



Divergent *Hd1*, *Ghd7*, and *DTH7* Alleles Control Heading Date and Yield Potential of *Japonica* Rice in Northeast China

Jing Ye^{1,2†}, Xiaojun Niu^{1†}, Yaolong Yang¹, Shan Wang¹, Qun Xu¹, Xiaoping Yuan¹, Hanyong Yu¹, Yiping Wang¹, Shu Wang^{2*}, Yue Feng^{1*} and Xinghua Wei^{1*}

¹ State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou, China, ² College of Agronomy, Shenyang Agricultural University, Shenyang, China

OPEN ACCESS

Edited by:

Chengdao Li,
Murdoch University, Australia

Reviewed by:

Jinjin Jiang,
Yangzhou University, China
Xiang-Yang Lou,
University of Alabama at Birmingham,
United States

*Correspondence:

Xinghua Wei
weixinghua@caas.cn
Yue Feng
fy_555500@163.com
Shu Wang
wangshu1@126.com

[†] These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 14 July 2017

Accepted: 09 January 2018

Published: 26 January 2018

Citation:

Ye J, Niu X, Yang Y, Wang S, Xu Q,
Yuan X, Yu H, Wang Y, Wang S,
Feng Y and Wei X (2018) Divergent
Hd1, *Ghd7*, and *DTH7* Alleles Control
Heading Date and Yield Potential
of *Japonica* Rice in Northeast China.
Front. Plant Sci. 9:35.
doi: 10.3389/fpls.2018.00035

The heading date is a vital factor in achieving a full rice yield. Cultivars with particular flowering behaviors have been artificially selected to survive in the long-day and low-temperature conditions of Northeast China. To dissect the genetic mechanism responsible for heading date in rice populations from Northeast China, association mapping was performed to identify major controlling loci. A genome-wide association study (GWAS) identified three genetic loci, *Hd1*, *Ghd7*, and *DTH7*, using general and mixed linear models. The three genes were sequenced to analyze natural variations and identify their functions. Loss-of-function alleles of these genes contributed to early rice heading dates in the northern regions of Northeast China, while functional alleles promoted late rice heading dates in the southern regions of Northeast China. Selecting environmentally appropriate allele combinations in new varieties is recommended during breeding. Introducing the early *indica* rice's genetic background into Northeast *japonica* rice is a reasonable strategy for improving genetic diversity.

Keywords: heading date, *japonica* rice, Northeast China, GWAS, genetic diversity, selection

INTRODUCTION

Rice (*Oryza sativa* L.) is a short-day (SD) plant in which flowering time (or heading date) is severely delayed under long-day (LD) conditions (Shimamoto and Yokoi, 2005). However, rice is widely cultivated in high-latitude regions, such as Italy, Japan, and Northeast China (Shrestha et al., 2014). Northeast China, with its fertile soil, vast plains and relatively low population density, contains 4.5 million hectares of farmland and, in 2014, produced 32.9 million tons of rice (16% of the total rice production of China). This region is the main rice production area and an important commercial food base in China (Wei et al., 2008). Since 2012, the area for rice cultivation has expanded in Huma (51°8'N), which is the most northern region for rice cultivation in the world and has 16 h days during the hottest month of the year. After domestication and with artificial selection, rice's adaptability to higher latitudes has improved because of its reduced photoperiod sensitivity (PS) (Fujino et al., 2013; Yan et al., 2013).

Flowering time (or heading date), which determines the beginning of the reproductive cycle, is greatly affected by environmental conditions (e.g., day length and temperature) and is mainly determined by PS, temperature sensitivity and vegetative-growth duration (Shang et al., 2012). The photoperiod-related gene network has been widely studied and is the major genetic pathway controlling heading date in rice (Gao et al., 2014). Rice as a model plant for

molecular genetic studies shares many homologous genes with *Arabidopsis*. For example, *HEADING DATE 3a (Hd3a)* and *RICE FLOWERING LOCUS T (RFT1)*, which are homologs of *FLOWERING LOCUS T (FT)* in *Arabidopsis*, act as florigen genes to accelerate flowering (Tamaki et al., 2007; Komiya et al., 2008). *Hd3a* promotes heading under SD conditions, whereas *RFT1* is a major floral activator under LD conditions (Kojima et al., 2002; Komiya et al., 2009). The florigen-regulated network is mainly induced by *HEADING DATE 1 (Hd1)* and *EARLY HEADING DATE 1 (Ehd1)* (Sun et al., 2014). *Hd1*, which is a homolog of *CONSTANS* in *Arabidopsis*, promotes flowering under SD conditions and represses flowering under LD conditions by regulating the expression of *Hd3a* (Putterill et al., 1995; Yano et al., 2000; Ishikawa et al., 2011). *Ehd1*, encoding a B-type response regulator without any ortholog in the *Arabidopsis* genome, promotes flowering by inducing the expression of two florigen genes under both SD and LD conditions (Doi et al., 2004).

Several upstream regulatory factors controlling flowering under LD conditions have been identified. Among them, *GRAIN NUMBER, PLANT HEIGHT AND HEADING DATE 7 (Ghd7)*, which is an LD-specific repressor of *Ehd1* expression, plays a crucial role in controlling yields, plant heights and heading dates in rice, simultaneously (Xue et al., 2008; Itoh et al., 2010). *DAYS TO HEADING 8 (DTH8)*, encoding an OsHAP3 subunit of a CCAAT-box binding protein (HAP complex), delays flowering by inhibiting expression of *Ehd1* under LD conditions (Wei et al., 2010b; Yan et al., 2011; Dai et al., 2012). *Heading date 17 (Hd17)*, which is an ortholog of *Arabidopsis EARLY FLOWERING 3*, down-regulates *Ehd1* expression by decreasing *Ghd7* expression and up-regulates *Hd1* expression by decreasing *OsGI* expression, resulting in accelerated flowering under LD conditions (Saito et al., 2012; Brambilla and Fornara, 2013; Yang et al., 2013). *Days to heading 7 (DTH7)* encodes a pseudo-response regulator protein (OsPRR37) and inhibits *Hd3a* expression to suppress flowering under LD conditions (Koo et al., 2013; Li et al., 2015).

With the rapid development of technology and decreasing cost of sequencing, single nucleotide polymorphism (SNP) markers are now used widely to construct genetic maps, establish DNA fingerprinting and analyze linkage disequilibrium (Lindblad-Toh et al., 2000; Syvänen, 2001; Primmer et al., 2002). A genome-wide association study (GWAS) is a powerful tool to identify genes associated with the quantitative variation of complex traits (Burton et al., 2007; Huang et al., 2010). Recently, GWAS has been applied to crops, particularly rice, resulting in the identification of several new genes (Gu et al., 2015; Dong et al., 2016; Yano et al., 2016). However, identifying unknown genes in rice is still a challenge because of spurious and indirect associations, which are generated by the strong population structure and low rate of linkage disequilibrium decay (Gupta et al., 2005; Hamblin et al., 2011).

In this study, 244 representative *japonica* rice cultivars from Northeast China with an enormous level of phenotypic diversity in flowering time, were grown in three environments having a wide geographical span (from 18°N to 41°N) to investigate the mechanistic responses to photoperiod. We attempted to identify flowering gene(s) involved in controlling the rice heading date in

Northeast China using GWAS. The association mapping between SNPs and heading dates showed that flowering times were largely determined by the natural variations in *Hd1*, *Ghd7*, and *DTH7*. Thus, *Hd1*, *Ghd7*, and *DTH7* were sequenced to estimate the distributions of different allelic combinations associated with rice adaptability in the different zones of high-latitude regions. The ability to predict heading dates and yield potentials could be used to produce hybrid progeny from parental rice varieties in this region and provide guidance for molecular breeding.

MATERIALS AND METHODS

Plant Materials and Growing Conditions

A set of 244 *japonica* rice varieties cultivated in Northeast China (including Heilongjiang, Jilin and Liaoning Provinces) were obtained from the National Medium Rice Genebank at the China National Rice Research Institute (Supplementary Figure S1 and Supplementary Table S1). Rice plant were grown in three natural environments at Lingshui (18°32'N, 110°01'E) (winter of 2015 and 2016), Hangzhou (30°15' N, 120°12'E) (2016) and Shenyang (41°48'N, 123°25'E) (2016) (Supplementary Figure S1A). A 6 × 6 block planting design was used for each cultivar, with three replicates.

Phenotype

The panicle numbers per plant, spikelets per panicle and grain yields per plant were measured from three plants in the middle of a planting block. Days to heading (DTH) were recorded as the period from the sowing date to the time when the first panicle appeared from the node (Koo et al., 2013). The PS index was calculated as $(DTH_{\text{shenyang}} - DTH_{\text{lingshui}}) / DTH_{\text{shenyang}}$ (Gao et al., 2014). Pearson's correlation coefficient calculations and a two-way analysis of variance for heading dates in the four environments were performed using SAS version 9.4.

Genotype

Total genomic DNA was extracted from the leaf samples of each accession, which had been grown in Lingshui in 2015. All the accessions were genotyped using Illumina custom designed arrays. Genotypes of these accessions were called using Genome Studio. After removing nucleotide variations with SNPs of low quality and minor allele frequencies < 0.05, we identified 21,198 SNPs, which were used further for GWAS analysis (Supplementary Table S2). We used MSU_version 6¹ to identify SNP and gene positions.

Population Genetic Analyses

To analyze the population structures of 244 varieties, the genetic components were calculated using Structure version 2.3 (Pritchard et al., 2000). In the analysis, the number of clusters (K) was set from 1 to 10 with a length of burn-in period of 10,000 steps and Markov chain Monte Carlo replications of 100,000 cycles, five runs independently for each K-value

¹http://rice.plantbiology.msu.edu/pub/data/Eukaryotic_Projects/o_sativa/annotation_dbs/pseudomolecules/version_6.0/

(Evanno et al., 2005). The optimal number of genetic clusters was identified when the ΔK was set at its maximum. Using the CLUMPP software, the results of all of the genetic components were combined to obtain the Q matrix, which was used for further association mapping (Jakobsson and Rosenberg, 2007). Major allele frequency, gene diversity, polymorphic information content and Nei's genetic distance (Nei, 1972) were assessed using PowerMarker version 3.25 (Liu and Muse, 2005). The construction of a neighbor-joining tree was based on pairwise Nei's genetic distances. The principal component analysis was estimated using the software NTSYSpc version 2.1 (Rohlf, 1988). The Kinship matrix was calculated based on SNP data using TASSEL version 4.0 (Bradbury et al., 2007). The heatmap of pairwise relative kinship values was constructed using the R package "Lattice."

Genome-Wide Association Study (GWAS)

For the GWAS, general linear (GLM) and mixed linear (MLM) models (with both Q and K matrices as covariates) were used. The observed *P*-values from marker-trait associations and the linkage disequilibrium parameter were calculated using TASSEL version 4.0 (Bradbury et al., 2007). The Manhattan plot and linkage disequilibrium block were constructed based on the results of TASSEL using Haploview 4.2 (Barrett et al., 2005). The Bonferroni-corrected threshold for the *P*-value was $1/21,198$ ($P = \alpha/n$, $\alpha = 1$) and $0.05/21,198$ ($\alpha = 0.05$), with corresponding $-\log_{10}(P)$ values of 4.33 and 5.63, respectively (Yang et al., 2014). The candidate genes at the identified loci were screened from the MSU Rice Genome Annotation Project Database².

Sequence Analysis

For PCR, primers for amplifying *Hd1*, *Ghd7*, and *DTH7* were designed based on the sequences of *japonica* cultivar Nipponbare from the Rice Genome Annotation Project Database³ using Primer3 (Rozen and Skaletsky, 2000). The primers for amplification and sequencing are listed in Supplementary Table S3. The 50- μ l total PCR reaction volume contained 25 μ l KOD FX PCR buffer, 10 μ l dNTPs, 1.5 μ l each 10 pmol/ μ l primer, 5 μ l template DNA, 6 μ l PCR grade water and 1 μ l KOD FX. The PCR cycle conditions were as follows: initial incubation at 94°C for 2 min, 35 cycles of 98°C for 10 s, 55°C for 30 s and 68°C at 1 kb/1 min, followed by a final extension at 68°C for 7 min⁴. Then, PCR products were purified by TaKaRa MiniBEST Agarose Gel DNA Extraction Kit Ver.4.0⁵ and sequenced in both directions with an ABI 3730xl DNA Analyzer⁶. The sequences were assembled using DNASTAR software. Sequences were aligned using the ClustalW method (Thompson et al., 2002), and a phylogenetic analysis using the neighbor-joining algorithm (Saitou and Nei, 1987) was constructed in MEGA version 5.05. Nucleotide diversity and two neutrality test parameters, Tajima's

D (Tajima, 1989), and Fu and Li's D (Fu and Li, 1993), were calculated with DnaSP version 5.0 (Rozas et al., 2003).

RESULTS

Genetic Characterization of *Japonica* Cultivars from Northeast China

We constructed a natural population composed of 244 *japonica* rice varieties that were selected from Northeast China. 50.8, 35.7, 10.7, and 2.9% of these varieties were selected from Liaoning, Heilongjiang, Jilin, and other provinces, respectively (Supplementary Figure S1C). The ancestry of these accessions can be traced back to four varieties, Liaojing5, Liaojing326, Songjing3, and Hejiang20 (Supplementary Table S1). After removing SNPs with minor allele frequencies less than 0.05 or greater than 20% missing data, only 21,198 SNPs were identified (Supplementary Table S2). These SNPs were distributed throughout all 12 chromosomes. The neighboring distances between 78.33% of the SNPs were less than 20 kb and only 3.6% of inter-SNP distances were greater than 80 kb (Supplementary Figure S2). The decay of the linkage disequilibrium with the physical distance between SNPs occurred at 550 kb (Supplementary Figure S3) when the cut off value of the linkage disequilibrium parameter was set to half its maximum (Huang et al., 2010).

Using these SNPs, we performed a genetic component analysis of each variety to quantify the population structure using Structure software. The ΔK achieved its maximum value when $k = 2$ (Figure 1A). Thus, the population was divided into two major subgroups and a mixed group based on the variety with *Q*-value >0.6 was appointed to corresponding Pop, and the rest were assigned to a mixed group. The geographic distribution analyses showed that all of the varieties of Heilongjiang and half varieties of Liaoning were assigned to Pop 1 (Figure 1B and Supplementary Table S4). Then, a neighbor-joining tree was constructed based on Nei's genetic distance. The resulting tree showed two major divergent subgroups separated, with some overlapping (Figure 1C). The principal component analysis showed that these accessions were clearly separated into three sub-groups, indicating that the 244 cultivars were appropriately divided into two major sub-clusters (Figure 1D). Furthermore, the population differentiation statistics value between Pop 1 and Pop 2 was 0.1, indicating that a low level of subpopulation differentiation existed in the 244 *japonica* varieties (Supplementary Table S5). Finally, the inner relative kinship of Pop 1 was higher than that of Pop 2, and the relative kinships of the two groups both represented a strong relatedness (Supplementary Figure S4). Moreover, the genetic diversity (0.25) of Pop 1 was lower than that (0.29) of Pop 2, and both indicated a low genetic diversity (Supplementary Table S6). Thus, there were close genetic relationships in the present populations, especially in Pop 1.

Phenotypic Variation in the Heading Date

All 244 *japonica* rice cultivars were grown in three distinct geographical locations, ranging from lower to higher latitudes, across China to investigate the mechanistic response to

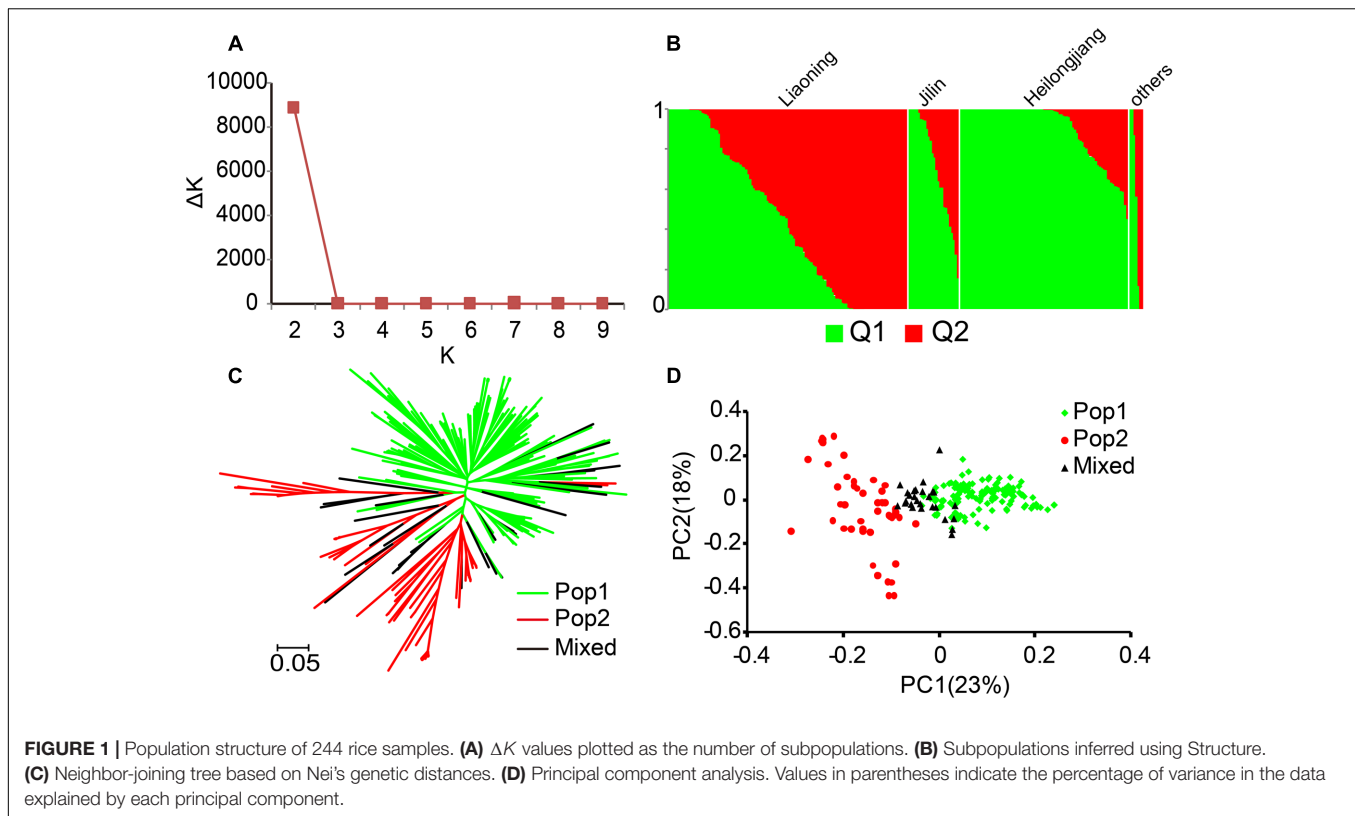
²<http://rice.plantbiology.msu.edu/index.shtml>

³<http://rapdb.dna.affrc.go.jp/>

⁴<http://www.toyobo-global.com>

⁵<http://www.takara.com.cn/>

⁶<http://www.appliedbiosystems.com/>



photoperiod (Supplementary Figure S1A). The correlation coefficients were both 0.80 under SD (2015 and 2016 in Lingshui) and LD (Hangzhou and Shenyang) conditions, while the remaining correlation coefficients were less than 0.7 (Supplementary Table S7). Two years of test data from Lingshui showed a similar pattern, thus we chose the 2015 test data without special mention for further analysis. A two-way analysis of variance revealed that the environment accounted for 78.15% of the variation in flowering time (Supplementary Table S8). Thus, environment had a dominant effect on flowering time.

Compared with accessions in Hangzhou (LD and high temperatures), the heading dates were significantly later in Shenyang (LD and low temperatures), and accessions tended to flower later in Shenyang than in Lingshui (SD and mid-range temperatures) (Figure 2A). Thus, the varieties were highly sensitive to the environment. In addition, part of Pop 2 was more sensitive to day length than part of Pop 1 (Figure 2B). The molecular mechanisms of PS are addressed below. Finally, spikelet number per panicle and grain yield per plant correlated positively and extremely significantly with flowering time (Figures 2C,D). However, flowering time influenced yield mainly through the spikelet number per panicle (Supplementary Table S9).

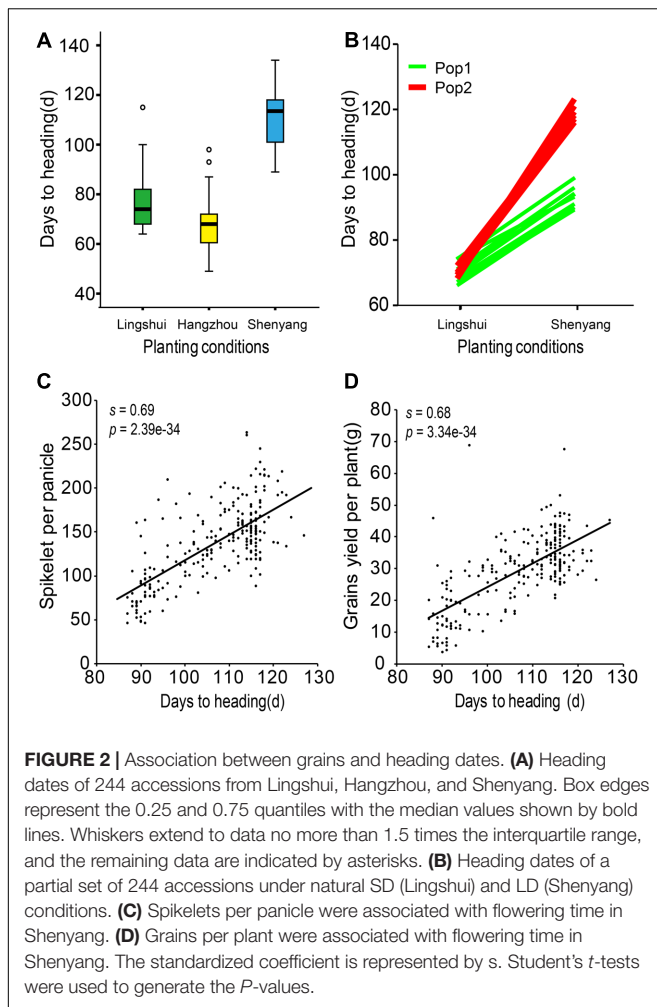
Association Mapping for Heading Date

To choose a better model for trait-marker associations, GLM and MLM were used to perform the association analysis. The results of quantile–quantile plots of estimated $-\log_{10}(P)$ showed that

the MLM was significantly better than the GLM in controlling false positive (Supplementary Figure S5). Thus, for heading date, we used an MLM for the GWAS. Only a significant SNP ($P < 2.36 \times 10^{-6}$) was found for heading date in three environments (Supplementary Table S10). It may be too strict to use $\alpha = 0.05$ for further analysis, so we examined the marginally significant association signals by using $\alpha = 1$. A total of nine loci were identified as having $P < 4.72 \times 10^{-5}$ ($\alpha = 1$) in a genome-wide scan of the three environments (two loci in Shenyang, six loci in Hangzhou, and one locus in Lingshui) (Figure 3 and Supplementary Table S10). Among the nine loci, three were known associations and the other six were newly identified. The GWAS produced a similar result using phenotypic data collected in 2016 and 2015 in Lingshui (Figure 3 and Supplementary Figure S6).

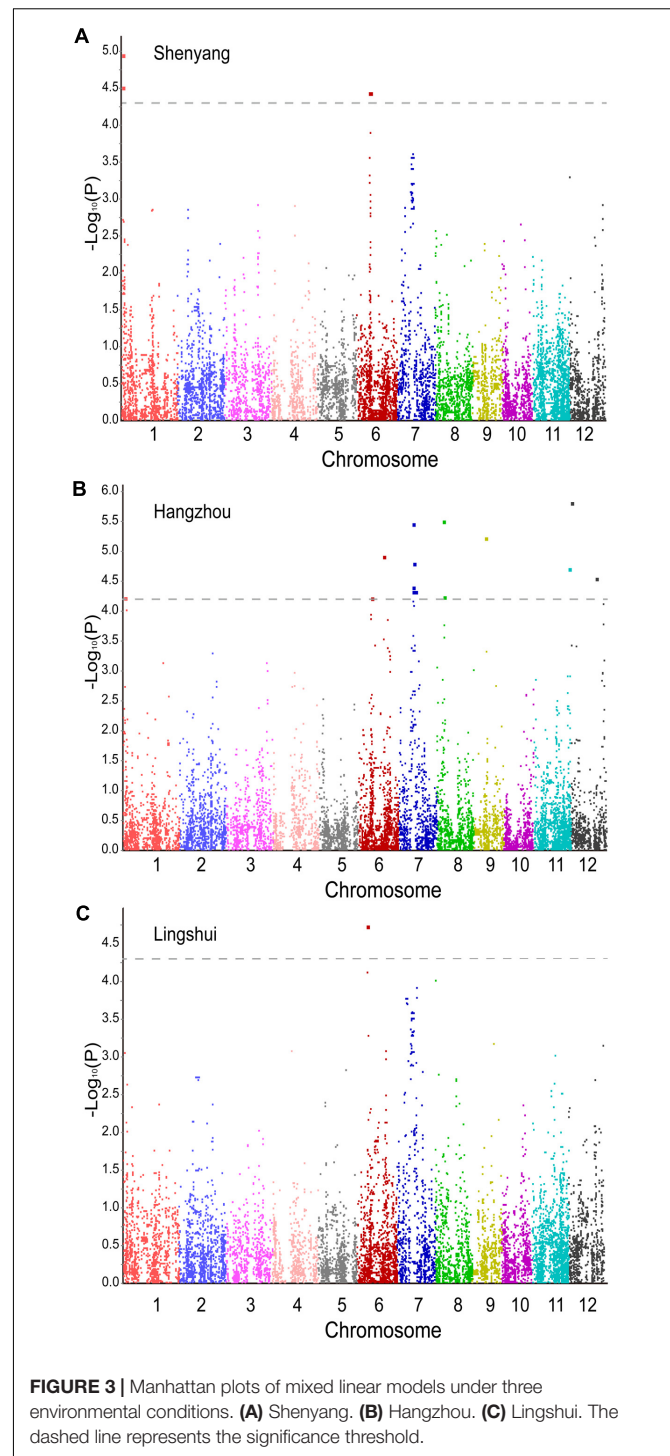
Indirect Association Owing to Allelic Heterogeneity

A peak on chromosome 6 was detected in both Shenyang and Lingshui, while the $-\log_{10}(P)$ of the peak (4.31) in Hangzhou was very close to the marginally significant threshold (4.33). Using a linkage disequilibrium block analysis we obtained a 106-kb candidate region (10.46–10.57 Mb) (Figure 4A). However, no good candidate gene was found in this region, which was consistent with a previous study (Yano et al., 2016). By checking the region nearby, we observed that *Hd1* was localized in the next linkage disequilibrium block (9.22–9.61 Mb). There are various haplotypes of *Hd1*, including null, intermediate and fully



functional alleles (Fujino et al., 2010), which may be the reason for the SNP markers could not achieve statistical significance in the region of *Hd1* (Yano et al., 2016).

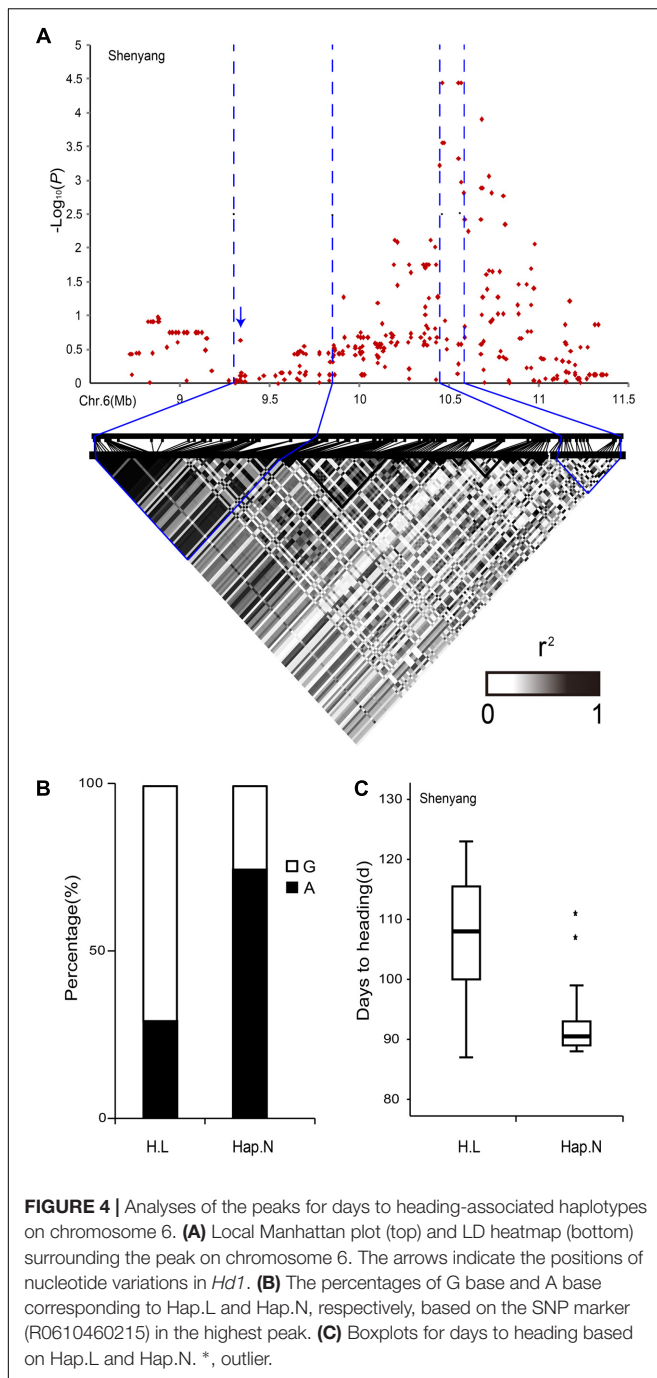
All 244 accessions were sequenced for *Hd1* to confirm the mis-association and whether it corresponded to previous studies or not. In total, 14 haplotypes were identified for *Hd1* and some of them caused frame-shift mutations, which were likely to be non-functional (Supplementary Figure S7A). The average phenotypic values were clearly separated between the frame-shift mutations (Hap.F and Hap.L) and non-frame-shift mutations (Hap.B and Hap.N). In this analysis, we ignored other varieties containing minor haplotypes (Supplementary Figure S7B). The 14 haplotypes were divided into two clades, haplotypes of clade 2 mainly had large insertion and deletion fragments at position 512 of *Hd1* compared with clade 1 (Supplementary Figure S7C). Hap.L mainly corresponded to guanine, while Hap.N mainly corresponded to adenine in the highest peak (R0610460215) based on SNPs (Figure 4B). Thus, Hap.L and Hap.N were clearly different. The accessions carrying Hap.N showed earlier heading dates than accessions carrying Hap.L (Figure 4C). These results indicated that *Hd1* was a causal gene in this region and demonstrated that high polymorphism level



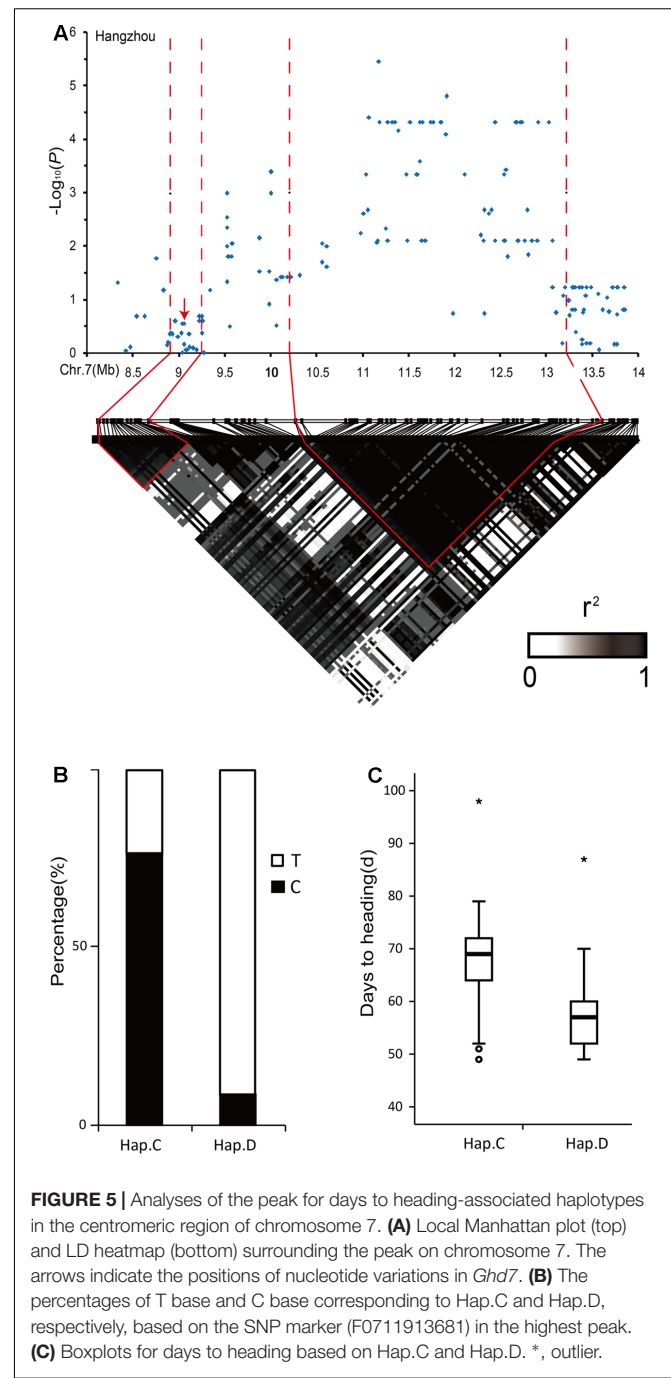
of a gene may lead to indirect associations (Galichon et al., 2012).

Indirect Association Owing to a High Linkage Disequilibrium Level

The lead SNP in the centromeric region of chromosome 7 was specifically identified in Hangzhou, while its $-\log_{10}(P)$



value exceeded 3 in Shenyang and Lingshui, and the value, as a compromised threshold, was used to detect the marginally significant association signals for GWAS (Feng et al., 2016). However, we could not find a good candidate gene in the region (10.22–13.21 Mb). When the candidate region covers a large distance (3 Mb) and is adjacent to the centromere, then the linkage relationships of the SNP markers are closer. Thus, we calculated the linkage disequilibrium block using a pairwise comparison of SNP markers separated by 5 Mb. This showed high level linkage relationship between the initial candidate region



and an adjacent linkage disequilibrium block (8.91–9.25 Mb) (Figure 5A). In the linkage disequilibrium block, *Ghd7* was a key quantitative trait loci that simultaneously controlled yield, plant height and heading date in rice (Xue et al., 2008).

We identified four haplotypes containing Hap.D with a premature stop codon in the coding sequence of *Ghd7* (Supplementary Figure S8A). The varieties of Hap.D had earlier heading dates than the others because of premature stop codon (Supplementary Figure S8B). The two clades were separated from these four haplotypes, and Hap.C and Hap.D in the 244

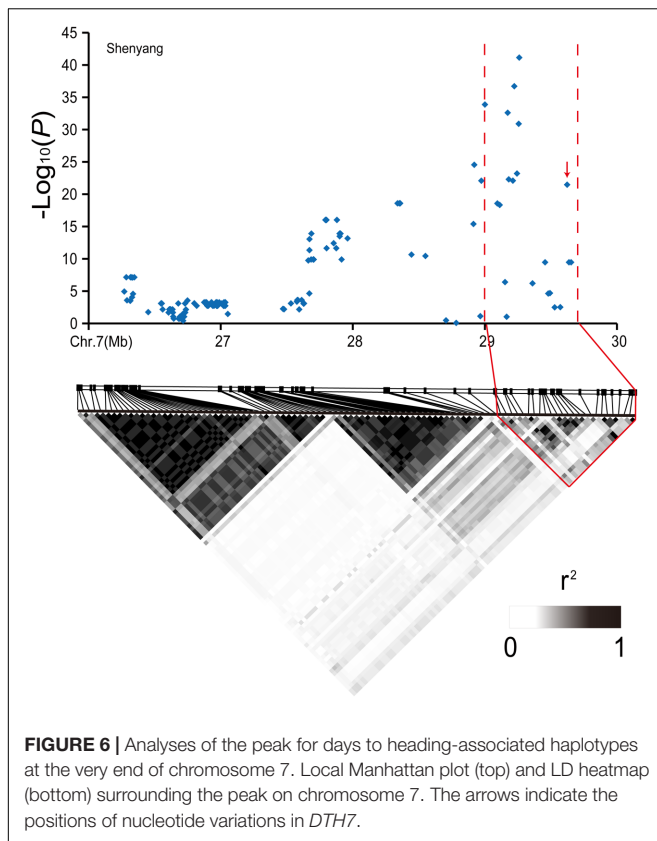


FIGURE 6 | Analyses of the peak for days to heading-associated haplotypes at the very end of chromosome 7. Local Manhattan plot (top) and LD heatmap (bottom) surrounding the peak on chromosome 7. The arrows indicate the positions of nucleotide variations in *DTH7*.

varieties were concentrated in clade 2 (Supplementary Figure S8C). Hap.C and Hap.D mainly corresponded to cytosine and thymine, respectively, in the SNP marker (F0711913681) at the highest peak in this linkage disequilibrium block (Figure 5B). The accessions carrying Hap.C had later heading dates than accessions carrying Hap.D (Figure 5C). These results indicated that *Ghd7* was confirmed as a causal gene in this region and demonstrated that high levels of linkage disequilibrium may also cause indirect associations (Grady et al., 2011).

Association Mapping Using GLM

Although the identification of heading date-related genes was effective using the MLM, some real genes may have been filtered out by the strict criteria. Therefore, association mapping was performed using the GLM. Two genes, *Hd1* and *Ghd7*, were identified using both MLM and GLM, while a locus on the distal end of chromosome 7 was only identified in Shenyang and Hangzhou using the GLM (Figure 6 and Supplementary Figure S9). In the linkage disequilibrium block of the candidate loci, *DTH7* is strongly regulated by day length (photoperiod) and temperature (Gao et al., 2014). There were five haplotypes with distinct phenotypes. Varieties carrying Hap.C had a later heading date than plants carrying Hap.B or Hap.D (Supplementary Figures S10A,B). The haplotypes of the cultivars were divided into two clades (Supplementary Figure S10C). The vast majority of varieties carrying Hap.B or Hap.C and both of the haplotypes belonged to the same clades.

Genetic Diversity among *Hd1*, *Ghd7*, and *DTH7*

To clarify whether *Hd1*, *Ghd7*, and *DTH7* had undergone selection during the expanded distribution and regional adaptation of *japonica* rice cultivars in Northeast China, we analyzed the nucleotide diversity (π) and the neutrality. The aligned lengths of exons that included insertion–deletions were as follows: 1,224 bp for *Hd1*, 774 bp for *Ghd7*, and 2,229 bp for *DTH7*. The π values for *Hd1* and *DTH7* were greater in Heilongjiang than in the other varieties, indicating more nucleotide diversity in Heilongjiang rice (Supplementary Table S11). However, *Ghd7* had similar π values in Heilongjiang (0.55) and Jilin (0.62). *Hd1* had the highest nucleotide diversity, especially in Heilongjiang ($\pi = 5.52$), while *DTH7* had the lowest variation in all regions. Tajima's D and Fu and Li's D statistics were not significant, except for those of *Hd1*, which were positive values.

Function of Allelic Variations of *Hd1*, *Ghd7*, and *DTH7*

Alleles of *Hd1* and *DTH7* were classified into functional and non-functional types (Supplementary Table S12). Hap.F, G, L and N of *Hd1* are non-functional alleles owing to frame-shift mutations or insertions of large fragments (Takahashi et al., 2009; Zhang et al., 2015; Yano et al., 2016). In addition to the haplotypes reported by previous studies, three new alleles of *Hd1*, Hap.C, D, and E, showed loss-of-functions because of frame-shift mutations. The alleles of Hap.B and Hap.D were equal in effect to *PRR37-2* and *PRR37-1*, respectively, in *DTH7*, while the alleles of Hap.C and Hap.E were the same as *PRR37-2a* and *PRR37-1a*, respectively, having no functions (Koo et al., 2013). The alleles of *Ghd7* could be classified into strongly and weakly functional, and non-functional types. Hap.B and Hap.C of *Ghd7* corresponded to *Ghd7-1* and *Ghd7-2*, respectively, which were confirmed to be strongly and weakly functional, respectively (Xue et al., 2008). Hap.D of *Ghd7* was the same as *Ghd7-0a*, with no function.

Regional Adaptation of the Three-Gene Combinations in Northeast China

The genetic effects of these three gene were clearly assigned to functional classifications. Thus, we investigated the effects of combinations of *Hd1*, *Ghd7*, and *DTH7* on heading dates in the present population (Li et al., 2015; Zhang et al., 2015). The combination of three non-functional loci (NNN) was associated with the earliest heading date under all conditions and was correlated with the lowest grain yield in Shenyang (Table 1). In general, the combinations functional *Hd1*, and non-functional *Ghd7* and *DTH7* (FNN), functional *Hd1*, weak functional *Ghd7*, and non-functional *DTH7* (FWN), and weak functional *Ghd7*, and non-functional *Hd1* and *DTH7* (NWN) exhibited similar heading dates, PS indices and grain yields under all three conditions. Functional *DTH7*, and non-functional *Hd1* and *Ghd7* (NNF) plants headed later and produced greater grain yields per plant than the plants with non-functional *DTH7* in Shenyang, where the day length is very long in the summer.

TABLE 1 | Heading date and grain yield associated with japonica rice cultivars from Northeast China combinations of *Hd1*, *Ghd7*, and *DTH7*.

Combinations			Heading date(d)			PS index	Grain yield(g)
<i>Hd1</i>	<i>Ghd7</i>	<i>DTH7</i>	Lingshui	Hangzhou	Shenyang		Shenyang
N	N	N	68.10 ± 2.28	56.67 ± 4.65	92.71 ± 5.7	0.264 ± 0.04	18.25 ± 8.19
F	N	N	71.83 ± 12.84	57.67 ± 14.47	93.33 ± 9.31	0.235 ± 0.06	19.97 ± 8.82
F	W	N	70.33 ± 3.21	61.67 ± 1.15	94.67 ± 2.31	0.257 ± 0.03	19.51 ± 8.44
N	W	N	72.06 ± 8.85	59.30 ± 7.61	95.24 ± 7.45	0.244 ± 0.05	19.05 ± 7.14
N	N	F	69.67 ± 3.08	59.33 ± 3.93	99.5 ± 5.79	0.299 ± 0.03	29.29 ± 7.14
N	W	F	77.82 ± 7.63	69.07 ± 4.87	111.42 ± 7.12	0.300 ± 0.07	33.46 ± 9.37
F	W	F	77.00 ± 12.79	75.33 ± 11.99	114.44 ± 10.49	0.319 ± 0.15	34.27 ± 5.94

N, non-functional haplotypes; F, functional haplotypes; W, weak functional haplotypes.

However, there was little difference between heading dates of NNF plants and non-functional *DTH7* plants in Lingshui and Hangzhou. The heading dates were extremely delayed and grain yields were significantly increased in both non-functional *Hd1*, weak functional *Ghd7* and functional *DTH7* (NWF) and weak functional *Ghd7*, and functional *Hd1* and *DTH7* (FWF) plants. Moreover, plants with functional *DTH7* displayed significantly higher PS index values than plants with non-functional *DTH7*, indicating that functional *DTH7* was associated with a strong PS.

The accumulated temperature gradually decreased from south to north in Northeast China, while it was divided into three zones in Heilongjiang Province (Supplementary Figure S1B). The different combinations of *Hd1*, *Ghd7*, and *DTH7* were examined in three provinces. The main combination found in Liaoning and Jilin Provinces was NWF, which is generally associated with a long lifespan and high grain yield (Figure 7A and Table 1). However, NNF plants were mainly distributed in Jilin Province. As expected, the FWF combination did not appear in Heilongjiang Province because of its later heading dates. Thus, we examined the geographical distributions of three-gene combinations in the accumulated temperature zones of Heilongjiang Province. As in Liaoning and Jilin Provinces, the major combination in the first accumulated temperature zone was NWF (Figure 7A). The NNN and NWN combinations, associated with a short lifespan and low PS were mainly distributed in the second and third accumulated temperature zones, and particularly in the latter. These results explained why part of Pop 2 had a greater PS index than Pop1 (Figure 2B).

To investigate the presence of different gene combinations in rice farming, we examined the gene distributions in cultivars of the top five varieties having planting areas. NWF plants played a core role and accounted for most of the cultivars in Liaoning, Jilin and the first accumulated temperature zone of Heilongjiang (Supplementary Table S13). In addition, NWN and NNN plants were grown as main rice cultivars in the second and third accumulated temperature zones of Heilongjiang, respectively. Notably, 'Kongyu131,' which had been planted on over 500,000 hectares a year, was the major rice variety in Heilongjiang Province from 2002 to 2012. After 2013, 'Longjing31' was cultivated on over 1 million

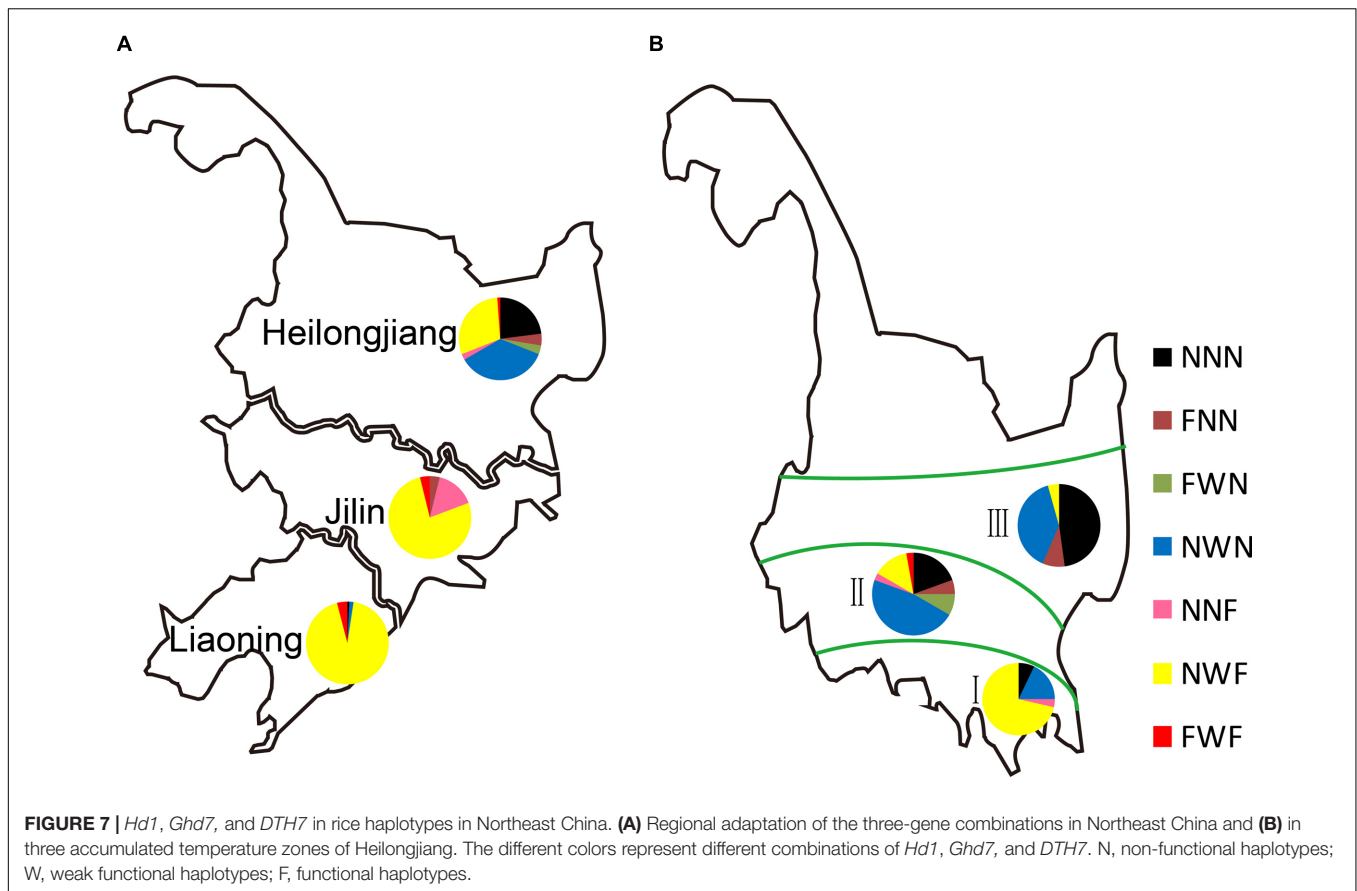
hectares a year, which is the largest planting area of a japonica rice variety in China. This is attributed to the combination NNN of two varieties that are high adaptable, which is affected by the day length and temperature in Heilongjiang Province.

DISCUSSION

Japonica Cultivars Are Closely Related in Northeast China

Rice is a staple food and is largely cultivated in Northeast China. In recent decades, rice cultivation areas in Northeast China have increased rapidly because of the good production quality and high economic benefit. It is of both academic and practical value to research the genetic architecture of rice germplasm resources (Xu et al., 2016). Thus, 244 rice materials originating from Northeast China, having a combined 21,198 SNPs that produce high quality genotyping results covering the whole genome, were used to investigate population structures and relationships (Supplementary Figure S1 and Supplementary Table S2). Limited SSR markers have been used in rice classifications in previous studies (Wei et al., 2009; Wang et al., 2014), while a large number of SNPs were used for clustering in this study, increasing the accuracy and reliability of the results. Similar results (low genetic distances and gene diversity levels) were obtained using Powermarker for the population structure analysis, indicating that japonica rice cultivars were limited in their hereditary base and had similar genetic backgrounds in Northeast China.

Cultivars have evolved a particular heading behavior in this region after long-term domestication and selection (Li et al., 2015). Since the day length is long and temperature is low, rice cultivars from other regions generally exhibit a long growth period and are unable to adapt to the unique conditions (Wei et al., 2008). As a consequence, the choice of materials is very limited for artificial breeding and it is difficult to introduce the genetic backgrounds of cultivars from other areas into those of the Northeast cultivars. Eventually, this resulted in the close relationship among japonica rice cultivars in Northeast China. The cultivars of Pop 2, which mostly originated from Liaoning Province, exhibited slight higher levels of variability



(Figure 1). The new genes from the rich biological resources of the southern region were easily introduced to the cultivars of Liaoning, which is geographically closer to the southern region.

For the GWAS, the population structure, which widely existed in the natural groups, usually increased linkage disequilibrium values between chromosomes, leading to false positives (Rafalski and Morgante, 2004). The relationships between individuals also had an impact on the results of the association analysis (Flint-Garcia et al., 2003, 2005). A close relationship of the population in this study exhibiting large extent of linkage disequilibrium may cause numerous SNPs associations on the different chromosomes. It is difficult to identify unknown genes owing to many candidate genes in a linkage disequilibrium block. But the QTLs with known genes can be easily identified in a local population with low genetic complexity and the genetic architecture (Fujino et al., 2015). The SNP associations may generate spurious associations between the phenotype and unlinked SNPs. According to a previous report, an association model with Q and K as covariates was more able than other models to reduce the rates of false positives (Yu et al., 2006). Likewise, the MLM was significantly better than the GLM in controlling corrections in this study (Supplementary Figure S5). Although false positives were well controlled, the MLM produced false negatives. A close relationship with strict parameters may cause true

loci to be eliminated in the MLM. Thus, the MLM was regarded as an ideal model to reduce false positive associations, while the GLM was necessary for the GWAS to avoid false negatives.

Hd1, *Ghd7*, and *DTH7* Combinations Determined the Different Rice Growing Regions

Grain yield is positively correlated with heading date, which is a vital factor in controlling rice yield (Gao et al., 2014). Varieties with an appropriate heading date can fully utilize the light and temperature resources to improve grain yields. *Ghd7* and *DTH7* play important roles in rice adaptation to LD conditions by reducing PS and causing early flowering (Xue et al., 2008; Koo et al., 2013). *Hd1*, a major PS quantitative trait locus, functions as a floral suppressor by forming a complex with *Ghd7* and *DTH7*, although *Ghd7* and *DTH7* can delay flowering without the functional *Hd1* allele under LD conditions (Yano et al., 2000; Nemoto et al., 2016; Goretti et al., 2017). In this study, *Ghd7* and *DTH7* as major repressors of flowering were detected by the GWAS (Figures 5, 6). Asian cultivated rice, which contains functional alleles of *Hd1*, *Ghd7*, and *DTH7* that have strong effects, originated from low-latitude regions with SD lengths. After long-term selection and domestication, the

alleles of *Hd1*, *Ghd7*, and *DTH7* with weak or no functions were generated to reduce the effects of prolonging heading in high-latitude regions with a LD lengths (Table 1 and Supplementary Table S12).

When rice was expanded to the most northerly regions, varieties harboring NNN were selected owing to their early heading dates and lack of PS (Figure 7). 'Kongyu131' and 'Longjing31,' carrying NNN, are mainly grown in the third accumulated temperature zone and possess the largest cropping area per year in China (Supplementary Table S13). FNN, NWN, and FWN plants showing phenotypes similar to that of NNN plants were widely grown in the second and third accumulated temperature zones. This phenomenon was because *Hd1* had no function by itself and *Ghd7* had a weak function. The varieties having NWF were widely planted in Northeast China because they had suitable flowering times that allow them to fully utilize the light and temperature resources. FWF plants, which exhibited a long flowering time, were mainly grown south of Liaoning. To summarize, loss-of-function alleles of *Hd1*, *Ghd7*, and *DTH7* existed in the rice grown in northerly regions because of their association with early heading dates, while functional alleles of *Hd1*, *Ghd7*, and *DTH7* existed in rice grown in southerly regions because of their association with late heading dates in Northeast China.

Selection of Flowering Genes as a Tool for Breeding

Heading date is a complex trait and is seriously affected by environmental factors (Sun et al., 2014). Local varieties are usually adapted to local environments, including proper heading dates. The functions of *Hd1*, *Ghd7*, and *DTH7* were clear, and their effects could be additive. Therefore, predicting heading dates in hybrid progenies could be possible based on their parental varieties. For instance, no significant segregation should be observed in the cross progenies of two parents with the same combination of these three genes. The late heading of F₁ hybrids might occur, and progeny will be different from the two parents with complementary *Hd1*, *Ghd7*, and *DTH7* genes. For example, an FWN plant crossed with an NNF plant. The optimized heading date could be designed using combinations of these three genes. The optimum combination would be selected to match the main cultivars in the locality. For instance, NNN plants would be selected in the third accumulated temperature zone.

The genetic diversity of varieties can improve the resistance of cultivars to disease (Zhu et al., 2000). Generally, the goals of crop breeding are high yield and quality to meet the supply demands, but this reduces the genetic diversity of the population, which, in turn, decreases the ability of crops to respond to stress. The long-term major aim of breeders in Northeast China is to produce higher and higher yields, resulting in many problems, including the loss of genetic diversity. In addition, the unique climate and single variety type (single-season rice) may lead to reductions in genetic diversity. Introducing genes from

varieties in other regions into the Northeast *japonica* varieties may result in a sustained development of rice production. The early *indica* rice, which was mainly cultivated in the Yangtze region and southern China, exhibits no or low PS and has a short growth duration. Part of cultivars exhibited no PS genes, having mutations in *Hd1*, *Ghd7*, and *DTH7*, while the remaining cultivars carried either functional *Hd1* or *Ghd7* genes (Xu et al., 2010). Thus, in terms of heading date, it is a reasonable strategy to breed new cultivars by selecting progenies from Northeast *japonica* rice and early *indica* rice, which can broaden the genetic base of the breeding materials and improve the genetic diversity of rice varieties in Northeast China.

Obvious heterosis has been found among *indica* × *japonica* crosses, which represent an important strategy to enhance rice yield potentials (Zheng and Yong Gen, 1999). The late heading dates in *indica* × *japonica* F₁ hybrids occur because of interactions between the heading date-related genes of the two parents, such as *Hd1* and *Ghd7* (Wei et al., 2010a). Heterosis utilization may become a reality in Northeast China, with one major problem: the long heading dates of F₁ hybrids. Thus, two parents with the same heading date-related gene or no PS gene (e.g., combination NNN) may be the key to obtaining suitable heading stages in F₁ hybrids of *indica* × *japonica*. Introducing early *indica* rice, and particularly those cultivars with no PS, might facilitate heterosis utilization in Northeast China. Those could provide guidance for molecular designed breeding of rice with specific heading date.

AUTHOR CONTRIBUTIONS

JY, XN, and XW: designed the research experiments. JY, YF, and ShaW: performed the phenotyping. QX, YF, and ShuW: carried out the genetic studies. XY, HY, and YW: managed the materials. XW: designed the overall project. JY and XN: analyzed the data and drafted the manuscript. ShuW and YF: helped to revise the manuscript. All of the authors read and approved the final manuscript.

ACKNOWLEDGMENTS

This study was supported by grants from the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (NSFC 31521064), the Chinese Academy of Agricultural Sciences (CAAS-ASTIP-201X-CNRRI) and the Natural Science Foundation of Zhejiang Province (LY15C130005).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.00035/full#supplementary-material>

REFERENCES

- 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. doi: 10.1093/bioinformatics/bth457
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Brambilla, V., and Fornara, F. (2013). Molecular control of flowering in response to day length in rice. *J. Integr. Plant Biol.* 55, 410–418. doi: 10.1111/jipb.12033
- Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., et al. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661–678. doi: 10.1038/nature05911
- Dai, X. D., Ding, Y. N., Tan, L. B., Fu, Y. C., Liu, F. X., Zhu, Z. F., et al. (2012). LHD1, an allele of *dth8/ghd8*, controls late heading date in common wild rice (*Oryza rufipogon*). *J. Integr. Plant Biol.* 54, 790–799. doi: 10.1111/j.1744-7909.2012.01166.x
- Doi, K., Izawa, T., Fuse, T., Yamanouchi, U., Kubo, T., Shimatani, Z., et al. (2004). Ehd1, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-like gene expression independently of Hd1. *Genes Dev.* 18, 926–936. doi: 10.1101/gad.1189604
- Dong, H., Zhao, H., Xie, W., Han, Z., Li, G., Yao, W., et al. (2016). A novel tiller angle gene, TAC3, together with TAC1 and D2 largely determine the natural variation of tiller angle in rice cultivars. *PLOS Genet.* 12:e1006412. doi: 10.1371/journal.pgen.1006412
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x
- Feng, Y., Lu, Q., Zhai, R., Zhang, M., Xu, Q., Yang, Y., et al. (2016). Genome wide association mapping for grain shape traits in indica rice. *Planta* 244, 819–830. doi: 10.1007/s00425-016-2548-9
- Flint-Garcia, S. A., Thornsberry, J. M., and Buckler, E. S. (2003). Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54, 357–374. doi: 10.1146/annurev.arplant.54.031902.134907
- Flint-Garcia, S. A., Thuillet, A. C., Yu, J., Pressoir, G., Romero, S. M., Mitchell, S. E., et al. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J.* 44, 1054–1064. doi: 10.1111/j.1365-3113X.2005.02591.x
- Fu, Y. X., and Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics* 133, 693–709.
- Fujino, K., Obara, M., Shimizu, T., Koyanagi, K. O., and Ikegaya, T. (2015). Genome-wide association mapping focusing on a rice population derived from rice breeding programs in a region. *Breed. Sci.* 65, 403–410. doi: 10.1270/jsbbs.65.403
- Fujino, K., Wu, J., Sekiguchi, H., Ito, T., Izawa, T., and Matsumoto, T. (2010). Multiple introgression events surrounding the *Hd1* flowering-time gene in cultivated rice, *Oryza sativa* L. *Mol. Genet. Genomics* 284, 137–146. doi: 10.1007/s00438-010-0555-2
- Fujino, K., Yamanouchi, U., and Yano, M. (2013). Roles of the Hd5 gene controlling heading date for adaptation to the northern limits of rice cultivation. *Theor. Appl. Genet.* 126, 611–618. doi: 10.1007/s00122-012-2005-5
- Galichon, P., Mesnard, L., Hertig, A., Stengel, B., and Rondeau, E. (2012). Unrecognized sequence homologies may confound genome-wide association studies. *Nucleic Acids Res.* 40, 4774–4782. doi: 10.1093/nar/gks169
- Gao, H., Jin, M., Zheng, X. M., Chen, J., Yuan, D., Xin, Y., et al. (2014). Days to heading 7, a major quantitative locus determining photoperiod sensitivity and regional adaptation in rice. *Proc. Natl. Acad. Sci. U.S.A.* 111, 16337–16342. doi: 10.1073/pnas.1418204111
- Goretti, D., Martignago, D., Landini, M., Brambilla, V., Gomez-Ariza, J., Gnesutta, N., et al. (2017). Transcriptional and post-transcriptional mechanisms limit heading date 1 (Hd1) function to adapt rice to high latitudes. *PLOS Genet.* 13:e1006530. doi: 10.1371/journal.pgen.1006530
- Grady, B. J., Torstenson, E. S., and Ritchie, M. D. (2011). The effects of linkage disequilibrium in large scale SNP datasets for MDR. *BioData Min.* 4:11. doi: 10.1186/1756-0381-4-11
- Gu, B., Zhou, T., Luo, J., Liu, H., Wang, Y., Shangguan, Y., et al. (2015). *An-2* encodes a cytokinin synthesis enzyme that regulates awn length and grain production in rice. *Mol. Plant* 8, 1635–1650. doi: 10.1016/j.molp.2015.08.001
- Gupta, P. K., Rustgi, S., and Kulwal, P. L. (2005). Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Mol. Biol.* 57, 461–485. doi: 10.1007/s11103-005-0257-z
- Hamblin, M. T., Buckler, E. S., and Jannink, J. L. (2011). Population genetics of genomics-based crop improvement methods. *Trends Genet.* 27, 98–106. doi: 10.1016/j.tig.2010.12.003
- Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42, 961–967. doi: 10.1038/ng.695
- Ishikawa, R., Aoki, M., Kurotani, K., Yokoi, S., Shinomura, T., Takano, M., et al. (2011). Phytochrome B regulates *Heading date 1 (Hd1)*-mediated expression of rice florigen *Hd3a* and critical day length in rice. *Mol. Genet. Genomics* 285, 461–470. doi: 10.1007/s00438-011-0621-4
- Itoh, H., Nonoue, Y., Yano, M., and Izawa, T. (2010). A pair of floral regulators sets critical day length for *Hd3a* florigen expression in rice. *Nat. Genet.* 42, 635–638. doi: 10.1038/ng.606
- Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806. doi: 10.1093/bioinformatics/btm233
- Kojima, S., Takahashi, Y., Kobayashi, Y., Monna, L., Sasaki, T., Araki, T., et al. (2002). *Hd3a*, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. *Plant Cell Physiol.* 43, 1096–1105. doi: 10.1093/pcp/pcf156
- Komiya, R., Ikegami, A., Tamaki, S., Yokoi, S., and Shimamoto, K. (2008). *Hd3a* and *RFT1* are essential for flowering in rice. *Development* 135, 767–774. doi: 10.1242/dev.008631
- Komiya, R., Yokoi, S., and Shimamoto, K. (2009). A gene network for long-day flowering activates *RFT1* encoding a mobile flowering signal in rice. *Development* 136, 3443–3450. doi: 10.1242/dev.040170
- Koo, B. H., Yoo, S. C., Park, J. W., Kwon, C. T., Lee, B. D., An, G., et al. (2013). Natural variation in *OsPRR37* regulates heading date and contributes to rice cultivation at a wide range of latitudes. *Mol. Plant* 6, 1877–1888. doi: 10.1093/mp/sst088
- Li, X. F., Liu, H. Z., Wang, M. Q., Liu, H. L., Tian, X. J., Zhou, W. J., et al. (2015). Combinations of *Hd2* and *Hd4* genes determine rice adaptability to Heilongjiang Province, northern limit of China. *J. Integr. Plant Biol.* 57, 698–707. doi: 10.1111/jipb.12326
- Lindblad-Toh, K., Winchester, E., Daly, M. J., Wang, D. G., Hirschhorn, J. N., Lavolette, J. P., et al. (2000). Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nat. Genet.* 24, 381–386. doi: 10.1038/74215
- Liu, K., and Muse, S. V. (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21, 2128–2129. doi: 10.1093/bioinformatics/bti282
- Nei, M. (1972). Genetic distance between populations. *Am. Nat.* 106, 283–292. doi: 10.1086/282771
- Nemoto, Y., Nonoue, Y., Yano, M., and Izawa, T. (2016). *Hd1a* *CONSTANS* ortholog in rice, functions as an *Ehd1* repressor through interaction with monocot-specific CCT-domain protein *Ghd7*. *Plant J.* 86, 221–233. doi: 10.1111/tpj.13168
- Primmer, C. R., Borge, T., Lindell, J., and Saetre, G. P. (2002). Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. *Mol. Ecol.* 11, 603–612. doi: 10.1046/j.0962-1083.2001.01452.x
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Putterlick, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (1995). The *CONSTANS* gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* 80, 847–857. doi: 10.1016/0092-8674(95)90288-0
- Rafalski, A., and Morgante, M. (2004). Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. *Trends Genet.* 20, 103–111. doi: 10.1016/j.tig.2003.12.002

- Rohlf, F. J. (1988). *NTSYS-pc 2.0. Numerical Taxonomy and Multivariate Analysis System*. New York, NY: Stony Brook University.
- Rozas, J., Sánchez-Delbarrio, J. C., Messeguer, X., and Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19, 2496–2497. doi: 10.1093/bioinformatics/btg359
- Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132, 365–386.
- Saito, H., Ogiso-Tanaka, E., Okumoto, Y., Yoshitake, Y., Izumi, H., Yokoo, T., et al. (2012). *Ej7* encodes an ELF3-like protein and promotes rice flowering by negatively regulating the floral repressor gene *Ghd7* under both short- and long-day conditions. *Plant Cell Physiol.* 53, 717–728. doi: 10.1093/pcp/pcs029
- Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.
- Shang, H.-X., Ye, S.-H., Deng, X.-M., Zhou, Y., Xiu, F.-L., Ji, X.-J., et al. (2012). Identification and fine mapping of heading date related mutant gene in rice. *Rice Sci.* 19, 269–276. doi: 10.1016/S1672-6308(12)60051-1
- Shimamoto, K., and Yokoi, S. (2005). *The Photoperiodic Control of Flowering in Rice, a Short-Day Plant*. Tokyo: Springer-Verlag. doi: 10.1007/4-431-27092-2_40
- Shrestha, R., Gomez-Ariza, J., Brambilla, V., and Fornara, F. (2014). Molecular control of seasonal flowering in rice, arabidopsis and temperate cereals. *Ann. Bot.* 114, 1445–1458. doi: 10.1093/aob/mcu032
- Sun, C., Chen, D., Fang, J., Wang, P., Deng, X., and Chu, C. (2014). Understanding the genetic and epigenetic architecture in complex network of rice flowering pathways. *Protein Cell* 5, 889–898. doi: 10.1007/s13238-014-0068-6
- Syvänen, A. C. (2001). Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat. Rev. Genet.* 2, 930–942. doi: 10.1038/35103535
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.
- Takahashi, Y., Teshima, K. M., Yokoi, S., Innan, H., and Shimamoto, K. (2009). Variations in *Hd1* proteins, *Hd3a* promoters, and *Ehd1* expression levels contribute to diversity of flowering time in cultivated rice. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4555–4560. doi: 10.1073/pnas.0812092106
- Tamaki, S., Matsuo, S., Wong, H. L., Yokoi, S., and Shimamoto, K. (2007). *Hd3a* protein is a mobile flowering signal in rice. *Science* 316, 1033–1036. doi: 10.1126/science.1141753
- Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2002). Multiple sequence alignment using ClustalW and ClustalX. *Curr. Protoc. Bioinformatics* 00:2.3.1-2.3.22 doi: 10.1002/0471250953.bi0203s00
- Wang, C. H., Zheng, X. M., Xu, Q., Yuan, X. P., Huang, L., Zhou, H. F., et al. (2014). Genetic diversity and classification of *Oryza sativa* with emphasis on Chinese rice germplasm. *Heredity* 112, 489–496. doi: 10.1038/hdy.2013.130
- Wei, X., Liu, L., Xu, J., Jiang, L., Zhang, W., Wang, J., et al. (2010a). Breeding strategies for optimum heading date using genotypic information in rice. *Mol. Breed.* 25, 287–298. doi: 10.1007/s11032-009-9332-5
- Wei, X., Xu, J., Guo, H., Jiang, L., Chen, S., Yu, C., et al. (2010b). DTH8 suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiol.* 153, 1747–1758. doi: 10.1104/pp.110.156943
- Wei, X., Yuan, X., Yu, H., Wang, Y., Xu, Q., and Tang, S. (2009). Temporal changes in SSR allelic diversity of major rice cultivars in China. *J. Genet. Genomics* 36, 363–370. doi: 10.1016/s1673-8527(08)60125-3
- Wei, X. J., Jiang, L., Xu, J. F., Zhang, W. W., Lu, G. W., Zhang, Y. S., et al. (2008). Genetic analyses of heading date of Japonica rice cultivars from Northeast China. *Field Crops Res.* 107, 147–154. doi: 10.1016/j.fcr.2008.01.008
- Xu, J., Wei, X., Jiang, L., Lu, G., Wang, H., Zhou, Z., et al. (2010). Genetic analysis of heading date of some early season indica rice cultivars and hybrid rice parents in China. *Chin. J. Rice Sci.* 24, 215–222.
- Xu, Q., Yuan, X., Wang, S., Feng, Y., Yu, H., Wang, Y., et al. (2016). The genetic diversity and structure of indica rice in China as detected by single nucleotide polymorphism analysis. *BMC Genet.* 17:53. doi: 10.1186/s12863-016-0361-x
- Xue, W., Xing, Y., Weng, X., Zhao, Y., Tang, W., Wang, L., et al. (2008). Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat. Genet.* 40, 761–767. doi: 10.1038/ng.143
- Yan, W., Liu, H., Zhou, X., Li, Q., Zhang, J., Lu, L., et al. (2013). Natural variation in *Ghd7.1* plays an important role in grain yield and adaptation in rice. *Cell Res.* 23, 969–971. doi: 10.1038/cr.2013.43
- Yan, W. H., Wang, P., Chen, H. X., Zhou, H. J., Li, Q. P., Wang, C. R., et al. (2011). A major QTL. *Ghd8*, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Mol. Plant* 4, 319–330. doi: 10.1093/mp/ssq070
- Yang, N., Lu, Y., Yang, X., Huang, J., Zhou, Y., Ali, F., et al. (2014). Genome wide association studies using a new nonparametric model reveal the genetic architecture of 17 agronomic traits in an enlarged maize association panel. *PLoS Genet.* 10:e1004573. doi: 10.1371/journal.pgen.1004573
- Yang, Y., Peng, Q., Chen, G.-X., Li, X.-H., and Wu, C.-Y. (2013). OsELF3 is involved in circadian clock regulation for promoting flowering under long-day conditions in rice. *Mol. Plant* 6, 202–215. doi: 10.1093/mp/sss062
- Yano, K., Yamamoto, E., Aya, K., Takeuchi, H., Lo, P. C., Hu, L., et al. (2016). Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat. Genet.* 48, 927–934. doi: 10.1038/ng.3596
- Yano, M., Katayose, Y., Ashikari, M., Yamanouchi, U., Monna, L., Fuse, T., et al. (2000). *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene *CONSTANS*. *Plant Cell* 12, 2473–2484. doi: 10.1105/tpc.12.12.2473
- Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208. doi: 10.1038/ng1702
- Zhang, J., Zhou, X., Yan, W., Zhang, Z., Lu, L., Han, Z., et al. (2015). Combinations of the *Ghd7*, *Ghd8* and *Hd1* genes largely define the ecogeographical adaptation and yield potential of cultivated rice. *New Phytol.* 208, 1056–1066. doi: 10.1111/nph.13538
- Zheng, G. Q., and Yong Gen, L. U. (1999). Breeding of the indica compatible japonica lines and their use in the breeding of super high yield hybrid rice. *Hybrid Rice* 6, 3–5. doi: 10.16267/j.cnki.1005-3956.1999.06.002
- Zhu, Y., Chen, H., Fan, J., Wang, Y., Li, Y., Chen, J., et al. (2000). Genetic diversity and disease control in rice. *Nature* 406, 718–722. doi: 10.1038/35021046

Conflict of Interest Statement: 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.

Copyright © 2018 Ye, Niu, Yang, Wang, Xu, Yuan, Yu, Wang, Wang, Feng and Wei. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.