# Key Genomic Regions of Rice Cultivar GuiHeFeng and Its Derivatives Revealed by Genome-Wide Analysis

Yu-Zhi Chen †, Xin-Yu Hao †, Yuexiong Zhang, Zeng-Feng Ma, Chi Liu, Xiao-Long Zhou, Min-Yi Wei, Bao-Xiang Qin, Yong Yan *, Dahui Huang *

Posted Date: 5 December 2025

doi: 10.20944/preprints202512.0238.v1

Keywords: rice; genome resequencing; single nucleotide polymorphisms; gene; breeding

 doi:10.20944/preprints202512.0238.v1



*Article*

# Key Genomic Regions of Rice Cultivar GuiHeFeng and Its Derivatives Revealed by Genome-Wide Analysis

**Yu-Zhi Chen [1,2,†], Xin-Yu Hao [1,2,†], Yue-Xiong Zhang [1], Zeng-Feng Ma [1], Chi Liu [1], Xiao-Long Zhou [1], Min-Yi Wei [1], Bao-Xiang Qin [2], Yong Yan [3,*] and Da-Hui Huang [1,*]**

[1] Rice Research Institute, Guangxi Academy of Agricultural Sciences/Guangxi Key Laboratory of Rice Genetics and Breeding, Nanning 530007, China

[2] College of Agriculture, Guangxi University, Nanning 530004, China

[3] Microbiology Research Institute, Guangxi Academy of Agricultural Sciences, Nanning 530007, China

* Correspondence: 14544286@qq.com (Y.Y.); hdh1103@ 163.com (D.-H.H.)

† These authors contributed equally to this work.

**Abstract**

Rice is a widely cultivated staple crop that serves as the primary source of carbohydrates for more than half of the global population. Elite parents with superior agronomic traits play a crucial role in rice breeding systems. In this study, we performed whole-genome resequencing of the rice cultivar GuiHeFeng and its nine derivative lines, identifying a total of 6,633,507 high-quality single-nucleotide polymorphisms (SNPs). The percentage of GuiHeFeng traceable blocks (GTBs) in the 9 derivatives ranged from 48.94% to 63.2%. Based on SNP analysis, we found 1310 key GuiHeFeng traceable blocks, which were derived from GuiHeFeng and present in all 9 derivatives. Moreover, 375 selective sweeps (SSWs) were identified, of which 20 were also located within the kGTBs. These 20 SSWs were regarded as key genomic regions for rice breeding. After the association test, 20 alleles including 17 genes were identified on the kGTBs, and 38 significant genes were found within the key genomic regions. A total of 295 SNPs related to agronomic traits were detected by GWAS analysis. This research identifies genomic segments and agronomically important genes/QTLs that will serve as essential targets for genomic selection in rice breeding.

**Keywords:** rice; genome resequencing; single nucleotide polymorphisms; gene; breeding

## 1. Introduction

Rice (Oryza sativa L.) is a vital food crop and the primary staple for over half of the global population [1]. It serves as a fundamental pillar of worldwide food security [2]. To address the rising food requirements of an expanding global populace, rice output has seen consistent growth over recent decades, largely attributed to the creation of new high-yielding varieties [3]. The introduction of semi-dwarf cultivars, a central element of the initial Green Revolution, led to a substantial boost in rice productivity during the 1960s. Hybrid rice varieties have further enhanced production by 9% over conventional types [4]. Nowadays rice production confronts challenges of rapid population growth, shrinking farmland, climate change and pest/disease pressure [5]. To ensure worldwide food security, there is an urgent need to create new rice varieties that offer higher yields and greater resilience to both biotic and abiotic stresses. Traditional breeding remains inefficient in developing new varieties due to limited understanding of genetic mechanisms and the time-consuming, labor-intensive process of selecting target traits [6]. According to the brand-new concept of 5G breeding, Genomic breeding (GB), which encompasses marker-assisted selection (MAS) and genomic selection (GS), appears to be a highly effective strategy for producing new high-yielding rice varieties capable of withstanding stressful conditions and unpredictable climate shifts [6].

 doi:10.20944/preprints202512.0238.v1



Several key characteristics of rice are governed by genes or quantitative trait loci (QTLs) with substantial effects. Marker-assisted selection (MAS) for major-effect genes/QTLs has been widely applied to improve agronomic traits such as yield, disease resistance, and stress tolerance. However, most agronomic traits are influenced by QTLs with minor phenotypic contributions [7]. Minor-effect QTLs are restrained from utilizing in marker assisted selection (MAS), mainly because of their uncertainty in different genetic backgrounds and growing environments [8]. It is necessary to identify a robust consensus genomic region for Minor-effect QTLs to improve their effectiveness in MAS [9,10]. Analysis of the key/conserved regions that contain the excellent alleles in elite germplasms as well as the foundation parents is a good alternative to identify these consensus genomic regions relevant to the important agronomic traits [11]. Identifying key genomic regions is fundamental to understanding the genetic basis of elite traits and accelerating the breeding of improved crop varieties [12–14].

Backbone parents, which carry accumulated beneficial agronomic traits, such as disease resistance, high yield, and adaptability, play a crucial role in modern crop breeding programs due to their ability to transmit desirable traits to offspring through selective breeding[15–17].These parents are foundational in crop breeding, as evidenced by their widespread use in major Chinese rice varieties (e.g., 70% derived from 35 backbone parents between 1950–2008)[18]. A large number of superior alleles were gathered and distributed on different genomic regions, due to selective sweeps pyramid in the long-time pedigree breeding progress of the backbone parents. Through large-scale genome sequencing in combination with pedigree analysis, some key genomic regions, which can stably inherit in different genetic backgrounds of the pedigree, have been found in the rice backbone parents such as Minghui63, Huanghuazhan, Shuhui527 and Jiayu253[5,18–20]. These key genomic regions are important for genomic selection such as genome-wide marker assisted selection to develop new rice cultivars [11]. These four backbone parents were developed or released more than two decades ago, in 1980, 1996, 2005 and 2005, for Minghui 63, Shuhui 527 Huanghuazhan and Jiayu 253, respectively. Nevertheless, the genomic structure of rice cultivars will evolve due to shifts in their growing conditions and production objectives [6]. Only a few rice cultivars were analyzed to identify critical genomic regions associated with important traits through genome sequencing. Moreover, little is known about the key genomic regions architecture of the rice varieties developed in recent years.

GuiHeFeng is an elite conventional rice cultivar released in 2015, showing increase of yield by 12.32% compared with the control cultivars LiuShaYouZhan202 in regional test, and from which more than 10 excellent cultivars have been derived. In this study, GuiHeFeng and its 9 derivatives were selected for Whole-genome resequencing (Table 1). Using this sequence information, we were able to uncover the key genomic regions of GuiHeFeng conserved in all derivatives. We further analyzed known loci related to rice important trait or unknown QTLs by GWAS analysis, revealing the basis for the excellent performance of GuiHeFeng and all its derivatives. This comprehensive study of genomic architecture of GuiHeFeng and its derivatives will provide key genomic regions and important agronomic genes/QTLs for rice high yield breeding by genomic selection (GS).

**Table 1.** Tables should be placed in the main text near to the first time they are cited.

| No. | Variety | Pedigree |
|---|---|---|
| 1 | GuiHeFeng | HeFengZhan/YueTaiZhan |
| 2 | GuiFeng18 | GuiHeFeng/MeiXiangZhan |
| 3 | HeFengDao445 | GuiHeFeng/GuiYu9Hao |
| 4 | NaFengZhan | GuiHeFeng//ZaoHui3Hao/GuiHui1561 |
| 5 | JingYouXiang139 | BaiXiang139/GuiHeFeng |
| 6 | GuiYaXiang | GuiHeFeng/XiangChangMang |
| 7 | GuiNongFeng | GuiHeFeng/YeXiangZhan |
| 8 | NaXiangSiMiao | GuiHeFeng//BaiXiang139/GuiHui110 |
| 9 | NaGuXiang | GuiHeFeng//BaiXiang139/HuangHuaZhan |

 doi:10.20944/preprints202512.0238.v1



| | | |
|---|---|---|
| 10 | HeXiFengZhan2Hao | HeXiXiang/GuiHeFeng |

## 2. Results

### *2.1. The Derivatives Exhibited Comparable Agronomic Trait Performance to GuiHeFeng*

Investigation of 11 agronomic traits was conducted for all the cultivars (Figure 1 and Table S1-1). Only the plant height of HeXiFengZhan2 and GuiYaXiang was higher and lower than that of GuiHeFeng at significant level respectively, and there was no significant difference of plant height between GuiHeFeng and the other 7 cultivars (Figure 1A). The effective panicle number (EPN) of NaFengZhan and GuiYaXiang was significantly higher than that of GuiHeFeng, and the EPN of GuiNongFeng was significantly lower than that of GuiHeFeng, and there was no significant difference of plant height between GuiHeFeng and the other 6 cultivars (Figure 1B). Just as the case of plant height, only the panicle length (PL) of NaXiangSiMiao and GuiYaXiang was significantly higher and lower than that of GuiHeFeng respectively, and there was no significant difference in PL between GuiHeFeng and the other 7 cultivars (Figure 1C). There was no significant difference of number of unfilled grains （NUG) and number of filled grains (NFG) between GuiHeFeng and all the other 9 derivatives ( Figure 1D, E). The seed setting rate (SSR) of JingYouXiang 139, NaXiangSi and NaGuXiang was significantly higher than that of GuiHeFeng, and there was no significant difference of SSR between GuiHeFeng and the other 6 cultivars (Figure 1F). Only the length of flag leaf of NaFengZhan and GuiNongFeng was significantly higher than that of GuiHeFeng, without significant difference between GuiHeFeng and the other 7 cultivares (Figure 1G). Except HeFengDao445, GuiNongFeng and NaXiangSiMiao, there was no significant difference of Width of flag leaf (WFL) between GuiHeFeng and the other 5 cultivars (Figure 1H). There was no significant difference of seed weight per plant （SWPP) between GuiHeFeng and all the other 9 derivatives ( Figure 1I). There was no significant difference in thousand-kernel weight (TKW) between GuiHeFeng and the other 4 cultivars, including HeFengDao445, JingYouXiang139, GuiYaXiang and NaXiangSiMiao, respectively (Figure 1J). Ratio of length and width (RLW) of GuiHeFeng and GuiYaXiang was significantly lower than that of other 8 cultivars (Figure 1K).

In addition, for NUG, NFG and SWPP, there was no significant difference between GuiHeFeng and its all 9 derivatives; for PH, PL and LFL, there was no significant difference between GuiHeFeng and other 7 cultivars; for EPN and SSR, the number was 6 cultivars; for WFL and TKW, the number was 5 cultivars and 4 cultivars, respectively. Over half of the derivatives closely resembled GuiHeFeng in the majority of the agronomic characteristics that were evaluated.

 doi:10.20944/preprints202512.0238.v1



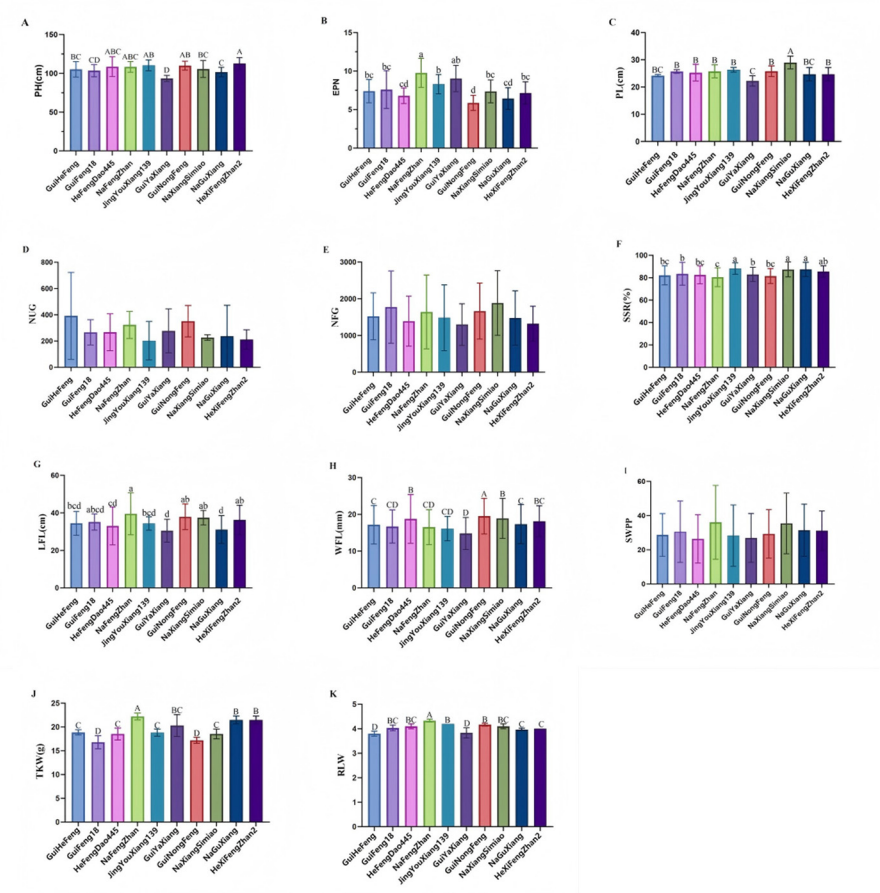


**Figure 1.** Agronomic traits. (A) Plant height (PH). (B) Effective panicle number (EPN). (C) Panicle length (PL). (D) Number of unfilled grains (NUG). (E) Number of filled grains (NFG). (F) Seed setting rate (SSR). (G) Length of flag leaf (LFL). (H) Width of flag leaf (WFL). (I) Seed weight per plant (SWPP). (J) Thousand-kernel weight (TKW). (K) Ratio of length and width (RLW) were investigated during all growth seasons. Statistical analysis was performed using LSD software. In the same comparison group, values with different small letter superscripts mean a significant difference ($P < 0.05$) and different capital letter superscripts mean a very significant difference ($P < 0.01$).

### *2.2. Sequence and SNPs Information Was Produced by Whole-Genome Resequencing*

GuiHeFeng and all its 9 derivatives were used to carry out whole-genome resequencing to get the basic sequence and SNPs information for further analysis. A total of 751.32 million (M) clean reads of 150 bp that include 111.18 GB data were generated from the 10 rice varieties with more than 19 × depth (Table 2). More than 98% of the clean reads of 150 bp were mapped to the Nipponbare genome, with average coverage ratio ranging from 82.42% to 89.69% (Table 2).

We used GATK v4.0 to call SNPs [21]. Overall, 6 633 507 SNPs were identified in these listed 10 cultivars. According to the genome annotation (MSU 6.1), 60.40% of all SNPs were found in intergenic regions, 5.12% in introns, 2.38% in UTRs and 3.91% in gene coding regions, 28.18% in other regions (Figure 2; Table S2).

 doi:10.20944/preprints202512.0238.v1



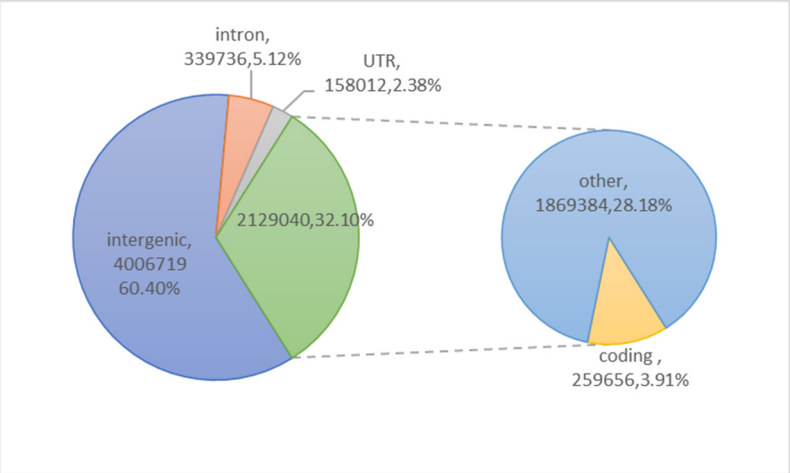


**Figure 2.** Distribution of SNPs localization in 10 cultivars.

**Table 2.** Resequencing of GuiHeFeng and its 9 derivatives.

| Variety | Reads（M） | Bases (G) | Map Reads (%) | Map Reads | Depth X | Cov_ratio (%) |
|---|---|---|---|---|---|---|
| GuiHeFeng | 124.74 | 18.59 | 98.59 | 122978517 | 51.78 | 89.69 |
| GuiFeng18 | 58.54 | 8.72 | 98.73 | 57794774 | 24.51 | 85.53 |
| HeFengDao445 | 66.84 | 9.93 | 98.77 | 66016096 | 27.99 | 86.67 |
| NaFengZhan | 70.66 | 10.53 | 98.62 | 69682802 | 29.65 | 86.98 |
| JingYouXiang139 | 84.91 | 12.60 | 98.70 | 83805328 | 35.4 | 87.92 |
| GuiYaXiang | 76.20 | 11.34 | 98.58 | 75121141 | 31.73 | 87.22 |
| GuiNongFeng | 71.49 | 10.66 | 98.61 | 70495489 | 29.83 | 87.66 |
| NaXiangSiMiao | 81.89 | 12.17 | 98.62 | 80756694 | 34.09 | 87.74 |
| NaGuXiang | 68.9 | 10.23 | 98.62 | 67954951 | 28.74 | 86.36 |
| HeXiFengZhan2 | 47.15 | 7.04 | 98.46 | 46426567 | 19.98 | 82.42 |
| Sum | 751.32 | 111.18 | | | | |

*2.3. Key GuiHeFeng Traceable Blocks Were Found in the Genome of Its Derivatives*

As the method described by Zhou et al [5], the rice genome was segmented into 7471 adjacent blocks with bin size of 50 kb (Table S3). Using a cut-off of more than 85% identity between GuiHeFeng and the derivatives to exploit the GuiHeFeng traceable blocks (GTBs). As shown in Figure 4, 63.2% genomic blocks of HeFengDao445 were identified as GTBs, 59.94% for GuiFeng18, 59.73% for NaGuXiang, 59.17% for NaXiangSiMiao, 52.63% for GuiNongFeng, 51.98% for JingYouXiang139, 50.59% for HeXiFengZhan2, 50.07% for GuiYaXiang and 48.94% for NaFengZhan, respectively. There were 1310 key GTBs (kGTBs), which were derived from GuiHeFeng and found in all the 9 derivatives (Table S4). These key GTBs were unevenly distributed on the hole genome of rice, chromosome 3 with the largest number of 192, and chromosome 11 with the lowest number of 22 (Figure 3).

 doi:10.20944/preprints202512.0238.v1



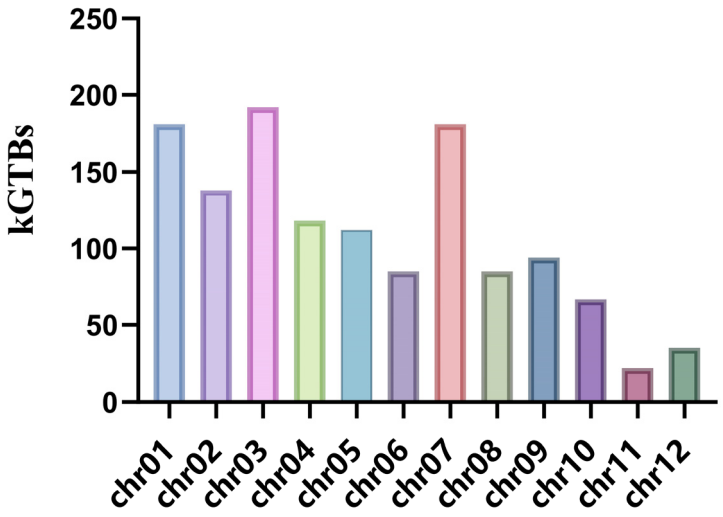


**Figure 3.** These key GTBs exhibited a non-uniform genomic distribution across the rice genome.(chr.1:181; chr.2:138; chr.3:192; chr.4:118; chr.5:112; chr.6:85; chr.7: 181; chr.8:85; chr.9:94; chr.10:67; chr.11:22; chr.12: 35).

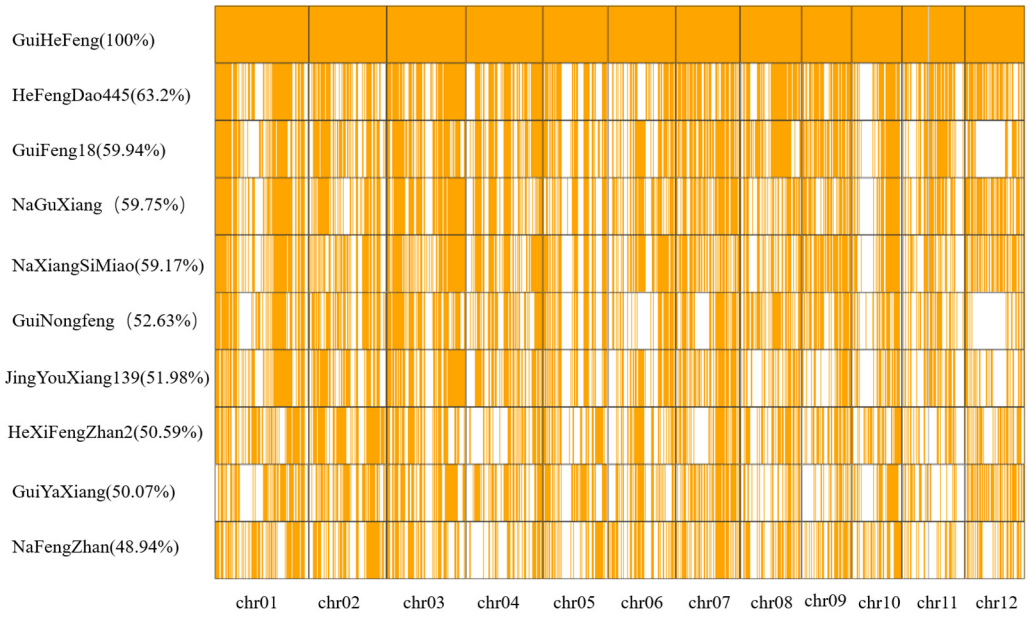


**Figure 4.** Key genomic regions of GuiHeFeng. The key genomic regions of GuiHeFeng stably inherited by its elite derivatives are represented in orange. The derivative name and the similarity of each derivative to the GuiHeFeng genome is shown in the left side.

*2.4. Key Genomic Regions Were Selected from kGTBs and Selection Sweeps*

Selective sweeps are the genomic region which probably contain excellent alleles relevant to the important agronomic traits and preferably selected by the breeder selective sweeps (SSWs) [11]. To exploit the selective sweeps (SSWs) of GuiHeFeng and its derivatives, $\theta\pi$, $\theta w$ and Tajima's D [22] were calculated with sliding window of 50 kb across 12 chromosomes with Variscan [23], with a cut-off of 5% of Tajima's D test ( Tajima's D≥1.94).We found 375 SSWs, totaling 18.75 Mb, distributing on all chromosomes (Figure 5; Table S5). Furthermore, we found only 20 SSWs were included in kGTBs, indicating all these 20 keys genomic regions were important for rice breeding and preferably selected by different breeders.

 doi:10.20944/preprints202512.0238.v1



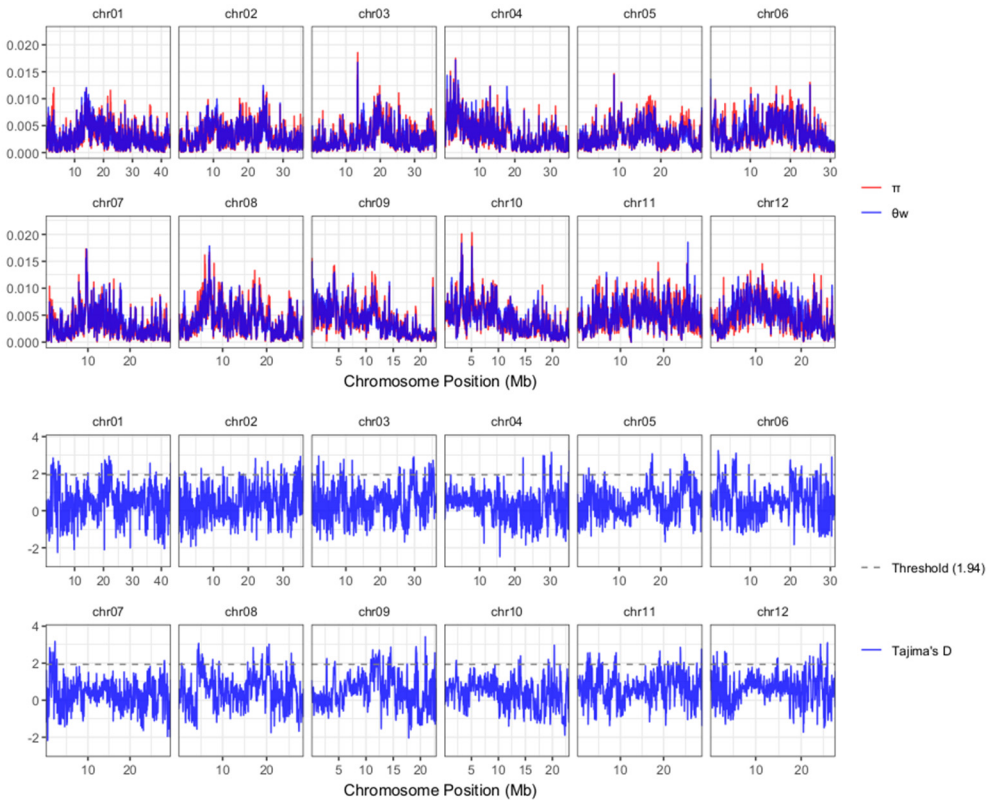


**Figure 5.** Distribution of SSWs on the chromosome.

*2.5. Excellent Alleles Were Exploited from kGTBs and Key Genomic Region*

Beyond key genomic regions, rice breeders are particularly interested in superior alleles located within these areas. To find the excellent alleles on kGTBs, adjacent SNPs with the same segregation pattern were combined to form a marker for association test with agronomic traits by PLINK analysis. As results (Table 3, Table S6) , 20 alleles including 17 genes were found on the kGTBs, 2 genes *Rd* and *OsCYP704A3*, associated with seed morphology, 2 genes *D2* and *TAC1*, linked to plant architecture, 6 genes (*Gnla*, *Rf3*, *OsLG3*, *DPL2*, *GLW7* and *HSA1b*) related to yield, 2 genes *Hd7* and *Hd1*, involved in heading date, 4 genes (*BET1*, *OsJAZ1*, *bZIP73* and *LHCB5*) for biotic stress, and one gene *OsUGT707A2* for secondary metabolism, respectively. There was the largest number of genes involved in yield regulation, while only one gene related to secondary metabolism. The SNPs polymorphism consistency between GuiHeFeng and all the derivatives of the 17 genes between were reconfirmed by gene chip analysis (Table S7-1). However, the derivatives showed difference between GuiHeFeng at some genes, such as *ALK*, *Badh2* and *Rf2* (Table S7-2, Table S7-3). We found no important genes on the 20 key genomic regions by PLINK analysis. So, we directly found the loci for key genomic regions (kGRs) in Nipponbare genome IRGSP-1.0 on Rice Gene Index (RGI; https://riceome.hzau.edu.cn) platform. As shown in Table 4 and Table S5, there were 38 genes on the key genomic regions, except the key genomic regions on chr.12. To our surprise, among the 38 genes, 29 genes are involved in the defense responses against biotic/abiotic stress, 4 genes for fertility, only 1 gene for yield components, and 4 genes for other functions (Table 4; Table S8).

**Table 3.** Important alleles relevant to agronomic traits on kGTBs.

| No. | Chr | Star | End | Gene | Function | Category | Genechip result |
|---|---|---|---|---|---|---|---|
| 1 | chr01 | 25383093 | 25383093 | *Rd/DFR/OsDfr* | red seed coat | Seed Morphology | T |
| 2 | chr01 | 5244076 | 5244076 | *D2/CYP90D2/SMG11* | larger tiller angle | Plant Architecture | T |
| 3 | chr01 | 5270928 | 5270928 | *Gn1a/OsCKX2* | increasing grain number | Yield components | T |

 doi:10.20944/preprints202512.0238.v1



| | | | | | | |
|---|---|---|---|---|---|---|
| 4 | chr01 5275530 | 5275530 | *Gn1a/OsCKX2* | increasing grain number | Yield components | T |
| 5 | chr01 5275544 | 5275544 | *Gn1a/OsCKX2* | increasing grain number | Yield components | T |
| 6 | chr01 5568692 | 5568692 | *Rf3/OsMADS3* | fertility restoration | Yield components | T |
| 7 | chr023009633030096330 | | *DTH2/Hd7* | delaying heading date under LD | Heading date | T |
| 8 | chr03 4353347 | 4353347 | *OsLG3* | increasing drought tolerance | Yield components | T |
| 9 | chr03 4353103 | 4353103 | *OsLG3* | increasing drought tolerance | Yield components | T |
| 10 | chr042388665923886659 | | *BET1* | Increasing boron-toxicity tolerance | Abiotic Stress | T |
| 11 | chr042889475328894753 | | *OsCYP704A3* | Longer seed size | Seed Morphology | T |
| 12 | chr043330491033304910 | | *OsJAZ1* | decreasing root length and weight | Abiotic Stress | T |
| 13 | chr06 4201227 | 4201227 | *DPL2* | hybrid incompatibility | Yield components | T |
| 14 | chr06 9338220 | 9338220 | *Hd1* | Promoting heading date under LD | Heading date | T |
| 15 | chr071906039819060398 | | *OsUGT707A2* | more 5-O-glucoside | Secondary metabolism | T |
| 16 | chr071910324919103249 | | *OsSPL13/GLW7* | increasing grain size | Yield components | T |
| 17 | chr091812285018122850 | | *bZIP73* | decreasing chilling tolerance | Abiotic Stress | T |
| 18 | chr092073184420731844 | | *TAC1* | Spread-out plant architecture | Plant Architecture | T |
| 19 | chr11 7659694 | 7659694 | *LHCB5* | increasing blast resistance | Biotic Stress | T |
| 20 | chr122466979724669797 | | *HSA1b* | hybrid incompatibility | Yield components | T |

**Table 4.** Important alleles relevant to agronomic traits on key genomic region.

| No. | Chr | Star | End | Gene | Function | Category |
|---|---|---|---|---|---|---|
| 1 | chr01 2053583 | | 2057638 | *LRK10L-2.1* | resistance gene analogs (RGAs) | Biotic stress |
| 2 | chr012866630928668106 | | | *Xa21* | bacterial blight resistance | Biotic stress |
| 3 | chr012866947928673568 | | | *OsLRR-RLK* | Regulate defence reaction | Biotic stress |
| 4 | chr021279834412804729 | | | Retrovirus-related Pol polyprotein from transposon RE1 | Increase the Resistance for Broad bean wilt virus 2 | Biotic stress |
| 5 | chr032695204826959200 | | | *OsTHIC* | positively REGULATE vitamin B 1 synthesis | Other |
| 6 | chr03 3489869 | | 3500130 | *TOP3α* | regulates meiotic recombination | Other |
| 7 | chr042236963222376812 | | | *OsABA1* | Positively regulate plant development and adaptation to abiotic and biotic stresses | Biotic/Abiotic Stress |
| 8 | chr042235370722355207 | | | *OsAP37* | Mediate the tolerance to drought | Abiotic Stress |
| 9 | chr042236223922367204 | | | *OsPT17* | Involved in Chilling Response and salt stress | Abiotic Stress |
| 10 | chr042238930322393831 | | | *OsPP65* | Decrease rice resistance to chilling | Abiotic Stress |
| 11 | chr043318581333186889 | | | *OsWAK54* | plays important roles in cell expansion, pathogen resistance | Biotic stress |
| 12 | chr043319262333196131 | | | *OsWAK55* | plays important roles in cell expansion, pathogen resistance | Biotic stress |
| 13 | chr043528778135289156 | | | *OsPR5* | increase pathogen resistance | Biotic stress |
| 14 | chr043527095235276805 | | | *OsSPARK2* | negatively regulation the tolerance | Biotic/Abiotic Stress |
| 15 | chr062894127128943704 | | | *OsRRK1* | Positively regulate brown planthopper resistance | Biotic stress |
| 16 | chr062890557728909089 | | | *OsLRR-RLK1* | initiates striped stem borer resistance | Biotic stress |
| 17 | chr063035769930361201 | | | *OsNPSN11* | Positively regulate the blast resistance | Biotic stress |

 doi:10.20944/preprints202512.0238.v1



| | | | | |
|---|---|---|---|---|
| 18 | chr08156955341570396O | Protein PHR1-LIKE 3 | enhances tolerance to Pi deficiency and salt stress in rice | Abiotic Stress |
| 19 | chr09131549431315583 2 | *OsSAP17* | Enhancing plant resistance to drought and salt | Abiotic Stress |
| 20 | chr09131813301318474 1 | *OsPHD38* | Mediate the tolerance to drought and salt stress | Abiotic Stress |
| 21 | chr09175569291755859 1 | *OsDjC69* | Mediate flowering and the tolerance to drought and salt stress | Abiotic Stress |
| 22 | chr09175654711756619 7 | *OsbHLH043* | Mediate the tolerance to drought and arsenic stress | Abiotic Stress |
| 23 | chr09209153012091980 8 | *OsMYB85* | cell wall regulators | Other |
| 24 | chr09211517362115435 8 | *OsCYP-24* | Mediate the tolerance to drought and salt stress | Abiotic Stress |
| 25 | chr09211559562115794 5 | *OsRNS4* | enhanced tolerance to high salinity | Abiotic Stress |
| 26 | chr09211716532117406 7 | *OsPAD1* | regulate pollen aperture formation | Fertility |
| 27 | chr09211893812119073 8 | *OsMYB31* | Increase yield | Yield components |
| 28 | chr09211975032119972 3 | *MS5* | regulate pollen formation | Fertility |
| 29 | chr09211997312120276 3 | *OsAPX9* | Increase the tolerance to drought, plant height and heading date | Abiotic Stress/Heading date/Plant Architecture |
| 30 | chr09226538492265704 6 | *Ohp2* | Positively Mediate the tolerance to salt stress | Abiotic Stress |
| 31 | chr09226663062267139 2 | *OsWD40-174* | take important role in rice-*Xoo* interactions | Biotic stress |
| 32 | chr10203550762035565 7 | *OsERF18* | enhances tolerance to Pi deficiency | Abiotic Stress |
| 33 | chr10203742522037545 5 | *OsEMSA1* | Involved in Embryo sac Development | Fertility |
| 34 | chr10203771952038023 5 | *OsNP1* | required for another cuticle formation and pollen exine patterning | Fertility |
| 35 | chr10203770312038609 6 | *OsPLDbeta1* | activates defense responses and increases disease resistance in rice | Biotic stress |
| 36 | chr11288042482880855 0 | *OsHSP70* | Induces the tolerance to high temperature stress | Abiotic Stress |
| 37 | chr11288276762882851 3 | *OsMT1a* | positively regulated rice resistance to blast | Biotic stress |
| 38 | chr11288459052885293 8 | *OsSCL57* | Regulate the phosphorus homeostasis of rice | Other |

*2.6. More SNPs (QTLs) Related to Important Traits Were Detected by GWAS Analysis*

To determine more SNPs related to putative traits, we tested the association between SNPs and the mean data of agronomic traits collected from early season of 2024 (Table S1-2), late season of 2024(Table S1-3) and early season of 2025(Table S1-4) and the average(Table S1-1) in Nanning using compressed general linear model (GLM) and mixed linear model (MLM) implemented in TASSEL.

A total of 255 significant association sites were detected using GLM (Table S9). Among these sites, 130 SNPs were found to be associated with PH, with 128 identified in the early season of 2024 and 2 in the early season of 2025. For EPN, WFL, and TKW, the numbers of associated SNPs were 10, 48, and 49, respectively, all detected in the early season of 2024. Additionally, 58 SNPs were found to be associated with SSR in the early season of 2025. Interestingly, 5 associated sites were identified across 4 principal areas of the GuiHeFeng traceable block (kGTBs) (Table S9). During the early part of the 2024 season, the SNP positions chr04:23737663 and chr04:23737670, which are associated with plant height, were found to be situated on the gene *Os4BGlu11*, and chr07: 23610137 located on *LOC_Os07g39410* (Figure 6A). In early-season of 2025, SNPs position chr06: 12623369 was located on *LOC_Os06g21850.1*, associated with SSR (Figure 6B). No gene was found for the position of chr07:4127113, which was associated with WFL in early season of 2024. *Os4BGlu11* encodes β-Glucosidases, which hydrolyzes abscisic acid glucose ester (ABA-GE), regulating the development of root [24]. *LOC_Os06g21850.1* encodes conserved hypothetical protein, and *LOC_Os07g39410* encodes

 doi:10.20944/preprints202512.0238.v1



retrotransposon protein, both with unknown functions. No significant association sites were in MLM and GLM analysis of late-season of 2024 and the average of the 3 seasons.

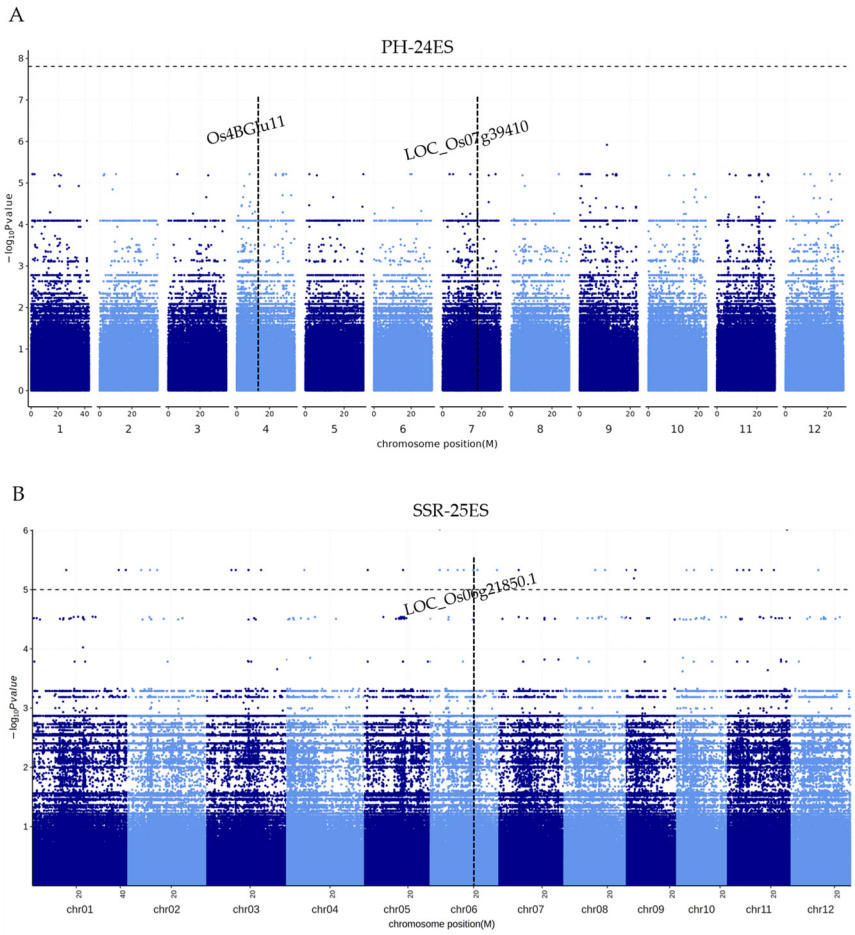


**Figure 6.** 6 SNPs detected in GWAS analysis. (A) In early-season of 2024: for plant height, SNPs position of chr04:23737663 chr04:23737670 were located on the gene *Os4BGlu11*, and chr07:23610137 located on *LOC_Os07g39410*; (B) In early-season of 2025, SNPs position chr06:12623369 was located on *LOC_Os07g39410*, associated with SSR.

## 3. Discussion

### *3.1. Important Genes Were Identified from GuiHeFeng and Its Derivatives*

Genes relevant to the critical agronomic traits play important role in rice breeding. For example, the 'Green Revolution' gene *sd1* has been used to develop a lot of rice cultivars and made a significant contribution increases in rice yields [25]. Exploiting and utilizing important genes from elite germplasm is the permanent target for rice breeders. Important genes such as *Xa21*[26], *Gn1a* [27], *Wx* [28], *GS5* [29]and *IPA1*[30] for resistance, grain yield, quality and plant type, were identified in an elite rice HuangHuaZhan through whole genome sequencing and pedigree analysis [5]. The important gene *TAC1*[31] was also found in HuangHuaZhan [6]. Six important genes—*sd1*[32], *LP* [33], *GW5*[34], *BC10*[35], *RL14*[36]and *OsNAC6*[37]—were discovered in another elite rice variety, 9311[19]. We identified 17 important genes in the kGTBs, which existed in both GuiHeFeng and the other 9 derivatives (Table 3). Among these 17 genes, 2 genes *Gn1a* and *TAC1* were also found the 7 genes identified in HuangHuaZhan. *Gn1a*, the first major QTL implicated in grain-number regulation per panicle, explained 44% of the phenotypic variance.[27]. *TAC1* is a major quantitative trait locus, positive controlling tiller angle in rice [31]. *D2* identified in GuiHeFeng also plays an important role in the regulation of tiller angle [38]. It seems that grain number per panicle and tiller angle, controlled

 doi:10.20944/preprints202512.0238.v1



by *Gn1a* and *TAC1/D2*, is a part of the most critical agronomic traits for rice breeder during breeding selection. *Rd* controls red coat of seed [39], and *CYP704A3* negatively controls the length of rice seed [40]. *Hd1*[41]and *DTH2*[42] both can delay heading date under long-day conditions. Longer heading date results in more biomass and higher yield. Maybe, this is the reason to explain the preference of breeder for *Hd1* and *DTH2* in rice breeding practice. *GLW7* increases both length and weight of rice grain [43]. Three seed production genes were found in GuiHeFeng and all the derivatives. *Rf3*[44] positively regulated the restoration of fertility, but *DPL2*[45] and *HSA1b* [46] both control hybrid incompatibility. The function of Rf3 contradicts the function of *DPL2* and *HSA1b*. However, GuiHeFeng and all the derivatives had high seed-setting rate, ranging from 80.5% to 88.2% (Table S1-2). Moreover, GuiHeFeng shows strong compatibility for both two-line and three-line male sterile line (Data unpublished). It is needed to carry out more research to illustrate the seed reproduction mechanism of the 3 genes for GuiHeFeng and the derivatives.

Previous research showed biotic and abiotic stress related genes were favored by breeders [5,6]. Our results were consentaneous with these previous findings. Among the 17 genes, 5 were stress related genes, *OsLG3*[47], *BET1*[48], *OsJAZ1*[49], *bZIP73*[50], *LHCB5*[51]. In addition, among the 38 genes in the key genomic regions, 29 genes are involved in the defense responses against biotic/abiotic stress (Table 4). Our results support the proposal：To maintain high yield and good quality of the target cultivars wherever cultivated, stress related genes would be spontaneously selected by different breeders to respond to varied environments in rice breeding.

### *3.2. kGTBs and Key Genomic Region Is Useful for Modern Rice Breeding*

Marker-assisted selection (MAS) has been successfully utilized to pyramid elite allele of important genes, improving the yield, quality and resistance of rice cultivars [52,53]. It is critical to assess the performance of target allele before its utilization in MAS. Due to uncertainty of genetic backgrounds and growing environments, it is difficult to detect the minor effect QTLs, especially for the abiotic stress related to QTLs, by traditional QTL analysis method [54,55]. A method named Meta-QTL analysis has been invented the detect the key genomic region, which contain the target allele and stably inherit in different genetic backgrounds and growing environments [9]. The emergence of high-throughput genome sequencing and the availability of pedigree analysis makes the finding of such key genomic region more precise and higher efficiency, and key genomic regions related to important agronomic traits have been found in the rice backbone parents, such as Minghui63, Huanghuazhan, Shuhui527 and Jiayu253 [5,18–20]. In the present study, 1310 key GuiHeFeng traceable blocks (Table S4), 375 selective sweeps, and 20 key genomic regions (Table S5) were identified from GuiHeFeng and the derivatives. Moreover 17 important genes were found on the kGTBs (Table 3), and 38 found on the 20 key genomic regions (Table 4). These key genomic regions could be used as important blocks for genomic selection (GS) in the future of rice breeding.

Some important genes, for example, most NLR genes are positionally clustered in a genomic region [13]. Some abiotic stress related QTLs are also clustered on the genomic region [56,57]. we found 3 alleles of *Gn1a* clustered on chr.1, 2 alleles of *OsLG3* clustered on chr.3 (Table 3). As shown in Table 4, 4 abiotic stress related genes *OsABA1*, *OsAP37*, *OsPT17* and *OsPP65* were clustered on chr.4; two resistance related genes *OsWAK54* and *OsWAK55* on chr.4; two resistance related genes *OsRRK1* and *OsLRR-RLK1* on chr.4 and 6; meanwhile, 9 abiotic stress related genes clustered on chr.9. It suggests that alleles of the same gene or QTLs, as well as gene/QTLs with similar functions, frequently cluster within specific genomic regions. In comparison with handling and utilizing individual alleles of gene/QTLs, key genomic regions that contain clusters of multiple elite alleles demonstrate greater effectiveness in rice breeding, particularly for minor-effect QTLs.

### *3.3. GuiHeFeng Is a Backbone Parent for Rice Breeding*

Backbone parents, as the carrier of multiple beneficial agronomic traits, are critical for rice breeding [18]. GuiHeFeng is typically a high yield rice cultivar, showing increase of yield by 12.32% in comparison with the control cultivars. So, it was used widely by different breeders to develop new

 doi:10.20944/preprints202512.0238.v1



rice cultivars. The percentage of GuiHeFeng traceable blocks in the derivatives ranged from 48.94% to 63.20%. However, more than half of the derivatives closely resembled GuiHeFeng in the majority of the agronomic traits that were evaluated (Figure 1). In addition, no derivative showed significant increase of seed weight per plant (SWPP) than that of GuiHeFeng. The results indicated that GuiHeFeng was dominant at large number of yield-related genes/QTLs, showing high heritability in yield performance. In addition to the maintaining of high yield performance of GuiHeFeng, 4 derivatives GuiNongFeng, NaXiangSiMiao, GuiYaXiang and JingYouXiang139 showed improvement of quality with fragrance Badh2 (Table S7-1). It is feasible to use GuiHeFeng as high yield backbone parent, to cross with the other unique parent to improve the quantity or resistance of rice cultivar.

Currently, biotic and abiotic stress tolerance has become a primary objective for rice breeding programs [6]. Our results showed that 18 of the 20 key genomic regions, which were identified from GuiHeFeng, contained more than one biotic or abiotic resistance related genes(Table S8). The results indicate that GuiHeFeng could be used as a stress resistance parent to develop new high-yield varieties of rice with resistance to stressful environments and unpredictable climate changes.

## 4. Materials and Methods

### *4.1. Plant Materials*

A total of 10 rice varieties were used for analysis in this study (Table 1). GuiHeFeng was one of the two parents of the other 9 derivatives. HeFengDao445 and HeXiFengZhan2Hao were collected from Hechi Agricultural Science Research Institute, JingYouXiang139 from Guangxi Boshiyuan Seed Industry Co., Ltd., GuiHeFeng and 9 derivatives from Rice Research Institute of Guangxi Academy of Agricultural Sciences. All varieties were planted in the experimental field of the Rice Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, China in early-season and Late-season of 2024, and early-season of 2025. Each variety was planted in three plots, 5 rows for each plot and 10 plants for each plot. The spacing between plants and plots was 20 cm×20 cm and 30 cm×30 cm respectively. The plots of all varieties were arranged in randomized complete block design.

### *4.2. Genome Resequencing and SNP Calling*

A single individual of each variety was selected for whole genome resequencing. Genomic DNA was extracted from young leaves using a DNA Extraction Kit (Qiagen, Hilden, Germany), sequenced on the Illumina X10 platform (150 bp reads and 300–500 bp insert). We removed the low-quality paired reads, including those (putative PCR duplicates, with >10 nucleotides aligned to the adapter, with ≥10% unidentified nucleotides (N) and >50% bases having phred quality <10) [58]. The clean reads were mapped to the reference genome of Nipponbare (MSU v7.0) by using Burrows–Wheeler Alignment (BWA) software (v0.7.12) [59]. The sequencing depth, genome coverage, and other information of each sample were calculated were calculated by SAMtools software [60]. GATK v4.0 software was used for identifying SNPs [21]. The SNPs were annotated using SnpEff (version 4.1) [61].

### *4.3. Construction of Genome Bins, Identification of Key Genomic Region and Selection Sweep Region*

The genome was segmented into non-overlapping bins of 50 kb length. The similarity between each sample of the 9 derivatives and GuiHeFeng is calculated to obtain the similarity matrix for each bin. If identity of according to bin of tested derivative and GuiHeFeng is larger than or equals to 0.85, then it was deemed as conserved blocks (GuiHeFeng traceable blocks, GTBs). Such GTBs found in all the 9 derivatives was considered as key GTB (kGTBs). To identify the selection sweeps (SSWs), θw and Tajima's D [22] were calculated with sliding window of 50 kb across 12 chromosomes with Variscan [23] using SNPs identified from resequencing. We used 5% as a cut-off of Tajima's D test (Tajima's D≥1.94) to identify top selective sweeps with high significance. Regions found in both

 doi:10.20944/preprints202512.0238.v1



kGTBs and top selective sweeps were identified as key genomic regions (kGRs) for rice breeding. Figures of key blocks or selection sweep regions were drawn using Perl script with GD module (www.perl.org).

### *4.4. Association Test and Gene Chip Analysis*

For kGTBs, adjacent SNPs with the same segregation pattern were combined to form a marker for association test [5]. PLINK was used to analyze the association between these markers and 11 agronomic traits in a linear model [62]. Important loci for agronomic traits were determined as those with FDR *p*-values less than 0.0001 from 100,000 permutation tests. The important loci for key genomic regions (kGRs) were found in Nipponbare genome IRGSP-1.0 on Rice Gene Index (RGI; https://riceome.hzau.edu.cn) platform. The whole-genome SNP array GSR40K was employed to analyze the variations of 164 functional genes. GSR40K analysis was performed at Wuhan Greenfafa Institute of Novel Gene chip R&D Co., LTD (Wuhan, China) (https : //green fafa.com/), according to the Infinium HD Assay Ultra Protocol (HYPERLINK: https://www.illumina.com).

### *4.5. Association Test and Gene Chip Analysis*

To determine more SNPs related to putative traits, we tested the association between SNPs and the mean data of agronomic traits collected from early season of 2024, late season of 2024 and early season of 2025 in Nanning using compressed general linear model (GLM) and mixed linear model (MLM) implemented in TASSEL.

### *4.6. Agronomic Trait Investigation*

Agronomic traits, including plant height (PH), effective panicle number (EPN) , panicle length (PL), number of unfilled grains (NUG), number of filled grains (NFG)， seed setting rate (SSR), length of flag leaf (LFL), width of flag leaf (WFL), seed weight per plant （SWPP), thousand-kernel weight (TKW), ratio of length and width (RLW) were investigated during all growth seasons. Statistical analysis was performed using LSD software.

## 5. Conclusions

Through in-depth analysis of key genomic regions in Guifeng rice using SNP data, this study integrated kGTB and SSW strategies to pinpointe critical genomic regions and identify superior alleles. It elucidated the potential molecular basis for yield traits and key regions underpinning stable inheritance in GuiHeFeng rice. These findings provide new insights for advancing molecular design, breeding and genomic selection in the GuiHeFeng rice background.

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Table S1-1: The Mean±SD of 11 agronomic traits investigated over three years for all cultivars; Table S1-2: The mean data of agronomic traits collected from early season of 2024; Table S1-3: The mean data of agronomic traits collected from late season of 2024; Table S1-4: The mean data of agronomic traits collected from early season of 2025; Table S2: Results of SNPs annotation; Table S3: All Bins. from. GuiHeFeng. matrix; Table S4: Key Region all Sample inherited from GuiHeFeng; Table S5: Selective Sweep. result. top; Table S6: All inherited from GuiHeFeng. Known Trait; Table S7-1-S7-3: Differentially expressed genes (DEGs) identified with the GuiHeFeng rice gene-chip; Table S8: Gene of Key region-4; Table S9: GWAS-SNPS.

**Author Contributions:** Writing—original draft, Y.-Z. C and X.-Y. H; formal analysis, Y.-Z. C and X.-L. Z; funding acquisition, D.-H. H; methodology, Y.-X. Z; data curation, Z.-F. M; resources, L.C; investigation, M.-Y. W; project administration, B.-X. Q; interpretation, writing—review and editing, Y.Y. and D.-H. H; supervision, D.-H. H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by Guangxi Science and Technology Projects (Grant number AB2506910035; AA23062051).

 doi:10.20944/preprints202512.0238.v1



**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author(s).

**Acknowledgments:** We would like to acknowledge Hechi Research Institute of Agricultural Sciences for providing HeFengDao445 and HeXiFengZhan2； acknowledge Guangxi Boshiyuan Seed Co., Ltd.

**Conflicts of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interests.

## References

1. Long-ping, Y. Development of hybrid rice to ensure food security. Rice science. 2014, 21, 1.
2. Habde, S.V.; Singh, S.K.; Singh, D.K.; Singh, A.K.; Sah, R.P.; Korada, M.; Khaire, A.R.; Majhi, P.K.; Singh, U.M.; Singh, V.K. QTL mapping reveals different set of candidate genes governing stable and location specific QTLs enhancing zinc and iron content in rice. Euphytica. 2024, 220, 179.
3. Wang, H.; Deng, X.W. Development of the "third-generation" hybrid rice in China. Genomics, proteomics & bioinformatics. 2018, 16, 393-396.
4. Peng, S.; Khush, G.S.; Virk, P.; Tang, Q.; Zou, Y. Progress in ideotype breeding to increase rice yield potential. Field Crops Research. 2008, 108, 32-38.
5. Zhou, D.; Chen, W.; Lin, Z.; Chen, H.; Wang, C.; Li, H.; Yu, R.; Zhang, F.; Zhen, G.; Yi, J. Pedigree-based analysis of derivation of genome segments of an elite rice reveals key regions during its breeding. Plant Biotechnology Journal. 2016, 14, 638-648.
6. Chen, S.; Lin, Z.; Zhou, D.; Wang, C.; Li, H.; Yu, R.; Deng, H.; Tang, X.; Zhou, S.; Wang Deng, X. Genome-wide study of an elite rice pedigree reveals a complex history of genetic architecture for breeding improvement. Scientific reports. 2017, 7, 45685.
7. Abe, A.; Kosugi, S.; Yoshida, K.; Natsume, S.; Takagi, H.; Kanzaki, H.; Matsumura, H.; Yoshida, K.; Mitsuoka, C.; Tamiru, M. Genome sequencing reveals agronomically important loci in rice using MutMap. Nature biotechnology. 2012, 30, 174-178.
8. Devanna, B.N.; Sucharita, S.; Sunitha, N.; Anilkumar, C.; Singh, P.K.; Pramesh, D.; Samantaray, S.; Behera, L.; Katara, J.L.; Parameswaran, C. Refinement of rice blast disease resistance QTLs and gene networks through meta-QTL analysis. Scientific Reports. 2024, 14, 16458.
9. Khahani, B.; Tavakol, E.; Shariati, V.; Rossini, L. Meta-QTL and ortho-MQTL analyses identified genomic regions controlling rice yield, yield-related traits and root architecture under water deficit conditions. Scientific reports. 2021, 11, 6942.
10. Tanin, M.J.; Saini, D.K.; Sandhu, K.S.; Pal, N.; Gudi, S.; Chaudhary, J.; Sharma, A. Consensus genomic regions associated with multiple abiotic stress tolerance in wheat and implications for wheat breeding. Scientific reports. 2022, 12, 13680.
11. ZHANG, N.; Jun, J.; Fa, C. Dissecting the key genomic regions underlying high yield potential in common wheat variety 'Kenong 9204'. Journal of Integrative Agriculture. 2023, 22, 2603-2616.
12. Sansaloni, C.; Franco, J.; Santos, B.; Percival-Alwyn, L.; Singh, S.; Petroli, C.; Campos, J.; Dreher, K.; Payne, T.; Marshall, D. Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. Nature communications. 2020, 11, 4572.
13. Stein, J.C.; Yu, Y.; Copetti, D.; Zwickl, D.J.; Zhang, L.; Zhang, C.; Chougule, K.; Gao, D.; Iwata, A.; Goicoechea, J.L. Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza. Nature genetics. 2018, 50, 285-296.
14. Tanger, P.; Klassen, S.; Mojica, J.P.; Lovell, J.T.; Moyers, B.T.; Baraoidan, M.; Naredo, M.E.B.; McNally, K.L.; Poland, J.; Bush, D.R. Field-based high throughput phenotyping rapidly identifies genomic regions controlling yield components in rice. Scientific reports. 2017, 7, 42839.
15. Breseghello, F.; Coelho, A.S.G. Traditional and modern plant breeding methods with examples in rice (Oryza sativa L.). Journal of agricultural and food chemistry. 2013, 61, 8277-8286.
16. Hoque, A.; Anderson, J.V.; Rahman, M. Genomic prediction for agronomic traits in a diverse Flax (Linum usitatissimum L.) germplasm collection. Scientific Reports. 2024, 14, 3196.

 doi:10.20944/preprints202512.0238.v1




17. Li, A.; Hao, C.; Wang, Z.; Geng, S.; Jia, M.; Wang, F.; Han, X.; Kong, X.; Yin, L.; Tao, S. Wheat breeding history reveals synergistic selection of pleiotropic genomic sites for plant architecture and grain yield. Molecular Plant. 2022, 15, 504-519.
18. Chen, J.; Zhang, H.; Deng, S.; Du, H.; Chen, Z.; Zhao, Y.; Dai, D.; Liang, C.; Li, X.; Liang, C. A backbone parent contributes core genomic architecture to pedigree breeding of early-season indica rice. Journal of genetics and genomics= Yi chuan xue bao. 2021, 48, 1040-1043.
19. Huang, J.; Li, J.; Zhou, J.; Wang, L.; Yang, S.; Hurst, L.D.; Li, W.-H.; Tian, D. Identifying a large number of high-yield genes in rice by pedigree analysis, whole-genome sequencing, and CRISPR-Cas9 gene knockout. Proceedings of the National Academy of Sciences. 2018, 115, E7559-E7567.
20. Yun, R.; Dan, C.; Wen-jie, L.; Luo, T.; Guo-qiang, Y.; Ye, C.; Xue-mei, L.; Qiming, D.; Shi-quan, W.; Jun, Z. Genome-wide pedigree analysis of elite rice Shuhui 527 reveals key regions for breeding. Journal of Integrative Agriculture. 2021, 20, 35-45.
21. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research. 2010, 20, 1297-1303.
22. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989, 123, 585-595.
23. Hutter, S.; Vilella, A.J.; Rozas, J. Genome-wide DNA polymorphism analyses using VariScan. BMC bioinformatics. 2006, 7, 409.
24. Kongdin, M.; Mahong, B.; Lee, S.-K.; Shim, S.-H.; Jeon, J.-S.; Ketudat Cairns, J.R. Action of multiple rice β-glucosidases on abscisic acid glucose ester. International Journal of Molecular Sciences. 2021, 22, 7593.
25. Terao, T.; Hirose, T. Control of grain protein contents through SEMIDWARF1 mutant alleles: sd1 increases the grain protein content in Dee-geo-woo-gen but not in Reimei. Molecular Genetics and Genomics. 2015, 290, 939-954.
26. Song, W.-Y.; Wang, G.-L.; Chen, L.-L.; Kim, H.-S.; Pi, L.-Y.; Holsten, T.; Gardner, J.; Wang, B.; Zhai, W.-X.; Zhu, L.-H. A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. science. 1995, 270, 1804-1806.
27. Ashikari, M.; Sakakibara, H.; Lin, S.; Yamamoto, T.; Takashi, T.; Nishimura, A.; Angeles, E.R.; Qian, Q.; Kitano, H.; Matsuoka, M. Cytokinin oxidase regulates rice grain production. Science. 2005, 309, 741-745.
28. Tian, Z.; Qian, Q.; Liu, Q.; Yan, M.; Liu, X.; Yan, C.; Liu, G.; Gao, Z.; Tang, S.; Zeng, D. Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. Proceedings of the National Academy of Sciences. 2009, 106, 21760-21765.
29. Li, Y.; Fan, C.; Xing, Y.; Jiang, Y.; Luo, L.; Sun, L.; Shao, D.; Xu, C.; Li, X.; Xiao, J. Natural variation in GS5 plays an important role in regulating grain size and yield in rice. Nature genetics. 2011, 43, 1266-1269.
30. Jiao, Y.; Wang, Y.; Xue, D.; Wang, J.; Yan, M.; Liu, G.; Dong, G.; Zeng, D.; Lu, Z.; Zhu, X. Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. Nature genetics. 2010, 42, 541-544.
31. Yu, B.; Lin, Z.; Li, H.; Li, X.; Li, J.; Wang, Y.; Zhang, X.; Zhu, Z.; Zhai, W.; Wang, X. TAC1, a major quantitative trait locus controlling tiller angle in rice. The Plant Journal. 2007, 52, 891-898.
32. Monna, L.; Kitazawa, N.; Yoshino, R.; Suzuki, J.; Masuda, H.; Maehara, Y.; Tanji, M.; Sato, M.; Nasu, S.; Minobe, Y. Positional cloning of rice semi dwarfing gene, sd-1: rice "green revolution gene" encodes a mutant enzyme involved in gibberellin synthesis. DNA research. 2002, 9, 11-17.
33. Li, M.; Tang, D.; Wang, K.; Wu, X.; Lu, L.; Yu, H.; Gu, M.; Yan, C.; Cheng, Z. Mutations in the F-box gene LARGER PANICLE improve the panicle architecture and enhance the grain yield in rice. Plant biotechnology journal. 2011, 9, 1002-1013.
34. Wan, X.; Weng, J.; Zhai, H.; Wang, J.; Lei, C.; Liu, X.; Guo, T.; Jiang, L.; Su, N.; Wan, J. Quantitative trait loci (QTL) analysis for rice grain width and fine mapping of an identified QTL allele gw-5 in a recombination hotspot region on chromosome 5. Genetics. 2008, 179, 2239-2252.
35. Zhou, Y.; Li, S.; Qian, Q.; Zeng, D.; Zhang, M.; Guo, L.; Liu, X.; Zhang, B.; Deng, L.; Liu, X. BC10, a DUF266-containing and Golgi-located type II membrane protein, is required for cell-wall biosynthesis in rice (Oryza sativa L.). The Plant Journal. 2009, 57, 446-462.

 doi:10.20944/preprints202512.0238.v1




36. Fang, L.; Zhao, F.; Cong, Y.; Sang, X.; Du, Q.; Wang, D.; Li, Y.; Ling, Y.; Yang, Z.; He, G. Rolling-leaf14 is a 2OG-Fe (II) oxygenase family protein that modulates rice leaf rolling by affecting secondary cell wall formation in leaves. Plant biotechnology journal. 2012, 10, 524-532.
37. Nakashima, K.; Tran, L.S.P.; Van Nguyen, D.; Fujita, M.; Maruyama, K.; Todaka, D.; Ito, Y.; Hayashi, N.; Shinozaki, K.; Yamaguchi-Shinozaki, K. Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice. The Plant Journal. 2007, 51, 617-630.
38. Dong, H.; Zhao, H.; Xie, W.; Han, Z.; Li, G.; Yao, W.; Bai, X.; Hu, Y.; Guo, Z.; Lu, K. A novel tiller angle gene, TAC3, together with TAC1 and D2 largely determine the natural variation of tiller angle in rice cultivars. PLoS genetics. 2016, 12, e1006412.
39. Furukawa, T.; Maekawa, M.; Oki, T.; Suda, I.; Iida, S.; Shimada, H.; Takamure, I.; Kadowaki, K.i. The Rc and Rd genes are involved in proanthocyanidin synthesis in rice pericarp. The Plant Journal. 2007, 49, 91-102.
40. Tang, W.; Wu, T.; Ye, J.; Sun, J.; Jiang, Y.; Yu, J.; Tang, J.; Chen, G.; Wang, C.; Wan, J. SNP-based analysis of genetic diversity reveals important alleles associated with seed size in rice. BMC plant biology. 2016, 16, 93.
41. Leng, Y.; Gao, Y.; Chen, L.; Yang, Y.; Huang, L.; Dai, L.; Ren, D.; Xu, Q.; Zhang, Y.; Ponce, K. Using Heading date 1 preponderant alleles from indica cultivars to breed high-yield, high-quality japonica rice varieties for cultivation in south China. Plant Biotechnology Journal. 2020, 18, 119-128.
42. Wu, W.; Zheng, X.-M.; Lu, G.; Zhong, Z.; Gao, H.; Chen, L.; Wu, C.; Wang, H.-J.; Wang, Q.; Zhou, K. Association of functional nucleotide polymorphisms at DTH2 with the northward expansion of rice cultivation in Asia. Proceedings of the National Academy of Sciences. 2013, 110, 2775-2780.
43. Si, L.; Chen, J.; Huang, X.; Gong, H.; Luo, J.; Hou, Q.; Zhou, T.; Lu, T.; Zhu, J.; Shangguan, Y. OsSPL13 controls grain size in cultivated rice. Nature genetics. 2016, 48, 447-456.
44. Cai, J.; Liao, Q.; Dai, Z.; Zhu, H.; Zeng, R.; Zhang, Z.; Zhang, G.-Q. Allelic differentiations and effects of the Rf3 and Rf4 genes on fertility restoration in rice with wild abortive cytoplasmic male sterility. Biologia plantarum. 2013, 57, 274-280.
45. Mizuta, Y.; Harushima, Y.; Kurata, N. Rice pollen hybrid incompatibility caused by reciprocal gene loss of duplicated genes. Proceedings of the National Academy of Sciences. 2010, 107, 20417-20422.
46. Kubo, T.; Takashi, T.; Ashikari, M.; Yoshimura, A.; Kurata, N. Two tightly linked genes at the hsa1 locus cause both F1 and F2 hybrid sterility in rice. Molecular plant. 2016, 9, 221-232.
47. Xiong, H.; Yu, J.; Miao, J.; Li, J.; Zhang, H.; Wang, X.; Liu, P.; Zhao, Y.; Jiang, C.; Yin, Z. Natural variation in OsLG3 increases drought tolerance in rice by inducing ROS scavenging. Plant Physiology. 2018, 178, 451-467.
48. Ochiai, K.; Shimizu, A.; Okumoto, Y.; Fujiwara, T.; Matoh, T. Suppression of a NAC-like transcription factor gene improves boron-toxicity tolerance in rice. Plant physiology. 2011, 156, 1457-1463.
49. Fu, J.; Wu, H.; Ma, S.; Xiang, D.; Liu, R.; Xiong, L. OsJAZ1 attenuates drought resistance by regulating JA and ABA signaling in rice. Frontiers in plant science. 2017, 8, 2108.
50. Liu, C.; Schläppi, M.R.; Mao, B.; Wang, W.; Wang, A.; Chu, C. The bZIP 73 transcription factor controls rice cold tolerance at the reproductive stage. Plant Biotechnology Journal. 2019, 17, 1834-1849.
51. Liu, M.; Zhang, S.; Hu, J.; Sun, W.; Padilla, J.; He, Y.; Li, Y.; Yin, Z.; Liu, X.; Wang, W. Phosphorylation-guarded light-harvesting complex II contributes to broad-spectrum blast resistance in rice. Proceedings of the National Academy of Sciences. 2019, 116, 17572-17577.
52. Liu, M.; Fan, F.; He, S.; Guo, Y.; Chen, G.; Li, N.; Li, N.; Yuan, H.; Si, F.; Yang, F. Creation of elite rice with high-yield, superior-quality and high resistance to brown planthopper based on molecular design. Rice. 2022, 15, 17.
53. Reinke, R.; Kim, S.-M.; Kim, B.-K. Developing Japonica rice introgression lines with multiple resistance genes for brown planthopper, bacterial blight, rice blast, and rice stripe virus using molecular breeding. Molecular Genetics and Genomics. 2018, 293, 1565-1575.

 doi:10.20944/preprints202512.0238.v1




54. Lei, L.; Zheng, H.; Wang, J.; Liu, H.; Sun, J.; Zhao, H.; Yang, L.; Zou, D. Genetic dissection of rice (Oryza sativa L.) tiller, plant height, and grain yield based on QTL mapping and meta-analysis. Euphotic. 2018, 214, 109.
55. Yang, J.; Hu, C.; Hu, H.; Yu, R.; Xia, Z.; Ye, X.; Zhu, J. QTL Network: mapping and visualizing genetic architecture of complex traits in experimental populations. Bioinformatics. 2008, 24, 721-723.
56. Fu, Y.; Zhong, X.; Pan, J.; Liang, K.; Liu, Y.; Peng, B.; Hu, X.; Huang, N. QTLs identification for nitrogen and phosphorus uptake-related traits using ultra-high density SNP linkage. Plant Science. 2019, 288, 110209.
57. Gu, R.; Chen, F.; Long, L.; Cai, H.; Liu, Z.; Yang, J.; Wang, L.; Li, H.; Li, J.; Liu, W. Enhancing phosphorus uptake efficiency through QTL-based selection for root system architecture in maize. Journal of Genetics and Genomics. 2016, 43, 663-672.
58. Han, B.; Cui, D.; Ma, X.; Cao, G.; Zhang, H.; Koh, H.J.; Han, L. Evidence for evolution and selection of drought-resistant genes based on high-throughput resequencing in weedy rice. Journal of Experimental Botany. 2022, 73, 1949-1962.
59. Li, H.; Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics. 2010, 26, 589-595.
60. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; Subgroup, G.P.D.P. The sequence alignment/map format and SAMtools. bioinformatics. 2009, 25, 2078-2079.
61. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. fly. 2012, 6, 80-92.
62. Crossa, J.; Campos, G.d.l.; Pérez, P.; Gianola, D.; Burgueno, J.; Araus, J.L.; Makumbi, D.; Singh, R.P.; Dreisigacker, S.; Yan, J. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics. 2010, 186, 713-724.