Theoretical and Applied Genetics (2025) 138:58
https://doi.org/10.1007/s00122-025-04850-x

ORIGINAL ARTICLE

QTN detection and candidate gene identification for improved eating and cooking quality in rice using GWAS and PLS regression analysis

Kiyosumi Hori[1,2] · Matthew Shenton[1] · Kenta Mochizuki[1,3] · Keitaro Suzuki[1] · Ken Iijima[1] · Noriyuki Kuya[1] · Koka Shu[1] · Kosuke Ono[2] · Yuji Kinoshita[3] · Kazuhiko Sugimoto[1] · Takayuki Umemoto[1] · Jun-ichi Yonemaru[1] · Masanori Yamasaki[4,5] · Yoshinobu Takeuchi[1] · Kaworu Ebana[1] · Yoshimasa Tsujii[3]

Received: 19 October 2024 / Accepted: 7 February 2025 / Published online: 26 February 2025

Abstract

Key message **We performed GWAS for starch properties and eating, cooking and appearance quality characteristics traits in rice and then used PLS regression to show importance of different loci for different food applications.**

Abstract We performed a genome-wide association study for appearance, eating and cooking quality traits in grain of *japonica* rice cultivars and identified candidate genes for adhesiveness of cooked rice grains, amylopectin composition and β-glucanase activity in rice endosperm among a total of 525 quantitative trait nucleotide (QTN) loci. The study used 1,054,635 single-nucleotide polymorphisms (SNPs) based on genome sequence data of 150 rice cultivars and 89 grain appearance, eating and cooking quality traits. These included grain shape, protein content, amylose content, amylopectin chain length, starch viscosity properties, starch degradation enzyme activities, and physicochemical characteristics of cooked rice grains analyzed in three years. Cluster regions of genetic loci on rice chromosomes 1, 4, 5, 6, 8, 9, 10 and 11 were detected, with several regions co-located with starch biosynthesis and degradation genes. Partial least squares (PLS) regression analysis revealed that the QTN genotypes were unevenly distributed in subpopulations of rice cultivars classified by their primary application. We could therefore select and accumulate these QTNs to improve grain quality in further breeding programs by developing novel rice cultivars with appropriate phenotypes for each food usage: as high eating quality cooked rice, staple food rice, sushi rice, sake brewing rice, and high-yielding rice cultivars.

Communicated by Jiankang Wang.

Kiyosumi Hori, Matthew Shenton, and Kenta Mochizuki have contributed equally to this work.

✉ Kiyosumi Hori
horikiyo@affrc.go.jp

✉ Yoshimasa Tsujii
96tsujii@nodai.ac.jp

1 National Agriculture and Food Research Organization (NARO), Tsukuba, Ibaraki, Japan

2 Department of Integrated Biosciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba, Japan

3 Department of Agricultural Chemistry, Faculty of Applied Bioscience, Tokyo University of Agriculture, Setagaya, Tokyo, Japan

4 Food Resources Education and Research Center, Kobe University, Kasai, Hyogo, Japan

5 Present Address: Faculty of Agriculture, Niigata University, Niigata, Niigata, Japan

Introduction

Rice (*Oryza sativa* L.) is one of the most important food crops, a staple food for over half of the world's population. World rice production exceeds 500 million metric tons, although rice consumption per capita is decreasing steadily in several countries due to diversification of dietary habits (Shahbandeh 2024). Recent economic growth in Asian countries including China, Korea, and Japan has led consumers to demand high grain quality and good eating quality in rice (Champagne et al. 1999; Fitzgerald et al. 2009; Hori and Sun 2022; Li et al. 2022). At the same time, increasing population pressure accompanying economic growth has led to decrease in the area of agricultural land and fewer farmers (UNICEF 2024), with climate change also predicted to have a severe impact on agricultural production (Kole 2020). Therefore, it is important to develop novel rice cultivars with the high grain quality and good eating quality demanded by consumers, in addition to high grain yield under variable cultivation conditions.

Consumer preferences for rice grains vary widely by region (Calingacion et al. 2014; Custodio et al. 2019; Sreenivasulu et al. 2022). For example, consumers in parts of South-East and South-West Asia prefer long and slender grains of *indica* rice cultivars. Within those regions, firm and separate cooked rice grains are preferred in Malaysia, Philippines, India, Iran, and Pakistan, while consumers in Thailand and Cambodia prefer soft and very sticky cooked grains. *Japonica* rice grains are generally short in size and sticky when cooked; *japonica* cultivars are predominantly grown in Japan, Korea, some regions of China, Southern Europe, Australia, and California (USA) (Hori and Yano 2013; Park et al. 2019). The cultivation area of *japonica* rice cultivars in China has expanded in recent decades, growing from 11% in 1980 to 29% of the total rice cultivation area in 2000. Consumer preferences and demands have resulted in increased *japonica* rice production in both northern and southern China (Meng et al. 2022). Increasingly, health-conscious consumers worldwide favor rice grains and rice flours because they can be used in gluten-free foods. Several rice cultivars have been developed and released for use in making rice flours, breads, and noodles (Nakanishi et al. 2022; Farooq et al. 2021; Hori et al. 2022). Maintaining or increasing yields of high-quality *japonica* rice during climate change depends on an understanding of the genetic basis of cooking and eating quality traits.

Previous studies have reported some of the gene loci underlying eating and cooking quality traits in rice. QTL (Quantitative Trait Loci) mapping has shown that a few major genes control most of the starch physicochemical properties, such as amylose content, gelatinization temperature, and viscosity. The *GBSSI* (Granule-bound starch synthase I; *Waxy* or *Wx*) gene has been detected as a major effect QTL for amylose content, gel consistency and starch viscosity characteristics (Wang et al. 1995; Bao et al. 2000; Fan et al. 2005; Deng et al. 2022). The functional alleles Wx^a and Wx^b are associated with high (22–29%) and low (12–19%) amylose content, respectively, and rice cultivars with the non-functional allele, *wx*, are almost completely lacking (close to 0%) in amylose (glutinous rice). The *SSIIa* (Soluble starch synthase IIa; *Alk*) gene was detected as another major effect QTL for gelatinization temperature and gel consistency (Tian et al. 2009, Umemoto et al. 2002, Wang et al. 2007). Although other starch biosynthesis genes including *SSIIIa* (Soluble starch synthase IIIa), *SSI* (Starch synthase I), and *SBEI* (Starch branching enzyme I) have also been identified as minor effect QTLs, previous genetic studies have identified only a few genes controlling eating and cooking quality traits (Li et al. 2022; Ren et al. 2023). A series of eating quality QTLs including *qOE3* (Wada et al. 2015; Hori et al. 2021) that affect the stickiness or surface adhesiveness of cooked rice grains have been identified. The segregating populations were derived from crosses between Japanese rice cultivars and the non-glutinous *temperate japonica* cultivar Koshihikari (Kobayashi and Tomita 2008; Takeuchi et al. 2008; Wada et al. 2008). Koshihikari (Koshi) has been a top rice cultivar in Japan for more than 40 years, and its grains possess superior eating quality characteristics. The Koshi *qOE3* region on chromosome 3 is a major effect QTL for eating quality and scores highly for stickiness and surface adhesiveness of cooked rice grains, making this QTL allele the most important genetic factor controlling eating quality in Japanese rice cultivars to date (Hori and Yano 2013). However, the causal gene has not been reported.

Advances in genome sequencing technology have made it possible to obtain whole genome sequence data from large numbers of rice cultivars (Yano et al. 2016; Wang et al. 2018). Using whole genome resequence information, genome-wide association studies (GWAS) have been used to identify associated genetic loci and several important genes controlling agronomic traits such as grain yield, stress tolerance, and nutrition uptake efficiency (Bollinedi et al. 2020; Nayyeripasand et al. 2021; Kham et al. 2024). Several GWAS studies focused on grain quality traits have been reported recently (Mogga et al. 2018, Wang et al. 2019, Verma et al. 2022, Xu et al. 2022, Yoshida et al. 2023, Zhang et al. 2024). These studies detected and confirmed major QTLs that include the *GBSSI* and *SSIIa* genes. However, most of the genes controlling eating and cooking quality traits remain to be identified, especially those concerning the eating quality of cooked rice grains in *japonica* rice cultivars. In the present study, we used 1,054,635 nucleotide variants (SNPs and InDels) in 150 rice cultivars to evaluate a total of 89 cooking and eating quality traits of rice grains; to detect associated loci and identify candidate genes involved in the control of these traits by GWAS; and to select loci that make a large contribution during breeding selection by partial least squares (PLS) regression analysis.

Material and method

Plant materials

150 rice cultivars were selected from germplasm collections maintained at the Genetic Resources Center of the National Agriculture and Food Research Organization (NARO) (Supplementary Table S1). The rice cultivar set included both landraces and improved rice cultivars, selected as a representative set based on their geographic origin and the results of cluster analysis of genetic variation by means of genome-wide DNA polymorphisms (Ebana et al. 2008; Yamasaki and Ideta 2013). There were 137 non-glutinous and 13 glutinous rice cultivars, the majority (138) were *temperate japonica*, and 12 were *tropical japonica* cultivars. Twenty-four plants per cultivar were grown in double rows with 18 cm between

plants and 36 cm between rows in paddy fields at NARO, Tsukuba, Japan (36.03°N, 140.11°E) in 2012, 2014 and 2015. Seeds were sown in the middle of April, and seedlings were transplanted into the fields in the middle of May. Rice grains were harvested when each cultivar was at the grain maturing stage in September and October. Cultivation management followed the standard procedures for rice at NARO.

Evaluation of grain appearance, eating and cooking quality

We evaluated 89 grain quality and cooking and eating quality traits in the 150 rice cultivars using data from three years, 2012, 2014 and 2015 (Supplementary Table S1), resulting in a total of 249 trait datasets in this study. The 89 traits consisted of two grain component traits (amylose content and protein content of polished rice grains); 28 grain size and shape traits (length, width, perimeter, area, their ratios, circumference, elliptical axial ratio, long elliptical axis, short elliptical axis, elliptical circumference and 1000-grain weight of brown, polished and cooked rice grains); nine traits obtained in a rice cooking characteristics test (water absorption rate after boiling, expansion rate of boiled rice grains, total leached sugar content, leached reducing sugar content, total reducing sugar ratio, maximum absorption wavelength, absorbances at 540 nm, 600 nm and 660 nm of boiled rice water); five traits obtained using the Cooked Rice Taste Analyzer (eating quality, stickiness, hardness, appearance and balance degree of cooked rice grains); seven traits obtained using the Rapid Visco Analyzer (maximum viscosity, minimum viscosity, breakdown, final viscosity, setback, peak time, gelatinization temperature of rice flour); 19 traits for physical properties obtained using the Tensipresser (hardness, stickiness, adhered mass, adhesiveness, their balance degrees, sample thickness, tenderness, pliability, toughness and brittleness of surface and whole cooked rice grains); eight traits for starch and cell wall degradation enzyme activities in rice endosperm (α-amylase, β-amylase, α-glucosidase, α-galactosidase, β-galactosidase, β-glucanase, α-mannosidase and β-xylanase of rice flour); and seven traits for amylopectin chain length distribution (degrees of polymerization 5–12 (a), 13–24 (b_1), 25–36 (b_2) and ≥ 37 (b_3), chain length ratios of a/b_3, b_1/b_3 and $a+b_1/b_2+b_3$ of polished rice grains).

These grain appearance, eating and cooking quality traits were evaluated by the instrumental methods reported in previous studies (Hori et al. 2016, 2021; Iijima et al. 2019). The grain shape of brown rice, polished rice and cooked rice was scanned using an FB1210U scanner (Canon Inc., Tokyo, Japan), and each trait was evaluated using Image-Pro Plus software (Media Cybernetics Inc. DC, USA). Apparent amylose content was determined by absorbance measurement of purified starch using an SH-1000 microplate reader (Hitachi High-Tech Co. Ltd., Tokyo, Japan). Protein content was determined by near-infrared spectrometry (Inframatic 9520 spectrometer, Perten Instruments, Hägersten, Sweden). Amylopectin chain length distribution was evaluated by high performance anion exchange chromatography using a pulsed amperometric detector (HPAEC-PAD; Umemoto et al. 2002). Rapid Visco Analyzer 4800 (Perten Instruments, Hägersten, Sweden) was used to investigate starch characteristics during and after gelatinization according to American Association of Cereal Chemists Standard Method 61.02 (Aoki et al. 2012). Absorbances of eluted sugars and starch from boiled rice grains were evaluated in a rice cooking characteristics test (Shiraishi 1994). Eating quality scores were measured using an STA1A Cooked Rice Taste Analyzer (Satake Co. Ltd., Hiroshima, Japan). Physical properties of cooked grains were measured using the high-compression/low-compression method with a Tensipresser MyBoy texture analyzer (Takemoto Electric Co., Tokyo, Japan). These instrumental methods show significant correlations with eating quality scores obtained by sensory tests (Okadome 2005; Mikami 2009; Kwon et al. 2011; Hori et al. 2016). Activities of starch or cell wall degrading enzymes were evaluated in rice flours obtained from polished grains (Iijima et al. 2019). β-(1–3),(1–4)-glucan content was measured using the β-glucan assay kit (K-BGLU, Megazyme Ltd., Bray, Ireland) according to the manufacturer's instructions.

Genome sequence assembly and analysis of single-nucleotide polymorphisms (SNPs)

We collected whole genome sequence data of 150 rice cultivars that were published in previous studies (Yabe et al. 2018; Tanaka et al. 2020). These Illumina short-read sequence data were mapped against the IRGSP1.0 Nipponbare rice genome (Kawahara et al. 2013) using bwa mem (Li et al. 2019), and sequence variants for each cultivar were called essentially following the GATK Best practices for germline SNP/Indel discovery (Van der Auwera et al. 2013). Variants were first called on a by sample basis using GATK HaplotypeCaller, and then, variants were consolidated in a joint calling step with GenotypeGVCFs (Poplin et al. 2018). GATK version 4.0.11.0 was employed for all steps. Variants were then filtered using bcftools view (Li 2011) with the parameters: -m2 -M2 -g hom –output-type z –exclude-uncalled -e "MAF < 0.05 || N_MISSING > 0 || QD < 5.0 || FS > 50.0 || SOR > 3.0 || MQ < 50.0 || MQRankSum < − 2.5 || ReadPosRankSum < − 1.0 || ReadPosRankSum > 3.5", resulting in a set of variants where no position had missing data or a minor allele frequency of < 0.05. Nucleotide polymorphisms were categorized for their potential effects on protein coding genes using SnpEff 4.3t (Cingolani et al. 2012) with the *O. sativa* database.

For detailed haplotype analysis of candidate gene regions, we made a high-quality genome assembly using publicly available datasets for the rice cultivar Koshihikari (Koshi). Pacbio (SRR14366946) and Nanopore (SRR14269765) long reads were assembled together using the Canu assembler (Koren et al. 2017) with Illumina short reads (SRR1630927) for error correction. Genome annotation was performed using Braker (Stanke et al. 2006) with RNAseq samples SRR13165580, SRR13165641, SRR14140390 and SRR14140391. To compare regions harboring Hayamasari (Haya) or Khau Mac Kho chromosome segments, we used de novo assembly of paired Illumina short reads (DRR190941 and DRR289340, respectively) using Megahit (Li et al. 2015). After de novo assembly, the scaffolding program Ragtag (Alonge et al. 2022) was employed with default parameters, using the Nipponbare IRGSP-1.0 genome assembly (Kawahara et al. 2013) as the basis to order and assemble contigs.

Definition of population structure

The SNP dataset was analyzed for population structure using the fastSTRUCTURE software (Raj et al. 2014). Principal component analysis was performed on the SNP dataset using the R package SNPRelate (Zheng et al. 2012).

Genome-wide association study (GWAS)

For GWAS, we used mixed linear models (MLM; Yu et al. 2006). Association studies were performed using GEMMA (Zhou and Stephens 2012). Visualization used scripts from the R package qqman (Turner 2018). We removed variants with MAF < 0.05 from the relevant dataset in GWAS when using populations of all cultivars, *temperate japonica* rice cultivars, non-glutinous rice cultivars, or *temperate japonica* and non-glutinous rice cultivars. A significance threshold was derived by using the CalcThreshold function in the R package RAINBOWR (Hamazaki 2020) by the Benjamini–Hochberg method with a significance level of 0.05. Markers exceeding the threshold for each experiment were defined as significant markers.

Confirmation of genetic effects for detected QTNs

We used the substitution line SL937 that contains a rice cultivar Haya segment in the Koshi genetic background and SL2714 having a rice cultivar Khau Mac Kho segment in the Koshi genetic background. These substitution lines were developed by Nagata et al. (2023). Mutant lines were developed by sodium azide (Az) or N-methyl-N-nitrosourea (ENU) treatments of Toyomeki (Toyo), and by Az, ENU, diepoxybutane (DEB), or N-methyl-N-nitrosourea (MNU) treatments of Koshi. SL937 and Toyo mutant lines were used to confirm genetic effects of the QTN for amylose content, amylopectin chain length ratio and absorbance at 540 nm. SL2714 and Koshi mutant lines were used to confirm a genetic effect of the QTN for β-glucanase activity. The substitution lines and mutant lines were planted and harvested using the same cultivation and field management methods as the 150 rice cultivars in the years 2019–2023. Phenotypes of the substitution lines and mutant lines were compared with those of each wild-type parent Koshi and Toyo by using Student's *t* test.

RNAi knockdown lines were developed by using the methods in previous study (Hori et al. 2013). We generated transgenic rice plants in Koshi and Nipponbare (Nip) by means of an *Agrobacterium*-mediated transformation method using strain EHA105. For each candidate gene, three homozygous T_2 lines derived from independent single-copy T_0 transformants were grown at greenhouses and their phenotypes were scored in the years 2016 and 2018. Phenotypes of the knockdown lines were compared with those of wild-type parent Koshi and Nip by using the Dunnett's multiple comparison procedure provided by the JMP 11 software (SAS Institute Inc., NC, USA). RT-PCR and RealTime-PCR analysis in the knockdown lines were performed according to the methods in a previous study (Hori et al. 2013). Transcription levels of candidate genes, actin and ubiquitin were quantified according to their PCR product abundances. The results represent the means of at least three biological replicates, with three technical repeats for each biological replicate.

Candidate gene identification for detected QTNs

Consecutive SNPs that were closely linked with an $r^2 \geq 0.6$ in the QTN regions were considered as the conserved LD interval according to Yano et al. (2016) and identified using the clump function of plink 1.9 (Purcell et al. 2007) using the calculated Benjamini–Hochberg *p* value threshold for the parameter p1 and 10 times the p1 value as the value for p2 (i.e., a difference of 1 in $-\log_{10}P$ in a Manhattan plot) (Chang et al. 2015). Heatmaps of linkage disequilibrium were drawn with the R package LDheatmap (Shin et al. 2006). Investigation of the effects of individual SNP genotypes on the phenotypes was performed by comparing gene and amino acid sequences between each cultivar and the reference annotations of Nip in the Rice Annotation Project Database (RAP-DB, https://rapdb.dna.affrc.go.jp/index.html) (Sakai et al. 2013) and by using the R packages Haplotypes (Aktas 2020) and VariantAnnotation (Obenchain 2014).

Partial least squares (PLS) regression analysis

We selected 50 rice cultivars that were grown in a large cultivation area in Japan. The cultivars were divided into

five sub-groups according to their food applications (Japan Rice Market 2023): good eating quality rice, staple food rice, sushi rice, sake brewing rice, and high yielding rice (including rice flour used for making gluten-free bread). PLS regression analysis was performed using the multivariate analysis software Pirouette Ver. 4.5 (Infometrix Inc., WA. USA, https://infometrix.com/tag/pirouette/). We used scores of 1 or 0 for food applications as the objective variable and evaluation scores of each trait as the explanatory variable.

Results

Identification of genetic loci affecting grain appearance, eating and cooking quality traits by GWAS

Our study population comprised 150 rice cultivars including both landraces and improved cultivars (Supplementary Table S1). The Japanese rice core collection (JRC) comprises 50 cultivars that were selected as a representative population of Japanese landrace cultivars. We selected 44 of the 50 that were *temperate* or *tropical japonica* landraces. The remaining 106 cultivars were selected as a representative set of improved rice cultivars developed from the 1920s to the 2000s. A set of single-nucleotide polymorphism variants (SNPs) was chosen with a minimum minor allele frequency of 0.05 and with no missing genotypes among the 150 cultivars. This SNP set comprised 1,054,635 SNPs distributed throughout all 12 rice chromosomes, and we used them to further study population structure and to detect genetic loci.

The population structure of the 150 rice cultivars was investigated using fastSTRUCTURE with a K value (number of population groups) ranging from 2 to 7 (Supplementary Fig. S1A). Analysis using the choose K function in fastSTRUCTURE identified that the most suitable model component to explain structure in the data was 5, while the complexity that maximized marginal likelihood was 3, suggesting that the optimal number of groups is between three and five. At $K=2$, the 150 cultivars were mostly divided into *temperate japonica* rice and *tropical japonica* rice. At $K=5$, in addition to one *tropical japonica* group including 17 cultivars, the *temperate japonica* cultivars were further divided into four groups comprising 84, 23, 13 and 13 cultivars, respectively. Glutinous rice cultivars were distributed throughout all five groups, suggesting that glutinous rice cultivars have been created by mutation or introgression into all of the population groups used in this study.

3D scatter plots of the results of principal component analysis (PCA) were drawn based on the genetic relationship matrix derived from the 1,054,635 SNPs after pruning based on linkage disequilibrium (Supplementary Fig. S1B–E). Among all the rice cultivars, there were two major clusters clearly distinguished along the axis of principal component 1 (PC1) between the *temperate japonica* and *tropical japonica* groups (Supplementary Fig. S1B). Glutinous rice cultivars were included both of these clusters as shown in the population structure analysis in Supplementary Fig. S1A and in Supplementary Fig. S1C. Among the *temperate japonica* rice cultivars, *K*-means clustering of the genetic relationship matrix with a K of 4 derived similar groups as the structure analysis (Supplementary Fig. S1C, D). A phylogenetic tree based on the whole-genome SNPs illustrates the separation of *tropical* and *temperate japonica* cultivars, and the distribution of glutinous cultivars throughout the populations (Supplementary Fig. S1F).

We evaluated 89 grain appearance quality and cooking and eating quality traits among the 150 rice cultivars using data from 3 years, 2012, 2014 and 2015 (Supplementary Table S1) as detailed in the Materials and Methods section. The main categories of traits were: grain size and shape, grain components, cooking characteristics including absorbance and sugar content eluted in the cooking water, taste analysis, starch viscosity, physical properties of cooked rice, endosperm starch and cell wall degradation enzyme assays, and amylopectin chain length distribution.

We carried out GWAS for these grain appearance, eating and cooking quality traits using all 150 rice cultivars and detected a total of 525 regions where significant GWAS peaks were detected (Supplementary Table S2, S3, Fig. S2). The detected QTNs were localized on all 12 rice chromosomes. Among these QTNs, cluster regions of genetic loci were found on rice chromosomes 1, 4, 5, 6, 8, 9, 10 and 11 (Supplementary Fig. S2). Several cluster regions included starch biosynthesis and degradation genes, including *SSIVa*, *PUL*, *PPDKB*, *AGPL1*, *SSIVb*, *GBSSI*, *SSI*, *SSIIa*, *Amy2A*, *BEI*, *Amy3E* and *ISAI*. In GWAS for amylose content, we detected a significant peak on the short arm of chromosome 6 (Supplementary Fig. S3A). The peak was localized with the amylose biosynthesis gene *GBSSI*, indicating that GWAS detected allelic differences of Wx^b and *wx* at the *GBSSI* gene among the 150 rice cultivars. In GWAS for grain width of brown rice, we detected significant peaks on chromosomes 3 and 4. Significant peaks were localized near previously identified genes such as *GS3, BBS1 and NGLF* that are implicated in controlling grain size and grain width (Supplementary Fig. S3B), indicating that there are allelic differences associated with these grain size genes among the 150 rice cultivars. These results suggested that this population of rice cultivars and sequence polymorphisms had the potential to detect genetic loci associated with grain traits by GWAS.

GWAS was also carried out using different cultivar sets within the original 150: 137 non-glutinous rice cultivars, 138 *temperate japonica* rice cultivars and 129 *temperate japonica* non-glutinous rice cultivars. Compared with the using all 150 cultivars different numbers of significant

peaks were observed on several grain appearance, eating and cooking quality traits in these cultivar sets. For example, on absorbance score at 540 nm of rice cooking characteristics test, two major significant peaks were detected on chromosomes 6 and 11 in all 150 rice cultivars, while only one significant peak was detected on chromosome 11 in 138 *temperate japonica* rice cultivars (Fig. 2A, Supplementary Fig. S3C, D). For amylopectin chain length short/long chain distribution (a + b1/b2 + b3 chain ratio; Supplementary Table S1), one significant peak was observed on chromosome 3 in 2015 using all 150 cultivars, while two significant peaks were observed on chromosomes 3 and 11 in 129 non-glutinous *temperate japonica* rice cultivars (Supplementary Fig. S3E, F). This suggests that loci influencing grain characteristics are distributed, and allelic differences of their responsible genes exist both in *temperate* and *tropical japonica* and in glutinous and non-glutinous rice cultivars. We focused on GWAS peaks that were detected in at least two different years for candidate gene analysis.

Candidate gene identification for adhesiveness of cooked rice grains

Surface adhesiveness of cooked rice grains is an important component of eating quality. We detected a GWAS peak for surface adhesiveness of cooked rice grains on the short arm of rice chromosome 3 using the 129 cultivars of *temperate japonica* non-glutinous rice population (Fig. 1A) and as a lesser peak using the full 150 cultivar population including glutinous rice, where the *GBSSI* genotype on chromosome 6 also strongly affects surface adhesiveness (Supplementary Fig. S3G).

Based on local LD structure and examination of Manhattan plots for this genetic locus on chromosome 3, we identified an associated genomic region including thirty-three genes from 1.17 Mbp to 1.50 Mbp (Fig. 1B, Supplementary Table S2). This region is adjacent to, but distinct from the eating quality QTL *qOE3* (0.41–1.23 Mb for *qOE3*) documented in previous reports (Wada et al. 2015; Hori et al. 2021). Among the thirty-three genes within the narrowed down region, we found three candidate genes that have insertion/deletions or nonsynonymous substitution polymorphisms between Koshihikari (Koshi) and the rice cultivar Nipponbare (Nip) which has poorer eating characteristics (Fig. 1C). These three candidate genes had several polymorphisms among the 150 rice cultivars, with six haplotypes of Os03g0124100 encoding a protein containing a DUF604 domain, seven haplotypes of Os03g0125300 encoding a protein containing a D111/G-patch domain, and 13 haplotypes of Os03g0125800 encoding a protein containing a cystathionine beta-synthase domain among the population. The variants in the gene regions and phenotypic differences observed among the haplotype groups are shown in Supplementary Figs. S4, S5, and S6. Perhaps the most impactful sequence change is in Os03g0124100 where a 17 bp deletion changes the N terminus of the predicted protein and indicates an alternative start codon (Fig. 1C; Supplementary Fig. S7). The effect of either Nip or Koshi haplotypes on the phenotype is shown in Supplementary Fig. S8 for each gene. In each case, haplotypes containing the Koshi allele at the representative SNP show superior eating quality scores as well as increased surface adhesiveness. We detected mRNA of the three candidate genes by RT-PCR in rice endosperm at the grain filling stage in both Koshi and Nip (Supplementary Fig. S9), supported by gene expression data in the expression database RiceXPro (https://ricexpro.dna.affrc.go.jp/). We developed three independent RNAi knockdown transformant lines for each of these three candidate genes in the Koshi and Nip genetic backgrounds. In the RNAi knockdown transformant lines, mRNA expression of the three candidate genes was decreased in rice endosperm at the grain filling stage (Fig. 1D). The RNAi knockdown transformant lines of Os03g0124100 showed significantly weaker adhesiveness on the surface of cooked rice grains and low levels for eating quality scores in both the Nip and Koshi backgrounds, compared with the respective wild type plants, consistently in two years (Fig. 1E). The RNAi knockdown transformant line of Os03g0125300 did not show significant change in adhesiveness in the Koshi background, but the eating quality score was significantly increased in both years. In the Nip background, adhesion and eating quality were increased in one year only. For Os03g0125800 eating quality score was significantly decreased in both backgrounds consistently in two years, while adhesion significantly changed only in one year, and in opposite directions in the different backgrounds. Thus, each of the three candidate genes has the capacity to influence eating quality and adhesiveness on the surface of cooked rice grains, with Os03g0124100 showing the most consistent effects.

Candidate gene identification at novel loci for amylose and amylopectin composition

Milled rice contains up to approximately 90% starch (Patindol et al. 2015), so starch composition is fundamental to cooking and eating properties. We measured amylose in purified starch by spectrophotometry and evaluated amylopectin chain length distribution by high performance anion exchange chromatography. We also measured the absorbance of the eluates of rice boiled in 50-ml tubes—spectrophotometry in the visible range has been used to estimate the relative concentrations of amylose and amylopectin, which have overlapping absorption spectra (Chikubu et al. 1960; Shiraishi 1994). Absorbance at 540 nm (A_{540}) coincides with the peak absorption wavelength of amylopectin, while amylose also absorbs to a lesser degree at the same wavelength.

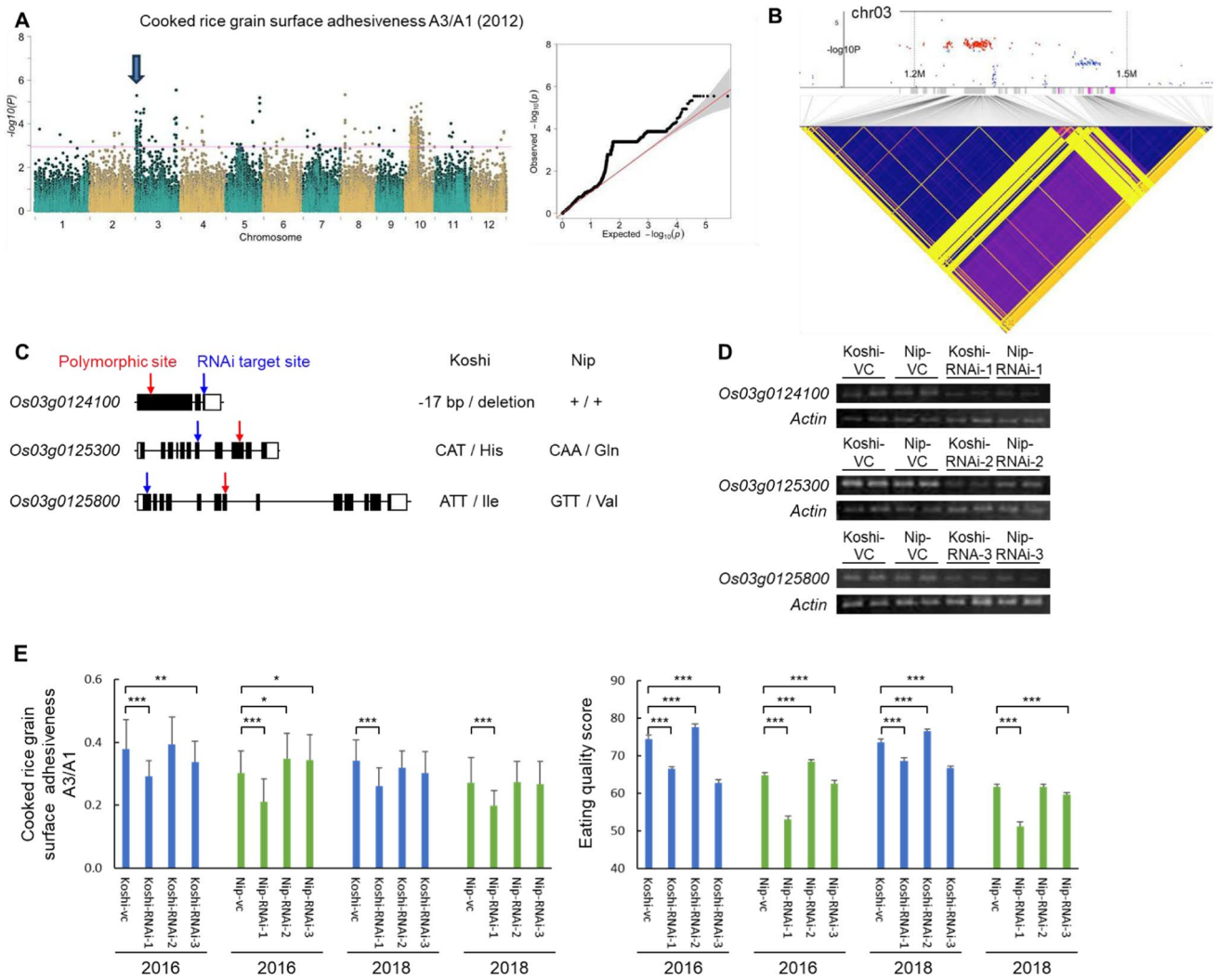

Fig. 1 GWAS results for adhesiveness balance degree A3/A1 on the surface of cooked rice grains in 129 *temperate japonica* non-glutinous rice cultivars (**A**) The chromosome 3 peak is indicated by the arrow. Candidate gene analysis with local Manhattan plot and linkage disequilibrium (LD) heatmap around the genetic loci on the short arm of chromosome 3. Significant SNPs are shown in red; examined candidate genes are indicated in pink. The horizonal black line above the plot indicates the local LD region. (**B**). Polymorphic sites between Nipponbare (Nip) and Koshihikari (Koshi), and RNAi target site on structure of candidate genes (**C**). mRNA expression of the candidate genes Os03g0124100, Os03g0125300 and Os03g0125800 with the *Actin* gene in endosperm at grain filling stage at two weeks after flowering in Koshi, Nip, three RNAi knockdown lines in the Koshi genetic background, and three RNAi knockdown lines in the Nip genetic background (**D**). Adhesiveness balance degree A3/A1 of cooked rice grains and eating quality scores in RNAi knockdown lines in Koshi and Nip genetic backgrounds analyzed in two years (**E**)

We detected a novel peak on chromosome 11 that appeared in several GWAS analyses related to starch composition. The peak appeared in A_{540} of boiled rice water eluate (Fig. 2A), in chain length distribution of amylopectin a/b3 (Fig. 2B) and in amylose content (Supplementary Fig. S3A). The presence of the peak was variable depending on the population. For A540, its presence depended on very low amylose cultivars, as the peak disappeared when only non-glutinous cultivars were used in GWAS (Supplementary Fig. S2C, D, E). However, the peak was detected using only non-glutinous cultivars for a/b3 amylopectin chain length ratio (Supplementary Fig. S3H, I). Thus, a genomic peak region on chromosome 11 was associated both with amylose content and with amylopectin chain length ratio.

The A_{540} chromosome 6 peak, as for the chromosome 6 amylose content peak, was localized in the same genomic region as the *GBSSI* gene (Fig. 2C, Supplementary Fig. S3). The novel peak for A_{540} of boiled rice water on the long arm of chromosome 11 was detected both using all 150 rice cultivars and also using only the 138 cultivars of *temperate japonica* rice population where it was the largest peak (Fig. 2A, Supplementary Fig. S3C, D). Analyzing local

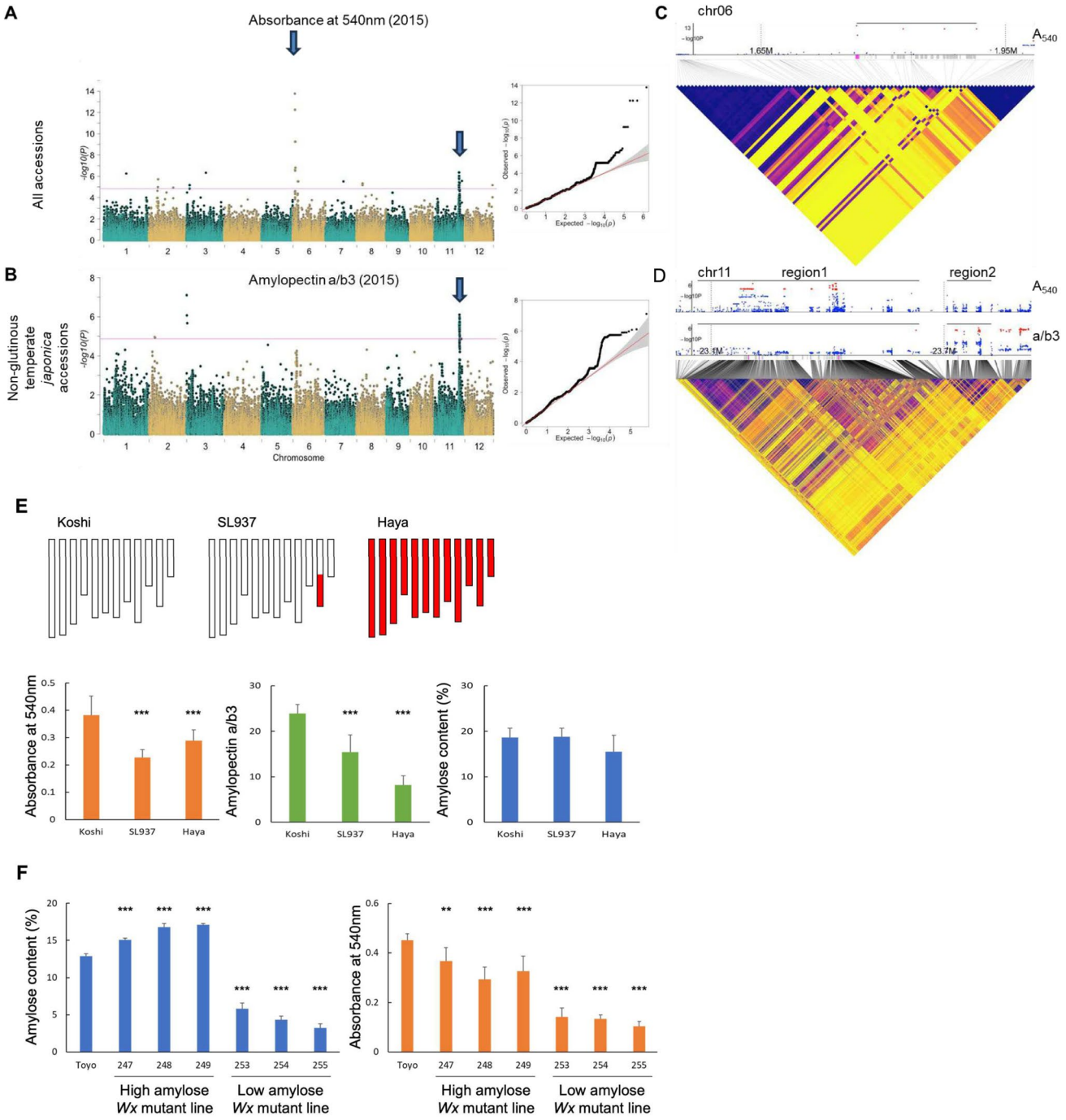

Fig. 2 GWAS results for absorbance at 540 nm of boiled rice water in 150 *japonica* rice cultivars (**A**). GWAS results for amylopectin chain length distribution of a/b3 in 129 *temperate japonica* non-glutinous rice cultivars. Arrows indicate the chromosome 6 peak at *GBSSI* and the novel chromosome 11 peak (**B**). Candidate gene analysis with local Manhattan plot and linkage disequilibrium (LD) heatmap around the genetic loci on the short arm of chromosome 6. Significant SNPs are shown in red; examined candidate genes are indicated in pink. The horizonal black line above the plot indicates the local LD region (**C**). Candidate gene analysis with local Manhattan plot and linkage disequilibrium (LD) heatmap around the genetic loci on the long arm of chromosome 11. Significant SNPs are shown in red; examined candidate genes are indicated in pink. The horizonal black line above the plot indicates the local LD region (**D**). Confirmation of genetic loci for absorbance at 540 nm with genotype and grain appearance in a substitution line of SL937 having a Hayamasari (Haya) chromosome segment in the Koshihikari (Koshi) genetic background (**E**). Confirmation of genetic loci for amylopectin chain length distribution of a/b3 and amylose content in a substitution line SL937 (F). Amylose content, absorbance at 540 nm, and grain appearance in six mutant lines for the *GBSSI* (*Wx*) gene in the genetic background of Toyomeki (Toyo)

linkage disequilibrium (LD) and Manhattan plots in this region, we identified an associated genome region within a broad peak spanning 23.06–23.63 Mbp on chromosome 11 that included thirty-four genes in region 1 (Fig. 2D, Supplementary Table S2). Adjacent, and linked to this peak, was the peak for amylopectin a/b3 chain length ratio detected in the full population and notably also in the 129 *temperate japonica* non-glutinous rice cultivars in region 2 (Fig. 2D; 23.62–23.99 Mbp, Supplementary Table S2). To confirm the genetic effects of these candidate regions on chromosome 11, we evaluated a substitution line SL937 (CSSL) that has a segment from the rice cultivar Haya in the Koshi genetic background (Fig. 2E). Haya had decreased A_{540} of boiled rice water compared with Koshi. A_{540} was also decreased in SL937, indicating that we confirmed the presence of a QTL associated with A_{540} in this chromosome region. We also confirmed the genetic effect of this QTL on amylopectin chain length a/b3 distribution in SL937 (Fig. 2E)—Haya genotype in this region reduced the proportion of a/b3 chains. Furthermore, the amylose content of SL937 grains was not significantly different compared with those of Koshi or Haya. These results indicated that the region on chromosome 11 is involved in controlling amylopectin biosynthesis and absorbance in eluted boiled water at 540 nm, independently of the control of amylose biosynthesis in rice grains.

The A_{540} chromosome 11 peak was not detected when all glutinous cultivars were excluded (Supplementary Fig. S3C, D) and we examined six nonsynonymous *GBSSI* mutant lines showing higher or lower (but not completely deficient) amylose contents, to clarify the relationship between allelic differences in *GBSSI*, amylose content in the grain and A_{540} in the boiled eluate (Fig. 2F). The A_{540} scores of the low amylose mutant lines (253, 254 and 255) were greatly decreased compared with those of the high amylose mutant lines, corresponding with a two–threefold decrease in grain amylose content. Compared with the mutant lines background of cultivar Toyomeki (Toyo), three high amylose mutant lines (247, 248 and 249) also showed decreased A_{540} of boiled rice water eluate. These results indicated that allelic differences in the *GBSSI* gene affect amylose content, but the relationship with A_{540} is complex, as both increase and decrease in amylose caused a decrease in the A_{540}.

According to gene annotation information in RAP-DB, there were four candidate genes in the chromosome 11 region that could be involved in starch biosynthesis (Fig. 2D, Supplementary Table S2). Three candidate genes in region 1 had several polymorphisms among the 150 rice cultivars, and the rice cultivars harbored 11 haplotypes of Os11g0602800 encoding a protein containing a galactose oxidase domain, 13 haplotypes of Os11g0603000 encoding a Basic helix-loop-helix (bHLH) transcription factor, and 19 haplotypes of Os11g0607100 encoding a protein containing a pentatricopeptide repeat domain (Supplementary Fig. S10, S11, S12). In region 2, we found 7 haplotypes of Os11g0612950 encoding an inorganic diphosphatase with varying phenotypes (Supplementary Fig. S13).

The three candidate genes in the first LD region were expressed in the developing endosperm and mature seed according to the RiceXPro database (Supplementary Figure S14). Because there were many variant nucleotides among the haplotypes for the genes in the peak region on chromosome 11, we considered that the variation might be due to structural differences or large introgressions among cultivars. We compared local genome alignments in the region for Koshi, Haya and Nip. For both Os11g0602800 and Os11g0607100, the Haya region was divergent in the promoter region from the Koshi sequence, suggesting the possibility of divergent expression or other variation depending on non-coding sequences between Koshi and Haya (Supplementary Fig. S15).

Considering Os11g0612950 in region 2, detected in the peak for amylopectin a/b3 chain length ratio, all haplotypes except haplotype 2 had a non-synonymous mutation resulting in an amino acid change in the predicted protein sequence (Supplementary Figure S13). Comparing the Haya (Haplotype 5) and Koshi (Haplotype 2—the same as Nip) haplotypes, the Koshi haplotype was associated with higher amylopectin a/b3 ratio in the whole population and in the non-glutinous cultivars (Supplementary Figs. S13C, S16A).

As region 1 and region 2 are linked, we also checked the mean A_{540} for Os11g0612950. A_{540} was slightly higher with the Haya allele, and the lowest A_{540} values were derived from with glutinous cultivars (Supplementary Fig. S16B). This was consistent with the amylose and amylopectin contents affecting the A_{540} in boiled rice water eluate in a complex manner, as in the *GBSSI* alleles where both higher and lower amylose contents caused decreased A_{540} in the boiled eluate (Fig. 2F).

We compared the Koshi genome assembly with the IRGSP1.0 Nip genome and with the de novo assembled Haya contigs in the region of Os11g0612950 (Supplementary Fig. S17). According to this analysis, it is possible that the gene could be deleted in Koshi—the haplotype recorded as the same as Nip in the short-read mapping analysis could result from a lack of mapped reads from Koshi in the gene region, and thus a lack of SNP calls. Therefore, it is possible that the phenotypic changes associated with the SNPs in this gene actually reflect presence or absence of the gene rather than a sequence change between Haya and Koshi alleles.

Candidate gene identification for β-glucanase activity in rice endosperm

We identified a GWAS peak for β-glucanase activity in rice endosperm on the long arm of rice chromosome 3 (Fig. 3A). The peak was detected both in the 150 rice cultivars and

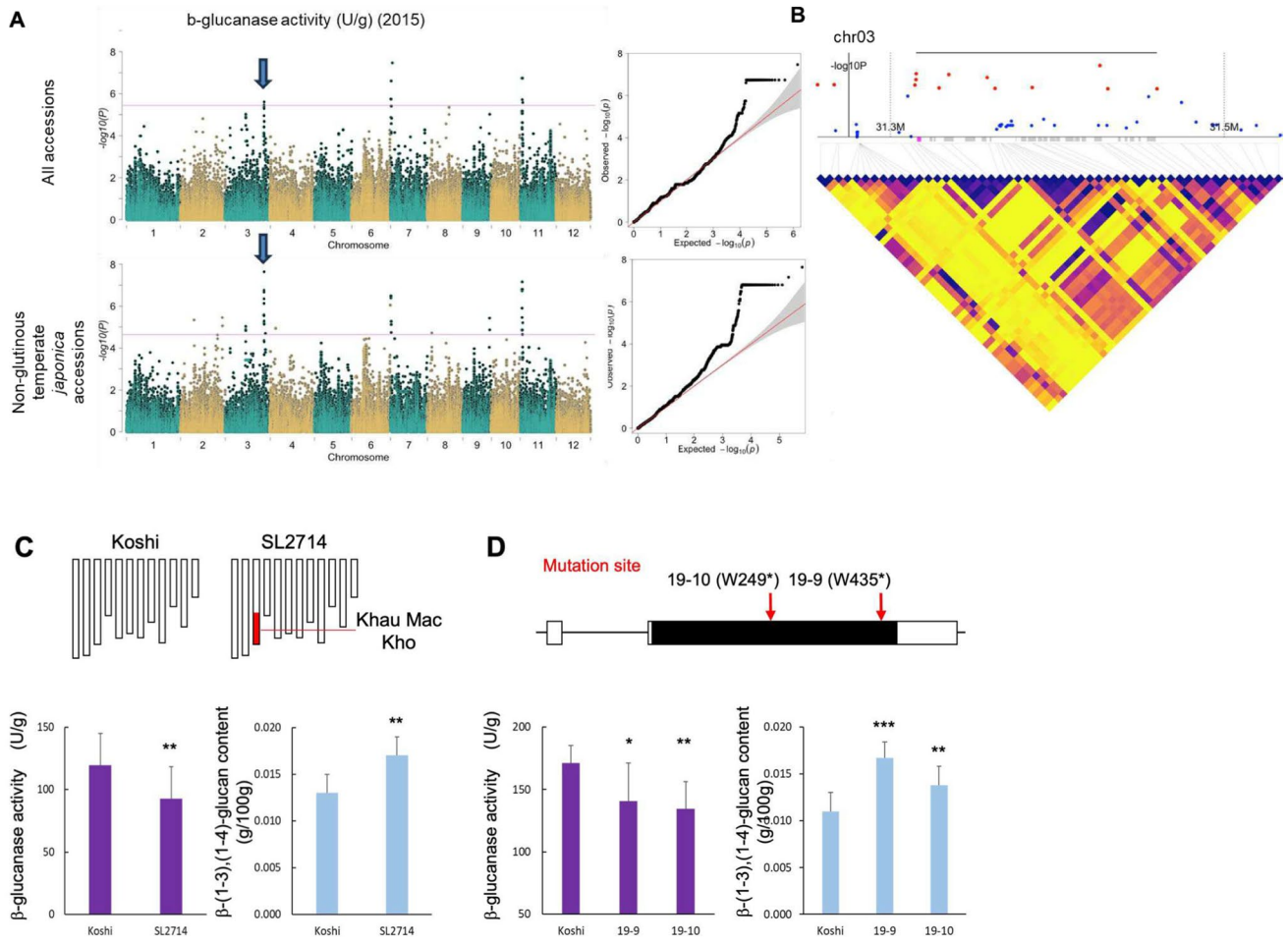

Fig. 3 GWAS results for β-glucanase activity in 150 rice cultivars and 129 *temperate japonica* non-glutinous rice cultivars (**A**). Candidate gene analysis with local Manhattan plot and linkage disequilibrium (LD) heatmap around the genetic loci on the long arm of chromosome 3 (**B**). Confirmation of genetic loci for β-glucanase activity and β-(1–3),(1–4)-glucan content with genotype and grain appearance in a substitution line of SL2714 having a Khau Mac Kho chromosome segment in the Koshihikari (Koshi) genetic background. Because Khau Mac Kho showed extremely late heading date in our fields, we were unable to evaluate its β-glucanase activity and β-(1–3),(1–4)-glucan content in this study (**C**). Mutation site on gene structure, grain appearance, β-glucanase activity and β-(1–3),(1–4)-glucan content in two mutant lines for the candidate gene Os03g0757900 in the genetic background of Koshi (**D**)

the 129 *temperate japonica* non-glutinous rice cultivars. We checked local Manhattan plots and LD structure for this locus and narrowed down the genome region from 31.31 to 31.45 Mbp including twenty-nine genes (Fig. 3B, Supplementary Table S2). To confirm genetic effects of the QTN on chromosome 3, we used a substitution line SL2714 which has a segment from rice cultivar Khau Mac Kho in the Koshi genetic background (Fig. 3C). SL2714 showed decreased β-glucanase activity compared with Koshi, which likely resulted in the observed increase in β-(1–3),(1–4)-glucan content in SL2714. These results indicated presence of a QTN associated with β-glucanase activity in this chromosome region. Among the 29 genes within this genome region, we found one candidate gene related to saccharide dehydrogenase activity according to the RAP-DB gene annotation (Supplementary Table S2). The candidate gene product of Os03g0757900 has a similar amino acid sequence to UDP-glucose 6-dehydrogenases. This gene had several polymorphisms among the 150 rice cultivars, and the rice cultivars harbored 7 haplotypes within the coding region; however, we could not identify significant phenotype differences among these seven haplotypes. Extending the examined region into the promoter region (Supplementary Fig. S18A), we identified ten haplotypes, and one SNP at position 31,315,421, where the Khau Mac Kho allele was strongly associated with higher β-glucanase activity, suggesting that there may be variation in expression level among the cultivars (Supplementary Fig. S18D). Examination of the de novo assembled contigs for this region in Khau Mac Kho in comparison with the Koshi genome assembly did not show major changes in the region of the candidate gene (Supplementary Fig. S18E). According to the RiceXPro

database, the candidate gene was expressed in developing endosperm and mature seed (Supplementary Fig. S18C). We developed two mutant lines where the mutation resulted in a stop codon gain in the coding region of Os03g0757900 in the Koshi genetic background. Both mutant lines showed significantly decreased β-glucanase activity compared with the wild type of cultivar Koshi (Fig. 3D). These results indicated that the candidate gene was involved in controlling differences in β-glucanase activity in developing grains among the rice cultivars.

Detection of gene loci combinations for improving grain appearance, eating and cooking quality traits

Among the detected genetic loci including QTNs and their candidate genes, we tried to elucidate important genetic loci for rice breeding programs for each food application. The 150 rice cultivars were divided into five application groups: good eating quality rice, staple food rice, sushi rice, high yielding rice, and sake brewing rice based on data in Japan Rice Market (2023). The good eating quality rice groups contain the top 10 leading rice cultivars in Japan such as Koshi, Hitomebore and Hinohikari showing high eating quality scores. Staple food rice shows high grain yield and is used for general cooking and eating. Sushi rice including Sasanishiki is mainly used for making sushi. High yielding rice and Sake brewing rice have large grains, high chalkiness and many white core grains and are not suitable for general eating.

We selected 50 rice cultivars based on their pedigree and breeding history (Supplementary Fig. S19). The 50 rice cultivars consisted of 16 good eating quality cultivars, 12 staple food cultivars, 10 sushi cultivars, six high yielding cultivars and six sake brewing cultivars as a calibration set to construct a calibration curve in partial least squares (PLS) regression analysis (Figs. 4A, B). The remaining 100

A

Good eating quality rice	Staple food rice	Sushi rice	High yielding rice	Sake brewing rice
Koshihikari	Goropikari	Sasanishiki	Itadaki	Yamadanishiki
Tsuyahime	Daichinokaze	Aichinokaori	Fukihibiki	Gohyakumangoku
Akitakomachi	Chiyonishiki	Asahi	Oochikara	Omachi
Hitomebore	Dontokoi	Kinuhikari	Akihikari	Wataribune
Nanatsuboshi	Notohikari	Nipponbare	Mutsuhomare	Shinriki
Haenuki	Hananomai	Domannaka	Todorokiwase	Nakateshinsenbon
Hinohikari	Haruru	Hatsushimo		
Yumetsukushi	Hohohonoho	Hanaechizen		
Nikomaru	Manamusume	Menkoina		
Tsugaruroman	Mutsuhomare	Yukinosei		
Asahinoyume	Yomehitachi			
Fusaotome	Yukihikari			
Kinumusume				
Aichinokaori				
Sakihikari				
Koshiibuki				

B

Food application subgroup	R^2 score
Good eating quality rice	0.543
Staple food rice	0.403
Sushi rice	0.330
High yeilding rice	0.514
Sake brewing rice	0.402

C

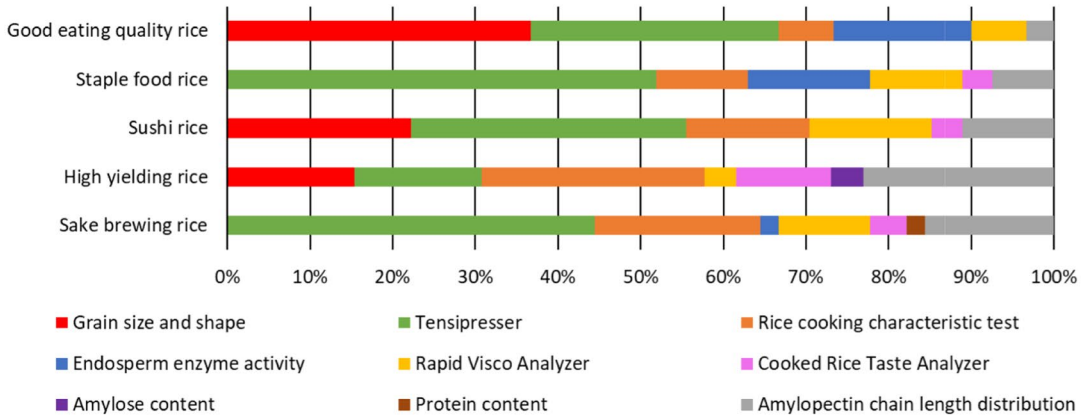

Fig. 4 Calibration set of 50 rice cultivars that are divided into five groups based on their food applications in partial least squares (PLS) regression analysis (**A**). Results of predicting the application of unclassified 100 rice cultivars according the PLS regression analysis (**B**). Combination of QTNs or SNPs showing large regression coefficient scores based on calibration curve among each food application group (**C**)

rice cultivars were unclassified and used as a prediction set. SNP genotypes that were coincident with individual QTNs were used as explanatory variables, and food applications of rice cultivars were used as response variables. The R^2 scores obtained for each group are presented in Fig. 4B and ranged from 0.330 to 0.543. This PLS regression analysis revealed that SNP sets showing large regression coefficient scores based on the calibration curve were different among rice cultivar groups with each food application (Fig. 4C). In the good eating quality rice cultivars including Koshi, 124 QTNs for grain size and shape, physical properties of cooked rice grains and endosperm enzyme activities showed large regression coefficient scores, indicating that frequencies of SNP alleles of these QTNs were different with other application and characteristics groups. The good eating quality cultivars showed large grains, high stickiness and low hardness scores for cooked rice grains, and high enzyme activities compared with other cultivar groups. In the staple food rice cultivars, 142 QTNs for physical properties of cooked rice grains showed large regression coefficient scores, indicating that this cultivar group had a tendency for physical properties of cooked rice grains with low stickiness and high hardness scores. In the sushi rice cultivars, 79 QTNs for grain size and shape, physical properties of cooked rice grains and gelatinization viscosity showed large regression coefficient scores, indicating that large grains, weak stickiness of cooked rice grains and resistance to starch aging were important in this cultivar group. In the high yielding rice cultivars, 165 QTNs for protein content and all of the grain appearance, eating and cooking quality traits except for endosperm enzyme activities were detected, indicating that these cultivars had different SNP alleles for the QTNs compared with other cultivar groups. For example, the high yielding rice cultivars had a tendency for high protein content with low stickiness and high hardness scores in cooked rice grains. In the sake brewing rice cultivars, 93 QTNs for protein content and physical properties of cooked rice grains showed large regression coefficient scores, indicating that low protein content and strong surface hardness combined with softness inside cooked rice grains were important properties in this cultivar group. The PLS regression analysis selected combinations of QTNs (SNP genotypes) showing good performance for each food application.

Discussion

Rice cultivars are required to show both high yield and high grain quality. In particular, grain quality in rice has recently become the primary consideration of breeders and farmers, because grain quality in rice largely determines the market price and consumer acceptance (Champagne et al. 1999; Fitzgerald et al. 2009; Hori and Sun 2022; Li et al. 2022). Grain quality is a major target in present rice breeding programs in many countries including China, India, Thailand, Indonesia, the Philippines, Korea and Japan. An understanding of the genetic factors that are involved in the control of rice grain quality is important for the efficient development of novel rice cultivars with high grain quality. In this study, GWAS detected a total of 525 QTN peaks associated with grain appearance, eating and cooking quality traits in rice cultivars. Several associated regions were detected in the whole 150 rice cultivar population including both of *tropical* and *temperate japonica* and both glutinous and non-glutinous cultivars; other QTNs were detected only among the 129 *temperate japonica* non-glutinous rice cultivars. The different QTNs detected in each cultivar likely depend on differences in allelic representation between *tropical* and *temperate japonica* subspecies or between glutinous and non-glutinous types. Among the detected QTNs, PLS regression analysis revealed 124 QTNs that could be important for the development of novel rice cultivars with good eating and cooking characteristics in further breeding programs. Some of the genotypes may achieve superior eating quality when compared with current cultivars.

In this study, we confirmed the genetic effects of QTNs for adhesiveness on surface of cooked rice grain on the short arm of chromosome 3; amylose content, amylopectin composition and A_{540} on the short arm of chromosome 6 and on the long arm of chromosome 11; and for β-glucanase activity in rice endosperm on the long arm of chromosome 3. Detailed haplotype analysis identified eight candidate genes. Among the eight candidate genes, the gene functions of five candidate genes were investigated by development and evaluation of mutant and knockdown transgenic lines.

For the chromosome 3 region associated with adhesiveness on the surface of cooked rice grains, we identified a region that adjoins the *qOE3* QTL region (Hori and Yano 2013), but is distinct and thus represents a novel QTL locus. We examined knockdown transgenic lines of Os03g0124100, Os03g0125300 and Os03g0125800, which each affected the surface adhesiveness and eating quality in the transgenic lines. Thus, all three are potential causal genes for adhesiveness of cooked rice grain. All knockdown lines also showed similar grain appearance, indicating that these candidate genes could improve eating and cooking quality traits without grain appearance alterations. Based on its consistent knockdown phenotype and expression pattern and its position in the GWAS peak, Os03g0124100 may be the most favorable candidate for rice improvement. This region is particularly interesting because the knockdown lines could achieve eating quality scores higher than Koshi—widely considered as the consumer gold standard for *japonica* rice. The region also adjoins a known eating quality QTL—the use of GWAS allows a resolution greater than that achievable by biparental mapping.

The multi-trait-associated locus on chromosome 11 related to starch properties revealed a number of candidate genes, although it was difficult to identify non-synonymous nucleotide changes that would clearly affect protein activities. However, this locus does appear to be important for rice grain quality based on the number of QTNs detected among all of the GWAS experiments (Supplementary Fig. S1). The mutant lines with high or low amylose content clearly indicated that the *GBSSI* gene was the main causal gene for amylose content, amylopectin composition and absorbance at 540 nm, but the chromosome 11 locus also obviously has some influence, particularly concerning amylopectin chain length. The chromosome 11 locus may be important for variation in starch composition, particularly in temperate non-glutinous *japonica* rice cultivars. Identification of such additional loci that can potentially be used for control of starch composition provides additional tools for breeders to improve rice grains.

The knockout mutant lines of the β-glucanase gene Os03g0757900 clearly showed decreased enzyme activities in rice endosperm. The adhesiveness genes and the β-glucanase gene on chromosome 3 are shown to be novel eating and cooking quality genes, because there are no previous studies that reported gene functions for these traits (Li et al. 2022; Ren et al. 2023). Adhesiveness is an important component of eating quality (Hori et al. 2016), and endosperm β-glucanase activity is lower in recently developed cultivars than in landraces or older cultivars, negatively correlated with eating quality scores (Iijima et al. 2019). Our results provide support for the use of GWAS and haplotype analysis as good experimental tools to collect and identify novel eating and cooking quality genes.

For each of the three loci analyzed in this study, knockdowns or mutant lines in the newly identified loci did not significantly alter the appearance of the polished rice grains (Supplementary Figs. S8F, S14D, E and S18E, F). This is an important consideration for the production of new cultivars. For example, an increased proportion of white-centered grains may indicate increased chalkiness. Degree of grain chalkiness and aroma are important grain quality traits in rice breeding programs in many rice cultivation areas (Champagne et al. 1999; Fitzgerald et al. 2009; Hori and Sun 2022), although we did not evaluate them directly in this study. Among our 150 rice cultivars, there are not large phenotypic differences for fragrance or grain chalkiness, as all Japanese rice cultivars lack aroma and have a low degree of grain chalkiness. However, global warming has a major impact on grain quality including fragrance and chalkiness (Matsutomi et al. 2019, Yoshimoto et al. 2021) and is likely to increase chalkiness to the detriment of grain quality. Therefore, our GWAS and haplotype procedures should also be used to identify causal genes for aroma and grain chalkiness traits in future using the current population and additional populations as the environment changes. Introduction of newly identified variants may contribute to mitigation of detrimental grain quality changes induced by higher environmental temperatures.

In addition to GWAS using whole genome sequences of a large numbers of rice cultivars, other omics technologies such as transcriptome, metabolome and ionome analyses are also useful to accelerate gene identification for grain appearance, eating and cooking quality traits. Previous genetic studies have mainly identified starch biosynthesis or degradation genes as grain quality, and eating and cooking quality trait genes in rice. However, several studies have reported that other genes encoding transcription factors and other proteins directly interacted with starch biosynthesis genes. For example, *Flo6*, *FGR1* and *OsPK2* are involved in the control of the grain appearance and eating and cooking quality traits (Peng et al. 2014; Cai et al. 2018; Hao et al. 2019; Li et al. 2022; Ren et al. 2023). Genome editing at cis-regulatory elements of starch biosynthesis genes or nutrition uptake genes has been successful in altering targeted gene expression levels and changing grain components (Ding et al. 2021). These results have indicated that it is important to understand interactions among combinations of genes such as control of expression levels of individual genes in the development to novel high grain quality cultivars in rice. Metabolome and ionome analyses have revealed many kinds of grain components including saccharides, amino acids, lipids, carotenoids, organic acids and minerals other than starch (amylose and amylopectin) as well as seed storage proteins. Several genetic studies have identified transcription factor and transporter genes for these grain components such as *OsbZIP18*, *OsAAP6* and *OsNramp5* so far (Peng et al. 2014; Sun et al. 2020; Yu et al. 2022). Additional gene identifications are required to improve the grain appearance, eating and cooking quality traits, and our study demonstrates that GWAS is one viable method to identify the necessary alleles and loci.

Grain appearance, eating and cooking quality represents a set of complex and quantitative traits that are controlled by multiple genes or quantitative trait loci (QTLs). These traits are strongly influenced by environmental factors (Champagne et al. 1999; Fitzgerald et al. 2009; Hori and Yano 2013; Hori and Sun 2022; Li et al. 2022). Therefore, pyramiding of multiple genes with both major and minor genetic effects is likely to be effective to improve grain appearance, eating and cooking quality traits. Our PLS regression analysis revealed important QTNs for developing further rice cultivars for each food application. And genomic prediction can estimate phenotypic variations for targeted traits in progenies derived from certain cross-combinations between two parental cultivars (Sekine and Yabe 2020). Successful genomic prediction and subsequent breeding selection have been reported in many crop species

such as wheat, maize, cassava, rice, chickpea, and groundnut (Budhlakoti et al. 2022; Onogi 2023). By using the same genome sequence information and phenotype data as used in GWAS in this study, we could perform genomic prediction for grain appearance, eating and cooking quality traits in the present rice cultivars. Based on the GWAS, PLS regression and genomic prediction procedures in further breeding programs, breeders could cross appropriate parental cultivars and select progenies having desirable genotypes. Accumulation of mutant gene alleles and creation of novel functional gene alleles by genome editing technology are also powerful tools to improve quantitative traits such as grain yield and grain quality. The combination of these genome and gene-based breeding designs will contribute greatly to the efficient development of novel rice cultivars.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00122-025-04850-x.

Acknowledgements We are grateful to the technical staff of the Field Management Division at the NARO for their assistance in paddy field work, and to Ms. Y. Nakamura, Ms. C. Yazawa, Ms. T. Takahashi and Ms. M. Iizumi at the National Institute of Crop Science, NARO for their kind assistance in all experiments.

Author contribution statement KH, KeS and YoshimT designed the experiments. KH, KeS, KM, KoS, KI, NK, KaS, KO, YK, TU, YoshinT and KE developed plant materials, cultivated the plants and evaluated their phenotypes. MS, KM, JY and MY performed GWAS and PLS analysis. KH, MS, KM and YoshimT analyzed the whole data and wrote the manuscript. All authors read the manuscript and approved it for publication.

Funding This study was supported by grants from the Ministry of Agriculture, Forestry, and Fisheries of Japan (grant numbers QTL4010 to YT, KS and KH, NVR1001 to JY and EK, RBS2011 to YT, KS and KH, IVG2003 to JY and EK, BAC2001 to YT and MY, BAC2002 to KH, DIT2001 to KH and KS, JP J012037 (SAM1-2–3) to KH, MS and KS), by the Project of the NARO Bio-oriented Technology Research Advancement Institution (25035B, 28014B to KH and KI) and by the Strategic Innovation Promotion Program from the Cabinet Office (DDB2007 to KH, KS and NK). This study was also supported by Grants-in-Aid for Scientific Research (C) (grant numbers 16K07564 to KH and 19K05984 to KH, KM, KO, YK and YT) from the Japan Society for the Promotion of Science (JSPS).

Data availability All datasets supporting the conclusions of this article are included in the article and supplementary files. NGS data of each rice cultivar and plant materials are available from public databases. NGS data are all from previous studies and are available at DNA Data bank of Japan (DDBJ; https://www.ddbj.nig.ac.jp) with the following accession numbers: PRJDB7213, DRR240814 to DRR240859, DRR 095338, DRR095344, DRR095348, DRR095346, SRR14366946, SRR14269765, SRR1630927, SRR13165580, SRR13165641, SRR14140390, SRR14140391, DRR289340 and DRR190941.

Declarations

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors have no relevant financial or non-financial interests to disclose.

References

Aktas C (2020) haplotypes: Manipulating DNA sequences and estimating unambiguous haplotype network with statistical parsimony. Comprehensive R Archive Network

Alonge M, Lebeigle L, Kirsche M, Jenike K, Ou S, Aganezov S, Wang X, Lippman ZB et al (2022) Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. Genome Biol 23:258

Aoki N, Umemoto T, Hamada S, Suzuki K, Suzuki Y (2012) The amylose content and amylopectin structure affect the shape and hardness of rice bread. J Appl Glycosci 59:75–82

Bao JS, Zheng XW, Xia YW, He P, Shu QY, Lu X, Chen Y et al (2000) QTL mapping for the paste viscosity characteristics in rice (*Oryza sativa* L.). Theor Appl Genet 100:280–284

Bollinedi H, Yadav AK, Vinod KK, Gopala Krishnan S, Bhowmick PK, Nagarajan M, Neeraja CN et al (2020) Genome-wide association study reveals novel marker-trait associations (MTAs) governing the localization of Fe and Zn in the rice grain. Front Genet 11:213

Budhlakoti N, Kushwaha AK, Rai A, Chaturvedi KK, Kumar A, Pradhan AK, Kumar U, Kumar RR et al (2022) Genomic selection: A tool for accelerating the efficiency of molecular breeding for development of climate-resilient crops. Front Genet 13:832153

Cai Y, Li S, Jiao G, Sheng Z, Wu Y, Shao G, Xie L, Peng C et al (2018) Os2 encodes a plastidic pyruvate kinase involved in rice endosperm starch synthesis, compound granule formation and grain filling. Plant Biotechnol J 16:1878–1891

Calingacion M, Laborte A, Nelson A, Resurreccion A, Concepcion JC, Daygon VD, Mumm R, Reinke R et al (2014) Diversity of global rice markets and the science required for consumer-targeted rice breeding. PLoS ONE 9:e85106

Champagne ET, Bett KL, Vinyard BT, McClung AM, Barton FE, Moldenhauer K, Linscombe S, McKenzie K (1999) Correlation between cooked rice texture and rapid visco analyser measurements. Cereal Chem 76:764–771

Chang ALS, Raber I, Xu J, Li R, Spitale R, Chen J, Kiefer AK, Tian C et al (2015) Assessment of the genetic basis of rosacea by genome-wide association study. J Invest Dermatol 135:1548–1555

Chikubu S, Iwasaki T, Tani T (1960) Studies on cooking and eating qualities of white rice. (Part 1). J Jpn Soci Food Nutr 13:137–140

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X et al (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly 6:80–92

Custodio MC, Cuevas RP, Ynion J, Laborte AG, Velasco ML, Demont M (2019) Rice quality: how is it defined by consumers,

industry, food scientists, and geneticists? Trends Food Sci Technol 92:122–137
Deng Z, Liu Y, Gong C, Chen B, Wang T (2022) Waxy is an important factor for grain fissure resistance and head rice yield as revealed by a genome-wide association study. J Exp Bot 73:6942–6954
Ebana K, Kojima Y, Fukuoka S, Nagamine T, Kawase M (2008) Development of mini core collection of Japanese rice landrace. Breed Sci 58:281–291
Fan CC, Yu XQ, Xing YZ, Xu CG, Luo LJ, Zhang Q (2005) The main effects, epistatic effects and environmental interactions of QTLs on the cooking and eating quality of rice in a doubled-haploid line population. Theor Appl Genet 110:1445–1452
Farooq MA, Murtaza MA, Aadil RM, Arshad R, Rahaman AAO, Siddique R, Hassan S, Akhtar HMS et al (2021) Investigating the structural properties and in vitro digestion of rice flours. Food Sci Nutr 9:2668–2675
Fitzgerald MA, McCouch SR, Hall RD (2009) Not just a grain of rice: the quest for quality. Trends Plant Sci 14:133–139
Hamazaki K, Iwata H (2020) RAINBOW: Haplotype-based genome-wide association study using a novel SNP-set method. PLOS Comput Biol 16:e1007663
Hao Y, Wang Y, Wu M, Zhu X, Teng X, Sun Y, Zhu J, Zhang Y et al (2019) The nuclear-localized PPR protein OsNPPR1 is important for mitochondrial function and endosperm development in rice. J Exp Bot 70:4705–4720
Hori K, Sun J (2022) Rice grain size and quality. Rice 15:33
Hori K, Yano M (2013) Genetic improvement of grain quality in *japonica* rice. Wiley Blackwell, Hoboken
Hori K, Ogiso-Tanaka E, Matsubara K, Yamanouchi U, Ebana K, Yano M (2013) *Hd16*, a gene for *casein kinase I*, is involved in the control of rice flowering time by modulating the day-length response. Plant J 76:36–46
Hori K, Suzuki K, Iijima K, Ebana K (2016) Variation in cooking and eating quality traits in Japanese rice germplasm cultivars. Breed Sci 66:309–318
Hori K, Suzuki K, Ishikawa H, Nonoue Y, Nagata K, Fukuoka S, Tanaka J (2021) Genomic regions involved in differences in eating and cooking quality other than Wx and Alk genes between *indica* and *japonica* rice cultivars. Rice 14:8
Hori K, Okunishi T, Nakamura K, Iijima K, Hagimoto M, Hayakawa K, Shu K, Ikka T et al (2022) Genetic background negates improvements in rice flour characteristics and food processing properties caused by a mutant allele of the PDIL1-1 seed storage protein gene. Rice 15:13
Iijima K, Suzuki K, Hori K, Ebana K, Kimura K, Tsujii Y, Takano K (2019) Endosperm enzyme activity is responsible for texture and eating quality of cooked rice grains in Japanese cultivars. Biosci Biotechnol Biochem 83:502–510
Japan Rice Market (2023) Kome hinshu taizen 7. Rice Databank Company, Tokyo
Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, Schwartz DC, Tanaka T et al (2013) Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. Rice 6:1–10
Kham NM, Kanamori H, Wu J, Matsumoto T, Fujita D, Yasui H, Yoshimura A, Yamagata Y (2024) Resistance haplotypes to green rice leafhopper (*Nephotettix cincticeps* Uhler) estimated in genome-wide association study in Myanmar *indica* rice landraces. Breed Sci 74:366–381
Kobayashi A, Tomita K, Yu F, Takeuchi Y, Yano M (2008) Verification of quantitative trait locus for stickiness of cooked rice and amylose content by developing near-isogenic lines. Breed Sci 58:235–242
Kole C (2020) Genomic designing of climate-smart cereal crops. Springer Nature, Cham
Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM (2017) Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–736
Kwon SW, Cho YC, Lee JH, Suh JP, Kim JJ, Kim MK, Choi IS, Hwang HG et al (2011) Identification of quantitative trait loci associated with rice eating quality traits using a population of recombinant inbred lines derived from a cross between two temperate *japonica* cultivars. Mol Cells 31:437–445
Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–2993
Li D, Liu CM, Luo R, Sadakane K, Lam TW (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:1674–1676
Li P, Chen YH, Lu J, Zhang CQ, Liu QQ, Li QF (2022) Genes and their molecular functions determining seed structure, components, and quality of rice. Rice 15:18
Li X, Xu L, Zhang J (2019) Improving the thread scalability and parallelism of BWA-MEM on Intel HPC platforms. In Proceedings of the 2019 IEEE 21st International Conference on High Performance Computing and Communications, pp1858–1865
Masutomi Y, Takimoto T, Shimamura M, Manabe T, Arakawa M, Shibota N, Ooto A, Azuma S et al (2019) Rice grain quality degradation and economic loss due to global warming in Japan. Environ Res Commun 1:121003
Meng T, Zhang X, Ge J, Chen X, Zhu G, Chen Y, Zhou G, Wei H et al (2022) Improvements in grain yield and nutrient utilization efficiency of *japonica* inbred rice released since the 1980s in eastern China. Field Crops Res 277:108427
Mikami T (2009) Development of evaluation systems for rice taste quality. Jpn J Food Eng 10:191–197
Mogga M, Sibiya J, Shimelis H, Lamo J, Yao N (2018) Diversity analysis and genome-wide association studies of grain shape and eating quality traits in rice (*Oryza sativa* L.) using DArT markers. PLoS ONE 13:e0198012
Nagata K, Nonoue Y, Matsubara K, Mizobuchi R, Ono N, Shibaya T, Ebana K, Ogiso-Tanaka E et al (2023) Development of 12 sets of chromosome segment substitution lines that enhance allele mining in Asian cultivated rice. Breed Sci 73:332–342
Nakanishi A, Tamura K, Kataoka T, Sato H, Tamura Y, Sakai M, Fushimi T, Takeuchi Y (2022) 'Emitawawa', a new rice cultivar with good properties for rice bread and high grain yield for the warm and temperate regions in Japan. Breed Res 24:160–167
Nayyeripasand L, Garoosi GA, Ahmadikhah A (2021) Genome-wide association study (GWAS) to identify salt-tolerance QTLs carrying novel candidate genes in rice during early vegetative stage. Rice 14:9
Obenchain V, Lawrence M, Carey V, Gogarten S, Shannon P, Morgan M (2014) VariantAnnotation : a bioconductor package for exploration and annotation of genetic variants. Bioinformatics 30:2076–2078
Okadome H (2005) Application of instrument-based multiple texture measurement of cooked milled-rice grains to rice quality evaluation. JARQ 39:261–268
Onogi A (2023) A Bayesian model for genomic prediction using metabolic networks. Bioinform Adv 3:vbad106
Park SG, Park HS, Baek MK, Jeong JM, Cho YC, Lee GM, Lee CM, Suh JP et al (2019) Improving the glossiness of cooked rice, an important component of visual rice grain quality. Rice 12:87
Patindol JA, Siebenmorgen TJ, Wang YJ (2015) Impact of environmental factors on rice starch structure: a review. Starch 67:42–54
Peng C, Wang Y, Liu F, Ren Y, Zhou K, Lv J, Zheng M, Zhao S et al (2014) FLOURY ENDOSPERM6 encodes a CBM48 domain-containing protein involved in compound granule formation and starch synthesis in rice endosperm. Plant J 77:917–930

Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, Kling DE, Gauthier LD et al. (2018) Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv 201178

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575

Raj A, Stephens M, Pritchard JK (2014) fastSTRUCTURE: Variational inference of population structure in large SNP data sets. Genetics 197:573–589

Ren D, Ding C, Qian Q (2023) Molecular bases of rice grain size and quality for optimized productivity. Sci Bull (Beijing) 68:314–350

Sakai H, Lee SS, Tanaka T, Numa H, Kim J, Kawahara Y, Wakimoto H, Yang C et al (2013) Rice annotation project database (RAP-DB): an integrative and interactive database for rice genomics. Plant Cell Physiol 54:e6–e6

Sekine D, Yabe S (2020) Simulation-based optimization of genomic selection scheme for accelerating genetic gain while preventing inbreeding depression in onion breeding. Breed Sci 70:594–604

Shahbandeh M (2024) Total rice consumption worldwide from 2008/2009 to 2023/2024. Statistica https://www.statista.com/statistics/255971/top-countries-based-on-rice-consumption-2012-2013/. Accessed 2024–09–06

Shin JH, Blay S, McNeney B, Graham J (2006) LDheatmap: An R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. J Stat Softw 16:1–9

Shiraishi M (1994) Study on breeding for amylose content in rice endosperm. Bull Oita Pref Agri Res Cent 24:91–135

Sreenivasulu N, Zhang C, Tiozon RN, Liu Q (2022) Post-genomics revolution in the design of premium quality rice in a high-yielding background to meet consumer demands in the 21st century. Plant Commun 3:100271

Stanke M, Schöffmann O, Morgenstern B, Waack S (2006) Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC Bioinform 7:1–11

Sun Y, Shi Y, Liu G, Yao F, Zhang Y, Yang C, Guo H, Liu X et al (2020) Natural variation in the OsbZIP18 promoter contributes to branched-chain amino acid levels in rice. New Phytol 228:1548–1558

Takeuchi Y, Hori K, Suzuki K, Nonoue Y, Takemoto-Kuno Y, Maeda H, Sato H, Hirabayashi H et al (2008) Major QTLs for eating quality of an elite Japanese rice cultivar, Koshihikari, on the short arm of chromosome 3. Breed Sci 58:437–445

Tanaka N, Shenton M, Kawahara Y, Kumagai M, Sakai H, Kanamori H, Yonemaru J, Fukuoka S et al (2020) Investigation of the genetic diversity of a rice core collection of Japanese landraces using whole-genome sequencing. Plant Cell Physiol 61:2087–2096

Tian Z, Qian Q, Liu Q, Yan M, Liu X, Yan C, Liu G, Gao Z et al (2009) Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. Proc Natl Acad Sci USA 106:21760–21765

Turner SD (2018) qqman: an R package for visualizing GWAS results using QQ and manhattan plots. J Open Source Softw 3:731

Umemoto T, Yano M, Satoh H, Shomura A, Nakamura Y (2002) Mapping of a gene responsible for the difference in amylopectin structure between *japonica*-type and *indica*-type rice varieties. Theor Appl Genet 104:1–8

UNICEF (2024) The State of Food Security and Nutrition 2024. http://icsfarchives.net/20253/. Accessed 2024–09–06

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, Jordan T, Shakir K et al (2013) From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Curr Protoc Bioinformatics 43:11.10.1–11.10.33

Verma RK, Chetia SK, Sharma V, Baishya S, Sharma H, Modi MK (2022) GWAS to spot candidate genes associated with grain quality traits in diverse rice cultivars of North East India. Mol Biol Rep 49:5365–5377

Wada T, Ogata T, Tsubone M, Uchimura Y, Matsue Y (2008) Mapping of QTLs for eating quality and physicochemical properties of the *japonica* rice Koshihikari. Breed Sci 58:427–435

Wada T, Miyahara K, Sonoda J, Tsukaguchi T, Miyazaki M, Tsubone M, Ando T, Ebana K et al (2015) Detection of QTLs for white-back and basal-white grains caused by high temperature during ripening period in *japonica* rice. Breed Sci 65:216–225

Wang C, Shu Q (2007) Fine mapping and candidate gene analysis of purple pericarp gene Pb in rice (*Oryza sativa* L.). Sci Bull 52:3097–3104

Wang ZY, Zheng FQ, Shen GZ, Gao JP, Snustad DP, Li MG, Zhang JL, Hong MM (1995) The amylose content in rice endosperm is related to the post-transcriptional regulation of the waxy gene. Plant J 7:613–622

Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z, Li M, Zheng T et al (2018) Genomic variation in 3,010 diverse cultivars of Asian cultivated rice. Nature 557:43–49

Wang H, Zhu S, Dang X, Liu E, Hu X, Eltahawy MS, Zaid IU, Hong D (2019) Favorable alleles mining for gelatinization temperature, gel consistency and amylose content in *Oryza sativa* by association mapping. BMC Genet 20:34

Xu X, Ye J, Yang Y, Li R, Li Z, Wang S, Sun Y, Zhang M, Xu Q, Feng Y, Wei X, Yang Y (2022) Genetic diversity analysis and GWAS reveal the adaptive loci of milling and appearance quality of *japonica* rice (*Oryza sativa* L.) in Northeast China. J Integr Agric 21:1539–1550

Yabe S, Yoshida H, Kajiya-Kanegae H, Yamasaki M, Iwata H, Ebana K, Hayashi T, Nakagawa H (2018) Description of grain weight distribution leading to genomic selection for grain-filling characteristics in rice. PLoS ONE 13:e0207627

Yamasaki M, Ideta O (2013) Population structure in Japanese rice population. Breed Sci 63:49–57

Yano K, Yamamoto E, Aya K, Takeuchi H, Lo P, Hu L, Yamasaki M, Yoshida S et al (2016) Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. Nat Genet 48:927–934

Yoshida H, Okada S, Wang F, Shiota S, Mori M, Kawamura M, Zhao X, Wang Y et al (2023) Integrated genome-wide differentiation and association analyses identify causal genes underlying breeding-selected grain quality traits in *japonica* rice. Mol Plant 16:1460–1477

Yoshimoto M, Sakai H, Ishigooka Y, Kuwagata T, Ishimaru T, Nakagawa H, Maruyama A, Ogiwara H, Nagata K (2021) Field survey on rice spikelet sterility in an extremely hot summer of 2018 in Japan. J Agric Meteorol 77:262–269

Yu J, Pressoir G, Briggs WH, Vroh BI, Yamasaki M, Doebley JF, McMullen MD, Gaut BS et al (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208

Yu E, Wang W, Yamaji N, Fukuoka S, Che J, Ueno D, Ando T, Deng F et al (2022) Duplication of a manganese/cadmium transporter gene reduces cadmium accumulation in rice grain. Nat Food 3:597–607

Zhang L, Deng B, Peng Y, Gao Y, Hu Y, Bao J (2024) Population structure and genetic diversity of Shanlan landrace rice for GWAS of cooking and eating quality traits. Int J Mol Sci 25:3469

Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics 28:3326–3328

Zhou X, Stephens M (2012) Genome-wide efficient mixed-model analysis for association studies. Nat Genet 44:821–824