

Article

Development of an *Aus*-Derived Nested Association Mapping (*Aus*-NAM) Population in Rice

Justine K. Kitony^{1,†} , Hidehiko Sunohara^{1,2,†}, Mikako Tasaki¹, Jun-Ichi Mori¹, Akihisa Shimazu¹, Vincent P. Reyes¹ , Hideshi Yasui³, Yoshiyuki Yamagata³ , Atsushi Yoshimura³, Masanori Yamasaki⁴ , Shunsaku Nishiuchi¹ and Kazuyuki Doi^{1,*} 

¹ Graduate School of Bioagricultural Sciences, Nagoya University, Chikusa-ku, Furo-cho, Nagoya 464-8601, Japan; kitony.kipruto.justine@g0.nagoya-u.jp (J.K.K.); taichung65@yahoo.co.jp (H.S.); tasakimikako21@gmail.com (M.T.); mjun.turuturu@gmail.com (J.-I.M.); pc0wc3874w2vj052x@yahoo.co.jp (A.S.); reyes.vincent.pamugas@f.mbox.nagoya-u.ac.jp (V.P.R.); s_nishi@agr.nagoya-u.ac.jp (S.N.)

² Environmental Control Center Co., Ltd., Hachioji 193-0832, Japan

³ Faculty of Agriculture, Kyushu University, 744 Motooka, Fukuoka 819-0395, Japan; hyasui@agr.kyushu-u.ac.jp (H.Y.); yoshiyuk@agr.kyushu-u.ac.jp (Y.Y.); ayoshi@agr.kyushu-u.ac.jp (A.Y.)

⁴ Food Resources Education and Research Center, Graduate School of Agricultural Science, Kobe University, Kasai 675-2103, Japan; yamasakim@tiger.kobe-u.ac.jp

* Correspondence: kdoi@agr.nagoya-u.ac.jp

† These two authors contributed equally to this work.



Citation: Kitony, J.K.; Sunohara, H.; Tasaki, M.; Mori, J.-I.; Shimazu, A.; Reyes, V.P.; Yasui, H.; Yamagata, Y.; Yoshimura, A.; Yamasaki, M.; et al. Development of an *Aus*-Derived Nested Association Mapping (*Aus*-NAM) Population in Rice. *Plants* **2021**, *10*, 1255. <https://doi.org/10.3390/plants10061255>

Academic Editors: Yong-Gu Cho and Nidhi Rawat

Received: 15 May 2021
Accepted: 15 June 2021
Published: 21 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: A genetic resource for studying genetic architecture of agronomic traits and environmental adaptation is essential for crop improvements. Here, we report the development of a rice nested association mapping population (*aus*-NAM) using 7 *aus* varieties as diversity donors and T65 as the common parent. *Aus*-NAM showed broad phenotypic variations. To test whether *aus*-NAM was useful for quantitative trait loci (QTL) mapping, known flowering genes (*Ehd1*, *Hd1*, and *Ghd7*) in rice were characterized using single-family QTL mapping, joint QTL mapping, and the methods based on genome-wide association study (GWAS). *Ehd1* was detected in all the seven families and all the methods. On the other hand, *Hd1* and *Ghd7* were detected in some families, and joint QTL mapping and GWAS-based methods resulted in weaker and uncertain peaks. Overall, the high allelic variations in *aus*-NAM provide a valuable genetic resource for the rice community.

Keywords: *aus*-NAM; GBS; QTL; *Ehd1*; *Hd1*; *Ghd7*; joint linkage mapping; GWAS

1. Introduction

Improvement of rice (*Oryza sativa* L.) production has been achieved by development of new varieties and optimization of cultural practices. The majority of phenotypic variations of agricultural traits is determined by many loci with small effects (quantitative trait loci, QTL). Understanding the genetic basis of quantitative traits is essential for crop improvement. Application of DNA markers has become a common method in breeding programs. In addition, the recent advances in DNA marker technologies have enabled a systematic approach to identification of QTL and marker-assisted selection [1]. In rice, genome sequencing has enhanced identification of causal genes related to yield [2]. Most of these genes/QTL were identified using bi-parental populations, combined with the development of backcrossed progeny [3]. The backcross progeny enables precise estimation of allelic effects and fine-mapping of the target loci. However, limited allele richness is a disadvantage in these types of mapping populations.

On the other hand, development of next-generation sequencing (NGS) technology enabled direct detection of genetic loci associated with traits through genome-wide association studies (GWAS) [4]. In plants, GWAS can utilize a variety of germplasm collections such as landraces, breeding lines, and varieties that have accumulated recombination

events both recent and historic. This attribute gives GWAS a higher gene mapping resolution compared with bi-parental populations [5–7]. However, unaccounted population structures/stratifications make GWAS prone to false associations. Additionally, the absence of pedigree information in diversity panels prohibits classical pedigree-based haplotype mapping, consequently reducing statistical power [8–10].

To combine the advantages of bi-parental populations and diversity panels, multi-cross mating designs consisting of diverse donors have been proposed. Examples of multi-cross designs include nested association mapping (NAM) [11], multi-parent advanced generation inter-crosses (MAGIC) [12], and random-open-parent association mapping (ROAM) [13]. Multi-parent populations have advantages, such as (1) allele richness coming from the diversity donors, (2) known population structures, and (3) flexibility to be used as a breeding platform.

The first NAM population was reported in maize [14]. The population was successfully used to profile the genetic architecture of agronomic traits, such as flowering time, leaf architecture, stalk strength, and plant height [15–17]. Following the successes in maize, NAM populations were developed in other crops, such as rice [18], wheat [19], barley [20], soybean [21], sorghum [22], and rapeseed [23].

Asian rice (*O. sativa*) is classified into five major varietal groups, namely, *temperate japonica*, *tropical japonica*, *indica*, *aus*, and *aromatic* [24]. *Aus* rice varieties are considered to have evolved from annual *Oryza nivara* found in Bangladesh, Northeast India, Nepal, and northern Myanmar. Most *aus* varieties exhibit photoperiod insensitivity, a source of local environmental adaptation. Moreover, *aus* is known to possess drought tolerance and rice blast resistance [7,25–27].

In this study, we describe the development of a rice NAM population using *aus* varieties as diversity donors (*aus*-NAM) and test its QTL mapping powers using known flowering-time genes. The set of recombinant inbred lines (RILs) utilized was generated using single-seed descent. Genotypes of each line were determined using NGS. For GWAS, parental DNA variants were projected onto the progeny genotypes, and mapping of days to heading (DTH) QTL/genes was performed.

2. Results

2.1. Development of *Aus*-NAM Population

Out of the seven families used in this study, five (WNAM02, WNAM29, WNAM31, WNAM35, and WNAM39) were newly developed. The numbers of plants in F₂ and F₅ are listed in Table 1. Because of hybrid sterility and late heading, a substantial number of the plants in F₂, F₃, and F₄ could not be harvested. The residual rate for families at F₅ ranged from 46.8% to 74.4% (Table 1). In total, 895 RILs ranging from 107 to 163 per family were obtained (Table 1).

Table 1. List of recombinant inbred lines in the *aus* nested association mapping population.

Family Name	Founder's Name	WRC No. ¹	F ₂	F ₅	Residual Rate
WNAM02	Kasalath	WRC02	233	109	46.78%
WNAM29	Kalo Dhan	WRC29	219	163	74.43%
WNAM31	Shoni	WRC31	174	121	69.54%
WNAM35	ARC5955	WRC35	229	137	59.83%
WNAM39	Badari Dhan	WRC39	213	126	59.15%
WNAM72	DV85	-	-	107	-
WNAM73	ARC10313	-	-	132	-

¹ World Rice Core Collection number [28].

2.2. Phenotypic Characteristics

The mean values of DTH in each family varied from 88 days to 105 days with potential transgressive segregation in some families, and some of the lines showed extreme values

compared to the parents (Figure S1). Analysis of variance showed a statistically significant difference among the RIL families with an F value of 73.52 and a p -value $< 2 \times 10^{-16}$.

2.3. Linkage Map and Projection of Parental Variants

The number of clean SNPs obtained from GBS in each family ranged from 2868 to 4285 (Table S1). Upon filtering lines with excess heterozygosity (>0.125), 887 RILs were retained for subsequent genetic analyses (Table S1). For QTL mapping, 1786 non-redundant SNP markers were used. They covered all the 12 rice chromosomes with an average distance between markers ranging from 0.41 cM to 0.86 cM (Table S2).

For GWAS, parental variants obtained from whole-genome resequencing (4,643,123 SNPs) were firstly filtered then projected onto each of the individual family skeleton linkage maps. A total of 41,561 SNPs were included in the association analysis.

2.4. Population Structure

Estimation of population structure using probabilistic principal component analysis (PPCA) showed fairly controlled stratification. The R^2 values were 0.067, 0.044, 0.041, 0.038, 0.037, and 0.034 for PC1, PC2, PC3, PC4, PC5, and PC6, respectively. (Figure 1, Figure S2, and Table S3). WNAM29 formed an isolated group from other families (Figure 1).

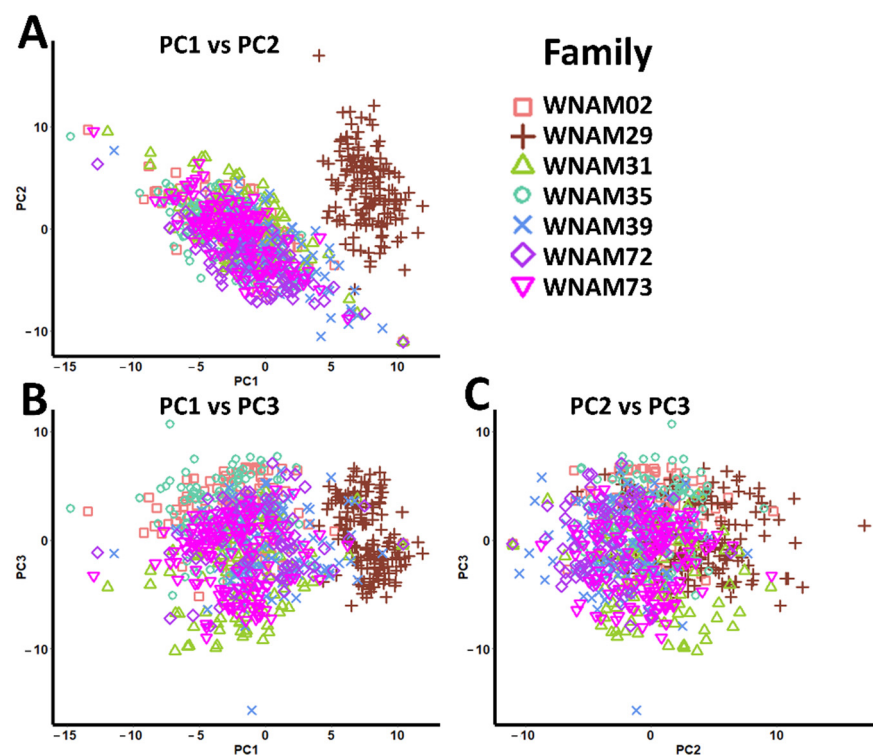


Figure 1. Probabilistic principal component analysis (PPCA) showing population structure in the *aus* nested association mapping (*aus*-NAM) population. (A) PC1 vs. PC2, (B) PC1 vs. PC3, and (C) PC2 vs. PC3. Different shapes and colors represent different *aus*-NAM families.

2.5. QTL Detected by Single-Family Analysis, Joint Analysis, and GWAS-Based Methods

Single-family QTL analysis detected a total of 14 significant additive QTL on chromosomes 5, 6, 7, and 10 (Figure 2 and Table S4). The QTL individually explained 12% to 36% of the trait variances. Among the QTL detected, eight contained alleles increasing days to heading from *aus* and six from T65, respectively. The QTL on chromosome 10 was detected across all families, with WNAM73 possessing the highest LOD score (11.81) (Table S4).

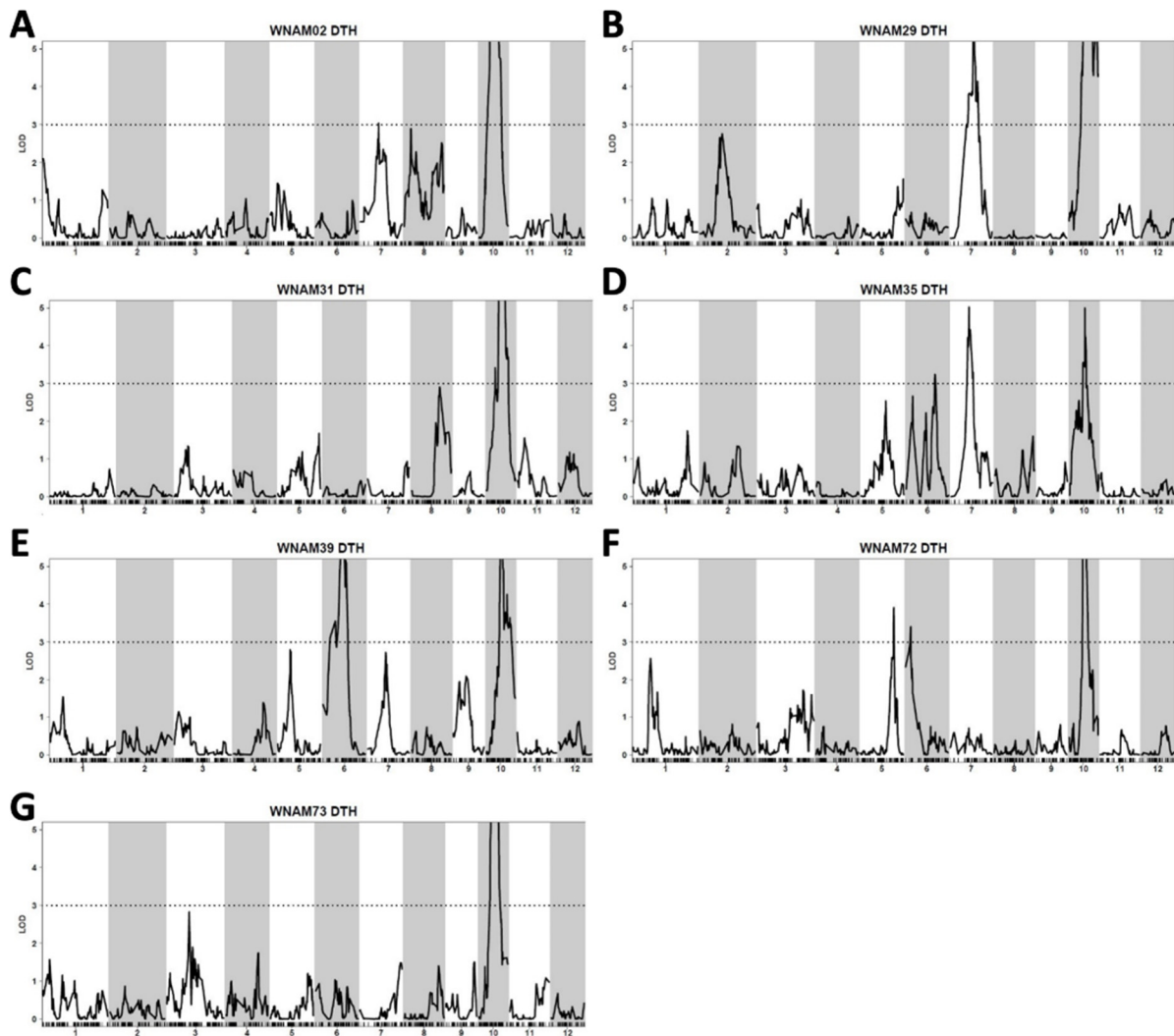


Figure 2. Interval mapping showing the locations of days to heading QTL in *aus* nested association mapping (*aus*-NAM) families using R/QTL scanone function. (A) WNAM02, (B) WNAM29, (C) WNAM31, (D) WNAM35, (E) WNAM39, (F) WNAM72, and (G) WNAM73. The black lines correspond to LOD score profile (Y-axis) as a function of distance in cM across each chromosome (X-axis). Horizontal dotted lines in all panels indicate the LOD threshold value of 3.

Joint linkage mapping identified a total of 19 QTL (Figure 3 and Table S5). Some of the joint QTL overlapped with single-family QTL, such as QTL on chromosomes 6, 7, and 10. The peak on chromosome 10 was detected as a single peak. The peaks on chromosomes 6 and 7 appeared to combine peak locations of the seven populations, resulting in two separate peaks on each chromosome. In addition, 14 putative QTL spanning a relatively wide region were detected on chromosomes 1, 2, and 3 (Figure 3). On the other hand, a significant QTL on chromosome 5 in WNAM72 (Figure 2F) was not detected in the joint QTL analysis.

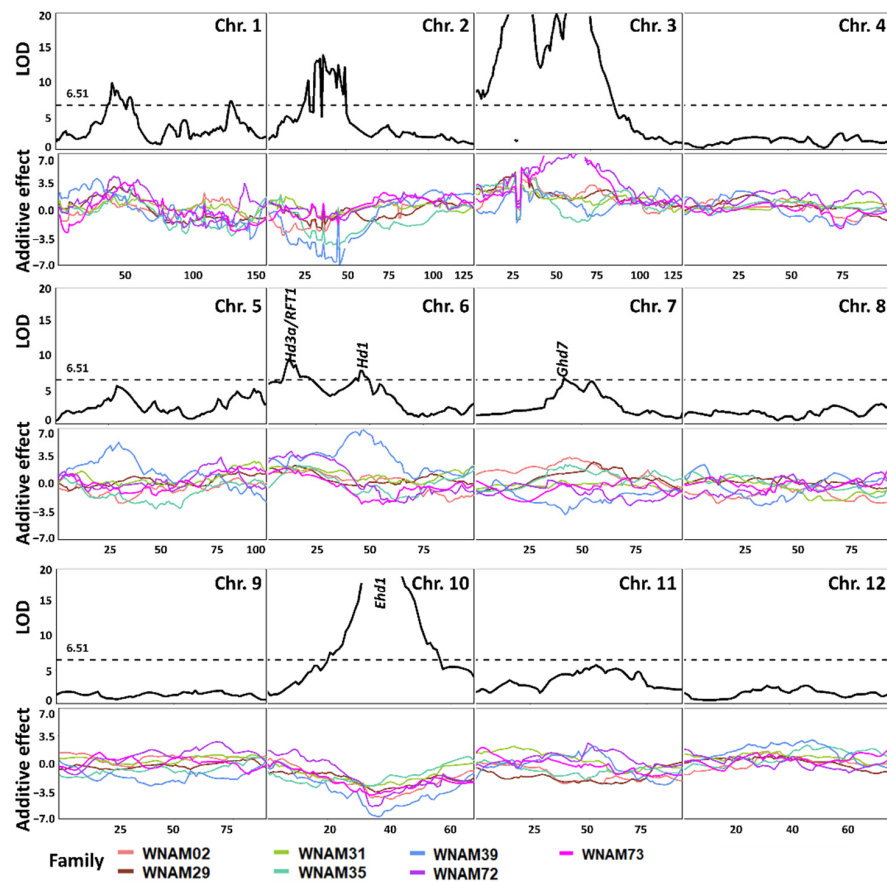


Figure 3. LOD profile of QTL (top of each panel) and its additive effects (bottom of each panel) detected by joint inclusive composite interval mapping (JICIM). Scanning step was 1 cM, and the dotted horizontal line (6.51) represents thresholds obtained from 1000 permutation tests with type I errors of 0.05. Different line colors in additive effect panels represent different families. The positions of known loci (*RFT*, *Hd3a*, *Hd1*, *Ehd1*) are shown in the panels of chromosomes 6, 7, and 10.

Association mapping by a naive general linear model revealed significant QTL signals in all chromosomes (Figure S3), while mixed linear model analysis identified significant SNPs on chromosomes: 6, 7, and 10 (Figure 4), and a total of 188 SNPs met the negative logarithm P value ($-\log_{10}P$) of 5.9 Bonferroni threshold at alpha 0.05.

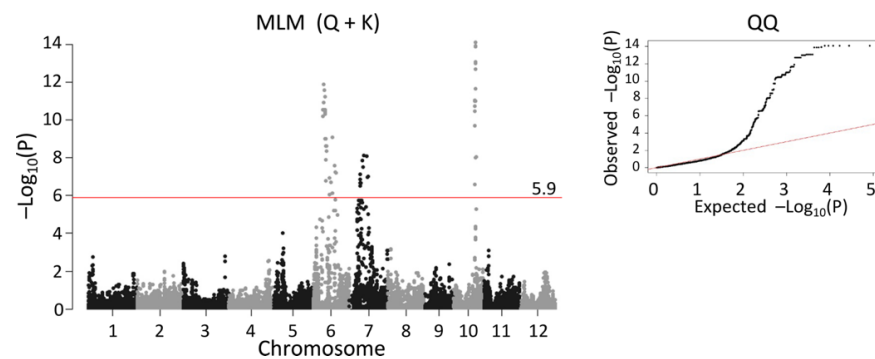


Figure 4. Manhattan plot of days to heading using mixed linear model with principal components (Q) and genomic kinship (K) as covariates (MLM (Q + K)). The red horizontal line marks the threshold for genome-wide significance (5.9) on a $-\log_{10}$ scale. A quantile–quantile (QQ) plot is shown in the right panel, where the observed p -values (Y-axis) against the expected p -values (X-axis) under the null hypothesis of no association are plotted on a $-\log_{10}$ scale. Each black dot indicates an SNP.

2.6. Evaluation of Mapping Accuracy

The QTL commonly detected in the three methods on chromosomes 6, 7, and 10 included the regions of *RFT1* [29], *Hd3a* [30], *Hd1* [31], *Ghd7* [32], and *Ehd1* [33]. Assuming that *Hd1*, *Ghd7*, and *Ehd1* were the genes underlying the detected QTL, these loci were used to evaluate the accuracy of gene mapping using *aus*-NAM.

A major QTL on chromosome 10 was identified in every individual family (Figure 2 and Table S4). This QTL corresponded to *Ehd1* (Os10g0463400) [33], which is located between 17,076,098 bp and 17,081,344 bp on chromosome 10. The peak was detected at the marker position from 16,626,134 bp to 17,367,103 bp in single-family QTL mapping (Figure 2), while joint QTL analysis identified marker position 16,772,764 bp as the peak (Figure 3). MLM (Q + K) detected a peak spanning from 17,095,439 bp to 17,164,368 bp (Figure 5C). Thus, all of the statistical methods successfully detected *Ehd1*.

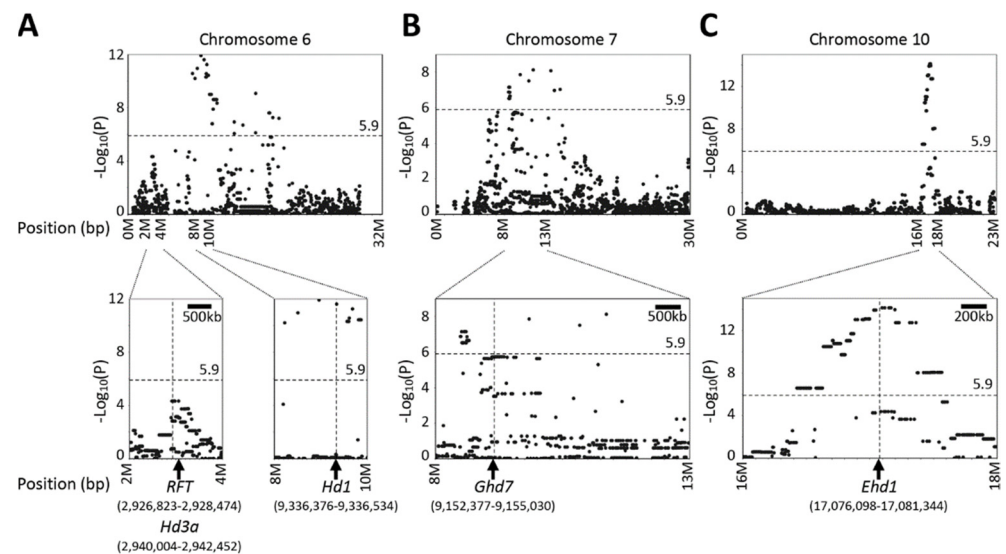


Figure 5. Manhattan scatter plots showing local association of days to heading in the *aus* nested association mapping (*aus*-NAM) population: (A) chromosome 6, (B) chromosome 7, and (C) chromosome 10. The panels at the bottom are magnification around *RFT1*/*Hd3a*, *Hd1*, *Ghd7*, and *Ehd1* loci.

A significant QTL was detected in WNAM39 (Badari Dhan), close to *Hd1* on chromosome 6 (9,335,377 bp to 9,337,570 bp) (Figure 2E). This peak was flanked by the markers S06_8837126 and S06_9318022 in joint linkage analysis with an LOD score of 7.84 (Figure 3). In MLM (Q + K), marker S06_9338330 was the most significant ($-\text{Log}_{10}P$ value of 11.6) (Figure 5A). To better understand the individual family contributions to the joint peak on chromosome 6 (Figure 3), additive effects at this locus were further analyzed. The Badari Dhan allele had the highest additive effect value (6.72 days) (Table S5). Multiple alignment of the amino acid sequences deduced using the genomic sequences confirmed loss-of-function in *aus* varieties except for Badari Dhan (Figure S4). Another signal in the vicinity of *RFT1* and/or *Hd3a* on chromosome 6 was also detected by GWAS (Figures 4 and 5), although the signal did not reach the significant threshold.

Ghd7 is located between 9,184,534 bp and 9,187,187 bp on chromosome 7 [32,34]. Single-family QTL analysis showed that WNAM02 and WNAM35 families had a significant QTL in the vicinity of *Ghd7* (Figure 2A,D). Therefore, it was hypothesized that Kasalath (WNAM02) and ARC5955 (WNAM35) possess functional (late) alleles for *Ghd7*. The peak corresponding to *Ghd7* in joint QTL was from 8.93Mbp to 9.35Mbp on chromosome 7, which contained the *Ghd7* locus (Figure 3). MLM (Q + K) detected a cluster of markers around the *Ghd7* locus with $-\text{Log}_{10}P$ values greater than 5, surrounded by further markers that showed higher $-\text{Log}_{10}P$ values than the threshold. Thus, the position of *Ghd7* was not clear from this analysis. In WNAM39, a QTL near *Ghd7* was detected where the *aus*

allele had a negative additive effect that was opposite from other families, and this peak was also detected in the joint QTL analysis.

3. Discussion

3.1. Development of *aus*-NAM Population

The NAM population brings the advantage of being able to apply both joint linkage analysis and GWAS in dissecting the genetic basis of complex traits [11,14,16]. In this study, we developed and characterized the first rice NAM population using the *aus* varietal group as diversity donors. Because of hybrid sterility and late heading, a substantial number of lines were lost in the process of single seed descent. However, the population size allowed detection of both known and novel QTL (Figure 2).

While examining the population structure of *aus*-NAM, PCCA showed only a weak stratification, with half-sib RILs dispersed around the T65 (common female parent) and *aus* (diversity donor parent) (Figure 1 and Figure S2). A similar population structure was reported in oilseed rape NAM [23] and sorghum NAM [22]. This confirmed that *aus*-NAM retained the genetic diversity of donors, and the population structure was suitable for analyzing the genetic architecture of complex traits. Further expansion of the seven *aus*-NAM RIL populations will fill the gaps and enhance the statistical power of *aus*-NAM.

3.2. Genotyping of *aus*-NAM Population

A GBS method [35,36] was used to obtain marker genotypes of *aus*-NAM. The potential number of markers obtained by the restriction enzyme used in this study (*KpnI*-*MspI*) was estimated at approximately 5000 loci. This number is sufficient to model phenotypes based on best linear unbiased prediction [37], hence, enough for gene mapping. Therefore, the choice of the enzyme was appropriate for *aus*-NAM. The advancement of next-generation sequencing technology will allow the use of more “frequent” cutters instead of *KpnI* or even whole-genome shotgun for all of the lines.

3.3. Accuracy of QTL Mapping Using *aus*-NAM Population

To date, over 40 flowering QTL have been identified in rice [34]. In this study, *Ehd1* [33] was detected as the most common major QTL. The mapped positions of the QTL on chromosome 10 corresponded to the actual position of *Ehd1* (Figure 2, Figure 3, Figure 4, and Figure 5C). Another heading time locus, *Hd1* [31], was detected only in WNAM39, and analysis of the deduced amino acid sequence confirmed that the founder of WNAM39 (Badari Dhan) was the only variety possessing the functional allele of *Hd1* (Figure S4). The effect of a functional allele of *Hd1* in the environment of this study (long day) was to delay heading, and it matched the observed result. However, the *Hd1* peak in Figure 5A was not surrounded by markers with smaller $-\text{Log}_{10}(P)$ values such as *Ehd1*. This was probably because the sequence differences between Badari Dhan signals of linked markers were diluted by other families. Without prior information, *Hd1* would not be mapped to a precise position using MLM (Q + K).

Unlike *Ehd1* and *Hd1*, it was not possible to discriminate alleles at *Hd3a*, *RFT1*, and *Ghd7* despite the previous reports [29,30,32]. However, it should be noted that joint QTL precisely mapped the peaks of *Ehd1*, *Hd1*, *Ghd7*, and a combined peak of *RFT1/Hd3a* (Figure 3). The joint QTL mapping approach has been reported to amplify small effects signals found on individual family RILs [18]. The results in the present study indicated that joint QTL mapping is advantageous in the precision of QTL positioning compared with MLM (Q + K).

A QTL tightly linked to *Ghd7* was detected in WNAM02 and WNAM35. Another QTL near, but not tightly linked to *Ghd7* (11.0Mb on chromosome 7, 2.08Mb apart from *Ghd7* (8.93Mb)), was detected in WNAM39, where the functional allele of *Hd1* segregated. This QTL showed an opposite additive effect as that expressed in WNAM02 and WNAM35. *Ghd7* was reported to have the ability to switch its additive effects by an epistasis with

Hd1 [38]. The underlying gene in the QTL detected in WNAM39 remains unclear. A well-saturated linkage map will facilitate the characterization of genes underpinning this QTL.

3.4. Prospects

In the modeling used for genomic selection in practical breeding programs, population size is more important than marker density [37]. Further expansion of the population being conducted by the authors will enable accurate association analysis by eliminating the limitations of population size. The new population, with more founders, will have a higher representation of rare alleles, which is important for crop improvements [39–41]. In addition, the immortal nature of NAM populations allows for evaluations in various environments. Thus, they could be used to reveal heading time gene networks in detail. In conclusion, our results elucidated that *aus*-NAM will be a valuable genetic resource for QTL mapping and may be used for genomic selection.

4. Materials and Methods

4.1. Plant Materials

An *aus*-NAM was built using a temperate japonica variety, Taichung 65 (T65), as the common female parent. The five *aus* cultivars Kasalath, Kalo Dhan, Shoni, ARC5955, and Badari Dhan (kindly supplied by the National Agricultural Research Organization (NARO) Genebank, Tsukuba, Japan) were used as diversity donor parents (founders) [28]. The five *aus* varieties were crossed to T65, and RILs were derived from the F₂ generation using single-seed descent (SSD) to obtain F₅ in 2015. The RILs (F₁₃) from T65 × DV85 and T65 × ARC10313 were generated at Kyushu University, Fukuoka, Japan, using SSD, and were kindly provided through National Bioresource Project. The 7 families of RILs were designated as WNAM02 (Kasalath), WNAM29 (Kalo Dhan), WNAM31 (Shoni), WNAM35 (ARC5955), WNAM39 (Badari Dhan), WNAM72 (DV85), and WNAM73 (ARC10313).

4.2. Trait Evaluation and Statistical Analysis

Field trials were conducted in 2015 at Togo Field, Nagoya University, Aichi, Japan (35°06′36.5″ N, 137°05′06.3″ E). Twelve to fifteen seeds per line of F₅ were sown on 27 May, and four plants per line were transplanted on 2 July, in a randomized design without replication in a single plot. The spacing was 20 cm between the hills and 30 cm between rows. Standard agronomic management was followed during the trials, except no fertilizer was applied.

Days to heading (DTH) was calculated as the difference between the date of emergence of inflorescence and sowing. The average standard deviation within a line was approximately 2.64 days; thus, most of the lines can be regarded as genetically fixed. The second plant in each line was used for both of phenotyping and genotyping. Phenotype value distributions across subpopulations were examined. To find which trait means were significantly different among *aus*-NAM families, a one-way analysis of variance (ANOVA) followed by Tukey HSD with a 95% confidence level was performed. All statistical analyses and visualizations were performed using R version 4.0.3 [42].

4.3. Genotyping By Sequencing (GBS)

For genotyping, approximately 5 cm of leaf tissues from each line was sampled into paper envelopes. The samples were dried in an oven at 53 °C overnight and then stored at 6 °C. Total DNA of RILs and founders were extracted using a modified Dellaporta method [43]. DNA qualities were checked by electrophoresis on a 0.6% agarose gel in 1 × Tris/Borate/EDTA (TBE; 40 mmol/L Tris, 20 mmol/L acetic acid, and 0.5 mmol/L EDTA-2Na). The Quantiflour dsDNA system (Promega, Madison, WI, USA) was used for the quantification of the extracted total double-stranded DNA. GBS libraries were prepared using reported protocols [35,36]. Briefly, 200 ng (20 ng × 10 µL) of individual samples of DNA was double-digested with *KpnI* and *MspI* enzymes (New England Biolabs Inc., Ipswich, MA, USA), ligated to barcode adaptors, pooled (multiplexed), and purified using

a QIAquick PCR Purification Kit (Qiagen Sciences, Germantown, MD, USA). Flowcell primers were added to the pooled samples and amplified. The library was sequenced using Illumina MiSeq (Illumina, San Diego, CA, USA).

Raw sequences were processed using the TASSEL-GBS pipeline [44] with default parameters, except (1) minimum allele frequency higher than 0.02, (2) minimum locus coverage set to 0.3, and (3) heterozygous sites and taxa that exceeded 0.125 were filtered out. Os-Nipponbare-IRGSP-1.0 [45] was used as the reference for SNP identification. SNPs were further filtered based on parental polymorphism. Only loci that were polymorphic between parents but monomorphic in each parent were included. Additionally, missing data were imputed using the FSFHap algorithm [46].

4.4. Whole-Genome Resequencing of T65 and *aus* Founders

DNA of the founders was extracted using the cetyl trimethylammonium bromide (CTAB) method, then fragmented using Covaris Model S2 (Covaris, Woburn, MA, USA) and used to construct a sequencing library using a TruSeq DNA LT kit (Illumina, San Diego, CA, USA). Sequencing was conducted using Illumina MiSeq with a MiSeq Reagent Kit v3 (600 cycles). Variant calling was conducted following the standard protocol of the Genome Annotation ToolKit (GATK) [47] using Os-Nipponbare-IRGSP-1.0 [45] as the reference.

4.5. Projection of Parental Variants and Population Structure Estimation

SNPs from parental read sequences were projected onto each RIL. Projections were performed in two steps: (1) employ GBS markers as skeletons and (2) check adjacent skeleton markers. If homozygous and possessing the common allele type as one of the parents, project the parental SNPs onto the intervals; otherwise, set the intervals as missing. Population stratifications were estimated using a probabilistic PCA (PPCA) algorithm in the Bioconductor package PCA methods [48] implemented in R.

4.6. Simple QTL Mapping

Genotype files in hapmap format were converted to parent-based format, where A represented T65 and B represented *aus* genotype; heterozygotes and missing were represented by H and “-”, respectively. Kosambi mapping function in R/qtl package [49] was used to obtain genetic distances in cM. QTL mapping was performed based on interval mapping using the “hk” method implemented in R/qtl. The additive effects of a marker were calculated as ((average of *aus*) – (average of T65))/2. Positive and negative additive effects values implied that *aus* and T65 alleles increased trait values respectively. A logarithm of odds (LOD) value of 3 was used as the threshold, while 3.04 was obtained as the empirical threshold (type I error of 0.05) based on 1000 permutation tests [50].

4.7. Joint QTL Mapping

For joint linkage mapping, genotype information in a common genetic map and DTH were subjected to joint inclusive composite interval mapping (JICIM) [51]. Missing phenotypes were replaced by the mean of the trait, 1 cM step was selected, and LOD threshold was obtained from 1000 permutation tests with a type I error of 0.05. The genotype file was converted into a numeral format where the T65 genotype was represented by 0, and *aus* genotype was represented by 2; heterozygous genotypes and missing genotypes were represented by 1 and –1, respectively. Positive and negative additive effect values mean that *aus* and T65 alleles increased trait values respectively.

4.8. Genome-Wide Association Analysis

TASSEL (trait analysis by association, evolution, and linkage) software [52] was used for QTL mapping using a GWAS-based method. General linear model (GLM) and mixed linear model (MLM) with principal components (Q) to account for population structure and genomic kinship (K) as covariates (MLM (Q + K)) were used. In MLM (Q + K), pedigree information (family) was used by TASSEL software as covariates. The threshold

for declaring significance was determined using Bonferroni with the equation: $p \leq 1/N$ ($\alpha = 0.05$), where N is the number of markers [53].

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/plants10061255/s1>, Figure S1: Violin plots showing the frequency distributions for days to heading (DTH) in *aus* nested association mapping population. Yellow and blue dots represent the population founders, *aus* and T65 respectively. Black dots show the recombinant inbred lines average in each family. Groups with no significant difference by Tukey HSD with a 95% confidence level are represented by the same letters shown above the plots; Figure S2: Probabilistic principal component analysis (PPCA) showing population structure in *aus*-NAM. (A) PC2 vs. PC3 (B) PC3 vs. PC4 (C) PC5 vs. PC6. Different shapes and colors represent different *aus*-NAM families; Figure S3: Manhattan plot of days to heading using general linear model (GLM). The red horizontal line marks the threshold for genome wide significance (5.9) on a $-\log_{10}$ scale. A quantile-quantile (QQ) plot is shown in the right panel, where the observed P-values (Y-axis) against the expected p -values (X-axis) under the null hypothesis of no association are plotted on a $-\log_{10}$ scale. Each black dot indicates an SNP; Figure S4: Alignment of amino-acid sequences of *Hd1* in functional alleles of Nipponbare and Ginbouzu, and Badari Dhan (WRC39). The amino acid sequences were deduced from genomic sequences. Regions of the 2 ZF-B box and CCT motif are indicated. The sequence of Badari Dhan contained 6 non-synonymous mutations, but it was considered that the allele retains function; Table S1: Number of SNP markers and lines in *aus* nested association mapping (*aus*-NAM) population; Table S2: Joint linkage map statistics in *aus* nested association mapping population; Table S3: R^2 values obtained by probabilistic principal component analysis (PPCA); Table S4: Days to heading QTL detected using single-family QTL analysis; Table S5: Days to heading QTL and additive effects detected using joint linkage analysis.

Author Contributions: Conceptualization, K.D., H.S., and M.Y.; formal analysis, J.K.K., H.S., and K.D.; investigation J.K.K., H.S., M.T., J.-I.M., A.S., V.P.R., and K.D.; data curation, J.K.K. and K.D.; writing—original draft preparation, J.K.K. and H.S.; writing—review and editing, J.K.K., M.Y., S.N., and K.D., supervision, H.Y., Y.Y., A.Y., and K.D.; funding acquisition, K.D., M.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Cross-ministerial Strategic Innovation Promotion Program (SIP), National Bioresource Project (NBRP), RIKEN Cluster for Science, Technology and Innovation Hub (RCSTI), and JSPS KAKENHI Grant Number 21K05522/20KK0138.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: *Aus*-NAM germplasm and genotypes are available by request via kdoi@agr.nagoya-u.ac.jp.

Acknowledgments: JKK is grateful to JICA Agri Net program for the scholarship. The authors would like to thank NIAS Genebank for providing the founder materials.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Reyes, V.P.; Angeles-Shim, R.B.; Mendioro, M.S.; Manuel, M.C.C.; Lapis, R.S.; Shim, J.; Sunohara, H.; Nishiuchi, S.; Kikuta, M.; Makihara, D.; et al. Marker-Assisted Introgression and Stacking of Major QTLs Controlling Grain Number (*Gn1a*) and Number of Primary Branching (*WFP*) to NERICA Cultivars. *Plants* **2021**, *10*, 844. [CrossRef]
2. Ikeda, M.; Miura, K.; Aya, K.; Kitano, H.; Matsuoka, M. Genes offering the potential for designing yield-related traits in rice. *Curr. Opin. Plant Biol.* **2013**, *16*, 213–220. [CrossRef] [PubMed]
3. Yano, M.; Sasaki, T. Genetic and molecular dissection of quantitative traits in rice. *Plant Mol. Biol.* **1997**, *35*, 145–153. [CrossRef]
4. Ogura, T.; Busch, W. From phenotypes to causal sequences: Using genome wide association studies to dissect the sequence basis for variation of plant development. *Curr. Opin. Plant Biol.* **2015**, *23*, 98–108. [CrossRef]
5. Mogga, M.; Sibiya, J.; Shimelis, H.; Lamo, J.; Yao, N. Diversity analysis and genome-wide association studies of grain shape and eating quality traits in rice (*Oryza sativa* L.) using DArT markers. *PLoS ONE* **2018**, *13*, e0198012. [CrossRef]
6. Yano, K.; Yamamoto, E.; Aya, K.; Takeuchi, H.; Lo, P.C.; Hu, L.; Yamasaki, M.; Yoshida, S.; Kitano, H.; Hirano, K.; et al. Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat. Genet.* **2016**, *48*, 927–934. [CrossRef] [PubMed]

7. Norton, G.J.; Travis, A.J.; Douglas, A.; Fairley, S.; Alves, E.P.; Ruang-Areerate, P.; Naredo, M.E.B.; McNally, K.L.; Hossain, M.; Islam, M.R.; et al. Genome wide association mapping of grain and straw biomass traits in the rice Bengal and Assam Aus Panel (BAAP) grown under alternate wetting and drying and permanently flooded irrigation. *Front. Plant Sci.* **2018**, *9*, 1223. [[CrossRef](#)] [[PubMed](#)]
8. Xiao, Y.; Liu, H.; Wu, L.; Warburton, M.; Yan, J. Genome-wide association studies in maize: Praise and stargaze. *Mol. Plant* **2017**, *10*, 359–374. [[CrossRef](#)]
9. Zhu, C.; Gore, M.; Buckler, E.S.; Yu, J. Status and prospects of association mapping in plants. *Plant Genome* **2008**, *1*, 5–20. [[CrossRef](#)]
10. Korte, A.; Farlow, A. The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* **2013**, *9*, 29. [[CrossRef](#)] [[PubMed](#)]
11. Yu, J.; Holland, J.B.; McMullen, M.D.; Buckler, E.S. Genetic design and statistical power of nested association mapping in maize. *Genetics* **2008**, *178*, 539–551. [[CrossRef](#)]
12. Dell’Acqua, M.; Gatti, D.M.; Pea, G.; Cattonaro, F.; Coppens, F.; Magris, G.; Hlaing, A.L.; Aung, H.H.; Nelissen, H.; Baute, J.; et al. Genetic properties of the MAGIC maize population: A new platform for high definition QTL mapping in *Zea mays*. *Genome Biol.* **2015**, *16*, 167. [[CrossRef](#)]
13. Xiao, Y.; Tong, H.; Yang, X.; Xu, S.; Pan, Q.; Qiao, F.; Raihan, M.S.; Luo, Y.; Liu, H.; Zhang, X.; et al. Genome-wide dissection of the maize ear genetic architecture using multiple populations. *New Phytol.* **2016**, *210*, 1095–1106. [[CrossRef](#)]
14. McMullen, M.D.; Kresovich, S.; Villeda, H.S.; Bradbury, P.; Li, H.; Sun, Q. Supporting online material for: Genetic properties of the maize nested association mapping population. *Science* **2009**, *325*, 737–741. [[CrossRef](#)] [[PubMed](#)]
15. Buckler, E.S.; Holland, J.B.; Bradbury, P.J.; Acharya, C.B.; Brown, P.J.; Brown, C.; Ersoz, E.; Flint-Garcia, S.; Garcia, A.; Glaubitz, J.C.; et al. The genetic architecture of maize flowering time. *Science* **2009**, *325*, 714–718. [[CrossRef](#)] [[PubMed](#)]
16. Tian, F.; Bradbury, P.J.; Brown, P.J.; Hung, H.; Sun, Q.; Flint-Garcia, S.; Rocheford, T.R.; McMullen, M.D.; Holland, J.B.; Buckler, E.S. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.* **2011**, *43*, 159–162. [[CrossRef](#)] [[PubMed](#)]
17. Peiffer, J.A.; Romay, M.C.; Gore, M.A.; Flint-Garcia, S.A.; Zhang, Z.; Millard, M.J. The genetic architecture of maize height. *Genetics* **2014**, *196*, 1337–1356. [[CrossRef](#)] [[PubMed](#)]
18. Fragoso, C.A.; Moreno, M.; Wang, Z.; Heffelfinger, C.; Arbelaez, L.J.; Aguirre, J.A.; Franco, N.; Romero, L.E.; Labadie, K.; Zhao, H.; et al. Genetic architecture of a rice nested association mapping population. *G3* **2017**, *7*. [[CrossRef](#)]
19. Bajgain, P.; Rouse, M.N.; Tsilo, T.J.; Macharia, G.K.; Bhavani, S.; Jin, Y.; Anderson, J.A. Nested association mapping of stem rust resistance in wheat using genotyping by sequencing. *PLoS ONE* **2016**, *11*, e0155760. [[CrossRef](#)]
20. Maurer, A.; Draba, V.; Jiang, Y.; Schnaithmann, F.; Sharma, R.; Schumann, E.; Kilian, B.; Reif, J.C.; Pillen, K. Modelling the genetic architecture of flowering time control in barley through nested association mapping. *BMC Genom.* **2015**, *16*, 290. [[CrossRef](#)]
21. Song, Q.; Yan, L.; Quigley, C.; Jordan, B.D.; Fickus, E.; Schroeder, S.; Song, B.H.; Charles An, Y.Q.; Hyten, D.; Nelson, R.; et al. Genetic characterization of the soybean nested association mapping population. *Plant Genome* **2017**, *10*. [[CrossRef](#)] [[PubMed](#)]
22. Bouchet, S.; Olatoye, M.O.; Marla, S.R.; Perumal, R.; Tesso, T.; Yu, J.; Tuinstra, M.; Morris, G.P. Increased power to dissect adaptive traits in global sorghum diversity using a nested association mapping population. *Genetics* **2017**, *206*, 573–585. [[CrossRef](#)]
23. Hu, J.; Guo, C.; Wang, B.; Ye, J.; Liu, M.; Wu, Z.; Xiao, Y.; Zhang, Q.; Li, H.; King, G.J.; et al. Genetic properties of a nested association mapping population constructed with semi-winter and spring oilseed rapes. *Front. Plant Sci.* **2018**, *9*, 1740. [[CrossRef](#)] [[PubMed](#)]
24. Garris, A.J.; Tai, T.H.; Coburn, J.; Kresovich, S.; McCouch, S. Genetic structure and diversity in *Oryza sativa* L. *Genetics* **2005**, *169*, 1631. [[CrossRef](#)]
25. Travis, A.J.; Norton, G.J.; Datta, S.; Sarma, R.; Dasgupta, T.; Savio, F.L.; Macaulay, M.; Hedley, P.E.; McNally, K.L.; Sumon, M.H.; et al. Assessing the genetic diversity of rice originating from Bangladesh, Assam and west Bengal. *Rice (New York)* **2015**, *8*, 35. [[CrossRef](#)]
26. Kihupi, A.; Angeles, E.; Khush, G. Genetic analysis of resistance to bacterial blight, *Xanthomonas oryzae* pv. *oryzae*, in rice, *Oryza sativa* L. *Euphytica* **2001**, *117*, 39–46. [[CrossRef](#)]
27. Takehisa, H.; Yasuda, M.; Fukuta, Y.; Kobayashi, N.; Hayashi, N.; Nakashita, H.; Abe, T.; Sato, T. Genetic analysis of resistance genes in an *Indica*-type rice (*Oryza sativa* L.), Kasalath, using DNA markers. *Breed. Sci.* **2009**, *59*, 253–260. [[CrossRef](#)]
28. Kojima, Y.; Ebana, K.; Fukuoka, S.; Nagamine, T.; Kawase, M. Development of an RFLP-based rice diversity research set of germplasm. *Breed. Sci.* **2005**, *55*, 431–440. [[CrossRef](#)]
29. Izawa, T.; Oikawa, T.; Sugiyama, N.; Tanisaka, T.; Yano, M.; Shimamoto, K. Phytochrome mediates the external light signal to repress *FT* orthologs in photoperiodic flowering of rice. *Genes Dev.* **2002**, *16*, 2006–2020. [[CrossRef](#)]
30. Kojima, S.; Takahashi, Y.; Kobayashi, Y.; Monna, L.; Sasaki, T.; Araki, T.; Yano, M. *Hd3a*, a rice ortholog of the Arabidopsis *FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. *Plant Cell Physiol.* **2002**, *43*, 1096–1105. [[CrossRef](#)]
31. Yano, M.; Katayose, Y.; Ashikari, M.; Yamanouchi, U.; Monna, L.; Fuse, T.; Baba, T.; Yamamoto, K.; Umehara, Y.; Nagamura, Y.; et al. *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the arabidopsis flowering time gene *CONSTANS*. *Plant Cell* **2000**, *12*, 2473–2483. [[CrossRef](#)] [[PubMed](#)]
32. Xue, W.; Xing, Y.; Weng, X.; Zhao, Y.; Tang, W.; Wang, L.; Zhou, H.; Yu, S.; Xu, C.; Li, X.; et al. Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat. Genet.* **2008**, *40*, 761–767. [[CrossRef](#)]

33. Doi, K.; Izawa, T.; Fuse, T.; Yamanouchi, U.; Kubo, T.; Shimatani, Z.; Yano, M.; Yoshimura, A. *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.* **2004**, *18*, 926–936. [[CrossRef](#)] [[PubMed](#)]
34. Yamamoto, E.; Yonemaru, J.-i.; Yamamoto, T.; Yano, M. OGRO: The overview of functionally characterized genes in rice online database. *Rice* **2012**, *5*, 26. [[CrossRef](#)]
35. Poland, J.A.; Brown, P.J.; Sorrells, M.E.; Jannink, J.-L. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* **2012**, *7*, e032253. [[CrossRef](#)]
36. Furuta, T.; Ashikari, M.; Jena, K.K.; Doi, K.; Reuscher, S. Adapting genotyping-by-sequencing for rice F2 populations. *G3 (Bethesda)* **2017**, *7*, 881–893. [[CrossRef](#)]
37. Xu, Y.; Wang, X.; Ding, X.; Zheng, X.; Yang, Z.; Xu, C.; Hu, Z. Genomic selection of agronomic traits in hybrid rice using an NCII population. *Rice* **2018**, *11*, 32. [[CrossRef](#)]
38. Fujino, K.; Yamanouchi, U.; Nonoue, Y.; Obara, M.; Yano, M. Switching genetic effects of the flowering time gene *Hd1* in LD conditions by *Ghd7* and *OsPRR37* in rice. *Breed Sci.* **2019**, *69*, 127–132. [[CrossRef](#)]
39. Valluru, R.; Gazave, E.E.; Fernandes, S.B.; Ferguson, J.N.; Lozano, R.; Hirannaiah, P.; Zuo, T.; Brown, P.J.; Leakey, A.D.B.; Gore, M.A.; et al. Deleterious Mutation Burden and Its Association with Complex Traits in Sorghum (*Sorghum bicolor*). *Genetics* **2019**, *211*, 1075. [[CrossRef](#)] [[PubMed](#)]
40. Kremling, K.A.G.; Chen, S.Y.; Su, M.H.; Lepak, N.K.; Romay, M.C.; Swarts, K.L.; Lu, F.; Lorant, A.; Bradbury, P.J.; Buckler, E.S. Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. *Nature* **2018**, *555*, 520–523. [[CrossRef](#)]
41. Gage, J.L.; Monier, B.; Giri, A.; Buckler, E.S. Ten Years of the Maize Nested Association Mapping Population: Impact, Limitations, and Future Directions. *Plant Cell* **2020**, *32*, 2083–2093. [[CrossRef](#)]
42. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020. Available online: <https://www.R-project.org/> (accessed on 1 March 2021).
43. Dellaporta, S.L.; Wood, J.; Hicks, J.B. A plant DNA miniprep: Version II. *Plant Mol. Biol. Rep.* **1983**, *1*, 19–21. [[CrossRef](#)]
44. Glaubitz, J.C.; Casstevens, T.M.; Lu, F.; Harriman, J.; Elshire, R.J.; Sun, Q.; Buckler, E.S. TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS ONE* **2014**, *9*, e090346. [[CrossRef](#)]
45. Kawahara, Y.; de la Bastide, M.; Hamilton, J.P.; Kanamori, H.; McCombie, W.R.; Ouyang, S.; Schwartz, D.C.; Tanaka, T.; Wu, J.; Zhou, S.; et al. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice* **2013**, *6*, 4. [[CrossRef](#)]
46. Swarts, K.; Li, H.; Romero Navarro, J.A.; An, D.; Romay, M.C.; Hearne, S.; Acharya, C.; Glaubitz, J.C.; Mitchell, S.; Elshire, R.J.; et al. Novel methods to optimize genotypic imputation for low-coverage, next-generation sequence data in crop plants. *Plant Genome* **2014**, *7*. [[CrossRef](#)]
47. DePristo, M.A.; Banks, E.; Poplin, R.; Garimella, K.V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; del Angel, G.; Rivas, M.A.; Hanna, M.; et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **2011**, *43*, 491–498. [[CrossRef](#)]
48. Stacklies, W.; Redestig, H.; Scholz, M.; Walther, D.; Selbig, J. PCAMethods—A bioconductor package providing PCA methods for incomplete data. *Bioinformatics* **2007**, *23*, 1164–1167. [[CrossRef](#)] [[PubMed](#)]
49. Arends, D.; Prins, P.; Jansen, R.C.; Broman, K.W. R/qtl: High-throughput multiple QTL mapping. *Bioinformatics* **2010**, *26*, 2990–2992. [[CrossRef](#)]
50. Churchill, G.A.; Doerge, R.W. Empirical threshold values for quantitative trait mapping. *Genetics* **1994**, 963–971. [[CrossRef](#)]
51. Li, H.; Bradbury, P.; Ersoz, E.; Buckler, E.S.; Wang, J. Joint QTL linkage mapping for multiple-cross mating design sharing one common parent. *PLoS ONE* **2011**, *6*, e17573. [[CrossRef](#)]
52. Bradbury, P.J.; Zhang, Z.; Kroon, D.E.; Casstevens, T.M.; Ramdoss, Y.; Buckler, E.S. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **2007**. [[CrossRef](#)] [[PubMed](#)]
53. Haynes, W. Bonferroni Correction. In *Encyclopedia of Systems Biology*; Dubitzky, W., Wolkenhauer, O., Cho, K.-H., Yokota, H., Eds.; Springer: New York, NY, USA, 2013; p. 154.