agronomy

MDPI

Article

Genome-Wide Association Study Using Genotyping by Sequencing for Bacterial Leaf Blight Resistance Loci in Local Thai Indica Rice

Chananton Danaisilichaichon [1,2], Phanchita Vejchasarn [2], Sujin Patarapuwadol [3], Alessandro Tondelli [4], Giampiero Valè [5], Theerayut Toojinda [6] and Chatchawan Jantasuriyarat [1,7,*]

1 Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand
2 Rice Department, Ministry of Agriculture, Bangkok 10900, Thailand
3 Department of Plant Pathology, Faculty of Agriculture at Kamphaeng Saen, Kasetsart University, Nakhon Pathom 73140, Thailand
4 Council for Agricultural Research and Economics, Research Centre for Genomics and Bioinformatics, 29017 Fiorenzuola d'Arda, Italy
5 Department for Sustainable Development and Ecological Transition–DiSSTE, University of Piemonte Orientale, Piazza S. Eusebio 5, 13100 Vercelli, Italy
6 National Center for Genetic Engineering and Biotechnology (BIOTEC), 113 Thailand Science Park, Pahonyothin Road, Khlong Nueng, Khlong Luang, Pathumthani 12120, Thailand
7 Center for Advanced Studies in Tropical Natural Resources, National Research University-Kasetsart (CASTNAR, NRU-KU), Kasetsart University, Bangkok 10900, Thailand
* Correspondence: fscicwj@ku.ac.th; Tel.: +66-81627-4635

Abstract: Bacterial leaf blight (BLB) is a devastating disease caused by *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), which poses a significant threat to global rice production. In this study, a genome-wide association study (GWAS) was conducted using the genotyping-by-sequencing (GBS) approach to identify candidate single nucleotide polymorphisms (SNPs) associated with BLB resistance genes. The study utilized 200 indica rice accessions inoculated with seven distinct *Xoo* isolates and filtered highly significant SNPs using a minor allele frequency (MAF) of >5% and a call rate of 75%. Four statistical models were used to explore potential SNPs associated with BLB resistance, resulting in the identification of 32 significant SNPs on chromosomes 1–8 and 12 in the rice genome. Additionally, 179 genes were located within ±100 kb of the SNP region, of which 49 were selected as candidate genes based on their known functions in plant defense mechanisms. Several candidate genes were identified, including two genes in the same linkage disequilibrium (LD) decay as the well-known BLB resistance gene (*Xa1*). These findings represent a valuable resource for conducting further functional studies and developing novel breeding strategies to enhance the crop's resistance to this disease.

Keywords: genome-wide association; bacterial leaf blight; *Xanthomonas oryzae* pv. *oryzae*; resistance gene; indica rice

Citation: Danaisilichaichon, C.; Vejchasarn, P.; Patarapuwadol, S.; Tondelli, A.; Valè, G.; Toojinda, T.; Jantasuriyarat, C. Genome-Wide Association Study Using Genotyping by Sequencing for Bacterial Leaf Blight Resistance Loci in Local Thai Indica Rice. *Agronomy* **2023**, *13*, 1286. https://doi.org/10.3390/agronomy13051286

Academic Editors: Niharika Sharma, Davinder Singh and Kalenahalli N. Yogendra

Received: 22 March 2023
Revised: 26 April 2023
Accepted: 27 April 2023
Published: 29 April 2023

1. Introduction

Rice (*Oryza sativa* L.) is one of the world's most popular staple foods and is consumed by more than half of the world's population. In 2021, almost 780 million tons of rice were produced, with about 90% of the production occurring in Asia [1]. However, current production levels are still insufficient to meet the demand for rice in Asia. Food security can be assured by preventing the impacts of climate change, such as pathogen spread. This is one of the fundamental problems in rice production resulting from biotic stresses, which were identified as a major problem [2–4]. World rice production is estimated to incur losses due to pests and pathogens, with bacterial leaf blight (BLB) being among the top five pathogens responsible for these losses [5].

BLB disease, caused by *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), is one of the most serious diseases affecting rice production worldwide. In the past, there was confusion between

two diseases: leaf streak and bacterial leaf blight. This was the case until 1957, when they were clearly classified. In 1978, *Xanthomonas campestris* pv. *oryzae* was identified as the causal agent of both diseases. However, in 1990, to alleviate confusion, the pathogen was renamed *Xanthomonas oryzae* pv. *oryzae*. [4,6]. The first report of BLB in rice was published in Japan in 1884 [5,7]. The disease then spread to other parts of Asia, including China, Korea, India, Indonesia, the Philippines, Taiwan, Thailand, and Vietnam, before spreading worldwide [6]. The first report of *Xoo* in Thailand was found in Bangkok, where it infected several Thai rice varieties, including Hom Set Ti, Khao Ta Hang, and Niew Kan Plu [8]. As a vascular disease, BLB can adversely affect rice yields. The losses could depend on the variety of rice, environmental conditions, and rice development stage [9–11]. The infection resulted in a systemic condition in which the plant leaves presented a tannish grey-to-white appearance with lesions occurring along their veins. The symptoms can be divided into three phases based on their characteristics, including the leaf blight phase, kresek phase, and pale-yellow phase, each with its own infection stage. Leaf blight is the most common symptom observed after plant transplanting and usually appears within the first 4–6 weeks following planting. The lesions appear as yellow or straw-colored stripes along the leaf veins, and the leaf edges are also affected, forming a curved margin from the tips to the leaf bases [11]. A small yellow droplet of bacterial ooze may be observed at the beginning of the blight phase. The second phase of seedling infection is known as the kresek phase, causing wilting of young plants during the early stages of infection. Once leaves become infected, they turn yellow or green and begin to roll up. Another phase involves pale-yellow leaves on the youngest parts of the plant, which become white or pale yellow in color [11]. As a result of the lesions on the leaf blade or leaf tips, BLB symptoms, such as yellowing, wilting, and drying, were easily identified.

The simplest method of preventing BLB is using chemical pesticides; however, widespread application of these chemicals poses a threat to rice food safety. Additionally, any chemical control implemented in a monsoon climate would be completely ineffective. In order to control this disease and ensure food security, resistant rice cultivars containing *R* genes must be developed and implemented [12–14]. According to reports in recent years, there have been more than 45 genes that confer resistance to BLB, which has increased over the years [15,16]. As far as *R* genes are concerned, they are mostly found on chromosome 11 [17–20]. However, *R* genes are also present on other rice chromosomes except for 9 and 10 [19,21–24]. Overall, several GWAS studies have been conducted to identify genes associated with BLB resistance in rice. Shu et al. [25] identified 58 candidate genes on chromosome 7, and a GWAS using 259 rice accessions revealed significant single nucleotide polymorphisms (SNPs) near several known resistance genes including *xa25*, *Xa26*, *Xa21*, *Xa23*, *Xa33*, *xa13*, and *Os-11N3* genes. Dilla-Ermita et al. [17] conducted a GWAS on 285 rice accessions and identified novel SNPs linked to known *Xa* genes. The results found that five genes (*xa5*, *Xa7*, *xa13*, *xa24*, and *xa34*) originate from the *aus* genotype, while ten genes (*Xa2*, *Xa4*, *Xa11*, *Xa14*, *Xa16*, *Xa18*, *xa25*, *Xa26*, *xa28*, and *Xa39*) are derived from *indica*. Furthermore, a novel BLB resistance gene was identified on chromosome 11, designated as *Xa43(t)* within the 27.83–27.95 Mbp region by Kim and Reinke [23]. Moreover, Korinsak et al. [18] identified 147 genes on chromosome 1–6, 8, 9, and 11 using GWAS by phenotypic testing against 20 *Xoo* isolates. Among the identified genes, ten genes on chromosome 11 were found to be located close to well-known genes within the 17–29 Mb region.

Genotyping-by-sequencing (GBS) is a well-established method for generating genotypic information for GWAS. This technique provides a large number of SNPs for species diversity classification, haplotype map construction, the performance of GWAS, and the discovery of candidate SNPs [24–26]. Due to its effectiveness and low cost, this technique has been successfully incorporated into rice using restriction enzymes as the primary method for reducing the complexity of DNA samples and generating high-quality markers [25,27,28]. Many research studies of rice have incorporated this technique into their GWAS methods [17,29–32]. GWAS has become a powerful approach for unraveling the molecular genetics associated with natural phenotypic variation in rice. Various traits have

been investigated using this technique, including agronomic traits, quality, and resistance to biotic and abiotic stresses [26,33–36]. Data collected from a variety of populations, plant growth stages, and different *Xoo* variations have been utilized in GWAS for gene mining, QTL mapping, and searches for candidate SNPs for BLB resistant traits in rice. In 2021, Shu et al. [25] conducted 259 rice lines including 240 indica and 19 japonica rice lines against two *Xoo* races (P3 and P6 races) and then discovered 2 NLR protein-encoding genes that exhibited a significantly high expression in the rice resistance lines. In 2022, Lu et al. [37] identified 13 quantitative trait loci (QTL) associated with BLB resistance (eight novel QTLs and five QTLs harboring known genes) located on chromosomes 5 and 11.

According to the World Agricultural Outlook Board's report on rice from the US Department of Agriculture, Thailand ranks sixth worldwide in rice production and second in rice exports [38,39]. In 2018–2022, Thailand exported more than 35 million tons of rice [40]. In this regard, Thailand is one of the principal production regions for rice in the world. It has abundant rice genetic resources and a diversity of ecologies such as irrigated, rain-fed lowland, deep water, and upland ecosystems [41,42]. Thailand's rice germplasm consists of improved cultivars, landraces, and local cultivars [43]. However, only one report has described the identification of a BLB resistance gene in Thailand's germplasm using GWAS [18]. Therefore, the objective of this study is to identify candidate SNPs for BLB resistance genes in 200 accessions of Thailand's local *indica* rice at the tillering stage against Thai *Xoo* isolates. To achieve this, we conducted a GWAS analysis and phenotypic evaluation by analyzing the average lesion length of rice panels against seven Thai *Xoo* isolates, along with genotyping data from 25,338 SNPs using the GBS method.

2. Materials and Methods

2.1. Plant Materials

Two hundred Thai landrace rice accessions were utilized to represent the diverse geographical regions and ecosystems found in Thailand. This comprehensive representation of rice diversity is crucial for breeding programs, genetic studies, and conservation efforts. The accessions included lowland, upland, and floating varieties, sourced from various regions of Thailand, with 70 from the northern region, 39 from the central region, 61 from the northeastern region, four from the western region, 14 from the eastern region, and six from the southern region. To grow the accessions for evaluation, the study was conducted under greenhouse conditions at the Thailand Rice Science Institute (TRSI) from October 2021 to January 2022, where the temperature ranged from 24 to 35 °C during the day and from 17 to 33 °C at night. The seed accessions were germinated on tissue paper for 48 h, and one week later were transplanted into pots containing three plants each, for a total of six plants per accession. Fertilization of the plants occurred at 15 and 21 days of age, using a 15-15-15 fertilizer and 46-0-0 fertilizer, respectively [44,45]. In addition to the study accessions, five standard rice varieties known for their resistance to the disease (DV85, IRBB5, IRBB21, RD7, and RD31) were included, as well as three susceptible rice varieties (TN1, IR24, and KDML105) as a comparison for the evaluation of disease resistance.

2.2. Evaluation of Bacterial Leaf Blight Resistance in Rice Plant

Seven *Xoo* isolates were selected for artificial inoculation of 200 Thai landrace rice accessions. The isolates are named SK2-3-F, XORE1-1, XONS2-1, 3XOBR2-2, 59XOCRCS7-3, 61XOSPSJ1-10, and 61XOSPSJ2-10 They were selected from a diversity of genetic and pathotype representations of *Xoo*. The *Xoo* cultures were multiplied for 48 h at 28 °C on nutrient agar medium and then suspended in sterilized water to a final concentration of 10^9 CFU/mL. Each rice accession was inoculated with six plants using the leaf-clipping method at the tillering stage, with four fully expanded leaves clipped [46–49]. The lesion length (LL) was measured 14 days after inoculation or until a stable lesion was visible on the susceptible standard check variety [17,50]. The LL measurement was taken from the cutting edge of the leaf to the necrotic blight lesion in the middle vein of the leaf with the unit in centimeters [48]. The scoring was based on the standard method of the International

Rice Research Institute (IRRI) [17,50–53]. Four groups of scoring were established based on LL, with lesion lengths ranging from 0–5 cm considered resistant, >5–10 cm moderately resistant, >10–15 cm moderately susceptible, and more than 15 cm susceptible [17,50,54]. The average value of lesion length was used as the phenotypic data for the GWAS analysis.

2.3. Genotyping Preparation

The study utilized two hundred BAM files, provided by the Ubon Ratchathani Rice Research Center in the Rice Department of Thailand, to obtain genotypic information on the rice varieties involved. The DNA library was generated using a *Ape*KI restriction enzyme, followed by ligation with adaptors specific to the Ion S5™ XL Sequence (Thermo Fisher Scientific, Waltham, MA, USA). Using E-Gel™ SizeSelect™ agarose gels (Invitrogen, Waltham, MA, USA), 500–300 base pairs of fragments were selected. All sequence data were aligned to the Nipponbare genome as a reference genome (Nipponbare Reference-IRGSP-1.0) using the Ion Torrent™ Suite Software Alignment Plugin version 5.2.2 (Thermo Fisher Scientific, Waltham, MA, USA) and then BAM files were automatically generated. Duplicate reads in the BAM format were removed using Picard Tools version 2.27.0 [55]. The BAM files were created as fastq files using Samtools v1.9 [56] and realigned with the japonica reference genome using a Burrow–Wheeler Aligner (BWA) v 0.7.17 [57] and SAMtools. Variant were called using the Genome Analysis Toolkit (GATK) software v 4.2.3.0 [42,58] where SNPs and insertions-deletions (indels) were named using the GVCF analysis, and variants were grouped using GATK's CombineGVCFs function [59]. The joint call step was then performed using genotypeGVCFs of the GATK software [60], followed by SNP extraction using SelectVariantsGATK based on the reference genome. To improve the quality of the SNPs, Tassel v.5 was utilized to remove low quality SNPs that possessed a minor allele frequency less than 5% and a call rate below 75% [29,61–63]. Genetic datasets generated using GBS technology may contain missing values in SNP data due to low coverage or other technical issues. Imputation is a common method used to impute these missing values based on the linkage disequilibrium between SNPs in the dataset. Beagle version 5.4 is a widely used software tool for imputing genotypes in such datasets and is well-suited for use with GBS data. Therefore, in this study, we used Beagle version 5.4 to impute the missing values in the SNP data obtained from GBS genotyping. The algorithm is consistently high in accuracy within 90–100% for the major homozygous genotype, even with increasing missing rates [64–66]. As a result of filtering, a total of 25,338 high-quality SNPs were included in the GWAS analysis.

2.4. Population Genetics Structure Analysis, Principal Component Analysis, and Linkage Disequilibrium Assessment

The estimation of subgroup numbers in the rice panel was accomplished through the application of population structure analysis and principal component analysis (PCA). Following the imputation and pruning of data obtained from the panel of 200 rice accessions, 25,338 SNPs with a minor allele frequency greater than 5% were obtained. The STRUCTURE 2.3.4 program was utilized for estimating the population structure and K values were determined by setting the range from 1 to 10 to find the lowest coefficient of variation [28,67,68]. Both PCA and linkage disequilibrium (LD) analysis were performed using the Tassel [61]. The PCA was conducted on 25,338 SNPs obtained from a total of 4,086,284 SNPs within the rice panel through an initial analysis of all 200 rice accessions. Subpopulation verification was performed by combining and naming all the 200 BAM files with a representative *japonica* rice, followed by plotting using the RStudio program. Using the 25,338 highly accurate SNPs, Hapmap files were generated through the Tassel software. The LD was determined by calculating the pairwise squared correlation (r^2) between SNP markers for each pair of loci. The LD analysis was conducted following the method described by Hill and Weir [69], based on a sliding window of 50 SNPs with a minor allele frequency of 0.05. Subsequently, a scatter plot with smoothing spline regression lines was developed. In addition to these values, the LD decay rate was estimated based

on the physical distance between markers and the r^2 values and was determined when the r^2 value declined to half its average maximum value [70].

2.5. Genome-Wide Association Analysis

In this GWAS, multiple statistical methods were used to control for false-negative and false-positive results [71,72]. The Genomic Association and Prediction Integrated (GAPIT) tool was employed to perform the analysis, which was run using the R version 4.1.2 [73,74]. The 25,338 SNP loci were analyzed in relation to seven *Xoo* isolates to identify the association with BLB resistance. The four statistic methods were used as the mixed linear models (MLM) [75], Settlement of MLM Under Progressively Exclusive Relationship (SUPER) [76], fixed and random model circulation probability unification (FarmCPU) [77], and the Bayesian-information and linkage disequilibrium iteratively nested keyway (BLINK) method [78] were used to control for false-negative results [71]. The threshold for identifying significant SNPs was determined using the False Discovery Rate (FDR) [18,25,36,79,80]. Gene annotations within a 100 kb region surrounding the highly associated SNPs were taken into consideration for SNP identification. To visualize the results of the GWAS, Manhattan plots and quantile-quantile (Q-Q) plots were generated using the GAPIT tool in R. The scatter plot of the Manhattan plot provides a summary of the results by plotting the negative logarithm of the *p*-value obtained from each GWAS model against the genomic region of each SNP on the 12 rice chromosomes. On the other hand, the Q-Q plot visualizes the observed distribution of the negative logarithm of the *p*-values in comparison to their expected distribution under the null hypothesis.

2.6. Candidate SNP Identification

In order to explore candidate genes underlying BLB resistance, SNPs with a *p*-value less than 0.05 were considered as genomic regions carrying candidate genes. All gene loci within ±100 kb of the flanking regions of each SNP were extracted from the annotation of the Oryza sativa reference sequence (OsNippon-bare-Reference-IRGSP-1.0) [81]. To validate the results of GWAS, all annotated genes contained within the genomic regions were compared with genes known to be related to the phenotypic traits analyzed, as available in the Oryzabase database, the rice genome annotation project, and The Rice Annotation Project Database [20,82–85], or present in the literature as QTLs or genes. The Haploview 4.2 software was used to estimate the local LD block within ±100 kb of the genomic region containing significant SNPs [86]. Subsequent to identifying the candidate genes, haplotype analysis was conducted using all SNPs within their coding regions. A major haplotype was determined to be present when a minimum of 10 accessions were detected [37].

3. Results

3.1. Phenotypic Evaluation among Rice Panel

Through the inoculation of 200 rice varieties with seven *Xoo* isolates, distinct symptoms were identified between the various rice accessions. The lesion length distribution among *Xoo* isolates SK2-3-F, XORE1-1, XONS2-1, 3XOBR2-2, 59XOCRCS7-3, and 61XOSPSJ1-10 are presented in the bar graph (Figure 1a–g). The lesion length distribution is continuous and ranges from 1 to 35 cm. Thirty-one percent of 200 Thai landrace rice accessions showed a resistant reaction to *Xoo* isolate 59XOCRCS7-3, which was the highest percentage among the seven *Xoo* isolates used in this study (Figure 1h). While *Xoo* isolate SK2-3-F had 20% of the examined rice accessions, which showed a susceptible reaction.

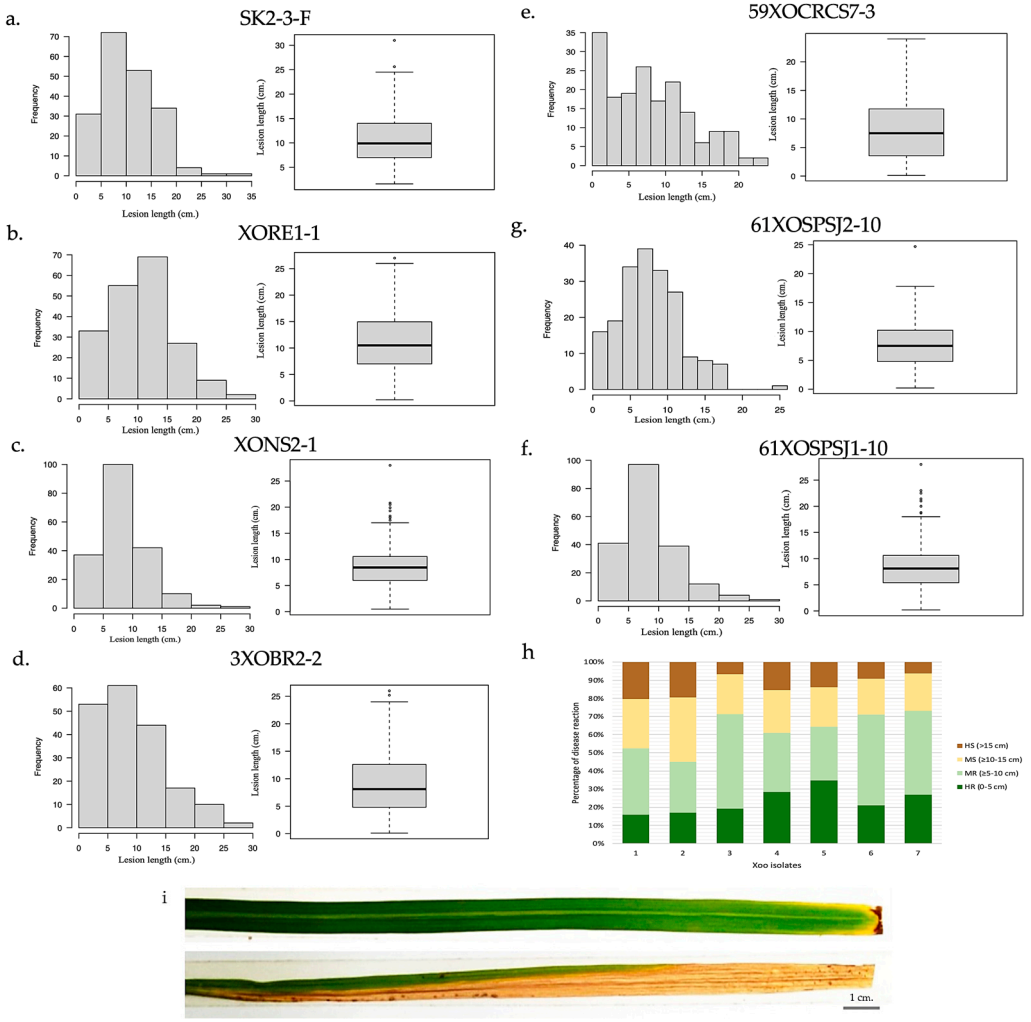

Figure 1. Pathogenicity of 200 Thai rice accessions against seven different isolates of *Xoo*. (**a**–**g**) Bar graphs and box plots show the distributions of phenotypic traits in response to *Xoo* isolates SK2–3–F9, XORE1–1, XONS2–1, 3XOBR2–2, 59XOCRCS7–3, 61XOSPSJ1–10, and 61XOSPSJ2–10. The LL caused by each *Xoo* isolate is represented on the X-axis of the bar graph and Y-axis of the box plot in centimeters. (**h**) Proportions of rice reactions among *Xoo* isolates show highly resistant, moderately resistant, moderately susceptible, and highly susceptible categories. (**i**) Lesion lengths of standard resistant and susceptible rice varieties, bar = 1 cm.

Most rice accessions showed moderate resistance, as evidenced by a slight left skew distribution of the bar graphs. However, two accessions, namely ARDA-174 and ARDA 262, demonstrated broad-spectrum resistance to all *Xoo* isolates. The proportion of resistant to susceptible reactions were calculated and are shown in Figure 1h. Notably, the symptoms observed between the standard checks for resistance and susceptibility were significantly distinguishable (Figure 1i). The results showed a range of highly resistant reactions (HR: 0–5 cm), from 15.5% to 31%, while the range of susceptible reactions (moderately susceptible (MS) and highly susceptible (HS): ≥10) varied from 26% to 46.5% (Figure 1).

3.2. Genotypic Variation and Population Structure in Rice Panel

Imputation and error correction of the GBS data were performed. The VCF file of 4,086,620 SNPs from 200 Thai rice accessions was subjected to quality control procedures using Tassel version 5. SNPs with MAF less than 0.05% and a coverage of less than 75% were filtered out (Figure 2). These results suggested that imputation was an essential step in improving the quality of GBS data, and the use of Beagle 5.4 can be an effective

method for imputing missing genotypic data. Additionally, heterozygosity alleles and indel sequences were removed before imputation to reduce the missing data rate [46]. The GWAS was conducted using 25,338 high-quality SNPs, which were distributed evenly across the 12 chromosomes. Chromosome 1 had the largest number of SNP loci, while chromosome 9 had the smallest. This result was consistent with the research of Lu et al. [37], who conducted GWAS of 421 rice accessions derived from the 3000 rice genome project (3KRGP), and the results showed that the highest and lowest number of SNPs were identified on chromosomes 1 and 9, respectively.

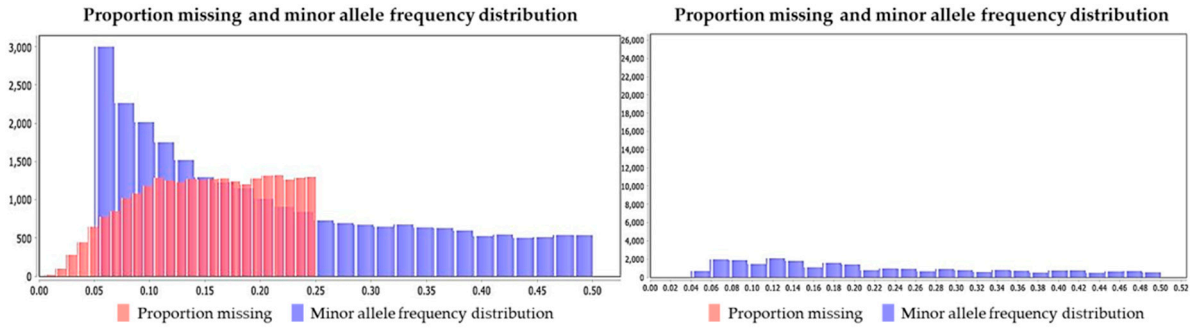

Figure 2. Imputation and error correction in GBS data. The left panel shows the original dataset without imputation, where a high rate of missing data is represented by the red bar and a significant decrease in the minor allele frequency is represented by the purple bar. The right panel shows the same dataset after imputation using Beagle 5.4. With imputation, missing data was reduced, leading to an increase in the number of genotyped sites and a more uniform distribution of minor allele frequency. This correction helps to reduce the bias in downstream analyses and improve the accuracy of genetic diversity estimates.

The analysis of 25,338 high-quality SNPs revealed two subpopulations in the rice panel, as supported by the estimation of the K value from the DeltaK plot generated by STRUCTURE 2.3.4 (Figure 3). A DeltaK plot identifies possible subpopulations in a dataset by detecting the peak in log-likelihood values at different values of K. In this study, the DeltaK plot revealed that the highest peaks were observed at K = 2 and a second peak at K = 6 (Figure 3), suggesting the presence of two or six distinct clusters within the 200 Thai rice accession panel. These results were consistent with a previous investigation [87,88], which classified the indica group into two major subpopulations (K = 2), but which reported four subgroups within those two subpopulations (K = 4).

The majority (86.8%) of pairwise kinship coefficients were found to be within the range of 0 to 0.5, while a minority (approximately 13.19%) of kinship values exceeded 0.5 (Figure 4a). These results indicated low levels of relatedness within the rice indica panel [87]. The eigenvalues associated with each PC represent the amount of variation explained by that component. In Figure 4b, we presented the eigenvalue plot obtained from the PCA. The plot showed a clear elbow point at PC = 2, indicating that the first two PCs capture most of the variance in the data. This suggested that the data can be effectively summarized and analyzed using only the first two PCs. Therefore, we retained the first two PCs for further analysis. A PCA was conducted to explore the genetic diversity and structure within the panel of rice accessions. The first two PCs (PC1 and PC2) were visualized to reveal the genetic variation among the accessions. However, the PCA plot indicated an ambiguous subpopulation structure (Figure 4c). PC1 and PC2 jointly accounted for roughly 14% of the total genetic variation, with PC1 and PC2 accounting for 8.5% and 5.5%, respectively. The two principal components captured some of the genetic variation among the accessions, but other factors may contribute to the overall genetic diversity within the panel. The findings demonstrated that genetic differentiation among the six geographic populations was not substantial, which may be attributed to genetic exchange between populations [88]. Based on the geographical origin of the rice accessions, the 200 rice accessions were separated along PC1, with accessions from central and eastern

regions exhibiting negative values, while those from northeastern and northern regions had positive values. These findings agree with previous reports suggesting that indica rice varieties tend to cluster according to their geographical location [42,88,89].

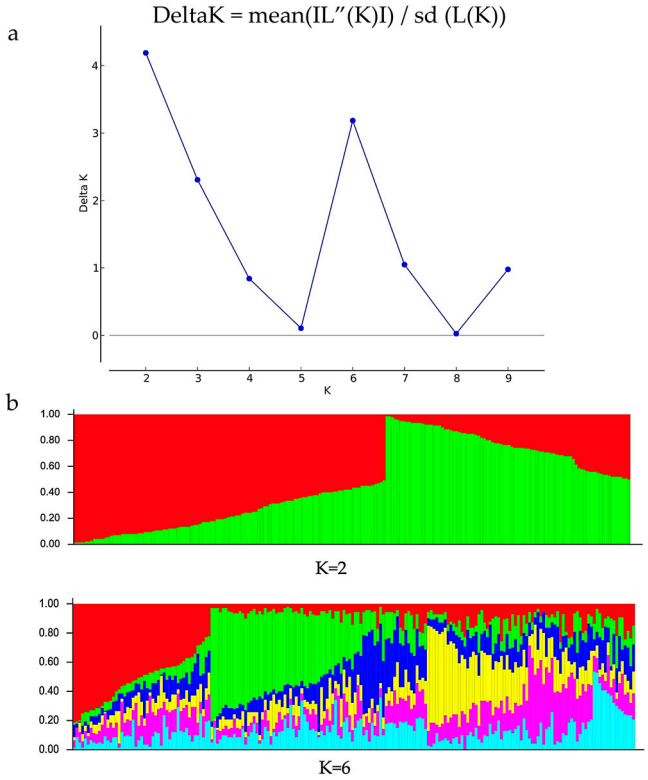

Figure 3. The population structure of 200 rice accessions. (**a**) A plot of DeltaK values used to classify subpopulations within the rice panel. The plot shows the highest peak at K = 2, indicating the presence of two subpopulations. (**b**) The estimated subpopulation components determined by the STRUCTURE 2.3.4 procedure, which were distributed based on the DeltaK method used to determine the most probable number of subpopulations at K = 2 and K = 6. The plot confirms the presence of two subpopulations in the rice panel, as supported by the DeltaK method. Different colors represent different populations of rice accessions.

To confirm the sub-species of rice, a panel of rice samples was utilized to perform SNP recall for 31,300 loci combined to 10 representative SNPs selected for the *japonica* rice group. Figure 4d shows a distinct separation between the indica and ten *japonica* rice groups. Out of the 200 rice samples analyzed, 18 were categorized as belonging to the admixture group, while only one sample was identified as being of the *japonica* group. The genetic variation in the sample population is geographically dependent, with the majority of the samples from the north and northeast located on the positive side of PC2, while the rest were located on the negative side of PC2. This suggests that most of the samples in the study are *indica* rice and admixture groups [42].

The current investigation estimated LD decay by evaluating 25,338 SNPs across a genomic region spanning 12 chromosomes. The findings revealed that the maximum r^2 value for the entire genome drops by half when the distance between SNPs reaches 118,257 base pairs, as depicted in Figure 5. The shortest LD decay distance was 72 kb, observed on chromosome 11, while the longest distance of LD decay is 225.8 kb, identified on chromosome 10, as indicated. These outcomes corresponded to prior studies, which suggested that the average LD decay distance for Indica rice is approximately 100 kb [18,70,87,90]. Since most marker distances are below 100 kb, it is expected that the study would have reasonable power to identify common large-effect variations in the *Indica* panel.

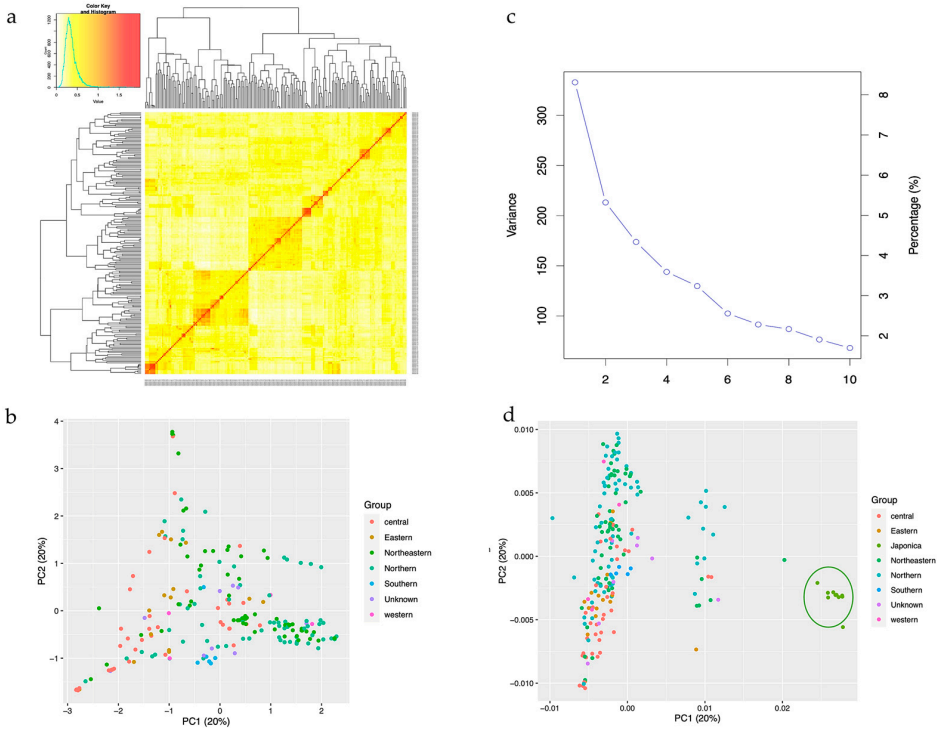

Figure 4. Kinship matrix and PCA of 200 rice accessions. (**a**) Heatmap of the kinship matrix, which represents the degree of relatedness between each pair of rice accessions in the dataset. Darker colors indicate greater relatedness. (**b**) PCA eigenvalue plot showing the number of principal components versus the percentage of variance explained by each component. The plot indicates that the first two principal components account for the largest amount of variance in the dataset. (**c**) PCA of 200 rice accessions based on the first two principal components. Each data point represents a single rice accession, and the different colors indicate different groups within the dataset base on geographical regions. (**d**) PCA of the rice panel combined into 10 representative *japonica* rice groups (green circle) to confirm the groups of sample tests. Two hundred rice accessions used in this study are significantly separated from representative japonica rice which confirm that most of them are *indica* rice with 2 subgroups in the population.

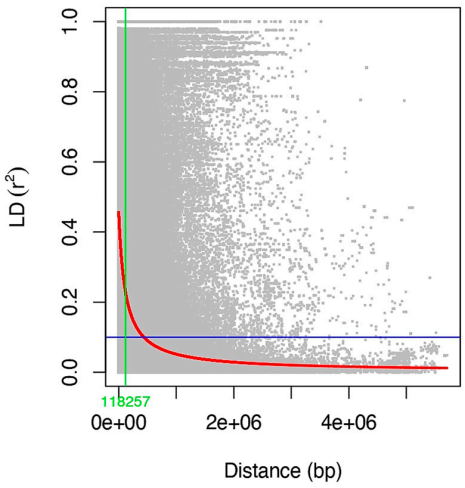

Figure 5. The decay of LD by distance in the panel of 200 Thai rice accessions. The LD decay plots illustrate the relationship between the distance (in base pairs) separating SNPs and the corresponding LD value (r^2). Red line represents the genome-wide LD decay. The LD decay distance and half LD values are indicated by the vertical (**green**) and horizontal lines (**dark blue**), respectively. LD decay = 118 kb.

3.3. GWAS Analysis of Bacterial Leaf Blight Resistance Traits

Various statistical models were used to identify SNPs associated with BLB disease caused by *Xoo* in rice. The results showed statistical significance at a threshold of 1.97×10^{-6} and a false discovery rate (FDR) cutoff at 0.05. The MLM model, which considers fixed and random effects, identified four SNPs significantly associated with the XORE1-1 isolate on chromosome 2. Meanwhile, the SUPER method designed to control false positives detected 15 SNPs linked to SK2-3-F, XORE1-1, and 59XOCRCS7-3 isolates on chromosomes 1, 2, 4, and 7. Similarly, multi-locus models such as FarmCPU and BLINK identified 11 and 10 significant SNPs linked to SK2-3-F and 59XOCRCS7-3 isolates, respectively (Table 1). Notably, all models discovered significant SNPs associated with only three *Xoo* isolates (SK2-3-F, XORE1-1, and 59XOCRCS7-3), emphasizing the significance of isolates-specific effects in identifying SNP-trait associations. Overall, the findings contributed insights into the genetic basis of BLB resistance in rice and demonstrate the usefulness of different statistical models in SNP-trait association studies.

Table 1. The significant SNPs of 200 Thai rice accessions against 7 *Xoo* isolates using four statistical models in GWAS.

Method/Traits	SNP	Chr.	Allele	Pos	*p*-Value	MAF	H&B. *p*-Value	q-Value
SUPER.Xoo1	S1_27143044	1	A/C	27,143,044	1.97×10^{-6}	0.26531	0.00919	5.71
SUPER.Xoo1	S1_27155997	1	A/G	27,155,997	7.17×10^{-7}	0.26020	0.00454	6.14
SUPER.Xoo1	S7_3529421	7	A/G	3,529,421	4.70×10^{-7}	0.15816	0.00397	6.33
SUPER.Xoo1	S7_3537960	7	T/C	3,537,960	4.70×10^{-7}	0.15816	0.00397	6.33
SUPER.Xoo1	S7_18664455	7	A/G	18,664,455	4.47×10^{-7}	0.21939	0.00397	6.35
FarmCPU.Xoo1	S1_27155997	1	A/G	27,155,997	3.52×10^{-8}	0.26020	0.00083	7.45
FarmCPU.Xoo1	S2_935021	2	A/T	935,021	1.34×10^{-7}	0.32653	0.00113	6.87
FarmCPU.Xoo1	S5_18901969	5	C/A	18,901,969	1.64×10^{-6}	0.13776	0.01036	5.79
FarmCPU.Xoo1	S7_3529421	7	A/G	3,529,421	6.57×10^{-8}	0.15816	0.00083	7.18
FarmCPU.Xoo1	S8_6216184	8	A/G	6,216,184	4.19×10^{-6}	0.46429	0.02126	5.38
FarmCPU.Xoo1	S4_28336644	4	C/T	28,336,644	1.22×10^{-5}	0.18878	0.05167	4.91
FarmCPU.Xoo1	S2_23903211	2	G/T	23,903,211	1.44×10^{-5}	0.10459	0.05223	4.84
BLINK.Xoo1	S2_935021	2	A/T	935,021	5.78×10^{-8}	0.32653	0.00073	7.24
BLINK.Xoo1	S7_3529421	7	A/G	3,529,421	6.07×10^{-10}	0.15816	0.00002	9.22
BLINK.Xoo1	S7_16039418	7	A/T	16,039,418	7.35×10^{-7}	0.29592	0.00465	6.13
BLINK.Xoo1	S8_6216184	8	A/G	6,216,184	2.57×10^{-7}	0.46429	0.00217	6.59
MLM.Xoo2	S2_25400454	2	G/A	25,400,454	1.09×10^{-6}	0.43333	0.01181	5.96
MLM.Xoo2	S2_25408622	2	G/A	25,408,622	1.77×10^{-6}	0.43846	0.01181	5.75
MLM.Xoo2	S2_25413622	2	G/A	25,413,622	1.77×10^{-6}	0.43846	0.01181	5.75
MLM.Xoo2	S2_25400408	2	T/C	25,400,408	2.80×10^{-6}	0.43590	0.01181	5.55
MLM.Xoo2	S2_25400480	2	A/C	25,400,480	2.80×10^{-6}	0.43590	0.01181	5.55
MLM.Xoo2	S2_25400530	2	A/C	25,400,530	2.80×10^{-6}	0.43590	0.01181	5.55
MLM.Xoo2	S2_25400580	2	C/T	25,400,580	6.24×10^{-6}	0.43077	0.02260	5.20
SUPER.Xoo2	S2_25400454	2	G/A	25,400,454	4.77×10^{-6}	0.43333	0.03765	5.32
SUPER.Xoo2	S2_25408622	2	G/A	25,408,622	6.73×10^{-6}	0.43846	0.03765	5.17
SUPER.Xoo2	S2_25413622	2	G/A	25,413,622	6.73×10^{-6}	0.43846	0.03765	5.17
SUPER.Xoo2	S2_25400408	2	T/C	25,400,408	8.92×10^{-6}	0.43590	0.03765	5.05
SUPER.Xoo2	S2_25400480	2	A/C	25,400,480	8.92×10^{-6}	0.43590	0.03765	5.05
SUPER.Xoo2	S2_25400530	2	A/C	25,400,530	8.92×10^{-6}	0.43590	0.03765	5.05
SUPER.Xoo2	S4_31500482	4	C/G	31,500,482	1.43×10^{-5}	0.45385	0.04039	4.84
SUPER.Xoo2	S4_31505231	4	T/G	31,505,231	1.43×10^{-5}	0.45385	0.04039	4.84
SUPER.Xoo2	S4_23423399	4	G/A	23,423,399	1.43×10^{-5}	0.46667	0.04039	4.84
BLINK.Xoo2	S2_25400454	2	G/A	25,400,454	3.96×10^{-12}	0.43333	0.00000	11.40
BLINK.Xoo2	S4_31770194	4	C/G	31,770,194	7.04×10^{-8}	0.46410	0.00089	7.15
SUPER.Xoo5	S1_27155997	1	A/G	27,155,997	9.82×10^{-8}	0.36592	0.00249	7.01
SUPER.Xoo5	S1_27143044	1	A/C	27,143,044	2.55×10^{-7}	0.36313	0.00322	6.59

Table 1. *Cont.*

Method/Traits	SNP	Chr.	Allele	Pos	*p*-Value	MAF	H&B. *p*-Value	q-Value
FarmCPU.Xoo5	S12_15567690	12	G/A	15,567,690	3.45×10^{-9}	0.39385	0.00009	8.46
FarmCPU.Xoo5	S1_22492602	1	G/A	22,492,602	6.15×10^{-8}	0.45810	0.00078	7.21
FarmCPU.Xoo5	S3_438061	3	A/G	438,061	6.39×10^{-7}	0.43017	0.00540	6.19
FarmCPU.Xoo5	S5_2313550	5	G/T	2,313,550	2.61×10^{-6}	0.39944	0.01656	5.58
FarmCPU.Xoo5	S6_1362133	6	C/T	1,362,133	3.53×10^{-6}	0.37989	0.01789	5.45
FarmCPU.Xoo5	S12_18758444	12	T/C	18,758,444	8.17×10^{-6}	0.39944	0.03042	5.09
FarmCPU.Xoo5	S1_30147885	1	T/C	30,147,885	8.40×10^{-6}	0.37709	0.03042	5.08
BLINK.Xoo5	S1_27155997	1	A/G	27,155,997	4.55×10^{-11}	0.36592	0.00000	10.30
BLINK.Xoo5	S8_26134412	8	C/A	26,134,412	7.65×10^{-8}	0.42737	0.00097	7.12
BLINK.Xoo5	S2_25453856	2	A/C	25,453,856	5.97×10^{-7}	0.44413	0.00504	6.22
BLINK.Xoo5	S2_18882825	2	A/G	18,882,825	5.00×10^{-6}	0.45810	0.03167	5.30

Based on the Q-Q plots, the MLM, FarmCPU, and BLINK models exhibited a linear pattern with a distinct upward deviation in the tail region, suggesting effective control of both false positives and false negatives. However, MLM appeared to have a limited ability to detect the associated SNPs. On the other hand, the SUPER model displayed a non-normal distribution on the lower left corner, which might be indicative of a ceiling effect or right-skewness in the data. Through GWAS, we identified significant SNPs in three *Xoo* isolates using both SUPER and BLINK analysis. However, we found that the Q-Q plots of FarmCPU and BLINK provided a better fit to the data, as the plotted points followed a straight line closely along the identity line with a sharply deviated tail (Figures 6–8). This distinctive pattern was more apparent than that observed in other models. The candidate genes containing SNPs significantly associated with bacterial leaf blight resistance trait were identified by the gene annotation [81] and the Oryzabase database [82]. In summary, GWAS revealed 32 significant SNPs associated with BLB resistance, which were located on chromosomes 1–8 and 12. The SNP with the highest level of significance was identified at S7_3529421 on chromosome 7, with a *p*-value of 6.07×10^{-10}. This result was obtained using the BLINK method against SK2-3-F, and the details are provided in Table 1.

The Manhattan plots showed a set of highly significant SNPs that were associated with the SK2-3-F isolate, as identified by four distinct methods (Figure 6). These results were generated using the SUPER model with the genomic coordinates S1_2714304, S1_2715599, S7_3529421, S7_3537960, and S7_1866445 on chromosomes 1 and 7. The corresponding $-\log_{10}$ (*p*-values) for these SNPs ranged from 5.70 to 6.34. Notably, the FarmCPU model generated the largest number of significant SNPs, spanning across chromosomes 1, 2, 4, 5, 7, and 8, with $-\log_{10}$ (*p*-values) ranging from 4.84 to 7.45 at FDR threshold of 0.05. The BLINK model identified a limited number of significant SNPs associated with resistance to the SK2-3-F isolate. Specifically, this model identified four significant SNPs located on chromosomes 2, 7, and 8. However, it demonstrated the highest peak of significance among all methods, with a $-\log_{10}$ (*p*-value) of 9.22.

Manhattan plot showed the significant SNPs associated with BLB inoculated with XORE1-1 isolate (Figure 7). The findings revealed that only two out of the four statistical models generated significant SNPs associated with the BLB resistance gene. The MLM models produced the highest number of significant SNPs, with most located on chromosome 2. In contrast, the BLINK models only identified two significant SNPs on chromosomes 2 and 4. In addition, the Manhattan plot showed the significant SNPs linked to BLB resistance against the 59XOCRCS7-3 isolate (Figure 8). The results generated from the SUPER, FarmCPU, and BLINK models showed highly significant SNPs along chromosomes 1, 2, 3, 5, 6, 8, and 12. These findings indicated that different statistical models can reveal significant SNPs associated with the resistance to BLB caused by different *Xoo* strains. Therefore, the use of multiple models can provide a more comprehensive understanding of the genetic factors underlying BLB resistance in rice. Notably, both the BLINK and SUPER models were able to detect significant SNPs in all three *Xoo* isolates (SK2-3-F, XORE1-1, and 59XOCRCS7-3).

Among the four models, BLINK method performed the best in controlling false positive and false negative errors, as evidenced by an analysis of the Q-Q plots, which is a common approach to assess model performance. The Q-Q plot generated by the SUPER model showed an inflated upward line, indicating a high number of false positives. In contrast, the BLINK model produced a straight line with a slight tail, suggesting that it effectively controlled both false positives and false negatives.

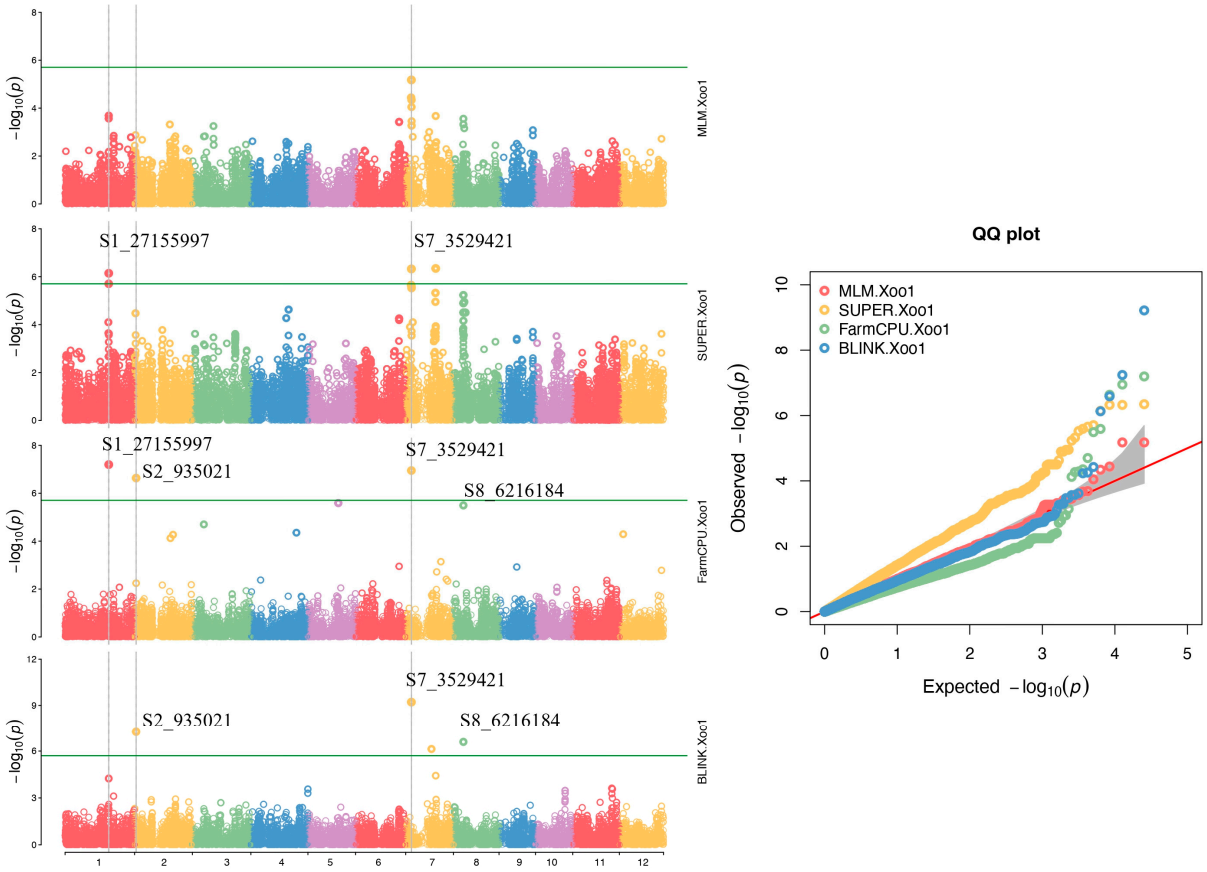

Figure 6. Manhattan and Q-Q plots for four different methods of GWAS analysis using a panel of 200 rice accessions that are associated with SK2-3-F isolate. The significant 11 SNPs were found to be located on chromosomes 1, 2, 4, 5, 7, and 8. Different colors represent different rice chromosomes from chromosome 1 to chromosome 12.

Out of the 581 genes that were located surrounding the significant SNPs, 179 genes were found within the ±100 kb region surrounding the SNP of interest. Among these 179 genes, 49 were selected as candidate genes based on their known functions in plant defense mechanisms. In addition, they were selected for their proximity to significant SNPs in the gene region. This study focuses on genes associated with several plant defense functions. These genes include the serine/threonine protein kinase-related domain-containing protein, the WRKY transcription factor, the cell cycle-associated protein kinase, the Leucine-rich repeat, the NBS-LRR disease resistance protein, disease resistance protein RGA2, and the MYB transcription factor.

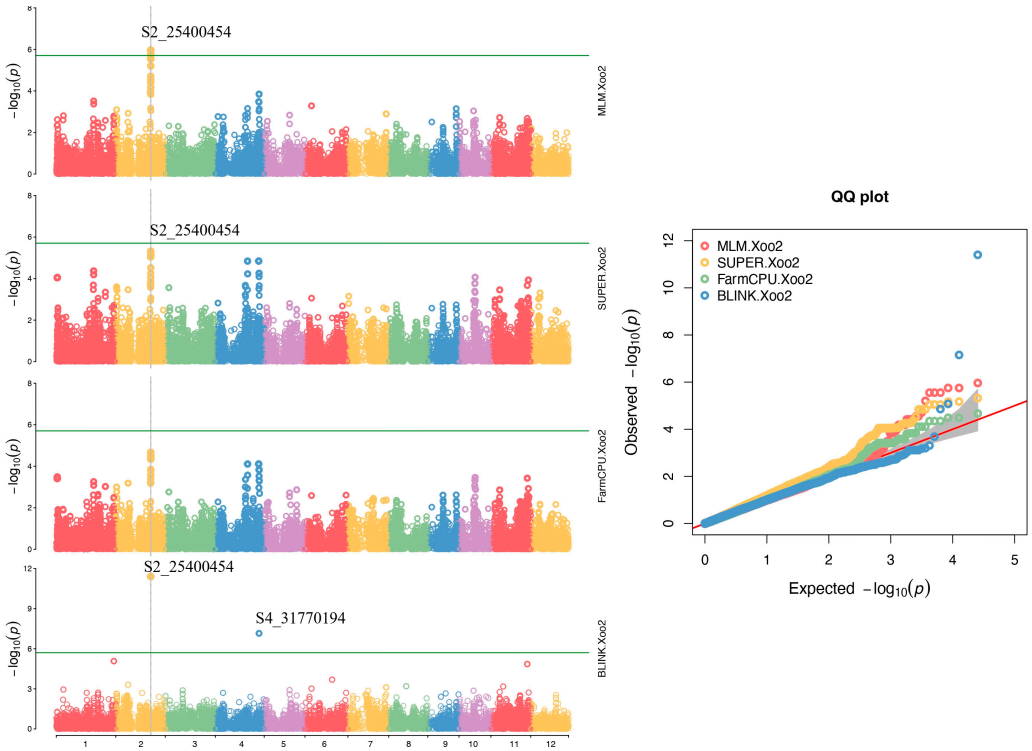

Figure 7. Manhattan and Q-Q plots for four different methods of GWAS analysis using a panel of 200 rice accessions that are associated with XORE1-1 isolate. Different colors represent different rice chromosomes from chromosome 1 to chromosome 12.

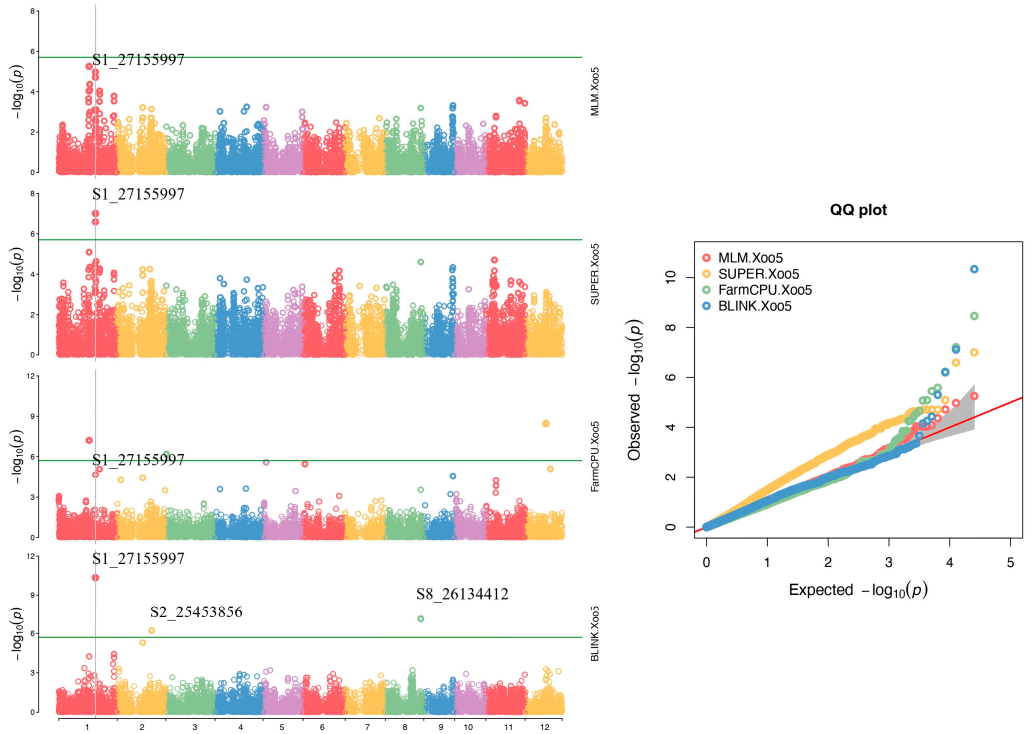

Figure 8. Manhattan and Q-Q plots for four different methods of GWAS analysis using a panel of 200 rice accessions that are associated with 59XOCRCS7-3 isolate. Different colors represent different rice chromosomes from chromosome 1 to chromosome 12.

3.4. Analysis of Haplotypes and Linkage of Candidate Genes

In order to perform a haplotype analysis to detect significant differences between the resistance and susceptibility, 16 candidate SNP loci were selected based on the gene annotation as a plant defense mechanism located on chromosome 2, 4, and 8. One of the significant SNP loci, S2_25453856, had two candidate genes surrounding its region (±100 kb), namely *LOC_Os02g42310.1* and *LOC_Os02g42314.2*, which were located within the LD decay. These genes started from the position 25442875 to 25450093 and 25450313 to 25454074, respectively.

The *LOC_Os02g42310.1* gene is composed of 6 SNPs with 5 haplotypes, and the function of the gene is predicted to be a serine carboxypeptidase II-like protein. Serine carboxypeptidases are a group of enzymes that are involved in cleaving C-terminal amino acids from peptide substrates. These enzymes have been shown to play important roles in various biological processes in plants, including cell wall metabolism and defense responses against pathogens. One specific serine carboxypeptidase gene has been identified as being significantly associated with sheath blight resistance [91]. Six SNPs identified within this gene have the potential to affect the activity or stability of the protein, which could, in turn, impact the plant's resistance or susceptibility to the disease. These findings suggest that the identification and characterization of specific genetic markers, such as SNPs in this gene, could be a valuable tool for developing more effective strategies for breeding rice varieties with enhanced disease resistance. The other gene locus, *LOC_Os02g42314.2*, is composed of three SNPs that can produce seven different haplotypes (Hap1-Hap7). The protein encoded by this gene is a Ubiquitin-conjugating enzyme/RWD-like domain-containing protein.

Moreover, two gene loci in the same linkage disequilibrium decay with S8_6216184 on chromosome 8 were identified, *LOC_Os08g10440.1*, which encodes a CC-NBS-LRR protein known to confer *Pyricularia oryzae* resistance-*33(t)* in rice [92]. This gene, located on chromosome 8, consists of seven different genetic patterns. Interestingly, haplotypes Hap7 and Hap8 of this gene were found to confer a moderated resistance (MR) phenotype, where the lesion length (LL) was greater than or equal to 10 cm. *LOC_Os08g10560.1* is responsible for encoding the nuclear factor Y C subunit 5 (NF-YC5) in rice, which plays a crucial role in the processes of DNA replication and repair. The gene ontology annotation suggested that the *OsNF-Y* genes are co-expressed with stress response, seed storage reserve accumulation, and plant development genes [93]. NF-Y is a heterotrimeric transcription factor that regulates a variety of biological processes in plants, including defense mechanisms against both biotic and abiotic stresses. In particular, the NF-YC5 subunit of the NF-Y complex is known to regulate the plant defense response in various species, such as *Arabidopsis thaliana* and Cassava (*Manihot esculenta*). The expression of *NF-YC5* is induced upon infection by a range of plant pathogens, including bacteria, fungi, and viruses. [94,95]. The gene comprises 13 SNPs, resulting in 8 haplotypes. Among these, Hap1, Hap4, and Hap9 demonstrated a moderated resistance reaction, as evidenced by a lesion length (LL) of $5 \geq 10$ cm.

LOC_Os04g52900 was associated with SNP S4_31500482 and S4_31505231, which encode for ATP binding cassette (ABC) transporters, a well-studied and highly conserved family of transporter proteins. Previous research has demonstrated that certain ABC transporters play a role in conferring resistance against bacterial pathogens, such as the HpaABC transporter which is involved in bacterial blight resistance [92,96]. In addition, the genetic locus *LOC_Os04g52900* contains several SNP loci, resulting in a genetic variant with three distinct haplotypes (Figure 9). Of these haplotypes, Hap2 exhibited a significantly enhanced level of resistance, as evidenced by a lesion length of less than or equal to 5 cm. Additionally, our investigation suggested that genetic variation within the *LOC_Os04g53350* gene, annotated as a conserved hypothetical protein, may also play a crucial role in providing disease resistance in plants. This gene demonstrated a highly resistant reaction (LL = 4 cm) against the SK2-3-F pathogen and is located on chromosome 4 within a gene region containing three genes. Our findings highlight the potential of these genes as targets for improving plant disease resistance through breeding programs.

a

LOC_Os04g53350.1	S4_31769932	S4_31770175	S4_31770194	S4_31770346	S4_31770616	S4_31770769	S4_31770901	LL(cm.)	Frequency
Hap1	C	C	C	C	G	A	C	12	24
Hap2	C	C	C	T	G	A	A	13	15
Hap3	C	C	C	T	G	A	C	11	120
Hap4	C	C	G	C	A	G	C	4	14
Hap5	C	T	C	T	G	A	C	10	13
Hap6	T	C	C	T	G	A	C	13	14

b

LOC_Os04g52900.1	S4_31505186	S4_31505197	S4_31505231	LL(cm.)	Frequency
Hap1	A	A	T	12	30
Hap2	A	T	G	5	18
Hap3	A	T	T	11	91
Hap4	G	T	T	11	61

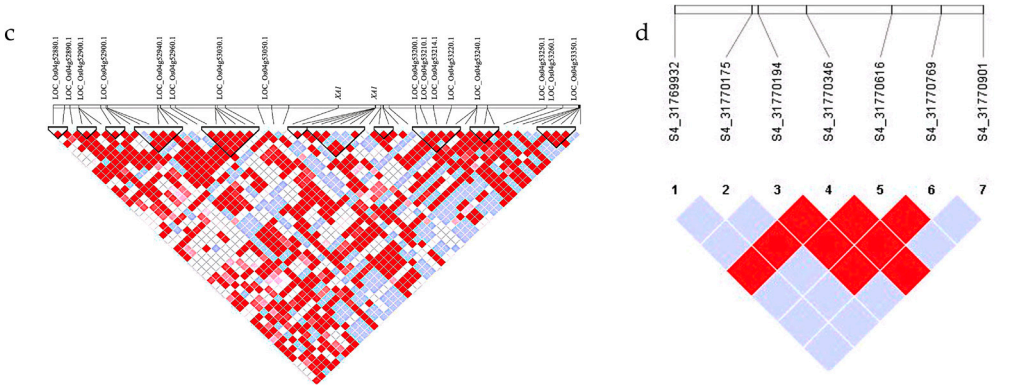

Figure 9. The identification of alternative haplotypes in a chromosomal region on chromosome 4 through GWAS analysis and LD heatmaps. Haplotype analysis for two MSU genes, *LOC_Os04g53350.1* and *LOC_O4g52900.1*, was shown in (**a**,**b**), while (**c**) displays the LD heatmap for a specific region (31.3–32.7 Mb) indicating positions of well-known genes, including *Xa1* and 15 other genes within ±100 kb of significant SNPs. White and gray areas suggests low levels of LD within this region, while red areas suggests high levels of LD within the region. (**d**) shows the significant LD block within *LOC_Os04g53350.1*.

4. Discussion

BLB, caused by *Xoo* bacteria, have reported yield losses ranging from 80–100% in Asia [4,5,12]. GBS is an advanced technique that has emerged as a valuable tool for identifying genetic variants linked to BLB resistance in crops through GWAS. Despite the challenges associated with GBS, it is still a cost-effective, high-throughput, and minimally invasive method, making it suitable for large-scale studies of diverse crop populations. GBS can capture rare and low-frequency genetic variants that may not be detected by other genotyping methods, providing a more comprehensive view of the genetic architecture underlying BLB resistance [29,60,97].

In this study, we explored the genetic basis of BLB resistance in Thai local rice. To do this, we conducted a GWAS on a panel of 200 *indica* rice accessions collected from six different regions in Thailand. We can classify the rice panel into two subpopulations, which is consistent with the previous report that used *indica* rice for genetic analysis [42,87]. We assessed BLB resistance by infecting all 200 rice accessions with seven Thai *Xoo* isolates. The four methods of statistical analysis were used for GWAS analysis. The Q-Q plots revealed a non-uniform distribution of *p*-values in the SUPER models of all empirical traits, consistent with previous research, indicating their unsuitability for trait-associated mapping due to the possibility of false marker-trait associations [72]. To correct population structure and family relatedness, complex models such as MLM, FarmCPU, and BLINK have been suggested [72,76–79]. MLM and FarmCPU Q-Q plots displayed a straight line with slightly deviated tails, indicating a decrease in false positives but an increase in false negatives due to overfitting, which is in line with other studies. However, the Q-Q plot of the multi-locus model, BLINK, indicated that it controls both false positives and false negatives for all empirical traits in both crops, as shown by a straight line with a sharp deviated tail. As a result of our GWAS analysis, in this study, we have identified a novel region that con-

tains SNPs associated with three *Xoo* isolates. These SNPs were found to be located on chromosomes 1–8 and 12 but not on chromosomes 9–11. It is worth noting that previous reports have mostly identified the BLB resistance gene on chromosome 11 [17,18,36,98]. However, our results indicate that SNPs on chromosome 11 were only detected at a low level of significance, with a *q*-value threshold of 5.71 and an FDR of 0.01 for the three *Xoo* isolates. In addition, our investigation revealed 15 MSU genes flanking the well-known resistance gene *Xa1*, which were also located within the same LD block. Previous studies have reported that *Xa2*, *Xa12*, *Xa14*, *Xa25*, *Xa31(t)*, and *Xa38* were localized to the region near *Xa1* and conferred race-specific resistance against *Xoo* [99]. Our finding indicated that the two significant SNPs, S4_31500482 and S4_31505231, consist of 26 MSU genes situated around them. Out of 15 MSU genes, *LOC_Os04g52890*, *LOC_Os04g52870*, and *LOC_Os04g52900* have been identified as genes for a blast resistance gene *Pikahei-1(t)* mapped to chromosome 4 [17]. We also identified a gene, *LOC_Os04g53160*, which is located within the same LD decay as *Xa1* and encodes BED-NLBs. It is a DNA-binding domain present in many transposases and chromatin-boundary element-binding proteins and acts as an integrated decoy domain. This gene was previously reported as an *R* gene analog named *RGAg* (*Os04g53160*) [100]. Additionally, we found that *LOC_Os04g53160* was located near *LOC_Os04g53170* and *LOC_Os04g53180*, which were identified as a putative *Xa1* gene and a bromodomain-containing protein, respectively [100]. These findings provide valuable insights into the potential roles of these genes in BLB resistance and suggest new avenues for further exploration in developing BLB-resistant rice varieties. Furthermore, we found that this region contains *LOC_Os04g53210* and *LOC_Os04g53214*, which encode for a glycolate oxidase (GLO) enzyme. Previous studies have reported that a reduction in the expression of GLO can lead to an increase in the production of reactive oxygen species (ROS), which can ultimately enhance the plant's ability to defend against various stresses [101]. In addition, this study has identified seven NBS-LRR coding genes (*LOC_Os02g02660.1*, *LOC_Os02g02670.1*, *LOC_Os02g39660.1*, *LOC_Os04g53050.1*, *LOC_Os06g03500.1*, *LOC_Os07g31500.1*, and *LOC_Os08g10440.1*) located on chromosomes 2, 4, 6, 7, and 8, which may be crucial for the recognition and constant surveillance of invading elicitors and play a significant role in the plant defense mechanism [72,86,102]. Among these genes, *LOC_Os04g53050* was identified as an *Xa38* gene associated with S4_31500482 and S4_31505231 [4]. In a previous study, Dilla-Ermita et al. [17] identified eight SNPs located on chromosome 8 that were adjacent to and flanked the *xa13* gene, including S8_26013849, S8_26044514, S8_26044528, S8_27319974, S8_27520607, S8_27536844, S8_27641374, and S8_27663433. Interestingly, these SNPs also flank our S8_26134412, which is located on the *LOC_Os08g41380* gene that encodes a P-type PPR protein and has been implicated in responding to biotic and abiotic stresses [103]. Shu et al. [25] also reported the identification of S12_175769641 and S8_27163888, which are located near the clone *R* gene *xa25* and *Xa23*, respectively. Interestingly, these two significant SNPs are located in close proximity to our own SNP results.

The SNPs with the highest level of significance associated with BLB resistance are located on chromosome 2 (Table 1) with a *p*-value of 3.69×10^{-12}. Among these, the candidate SNP S2_25400454 is located on non-coding regions [84], which are linked to several genes including *LOC_Os02g42110* (cell cycle-associated protein kinase), *LOC_Os02g42150* (wall-associated kinase, which positively regulates rice blast resistance), *LOC_Os02g42200* (BB-block binding subunit of TFIIIC domain containing protein), and *LOC_Os02g42310* (similar to serine carboxypeptidase II-like protein), all of which are in the same LD decay block ($\pm$ 100 kb from the significant SNP). Finally, our analysis of flashpoint regions revealed that the *LOC_Os04g53350* gene contains six haplotypes, among which Hap4 (CCGCAGC) showed a strong resistance reaction. This gene is located on chromosome 4 within a gene region that contains three other genes. Our findings suggest that *LOC_Os04g53350* may play a crucial role in providing disease resistance in plants and could be a promising target for improving plant disease resistance through breeding programs.

The identification of a candidate gene in this study highlights its potential role in regulating immunity to BLB resistance in rice. The results provide valuable insights into the underlying genetic mechanisms of BLB resistance, which can help to improve breeding programs. Further investigations into the specific roles of these genes in enhancing rice's resistance to this devastating disease are essential for the development of effective management strategies. These findings underscore the importance of continued research in this area to ensure the sustainability of rice production and food security worldwide.

5. Conclusions

The Genome-Wide Association Study (GWAS) approach was adopted to identify Single Nucleotide Polymorphisms (SNPs) associated with resistance to each of the seven *Xoo* isolates. To analyze the results, Manhattan and Q-Q plots were used, and it was found that the BLINK method outperformed other methods in terms of smoother test lines on the Q-Q plots. The study also identified candidate genes that could potentially confer resistance against *Xoo*. Two genes, *LOC_Os04g53350* and *LOC_Os04g52900*, were found to exhibit highly resistant action against *Xoo*, with lesion lengths of 4 and 5 cm (Hap4 and Hap2 of each gene), respectively. Additionally, ten candidate genes were analyzed for haplotypes. The results of this research could be used to develop more efficient and effective breeding programs aimed at enhancing the resistance of rice against *Xoo*, ultimately leading to improved crop productivity and food security.

Author Contributions: C.J. and S.P.; designed and supervised the research, P.V.; provided material and sequenced of the rice accessions resources C.D.; performed the experiments and analyzed the data, C.J. and C.D.; interpreted the results and writing original draft preparation, C.D., A.T. and G.V.; data curation, T.T.; funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This project is funded by National Research Council of Thailand (NRCT): project no. N41A640140 and NRCT-RTA/812/2563 and Kasetsart University Research and Development Institute (KURDI). C.D. received the Chalermprakiat 70 years of reign Ph.D. scholarship from the Agricultural Research Development Agency (ARDA), Thailand, grant number HRD6301011.

Data Availability Statement: The data presented in this study.

Acknowledgments: The authors would like to express our gratitude to the Thai Rice Department and the Ministry of Agriculture and Cooperatives for their support in providing the necessary greenhouse and germplasm for the experiments. We also extend our appreciation to their colleagues who provided suggestions and support for this research.

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

References

1. FAOSTAT. Food and Agriculture Organization of the United, Production/Yield Quantities of Rice in World + (Total). 2021. Available online: https://www.fao.org/faostat/en/#data/QCL/visualize (accessed on 26 February 2023).
2. Asibi, A.E.; Chai, Q.; Coulter, J.A. Rice blast: A disease with implications for global food security. *J. Agron.* **2019**, *9*, 451. [CrossRef]
3. Kumar, A.; Kumar, R.; Sengupta, D.; Das, S.N.; Pandey, M.K.; Bohra, A.; Sharma, N.K.; Sinha, P.; Sk, H.; Ghazi, I.A.; et al. Deployment of genetic and genomic tools toward gaining a better understanding of rice-*Xanthomonas oryzae* pv. *oryzae* interactions for development of durable bacterial blight resistant rice. *Front. Plant Sci.* **2020**, *4*, 1152. [CrossRef] [PubMed]
4. Fiyaz, R.A.; Shivani, D.; Chaithanya, K.; Mounika, K.; Chiranjeevi, M.; Laha, G.S.; Viraktamath, B.C.; Subba Rao, L.V.; Sundaram, R.M. Genetic improvement of rice for bacterial blight resistance: Present status and future prospects. *Rice Sci.* **2022**, *29*, 118–132. [CrossRef]
5. Savary, S.; Willocquet, L.; Pethybridge, S.J.; Esker, P.; McRoberts, N.; Nelson, A. The global burden of pathogens and pests on major food crops. *Nat. Ecol. Evol.* **2019**, *3*, 430–439. [CrossRef] [PubMed]
6. Ou, S.H. *Rice Diseases*, 2nd ed.; Commonwealth Mycological Institute: Great, Britain, 1985.
7. Shekhar, S.; Sinha, D.; Kumari, A. An overview of bacterial leaf blight disease of rice and different strategies for its management. *Int. J. Curr. Microbiol. App. Sci.* **2020**, *9*, 2250–2265. [CrossRef]
8. EPPO (European and Mediterranean Plant Protection Organization). EPPO Bulletin: *Xanthomonas oryzae*. Available online: https://doi.org/10.1111/j.1365-2338.2007.01162.x (accessed on 2 February 2022).

9. Doucoure, H.; Perez-Quintero, A.L.; Reshetnyak, G.; Tekete, C.; Auguy, F.; Thomas, E. Functional and genome sequence-driven characterization of tal effector gene repertoires reveals novel variants with altered specificities in closely related malian *Xanthomonas oryzae* pv. *oryzae* strains. *Front. Microbiol.* **2018**, *9*, 1657. [CrossRef]
10. Jiang, N.; Yan, J.; Liang, Y.; Shi, Y.; He, Z.; Wu, Y.; Zeng, Q.; Liu, X.; Peng, J. Resistance genes and their interactions with bacterial blight/leaf streak pathogens (*Xanthomonas oryzae*) in rice (*Oryza sativa* L.)—An updated review. *Rice* **2021**, *13*, 3. [CrossRef]
11. Naqvi, S.A.H. Bacterial leaf blight of rice: An overview of epidemiology and management with special reference to Indian sub-continent. *Pak. J. Agric. Sci.* **2019**, *32*, 359–380. [CrossRef]
12. Pradhan, S.K.; Nayak, D.K.; Mohanty, S.; Behera, L.; Barik, S.R.; Pandit, E.; Lenka, S.; Anandan, A. Pyramiding of three bacterial blight resistance genes for broad-spectrum resistance in deepwater rice variety, Jalmagna. *Rice* **2015**, *8*, 19. [CrossRef]
13. Gnanamanickam, S.S. Biological control of bacterial blight of rice. In *Biological Control of Rice Diseases*; Springer: Dordrecht, The Netherlands, 2009; pp. 67–68. [CrossRef]
14. Sundin, G.W.; Castiblanco, L.F.; Yuan, X.; Zeng, Q.; Yang, C.H. Bacterial disease man-agement: Challenges, experience, innovation and future prospects. *Mol. Plant Pathol.* **2016**, *17*, 1506–1518. [CrossRef]
15. Ullah, I.; Ali, H.; Mahmood, T.; Khan, M.N.; Haris, M.; Shah, H.; Mihoub, A.; Jamal, A.; Saeed, M.F.; Mancinelli, R.; et al. Pyramiding of four broad spectrum bacterial blight resistance genes in cross breeds of basmati rice. *Plants* **2023**, *12*, 46. [CrossRef] [PubMed]
16. Wang, S.; Liu, W.; Lu, D.; Lu, Z.; Wang, X.; Xue, J.; He, X. Distribution of bacterial blight resistance genes in the main cultivars and application of *Xa23* in rice breeding. *Front. Plant Sci.* **2020**, *11*, 555228. [CrossRef] [PubMed]
17. Dilla-Ermita, C.J.; Tandaya, E.; Juanillas, V.M.; Detras, J.; Lozada, D.N.; Dwiyanti, M.S.; Cruz, C.V.; Mbanjo, E.G.N.; Ardales, E.; Diaz, M.G.; et al. Genome-wide association analysis tracks bacterial leaf blight resistance loci in rice diverse germplasm. *Rice* **2017**, *10*, 8. [CrossRef] [PubMed]
18. Korinsak, S.; Darwell, C.T.; Wanchana, S.; Praphaisal, L.; Korinsak, S.; Thunnom, B.; Patarapuwadol, S.; Toojinda, T. Identification of bacterial blight resistance loci in rice (*Oryza sativa* L.) against diverse *Xoo* Thai strains by genome-wide association study. *Plants* **2021**, *10*, 518. [CrossRef] [PubMed]
19. Chukwu, S.C.; Rafii, M.Y.; Ramlee, S.I.; Ismail, S.I.; Hasan, M.M.; Oladasu, Y.A.; Magaji, U.G.; Akos, I.; Olalekan, K.K. Bacterial leaf blight resistance in rice: A review of conventional breeding to molecular approach. *Mol. Biol. Rep.* **2019**, *46*, 1519–1532. [CrossRef]
20. Kawahara, Y.; 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]
21. Chukwu, S.C.; Rafii, M.Y.; Ramlee, S.I.; Ismail, S.I.; Oladosu, Y.; Okporie, E.; Onyishi, G.; Utobo, E.; Ekwu, L.; Swaray, S.; et al. Marker-assisted selection and gene pyramiding for resistance to bacterial leaf blight disease of rice (*Oryza sativa* L.). *Biotechnol. Biotechnol. Equip.* **2019**, *33*, 440–455. [CrossRef]
22. Kim, S.M. Identification of novel recessive gene xa44(t) conferring resistance to bacterial blight races in rice by QTL linkage analysis using an SNP chip. *Appl. Genet.* **2018**, *131*, 2733–2743. [CrossRef]
23. Kim, S.M.; Reinke, R. A novel resistance gene for bacterial blight in rice, *Xa43(t)* identified by GWAS, confirmed by QTL mapping using a bi-parental population. *PLoS ONE* **2019**, *14*, e0211775. [CrossRef]
24. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **2011**, *6*, e19379. [CrossRef]
25. Shu, X.; Wang, A.; Jiang, B.; Jiang, Y.; Xiang, X.; Yi, X.; Li, X.; Deng, Q.; Wang, S.; Zhu, S.; et al. Genome-wide association study and transcriptome analysis discover new genes for bacterial leaf blight resistance in rice (*Oryza sativa* L.). *BMC Plant Biol.* **2021**, *21*, 255. [CrossRef] [PubMed]
26. Reyes, V.P.; Kitony, J.K.; Nishiuchi, S.; Makihara, D.; Doi, K. Utilization of genotyping-by-sequencing (GBS) for rice pre-breeding and improvement: A review. *Life* **2022**, *12*, 1752. [CrossRef] [PubMed]
27. Verdeprado, H.; Kretzschmar, T.; Begum, H.; Raghavan, C.; Joyce, P.; Lakshmanan, P.; Cobb, J.N.; Collard, B.C.Y. Association mapping in rice: Basic concepts and perspectives for molecular breeding. *Plant Prod. Sci.* **2018**, *21*, 159–176. [CrossRef]
28. Zeng, P.; Zhao, Y.; Qian, C.; Zhang, L.; Zhang, R.; Gou, J.; Liu, J.; Liu, L.; Chen, F. Statis-tical analysis for genome-wide association study. *J. Biomed. Res.* **2015**, *29*, 285–297. [CrossRef] [PubMed]
29. Ayalew, H.; Anderson, J.D.; Krom, N.; Tang, Y.; Butler, T.J.; Rawat, N.; Tiwari, V.; Ma, X.F. Genotyping-by-sequencing and genomic selection applications in hexaploid triticale. *G3 Genes | Genomes | Genet.* **2022**, *12*, jkab413. [CrossRef]
30. Begum, H.; Spindel, J.E.; Lalusin, A.; Borromeo, T.; Gregorio, G.; Hernandez, J.; Virk, P.; Collard, B.; McCouch, S.R. Genome-wide association mapping for yield and other agronomic traits in an elite breeding population of tropical rice (*Oryza sativa*). *PLoS ONE* **2015**, *10*, e0119873. [CrossRef]
31. Tekeu, H.; Ngonkeu, E.L.M.; Belanger, S.; Djocgoue, P.F.; Abed, A.; Torkamaneh, D.; Boyle, B.; Tsimi, P.M.; Tadesse, W. GWAS identifies an ortholog of the rice D11 gene as a candidate gene for grain size in an international collection of hexaploid wheat. *Sci. Rep.* **2021**, *11*, 19483. [CrossRef]
32. Bandillo, N.; Raghavan, C.; Muyco, P.A.; Sevilla, M.A.L.; Lobina, I.T.; Dilla-Ermita, C.J.; Tung, C.W.; McCouch, S.; Thomoson, M.; Mauleon, R.; et al. Multi-parent advanced generation inter-cross (MAGIC) populations in rice: Progress and potential for genetics research and breeding. *Rice* **2013**, *6*, 11. [CrossRef]

33. Du, M.; Xiong, M.; Chang, Y.; Liu, Z.; Wang, R.; Lin, X.; Zhou, Z.; Lu, M.; Liu, C.; Liu, E. Mining candidate genes and favorable haplotypes for flag leaf shape in rice (*Oryza sativa* L.) based on a genome-wide association study. *Agronomy* **2022**, *12*, 1814. [CrossRef]
34. Xiang, J.; Zhang, C.; Wang, N.; Liang, Z.; Zhenzhen, Z.; Liang, L.; Yuan, H.; Shi, Y. Genome-wide association study reveals candidate genes for root-related traits in rice. *Curr. Issues Mol. Biol.* **2022**, *44*, 4386–4405. [CrossRef]
35. Zhang, G.; Wang, R.; Ma, J.; Gao, H.; Deng, L.; Wang, N.; Wang, Y.; Zhang, J.; Li, K.; Zhang, W.; et al. Genome-wide association studies of yield-related traits in high-latitude japonica rice. *BMC Genom. Data.* **2021**, *22*, 39. [CrossRef] [PubMed]
36. Sompop, P.; Ruengchaijatuporn, N.; Sriswasdi, S.; Buaboocha, T.; Chadchawan, S.; Chaiwanon, J. Hyperspectral and genome-wide association analyses of leaf phosphorus status in local Thai indica rice. *PLoS ONE* **2022**, *17*, e0267304. [CrossRef]
37. Lu, J.; Li, Q.; Wang, C.; Wang, M.; Zeng, D.; Zhang, F.; Zhai, W.; Zhou, Y. Identification of quantitative trait loci associated with resistance to *Xanthomonas oryzae* pv. *oryzae* pathotypes prevalent in South China. *Crops J.* **2022**, *10*, 498–507. [CrossRef]
38. United States Department of Agriculture Foreign Agricultural Service, Grain: World Markets and Trade. 2022. Available online: https://usda.library.cornell.edu/concern/publications/zs25x844t?locale=en (accessed on 24 April 2023).
39. Principal Rice Exporting Countries Worldwide in 2022/2023. US Department of Agriculture. 2023. Available online: https://www.statista.com/statistics/255947/top-rice-exporting-countries-worldwide-2011/ (accessed on 24 April 2023).
40. Department of Foreign Trade. 2023. Available online: https://www.dft.go.th/th-th/dft-service-data-productgroup (accessed on 21 January 2023).
41. Chakhonkaenm, S.; Pitnjam, K.; Saisuk, W.; Ukoskit, K.; Muangprom, A. Genetic structure of Thai rice and rice accessions obtained from the International Rice Research Institute. *Rice* **2012**, *5*, 19. [CrossRef]
42. Vejchasarn, P.; Shearman, J.R.; Chaiprom, U.; Phansenee, Y.; Suthanthangjai, A.; Jairin, J.; Chamarerk, V.; Tulyananda, T.; Amornbunchornvej, C. Population structure of nation-wide rice in Thailand. *Rice* **2021**, *14*, 88. [CrossRef]
43. Fongfon, S.; Pusadee, T.; Prom-u-thai, C.; Rerkasem, B.; Jamjod, S. Diversity of purple rice (*Oryza sativa* L.) landraces in northern Thailand. *Agronomy* **2021**, *11*, 2029. [CrossRef]
44. Luangmanee, J.; Banterng, P.; Wongcharoen, A. Blast disease of black glutinous rice germplasms under inoculation at seeding and tillering stages. *Turk. J. Field Crops* **2016**, *21*, 131–138. [CrossRef]
45. Nath, D.; Nath, A.; Majumder, K.; Kundagrami, S. Phenotypic and molecular screening of rice genotypes inoculated with three widely available pathogenic strains of *Xanthomonas oryzae* pv. *oryzae*. *Agric. Biol. Res.* **2022**, *38*, 278–284. [CrossRef]
46. Kauffman, H.E.; Reddy, A.P.K.; Hsieh, S.P.Y.; Merca, S.D. Improved technique for evaluating resistance of rice varieties to *Xanthomonas oryzae*. *Plant Dis. Rep.* **1973**, *57*, 537–541.
47. Pradhan, K.C.; Barik, S.R.; Mohapatra, S.; Nayak, D.K.; Pandit, E.; Jena, B.K.; Sangeeta, S.; Pradhan, A.; Samal, A.; Meher, J.; et al. Incorporation of two bacterial blight resistance genes into the popular rice variety, ranidhan through marker-assisted breeding. *Agriculture* **2022**, *12*, 1287. [CrossRef]
48. Ke, Y.; Hui, S.; Yuan, M. *Xanthomonas oryzae* pv. *oryzae* inoculation and growth rate on rice by leaf clipping method. *Bio-Protocol* **2017**, *7*, e2568. [CrossRef]
49. Wonglom, P.; Watcharachaiyakup, J.; Patarapuwadol, S.; Kositratana, W. Assessment of diversity among pathotype of *Xanthomonas oryzae* pv. *oryzae* prevalent in Thailand. *Agric. Sci. J.* **2015**, *46*, 165–175.
50. Sombunjitt, S.; Sriwongchai, T.; Kuleung, C.; Hongtrakul, V. Searching for and analysis of bacterial blight resistance genes from Thailand rice germplasm. *Agric. Nat. Resour.* **2017**, *51*, 365–375. [CrossRef]
51. Lu, J.; Wang, C.; Zeng, D.; Li, J.; Shi, X.; Shi, Y.; Zhou, Y. Genome-wide association study dissects resistance loci against bacterial blight in a diverse rice panel from the 3000 rice genomes project. *Rice* **2021**, *14*, 22. [CrossRef] [PubMed]
52. Chen, L.; Yin, F.; Zhang, D.; Xiao, S.; Zhong, Q.; Wang, B.; Ke, X.; Ji, Z.; Wang, L.; Zhang, Y.; et al. Unveiling a novel source of resistance to bacterial blight in medicinal wild rice, *Oryza officinalis*. *Life* **2022**, *12*, 827. [CrossRef]
53. International Rice Research Institute (IRRI). *Standard Evaluation System for Rice*, 5th ed.; International Rice Research Institute (IRRI): Manila, Philippines, 2013.
54. Chumpol, A.; Monkham, T.; Saepaisan, S.; Sanitchon, J.; Falab, S.; Chankaew, S. Phenotypic broad spectrum of bacterial blight disease resistance from Thai indigenous upland rice germplasms implies novel genetic resource for breeding program. *Agronomy* **2022**, *12*, 1930. [CrossRef]
55. Broad Institute. Picard. 2022. Available online: https://broadinstitute.github.io/picard/ (accessed on 15 January 2022).
56. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef]
57. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997v2.
58. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The genome analysis toolkit: A map reduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [CrossRef]
59. Broad Institute. Genome Analysis Toolkit: GATK. 2022. Available online: https://gatk.broadinstitute.org/hc/en-us/categories/36000231059 (accessed on 15 February 2022).
60. Wei, Y.; Liu, S.; Xiong, D.; Xiong, Z.; Zhang, Z.; Wang, F.; Huang, J. Genome-wide association study for non-photochemical quenching traits in *Oryza sativa* L. *Agronomy* **2022**, *12*, 3216. [CrossRef]

61. 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**, *23*, 2633–2635. [CrossRef] [PubMed]
62. Cruz, M.; Arbelaez, J.D.; Loaiza, K.; Cuasquer, J.; Rosas, J.; Graterol, E. Genetic and phenotypic characterization of rice grain quality traits to define research strategies for improving rice milling, appearance, and cooking qualities in Latin America and the Caribbean. *Plant Genome* **2021**, *14*, e20134. [CrossRef] [PubMed]
63. Browning, B.L.; Tian, X.; Zhou, Y.; Browning, S.R. Fast two-stage phasing of large-scale sequence data. *Am. J. Hum. Genet.* **2021**, *108*, 1880–1890. [CrossRef]
64. Browning, B.L.; Zhou, Y.; Browning, S.R. A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* **2018**, *103*, 338–348. [CrossRef]
65. Munyengwa, N.; Guen, V.L.; Bille, H.N.; Souza, L.M.; Demange, A.C.; Mournet, P.; Masson, A.; Soumahoro, M.; Kouassi, D.; Cros, D. Optimizing imputation of marker data from genotyping-by-sequencing (GBS) for genomic selection in non-model species: Rubber tree (*Hevea brasiliensis*) as a case study. *Genomics* **2021**, *113*, 655–668. [CrossRef] [PubMed]
66. Spence, J.P.; Sinnott-Armstrong, N.; Assimes, T.L.; Pritchard, J.K. A flexible modeling and inference framework for estimating variant effect sizes from GWAS summary statistics. *bioRxiv* **2022**. [CrossRef]
67. Xu, X.; Junhua, Y.; Yang, Y.; Zhang, M.; Xu, Q.; Feng, Y.; Yuan, X.; Yu, H.; Wang, Y.; Yang, Y.; et al. Genome-wide association study of rice rooting ability at the seedling stage. *Rice* **2020**, *13*, 59. [CrossRef]
68. Hill, W.; Weir, B. Variances and covariances of squared linkage disequilibria in finite populations. *Theor. Popul. Biol.* **1988**, *33*, 54–78. [CrossRef]
69. Huang, X.; Wei, X.; Sang, T.; Zhao, Q.; Feng, Q.; Zhao, Y.; Li, C.; Zhu, C.; Lu, T.; Zhang, Z.; et al. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* **2010**, *42*, 961–967. [CrossRef]
70. Zhang, Y.M.; Jia, Z.; Dunwell, J. Editorial: The applications of new multi-locus GWAS methodologies in the genetic dissection of complex traits. *Front. Plant Sci.* **2019**, *10*, 100. [CrossRef]
71. Kaler, A.S.; Gillman, J.D.; Beissinger, T.; Purcell, L.C. Comparing different statistical models and multiple testing corrections for association mapping in soybean and maize. *Front. Plant Sci.* **2020**, *10*, 1794. [CrossRef] [PubMed]
72. Lipka, A.E.; Tian, F.; Wang, Q.; Peiffer, J.; Li, M.; Bradbury, P.J.; Gore, M.J.; Buckler, E.S.; Zhang, Z. GAPIT: Genome association and prediction integrated tool. *Bioinformatics* **2012**, *28*, 2397–2399. [CrossRef] [PubMed]
73. R Core Team. R: A Language and Environment for Statistical Computing. 2022: R Foundation for Statistical Computing, Vienna, Austria. Available online: https://www.R-project.org/ (accessed on 19 February 2022).
74. Yu, J.; Pressoir, G.; Briggs, W.H.; Vroh, B.I.; Yamasaki, M.; Doebley, J.F.; Deng, Y.; Dai, L.; Zhou, Y.; Zhang, X.; et al. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **2002**, *296*, 79–92. [CrossRef]
75. Wang, Q.; Tian, G.; Pan, Y.; Buckler, E.S.; Zhang, Z. A SUPER powerful method for genome wide association study. *PLoS ONE* **2014**, *9*, e107684. [CrossRef] [PubMed]
76. Liu, X.; Huang, M.; Fan, B.; Buckler, E.S.; Zhang, Z. Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet.* **2016**, *12*, e1005767. [CrossRef]
77. Huang, M.; Liu, X.; Zhou, Y.; Summers, R.M.; Zhang, Z. BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* **2019**, *8*, giy154. [CrossRef]
78. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **1995**, *57*, 289–300. [CrossRef]
79. Subedi, S.R.; Sandhu, N.; Singh, V.K.; Sinh, P.; Kumar, S.; Singh, S.P.; Ghimire, S.K.; Pandey, M.; Yadaw, R.B.; Varshney, R.K.; et al. Genome-wide association study reveals significant genomic regions for improving yield, adaptability of rice under dry direct seeded cultivation condition. *BMC Genom.* **2019**, *20*, 471. [CrossRef]
80. International Rice Genome Sequencing Project, The Rice Annotation Project Database (rap-dp). Available online: https://rapdb.dna.affrc.go.jp/download/irgsp1.html (accessed on 18 February 2022).
81. National Bio-Resource Project: NBRP. Oryzabase: Integrated Rice Science Database. 2023. Available online: https://shigen.nig.ac.jp/rice/oryzabase/ (accessed on 1 December 2022).
82. Sakai, H.; Lee, S.S.; Tanaka, T.; Numa, H.; Kim, J.; Kawahara, Y.; Wakimoto, H.; Yang, C.C.; Iwamoto, M.; Abe, T.; et al. Rice annotation project database (RAP_DB): An integrative and Interactive database for rice genomes. *Plant Cell Physiol.* **2013**, *54*, e6. [CrossRef]
83. The Rice Annotation Project Database. Available online: https://rapdb.dna.affrc.go.jp/search/?q=Os08g0205200 (accessed on 15 February 2022).
84. Rice Genome Annotation Project. Available online: http://rice.uga.edu (accessed on 15 February 2022).
85. Barrett, J.C.; Fry, F.; Maller, J.; Daly, M.J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **2005**, *21*, 263–265. [CrossRef]
86. Lu, Q.; Zhang, M.; Niu, X.; Wang, S.; Xu, Q.; Feng, Y.; Wang, C.; Deng, H.; Yuan, X.; Yu, X.; et al. Genetic variation and association mapping for 12 agronomic traits in indica rice. *BMC Genom.* **2015**, *16*, 1067. [CrossRef] [PubMed]
87. Aesomnuk, W.; Ruengphayak, S.; Ruanjaichon, V.; Sreewongchai, T.; Malumpong, C.; Vanavichhit, A.; Toojinda, T.; Wanchana, S.; Arikit, S. Estimation of the genetic diversity and population structure of Thailand's rice landraces using SNP markers. *J. Agron.* **2021**, *11*, 995. [CrossRef]

88. Pusadee, T.; Wongtamee, A.; Rerkasem, B.; Olsen, K.M.; Jamjod, S. Farmers drive genetic diversity of Thai purple rice (*Oryza sativa* L.) landraces. *Econ. Bot.* **2019**, *73*, 76–85. [CrossRef]
89. Kaewcheenchai, R.; Vejchasarn, P.; Hanada, K.; Shirai, K.; Jantasuriyarat, C.; Juntawong, P. Genome-wide association study of local Thai indica rice seedlings exposed to excessive Iron. *Plants* **2021**, *10*, 798. [CrossRef] [PubMed]
90. Fu, D.; Zhong, K.; Zhong, Z.; Hu, G.; Zhang, P.; Tong, H. Genome-wide association study of sheath blight resistance within a core collection of rice (*Oryza sativa* L.). *Agronomy* **2022**, *12*, 1493. [CrossRef]
91. Divya, B.; Biswas, A.; Robin, S.; Rabindran, R.; Joel, A.J. Gene interactions and genetics of blast resistance and yield attributes in rice (*Oryza sativa* L.). *J. Genet.* **2014**, *93*, 415–424. [CrossRef]
92. Zanetti, M.E.; Carolina, R.; Andreas, N. Plant NF-Y transcription factors: Key players in plant-microbe interactions, root development and adaptation to stress. *BBA-Gene Regul. Mech.* **2016**, *1860*, 645–654. [CrossRef]
93. He, X.; Liu, G.; Li, B.; Xie, Y.; Wei, Y.; Shang, S.; Shi, H. Functional analysis of the heterotrimeric NF-Y transcription factor complex in cassava disease resistance. *Ann. Bot.* **2019**, *124*, 1185–1197. [CrossRef]
94. Zhao, H.; Wu, D.; Kong, F.; Lin, K.; Zhang, H.; Li, G. The *Arabidopsis thaliana* nuclear factor Y transcription factors. *Front. Plant Sci.* **2016**, *7*, 2045. [CrossRef] [PubMed]
95. Buttner, D.; Nennstiel, D.; Klusener, B.; Bonas, U. Functional analysis of HrpF, a putative type III translocon protein from *Xanthomonas campestris* pv. *vesicatoria*. *J. Bacteriol. Res.* **2002**, *184*, 2389–2398. [CrossRef]
96. Yang, F.; Zhang, J.; Zhang, H.; Ji, G.; Zeng, L.; Li, Y.; Yu, C.; Fernando, W.G.D.; Chen, W. Bacterial blight induced shifts in endophytic microbiome of rice leaves and the enrichment of specific bacterial strains with pathogen antagonism. *Front. Plant Sci.* **2020**, *11*, 963. [CrossRef] [PubMed]
97. Brodie, A.; Azaria, J.R.; Ofran, Y. How far from the SNP may the causative genes be? *Nucleic Acids Res.* **2016**, *44*, 6046–6054. [CrossRef] [PubMed]
98. Lu, Y.; Zhong, Q.; Xiao, S.; Wang, B.; Ke, X.; Zhang, Y.; Yin, F.; Zhang, D.; Jiang, C.; Liu, L.; et al. A new NLR disease resistance gene *Xa47* confers durable and broad-spectrum resistance to bacterial blight in rice. *Front. Plant Sci.* **2022**, *13*, 1037901. [CrossRef] [PubMed]
99. Zhang, B.; Zhang, H.; Li, F.; Ouyang, Y.; Yuan, M.; Li, X.; Xiao, J.; Wang, S. Multiple alleles encoding atypical NLRs with unique central tandem repeats in rice confer resistance to *Xanthomonas oryzae* pv. *oryzae*. *Plant Commun.* **2020**, *1*, 100088. [CrossRef]
100. Wang, L.L.; Ma, Z.B.; Kang, H.X.; Gu, S.; Mukhina, Z.; Wang, C.H.; Wang, H.; Bai, Y.J.; Sui, G.M.; Zheng, W.J. A blast resistance gene Pi65 with LRR-RLK domain is required for resistance to *M. oryzae*. *Theor. Appl. Genet.* **2023**. [CrossRef]
101. Chern, M.; Bai, W.; Chen, X.; Canlas, P.E.; Ronald, P.C. Reduced expression of glycolate oxidase leads to enhanced disease resistance in rice. *PeerJ* **2013**, *1*, e28. [CrossRef]
102. Niño, M.C.; Cho, Y.G. Transcriptional modulation of resistance against *Xanthomonas oryzae* pv. *oryzae* Korean race K2 in japonica rice. *Agronomy* **2020**, *10*, 960. [CrossRef]
103. Qiu, T.; Zhao, X.; Feng, H.; Yang, J.; Peng, Y.L.; Zhao, W. OsNBL3, a mitochondrion-localized pentatricopeptide repeat protein, is involved in splicing nad5 intron 4 and its disruption causes lesion mimic phenotype with enhanced resistance to biotic and abiotic stresses. *Plant Biotechnol. J.* **2021**, *19*, 2277–2290. [CrossRef]