Genome-wide Association Study-based Identification Genes Influencing Agronomic Traits in Rice (Oryza Sativa L.)

aijun wang
State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China, Chengdu, China 2Rice Research Institute of Sichuan Agricultural University, Chengdu, China 3Key laboratory of Sichuan Crop Major Disease, Sichuan Agricultural University, Chengdu, China https://orcid.org/0000-0002-2985-5796

Xinyue Shu
Sichuan Agricultural University

yuqi Jiang
Sichuan Agricultural University

Li Ma
Sichuan Agricultural University

xiaomei jia
Sichuan Agricultural University

xiaoying ye
Sichuan Agricultural University

qiming deng
Sichuan Agricultural University

shuangcheng Li
Sichuan Agricultural University

shiquan Wang
Sichuan Agricultural University

jun zhu
Sichuan Agricultural University

yueyang liang
Sichuan Agricultural University

ting zou
Sichuan Agricultural University

Linxia Wang
Sichuan Agricultural University

jianqing zhu
Sichuan Agricultural University

Ping Li
Sichuan Agricultural University

Aiping Zheng (apzh0602@gmail.com)

State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China, Chengdu, China 2Rice Research Institute of Sichuan Agricultural University, Chengdu, China 3Key laboratory of Sichuan Crop Major Disease, Sichuan Agricultural University, Chengdu, China

---

Original article

Keywords: single nucleotide polymorphisms(SNPs), best linear unbiased prediction (BLUP), linkage disequilibrium (LD)

Posted Date: September 16th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-74749/v1

Version of Record: A version of this preprint was published at Genomics on May 1st, 2021. See the published version at https://doi.org/10.1016/j.ygeno.2021.03.016.

Abstract

Background

Rice (*Oryza sativa* L.) is one of the most important cereal crops, providing the daily dietary intake for approximately 50% of the global human population. To needs of the rapidly increasing human population worldwide, cultivation of rice varieties with high yield and quality, more genes or QTLs association with yield traits are required.

Results

Currently, correlations among different traits and gene interactions both affect the rice breeding. Here, we re-sequenced 259 rice accessions, generating 1, 371.65 Gb of raw data. Furthermore, we performed genome-wide association studies (GWAS) on 13 agronomic traits using 2.8 million single nucleotide polymorphisms (SNPs) characterized in 259 rice accessions. Phenotypic data and best linear unbiased prediction (BLUP) values of each of the 13 traits over two years of each trait were used for GWAS. The result showed that 816 SNP signals were significantly associated ($-\log10P \geq 5$) with the 13 agronomic traits. We detected candidate genes related to target traits within 200 kb upstream and downstream of the associated SNP loci, based on linkage disequilibrium (LD) blocks in the whole rice genome. These candidate genes were further identified though haplotype block construction.

Conclusions

This study provides an important genomic resource and valuable new information for breeding high yielding breeding rice cultivars through genomic selection.

Background

Rice (*Oryza sativa* L.) is a staple food crop that meets the daily dietary intake requirement of approximately 50% of global population (Chen et al., 2014). To meet the needs of the rapidly increasing human population worldwide, cultivation of rice varieties with high yield and quality is imperative (Li et al., 2006). Therefore, development of high-yielding cultivars has become one of the main goals of rice breeding. Grain yield is determined by multiple yield-related traits, such as grain size, grain number per panicle, and panicles number per plant. Furthermore, the size and shape of the flag leaf, which is the main source of carbohydrates, influence grain yield by affecting photosynthetic efficiency (Yang and Yang 1998; Yue et al., 2006). Therefore, studying yield-related agronomic traits is highly important for improving rice yield and quality as well as for understanding the rice domestication process.

Molecular biology tools are powerful and effective means of breeding highly-yielding rice cultivars; however, successful use of these tools requires a better understanding of the genetic architecture of key agronomic traits (Miura et al., 2011). Therefore, determining the genetic basis of agronomic traits is our top priority for accelerating the development of novel rice varieties. Several quantitative trait loci (QTLs) or genes affecting grain and plant shape, biotic and abiotic stresses responses, and growth habits have been identified in rice (Yano et al., 2016; Guo et al., 2018; Zhou et al., 2017; Liu et al., 2019). For instance, serine carboxypeptidase

encoded gene GS5, which positively regulates grain size (Duan et al., 2017). Two QTLs, qRES-5 and qRES-10, detected in the recombinant inbred line (RIL) population (derived from a cross between the wild rice accession W1944 and the *indica* variety 'Pei-Kuh'), explained 9.8% and 24.8% of the phenotypic variance, respectively, in the stigma exsertion rate (Uga et al., 2003a, 2003b). In a doubled-haploid population derived from a cross between Zhaiye Qing 8 and Jingxi 17, Li et al. (2000) detected 13 QTLs that explained 8.7%~18.5% of the phenotypic variation in the length, width, area, and length-width ratio of the flag leaf. However, our understanding of the genetic regulation of agronomic traits remains limited because most of these traits are naturally adapted into complex traits.

Genome-wide association study (GWAS), a powerful approach for mapping QTL associated with complex agronomic traits, has been successfully used in many plant species, including *Arabidopsis thaliana*, rice, cotton (*Gossypium hirsutum*), and soybean (*Glycine max*) (Du et al., 2018; Li et al., 2017; Zhang et al., 2017). Fang et al. (2017) identified 245 significant genetic loci associated with 84 agronomic traits through GWAS of 809 soybean accessions. Among these 245 loci, 95 showed genetic interactions with other loci, and 14 oil synthesis-related genes were found to be responsible for fatty acid accumulation in soybean. Similarly, in upland cotton, candidate genes of affecting fiber quality and yield were also detected by GWAS using 3.66 million SNPs obtained by sequencing 419 accessions (Ma et al., 2018). In rice, GWAS has been used to identify candidate genes of affecting disease resistance, yield, and plant type (Zhang et al., 2019; Zhan et al., 2019; Huang et al. 2010). Huang et al. (2010) identified four new genes associated with agronomic traits in rice using GWAS based on whole-genome sequencing. A core collention comprising 584 rice accessions was evaluated for resistance against three *Magnaporthe oryzae* strains, and the rice blast resistance gene, PiPR1, was identified through GWAS using 700,000 high-density SNPs (Liu et al., 2019).

Since different traits are correlated, resulting in heritable covariation, relying on the identification of the specific trait is insufficient for molecular breeding (Klingenberg, 2008; Wagner, 1996; Chen and Lubberstedt, 2010). For example, grain yield and protein content show a negative correlation, and simultaneously increasing these two desirable traits is difficult, if not impossible, in most crops (Duvick and Cassman, 1999; Rotundoa et al., 2009; Rharrabti et al., 2001). Therefore, quantifying and understanding the covariation among traits is essential for breeding projects that target multiple complex traits (Melo and Marroig, 2015).

In this study, we re-sequenced a worldwide collection of 259 rice germplasm accessions, thus generating approximately 2.88 million SNPs. Rice phenotyping was conducted over two years, and GWAS was performed on 13 agronomic traits (grain length, grain width, grain weight, panicle length, spikelet number, grain number per spike, percentage of full grains, plant height, flag-leaf length, flag-leaf width, flag-leaf angle, heading date, and awn presence or absence), which revealed that 816 SNPs were significantly associated with these traits. Our results provide important genomic resources for rice molecular breeding and lay a strong foundation for studying the basis of high yield and quality in rice.

Results

Genomic variation and population structure

To fine map the target trait-related genes, we sequenced 259 rice varieties at an average coverage of 12.16 × using the Illumina HiSeq platform (Figure 1A and Table S1), generating 1371.65 Gb of raw reads (Table S2). The sequence reads were mapped on to the Nipponbare reference genome sequence (ftp://ftp. ensemblgenomes. org/ pub/ plants/ release_36/ fasta/ oryza_indica/ dna/), and 2,888,332 high-confidence SNPs (missing data < 20%; minor allele frequency [MAF]>5%) were detected (Table S3). Of these, 464,911, 555,884, 1,146,191, 363,883, and 318,546 SNPs were located in introns, exons, intergenic regions, upstream regions, and downstream regions, respectively (Figure S1 and Table S3). SNP annotation revealed 2,850 splicing, 312,857 nonsynonymous, 1,085 stop-loss, and 16,416 stop gain SNPs in coding sequences (CDSs) (Table S3). The nonsynonymous/synonymous SNP ratio (Ks/Ks) was 1.39, this value is much higher than that reported in maize (*Zea mays*) (Jiao et al., 2012) and barley (*Hordeum vulgare*) (Russell et al., 2016), but lower than that in cotton (Ma et al., 2018). The chromosomal distribution of SNPs ranged from 825,160 SNPs on Chr. 1 to 519,313 SNPs on Chr. 9 (Table S4). Additionally, SNP frequency was the highest on Chr. 11 (25.67 SNPs/kb) (Figure S1 and Table S4). Thus, the SNP data set generated using 259 accessions serves as valuable resources for the molecular improvement of target traits in rice. Figure 1B shows the phylogenetic tree of 259 rice core verities. Population-structure analysis showed that these 259 rice accessions comprised three genetic groups (K=3): *japonica* cultivars (19), landrace (105) and improved *indica* cultivars (151) (Figure 1C and Table S1). Principal component analysis (PCA) showed the first two principal components explained 29.54% of the genetic variation within these 259 rice accessions (Figure 1D). Additionally, the nucleotide diversity of landraces ($\pi=1.22\times10^{-3}$) and improved *indica* cultivars ($\pi=1.26\times10^{-3}$) was significantly less than that of *japonica* cultivars ($\pi=6.91\times10^{-3}$) (Figure 1E) but were similar to the previously reported nucleotide diversity of *indica* cultivars ($\pi=1.6\times10^{-3}$) (Huang et al., 2010). The fixation index (F_{ST}) between landraces and improved indica cultivars was 0.09 (Figure 1E), which is higher than the F_{ST} between landraces and improved modern cultivars in cotton ($F_{ST}=0.04$) (Ma et al., 2018), and lower than that in maize ($F_{ST}=0.14$) (Jiao et al., 2012) and wheat (*Triticum aestivum*; $F_{ST} = 0.15$) (Hurst, 2009). These data indicate that rice accessions selected in this study represent abundant genetic variation. Based on the r^2 value, which declined to half of the maximum value, the linkage disequilibrium (LD) decay rate was estimated at 230 kb for *indica* cultivars, which was lower than that for *japonica* rice (Figure 1F), indicating that rice accessions used in this study exhibit moderate LD (Huang et al., 2010).

Phenotypic variation among rice accessions

We measured 13 agronomic traits in 259 accessions over two years at Chengdu, of China. These traits could be divided into three groups; grain shape-related traits (grain length, grain width, and grain weight), plant type-related traits (plant height, flag-leaf length, flag-leaf width, and flag-leaf angle), and panicle related traits (panicle length, spikelet number, grain number per spike, percentage of full grains, heading date, and awn presence or absence). All of these traits, except awn presence or absence, showed large variations and exhibited a normal distribution (Figure S2). Pearson's correlation coefficients of the same traits were significantly between the two years (0.83–1) (Figure 2). Among these traits, grain length was negatively correlated with grain width and percentage of full grains, but positively correlated with grain weight and panicle length (Figure 2). Plant height was positively correlated with flag-leaf length and flag-leaf angle but showed no significant correlation with flag-leaf width. Furthermore, several panicle-related traits showed significant correlation with plant type-related traits; for example, panicle length was positively correlated with

plant height, flag leaf length and flag leaf width; whereas grain number per spike was negatively correlated with flag leaf angle (Figure 2). Additionally, all 13 agronomic traits showed distinct distributions within the three sub-populations. Improved *indica* cultivars showed significantly longer and thinner grains, higher grain number per spike, and longer panicles than *japonica* cultivars and landraces cultivar of *indica* groups (Figure S3). Interestingly, variation in plant height within landraces *indica* cultivar or *japonica* varieties was larger than that in improved indica cultivars (Figure S3). These results indicate that rice accessions selected in this study show abundant genetic diversity, and artificial selection plays an important role in breeding high-yielding varieties.

GWAS of 13 agronomic traits

Based on 2,888,332 high-confidence SNPs, GWAS on 13 agronomic traits was performed using a mixed linear model (MLM). Phenotypic data and best linear unbiased prediction (BLUP) values of 13 agronomic traits over two years were used for GWAS. The results revealed that there are 28, 46, 18, 14, 11, 14, 16, 156, 198, 15, 16, 27, and 257 SNP loci (common to at least two values for each trait) were significantly associated ($-\log10P \geq 5$) with grain length, grain width, grain weight, panicle length, grain number per spike, flag-leaf length, flag-leaf width, plant height, flag-leaf angle, percentage of full grains, spikelet number, heading date, and awn traits, respectively (Figure S4–16, Table S5–17). Plant height showed the highest number of significantly associated SNPs, followed by heading date and awn, and the number of SNP loci association with the three grain shape-related traits was less than that associated with plant type- and panicle-related traits (Figure S4–16, Table S5–17). Additionally, we detected candidate genes within 200 kb upstream and downstream of the associated SNP loci, based on our results of the number and genome coverage of LD blocks in the whole genome (Figure 1F). The number of genes association with the 13 agronomic traits varied 422 to 7,820 (Table S5-17). Genes associated with important yield-related traits (grain length, grain width, and grain weight) were mostly located on Chr. 3 (Figure S4-6); however, those associated with panicle length was mostly located on Chr. 6, Chr. 9, and Chr. 12 (Figure S7).

We further compared the previously reported genes known to function in the 13 agronomic traits with our results, and found that several known trait-related genes were included in this study, such as genes related to grain length D2, GE, and GL7 (Table S5); grain width-related genes FUWA, GW2, BG1, qGL6, GW5 and OsPUP7 (Table S6); grain weight-related genes FUWA, GL3.2, d11, RSR1, APG, GSK2, OsPPKL2, OsGSR1, OsbZIP47, DEP3, GE, GL7, OsEIN2, DEP2, SG1, and Brd2 (Table S7); panicle length related gene LP1, Hd1, ASP1, and DEP1 (Table S8); grain number per spike and SPP1, Gn1a, and Ghd7 (Table S9); heading date gene Hd17, Hd3a, HGW, Hd1, DEP, and Ghd7 (Table S10). All of these genes were significantly associated with the corresponding trait. Notably, peak signals of GWAS loci often appeared near (but not within) the known genes because of local LD. These results were consistent with findings in previous studies and may result from multiple causal polymorphisms within a gene, coupled with the complex population structure (Huang et al., 2010). This shows that our data are accurate and reliable, indicating a high possibility of mining novel function genes.

Identification of candidate genes affecting grain-related traits

Grain length, width, and weight are important yield traits. A total of 29, 47, and 59 SNPs showed significant association with grain length, grain width, and grain weight, respectively, in this study. Among the 29 grain length-associated SNPs, the significant SNPs were located on Chr.3 and Chr. 7 (Figure 3A). The known grain length-related genes D2 (Fang et al., 2016), GL7 (Wang et al., 2015), and GE (Xu et al., 2015) were located near significant SNPs on Chr. 1 and Chr. 7 (Figure 3A). Furthermore, rice genes including LOC_Os03g37810 and LOC_Os03g38010, which encode an expressed protein and a Nuf2 family protein, respectively, were related to grain length (Figure 3A), which is consistent with the results of Huang et al. (2012). Based on these results, we further analyzed candidate genes harboring significant SNPs on Chr.3 (Chr.3-16682908, -Log10P=9.81) and Chr.7 (Chr.7-363388, -Log10P=6.26). Analysis of the haplotype block structure of 200 kb flanking Chr.3-16682908 on either side showed that candidate gene regions were included within a 30kb (16,864-16,894 Kb) haplotype block encompassing by 68 SNPs (Figure 3B) and three genes (LOC_Os03g29600, LOC_Os03g29614, and LOC_Os03g29630). Among these three genes, LOC_Os03g29630, which encodes the ulp1 protease family protein, contained three nonsynonymous SNPs (Chr.3_16890883, Chr.3_16891448, and Chr.3_16891826) that formed two haplotypes (Figure 3C). Rice accessions harboring the alternate CGG (reference) sequence showed significantly longer grains than those harboring the TCT (alternate) sequence (Figure 3D). In Arabidopsis, the ulp1 protease family protein plays a key role in the process of flowering-time control (Anjum et al., 2013), which is very important for grain development. Interestingly, we previously fine-mapped LOC_Os03g29630 using a bi-parental mapping population (Li et al. 2012), which further confirms that this locus is responsible for controlling grain size in rice. Similarly, we analyzed the haplotype block structure of 200 kb flanking Chr.7-363388 on either side to determine candidate genes. We focused on the locus mapped to a 177-kb region containing 10 significantly associated SNPs and 28 genes (Figure 3E). LOC_Os07g01600, which encodes the dirigent protein, contained three nonsynonymous SNPs (Chr.07_374057, Chr.07_374895, and Chr.07_374990) that formed two haplotypes (Figure 3F). Rice accessions carrying the CCC (alternate) allele showwed significantly shorter grains than those carrying the TTT (reference) allele (Figure 3G). This implies that LOC_Os07g01600 is also involved in the regulation of grain size in rice.

Furthermore, based on the results of the comparison of candidate genes identified in this study with the previously known grain width-related genes (Figure 3H), we selected a new significant SNP (Chr.3-8716055; -Log10P = 6.67) associated with grain width to analyze candidate genes. Haplotype block structure analysis of 200-kb sequence flanking Chr.3-8716055 on either side showed that the candidate gene region was included within a 202-kb haplotype block (8,738 -8,940 kb), containing 6 SNPs (Figure 3I) and 37 genes. Among these genes, LOC_Os03g15900, which encodes an SH3 domain-containing protein, contained a nonsynonymous SNP (Chr.3_8783467; T/G), which formed two hapllotypes (Figure 3J). Accessions carrying the GG (alternate) allele showed significantly lower grain width than those carrying the TT (reference) allele (Figure 3K). The SH3 domain-containing protein is involved in cell division in plants; for example, in Arabidopsis, the SH3 Domain-Containing Protein 2 (SH3P2) plays a crucial role in converting vesicles into the planar cell plate (Ahn G et al. 2017). We also performed haplotype block structure analysis of another SNP (Chr.02_24998871; G/A) significantly associated with grain weight, This locus did not contain any previously reported genes or QTLs related to grain weight (Figure 3L). Haplotype block structure analysis of this locus indicated that the candidate gene regions were included within a 260-kb (24,810-25,070 kb) haplotype block containing three SNPs (Figure 3M) and 27 genes. After the removal of genes encoding unknown proteins,

trotransposon, retrotransposons, 16 out of 27 genes were considered as candidates for further analysis. These 16 genes included LOC_Os02g41450 (no apical meristem protein, LOC_Os02g41460 (DUF640 domain-containing protein), LOC_Os02g41470 (class II tRNA synthetases domain-containing protein), LOC_Os02g41480 (OsWAK12), LOC_Os02g41500 (OsWAK) receptor-like cytoplasmic kinase, OsMYB30 (MYB transcription factor), LOC_Os02g41520 (glycosyl transferase 8 domain containing protein), LOC_Os02g41550 (FAD-binding domain of DNA photolyase domain-containing protein), LOC_Os02g41560 (pentatricopeptide), LOC_Os02g41580 (calcium/calmodulin depedent protein kinases CAMK_CAMK_like.14), LOC_Os02g41590 (kinase), LOC_Os02g41710 (cyclic nucleotide-gated ion channel), and OsPAL1-4 (phenylalanine ammonia lyases. Interestingly, the DUF640 domain-containing protein reportedly regulates the shape, size, and weight of rice grains by though controlling cell division and expansion (Wei et al. 2013; Cui et al. 2016). Therefore, we speculate that LOC_Os02g41460, encoding the DUF640 domain-containing protein, is directly involved in the regulation of grain weight in rice.

Identification of candidate genes affecting panicle-related traits

Panicle length, panicle number, total grain number per panicle, and percentage of full grains are important agronomic traits, as they directly affect crop yield in rice. Panicle length is one of the major components of rice panicle structure and is controlled by QTLs. Among the 14 panicle length-associated SNPs, the most significant SNPs were located on Chr. 1, Chr. 6, Chr. 9, and Chr. 12 (Figure 4A). The known panicle length-related genes, Hd1, DEP1, and LP1, coincided with significant SNPs on Chr. 6 and Chr. 9 (Figure 4A). Therefore, we analyzed candidate genes located near two significant SNPs on Chr. 1 (Chr. 1_19388090; -Log10P = 6.82) and Chr.12 (Chr.12_8864676; -Log10P = 8.28) (Figure 4A). Haplotype block structure analysis of 200-kb sequence flanking Chr. 1_19388090 on either side showed that the candidate region was included within 25-kb (19,382–19,407 kb) haplotype block (Figure 4B), which contained three genes (LOC_Os01g35030, LOC_Os01g35040, and LOC_Os01g35050) encoding an MDR-like ABC transporter (LOC_Os01g35030), a ZOS1-09-C2H2 zinc-finger protein (LOC_Os01g35040), and an early-responsive to dehydration protein (LOC_Os01g35050). According to previous reports, MDR/PGP-like ABC transporters play a crucial role in the transport of the plant hormone, auxin (Santelia et al. 2005). Haplotype block structure analysis of 200-kb sequence on either side of Chr.12_8864676 showed that the candidate region was included within a 60-kb (8,836 -8,896 kb) haplotype block (Figure 4C), which contained nine candidate genes (LOC_Os12g15480, LOC_Os12g15490, LOC_Os12g15500, LOC_Os12g15505, LOC_Os12g15510, LOC_Os12g15520, LOC_Os12g15530, LOC_Os12g15540, and LOC_Os12g15550). To investigate candidate genes affecting panicle number, we analyzed candidate genes located near the most significant SNP on Chr.6 (Chr.6_27177562; -Log10P = 7.41); Haplotype block structure analysis of 200kb flanking Chr.6_27177562 on either side showed that the candidate region was located within a 107-kb (27,118 -27,225 kb) haplotype block (Figure 4D), which contained 15 genes. Interestingly, one of those genes encoded the auxin efflux transporter OsPIN2. Overexpression of the OsPIN2 gene has been previously shown to reduce plant height and increase tiller number and tiller angle (Chen et al. 2012). Therefore, we speculate that this is involved in the regulation process of panicle number in rice.

Three previously known genes affecting grain number per panicle, SPP1, Gn1a, and Ghd7, were identified in this study (Figure 4E). We analyzed candidate genes located near a new significant SNP on Chr. 3 (Chr. 3_ 16776796; -Log10P = 10.09) (Figure 4E). Haplotype block structure analysis of 200-kb flanking Chr. 3_

16776796 on either side showed that the candidate region was located within a 120-kb (16,732–16,852 kb) haplotype block (Figure 4B). After the removal of genes encoding unknownproteins, trotransposon, and retrotransposons from this haplotype block, five genes were identified as candidate genes: LOC_Os03g29410, encoding a tyrosine protein kinase domain containing protein; LOC_Os03g29460, encoding the 60S ribosomal protein; LOC_Os03g29470, encoding the transcription initiation factor IID; LOC_Os03g29480, encoding the nodulation-signaling pathway 1 protein; LOC_Os03g29540, encoding the ATP-dependent protease. While the ATP-dependent protease is important for plant growth and development, it is particularly involved in maintaining mitochondrial homeostasis late in rosette development under short-day photoperiod (Janska 2010). To analyze candidate genes affecting the percentage of full grains, we selected the most significant SNP (Chr. 12_2272361; -Log10P = 7.39) for further analysis (Figure S10). Haplotype block structure analysis of 200-kb sequence flanking Chr. 12_2272361 on either side showed that the candidate gene region was included within a 30-kb (1,557 -1,587 kb) haplotype block (8,738 -8,940 kb), which contained five genes (Figure 4G), including LOC_Os12g03822, which encodes Rice Immature Pollen 1 (RIP1).It has been previously shown that RIP1 is necessary for pollen maturation and germination (Han et al. 2006), pollen germination is a basic of full grains. This indicates that RIP1 is function gene of this locus.

Identification of candidate genes affecting plant type-related traits

Some important plant type traits, such as flag-leaf length, width, and angle, also grain yield in rice by affecting photosynthesis (Yang and Yang 1998; Yue et al., 2006). Erect leaves increase both the growth rate and grain yield of rice plant (Zhang et al., 2015). Among the 14 SNPs that showed significant association with flag-leaf length, we analysis candidate genes located near a significant SNP on Chr.12 (Chr.12_23021423; -Log10P = 6.48) (Figure S12). Haplotype block structure analysis of Chr.12_23021423 showed that the candidate gene region was included within a 118-kb (22,906 -23,024 kb) haplotype block, containing 22 SNPs and 21 genes (Figure 5A). LOC_Os12g37430, which encodes PAPA-1-like conserved region family protein, harbored a nonsynonymous SNP (Chr.12_ 22973522; G/A) (Figure 5B). Accessions carrying the AA (alternate) allele showed significantly longer flag leaf than those carrying the GG (reference) allele (Figure 5C). To study flag-leaf width, we analyzed candidate genes located at a significant SNP on Chr.2 (Chr.2-29252196; -Log10P = 5.95). Haplotype block structure analysis of Chr.2-29252196 showed that the candidate gene region was included within a 230-kb (29,152 -29,382 kb) haplotype block encompassing eigth SNPs and 31 genes (Figure 5D). Further analysis showed that a SNP (Chr.2-29252252; C/G) was localized upstream of LOC_Os02g47820 (Figure 5E). Accessions carrying the TT (alternate) allele showed significantly wider flag leaf than those harboring the GG (reference) allele (Figure 5F). We also analyzed the haplotype block structure of Chr.12-19405426 (-Log10P = 7.86), which was significantly associated with flag-leaf width (Figure 5G). The results loci showed that the candidate gene region was located within 61-kb (19,379 -19,440 kb) haplotype block containing eight genes including LOC_Os12g32130 (encoding trehalose phosphatase), LOC_Os12g32150, LOC_Os12g32160, LOC_Os12g32170, LOC_Os12g32190, LOC_Os12g32200 (expressed proteins), LOC_Os12g32180 (cornichon protein), and LOC_Os12g32140 (transposon protein). Among these encoded proteins, trehalose phosphatase is linked to biological pathways involved in trehalose biosynthesis, which is part of glycan biosynthesis (Lannoo et al. 2014). Glycosylation is essential for the growth, development, and survival of all organisms, and plays important roles during the life cycle of a plant (Varki and Lowe, 2009). We

speculate that all of the eight abovementioned genes are directly involved in the regulation of flag leaf-related traits in rice.

Among the identified 27 heading date associated SNPs, the most significant SNPs were located on Chr.1 and Chr. 7 (Figure 5H). Because significant SNPs on Chr.7 coincided with known heading date-related genes, Ghd7 (Chr.7_ 9144304; -Log10P = 5.68) and DEP (Chr.7_472096; -Log10P = 8.35) (Figure 5H), we further analyzed candidate genes corresponding to a significant SNP on Chr.1 (Chr.1_22403310; -Log10P = 6.52). Haplotype block structure analysis of 200-kb sequence flanking Chr.1_22403310 on either side showed that the candidate region was included within 120-kb (22,408 -22,528 kb) haplotype block (Figure 5I), which contained 17 genes including LOC_Os01g39850, which encodes a heterotrimeric heme activator protein (HAP) family protein. In plants, HAP family proteins play a central role in the regulation of flowering; however, the number of HAP genes that regulate heading date in rice is unclear (Li et al., 2016).Bases on these data, we speculate that LOC_Os01g39850 is a functional gene at this locus, and is involved in the regulation of heading data in rice. Additionally, we mapped loci related to plant height and awn (Figure S11 and S17), and determined candidate genes underlying these loci (Table S12 and S17). Overall, these results enhance our understanding of the genetic basis of 13 agronomic traits in rice.

Allelic analysis suggests the impact of rice landraces and improved cultivars on yield traits

The large effects and broad impact of 13 agronomic traits-related QTLs identified in this study encouraged us to further detect their allelic variation during rice domestication and modern breeding. A marked impact of selection was observed at the identified candidate genes associated with grain length, LOC_Os03g29630, in *indica* landraces and improved_cultivars (Figure 6; Table S18). The CGG allele at LOC_Os03g29630, associated with long grain, and its amino acid profile was rare in *indica* landraces, which was successively selected during domestication and modern breeding. Furthermore, the candidate gene association with grain number per panicle, LOC_Os03g29540, also showed highly selected (Table S18). These results suggest that rice domestication and modern breeding extensively influenced the yield-related traits, especially grain size and number. This finding is consistent with the main objective of rice breeding programs, i.e., to select for yield-related traits. These results demonstrate that phenotypic selection for important yield-related traits has helped preserve elite alleles, in addition to providing molecular evidence of domestication and breeding activities.

Discussion

Rice breeding programs aim to develop new cultivars with high yield and superior quality. To achieve this aim, multiple desirable traits need to be integrated into a single variety (Chen et al. 2010). Understanding the genetic basis of yield-related traits and identification of candidate genes facilitates the development of effective strategies for crop improvement. Recent advances in genomics and molecular genetics have led to the identification of genes controlling yield and grain quality in rice (Jiang et al. 2012). However, trait correlations and genetic networks controlling these traits remain unclear in rice, and new genes that control important agronomic traits need to be identified.

We re-sequenced a large natural population of rice accessions and identified a group of SNPs and candidate genes significantly associated with yield and plant type. High SNP divergence was found among different chromosomes. The Ka/Ks ratio was relatively higher in rice than in other crops such as maize and barley. However, the F_{ST} between landraces and improved *indica* cultivars in rice was lower than that in maize and wheat, which is in agreement with previous results based on molecular markers and phenotypic traits (Yano et al. 2016). Recently, Huang et al. reported a GWAS on rice that revealed 112 loci associated with 14 agronomic traits under two different environments (Huang et al. 2010). These results provide a genomic basis for improving rice cultivars. In the current study, we re-sequence the genome of a core collection of rice accessions comprising a large natural population, including 19 *japonica* cultivars, 105 landrace, and 151 improved *indica* cultivars, at an average coverage of 12.16×, thus generating 2,888,332 SNPs.

GWAS is a powerful tool for investigating the genetic architecture of complex traits, and has been widely used to identify QTLs and genes associated with important traits in many crops (Zhang et al. 2015; Wen et al. 2018). In maize, Li et al. (2019) employed GWAS to demonstrate the role of an F-box protein-encoding gene, *ZmFBL41*, in the resistance to banded leaf and sheath blight. In this study, we detected loci associated with 13 important agronomic traits rice through the GWAS of using 2,888,332 high-confidence SNPs, and further detected candidate genes by haplotype block structure analysis of SNPs showing significant association with agronomic traits. The reliability of our results was verified by the comparison of candidate genes identified in this study with the functions of these genes reports previously. In addition, many genes identified in this study were previously uncharacterized; for example, the ulp1 protease family protein-encoding gene, LOC_Os03g29630, which controls grain length. Although this locus was previously fine-mapped by our team using a bi-parental mapping population (Li et al. 2012), however, genetic basis and variation of this gene have not been evaluated. Our results will be helpful for future functional analysis, and will provide valuable information for future gene cloning studies in rice.

Grain length, grain width, and grain weight are important traits that directly affect crop yield. The development of grain size is dependent on multiple pathways including the mitogen-activated protein kinase (MAPK) pathway (Xu et al. 2018). The SMG1 gene, which encodes mitogen-activated protein kinase kinase 4 (MAPKK4), is involved in the MAPK signaling pathway, and the smg1 mutant produces small and light-weight grains because of the reduction in cell number (Duan et al. 2014). OsMAPK6 also plays an important role in the regulation of grain size by limiting cell proliferation (Guo et al. 2018). Additionally, the ubiquitin–proteasome pathway also plays a key role in the regulation of grain size in rice (Gull et al. 2019). For example, GW2, which encodes a RING-type protein with E3 ubiquitin ligase activity, negatively regulating cell proliferation via the ubiquitin–proteasome pathway. The GW2 deficient mutant shows increased cell number, and consequently greater grain width and weight compared with the wild type (Song et al. 2007). In Arabidopsis, the ulp1 protease family is essential for processing the small ubiquitin-like modifier precursor protein, which is involved in the ubiquitin-like protein conjugation process (Elmore et al., 2011). Additionally, modification of small ubiquitin-like modifier proteins coordinates the expression of genes essential for growth and hormonal responses in plants by regulating transcription factor activity (Kerscher et al., 2006; Chen et al., 2009). Furthermore, SH3 domain containing proteins play a key role in plant cell division, and ubiquitin binding (Ahn et al. 2017). Therefore, based on our results, we speculate that the ulp1 protease family protein-encoding gene, LOC_Os03g29630, and the SH3 domain containing protein-encoding gene, LOC_Os03g15900,

regulate grain size though the ubiquitin–proteasome pathway. These results further provide a theoretical basis for future studies focusing on the mechanism of grain size regulation.

Leaf is the main photosynthetic organ in plants. In rice, the flag leaf is the most important and functionally efficient organ at the grain filling stage. Therefore, the shape of the flag leaf is one of the most essential traits for ideal plant type in super-rice breeding (Chen et al. 2001). In our study, flag-leaf length and flag-leaf width were showed a significant positive correlation with yield-related traits, grain width and panicle length. Furthermore, a significant positive correlation was also detected between flag-leaf angle and grain width, and between flag–leaf width and grain number per spike; however, negative correlation was detected between flag-leaf width and spikelet number. This is consistent with a previous study (Zhou et al. 2012), which showed that flag-leaf width was significantly correlated with panicle number and spikelet number per panicle. These results suggest that an increase in flag-leaf width increases the photosynthetic area, which enhances the energy supply. Therefore, appropriate increasing flag–leaf size may raise crop yield. In recent years, several QTLs or genes associated with leaf size and shape have been identified (Fujino et al. 2008; Qi et al. 2008; Zhang et al. 2009). Genes encoding the flavin-containing monooxygenase NAL7 and trypsin-like serine/cysteine protease NAL1 were fine-mapped on Chr. 3 and Chr. 4, respectively, and mutants of both these genes show a narrow leaf phenotype (Fujino et al. 2008; Qi et al. 2008). Jiang et al. (2010) identified three QTLs related to flag-leaf length on Chr. 3, 6, and 9 by constructing a genetic population using Shennong 265/LTH. The flag-leaf length-related QTL, qFLL6.2, was fine-mapped to a 62.1 kb region on Chr. 6 (Shen et al. 2012). In the current study, we detected 17, 26, and 43 QTLs related to flag-leaf length, flag-leaf width, and flag-leaf angle, respectively. We further identified candidate genes spanning SNPs showing significant association with traits through haplotype block structure analysis. The identified genes and QTLs can be effectively utilized for controlling flag-leaf size and shape and for breeding high yielding cultivas in rice.

In summary, future transgenic experiment are necessary to validate genes underlying the afronomic traits. Additionally, the genomic data generated in this study will improve our understanding of the allelic variation in the genetic resource collections and will also facilitate breeders to propose an efficient pipeline for variety improvement. This information will be helpful guidance for the breeders attempting to establish a clear strategy for variety development. Nonetheless, a strict background selection should be performed because alleles favorable for other traits should be maximally maintained.

Conclusions

In our study, we identified some candidate regions and genes of the 13 agronomic traits (grain length, grain width, grain weight, panicle length, spikelet number, grain number per spike, percentage of full grains, plant height, flag-leaf length, flag-leaf width, flag-leaf angle, heading date, and awn presence or absence) of rice by GWAS methods. Moreover, the SNPs found in these regions and genes could be used for future gene validation and marker-assisted selection. These results may provide useful information for rice high yield breeding.

Materials And Methods

Plant materials

A population of 259 rice accessions collected from multiple different locations including China, Senegal, Mexico, Malaysia, Colombia, and Brazil was used in this study (Supplementary Table S1). Of these 259 rice accessions, 146 were maintained by the International Rice Research Institute (IRRI), Philippines, and 113 were preserved by the Rice Research Institute of Sichuan Agricultural University, China. Seeds of all accessions were sown in an experimental field at Sichuan Agricultural University, China. Young leaves were harvested from 21-day-old seedlings of each accession for genomic DNA extraction. All 259 rice accessions were evaluated for 13 agronomic traits at Chengdu, China over two years (2018, 2019); the experiment was performed under field conditions.

Analysis of phenotypic data

Traits including grain length, grain width, grain weight, panicle length, and grain number per spike were quantified using the SC-G seed test analyzing system (Wanshen Test Technology Corporation, Hangzhou, China). Statistical tests including correlation analysis and analysis of variance (ANOVA) of all 13 agronomic traits of the each accession across different environments were performed using SPSS 23.0. *P*-values for Pearson's correlation coefficients were calculated with the two-tailed Student's *t*-test using the cor. test function in R (Ihaka and Gentleman, 1996).

DNA extraction and sequencing

Total genomic DNA was prepared using the cetyl trimethyl ammonium bromide (CTAB) method (Uzunova et al. 1995). DNA purity was determined using the NanoPhotometer spectrophotometer (IMPLEN, CA, USA), and DNA concentration was determined using the Qubit DNA Assay Kit on the Qubit 2.0 Fluorometer (Life Technologies, CA, USA). Genomic DNA of each accession was fragmented to a size of 350 bp by sonication, DNA fragments were then end-polished, A-tailed, and ligated with full-length adapters for Illumina sequencing with further PCR amplification. Subsequently, the Illumina Nova sequencing platform was used to generate 150-bp raw sequence reads. Reads with adapter sequences, stretches of -Ns, and low quality scores were removed from the raw data set. The remaining high-quality paired-end reads were mapped to the Nipponbare reference genome sequence (ftp://ftp. ensemblgenomes. org/ pub/ plants/ release_36/ fasta/ oryza_indica/ dna/) using the Burrows-Wheeler Aligner (BWA) software with the command 'mem -t 4 -k 32 –M' (Li et al. 2010). After alignment, genomic variants (in GVCF format for each accession) were identified with the Haplotype Caller module and GVCF model using the Genome Analysis Toolkit (GATK) software (Mckenna et al. 2010). All GVCF files were then merged together. A raw population genotype file with SNPs was created in the Haplotype Caller module, which was filtered based on the following parameters: depth for each individual ≥ 4; genotype quality for each individual ≥4; minor allele frequency (MAF) ≥ 0.01, miss rate ≤ 0.2. Consequently, a total of 2,888,332 SNPs were retained. The identified SNPs were further annotated using the ANNOVAR software (version 2013-05-20) (Wang et al. 2010), and divided into different groups based on reference genome annotations: intergenic SNPs, upstream SNPs (within 1kb upstream of the transcription start site), downstream SNPs (within 1kb downstream of the transcription stop site), CDS SNPs, and intron SNPs.

PCA,population genetics, and LD analysis

An individual-based neighbor-joining (NJ) tree was constructed based on the p-distance using genome-wide 2,888,332 SNPs with the TreeBest software (v1.9.2) and 1, 000 bootstrap replications (Vilella et al. 2009). The population genetic structure was examined using the ADMIXTURE (v1.23) program (Alexander et al. 2009), with the value of *K* ranging from 2 to 3. PCA was conducted using the GCTA software (Yang et al. 2011). First, the genetic relationship matrix (GRM) was obtained using the parameter '–make-grm'. Then, the top three principal components were estimated using the parameter '–pca3'. To estimate LD in the rice population, the value of squared correlation coefficient (r^2) between pairwise SNPs was computed using Pop LD decay (Zhang et al. 2019), with parameters in the program set as '-MaxDist 1000kb-MAF 0.05 -Miss 0.1'. The average r^2 value was calculated for pairwise markers in a 1 kb window, and values were averaged across the whole genome. Nucleotide diversity (π) and fixation index (F_{ST}) were estimated for each group (Indica-IC, Indica-Lan and Japonica) using VCFtools (v0.1.14) (Danecek et al. 2011).

Estimation of breeding value

BLUP was used to calculate the breeding values using the lme4 package in R (version 3.2.2) as follows (Poland et al. 2011):

$$Y=\mu+Line+Loc+(Line\times Loc)+(Rep\times Loc)+\varepsilon$$

where Y, μ, Line, and Loc represent phenotype, intercept, variety effects, and environmental effects, respectively; Rep represents the number of replications; ε represents random effects; Line×Loc represents the interaction between variety and environment; and Rep×Loc represents the interaction between replication and environment.

GWAS

Only SNPs with sequencing depth ≥ 4, missing rate < 0.2, and MAF ≥ 0.05 were used for GWAS. The EMMAX (beta version) software package was used to analyze GWAS data (Kang et al. 2010). The matrix of pairwise genetic distances, calculated by EMMAX, was used as the variance-covariance matrix of random effects. Significant P-value thresholds ($P < 10^{-5}$) were set to control the genome-wide type 1 error rate, which was calculated by 1/n (total SNPs) after rounding ≈ 5.

Haplotype block construction

Haplotype blocks were constructed using the confidence interval method (Gabriel et al., 2002) with the Haploview software (Barrett et al., 2005). The Hardy–Weinberg *P*-value cut-off was set at 0.001 and MAF was set at 0.05.

Abbreviations

GWAS: Genome-wide association study

LD: Linkage disequilibrium

PCA: Principle component analysis

SNP: Single nucleotide polymorphism

BLUP: Best linear unbiased prediction

CTAB: Cetyl trimethyl ammonium bromide

QTL: Quantitative trait loci

MAF: Minor allele frequency

MLM: Mixed linear model

Declarations

Availability of Data and Materials

All the sequences have been deposited in the Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) under the accession PRJNA598020.

Acknowledgements

We sincerely thank funding supports and all the participants.

Funding

This work was supported by the National Natural Science Foundation of China (No. 31400130).

Author information

Contributions

AW analyzed the data and drafted the manuscript; SL, QD, HL, SW, J.Z., YL, LW, TZ, JZ supervised the study; AW and XS phenotyped the germplasm; AW, XJ, and XY processed the genotypic and phenotypic data; AW, XS, LM, and YJ did the field experimental design; AW and AZ managed the project; AZ and PL revised the manuscript; All authors read and approved the final version of the manuscript.

Corresponding authors

Correspondence to Ping Li or Aiping Zheng.

Ethics declarations

Ethics Approval and Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Competing Interests

The authors declare that they have no competing interests.

References

1. Ahn, G., Kim, H., Kim, D.H., Hanh, H., Yoon, Y., Singaram, I., Wijesinghe, K.J., Johnson, K.A., Zhuang, X., Liang, Z., Stahelin, R.V., Jiang, L., Cho, W., Kang, B.H. and Hwang, I. (2017) SH3 domain-containing protein 2 plays a crucial role at the step of membrane tubulation during cell plate formation. PLANT CELL, 29(6):1388 – 1405.
2. Anjum, S.G., Xu, W.J., Nikkholgh, N., Basu, S., Nie, Y.C., Thomas, M., Satyamurti, M., Budnik, B.A., Ip, Y.T. and Veraksa, A. (2013) Regulation of toll signaling and inflammation by β-Arrestin and the SUMO protease Ulp1. GENETICS, 4: 1307-1317.
3. Alexander, D.H., Novembre, J. and Lange, K. (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome research, 19, 1655-1664.
4. Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics, 21, 263–265.
5. Chen, Y., Lubberstedt, T. (2010) Molecular basis of trait correlations. Trends Plant Sci. 15:454–61.
6. Chen, W., Gao, Y.Q., Xie, W.B., Gong, L., Lu, K., Wang, W.S., Li, Y., Liu, X.Q., Zhang, H.Y., Dong, H.X., Zhang, W., Zhang, L.J., Yu, S.B., Wang, G.W., Lian, X.M. and Luo, J. (2014) Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. Nature genetics. 46:714-21.
7. Chen, Y.N., Fan, X.R., Song, W.J., Zhang, Y.L. and Xu, G.H. (2012) Over-expression of OsPIN2 leads to increased tiller numbers, angle and shorter plant height through suppression of OsLAZY1. Plant Biotechnology Journal, 10(2): 139-149.
8. Cui, Y.C., Wang, M., Zhou, H., Li, M., Huang, L., Yin, X., Zhao, G., Lin, F., Xia, X., and Xu, G. (2016) OsSGL, a novel DUF1645 domain-containing protein, confers enhanced drought tolerance in transgenic rice and Arabidopsis. Front. Plant Sci. 7:2001.
9. Chen, Z.J.J., Sun, L.J.J. (2009) Nonproteolytic functions of ubiquitin in cell signaling. Mol Cell. 33: 275-286.
10. Chen, W.F., Xu, Z.J., Zhang, W.Z., Zhang, L.B., Yang, S.R. (2001) Creation of New Plant Type and Breeding Rice for Super High Yield. Yi Chuan Xue Bao. 27:665–72.
11. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., Depristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., Mcvean, G. and Durbin, R. (2011) The variant call format and VCFtools. Bioinformatics 27, 2156–2158.
12. Duan, P.G., Rao, Y.C., Zeng, D.L., Yang, Y.L., Xu, R., Zhang, B.L., Dong, G.J., Qian, Q. and Li, Y.H. (2014) SMALL GRAIN 1, which encodes a mitogen-activated protein kinase kinase 4, influences grain size in rice. The Plant Journal, 77(4): 547-557.
13. Duan, P.G., Xu, J.S., Zeng, D.L., Zhang, G.Z., Huang, K., Huang, L.J., Xu, R., Ge, S., Qian, Q. and Li, Y.H. (2017) Natural Variation in the Promoter of GSE5 Contributes to Grain Size Diversity in Rice. Molecular

Plant 10(5): 685-694.

14. Du X.M., Huang G., He S.P., Yang Z.E., Sun G.F,. Ma X.F,. Li, N., Zhang, X.Y., Sun, J.L., Liu, M., Jia, Y.H., Pan, Z.E., Gong, W.F., Liu, Z.H., Zhu, H.Q., Ma, L., Liu, F.Y., Yang, D.G., Wang, F., Fan, W., Gong, Q., Peng, Z., Wang, L.R., Wang, X.Y., Xu, S.J., Shang, H.H., Lu, C.R., Zheng, H.K., Huang, S.W., Lin, T., Zhu, Y.X. and Li, F.G. (2018) Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits. Nature genetics, 50(6): 796-802.
15. Duvick, D.N., Cassman, K.G. (1999) Post-green revolution trends in yield potential of temperate maize in the north-central United States. Crop Sci. 39:1622–30.
16. Elmore, Z.C., Donaher, M., Matson, B.C., Murphy, H., Westerbeck, J.W. and Kerscher, O. (2011) Westerbeck, J.W. and Kerscher, O. Sumo-dependent substrate targeting of the SUMO protease Ulp1. BMC Biology, 9:74.
17. Fang, C., Ma, Y.M., Wu, S.W., Liu, Z., Wang, Z., Yang, R., Hu, G.H., Zhou, Z.K., Yu, H., Zhang, M., Pan, Y., Zhou, G.A., Ren, H.X., Du, W.G., Yan, H.R., Wang, Y.P., Han, D.Z., Shen, Y.T., Liu, S.L., Liu, T.F., Zhang, J.X., Qin, H., Yuan, J., Yuan, X.H., Kong, F.J., Liu, B.H., Li, J.Y., Zhang, Z.W., Wang, G.D., Zhu, B.G. and Tian ZX. (2017) Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. Genome Biology 18:161.
18. Fang, N., Xu, R., Huang, L.J., Zhang, B.L., Duan, P.G., Li, N., Luo, Y.H., Li, Y.H. (2016) SMALL GRAIN 11 Controls Grain Size, Grain Number and Grain Yield in Rice. Rice, 9: 64.
19. Fujino, K., Matsuda, Y., Ozawa, K., Nishimura, T., Koshiba, T., Fraaije, M.W., Sekiguchi, H. (2008) NARROW LEAF 7 controls leaf shape mediated by auxin in rice. Mol Genet Genomics. 279:499–507.
20. Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J. DeFelice, M., Lochner, A., Faggart, M., Liu-Cordero, S.N., Rotimi, C., Adeyemo, A., Cooper, R., Ward, R., Lander, E.S., Daly, M.J. and Altshuler, D. (2002) The structure of haplotype blocks in the human genome. Science, 296, 2225–2229.
21. Guo, Z. L., Yang ,W.N., Chang, Y., Ma, X.S., Tu, H.F., Xiong, F., Jiang, N., Feng, H., Huang, C.L., Yang, P., Zhao, H., Chen, G.X., Liu, H.Y., Luo, L.J., Hu, H.H., Liu, Q. and Xiong, L.Z. (2018) Genome-wide association studies of image traits reveal genetic architecture of drought resistance in rice. Molecular Plant 11, 789–805.
22. Guo, T., Chen, K., Dong, N.Q., Shi, C.L., Ye, W.W., Gao, J.P., Shan, J.X., Lin, H.X. (2018) GRAIN SIZE AND NUMBER1 Negatively Regulates the OsMKKK10-OsMKK4-OsMPK6 Cascade to Coordinate the Trade-off between Grain Number per Panicle and Grain Size in Rice. The Plant Cell, 30(4): 871-888.
23. Gull, S., Haider, Z., Gu, H.W., Khan, R.A.R., Miao, J., Tan, W.C., Uddin, S., Ahmad, I. and Liang, G.H. InDel. (2019) Marker Based Estimation of Multi-Gene Allele Contribution and Genetic Variations for Grain Size and Weight in Rice (*Oryza sativa* L.). Int. J. Mol. Sci. 20(19), 4824.
24. Han, M.J., Jung, K.H., Yi, G.W., Lee, D.Y. and An, G. (2006) Rice Immature Pollen 1 (RIP1) is a Regulator of Late Pollen Development Plant and Cell Physiology, 47(11): 1457-1472.
25. Huang, X. H., Wei, X.H., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., Li, C.Y., Zhu, C.R., Lu, T.T., Zhang, Z.W., Li, M., Fan, D.L., Guo, Y.L., Wang, A.H., Wang, L., Deng, L.W., Li, W.J., Lu, W.Q., Weng, Q.J., Liu, K.Y., Huang, T., Zhou, T.Y., Jing, Y.F., Li, W., Lin, Z., Buckler, E.S., Qian, Q., Zhang, Q.F., Li, J.Y. and Han, B. (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nature genetics 42, 961-967.

26. Huang, X., Zhao, Y., Wei, X., Li, C.Y., Wang, A.H., Zhao, Q., Li, W.J., Guo, Y.L., Deng, L.W., Zhu, C.R., Fan, D.L., Lu, Y.Q., Weng, Q.J., Liu, K.Y., Zhou,T.Y., Jing, Y.F., Si, L.Z., Dong, G.J., Huang, T., Lu, T.T., Feng, Q., Qian, Q., Li, J.Y., Han, B. (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nature Genetics, 44, 32-41.
27. Hurst, L.D. (2009) Genetics and the understanding of selection. Nat. Rev. Genet. 10, 83–93.
28. Janska, H. (2010) ATP-dependent proteases in biogenesis and maintenance of Arabidopsis mitochondria. Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1797:107.
29. Jiang, Y., Cai, Z., Xie, W., Long, T., Yu, H., Zhang, Q. (2012) Rice functional genomics research: progress and implications for crop genetic improvement. Biotechnol Adv. 30:1059–70.
30. Jiao, Y. P., Zhao, H.N., Ren, L.H.、Song, W.B., Zeng, B., Guo, J.J., Wang, B.B., Liu, Z.P., Chen, J., Li, W., Zhang, M. ,Xie, S.J. and Lai, J.S. (2012) Genome-wide genetic changes during modern breeding of maize. Nature genetics. 44, 812-817.
31. Jiang, S.K., Zhang, X.J., Wang, J.Y., Chen, W.F., Xu, Z.J. (2010) Fine mapping of the quantitative trait locus qFLL9 controlling flag leaf length in rice. Euphytica. 176:341–7.
32. Klingenberg, P.C. (2008) Morphological integration and developmental modularity. Ann Rev Eco Evo Sys. 39:115–32.
33. Kang, H. M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S., Freimer, N.B., Sabattl, C. and Eskin, E. (2010) Variance component model to account for sample structure in genome-wide association studies. Nat Genet. 42, 348–54.
34. Kerscher, O., Felberbaum, R., Hochstrasser, M. (2006) Modification of proteins by ubiquitin and ubiquitin-like proteins. Annu Rev Cell Dev Biol. 22: 159-180.
35. Lannoo, N., Van, Damme, E.J.M., Albenne, C. and Jamet, E. (2014) Plant Glycobiology—a diverse world of lectins, glycoproteins, glycolipids and glycans. Front. Plant Sci. 5:604.
36. Li, C., Zhou, A. and Sang, T. (2006) Rice domestication by reducing shattering. Science 311, 1936–1939.
37. Liu, M.H., Kang, H.X., Xu, Y.C., Peng, Y., Wang, D., Gao, L.J., Wang, X.L., Ning, Y.S., Wu, J., Liu, W.D., Li, C.Y., Liu, B. and Wang, G.L. (2019) Genome-wide association study identifies an NLR gene that confers partial resistance to *Magnaporthe oryzae* in rice. Plant Biotechnology Journal pp. 1–8.
38. Li, S.G., He, P., Wang, Y.P., Li, H.Y., Chen, Y., Zhou, K.D., Zhu, L.H. (2000) Genetic analysis and gene mapping of the leaf traits in rice (*Oryza sativa* L.). Acta Genetica Sin. 26:261–5.
39. Li, W.T., Zhu, Z.W., Chern, M.S., Yin. J.J., Yang, C., Ran, L., Cheng, M.P., He, M., Wang, K., Wang, J., Zhou, X.G., Zhu, X.B., Chen, Z.X., Wang, J.C., Zhao, W., Ma, B.T., Qin, P., Chen, X.L., Wang, Y.P., Liu, J.L., Wang, W.M., Wu, X.J., Li, P., Wang, J.R., Zhu, L.H., Li, S.G. and Chen, X.W. (2017) A natural allele of a transcription factor in rice confers broad-spectrum blast resistance. Cell. 170, 114–126.
40. Li, S.C., Liu, M.W., Wang, S.Q., Deng, Q.M., Zheng, A.P., Liu, H.N., Wang, L.X., Zhu, J., Li, P. (2012) Fine mapping of a dominant minute-grain gene, Mi3, in rice. Mol Breeding, 30: 1045-1051.
41. Li, Q.P., Yan, W.H., Chen, H.X., Tan, C., Han, Z.M., Yao, W., Li, J.W., Yuan, M.Q., Xing, Y.Z. (2016) Duplication of OsHAP family genes and their association with heading date in rice. Journal of Experimental Botany, 67(6): 1759-1768.

42. Li, N., Lin, B., Wang, H., Li, X.M., Yang, F.F., Ding, X.H., Yan, J.B. and Chu, Z.H. (2019) Natural variation in ZmFBL41 confers banded leaf and sheath blight resistance in maize. Nature genetics. 51, 1540-1548.
43. Li, H., and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. 26, 589-95.
44. Ma, Z.Y., He, S.P., Wang, X.F., Sun, J.L., Zhang, Y., Zhang, G.Y., Wu, L.Q., Li, Z.K., Liu, Z.H., Sun, GF., Yan, Y.Y., Jia, Y.H., Yan, J., Pan, Z.E., Gu, Q.S., Li, X.Y., Sun, Z.W., Dai, P.H., Liu, Z.W., Gong, W.F., Wu, J.H., Wang, M., Liu, H.W., Feng, K.Y., Ke, H.F., Wang, J.D., Lan, H.Y., Wang, J.N., Peng, J., Wang, N., Wang, L.R., Pang, B.Y., Peng, Z., Li, R.Q., Tian, S.L., Du, X.M. (2018) Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield. Nature genetics, 50(6):803-813.
45. Melo, D., Marroig, G. (2015) Directional selection can drive the evolution of modularity in complex traits. Proc Natl Acad Sci USA. 112:470–5.
46. Miura, K., Ashikari, M. and Matsuoka, M. (2011) The role of QTLs in the breeding of high-yielding rice. Trends Plant Sci. 16, 319–326.
47. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. and Depristo, M.A. (2010) The genome analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297-1303.
48. Poland, J. A., Bradbury, P. J., Buckler, E. S., and Nelson, R. J. (2011) Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. Proc Natl Acad Sci USA. 108, 6893-6898.
49. Qi,J ., Qian, Q., Bu,Q.Y., Li,S.Y. ,Chen, Q., Sun,J.Q., Liang, W.X., Zhou,Y.H., Chu,C.C. , Li,X.G. , Ren, F.G., Palme, K .,Zhao, BR., Chen JF., Chen, MS., Li, C.Y. (2008) Mutation of the rice Narrow leaf1 gene, which encodes a novel protein, affects vein patterning and polar auxin transport. Plant Physiol. 147:1947–59.
50. Rharrabti, Y., Elhani, S., Martos-Nunez, V., Garcia Del Moral, L.F. (2001) Protein and lysine content, grain yield, and other technological traits in durum wheat under Mediterranean conditions. J Agric Food Chem. 49:3802–7.
51. Russell, J., Mascher, M., Dawson, I.K., Kyriakidis, S., Calixto, C., Freund, F., Bayer, M., Milne, I., Marshall-Griffiths, T., Heinen, S., Hofstad, A., Sharma, R., Himmelbach, A., Knauft, M., Zonneveld, M.V., Brown, J.W.S., Schmid, K., Kilian, B., Kilian, G.J., Stein N. and Waugh R. (2016) Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. Nature genetics. 48, 1024-1032.
52. Rotundoa, J.L., Borrása, L., Westgatea, M.E., Orfc, J.H. (2009) Relationship between assimilate supply per seed during seed filling and soybean seed composition. Field Crop Res. 112:90–6 .
53. Santelia, D., Vincenzetti, V., Azzarello, E., Bovet, L., Fukao, Y., Duchtig, P., Mancuso, S., Martinoia, E., Geisler, M. (2005) MDR-like ABC transporter AtPGP4 is involved in auxin-mediated lateral root and root hair development. FEBS Letters, 579 (24): 5399-5406.
54. Song, X.J., Huang, W., Shi, M., Zhu, M.Z. and Lin, H.X. (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nature Genetics, 39: 623-630.
55. Shen, B., Yu, W.D., Zhu, Y.J., Fan, Y.Y., Zhuang, J.Y. (2012) Fine mapping of a major quantitative trait locus, qFLL6.2, controlling flag leaf length and yield traits in rice (*Oryza sativa* L.). Euphytica, 184, 57–64.

56. Uga, Y., Fukuta, Y., Cai, H.W., Iwata, H., Ohsawa, R., Morishima, T., Fujimura, T. (2003) Mapping QTLs influencing rice floral morphology using recombinant inbred lines derived from a cross between *Oryza sativa* L. and *Oryza rufipogon* Griff. Theor. Appl. Genet. 107: 218-226.

57. Uga, Y., Fukuta, Y., Ohsawa, R., Fujimura, T. (2003) Variations of floral traits in Asian cultivated rice (*Oryza sativa* L.) and its wild relatives (*O. rufipogon* Griff.). Breed. Sci. 53: 345-352.

58. Uzunova, M., Ecke, W., Weissleder, K., Rbbelen, G. (1995) Mapping the genome of rapeseed (*Brassica napus* L.). I. Construction of an RFLP linkage map and localization of QTLs for seed glucosinolate content. Theoretical and Applied Genetics. 90, 194-204.

59. Varki, A. and Lowe, J. B. (2009) "Biological roles of glycans," in Essentials of Glycobiology, 2nd Edn., Cold Spring Harbor, eds A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart, M. E. Etzler (New York, NY: Cold Spring Harbor Laboratory Press), Chapter 6.

60. Vilella, A. J., Severin, J., Ureta-Vidal, A., Heng, L., Durbin, R., Birney, E. (2009) EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. Genome research. 19, 327-335 .

61. Wagner, G.P. (1996) Homologues, natural kinds and the evolution of modularity. Am Zool. 36:36–43.

62. Wang, K., Li, M. and Hakonarson, H. (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research. 38, e164-e164.

63. Wang, Y.X., Xiong, G.S., Hu, J., Jiang, L., Yu, H., Xu, J., Fang, Y.X., Zeng, L.J., Xu, E.B., Xu„ J., Ye, W.J., Meng, X.B., Liu, R.F., Chen, H.Q., Jing, Y.H., Wang, Y.H., Zhu, X.D., Li, J.Y. and Qian, Q. (2015) Copy number variation at the GL7 locus contributes to grain size diversity in rice. Nature Genetics, 47(8): 944-948.

64. Wen, Z.X., Tan, R.J., Zhang, S.C., Collins, P.J., Yuan, J.Z., Du, W.Y., Gu, C.H., Ou, S.J., Song, Q.J., An, Y.Q.C., Boyse, J.F., Chilvers, M.I., Wang, D.C. (2018) Integrating GWAS and gene expression data for functional characterization of resistance to white mould in soya bean. Plant Biotechnology Journal. 16, pp. 1825–1835 .

65. Xu, F., Fang, J., Ou, SJ., Gao, SP., Zhang, F.X., Du, L., Xiao, Y.H., Wang, H.R., Sun, X.H., Chu, J.F., Wang, G.D., Chu, C.C. (2015) Variations in CYP78A13 coding region influence grain size and yield in rice Plant, Cell & Environment, 38(4): 800-811.

66. Xu, R., Duan, P., Yu, H., Zhou, Z.K., Zhang, B.L., Wang, R.C., Li, J., Zhang, J.Z., Zhuang, S.S., Lyu, J., Li„N., Chai, T.Y., Tian, Z.X., Yao, S.G. and Li, Y.H. (2018) Control of grain size and weight by the OsMKKK10-OsMKK4-OsMAPK6 signaling pathway in rice. Molecular Plant, 11(6): 860-873.

67. Yano, K., Yamamoto, E., Aya, K., Takeuchi, H., Lo, P., Hu, L., Yamasaki, M., Yoshida, S., Kitano, H., Hirano, K. and Matsuoka, M. (2016) Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. Nature genetics, 48:927–934.

68. Yan, D.W., Zhou, Y., Ye, S.H., Zeng, L.J., Zhang, X.M. and He, Z.H. (2013) BEAK-SHAPED grain 1/Triangular hull 1, a DUF640 gene, is associated with grain shape, size and weight in rice. Science china life sciences, 56(3):275-283.

69. Yang, J., Lee, S. H., Goddard, M. E. and Visscher, P. M. (2011) GCTA: a tool for genome-wide complex trait analysis. American journal of human genetics. 88, 76-82.

70. Yano, K., Yamamoto, E., Aya, K., Takeuchi, H., Lo, P., Hu, L., Yamasaki, M., Yoshida, S., Kitano, H., Hirano, K. and Matsuoka, M. (2016) Genome-wide association study using whole-genome sequencing rapidly

identifies new genes influencing agronomic traits in rice. Nature genetics 48, 927–934.

71. Yang, RC., Yang, H.J. (1998) Progress of the research on new plant type rice at IRRI. Hybrid Rice. 13:29–31.
72. Yue, B., Xue, W.Y., Luo, L.J., Xing, Y.Z. (2006) QTL Analysis for flag leaf characteristics and their relationships with yield and yield traits in rice. Acta Genet Sin. 33:824–32.
73. Zhang, C., Dong, S.S., Xu, J.Y., He, W.M. and Yang, T.L. (2019) PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. Bioinformatics. 35, 1786-1788.
74. Zhang, J.P., Wang, X.Z., Lu, Y.M., Bhusal, S.J., Song, Q.J., Cregan, P.B., Yen, Y., Brown, M. and Jiang, G.L. (2017) Genome-wide scan for seed composition provides insights into soybean quality improvement and the impacts of domestication and breeding. Molecular Plant, 11, 460–472.
75. Zhan, C.F., Hu, J.X., Pang, Q., Yang, B., Cheng, Y.H., Xu, E.S., Zhu, P.W., Li, Y.Y., Zhang. H.S. and Cheng. J.P. (2019) Genome-wide association analysis of panicle exsertion and uppermost internode in rice (*Oryza sativa* L.). Rice, 12:72.
76. Zhang, F., Zeng, D., Zhang, C.S., Lu, J.L., Chen, T.J., Xie, J.P. and Zhou, Y.L. (2019) Genome-wide association analysis of the genetic basis for sheath blight resistance in rice. Rice, 12:93.
77. Zhou, H., Li, P.B., Xie, W.B., Hussain, S., Li, Y.B., Xia, D., Zhao, H., Sun, S.Y., Chen, J.X., Ye, H., Hou, J., Zhao, D., Gao, G.J., Zhang, Q.L., Wang, G.W., Lian, X.M., Xiao, J.H., Yu, S.B., Li, X.H. and He, Y.Q. (2017) Genome-wide association analyses reveal the genetic basis of stigma exsertion in rice. Mol. Plant. 10, 634–644.
78. Zhang, B., Ye, W.J., Ren, D.Y., Tian, P., Peng, Y.L., Gao, Y., Ruan, B.P., Wang, L., Zhang, G.H., Guo, L.B., Qian. Q. and Gao, Z.Y. (2015) Genetic analysis of flag leaf size and candidate genes determination of a major QTL for flag leaf width in rice. Rice, 8:2.
79. Zhang, J.P., Singh, A., Mueller, D.S., Singh, A.K. (2015) Genome-wide association and epistasis studies unravel the genetic architecture of sudden death syndrome resistance in soybean. The Plant Journal. 84, 1124–1136.
80. Zhou, L.H., Xie, Y.C., Chen, T., Zhang, Y.D., Zhu, Z., Zhao, Q.Y., Yao, S., Yu, X., Zhao, L. and Wang, C.L. (2012) Relations between flag leaf morphology and yield and QTL analysis of related traits. Jiangsu J Agr Sci. 28:1207–11.
81. Zhang, G.H., Xu, Q., Zhu, X.D., Qian, Q. and Xue, H.W. (2009) SHALLOT-LIKE1 is a KANADI transcription factor that modulates rice leaf rolling by regulating leaf abaxial cell development. Plant Cell. 21:719–35.

Figures

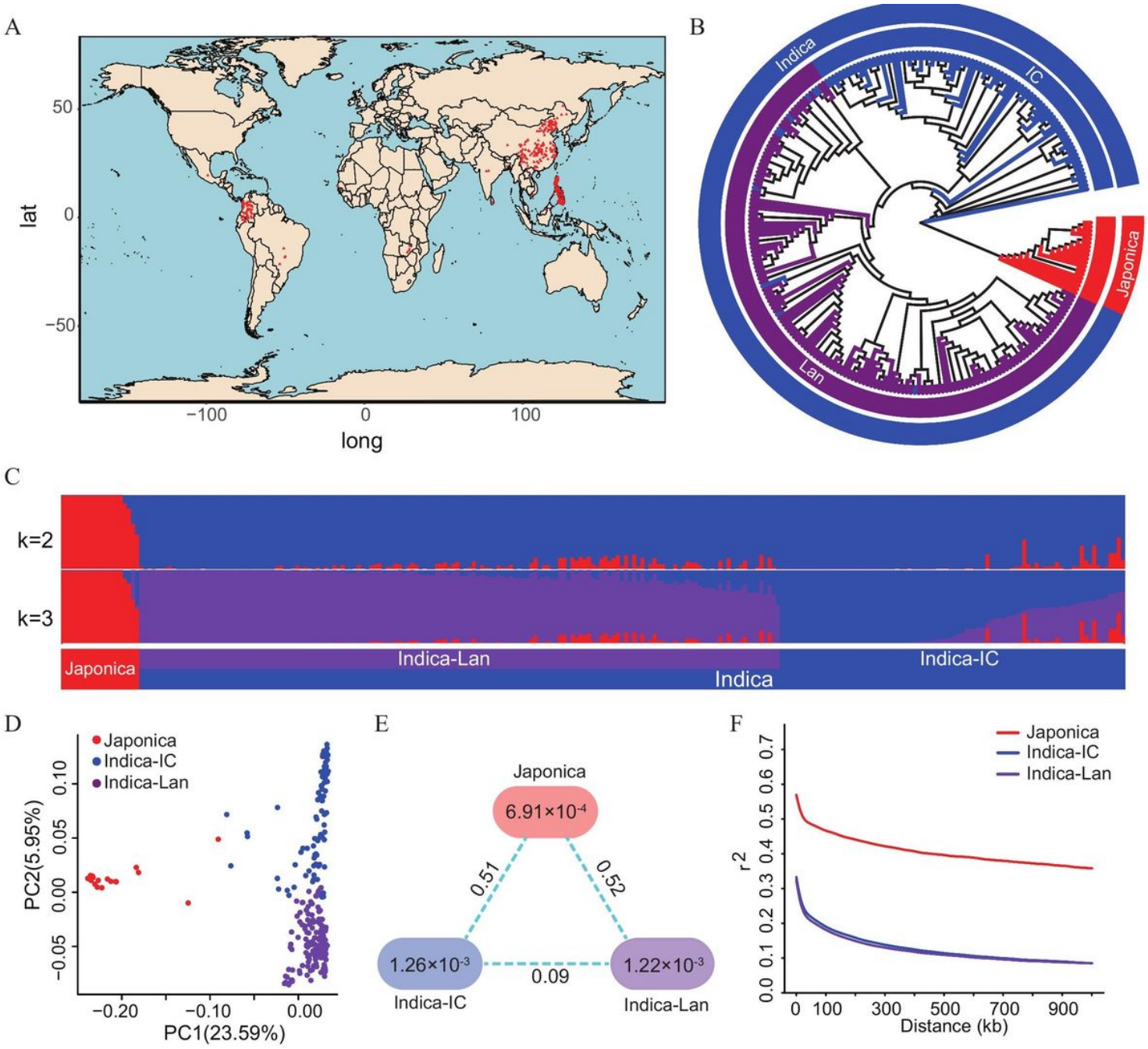

Figure 1

Population structure of 259 rice accessions. (A) Map showing the geographic distribution of 259 rice accessions. Each dot of a given color represents an individual rice accession; (B) Individual ancestry coefficients of 259 rice accessions determined using ADMIXTURE with the number of ancestry kinships (K) set at 2 or 3. Each accession is denoted by a vertical bar, and the proportion of different colors in each bar indicates the proportion of genetic background of each of the ancestral populations; (C) Neighbor-joining (NJ) phylogenetic tree constructed using 2,888,332 high-quality SNPs; (D) Principal component analysis (PCA) of 259 rice accessions. (E) Genetic diversity and population differentiation across the three groups. Values in circles represent the genetic diversity (π) of each group, and values between groups indicate population differentiation (FST); (F) Genome-wide average linkage disequilibrium (LD) decay rate of 259 rice accessions divided into three distinct groups.

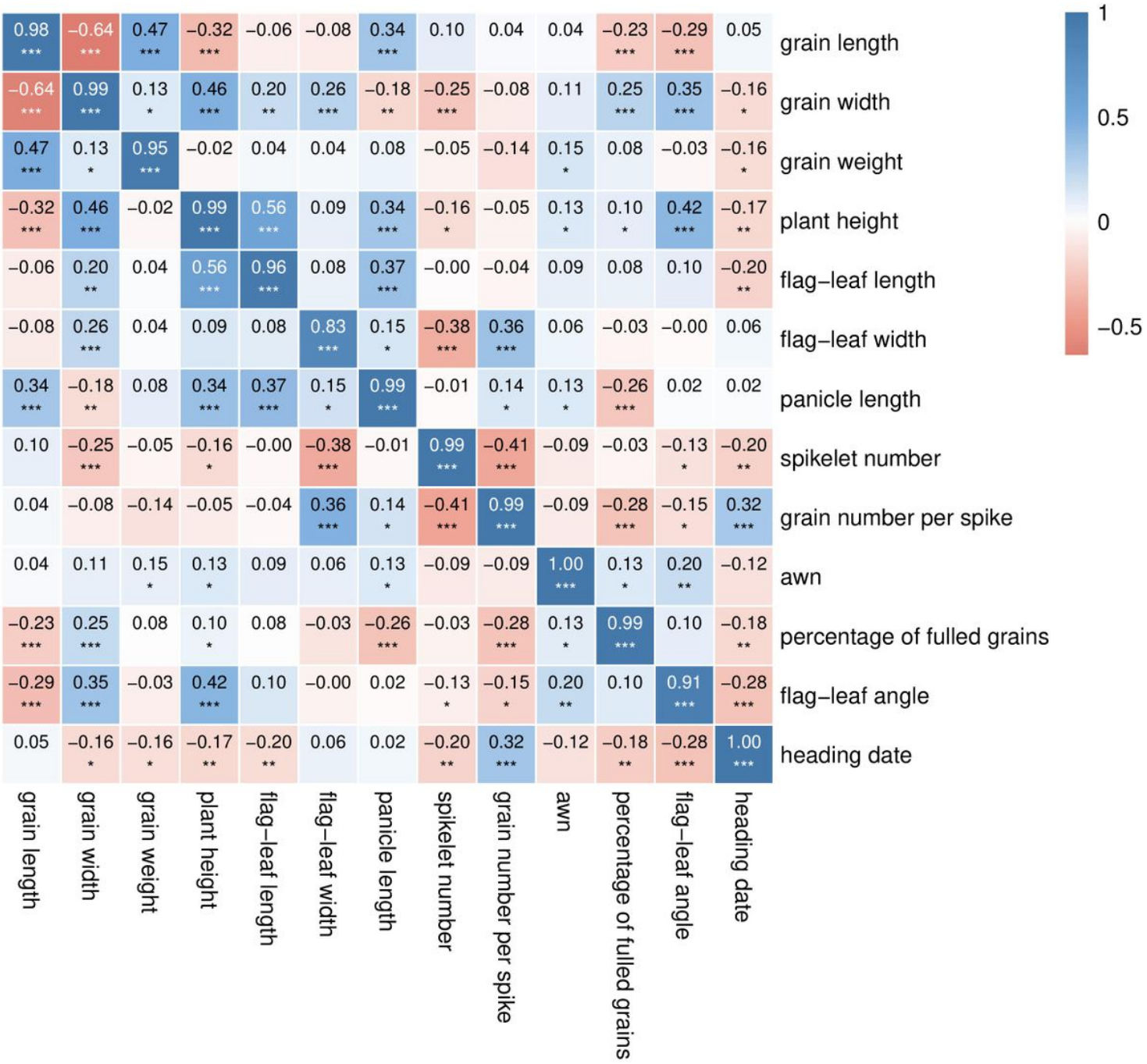

Figure 2

Phenotypic analysis reveals trait relationships.

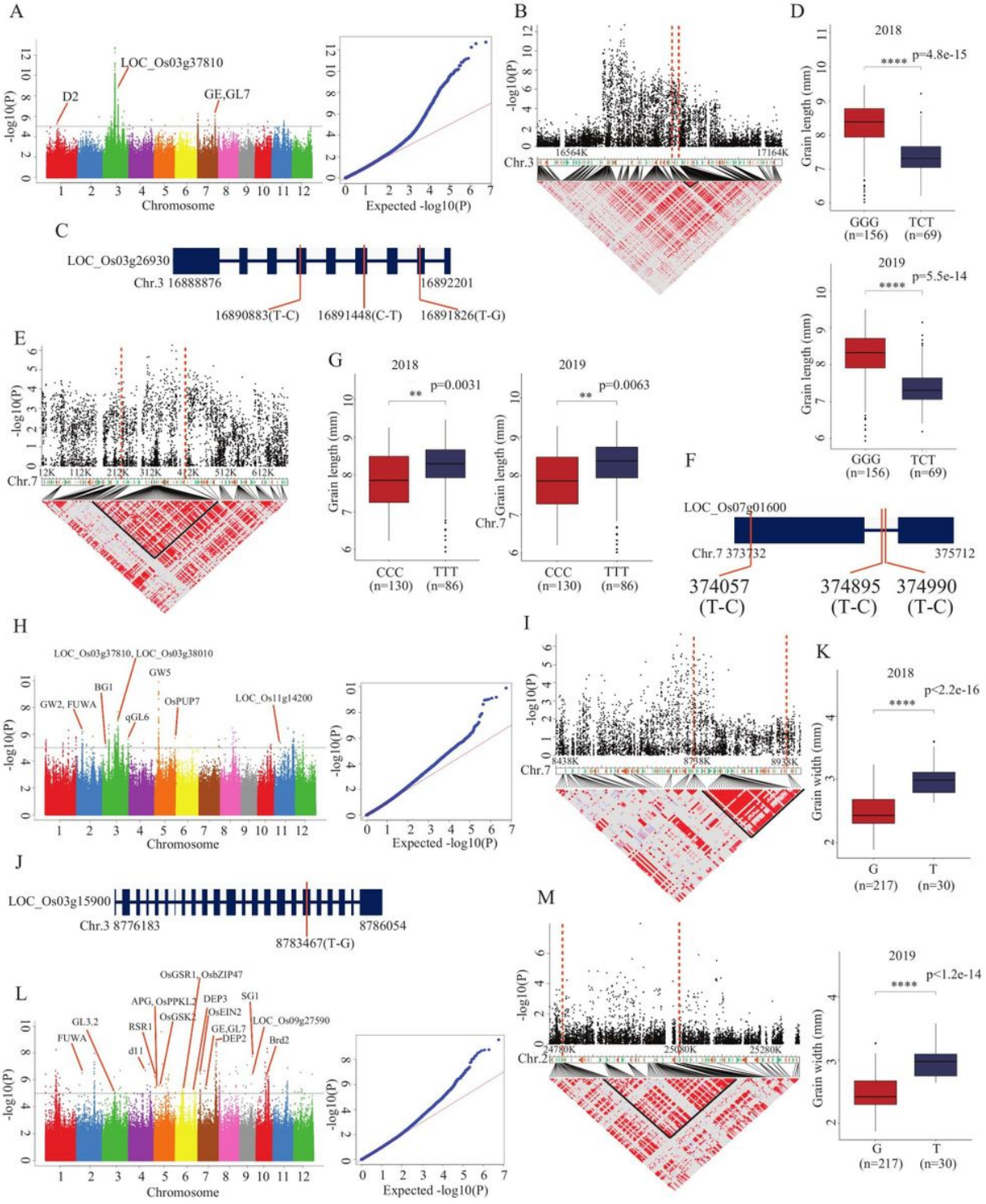

Figure 3

Identification of candidate genes related to grain length, width, and weight. (A) Genes with known function for rice grain length under candidate genes are shown in significant genome-wide association study analysis (GWAS) association positions; (B) Local Manhattan plot of single-polymorphism-based association (top) and LD heatmap (bottom) surrounding the peak on chromosome 3. Red dashed lines indicate the candidate region for associated SNPs; (C) Structure of LOC_Os03g26930 and corresponding SNPs; (D) Grain length based on haplotypes for LOC_Os03g26930 in 2018 (up) and 2019 (down). Differences between the

haplotypes were statistically analyzed using Tukey's test (****P < 0.0001); (E) Local Manhattan plot of single-polymorphism-based association (top) and LD heatmap (bottom) surrounding the peak on chromosome 7. Red dashed lines indicate the candidate region for associated SNPs; (F) Structure of LOC_Os07g01600 and corresponding SNPs; (G) Grain length based on the haplotypes for LOC_Os07g01600 in 2018 (left) and 2019 (right). Differences between the haplotypes were statistically analyzed using Tukey's test (**P < 0.01); (H) Genes with known function for rice grain width under candidate genes are shown in significant GWAS association positions; (I) Local Manhattan plot of single-polymorphism-based association (top) and LD heatmap (bottom) surrounding the peak on chromosome 7. Red dashed lines indicate the candidate region for associated SNPs; (J) Structure of LOC_Os03g15900 and corresponding SNPs; (K) Grain width based on haplotypes for LOC_Os03g15900 in 2018 (up) and 2019 (down). Differences between the haplotypes were statistically analyzed using Tukey's test (****P < 0.0001); (L) Genes with known function for rice grain weight under candidate genes are shown in significant GWAS association positions; (M) Local Manhattan plot of single-polymorphism-based association (top) and LD heatmap (bottom) surrounding the peak on chromosome 2. Red dashed lines indicate the candidate region for associated SNPs.

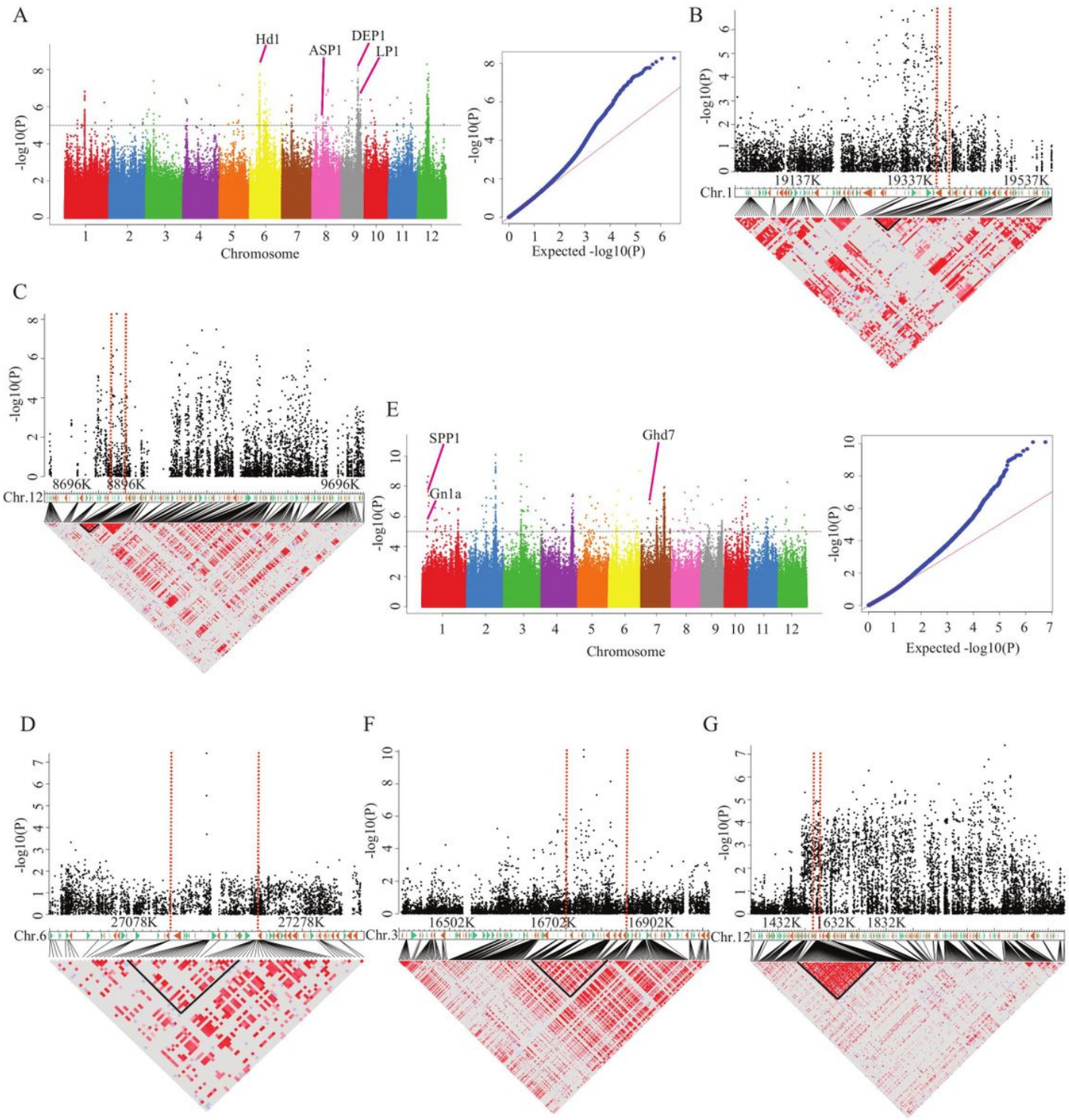

Figure 4

Identification of candidate genes related to panicle length, panicle number, total grain number per panicle, and percentage of full grains. (A) Genes with known function for rice panicle length under candidate genes are shown in significant GWAS association positions; (B) Local Manhattan plot of single-polymorphism-based association (top) and LD heatmap (bottom) surrounding the peak on chromosome 1. Red dashed lines indicate the candidate region for associated SNPs; (C) Local Manhattan plot of single-polymorphism-based association (top) and LD heatmap (bottom) surrounding the peak on chromosome 12. Red dashed lines

indicate the candidate region for associated SNPs; (D) Local Manhattan plot of single-polymorphism-based association (top), and LD heatmap (bottom) surrounding the peak on chromosome 6 (panicle number). Red dashed lines indicate the candidate region for associated SNPs; (E) Genes with known function for rice total grain number per panicle under candidate genes are shown in significant genome-wide association study analysis (GWAS) association positions; (F) Local Manhattan plot of single-polymorphism-based association (top) and LD heatmap (bottom) surrounding the peak on chromosome 3. Red dashed lines indicate the candidate region for associated SNPs; (G) Local Manhattan plot of single-polymorphism-based association (top) and LD heatmap (bottom) surrounding the peak on chromosome 12 (percentage of full grains). Red dashed lines indicate the candidate region for associated SNPs.

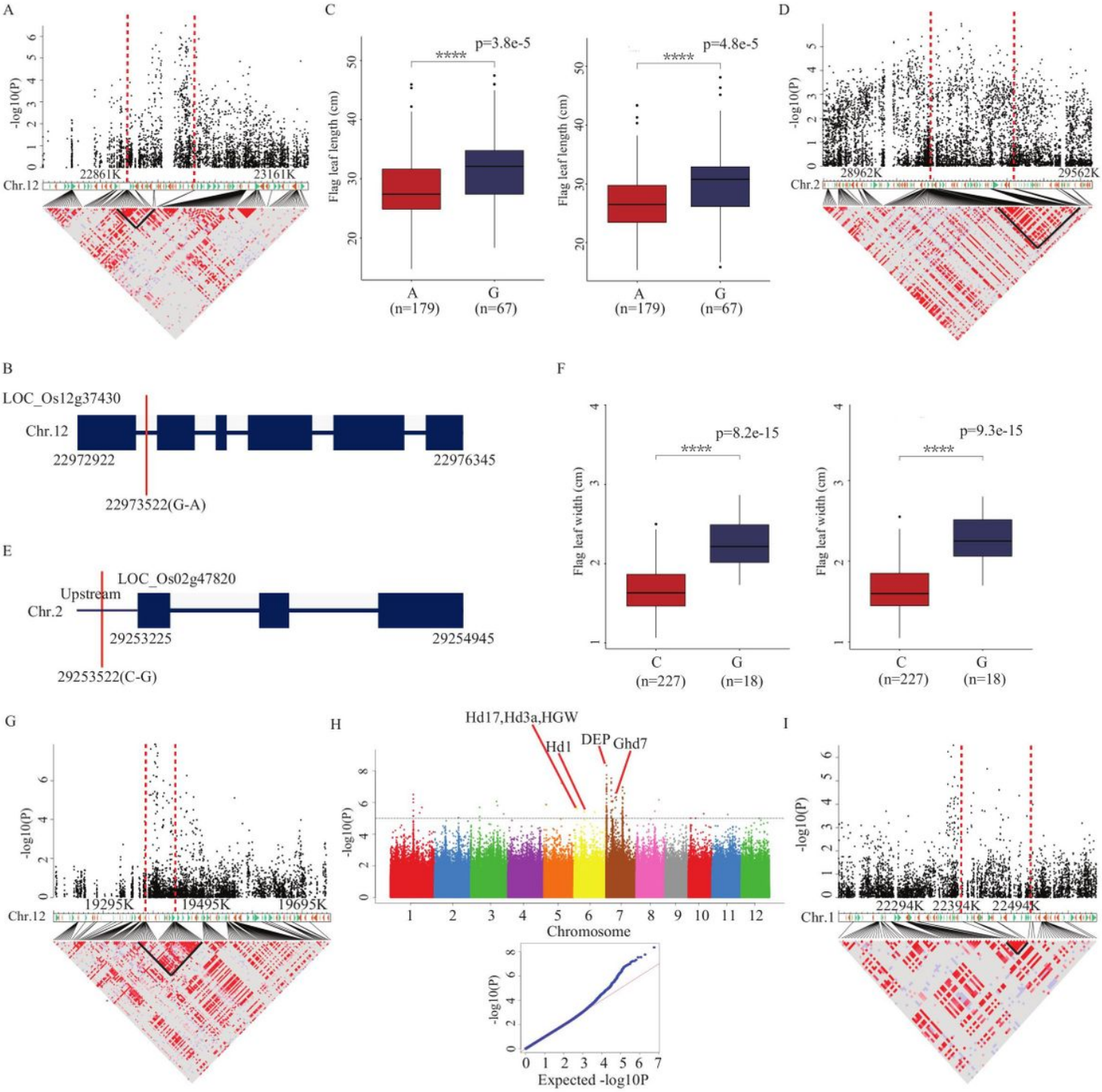

Figure 5

Identification of candidate genes related to flag-leaf length, flag-leaf width, flag-leaf angle, and heading date. (A) Local Manhattan plot of single-polymorphism-based association (top) and LD heatmap (bottom) surrounding the peak on chromosome 12 (flag-leaf length). Red dashed lines indicate the candidate region for associated SNPs; (B) Structure of LOC_Os12g37430 and corresponding SNPs; (C) Flag-leaf length based on the haplotypes for LOC_Os12g37430 in 2018 (left) and 2019 (right). Differences between haplotypes were statistically analyzed using Tukey's test (****P < 0.0001); (D) Local Manhattan plot of single-polymorphism-based association (top) and LD heatmap (bottom) surrounding the peak on chromosome 2 (flag-leaf width). Red dashed lines indicate the candidate region for associated SNPs; (E) Structure of LOC_Os02g47820 and corresponding SNPs; (F) Flag-leaf width based on the haplotypes for LOC_Os02g47820 in 2018 (left) and 2019 (right). Differences between haplotypes were statistically analyzed using Tukey's test (**P < 0.01); (G) Local Manhattan plot of single-polymorphism-based association (top) and LD heatmap (bottom) surrounding the peak on chromosome 12 (flag-leaf angle). Red dashed lines indicate the candidate region for associated SNPs; (H) Genes with known function for rice heading data under candidate genes are shown in significant GWAS association positions; (I) Local Manhattan plot of single-polymorphism-based association (top) and LD heatmap (bottom) surrounding the peak on chromosome 1. Red dashed lines indicate the candidate region for associated SNPs.

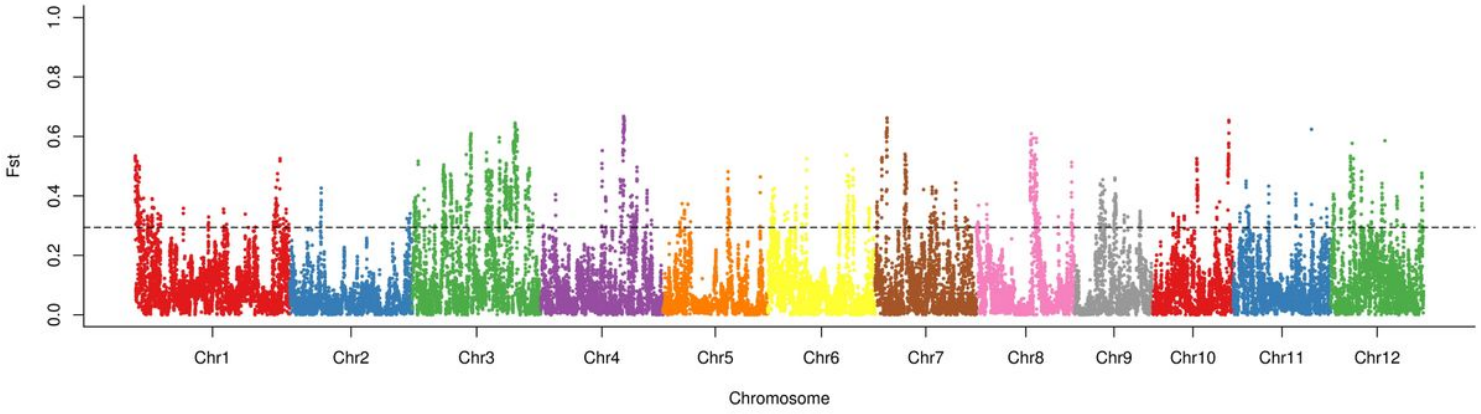

Figure 6

Genome-wide scanning and annotation of selected regions during 13 agronomic traits in rice (Improved_cultivar vs Landrace of indica). Genome-wide screening of selection signals based on FST values. Horizontal dashed lines indicate the significance threshold (FST=0.3).

Supplementary Files

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

- TableS18.xlsx
- TableS18.xlsx
- TableS18.xlsx

- TableS17.xlsx
- TableS17.xlsx
- TableS17.xlsx
- TableS16.xlsx
- TableS16.xlsx
- TableS16.xlsx
- TableS15.xlsx
- TableS15.xlsx
- TableS15.xlsx
- TableS14.xlsx
- TableS14.xlsx
- TableS14.xlsx
- TableS13.xlsx
- TableS13.xlsx
- TableS13.xlsx
- TableS12.xlsx
- TableS12.xlsx
- TableS12.xlsx
- TableS11.xlsx
- TableS11.xlsx
- TableS11.xlsx
- TableS10.xlsx
- TableS10.xlsx
- TableS10.xlsx
- TableS9.xlsx
- TableS9.xlsx
- TableS9.xlsx
- TableS8.xlsx
- TableS8.xlsx
- TableS8.xlsx
- TableS7.xlsx
- TableS7.xlsx
- TableS7.xlsx
- TableS6.xlsx
- TableS6.xlsx
- TableS6.xlsx

- TableS5.xlsx
- TableS5.xlsx
- TableS5.xlsx
- TableS4.xlsx
- TableS4.xlsx
- TableS4.xlsx
- TableS3.xlsx
- TableS3.xlsx
- TableS3.xlsx
- TableS2.xls
- TableS2.xls
- TableS2.xls
- TableS1.xlsx
- TableS1.xlsx
- TableS1.xlsx
- SupplementaryFigure.docx
- SupplementaryFigure.docx
- SupplementaryFigure.docx