Resequencing and variation identification of whole genome of the *japonica* rice variety "Longdao24" with high yield

Shukun Jiang *, Shichen Sun, Liangming Bai, Guohua Ding, Tongtong Wang, Tianshu Xia, Hui Jiang, Xijuan Zhang, Fengming Zhang

Cultivation and Farming Research Institute, Heilongjiang Academy of Agricultural Sciences, Harbin, P.R. China

* sk.jiang@hotmail.com

Abstract

*Japonica* rice mainly distributes in north of China, which accounts for more than half of the total *japonica* rice cultivated area of China. High yield, good grain quality and early heading date were the main breeding traits and commercial property in this region. We performed resequencing and genome wide variation analysis of one typical northern *japonica* rice variety Longdao24 and its parents (Longdao5 and Jigeng83) using the Illumina sequencing technology. 53.17 G clean bases were generated and more than 96.8% of the reads were mapped to the genomic reference sequence. An overall average effective depth of 43.67× coverage was achieved. We identified 420,475 SNPs, 95,624 InDels, and 14,112 SVs in Longdao24 genome with the genomic sequence of the *japonica* cultivar Nipponbare as reference. We identified 361,117 SNPs and 81,488 InDels between Longdao24 genome and Longdao5 genome. We also detected 428,908 SNPs and 97,209 InDels between Longdao24 genome and Jigeng83 genome. Twenty-two yield related genes, twenty-two grain quality related genes and thirty-nine heading date genes were analyzed in Longdao24. The alleles of *Gn1a*, *EP3*, *SCM2*, *Wx*, *ALK*, *OsLF* and *Hd17* came from the female parent Longdao5. The other alleles of *qGW8*, *SSIVa*, *SBE3*, *SSIIIb*, *SS1ic*, *DTH2*, *Ehd3* and *OsMADS56* came from the male parent Jigeng83. These results will help us to research the genetics basis of yield, grain quality and early heading date in northern rice of China.

Introduction

Given continuing population growth and increasing competition for arable land between food and energy crops, the next century may witness serious global food shortage problems. Thus, there is a need for an increase in rice yield, because rice is the world’s most important staple food crop. To substantially increase rice yield, Japan initiated a super high-yielding rice breeding program in 1981 that was targeted to raise rice yield 50% within 15 years. However, this target has not yet been realized [1, 2]. Next, the International Rice Research Institute (IRRI) launched a new plant-type breeding program to develop super rice in 1989, with the target of developing a super rice. However, this target was not reached either [1–3]. In order to meet the
food demand required by the Chinese people in the 21st century, a program to breed super rice through combining morphological improvement and the utilization of inter-subspecific heterosis was set up by the Ministry of Agriculture and the Ministry of Science and Technology in 1996 and 1997, respectively [1, 2, 4]. Today, nearly 80 super rice varieties have been released and some of them show high grain yields of 12–21 t/hm² in field experiments [4]. The core of super rice breeding is an effective use of germ plasma resources and favorable genes. For example, most of the northern japonica super rice varieties have DEP1 gene, which controlled large erect panicle [5, 6]. Although a large progress about the genetic basis for yield, grain quality and heading date of super rice had achieved [5–7], but the molecular mechanism is still poorly understood.

To meet the global food demands, achievements in molecular genetics of complex traits in rice will be important for increasing yield in the post-genomics era [8]. The draft genomic sequences of two rice subspecies, japonica (Nipponbare) and indica (93–11), were released [9, 10]; and later, the final genome sequence of Nipponbare was completed by International Rice Genome Sequencing Project [11]. These achievements provided a high quality reference genome for re-sequencing rice genomes using high-throughput sequencing technologies [12]. The widely used high-throughput sequencing technologies are also known as next-generation sequencing (NGS) technology. The NGS technology has been widely used in rice genomics and molecular breeding studies [13]. These characteristics have enabled researchers to perform accurate genetic polymorphism analysis and discovered important cloned genes of rice variety [14]. Several super rice or its parents including Liangyoupeii 9 [15], Shanyou 63 [16] and Shennong 265 [17] has been re-sequenced. However the Heilongjiang super rice varieties such as Longdao 5, Songgeng 9, Longgeng 31 and so on still not be researched. In this study, we re-sequenced a new bred elite japonica rice variety, Longdao24, with high quality, super high yield and early heading date in Heilongjiang province. Our results provided several candidates account for the super high yield, good quality and early heading date of rice in Heilongjiang province. The results therefore lay the groundwork for long-term efforts to uncover important genes and for future variety improvement.

Materials and methods

Plant materials

Longdao24 (Fig 1) was selected from the cross between two super japonica rice Longdao5 and Jigeng83. Longdao5 is a leading cultivar widely cultivated in the south and east part of Heilongjiang province and has high resistance to disease and super high yield. Jigeng83 is a leading cultivar widely cultivated in the middle part of Jilin province and has super high yield.

DNA isolation and genome sequencing

Genomic DNA was prepared from bulk harvested leaves of a single young plant of Longdao24, Longdao5 and Jigeng83 using a modified CTAB method [18]. Whole genome re-sequencing of the grape lines was performed on an Illumina HiSeq4000™ by Biomarker Technologies (Beijing, China). The procedure was performed as the standard Illumina protocol, including sample preparation and sequencing as follows: quantified DNA was extracted and DNA fragments were obtained using the ultrasonic wave technique, which was purified using the QIA quick PCR kit. End repair was performed with poly-A on the 3’ ends, then the adaptors were ligated, clusters were generated, agarose gel electrophoresis was used to select fragments, and PCR amplification was performed. Sequencing was performed by establishing a library with Illumina HiSeq4000™. The short reads were aligned using the Burrows–Wheeler transformation [19]. To ensure high quality, Low-quality reads (<20), reads with adaptor sequence and duplicated reads were filtered, and the remaining high-quality data were used in the mapping.
Mapping of reads to the reference

The clean sequencing reads were aligned to the temperate *japonica* Nipponbare reference genome—the unified-build release Os-Nipponbare-Reference-IRGSP-1.0 [20], using the BWA software under the default parameters with a small modification. The alignment results were then merged and indexed as BAM files [21]. Average sequencing depth and coverage were calculated using the alignment results. The mapped reads were then used to detect SNPs, InDels and SVs polymorphisms.

Detection of SNPs, InDels and SVs polymorphisms

Firstly, we used GATK tools software [22] to detect SNPs and InDels. Secondly, SVs were detected using BreakdancerMax.pl software [23] with its default parameters. In our result, the types of SVs include insertion (INS), deletion (DEL), inter chromosomal translocation (CTX), deletion including insertion (IDE), intra chromosomal translocation (ITX) and inversion (INV). And in our analysis, InDels were defined as the insertion or deletion the length of which was from 1 to 10 bp.

Annotation of SNPs, InDels and SVs

The annotation of SNPs and Indels were performed by SnpEff software [24]. The localization of SNPs, InDels and SVs were based on the annotation of reference genome databases. The polymorphisms in the gene region and other genome regions were annotated as genic and intergenic. The SNPs, InDels and SVs were classified according to their localization. SNPs in...
the CDS were further separated into synonymous and non-synonymous. The functional annotation data were achieved by blast each gene including non-synonymous SNPs or SVs to NR [25], SwissProt [25], GO [26], COG [27] and KEGG [28] database.

Results

Genome sequencing of Longdao24 and its parents

Genome sequencing of Longdao24, Longdao5 and Jigeng83 yielded 64.19, 53.93 and 70.60 million reads, most of them being 300 bp paired end reads. After cleaning, 19.17 G (Longdao24), 16.07 G (Longdao5) and 21.11 G (Jigeng83) clean bases were generated and 96.80%, 99.27% and 98.99% of the reads were mapped to the genomic reference sequence. The overall effective depth coverage of Longdao24, Longdao5 and Jigeng83 was 47×, 32× and 52×, respectively.

The 1×, 5× and 10× coverage ratio of Longdao24 to the reference genome was 97.87%, 96.11% and 95.08%, respectively. The 1×, 5× and 10× coverage ratio of Longdao5 to the reference genome was 97.21%, 95.44% and 93.19%, respectively. The 1×, 5× and 10× coverage ratio of Jigeng83 to the reference genome was 98.40%, 97.32% and 96.63%, respectively. The state information of Longdao24 and its parent were showed in Table 1.

Frequency distributions of different variation

We then examined genome-wide variations. 420,475 SNPs, 95,624 InDels, and 14,112 SVs were yielded from the Longdao24 genome. The frequency distributions of SNPs and InDels on Longdao24 genome were showed in Fig 2 and Fig 3. We detected ten abundant SNPs high region on chr 2(33–35 Mb), chr 4(6.5–12 Mb), chr 5(19-22Mb), chr 6(10-22Mb, 27-31Mb), chr 8(21.5–23 Mb), chr 10(2–6 Mb, 15-16Mb, 17–21 Mb) and chr12 (10–15 Mb). Moreover, the most abundant InDels region were identified on chr 2(33–35 Mb), chr 5(19-22Mb), chr 6(10-22Mb, 27-31Mb), chr 8(21.5–23 Mb) and chr 10(2–6 Mb, 15-16Mb, 17–21 Mb). Several chromosomal loci in the SNP high region were also found covered by the InDel high regions (Figs 2 and 3), which might suggest that those regions were the most polymorphic. 361,117 SNPs and 81,488 InDels were detected between the Longdao24 genome and Longdao5 genome. 428,908 SNPs and 97,209 InDels were identified between the Longdao24 genome and Jigeng83 genome. Frequency distributions of all different variation of Longdao24 genome was showed on Fig 4 via circos software.

Characteristics of SNPs, InDels and SVs

22,250, 65,858, 137,848 and 107,288 SNPs were found in intergenic region, intron region, upstream region and downstream region respectively. 9,054 SNPs were located in the 5’UTR

### Table 1. The re-sequencing state information of Longdao24 and its parent.

|                | Longdao24 | Longdao5 | Jigen83 |
|----------------|-----------|----------|---------|
| Raw_ Reads     | 64192132  | 53928143 | 70604131|
| Clean_ Reads   | 63958731  | 53582058 | 70340507|
| Clean_Base     | 19178366460| 16070000000| 21108274430|
| Mapped(%)      | 96.80     | 99.27    | 98.99   |
| Q30(%)         | 94.86     | 90.82    | 94.97   |
| Ave_depth      | 47        | 32       | 52      |
| Cov_ratio_1X(%)| 97.87     | 97.21    | 98.4    |
| Cov_ratio_5X(%)| 96.11     | 95.44    | 97.32   |
| Cov_ratio_10X(%)| 95.08    | 93.19    | 96.63   |

https://doi.org/10.1371/journal.pone.0181037.t001
and 3’UTR region. 1,958 SNPs were detected near the splice site. A total of 76,214 SNPs were located in the CDS regions, among which 31,308 were synonymous coding sequences (Syn CDS) and 42,115 were non-synonymous coding sequences (Nonsyn CDS) (Fig 5A).

5,623, 17,609, 34,717 and 27,132 InDels were found in intergenic region, intron region, upstream region and downstream region respectively. 4,465 InDels were located in the 5’UTR and 3’UTR region. 426 InDels were detected near the splice site. A total of 5,602 InDels were located in the CDS regions, among which 3,401 caused frame shift and 2,092 caused codon change (Fig 5B).

Fig 2. The frequencies of SNPs for Longdao24 at a 100 kb sliding window along each chromosome.

https://doi.org/10.1371/journal.pone.0181037.g002
A total of 14,112 Structure Variation (SVs) were detected in Longdao24 including 7,795 Insertion (INS), 3,273 Deletion (DEL), 217 Inversion (INV), 317 Intra-chromosomal Translocation (ITX) and 2,396 Inter-chromosomal Translocation (CTX). A total of 11,285 SVs were located in the CDS regions, among which 2,969 were detected in exon, 1,697 were located in intron and 6,619 were located in intergenic region. And 2,677 Copy number variation (CNVs) were also detected in Longdao24.

Fig 3. The frequencies of Indels for Longdao24 at a 100 kb sliding window along each chromosome.

https://doi.org/10.1371/journal.pone.0181037.g003
Variation annotation and gene categories

16,831 variations including 10,615 non-synonymous SNP mutations, 3,727 frame shifts and 2,489 SVs caused by indels, transversions, and transitions in 12,772 genes might influence the expression of the relevant protein. All tested and functional annotated genes were classified into gene ontology (GO) categories (Fig 6), which mainly divided the genes into three categories: cellular components (20 groups), molecular function (17 groups), and biological process (17 groups). Classification of gene variations compared with COG database by blast showed that most variations (1,935) were identified in replication, recombination and repair progress. Compared with the reference genome, most variations was screened in genetic information processing and metabolism.

Variation analysis on high yield genes

We investigated 22 cloned yield related genes [7] which might explain the high yield of Longdao24. A large number of SNPs and Indels (Table 2) were only detected in the CDS sequence of 4 genes related to high yield among Longdao24, Longdao5 and Jigeng83, such as Gn1a (related to spikelet number), EP3 (related to erect panicle and spikelet number), SCM2 (related to lodging resistance and spikelet number) and qGW8 (related to grain size and grain quality).
Four non-synonymous SNPs and a codon insertion InDel were identified in $Gn1a$ locus. Two non-synonymous SNPs and a codon change + deletion InDel were identified in $SCM2$ locus. Three non-synonymous SNPs and a codon deletion InDel were identified in $qGW8$ locus. One
frame shift InDel were identified in EP3 locus. The allele of Gn1a, DEP3 and SCM2 were come from the female parent Longdao5. The allele of qGW8 were come from the male parent Jigeng83.

**Variation analysis on starch quality genes**

We investigated 22 cloned grain quality related genes [29] in Longdao24 (Table 3). A total of fourteen SNPs and one Indels were detected in the CDS sequence of seven grain quality genes among Longdao24, Longdao5 and Jigeng83. Eleven SNPs and one Indels were detected in GBSSII (granule-bound starch synthase II), SSIIIb (Soluble starch synthase IIIb) and SSIIc (Soluble starch synthase Iic) locus between Longdao24 and Nipponbare. The GBSSII in Longdao24 had only one SNP. The SSIIIb and SSIIc of Longdao24 had six SNPs plus one Indel and four SNPs, respectively. The other three SNPs were detected in SSIVA and SBE3 locus among Longdao24, Longdao5 and Jigeng83. The allele of GBSSII locus was same among Longdao24, Longdao5 and Jigeng83, but different from Nipponbare. The other four allele of SSIVA, SBE3, SSIIIb and SSIIc were come from the male parent Jigeng83.
| Genes   | RAP NO. | MSU NO.       | Variation type | Variation region | chromosome | location | Nipponbare | Longdao24 | Longdao5 | Jigeng83 |
|----------|---------|---------------|----------------|------------------|------------|----------|------------|-----------|----------|----------|
| Gn1a     | Os01g0197700 | LOC_Os01g10110 | SNP: Non Synonymous | Exon Chr1 | 5270928 | A | C | C | A |
| Sd1      | Os01g083800 | LOC_Os01g66100 | SNP: Non Synonymous | Exon Chr1 | 5275176 | T | C | C | T |
| qGY2-1   | Os02g0154200 | LOC_Os02g05980 | SNP: Non Synonymous | Exon Chr1 | 5275210 | C | A | A | C |
| GW2      | Os02g0244100 | LOC_Os02g14720 | SNP: Non Synonymous | Exon Chr1 | 5275281 | A | ACGGCCG | ACGGCCG | A |
| EP3/LP   | Os02g0260200 | LOC_Os02g15950 | INDEL: Codon Insertion | Exon Chr2 | 5275362 | C | G | G | C |
| OsGRF4/ GS2/GL2 | Os02g0701300 | LOC_Os02g47280 | SNP: Non Synonymous | Exon Chr2 | 5275813 | ACGGCCG | ACGGCCG | ACGGCCG | A |
| GL3.1/qGL3-1/ qGL3/ OsPPKL | Os03g0646900 | LOC_Os03g44500 | SNP: Non Synonymous | Exon Chr2 | 5275813 | ACGGCCG | ACGGCCG | ACGGCCG | A |
| SCM3/ OsTB1/FC1 | Os03g0706500 | LOC_Os03g49880 | SNP: Non Synonymous | Exon Chr2 | 5275813 | ACGGCCG | ACGGCCG | ACGGCCG | A |
| GIF1/OsCIN2 | Os04g0413500 | LOC_Os04g33740 | SNP: Non Synonymous | Exon Chr2 | 5275813 | ACGGCCG | ACGGCCG | ACGGCCG | A |
| SPIKE/ LSHL4/ NAL1GPS | Os04g0615000 | LOC_Os04g52479 | SNP: Non Synonymous | Exon Chr2 | 5275813 | ACGGCCG | ACGGCCG | ACGGCCG | A |
| GS5      | Os05g0158500 | LOC_Os05g06660 | SNP: Non Synonymous | Exon Chr2 | 5275813 | ACGGCCG | ACGGCCG | ACGGCCG | A |
| qSW5     | Os05g0187500 | LOC_Os05g08520 | SNP: Non Synonymous | Exon Chr2 | 5275813 | ACGGCCG | ACGGCCG | ACGGCCG | A |
| HGW      | Os06g0160400 | LOC_Os06g06530 | SNP: Non Synonymous | Exon Chr2 | 5275813 | ACGGCCG | ACGGCCG | ACGGCCG | A |
| TGW6     | Os06g0623700 | LOC_Os06g41850 | SNP: Non Synonymous | Exon Chr2 | 5275813 | ACGGCCG | ACGGCCG | ACGGCCG | A |
| GW6a/OsHAT1 | Os06g0650300 | LOC_Os06g44100 | SNP: Non Synonymous | Exon Chr2 | 5275813 | ACGGCCG | ACGGCCG | ACGGCCG | A |
| SCM2/ APO1 | Os06g0665400 | LOC_Os06g45460 | SNP: Non Synonymous | Exon Chr6 | 27480424 | ACCGCGC | A | A | ACCGCGC |
| GL7/GW7  | Os07g0603300 | LOC_Os07g41200 | SNP: Non Synonymous | Exon Chr6 | 27480778 | G | C | C | G |

(Continued)
| Genes | RAP NO. | MSU NO. | Variation region | Variation type | Variation | Location | Nipponbare | Longdao24 | Jigeng83 |
|-------|---------|---------|------------------|----------------|-----------|----------|------------|-----------|---------|
| DEP1/EP1 | -      | LOC_Os07g16000 | 3'UTR | SNP: Non-Synonymous | G | 26501191 | - | - | - |
| - | Os07g0616000 | LOC_Os08g39890 | INDEL: critical polymorphism | - | - | - | - | - | - |
| - | Os08g0596000 | LOC_Os09g42410 | 5'UTR, 3'UTR, UTR | SNP: Non-Synonymous | - | - | - | - | - |
| - | Os09g0531600 | LOC_Os08g19440 | INDEL: critical polymorphism | - | - | - | - | - | - |
| - | Os09g0328999 | LOC_Os09g37780 | INDEL: critical polymorphism | - | - | - | - | - | - |
| TAWAWA1 | -      | LOC_Os09g441900 | Exon | SNP: Non-Synonymous | - | - | - | - | - |
| TAWAWA1 | -      | LOC_Os10g37780 | Exon | SNP: Non-Synonymous | - | - | - | - | - |
| https://doi.org/10.1371/journal.pone.0181037.t002 | -      | LOC_Os10g37780 | Exon | SNP: Non-Synonymous | - | - | - | - | - |
Table 3. Variations of 22 grain quality related genes in Longdao 24.

| Genes  | RAP NO. | MSU NO.               | Variation type       | Variation region | Variation chromosome | Variation location | Nipponbare | Longdao24 | Longdao5 | Jigeng83 |
|--------|---------|-----------------------|----------------------|------------------|---------------------|-------------------|------------|-----------|----------|----------|
| AGPL2  | Os01g0633100 | LOC_Os01g44220       | -                    | -                | -                   | -                 | -          | -         | -        | -        |
| SSIVa  | Os01g0720600 | LOC_Os01g52250       | SNP: Non Synonymous  | Exon             | Chr1                | 30038156          | A          | A         | G        | A        |
| SBE3   | Os02g0528200 | LOC_Os02g32660       | SNP: Non Synonymous  | Exon             | Chr2                | 19360942          | G          | G         | A        | G        |
| SSIb   | Os02g0744700 | LOC_Os02g51070       | SNP: Non Synonymous  | Exon             | Chr2                | 19361347          | T          | T         | C        | T        |
| AGPL1  | Os03g0735000 | LOC_Os03g52460       | SNP: Non Synonymous  | Exon             | Chr4                | 31752654          | G          | C         | G        | C        |
| SBE4   | Os04g0409200 | LOC_Os04g33460       | SNP: Non Synonymous  | Exon             | Chr4                | 31755480          | C          | T         | C        | T        |
| SSIb   | Os04g0624600 | LOC_Os04g53310       | SNP: Non Synonymous  | Exon             | Chr4                | 31757746          | G          | A         | G        | A        |
| ISA2   | Os05g0393700 | LOC_Os05g32710       | SNP: Non Synonymous  | Exon             | Chr4                | 31758185          | T          | G         | G        | G        |
| SSIb   | Os05g0533600 | LOC_Os05g53270       | SNP: Non Synonymous  | Exon             | Chr4                | 31758814          | G          | GA        | G        | GA       |
| AGPL3  | Os05g0555000 | LOC_Os05g50380       | SNP: Non Synonymous  | Exon             | Chr4                | 31758920          | -          | -         | -        | -        |
| GBSSI/Wx| Os06g0133000 | LOC_Os06g04200      | SNP: Non Synonymous  | Exon             | Chr4                | 31760234          | -          | -         | -        | -        |
| SSI    | Os06g0160700 | LOC_Os06g05660       | SNP: Non Synonymous  | Exon             | Chr4                | 31761345          | -          | -         | -        | -        |
| SSIa/ALK| Os06g0229800 | LOC_Os06g12450      | SNP: Non Synonymous  | Exon             | Chr4                | 31762346          | -          | -         | -        | -        |
| SBE1   | Os06g0726400 | LOC_Os06g51084       | SNP: Non Synonymous  | Exon             | Chr4                | 31763347          | -          | -         | -        | -        |
| AGPL4  | Os07g0243200 | LOC_Os07g13980       | SNP: Non Synonymous  | Exon             | Chr7                | 12917341          | -          | -         | -        | -        |
| GBSSI   | Os07g0412100 | LOC_Os07g22930       | SNP: Non Synonymous  | Exon             | Chr7                | 12917914          | A          | G         | G        | G        |
| SSIa   | Os08g0191433 | LOC_Os08g92230       | SNP: Non Synonymous  | Exon             | Chr7                | 15673341          | C          | T         | C        | T        |
| AGPS2  | Os08g0345800 | LOC_Os08g25734       | SNP: Non Synonymous  | Exon             | Chr7                | 15675357          | G          | A         | G        | A        |
| ISA1   | Os08g0509000 | LOC_Os08g40930       | SNP: Non Synonymous  | Exon             | Chr7                | 15675473          | G          | T         | G        | T        |
| AGPS1  | Os09g0298200 | LOC_Os09g12660       | SNP: Non Synonymous  | Exon             | Chr7                | 15680316          | C          | A         | C        | A        |

https://doi.org/10.1371/journal.pone.0181037.t003
| Genes          | RAP NO. | MSU NO.   | Variation type       | Variation region | Variation chromosome | Variation location | Nipponbare | Longdao24 | Longdao5 | Jigeng83 |
|---------------|---------|-----------|----------------------|------------------|----------------------|-------------------|------------|-----------|----------|----------|
| OsGI          | Os01g0182600 | LOC_Os01g06700 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| OsEF3         | Os01g0566100 | LOC_Os01g38530 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| OsLFL1        | Os01g0713600 | LOC_Os01g51610 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| OsCCT01       | Os01g0835700 | LOC_Os01g61900 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| OsMADS65      | Os01g0922800 | LOC_Os01g69850 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| Se13          | Os01g0949400 | LOC_Os01g72090 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| OsCOL4        | Os02g0610500 | LOC_Os02g39710 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| DTH2          | Os02g0734000 | LOC_Os02g49230 | SNP: Non Synonymous  | Exon             | Chr2                 | 30098026         | T          | T         | G        | T        |
| PPS           | Os02g0771100 | LOC_Os02g53140 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| rhd1          | Os02g0776900 | LOC_Os02g53690 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| HDR1          | Os02g0793900 | LOC_Os02g55080 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| Ehd4          | Os03g0112700 | LOC_Os03g02160 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| DTH3          | Os03g0122600 | LOC_Os03g03070 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| Se14          | Os03g0151300 | LOC_Os03g05680 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| OsHAP5C       | Os03g0251350 | LOC_Os03g14669 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| OsPRR73       | Os03g0284100 | LOC_Os03g17570 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| Hd6           | Os03g0762000 | LOC_Os03g55389 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| Hd16          | Os03g0793500 | LOC_Os03g57940 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| SPIN1         | Os03g0815700 | LOC_Os03g60110 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| OsRR1         | Os04g0423000 | LOC_Os04g36070 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| HAF1          | Os04g0648800 | LOC_Os04g55510 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| OsLF          | Os05g0541400 | LOC_Os05g46370 | SNP: Non Synonymous  | Exon             | Chr5                 | 26880725         | C          | G         | G        | C        |
|               |          |           | INDEL: Frame Shift   | Exon             | Chr5                 | 26880755         | C          | CGG       | CGG      | C        |
| Hd17          | Os06g0142600 | LOC_Os06g05060 | SNP: Non Synonymous  | Exon             | Chr6                 | 2235191          | A          | G         | A        | G        |
| RFT1          | Os06g0157500 | LOC_Os06g06300 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| Hd3a          | Os06g0157700 | LOC_Os06g06320 | SNP: Non Synonymous  | Exon             | Chr6                 | 2942192          | C          | G         | G        | G        |
| Hd1           | Os06g0275000 | LOC_Os06g16370 | INDEL: Codon Insertion | Exon             | Chr6                 | 9336861          | C          | 37 bp     | 37 bp    | 37 bp    |
|               |          |           | SNIP: Non Synonymous | Exon             | Chr6                 | 9336938          | A          | G         | G        | G        |
|               |          |           | INDEL: Codon change  | Exon             | Chr6                 | 9337012          | A          | 70 bp     | 70 bp    | 70 bp    |
|               |          |           | Codon Insertion      | Exon             | Chr6                 | 9337012          | A          | 70 bp     | 70 bp    | 70 bp    |
| Se5           | Os06g0603000 | LOC_Os06g40080 | -                    | -                | -                    | -                | -          | -         | -        | -        |
| Ghd7          | Os07g0261200 | LOC_Os07g15770 | -                    | -                | -                    | -                | -          | -         | -        | -        |

(Continued)
### Table 4. (Continued)

| Genes       | RAP NO.       | MSU NO.         | Variation type      | Variation region | Variation chromosome | Variation location | Nipponbare | Longdao24 | Longdao5 | Jigeng83 |
|-------------|---------------|-----------------|---------------------|------------------|---------------------|-------------------|------------|-----------|----------|----------|
| *Hd2*/Ghd7.1| Os07g0695100  | LOC_Os07g49460  | SNP: Non Synonymous | Exon             | Chr7                | 29617569          | G          | C         | C        | C        |
|             |               |                 | SNP: Non Synonymous | Exon             | Chr7                | 29623803          | G          | A         | A        | A        |
|             |               |                 | SNP: Non Synonymous | Exon             | Chr7                | 29628500          | T          | C         | C        | C        |
| *Ehd3*      | Os08g0105000  | LOC_Os08g01420  | SNP: Non Synonymous | Exon             | Chr8                | 274460            | G          | C         | G        | C        |
|             |               |                 | SNP: Non Synonymous | Exon             | Chr8                | 276145            | G          | A         | G        | A        |
|             |               |                 | SNP: Non Synonymous | Exon             | Chr8                | 276669            | G          | A         | G        | A        |
| *DTH8*      | Os08g0174500  | LOC_Os08g07740  | SNP: Non Synonymous | Exon             | Chr10               | 20872414          | T          | G         | T        | G        |
| *OsK4*      | Os08g0484600  | LOC_Os08g37800  | SNP: Non Synonymous | Exon             | Chr10               | 20872469          | G          | C         | G        | C        |
| *OsCO3*     | Os09g0240200  | LOC_Os09g06464  | SNP: Non Synonymous | Exon             | Chr10               | 20872638          | A          | G         | A        | G        |
| *Ehd2*      | Os10g0419200  | LOC_Os10g28330  | SNP: Non Synonymous | Exon             | Chr10               | 20872638          | A          | G         | A        | G        |
| *Ehd1*      | Os10g0463400  | LOC_Os10g32600  | SNP: Non Synonymous | Exon             | Chr10               | 20872638          | A          | G         | A        | G        |
| *OsMADS56*  | Os10g0536100  | LOC_Os10g39130  | SNP: Non Synonymous | Exon             | Chr10               | 20872638          | A          | G         | A        | G        |

https://doi.org/10.1371/journal.pone.0181037.t004
Variation analysis on heading date genes

Eight (DTH2, OsLF, Hd17, Hd3a, Hd1, Hd2, Ehd3 and OsMADS56) from 39 heading date genes had 24 variations (19 SNPs and 5 Indels) among Longdao24, Longdao5 and Jigeng83 (Table 4). The only one non-synonymous SNP in DTH2 allele from Longdao24 was same as Nipponbare and Jigeng83, but different from Longdao5. Hd1 allele in Longdao24 had 6 non-synonymous SNPs, one frame shift Indel and three long base pare insertion (36 bp, 69 bp and 51 bp). Hd2, Ehd3 and OsMADS56 in Longdao24 all be identified with 3 non-synonymous SNPs. OsLF in Longdao24 all had one non-synonymous SNP and one Indel. Only one non-synonymous SNP was detected in Hd3a and Hd17. The SNPs and Indels in Hd3a, Hd2 and Hd1 locus were detected between Nipponbare with the other three varieties. The allele of OsLF were come from the female parent Longdao5. The other fourth allele of DTH2, Hd17, Ehd3 and OsMADS56 were come from the male parent Jigeng83.

Development DNA markers for providing evidence of polymorphisms

Although the accuracy of deep sequencing has now been greatly improved, it still needs to be confirmed by other methods for the variations of interest. Therefore, we selected one InDel and one SNP polymorphisms to developed DNA markers (Table 5). We designed the InDel marker according the 9 bp deletion in SCM2 locus. A complete matching genotypes was observed in Longdao24 and its parents. Longdao24 and Longdao5 was found having this 9-bp

Table 5. The primers used for detection the polymorphisms in SCM2 and Hd17.

| Gene  | Marker type | Forward Prime | Reverse Primer | Enzyme |
|-------|-------------|---------------|----------------|--------|
| SCM2  | InDel       | CGACGGCGATGATGGAGG | CGGCCCTGCTCGGTTCCCT |        |
| Hd17  | CAPS        | CTACTTTTCGCCACCTGC   | AAACAAATCCGATATTATAGC | Hpy166II |

Development DNA markers for providing evidence of polymorphisms.

Fig 7. Development DNA markers for providing evidence of polymorphisms.

https://doi.org/10.1371/journal.pone.0181037.g007
deletion compared to Jigeng83 (Fig 7A). We also developed a CAPS marker to provide the SNP (2235191) in Hd17 locus. A complete matching genotypes was also observed in Longdao24 and its parents (Fig 7B).

Discussion

Breeding value of the re-sequencing data of super high yield rice variety

Marker-assisted selection (MAS) is expected to provide higher efficiency, reduced cost and shorter duration of the breeding scheme, compared with conventional methods. But most important complex traits in rice, including yield, quality and stress tolerance, are controlled by quantitative trait loci (QTLs). Isolating and characterizing genes involved in QTLs has been the key step for MAS. Whole genome re-sequencing technology has provided the possibility of genome wide variation analysis in rice breeding varieties. Longdao24 had been used widely as breeding parent, because of its high yield, good quality, early heading date and strong stress tolerance. Genome sequencing of Longdao24 with 47 × coverage effective depth and 95.08% 10 × coverage ratio was achieved. A total of 420,475 SNPs, 95,624 InDels, and 14,112 SVs were yielded from the Longdao24 genome. We identified 361,117 SNPs and 81,488 InDels between Longdao24 genome and Longdao5 genome. We also detected 428,908 SNPs and 97,209 InDels between Longdao24 genome and Jigeng83 genome. These different variations will be used for developing new molecular markers for genetics research and MAS.

The genetic basic of high yield in Longdao24

Developments of whole genome re-sequencing technology have provided new tools for discovering and tagging novel useful alleles for improving target traits and for manipulating those genes in rice breeding program through MAS. We identified four yield related genes (Gn1a, EP3, SCM2 and qGW8) having genetic variations in Longdao24. The first one is Gn1a, which controlled cytokinin accumulation in inflorescence meristems and increases the number of reproductive organs, resulting in enhanced grain yield [30]. Comparison of the DNA sequences between the cultivars revealed several nucleotide changes, including a 6-bp insertion in the first exon, and four nucleotide changes resulting in amino acid variation in the first and fourth exons of the Longdao24 and Longdao5 allele (Fig 8A). The Gn1a allele from Longdao24 and Longdao5 was similar to Habataki, which has been reported having the increasing Gn1a allele. The second one is EP3, which controlled panicle architecture and enhance the grain yield in rice [31]. Comparison of the DNA sequences between the cultivars revealed a 2-bp deletion in the first exon of the Longdao24 and Longdao5 allele (Fig 8B). The EP3 allele from Longdao24 and Longdao5 was a novel allele. The third one is SCM2, which showed enhanced culm strength and increased spikelet number [32]. Comparison of the DNA sequences between the cultivars revealed several nucleotide changes, including a 9-bp deletion in the second exon, and two nucleotide changes resulting in amino acid variation in the first exons of the Longdao24 and Longdao5 allele (Fig 8C). The SCM2 allele from Longdao24 and Longdao5 was same as Habataki, which has been reported having the increasing SCM2 allele. The last one is qGW8, which showed enhanced culm strength and increased spikelet number [33]. Comparison of the DNA sequences between the cultivars revealed several nucleotide changes, including a 10-bp insertion in the first exon, a 3-bp deletion in the first exon, and three nucleotide changes resulting in amino acid variation in the first and third exons of the Longdao24 and Jigeng83 allele (Fig 8D). The qGW8 allele from Longdao24 and Longdao5 was same as HJX74, which has been reported having the increasing yield haplotype. So, the super high yield formed by the large half erect panicle and lodging resistance of Longdao24 would be caused by these four increasing alleles of high yield genes including Gn1a, EP3, SCM2 and qGW8.
Fig 8. The yield related genes (Gn1a, EP3, SCM2 and qGW8) having genetic variations in Longdao24.

https://doi.org/10.1371/journal.pone.0181037.g008
A. **Hd3a**

![Diagram](image)

**MSU Os90 Rice Loci**

- LOC:0005064320

*dsFL2:* PTI-like or flowering locus T gene; contains Pfam profile PF01651: Phosphatidylinositol

**MSU Os90 Rice Gene Models**

- LOC:0005064320.1

|   |   |   |
|---|---|---|
| Nipponbare | C |   |
| Longdao24 | G |   |
| Jigeng68 | G |   |

B. **Hd1**

![Diagram](image)

**MSU Os90 Rice Loci**

- LOC:000645320

*GC/T-box zinc finger protein, putative, expressed*

**MSU Os90 Rice Gene Models**

- LOC:000645320.1

|   |   |   |   |
|---|---|---|---|
| Nipponbare | C | A | G | C |
| Longdao24 | 37bp | G | A | 70bp | A |
| Longdao5 | 37bp | G | A | 52bp | A |
| Jigeng68 | 37bp | G | A | 70bp | A |

C. **Hd2/DTH7/Ghd7.1**

![Diagram](image)

**MSU Os90 Rice Loci**

- LOC:000745660

*Response regulator receiver domain containing protein, expressed*

**MSU Os90 Rice Gene Models**

- LOC:000745660.1

|   |   |   |   |
|---|---|---|---|
| Nipponbare | A | G | T |
| Longdao24 | G | A | C |
| Longdao5 | G | A | C |
| Jigeng68 | G | A | C |

D. **DTH2**

![Diagram](image)

**MSU Os90 Rice Loci**

- LOC:000924920

*GT/C-box zinc finger protein, putative, expressed*

**MSU Os90 Rice Gene Models**

- LOC:000924920.1

|   |   |   |   |
|---|---|---|---|
| Nipponbare | T |   |   |
| Longdao24 | T |   |   |
| Longdao5 | G |   |   |
| Jigeng68 | T |   |   |

Fig 9. The heading related genes (*Hd3a*, *Hd1*, *Hd2* and *DTH2*) having genetic variations in Longdao24.

[https://doi.org/10.1371/journal.pone.0181037.g009](https://doi.org/10.1371/journal.pone.0181037.g009)
The genetic basic of early heading in Longdao24

Heading date is a major determinant of rice adaptability to regional and environmental conditions. It has been a major selecting target in rice breeding programs. To elucidate the key heading date gene in Longdao24 for adaptation to the northern limit of rice cultivation Heilongjiang Province (43°26'–53°33'), we analyzed 39 cloned heading date related genes. \textit{DTH2}, \textit{OsLF}, \textit{Hd17}, \textit{Hd3a}, \textit{Hd1}, \textit{Hd2}, \textit{Ehd3} and \textit{OsMADS56} had variations among Nipponbare, Longdao24, Longdao5 and Jigeng83. The alleles of \textit{Hd3a}, \textit{Hd1} and \textit{Hd2} in Longdao24, Longdao5 and Jigeng83 were different from that in Nipponbare. The might be the major reason for these three varieties adaptation to the northern limit of rice cultivation. \textit{Hd1} and \textit{Hd3a} are the two key genes controlled heading date in rice \cite{34}. Comparison of the DNA sequences of \textit{Hd3a} between the cultivars revealed only one nucleotide changes resulting in amino acid variation in the fourth exon of Longdao24, Longdao5 and Jigeng83 allele (Fig 9A). This \textit{Hd3a} allele was a novel allele, which was not be reported before. The only one SNP caused an amino acid change at the carboxyl end of the predicted protein: the Asn (AAG) in Nipponbare was changed to Lys (AAC). The \textit{Hd1} allele in Longdao24, Longdao5 and Jigeng83 was also a novel allele. Four InDels and six SNPs was identified by comparison of the DNA sequences of \textit{Hd1} between the cultivars. Three insertion and four nucleotide changes resulting in amino acid variation were identified in the first exon; and a 2-bp insertion and two nucleotide changes resulting in amino acid variation were identified in the second exon (Fig 9B). The \textit{Hd2} allele in Longdao24, Longdao5 and Jigeng83 all had the three nucleotide changes resulting in amino acid variation in the first, fourth and eighth exons (Fig 9C). These variations were found in many European and Asian rice cultivars, which flower extremely early under natural long-day conditions \cite{35}. The \textit{DTH2} allele in Longdao24 was same as Jigeng83 (Fig 9D). This SNP in the third exon was reported as a key functional differences in the rice landraces from northern limit of rice cultivation in the world \cite{36}. We also identified the allele of \textit{Hd17}, \textit{Ehd3} and \textit{OsMADS56} in Longdao24 was came from Jigeng83 (Table 4). But only allele of \textit{OsLF} was detected from Longdao5 (Table 4). All these information showed that the \textit{Hd3a}, \textit{Hd1} and \textit{Hd2} might be the major reason for the varieties adaptation to the northern rice cultivation region. The \textit{DTH2}, \textit{OsLF}, \textit{Hd17}, \textit{Ehd3} and \textit{OsMADS56} might formed different varieties adaptation to different region in northern rice cultivation location.

The genetic basic of grain quality in Longdao24

Starch is the major component of rice grain, which is mainly composed of amylose and amylopectin. The starch synthesis pathway is an ideal system for examining the evolution of biochemical pathways. Over 20 genes involved in the rice starch synthesis pathway have been identified so far \cite{37}. We investigated 22 cloned starch synthesis related genes in Longdao24 (Table 3). We only identified variations in \textit{SSIVa}, \textit{SBE3}, \textit{SSIIIb} and \textit{SSIIc} between the parents of Longdao24. All four allele of \textit{SSIVa}, \textit{SBE3}, \textit{SSIIIb} and \textit{SSIIc} were come from the male parent Jigeng83. These genes represented starch synthases (\textit{SSIVa}, \textit{SSIIIb} and \textit{SSIIc}) and branching enzymes (\textit{SBE3}). These enzymes are involved in amylopectin biosynthesis in rice endosperm. They had low to medium effects on variation in starch trait \cite{38}. So we need do more works to get enough data, for example polymorphisms or expression levels of these genes, to explore the quality genetic basis of Longdao24.

Author Contributions

\textbf{Conceptualization}: Liangming Bai.

\textbf{Data curation}: Shichen Sun, Guohua Ding, Hui Jiang, Xijuan Zhang.

\textbf{Formal analysis}: Shichen Sun, Tianshu Xia, Hui Jiang, Xijuan Zhang, Fengming Zhang.
Investigation: Tongtong Wang.

Methodology: Liangming Bai.

Project administration: Xijuan Zhang.

Resources: Guohua Ding.

Validation: Tongtong Wang.

Visualization: Tianshu Xia, Xijuan Zhang, Fengming Zhang.

Writing – original draft: Shukun Jiang.

Writing – review & editing: Shukun Jiang.

References

1. Chen WF, Xu ZJ, Zhang LB. Theories and Methods of Breeding Japonica Rice for Super High Yield. Journal of Shenyang Agricultural University. 2005; 1: 3–8. (in Chinese)
2. Yuan LP. Development of Super Hybrid Rice for Food Security in China. Engineering. 2015; 1: 13–14.
3. IRRI. IRRI redesigns rice plant to yield more grain. IRRI reporter. 1994; 2: 4–1.
4. Cao LY, Zhan XD, Chen SG, Feng Y, Wu WM, Shen XH, et al. Breeding Methodology and Practice of Super Rice in China. Rice Science. 2010; 2: 87–93.
5. Huang X, Qian Q, Liu Z, Sun H, He S, Luo D. Natural variation at the DEP1 locus enhances grain yield in rice. Nature Genetics. 2009; 4:494–497.
6. Wang J, Nakazaki T, Chen S, Chen W, Saito H, Tsukiyama T. Identification and characterization of the erect-pose panicle gene EP conferring high grain yield in rice (Oryza sativa L.). Theoretical and Applied Genetics. 2009; 119: 85–91. https://doi.org/10.1007/s00122-009-1019-0 PMID: 19407986
7. Rao YC, Li YY, Qian Q. Recent progress on molecular breeding of rice in China. Plant Cell Rep. 2014; 33: 551–564. https://doi.org/10.1007/s00299-013-1551-x PMID: 2442397
8. Miura K, Ashikari M, Matsuoka M. The role of QTLs in breeding of high-yielding rice. Trends Plant Sci. 2011; 6: 319–326.
9. Yu J, Hu SN, Wang J, Wong GK, Li SG, Liu B. A draft sequence of the rice genome (Oryza sativa L. ssp. indica). Science. 2002; 296: 79–92. https://doi.org/10.1126/science.1068037 PMID: 11935017
10. Goff SA, Ricke D, Lan TH, Presting G, Wang R L, Dunn M. A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science. 2002; 296: 92–100. https://doi.org/10.1126/science.1068275 PMID: 11935018
11. International Rice Genome Sequencing Project. The map-based sequence of the rice genome, Nature. 2005; 436: 793–800. https://doi.org/10.1038/nature03895 PMID: 16100779
12. Gao Q, Yue GD, Li WQ, Wang JY, Xu JH, Yin Y. Recent progress using high-throughput sequencing technologies in plant molecular breeding. J Integr Plant Biol. 2012; 4: 215–227.
13. Guo LB, Gao ZY, Qian Q. Application of resequencing to rice genomics, functional genomics and evolutionary analysis. Rice. 2014; 7: 4. https://doi.org/10.1186/s12284-014-0004-7 PMID: 2506357
14. Han B, Huang XH. Sequencing-based genome-wide association study in rice. Curr Opin Plant Biol. 2013; 16: 133–136. https://doi.org/10.1016/j.pbi.2013.03.006 PMID: 23562592
15. Gao ZY, Zhao SC, He WM, Guo LB, Peng YL, Wang JJ. Dissecting yield-associated loci in super hybrid rice by resequencing recombinant inbred lines and improving parental genome sequences. Proc Natl Acad Sci U S A. 2013; 35: 14492–14497.
16. Yu HH, Xie WB, Wang J, Xing YZ, Xu CG, Li XH. Grains in QTL detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. PLoS One. 2011; 3: e17595.
17. Li JY, Wang J, Zeigler RS. The 3,000 rice genomes project: new opportunities and challenges for future rice research. Giga Science. 2014; 3: 8. https://doi.org/10.1186/2047-217X-3-8 PMID: 24872878
18. Jiang SK, Zhong M, Xu ZJ. Classification of rice cultivars with RAPD molecular markers. J Shenyang Agric Univ. 2006; 37: 639–641. (in Chinese)
19. Li H., Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. 2010; 26: 589–595. https://doi.org/10.1093/bioinformatics/btp988 PMID: 2008505
20. Kawahara Y, De la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice*. 2013; 6:4. https://doi.org/10.1186/1939-8433-6-4 PMID: 24280374

21. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N. 1000 Genome Project Data Processing Subgroup: The Sequence Alignment/Map format and SAM tools. *Bioinformatics*. 2009; 25: 2078–2079. https://doi.org/10.1093/bioinformatics/btp352 PMID: 19505943

22. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010; 20:1297–303. https://doi.org/10.1101/gr.107524.110 PMID: 20644199

23. Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods*. 2009; 9: 677–681.

24. Cingolani P, Platts A, Wangle L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012; 2: 80–92.

25. Deng YY, Li JQ, Wu SF, Zhu YP, Chen YW, He FC. Integrated nr Database in Protein Annotation System and Its Localization. *Computer Engineering*. 2006; 5: 71–74.

26. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. *The Gene Ontology Consortium*. *Nat Genet*. 2000; 1: 25–29.

27. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res*. 2010; 1: 33–36.

28. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucleic Acids Res*. 2004; 32: D277–D280. https://doi.org/10.1093/nar/gkh063 PMID: 14681412

29. Tian ZX, Qian Q, Liu QQ, Yan MX, Liu XF, Yan CJ. Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. *Proc Natl Acad Sci U S A*. 2009; 106:21750–21755.

30. Takahashi Y, Teshima KM, Yokoi S, Innan H, Shimamoto K. 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*. 2009; 106(11): 4555–4560. https://doi.org/10.1073/pnas.0812092106 PMID: 19246394

31. Koo BH, Yoo SC, Park JW, Kwon CT, Lee BD, An G. Natural variation in *OsPRR37* regulates heading date and contributes to rice cultivation at a wide range of latitudes. *Mol Plant*. 2013; 6(6): 1877–1888. https://doi.org/10.1093/mp/ssq243 PMID: 23713079

32. Yu G, Olsen KM, Schaal BA. Molecular evolution of the endosperm starch synthesis pathway genes in rice (*Oryza sativa* L.) and its wild ancestor, *O. rufipogon* L. *Mol Biol Evol*. 2011; 28(1): 659–671. https://doi.org/10.1093/molbev/msq243 PMID: 20829346

33. Kharabian-Masouleh A, Waters DL, Reinke RF, Ward R, Henry RJ. SNP in starch biosynthesis genes associated with nutritional and functional properties of rice. *Nature Sci Rep*. 2012; 2: 557.