.nlm.nih.gov/26950093/).
- Smith, A.M., Denyer, K., Martin, C. 1997. The synthesis of the starch granule. *ANNU REV Plant Physiol Plant Mol Biol*, **48**(7),67-87. doi:[10.1146/annurev.arplant.48.1.67](https://doi.org/10.1146/annurev.arplant.48.1.67). PMID:[15012257](https://pubmed.ncbi.nlm.nih.gov/15012257/).
- Su, L., Xiang, X., Xu, Y.,Long, X.,Kang, C.,Xu, S. 2014. The Influence of *SSIIa* Genotype on the Calibration of Pasting Temperature of Rice and the Relationship between Gelatinization Temperature and Amylose Content. *Journal of Plant Genetic Resources*, **15**(2),348-353. doi:[10.13430/j.cnki.jpgr.2014.02.018](https://doi.org/10.13430/j.cnki.jpgr.2014.02.018).(Chinese)
- Sun X., Liu D., Zhang X., Li W., Liu H., Hong W., et al. 2013. SLAF-seq: an efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. *PLoS One*, **8**(3),e58700. doi:[10.1371/journal.pone.0058700](https://doi.org/10.1371/journal.pone.0058700). PMID: [23527008](https://pubmed.ncbi.nlm.nih.gov/23527008/).
- Sun, Y., Lv, Y., Dong, C., Wang, P., Huang, X., Deng, X. 2005. Genetic Relationship among *Wx* Gene, AC, GC and GT of Rice. *Acta Agronomica Sinica*, **31**(5),535-539. (Chinese)

- Takemoto-kuno, Y., Mitsueda, H., Suzuki, K., Hirabayashi, H., Ideta, O., Aoki, N., et al. 2015. *qAC2*, a novel QTL that interacts with *Wx* and controls the low amylose content in rice (*Oryza sativa* L.). *Theor Appl Genet*, **128**(4),563-573. doi: [10.1007/s00122-014-2432-6](https://doi.org/10.1007/s00122-014-2432-6). PMID: [25762131](https://pubmed.ncbi.nlm.nih.gov/25762131/).
- Varavinit, S., Shobsngob, S., Varayanond, W., Chinachoti, P., Naivikul, O. 2003. Effect of Amylose Content on Gelatinization, Retrogradation and Pasting Properties of Flours from Different Cultivars of Thai Rice. *Starch-Starke*. **55**(9), 410-415. doi:[10.1002/star.200300185](https://doi.org/10.1002/star.200300185).
- Wang, C.H., Yang, Y.L., Yuan, X.P., Xu, Q., Yue, F., Yu, H.Y., et al. 2014. Genome-wide association study of blast resistance in indica rice. *BMC Plant Biol*, **14**(1):311. doi:[10.1186/s12870-014-0311-6](https://doi.org/10.1186/s12870-014-0311-6). PMID:[25403621](https://pubmed.ncbi.nlm.nih.gov/25403621/).
- Wang, K., Diskin, S.J., Zhang, H., Attiyeh, E.F., Winter, C., Hou, C., et al. 2010. Integrative genomics identifies LMO1 as a neuroblastoma oncogene. *Nature*, **469**(7329),216-220. doi:[10.1038/nature09609](https://doi.org/10.1038/nature09609). PMID:[21124317](https://pubmed.ncbi.nlm.nih.gov/21124317/).
- Wang, Z.Y., Wu, Z.L., Xing, Y.Y., Zheng, F.G., Guo, X.L., Zhang, W.G., et al. 1990. Nucleotide sequence of rice waxy gene. *Nucleic Acids Res*, **18**(19),5898. doi:[10.1093/nar/18.19.5898](https://doi.org/10.1093/nar/18.19.5898). PMID: [216792](https://pubmed.ncbi.nlm.nih.gov/216792/).
- Wang, Z.Y., Zheng, F.Q., Shen, G.Z., Gao, J.P., Snustad, D.P., Li, M.G., Zhang, J.L., Hong, M.M. 1995. The amylose content in rice endosperm is related to the post-transcriptional regulation of the waxy gene. *Plant J*, **7**(4), 613-22. doi: [10.1046/j.1365-313X.1995.7040613.x](https://doi.org/10.1046/j.1365-313X.1995.7040613.x). PMID: [7742858](https://pubmed.ncbi.nlm.nih.gov/7742858/).
- Xu, J., Mo, Z., Ye, D., Wang, M., Liu, F., Jin, G., et al. 2012. Genome-wide

- association study in Chinese men identifies two new prostate cancer risk loci at 9q31.2 and 19q13.4. *Nat Genet*, **44**(11),1231-1237. doi:[10.1038/ng.2424](https://doi.org/10.1038/ng.2424). PMID:[23023329](https://pubmed.ncbi.nlm.nih.gov/23023329/).
- Xu, F., Zhang, G., Tong, C., Sun, X., Corke, H., Sun, M., et al(2013). Association Mapping of Starch Physicochemical Properties with Starch Biosynthesizing Genes in Waxy Rice (*Oryza sativa* L.). *J Agric Food Chem*, **61**(42), 10110-11117. doi:[10.1021/jf4029688](https://doi.org/10.1021/jf4029688). PMID:[24063600](https://pubmed.ncbi.nlm.nih.gov/24063600/).
- Yang, J., Wang, J., Fan, F.J., Zhu, J., Chen, T., Wang, C.L., et al. 2013. Development of AS-PCR marker based on a key mutation confirmed by resequencing of *Wx^{-mp}* in Milky Princess and its application in japonica soft rice (*Oryza sativa* L.) breeding. *Plant Breeding*, **132**(6),595-603. doi: [10.1111/pbr.12088](https://doi.org/10.1111/pbr.12088).
- Yano, M., Okuno,K., Satoh, H., Omura,T. 1998. Chromosomal location of genes conditioning low amylose content of endosperm starches in rice, *Oryza sativa* L. *Theor Appl Genet*. **76**(2), 183-189. doi:[10.1007/BF00257844](https://doi.org/10.1007/BF00257844). PMID: [24232104](https://pubmed.ncbi.nlm.nih.gov/24232104/).
- Yano, K., Yamamoto, E., Aya, K., Takeuchi, H., Lo, P.C., Hu, L., et al. 2016. Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat Genet*. **48**(8), 927-934. doi:[10.1038/ng.3596](https://doi.org/10.1038/ng.3596). PMID: [27322545](https://pubmed.ncbi.nlm.nih.gov/27322545/).
- Zeng, D., Yan, M., Wang, Y., Liu, X., Qian, Q., Li, J.(2007). *Du1*, encoding a novel Prp1 protein, regulates starch biosynthesis through affecting the splicing of *Wx^b* pre-mRNAs in rice (*Oryza sativa* L.). *Plant Mol Biol*, **65**(4), 501-509. doi:[10.1007/s11103-007-9186-3](https://doi.org/10.1007/s11103-007-9186-3). PMID:[17579813](https://pubmed.ncbi.nlm.nih.gov/17579813/).

- Zhang, G., Cheng, Z., Zhang, X., Guo, X., Su, N., Jiang, L., et al. 2011. Double repression of soluble starch synthase genes *SSIIa* and *SSIIIa* in rice (*Oryza sativa* L.) uncovers interactive effects on the physicochemical properties of starch. *Genome*, **54**(6),448-59. doi: [10.1139/g11-010](https://doi.org/10.1139/g11-010). PMID: [21595523](https://pubmed.ncbi.nlm.nih.gov/21595523/).
- Zhang, Z., Ersoz, E., Lai, C.Q., Todhunter, R.J., Tiwari, H.K., Gore, M.A., et al. 2010. Mixed linear model approach adapted for genome-wide association studies. *Nat Genet*, **42**(4),355-360. doi:[10.1038/ng.546](https://doi.org/10.1038/ng.546). PMID:[20208535](https://pubmed.ncbi.nlm.nih.gov/20208535/).

Draft

Fig.1. Histograms of average value of amylose content. (□) represents the amylose content from 0 to 2%, (□) 2-12%, (□) 12-20%, (□) 20-25%, (□) >25%.

Fig. 2. Chromosomal single-nucleotide polymorphisms (SNPs) distribution on chromosome.

Fig. 3. Group structure and clustering results. (a) The colours represent separate groups, with lines giving the group value. (b) The coefficient of variation (CV) value for each K-value.

Fig. 4. Genome-wide association studies of amylose content. (a) Manhattan plots of the general linear model for amylose content. Negative $\log_{10}$ -transformed P values from a genome-wide scan are plotted against position on 12 chromosomes. Blue horizontal dashed line indicates the genome-wide significance threshold. (b) Quantile-quantile plot of the general linear model for amylose content. (c) Mahattan plots of compresses mixed linear model for amylose content, as in (a). (d) Quantile-quantile plot of compresses mixed linear model for amylose content.

Fig.5. Relative expression of the 13 candidate genes and LOC_Os06g04200 on day 12 after flowering in kasalath, waxy and very low amyloscontent varieties by qRT-PCR. The x-axis indicates different amylose content types, CK= Kasalath, I = five very low amyloscontent varieties, II = five waxy varieties. The y-axis are scales of relative expression level. Error bars indicate standard deviations of independent biological replicates.

199x139mm (300 x 300 DPI)

250x187mm (300 x 300 DPI)

254x140mm (300 x 300 DPI)

561x381mm (300 x 300 DPI)

162x124mm (300 x 300 DPI)

Supplementary Table 1 Details of the primer sequences.

Gene ID	Primer Name	Sequense (5'to3 ')
LOC_Os06g10860	LOC_Os06g10860-qF	ACCGTAAATGCTGATCCTGTC
	LOC_Os06g10860-qR	CTCCCGATTGCCTTCTCC
LOC_Os06g10890	LOC_Os06g10890-qF	GCGAGGTTGGTAATAATAAG
	LOC_Os06g10890-qR	TCCAGCCATAACATCAGG
LOC_Os06g10960	LOC_Os06g10960-qF	CTGTTTCGTTTCGGATACTTC
	LOC_Os06g10960-qR	AGTTACTCGGTTGATTTCG
LOC_Os06g11010	LOC_Os06g11010-qF	GTTTACCTACTGTACCCGTCTTCA
	LOC_Os06g11010-qR	TATCTTCTCCGATCCGAGCC
LOC_Os06g11040	LOC_Os06g11040-qF	CAAACGCCGAAGCGACTCCT
	LOC_Os06g11040-qR	CACCCGAACACCACCACGAC
LOC_Os06g11130	LOC_Os06g11130-qF	CTTGTCGCTCGTCAGTTCAGG
	LOC_Os06g11130-qR	ATGTGCGTTGTGCTGGTGC
LOC_Os06g11140	LOC_Os06g11140-qF	GCCACTGCACAGGTAAGG
	LOC_Os06g11140-qR	CAGAAGAGGTCGGTCATT
LOC_Os06g11200	LOC_Os06g11200-qF	GCTCCCATCTCAAGCACC
	LOC_Os06g11200-qR	CTCAACACCGTCGAACCC
LOC_Os06g11210	LOC_Os06g11210-qF	GGGAGGAAGGCAACAAGG
	LOC_Os06g11210-qR	AAGCATCAGCAACCGAGT
LOC_Os06g11340	LOC_Os06g11340-qF	ACATCTGCTCGGATTAG
	LOC_Os06g11340-qR	CTTTGCGTAGTCCATTTA
LOC_Os06g11410	LOC_Os06g11410-qF	CACGCTCAGTTGACCTCAT
	LOC_Os06g11410-qR	TTGTAACAAGTTGCCACCC
LOC_Os06g11420	LOC_Os06g11420-qF	GGCAGACATCGGCGTTCAA
	LOC_Os06g11420-qR	TTCCAGTCGTCGTGGTCCTTC
LOC_Os06g11430	LOC_Os06g11430-qF	TCCCTTCTGTTTGGTTAA
	LOC_Os06g11430-qR	GCATTTGCTGGTTTGTAT
LOC_Os06g04200	LOC_Os06g04200-qF	TGAGGATGTTGTGTTCGTCTG
	LOC_Os06g04200-qR	CTGGTAGGAGATGTTGTGGATG

β -actin	β -qF	GAACTGGTATGGTCAAGGCTG
	β -qR	ACACGGAGCTCGTTGTAGAAG

Draft

Length (bp)	Product Size (bp)
21	246
18	
20	
18	109
19	
19	
24	122
20	
20	
20	158
20	
21	
19	270
18	
18	
18	154
18	
18	
18	160
18	
18	
18	159
18	
18	
18	216
18	
18	
18	141
18	
19	
19	249
19	
19	
21	187
18	
18	
18	124
21	
22	

21

250

21

Draft

Supplementary Table 2 The amylose content of 419 rice landraces

Serial number	Amylose content		Serial number	Amylose content	
	I (%)	II (%)		I (%)	II (%)
Line 1	24.9	25.4	Line 211	23.0	22.7
Line 2	23.8	24.5	Line 212	25.9	26.0
Line 3	21.3	21.2	Line 213	24.2	23.8
Line 4	25.3	25.4	Line 214	23.6	23.7
Line 5	21.9	22.3	Line 215	23.0	23.0
Line 6	23.8	24.9	Line 216	24.4	24.5
Line 7	26.8	25.8	Line 217	25.3	24.9
Line 8	22.7	22.9	Line 218	24.4	23.8
Line 9	26.9	25.9	Line 219	25.7	25.7
Line 10	24.6	24.5	Line 220	24.7	25.6
Line 11	20.8	20.9	Line 221	23.8	24.4
Line 12	23.8	22.9	Line 222	26.1	25.8
Line 13	23.5	22.9	Line 223	21.9	22.2
Line 14	23.7	23.4	Line 224	22.5	22.5
Line 15	22.4	21.5	Line 225	22.9	22.4
Line 16	24.0	24.1	Line 226	23.2	22.5
Line 17	20.9	20.0	Line 227	21.5	21.7
Line 18	19.5	18.7	Line 228	22.4	22.7
Line 19	17.8	18.1	Line 229	21.6	21.7
Line 20	1.8	2.0	Line 230	21.4	20.9
Line 21	1.5	1.5	Line 231	23.2	22.6
Line 22	24.4	24.0	Line 232	22.8	22.5
Line 23	0.2	0.7	Line 233	4.0	4.0
Line 24	0.0	0.4	Line 234	8.4	7.5
Line 25	1.5	1.5	Line 235	2.3	1.8
Line 26	0.1	0.1	Line 236	2.5	2.3
Line 27	22.8	22.0	Line 237	2.3	2.4
Line 28	1.5	1.5	Line 238	3.8	3.9
Line 29	20.5	20.4	Line 239	3.0	2.8
Line 30	22.9	21.9	Line 240	10.7	10.3
Line 31	1.5	1.7	Line 241	2.1	2.2
Line 32	0.3	0.0	Line 242	2.1	1.5
Line 33	24.3	24.6	Line 243	1.9	1.8
Line 34	0.0	0.2	Line 244	5.5	5.5
Line 35	23.5	24.5	Line 245	2.2	2.2
Line 36	23.9	24.8	Line 246	2.0	1.8
Line 37	2.0	2.0	Line 247	1.6	1.6
Line 38	25.2	24.8	Line 248	1.7	1.8
Line 39	23.3	23.2	Line 249	1.4	1.4
Line 40	0.6	0.0	Line 250	1.7	2.0
Line 41	22.3	22.7	Line 251	1.7	1.6
Line 42	1.1	0.6	Line 252	1.4	1.3
Line 43	0.0	0.6	Line 253	3.3	2.9
Line 44	23.3	24.2	Line 254	1.9	2.2
Line 45	1.8	1.8	Line 255	2.0	1.8

Line 46	22.7	22.2	Line 256	2.6	3.0
Line 47	1.1	1.1	Line 257	1.9	2.0
Line 48	1.9	1.8	Line 258	2.4	2.5
Line 49	23.8	24.8	Line 259	24.9	24.0
Line 50	25.1	26.0	Line 260	25.9	25.5
Line 51	1.3	1.4	Line 261	24.9	24.4
Line 52	2.0	2.1	Line 262	26.2	25.6
Line 53	3.9	4.3	Line 263	23.5	24.1
Line 54	14.3	14.5	Line 264	25.1	25.3
Line 55	2.3	2.0	Line 265	25.3	25.1
Line 56	11.4	10.4	Line 266	24.7	24.1
Line 57	23.7	24.2	Line 267	25.0	24.5
Line 58	23.0	23.9	Line 268	25.0	25.6
Line 59	0.4	0.3	Line 269	22.3	22.2
Line 60	1.5	1.6	Line 270	23.7	24.2
Line 61	2.8	2.8	Line 271	24.0	24.0
Line 62	1.5	1.8	Line 272	26.6	25.5
Line 63	23.1	22.8	Line 273	24.4	24.9
Line 64	23.1	22.9	Line 274	25.7	24.7
Line 65	24.4	23.9	Line 275	24.4	24.4
Line 66	23.0	23.2	Line 276	23.6	23.1
Line 67	22.8	23.0	Line 277	22.6	22.6
Line 68	24.6	24.0	Line 278	5.8	5.7
Line 69	22.3	23.1	Line 279	24.7	24.1
Line 70	23.2	22.4	Line 280	24.2	24.1
Line 71	23.2	23.6	Line 281	23.3	22.8
Line 72	24.2	23.5	Line 282	25.6	25.9
Line 73	20.1	20.3	Line 283	23.2	23.4
Line 74	21.0	21.1	Line 284	26.6	25.9
Line 75	24.6	24.7	Line 285	26.4	25.9
Line 76	23.2	23.0	Line 286	23.5	24.1
Line 77	22.7	21.7	Line 287	26.0	25.5
Line 78	20.9	21.5	Line 288	25.5	25.9
Line 79	22.4	22.3	Line 289	29.1	30.2
Line 80	22.9	21.9	Line 290	25.1	25.2
Line 81	21.3	20.3	Line 291	23.7	24.6
Line 82	25.0	24.8	Line 292	27.4	27.8
Line 83	20.5	20.7	Line 293	26.1	25.8
Line 84	24.2	23.4	Line 294	24.5	25.3
Line 85	4.8	4.9	Line 295	26.1	25.4
Line 86	22.6	21.6	Line 296	23.8	24.3
Line 87	22.0	21.9	Line 297	25.7	25.9
Line 88	25.6	24.5	Line 298	23.6	23.9
Line 89	23.5	23.3	Line 299	24.9	23.9
Line 90	23.1	22.8	Line 300	23.7	23.8
Line 91	23.7	22.7	Line 301	24.7	24.4
Line 92	23.3	22.5	Line 302	24.5	24.6
Line 93	3.6	3.4	Line 303	26.3	26.1
Line 94	24.7	23.8	Line 304	25.8	24.7
Line 95	22.6	22.6	Line 305	13.0	12.7

Line 96	23.1	22.4	Line 306	15.7	15.2
Line 97	22.8	23.5	Line 307	12.3	11.7
Line 98	22.8	22.1	Line 308	14.9	13.8
Line 99	23.8	22.8	Line 309	13.8	13.7
Line 100	14.6	14.2	Line 310	16.6	17.2
Line 101	22.3	22.5	Line 311	15.7	15.3
Line 102	22.8	23.7	Line 312	13.0	13.5
Line 103	23.7	22.7	Line 313	15.0	14.2
Line 104	24.6	24.4	Line 314	15.8	14.9
Line 105	22.3	21.2	Line 315	14.5	14.7
Line 106	24.1	25.1	Line 316	13.7	13.2
Line 107	22.2	22.0	Line 317	19.4	20.3
Line 108	22.2	22.5	Line 318	24.8	24.6
Line 109	24.3	23.5	Line 319	25.3	25.9
Line 110	23.8	23.3	Line 320	22.9	21.9
Line 111	26.3	25.5	Line 321	24.7	24.4
Line 112	21.2	20.8	Line 322	26.0	25.5
Line 113	23.7	22.9	Line 323	25.1	24.5
Line 114	23.0	22.6	Line 324	25.1	25.5
Line 115	24.8	24.6	Line 325	25.2	25.8
Line 116	22.0	22.9	Line 326	22.4	23.4
Line 117	22.8	22.7	Line 327	23.5	24.4
Line 118	24.2	23.7	Line 328	20.7	20.5
Line 119	22.6	21.7	Line 329	25.2	25.1
Line 120	22.6	23.5	Line 330	17.6	16.8
Line 121	24.6	25.4	Line 331	21.8	22.7
Line 122	23.2	22.8	Line 332	23.3	23.1
Line 123	3.1	3.2	Line 333	22.1	22.8
Line 124	22.6	22.4	Line 334	25.1	25.8
Line 125	24.7	23.6	Line 335	24.2	23.8
Line 126	24.0	25.0	Line 336	25.4	25.6
Line 127	24.6	25.6	Line 337	23.7	24.5
Line 128	19.8	19.2	Line 338	23.7	23.2
Line 129	23.3	23.8	Line 339	22.7	22.7
Line 130	25.1	25.8	Line 340	25.6	26.4
Line 131	25.7	26.2	Line 341	22.1	21.1
Line 132	2.2	2.6	Line 342	25.2	25.3
Line 133	2.5	2.5	Line 343	23.5	23.2
Line 134	2.3	2.4	Line 344	25.0	24.0
Line 135	23.8	24.1	Line 345	27.5	26.6
Line 136	2.1	2.4	Line 346	23.8	22.7
Line 137	23.4	23.2	Line 347	2.0	2.2
Line 138	3.0	2.8	Line 348	3.1	3.0
Line 139	7.2	7.3	Line 349	3.4	3.4
Line 140	2.8	2.7	Line 350	2.1	2.5
Line 141	1.2	1.2	Line 351	1.9	2.9
Line 142	14.3	14.2	Line 352	19.1	19.0
Line 143	25.9	24.9	Line 353	2.6	3.1
Line 144	25.8	26.4	Line 354	2.0	2.2
Line 145	24.6	24.4	Line 355	2.0	2.4

Line 146	25.2	24.2	Line 356	2.1	2.6
Line 147	21.1	21.1	Line 357	2.6	1.8
Line 148	25.8	26.0	Line 358	7.0	8.0
Line 149	19.2	19.3	Line 359	2.3	2.2
Line 150	22.6	23.4	Line 360	2.1	2.2
Line 151	24.5	24.2	Line 361	2.1	2.1
Line 152	25.7	26.7	Line 362	2.3	2.2
Line 153	25.0	25.1	Line 363	23.9	23.8
Line 154	23.5	23.7	Line 364	24.5	24.0
Line 155	24.6	23.7	Line 365	25.7	25.2
Line 156	26.8	26.9	Line 366	23.3	22.2
Line 157	25.7	25.2	Line 367	25.1	25.3
Line 158	24.7	25.4	Line 368	25.0	25.1
Line 159	21.9	22.5	Line 369	25.1	25.0
Line 160	25.2	24.9	Line 370	26.8	25.6
Line 161	23.8	23.5	Line 371	26.7	26.6
Line 162	21.8	22.0	Line 372	25.8	24.9
Line 163	24.1	25.1	Line 373	20.7	20.4
Line 164	23.5	23.3	Line 374	25.3	25.4
Line 165	22.5	22.9	Line 375	26.1	26.8
Line 166	24.8	25.6	Line 376	25.8	25.8
Line 167	23.5	23.6	Line 377	25.1	24.4
Line 168	26.9	27.9	Line 378	26.9	26.4
Line 169	25.4	25.4	Line 379	21.6	21.0
Line 170	26.7	26.1	Line 380	23.8	23.8
Line 171	24.7	23.7	Line 381	25.5	25.9
Line 172	26.9	26.3	Line 382	25.3	25.7
Line 173	24.4	25.0	Line 383	23.5	23.7
Line 174	25.6	25.2	Line 384	25.0	24.5
Line 175	25.0	24.1	Line 385	26.0	24.9
Line 176	23.2	22.3	Line 386	23.8	22.9
Line 177	26.5	25.9	Line 387	23.9	23.2
Line 178	23.8	23.2	Line 388	26.3	26.0
Line 179	23.9	23.9	Line 389	24.9	24.2
Line 180	21.3	22.2	Line 390	14.9	14.8
Line 181	22.4	22.4	Line 391	11.6	11.2
Line 182	12.3	12.3	Line 392	24.4	23.3
Line 183	2.1	1.6	Line 393	25.0	24.5
Line 184	1.8	1.5	Line 394	23.1	23.3
Line 185	24.0	24.3	Line 395	23.3	24.0
Line 186	23.1	23.5	Line 396	26.1	26.4
Line 187	3.2	2.6	Line 397	23.6	23.6
Line 188	23.8	23.6	Line 398	24.0	24.2
Line 189	1.7	1.8	Line 399	22.0	21.4
Line 190	2.5	2.4	Line 400	14.4	15.1
Line 191	4.1	3.2	Line 401	25.1	25.5
Line 192	2.7	2.8	Line 402	25.5	24.4
Line 193	21.1	21.9	Line 403	25.2	24.3
Line 194	24.7	24.3	Line 404	3.7	3.9
Line 195	1.8	2.2	Line 405	20.7	21.3

Line 196	1.7	1.8	Line 406	23.1	22.7
Line 197	14.1	14.1	Line 407	3.6	3.7
Line 198	1.5	1.5	Line 408	23.4	22.7
Line 199	1.9	1.8	Line 409	23.0	22.3
Line 200	12.6	12.9	Line 410	16.8	17.1
Line 201	24.1	23.7	Line 411	2.4	2.4
Line 202	1.3	1.2	Line 412	6.7	6.8
Line 203	11.6	11.0	Line 413	2.2	2.4
Line 204	11.3	10.9	Line 414	2.3	2.3
Line 205	21.0	22.0	Line 415	2.1	1.3
Line 206	21.7	21.4	Line 416	1.4	2.0
Line 207	18.9	18.5	Line 417	1.2	1.3
Line 208	17.2	17.3	Line 418	1.4	1.5
Line 209	14.6	14.5	Line 419	10.6	9.8
Line 210	20.0	19.5			

Draft

Supplementary Table 3 The summary of significantly associated signals from the general linear model.

Chr.	Position (bp)	P-value
2	24736366	3.82E-08
2	32816318	1.16E-12
2	32816330	1.40E-12
2	32816334	3.83E-11
2	32816559	1.88E-10
2	32816564	2.04E-08
3	16872869	3.96E-08
3	16873039	7.56E-10
3	17005327	3.05E-08
3	17189483	3.62E-08
3	17189700	3.62E-08
3	17189743	1.06E-09
3	17189763	5.55E-10
3	17189994	1.19E-10
3	17288134	1.97E-09
3	17288320	1.22E-10
3	17288321	1.03E-10
3	17308018	3.41E-08
3	17934255	7.30E-09
3	19027102	1.75E-09
3	19027120	8.57E-10
3	19923950	2.19E-08
3	23485149	1.16E-08
6	545785	1.85E-08
6	876324	9.81E-13
6	909090	4.27E-12
6	921039	2.47E-29
6	945908	3.89E-16
6	945924	3.89E-16
6	1179280	3.04E-12
6	1211363	6.11E-15
6	1211392	9.52E-16
6	1457896	3.68E-16

6	1460355	4.12E-08
6	1512757	7.35E-15
6	1512819	3.20E-15
6	1577337	5.72E-10
6	1592445	1.17E-27
6	1597045	2.20E-20
6	1607061	1.16E-66
6	1654904	3.81E-23
6	1661801	1.26E-14
6	1661829	2.10E-72
6	1684538	1.42E-24
6	1684543	1.81E-45
6	1684784	9.71E-33
6	1694761	1.16E-53
6	1694808	8.26E-52
6	1694818	2.01E-36
6	1703886	3.72E-70
6	1720764	2.87E-43
6	1720799	7.31E-70
6	1720821	9.84E-69
6	1745333	1.81E-26
6	1745339	5.70E-31
6	1754453	6.98E-66
6	1756202	2.99E-67
6	1756454	1.67E-71
6	1780293	2.19E-70
6	1780336	5.32E-72
6	1780519	3.02E-22
6	1792217	7.92E-80
6	1797551	4.51E-86
6	1807558	1.23E-38
6	1807737	2.52E-34
6	1807778	3.99E-42
6	1807797	2.00E-62
6	1822395	8.92E-63
6	1865861	5.23E-21
6	1958767	7.42E-22

6	2001358	2.02E-08
6	2001647	1.39E-08
6	2125248	5.37E-11
6	2256752	4.30E-20
6	5600598	2.67E-13
6	5600742	4.80E-14
6	5603254	7.85E-14
6	5603255	2.26E-15
6	5603256	2.50E-17
6	5603257	5.89E-16
6	5603263	7.46E-16
6	5661860	3.22E-10
6	5661861	1.51E-13
6	5662030	4.95E-13
6	5688919	4.74E-15
6	5724272	9.35E-14
6	5724307	1.07E-12
6	5759063	3.42E-17
6	5767465	2.97E-10
6	5767501	1.08E-10
6	5783531	1.61E-15
6	5843510	6.58E-22
6	5876935	2.44E-08
6	5877736	3.48E-15
6	5877743	7.26E-16
6	5877851	6.41E-16
6	5878166	1.12E-13
6	5878223	1.06E-12
6	5881943	1.10E-15
6	5881999	4.29E-18
6	5921101	1.94E-08
6	5928250	3.38E-10
6	5947329	1.28E-08
6	5947335	5.78E-09
6	5951179	1.47E-17
6	6017843	1.42E-16
6	6017859	1.50E-16

6	6018062	2.70E-17
6	6024319	3.83E-12
6	6035183	2.40E-20
6	6035194	9.54E-21
6	6063216	1.74E-08
6	6063260	1.05E-08
6	6086168	4.00E-17
6	6086201	4.46E-17
6	6086353	5.39E-15
6	6086354	1.91E-16
6	6098826	5.13E-17
6	6100096	8.34E-09
6	6100114	8.34E-09
6	6123158	1.48E-12
6	6152410	4.76E-12
6	6152443	3.72E-12
6	6188721	4.73E-20
6	6300137	2.40E-08
6	6465099	4.80E-09
6	6465228	2.06E-08
6	6465263	3.40E-08
6	6465280	2.06E-08
6	6465288	2.06E-08
6	6473671	5.91E-13
6	6475507	2.93E-12
6	6488717	3.48E-12
6	6491984	7.44E-10
6	6491993	4.72E-16
6	6538671	1.49E-13
6	6661283	2.74E-08
6	6667180	1.53E-15
6	6667418	1.16E-15
6	6693952	5.30E-09
6	6733142	9.62E-27
6	6733351	1.66E-27
6	6740370	4.35E-14
6	6767159	4.85E-18

6	6814292	2.49E-09
6	6857054	3.24E-13
6	6879482	3.38E-09
6	6879525	5.89E-17
6	6879531	1.52E-23
6	6879731	1.77E-23
6	6886001	8.70E-20
6	6919428	2.71E-16
6	6924742	5.63E-14
6	6926541	3.36E-14
6	6927719	1.11E-09
6	6927926	4.64E-14
6	6927931	2.77E-14
6	6938650	1.11E-08
6	6953368	3.42E-16
6	6957926	1.43E-12
6	6966838	6.91E-18
6	6967051	4.52E-16
6	6972778	4.36E-09
6	6972790	4.36E-09
6	7029657	8.13E-17
6	7075801	1.59E-12
6	7077571	1.21E-18
6	7077594	1.76E-19
6	7077599	1.76E-19
6	7080011	5.46E-21
6	7086118	7.76E-22
6	7086166	7.76E-22
6	7086184	7.76E-22
6	7088738	2.32E-17
6	7088886	2.46E-17
6	7098463	1.92E-20
6	7103272	5.25E-13
6	7607179	4.30E-08
7	22739134	7.64E-09
7	22739286	4.19E-09
7	22739336	1.07E-08

8	11459294	4.47E-08
8	19397003	7.13E-09
8	19397043	7.66E-09
8	22683080	8.30E-10
12	22040968	4.70E-08
12	22041422	9.07E-10
12	22092970	1.17E-09
12	22174462	5.84E-10
12	22174476	5.84E-10
12	22193984	5.92E-11
12	22194004	7.10E-11
12	22205922	2.21E-08
12	22205940	2.97E-08
12	22205944	3.24E-08
12	22213174	1.03E-08
12	22213211	1.16E-08
12	22213226	1.11E-08
12	22568586	3.42E-09

Supplementary Table 4 The summary of significantly associated signals from the compressed mixed linear model.

Chr.	Position (bp)	P-value
6	909090	1.54E-08
6	921039	4.97E-18
6	945908	2.17E-12
6	945924	2.17E-12
6	1179280	8.31E-10
6	1211363	4.67E-10
6	1211392	1.14E-10
6	1457896	1.73E-10
6	1512757	5.74E-10
6	1512819	2.61E-10
6	1592445	5.26E-18
6	1597045	2.27E-14
6	1607061	3.64E-30
6	1654904	6.05E-15
6	1661801	3.79E-10
6	1661829	5.57E-32
6	1684538	7.04E-14
6	1684543	3.51E-22
6	1684784	1.06E-17
6	1694761	2.03E-26
6	1694808	3.47E-25
6	1694818	1.25E-19
6	1703886	3.02E-32
6	1720764	2.26E-23
6	1720799	1.08E-32
6	1720821	2.38E-32
6	1745333	2.77E-14
6	1745339	7.50E-16
6	1754453	1.30E-30
6	1756202	1.04E-30
6	1756454	5.47E-32
6	1780293	7.23E-31
6	1780336	3.02E-31
6	1780519	6.79E-14

6	1792217	2.21E-33
6	1797551	3.97E-35
6	1807558	7.55E-20
6	1807737	8.59E-19
6	1807778	2.99E-21
6	1807797	8.26E-30
6	1822395	1.80E-28
6	1865861	8.27E-13
6	1958767	9.91E-16
6	2001358	4.71E-08
6	2001647	3.16E-08
6	2125248	4.41E-10
6	2256752	8.70E-14
6	5600742	3.45E-08
6	5603255	8.93E-09
6	5603256	2.06E-10
6	5603257	2.89E-09
6	5603263	1.99E-09
6	5661861	3.13E-09
6	5662030	6.72E-09
6	5688919	1.00E-09
6	5724272	9.00E-09
6	5759063	4.83E-10
6	5783531	1.00E-09
6	5835572	2.46E-12
6	5843510	3.77E-13
6	5877736	7.02E-10
6	5877743	3.62E-10
6	5877851	2.00E-10
6	5878166	2.24E-08
6	5881943	2.83E-10
6	5881999	1.03E-11
6	5951179	5.53E-12
6	6017843	1.07E-10
6	6017859	1.30E-10
6	6018062	6.24E-11
6	6024319	6.80E-09

6	6035183	9.73E-12
6	6035194	6.36E-12
6	6086168	5.04E-09
6	6086201	7.75E-09
6	6086354	2.49E-08
6	6098826	2.69E-10
6	6123158	9.98E-09
6	6152410	1.51E-08
6	6152443	1.44E-08
6	6188721	4.56E-11
6	6473671	9.10E-09
6	6475507	7.48E-09
6	6491993	8.89E-09
6	6538671	2.73E-08
6	6667180	7.93E-10
6	6667418	9.74E-10
6	6733142	9.54E-15
6	6733351	9.67E-15
6	6740370	3.72E-09
6	6767159	1.28E-10
6	6857054	8.68E-09
6	6879525	3.47E-10
6	6879531	2.94E-14
6	6879731	3.19E-14
6	6886001	2.46E-12
6	6919428	2.47E-10
6	6924742	2.99E-09
6	6926541	2.93E-09
6	6927926	1.53E-08
6	6927931	1.16E-08
6	6953368	4.15E-10
6	6966838	3.88E-10
6	6967051	4.16E-09
6	7029657	1.34E-09
6	7077571	2.32E-09
6	7077594	6.10E-10
6	7077599	6.10E-10

6	7080011	9.81E-11
6	7086118	1.67E-11
6	7086166	1.67E-11
6	7086184	1.67E-11
6	7088738	1.83E-08
6	7088886	2.00E-08
6	7098463	1.14E-11
6	7103272	1.20E-08

Draft

Supplementary Table 5 The information 116 significantly associated SNPs.

osome	Position	P-value	Marker_R2	Majorallele	Num
6	909090	1.54E-08	0.1305	C	462
6	921039	4.97E-18	0.2663	A	455
6	945908	2.17E-12	0.1706	G	455
6	945924	2.17E-12	0.1706	G	455
6	1179280	8.31E-10	0.1353	C	516
6	1211363	4.67E-10	0.1151	T	598
6	1211392	1.14E-10	0.1216	A	599
6	1457896	1.73E-10	0.1123	C	465
6	1512757	5.74E-10	0.1452	G	381
6	1512819	2.61E-10	0.1500	A	384
6	1592445	5.26E-18	0.2041	G	593
6	1597045	2.27E-14	0.1694	G	470
6	1607061	3.64E-30	0.3912	T	559
6	1654904	6.05E-15	0.2856	G	297
6	1661801	3.79E-10	0.1080	A	472
6	1661829	5.57E-32	0.4077	A	587
6	1684538	7.04E-14	0.2333	A	407
6	1684543	3.51E-22	0.4083	C	444
6	1684784	1.06E-17	0.3075	T	431
6	1694761	2.03E-26	0.3759	G	481
6	1694808	3.47E-25	0.3548	A	482
6	1694818	1.25E-19	0.2644	T	455
6	1703886	3.02E-32	0.4007	T	574
6	1720764	2.26E-23	0.2846	A	533
6	1720799	1.08E-32	0.4256	G	561
6	1720821	2.38E-32	0.4209	C	560
6	1745333	2.77E-14	0.3042	A	334
6	1745339	7.50E-16	0.3451	C	338
6	1754453	1.30E-30	0.3904	A	561
6	1756202	1.04E-30	0.3855	T	552
6	1756454	5.47E-32	0.4049	A	556
6	1780293	7.23E-31	0.4442	T	482
6	1780336	3.02E-31	0.4507	C	483
6	1780519	6.79E-14	0.1730	C	479
6	1792217	2.21E-33	0.4574	C	557
6	1797551	3.97E-35	0.4629	C	571
6	1807558	7.55E-20	0.4135	A	295
6	1807737	8.59E-19	0.3854	T	272
6	1807778	2.99E-21	0.4501	G	279
6	1807797	8.26E-30	0.3912	A	534
6	1822395	1.80E-28	0.4121	T	525
6	1865861	8.27E-13	0.1842	G	543
6	1958767	9.91E-16	0.2133	T	405
6	2001358	4.71E-08	0.0822	G	505
6	2001647	3.16E-08	0.0845	T	503
6	2125248	4.41E-10	0.1085	G	515

6	2256752	8.70E-14	0.2001	G	578
6	5600742	3.45E-08	0.0839	A	689
6	5603255	8.93E-09	0.0895	A	681
6	5603256	2.06E-10	0.1088	A	685
6	5603257	2.89E-09	0.0953	A	704
6	5603263	1.99E-09	0.0972	G	683
6	5661861	3.13E-09	0.1110	G	588
6	5662030	6.72E-09	0.1059	G	587
6	5688919	1.00E-09	0.1049	T	581
6	5724272	9.00E-09	0.0892	C	783
6	5759063	4.83E-10	0.1215	A	717
6	5783531	1.00E-09	0.1089	C	727
6	5835572	2.46E-12	0.1726	C	728
6	5843510	3.77E-13	0.1892	G	763
6	5877736	7.02E-10	0.1140	C	687
6	5877743	3.62E-10	0.1193	T	681
6	5877851	2.00E-10	0.1146	G	734
6	5878166	2.24E-08	0.0899	C	757
6	5881943	2.83E-10	0.1071	T	780
6	5881999	1.03E-11	0.1243	T	784
6	5951179	5.53E-12	0.1262	A	702
6	6017843	1.07E-10	0.1122	G	769
6	6017859	1.30E-10	0.1112	G	768
6	6018062	6.24E-11	0.1149	A	770
6	6024319	6.80E-09	0.1006	T	580
6	6035183	9.73E-12	0.1289	G	514
6	6035194	6.36E-12	0.1312	T	513
6	6086168	5.04E-09	0.1900	C	508
6	6086201	7.75E-09	0.1852	T	507
6	6086354	2.49E-08	0.1800	A	492
6	6098826	2.69E-10	0.1091	G	757
6	6123158	9.98E-09	0.0935	C	501
6	6152410	1.51E-08	0.1082	A	371
6	6152443	1.44E-08	0.1082	C	373
6	6188721	4.56E-11	0.2186	G	500
6	6473671	9.10E-09	0.1084	G	404
6	6475507	7.48E-09	0.1565	T	264
6	6491993	8.89E-09	0.0954	C	489
6	6538671	2.73E-08	0.1079	T	434
6	6667180	7.93E-10	0.1144	C	427
6	6667418	9.74E-10	0.1143	A	424
6	6733142	9.54E-15	0.3088	T	279
6	6733351	9.67E-15	0.3199	C	275
6	6740370	3.72E-09	0.1141	C	455
6	6767159	1.28E-10	0.1124	C	482
6	6857054	8.68E-09	0.1031	G	403
6	6879525	3.47E-10	0.1223	G	575
6	6879531	2.94E-14	0.1789	A	488
6	6879731	3.19E-14	0.1783	C	486
6	6886001	2.46E-12	0.1372	G	499

6	6919428	2.47E-10	0.1261	C	400
6	6924742	2.99E-09	0.1148	T	360
6	6926541	2.93E-09	0.0955	G	755
6	6927926	1.53E-08	0.1216	A	354
6	6927931	1.16E-08	0.1237	T	354
6	6953368	4.15E-10	0.1068	G	605
6	6966838	3.88E-10	0.1192	G	429
6	6967051	4.16E-09	0.1055	A	428
6	7029657	1.34E-09	0.1096	G	418
6	7077571	2.32E-09	0.1394	G	441
6	7077594	6.10E-10	0.1514	T	443
6	7077599	6.10E-10	0.1514	T	443
6	7080011	9.81E-11	0.1507	A	498
6	7086118	1.67E-11	0.1553	C	518
6	7086166	1.67E-11	0.1553	T	518
6	7086184	1.67E-11	0.1553	A	518
6	7088738	1.83E-08	0.1707	T	287
6	7088886	2.00E-08	0.1703	T	285
6	7098463	1.14E-11	0.1366	G	498
6	7103272	1.20E-08	0.0944	C	392

Draft

Minorallele	Num	Known loci
T	80	
T	75	
T	73	
A	73	
A	32	
C	142	
T	145	
A	339	
C	197	
G	198	
C	231	
A	236	
A	247	
T	161	
G	356	
C	241	
G	171	
T	136	
C	155	
C	231	
G	230	
C	251	
G	264	
G	269	
A	241	
T	240	
G	138	
T	134	
T	253	
C	264	
G	262	
C	248	
A	249	
T	251	
T	241	
T	253	Wx
G	159	
C	166	
T	159	
G	250	
C	219	
A	137	
C	211	
T	323	
C	317	
A	293	

Draft

A	34	
G	143	
T	149	
T	145	
T	126	
A	147	
A	38	
A	37	
G	211	
T	55	
T	45	
G	43	
A	45	
A	49	
T	49	
G	51	
A	46	
T	47	
C	52	
C	48	
G	40	
A	51	
A	52	
T	50	
C	38	
C	300	
G	301	
T	48	
C	49	
C	50	
T	51	
T	249	
G	231	
T	233	
T	58	
A	258	
C	186	
T	295	
A	244	
A	325	
G	324	
C	157	
T	155	<i>SSIIa</i>
A	263	
G	348	
A	321	
A	183	
G	272	
G	274	
A	321	

T	292
C	284
T	49
G	258
G	256
C	207
A	327
G	328
A	312
A	209
C	201
A	201
G	214
T	218
C	218
G	218
C	181
C	179
T	264
T	326

Draft

Supplementary Table 6 Detailed information on the DNA polymorphisms in the ca
Group I (1 polymorphism in 1 gene)

Chr.	Position (bp)	P-value	Majorallele	Minorallele	Gene ID
6	5688919	1.00E-09	T	G	LOC_Os06g10890
6	5835572	2.4596E-12	C	A	LOC_Os06g11130

Group II (15 polymorphisms in 8 genes)

Chr.	Position (bp)	P-value	Majorallele	Minorallele	Gene ID
6	5661861	3.1289E-09	G	A	LOC_Os06g10860
6	5662030	6.7235E-09	G	A	LOC_Os06g10860
6	5759063	4.825E-10	A	T	LOC_Os06g11010
6	5783531	1.0047E-09	C	G	LOC_Os06g11040
6	5877736	7.0246E-10	C	T	LOC_Os06g11200
6	5877743	3.6234E-10	T	G	LOC_Os06g11200
6	5877851	1.9971E-10	G	A	LOC_Os06g11200
6	5878166	2.2382E-08	C	T	LOC_Os06g11200
6	5881943	2.8309E-10	T	C	LOC_Os06g11210
6	5881999	1.0299E-11	T	C	LOC_Os06g11210
6	5951179	5.5295E-12	A	G	LOC_Os06g11340
6	6017843	1.0715E-10	G	A	LOC_Os06g11410
6	6017859	1.303E-10	G	A	LOC_Os06g11410
6	6018062	6.238E-11	A	T	LOC_Os06g11410
6	6024319	6.7966E-09	T	C	LOC_Os06g11420

Group III (2 polymorphisms in 1 gene)

Chr.	Position (bp)	P-value	Majorallele	Minorallele	Gene ID
6	6035183	9.7265E-12	G	C	LOC_Os06g11430
6	6035194	6.3571E-12	T	G	LOC_Os06g11430

Group IV (2 polymorphisms in 2 genes)

Chr.	Position (bp)	P-value	Majorallele	Minorallele	Gene ID
6	5724272	8.9988E-09	C	T	LOC_Os06g10960
6	5843510	3.7661E-13	G	A	LOC_Os06g11140

indicate region.

Region	Annotation
CDS	sterol carrier protein-2, putative, expressed
CDS	gibberellin receptor GID1L2, putative, expressed

Region	Annotation
promoter region	glucosyltransferase, putative, expressed
promoter region	glucosyltransferase, putative, expressed
promoter region	eukaryotic aspartyl protease domain containing protein, expressed
promoter region	expressed protein
promoter region	12-oxophytodienoate reductase, putative, expressed
promoter region	12-oxophytodienoate reductase, putative, expressed
promoter region	12-oxophytodienoate reductase, putative, expressed
promoter region	12-oxophytodienoate reductase, putative, expressed
promoter region	12-oxophytodienoate reductase, putative, expressed
promoter region	12-oxophytodienoate reductase, putative, expressed
promoter region	retrotransposon protein, putative, unclassified
promoter region	cyclin, putative, expressed
promoter region	cyclin, putative, expressed
promoter region	cyclin, putative, expressed
promoter region	SKIP/SNW domain containing protein, expressed

Region	Annotation
3'UTR	GRAM domain containing protein, expressed
3'UTR	GRAM domain containing protein, expressed

Region	Annotation
INTRON	xyloglucan fucosyltransferase, putative, expressed
INTRON	pyridine nucleotide-disulphide oxidoreductase domain containing protein, expressed

Draft

sscc