**Citation:** Kosmowski, F., Visaria, S., Stevenson, J., Gimode, D. & Platten, J. D. (2026). Farmers More Likely to Adopt Rice Varieties with Higher Density of Quantitative Trait Loci (QTL) in Viet Nam. Scientific Reports, doi: 10.1038/s41598-026-44331-9

Scientific Reports

https://doi.org/10.1038/s41598-026-44331-9

Article in Press

# Farmers more likely to adopt rice varieties with higher density of quantitative trait loci (QTL) in Viet Nam

Received: 27 October 2025

Accepted: 11 March 2026

Published online: 08 April 2026

Cite this article as: Kosmowski F., Visaria S., Stevenson J. *et al.* Farmers more likely to adopt rice varieties with higher density of quantitative trait loci (QTL) in Viet Nam. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-44331-9

**Frédéric Kosmowski, Sujata Visaria, James Stevenson, Davis Gimode & John Damien Platten**

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

# Farmers More Likely to Adopt Rice Varieties with Higher Density of Quantitative Trait Loci (QTL) in Viet Nam

Frédéric Kosmowski[a]*, Sujata Visaria[ac], James Stevenson[a], Davis Gimode[a], and John Damien Platten[b]

[a] CGIAR Standing Panel on Impact Assessment

[b] International Rice Research Institute

[c] City St. George's, University of London

* Email: f.kosmowski@cgiar.org

# Abstract

This paper offers the first nationwide analysis of Vietnamese rice varieties by combining DNA-based identification with quantitative trait loci (QTLs) and household survey data. Using nationally representative data from 2022, we found that 51% of rice farmers grew improved varieties. These varieties contained significantly more beneficial QTLs associated with yield, grain quality, and resistance to biotic and abiotic stresses than genetically unidentified varieties. On average, the improved varieties cultivated had been released 14 years prior to 2022. Farmer's socioeconomic characteristics correlated with adoption patterns: belonging to an ethnic minority or residing in a government-classified poor commune significantly reduced the likelihood of growing an improved variety. Among adopters, varietal traits were further associated with specific adoption choices. Each additional trait-related QTL was associated with a 0.9 percentage point increase in a province's adoption rate. Traits conferring tolerance to abiotic stress were positively associated with adoption, suggesting farmers may prefer varieties that enhance resilience to environmental stressors.

Keywords: Adoption; QTL; Marker-assisted selection; Molecular breeding; Rice; Viet Nam;

# Introduction

Rice has been cultivated in Viet Nam for over 4000 years. This staple crop underpins its economy and food security. Since the *Đổi Mới* reforms of 1986, a transition began that moved rice agriculture from subsistence farming to market-oriented production. Today Viet Nam is the world's second-largest rice exporter, and Vietnamese rice is exported to high-quality markets such as Japan, the European Union, and the United States. In 2023, Viet Nam exported 8.29 million tons of rice, 90% of which originated from the Mekong River Delta[1].

Since the early 1990s, molecular breeding, particularly the use of molecular markers, has played an increasingly important role in enhancing the traits in rice varieties that determine growth rates, yield and adaptability. Marker-Assisted Selection (MAS) and Marker-Assisted Backcrossing (MABC) allow breeders to identify and more precisely incorporate favorable genetic regions into elite cultivars[2]. Central to these approaches is the identification of Quantitative Trait Loci (QTLs)—specific genomic regions associated with variation in complex traits[3–5]. QTL analysis aims to determine the number, location, interactions, and effects of these loci[6].

Viet Nam has embraced these tools rapidly. Since 1984, 565 improved rice varieties have been certified nationally, of which 413 were released in the last two decades. Modern breeding programs in Viet Nam increasingly employ QTL pyramiding, a strategy that combines multiple QTLs associated with desirable traits into a single genetic background to enhance performance across the dimensions of productivity, quality, and resilience[7]. Breeding in Viet Nam has targeted yield-related traits, such as plant vigor, tillering, panicle length, and grains per panicle[8], as well as shorter growth durations to facilitate multiple cropping cycles. In addition, there has been a focus on quality traits, such as low amylose content and aroma, in line with consumer preferences both domestically and in export markets[9]. There has also been an effort to breed varieties that are resistant to pests (e.g., brown planthoppers) and diseases (e.g., blast, bacterial leaf blight)[10,11], and those that can tolerate abiotic stresses such as salinity, submergence[12], and drought[13].

However, we have limited evidence about Vietnamese farmers' adoption of improved rice varieties and their genetic composition. A report by the Ministry of Agriculture and Rural Development (MARD)[14] identified 12 major rice varieties accounting for 47% of Viet Nam's rice-growing area, led by IR50404 (13%), OM5451 (7%), and OM4900 (5%). Similarly, a 2021 MARD survey found that five varieties were cultivated on over 100,000 hectares in Southern Viet Nam. The General Statistics Office (GSO) data also highlight OM5451, Bac Thom 7, and IR50404 as the dominant varieties (GSO, 2022). Through expert elicitation in the Mekong River Delta [15] estimated that 44% of households planted salt-tolerant varieties[15]. Rather than relying on farmers' reports of the variety they planted, which is known to be prone to substantial measurement error, this study uses DNA-based varietal identification of the crops that they grow [16,17]. The DNA data also provides insights into the genetic architecture, particularly the QTL content, of the rice varieties in farmers' fields.

We examine whether farmers are more likely to be growing rice varieties that contain more QTLs. There are good reasons why this may be the case. QTL pyramiding enables the integration of multiple beneficial traits into a single variety, so that varieties containing more QTLs likely have many of the traits that farmers value[18,19]. Multi-QTL varieties generally offer greater yield stability and profitability and may provide better tolerance to biotic and abiotic stresses. In addition, if farmers' produce meets the quality traits associated with many of these QTLs, they are more likely to have access to premium markets[20]. A QTL count is therefore a reliable proxy for breeding success, as it signifies the cumulative enhancement of a variety's adaptive and productive capabilities.

Yet there are also good reasons to doubt whether farmers would preferentially plant varieties that contain more QTLs. A growing literature has documented the fact that especially in low and middle income countries, farmers are often unaware of the (promised) traits of the varieties they are cultivating[17,21]. This could be due to a lack of information but it could also occur if the phenotypic traits fail to appear in the plants—for instance, when specific genes are expressed only under particular environmental conditions. Varieties may also not be disseminated wide enough for farmers to have access to them.

This study addresses four research questions: (1) What is the current genetic composition of rice varieties cultivated by Vietnamese farmers?; (2) Which socioeconomic factors are associated with the adoption of trait-related Quantitative Trait Loci (QTLs)?; (3) Is the density of QTLs in rice varieties related to observed adoption patterns?; and (4) Which specific rice traits are associated with higher adoption levels?

By assessing farmers’ varietal adoption using the underlying genetic information about their varieties, this study provides the first robust assessment of the reach of genetic improvements in Viet Nam’s rice cultivation.

# Materials and Methods

## Survey design and rice sample collection

This study leverages data from the Viet Nam Household Living Standards Survey (VHLSS), conducted annually by Viet Nam’s General Statistics Office. The VHLSS follows a stratified two-stage cluster sampling design and is nationally and regionally representative. In 2022, the survey covered approximately 46,995 households, including 13,650 rice-growing ones.

Although the VHLSS typically does not collect plot-level data or conduct direct field observations, through a collaboration with the CGIAR’s Standing Panel on Impact Assessment, a new rice crop-sampling module was introduced in the 2022 survey. This module was administered to a random subsample of 832 rice-growing households, selected by odd-numbered GSO household IDs. Up to five households were selected per enumeration area.

In this exercise, trained enumerators identified plots that contained rice plants that were at least 20 days old. One plot was selected per household. One rice plant was selected from each of four randomly placed quadrants in this plot, and one leaf per plant was sampled using a hole punch. The samples from the four plants were then bulked, stored in silica gel. Each plot-level sample was barcoded, and samples were shipped to Hanoi.

Due to COVID-19-related delays in starting data collection, we were unable to spread the sample collection evenly across the year. The collected samples covered multiple cropping seasons: 73% from the winter–spring season, 11% from the summer–autumn season, and 16% from the autumn–winter season. The presented results represent household-level patterns of varietal use rather than production-weighted national averages.

In total, 832 samples were collected. Of these, 766 were successfully matched to household data. These samples were collected from 248 enumeration areas (EAs). The distribution across regions shows the Mekong River Delta most heavily sampled (n= 194 households), followed by the Red River Delta (n = 155), Northern Midlands & Mountain Areas (n = 178), North Central & Central Coastal Areas (n = 171), then smaller samples in the Central Highlands (n = 44) and Southeast (n = 24). This sample distribution appears aligned with the actual concentration of rice cultivation across Viet Nam.

## Reference library creation and genotyping protocol

To accurately identify a crop variety using genomic techniques, one requires a robust reference library. This allows a genetic matching of the rice samples from the farm to known cultivars[22]. For this study, we assembled a comprehensive reference set consisting of 122 unique rice cultivars officially released in Viet Nam and 174 breeding lines from the International Rice Research Institute (IRRI) in the Philippines.

To build this library, the team contacted 17 public and private Vietnamese institutions engaged in rice breeding and seed release and requested specific breeder seed samples. A total of 147 requests were made, and 99 samples were obtained from six organizations. The remaining requests were declined, primarily because the institutions had discontinued seed production or because the variety was temporarily unavailable due to seasonality. For cultivars whose breeder seeds could not be obtained, commercial seeds and grains were purchased from seed retailers. This added 23 unique cultivars to the reference library. The IRRI elite lines were genotyped in a previous project and incorporated into the library. A full list of the reference materials is provided in Table S1.

Genotyping of the collected farm-based rice leaf samples was conducted by Agriplex Genomics using the PlexSeq platform, which supports high-level multiplexing for mid-density Single Nucleotide Polymorphism (SNP) analysis. This study employed Version 4 of the IRRI Rice Custom Amplicon (RICA) SNP panel, as described by [23], comprising 1,024 SNP markers. This included 797 SNPs from the Cornell 6 K Infinium Rice Chip[24], 205 trait-linked SNPs, and 22 quality control SNPs. This high throughput genotyping approach supports precise varietal matching and enables downstream analysis of genetic traits linked to rice performance and adaptation.

## Variable measurements

### Measurement of rice varietal adoption

A key part of this study was the estimation of genomic relationships between rice samples from farmers' fields and known rice varieties present in reference materials. In total, 1,017 single nucleotide polymorphisms (SNPs) were used to genotype the rice samples, along with 297 reference samples designated for use as references in matching.

The genotypes were first filtered using a minimum call rate threshold of 0.5 to eliminate those with excessive missing data. The SNPs were then filtered to retain only robust and informative loci. This entailed removing markers with a low minor allele frequency (MAF < 0.01), excluding those with call rates below 0.8, and eliminating monomorphic markers. Additional quality metrics, including polymorphism information content (PIC), observed heterozygosity, and inbreeding coefficient, were evaluated to help understand the genetic properties of the samples and the information content of the markers used in downstream analyses.

After quality control, 903 SNPs were retained, including 772 general SNP markers and 131 trait-specific markers. However, these marker types were not distinguished during varietal assignments. Among the field samples, 789 passed quality control and were retained for downstream analysis. Reference samples that displayed high heterozygosity levels and low inbreeding coefficients were retained in the reference library since they may represent the true genetic makeup of some varieties in circulation.

To assign field samples to known varieties, pairwise genetic comparisons were conducted between the field and reference genotypes. This involved calculating percentage similarity (to assess purity) and identity-by-state (IBS) genetic distance (rendered as 1 minus IBS, with distances closer to zero indicating closer relationship). IBS analysis was done using the R package SNPRelate [25]. A field genotype was considered to be related to the reference if the IBS distance was <0.05. In cases where a sample matched multiple references (common due to close genetic similarity), the "Top Reference" was selected based on the lowest IBS score and the highest purity score, and this reference ID was assigned to the sample.

In total, 390 of the 766 field samples were successfully matched with a reference variety. For classification purposes, non-assigned samples (likely landraces) were grouped by region—Northern Midlands and Mountainous Area (NMMA); North Central Coast and Central Highlands (NCCCA); Red River Delta (RRD); Central Highlands (CH); Southeast Region (Southeast); and Mekong River Delta (MRD)—on the basis that these landraces are likely to share similar agro-ecological contexts, which may result in some degree of clustering.

## Measurement of rice QTLs

The presence of QTLs in the field samples was directly inferred from the allelic state of specific trait-linked SNP markers. The RICA V4 panel[23] was designed in collaboration with IRRI and includes 205 SNPs selected for their linkage to key major-effect QTL loci widely used in rice[26]. We assessed the presence of beneficial alleles based on these markers as diagnostic of trait-conferring QTLs in the background of improved varieties.

The QTLs we focused on in this study reflect the combined shifting priorities of rice breeding programs over the past several decades[8,27]. It relies on a standardized classification where the 'positive' [+] designation is assigned specifically to the allele or haplotype that confers the trait improvement desired by farmers. Trait-linked QTLs were grouped into categories based on the literature, as follows.

*Yield*: Eleven markers were used to assess traits related to grain number, panicle architecture, plant height, and flowering time. These included *Gn1a, Ehd1, RFT1, Hd1, GFR1, Hd2, Ghd7, NGR5, NAL1, Hd9/qDTY3.2* and *Hd3a*.

*Quality*: Key markers such as *SLG7, fgr, Alk, Chalk5, GS3, NAS3*, and *TGW6* play a role in determining grain quality by influencing its shape, starch composition, and nutritional content, thereby impacting consumer preference and market value.

*Biotic stress resistance*: These QTLs provide defense against major rice pathogens and pests, including rice blast (*Pita, qPi33*), bacterial blight (*Xa26* and *Xa4*), brown planthoppers (*Bph17* and *Bph32*), and tungro disease (*TSV1* and *STV11*).

*Abiotic stress tolerance*: These QTLs included drought tolerance (*qDTY1.1, DRO1, qDTY12.1, qDTY2.1, qDTY2.2, qDTY3.1, qDTY3.2 and qDTY4.1*), heat tolerance (*qHTSF4.1* and *TT1*), cold tolerance (*COLD1, qSCT1, qCTS10, and qPSST6*), salinity tolerance (*Saltol* and *qSIS1L*), submergence tolerance (conferred by *qSub1*, and *qAG3*, which specifically enables seeds to germinate and seedlings to establish under flooded, oxygen-deprived conditions).

To count the number of QTLs in each category that a household's rice plants had, genomic data were aggregated across samples per variety collected on farmers' fields. While trait-linked markers act as proxies for underlying functional genes, this approach has limitations when analyzing unassigned samples or landraces. The absence of specific marker alleles in landraces does not necessarily imply the absence of the associated traits. In landraces that have not undergone controlled breeding, natural recombination can disrupt marker–QTL associations, making direct comparisons with improved varieties more complex.

## Measurement of household's socio-economic characteristics

The information on socioeconomic characteristics of households was obtained from data collected by Viet Nam's General Statistics Office (GSO) through its VHLSS questionnaire modules. These characteristics included whether the household was headed by a female or an ethnic minority member, the age of the household head, and the highest education level completed by the household head. Additional indicators included annual household income (in million VND), the percentage of households located in poor communes, and the percentage of households with access to asphalt roads.

### Survey weights

In VHLSS 2022, rice crop-sampling was conducted with 766 randomly chosen households, without regard to the population size of rice-growing households within each commune, leading to a sample that is not representative of the population of rice growers. To address this issue, household weights in the VHLSS were adjusted to account for the fact that the rice leaf samples were taken from randomly chosen households, with a limit of 5 households per enumeration area (See Text S1).

## Empirical estimation

We aimed to understand whether farmers were more likely to adopt rice varieties that featured improved genetic traits, and if so, which traits were associated with adoption. To measure the genetic traits in the rice that farmers grew, we used the total number of trait-related QTLs (QTL count) from four trait categories: yield-related, grain quality, biotic stress resistance and abiotic stress tolerance. While our use of a count-based metric assumes that QTLs are equal despite differences in effect and environmental influence, in the absence of location-specific phenotype data, this serves as an indication of genetic enrichment as a result of allele pyramiding. This metric captures both QTL density and varietal improvements.

First, to examine whether the adoption of improved rice varieties is associated with households' socioeconomic characteristics, we estimated a probit regression model. The outcome variable included the seed status of the rice grown by the households (improved or not). Key independent variables included household-level characteristics.

Second, to test the hypothesis that varieties with a higher QTL count are more widely adopted by households, we estimated a series of linear regression models at the province-variety level. The dependent variable is the percentage of farming households in each province that adopted a given improved variety. Since adopters of improved rice varieties differ in both observable and unobservable characteristics from non-adopters, we restricted the following analysis to households that cultivated improved rice varieties. This restriction facilitates a more direct comparison with breeding outcomes, highlighting how modern genetic improvements relate to observed results.

In our baseline model, the only explanatory variable is the count of QTLs. We then sequentially add controls: regional fixed effects to account for agro-ecological and institutional differences, and next, province-level covariates. The province-level covariates—average years of education per household, average household income (in million VND), and a dummy for whether the main access road is asphalt—were aggregated from household data. All models were estimated using ordinary least squares (OLS) with the following specification:

$$\text{Adoption rate}_{vp} = \alpha + \beta \cdot \text{QTL Count}_{v} + \text{Region}_{p} + \gamma X_{p} + \epsilon_{vp}$$

where $v$ indexes rice varieties and $p$ indexes provinces. QTL $\text{Count}_v$ represents the number of specific trait loci identified per variety, $\text{Region}_p$ captures province-level fixed-effects, and $X_p$ is a vector of province-level controls.

Third, to assess which specific traits are linked to higher levels of adoption, we adopted the previous approach but instead used the number of QTLs linked to specific trait categories— yield-related traits, grain quality, biotic stress resistance, and abiotic stress tolerance as independent variables. The empirical analysis was performed using R version 4.4.2. [28]

All methods were carried out in accordance with relevant guidelines and regulations. The study protocol was reviewed and approved by the Institutional Review Board of Hanoi University of Public Health (Decision No. 732/QD-DHYTCC). Informed consent was obtained from all participants prior to data collection.

### Robustness checks

Robustness checks were conducted using two complementary approaches. First, we re-estimated the main specifications using only Winter-Spring season data, the largest sample in our dataset; the association of QTL count with adoption remained statistically significant ($p = .1$), with an effect size of similar magnitude (Table S5). This finding demonstrates that our main results are not driven by seasonality. Second, we used the average number of QTLs in the top three adopted varieties as an alternative dependent variable. The coefficient on QTL count remained positive and significant across all models (Table S6).

## Results

### Summary Statistics

We begin with summary statistics on rice varietal adoption in Viet Nam based on DNA fingerprinting data collected in tandem with VHLSS 2022. Figure 1 illustrates the share of different rice varieties grown by households in Viet Nam. The rice grown by nearly one-half (47%) of households could not be assigned to any of the references in the library. When visualized on a dendrogram (Figure S1), most of these formed distinct separate clusters, suggesting that they are not necessarily noisy versions of elite varieties, but likely genetically distinct traditional landraces or local varieties. In contrast, assigned samples which represented 53% of the households clustered with their corresponding reference varieties. These showed considerable heterogeneity, corresponding to 28 distinct improved rice varieties, with no single variety holding more than 8% of the share. The most widely cultivated varieties were BT7 (7.71%), OM4900 (7.01%), and Dai Thom 8 (5.91%), and these were followed by TBR225, Thien Uu 8, and KD18. Other varieties, such as OM5451, N98, and IR50404, represented smaller proportions. Most samples belonged to the indica subspecies, with only 1.8% classified within the japonica group[1]. Collectively, these data revealed considerable diversity in the improved rice cultivars currently planted in Viet Nam, with no mega varieties. By 2022, the improved rice varieties present in farmers' fields were released, on average, 14 years ago. Additional details, including age, pedigree, and origin of the identified cultivars, are provided in Table S2.

---

[1] These samples exhibited at least 90% similarity to the japonica varieties Asiminori and Taichung65.

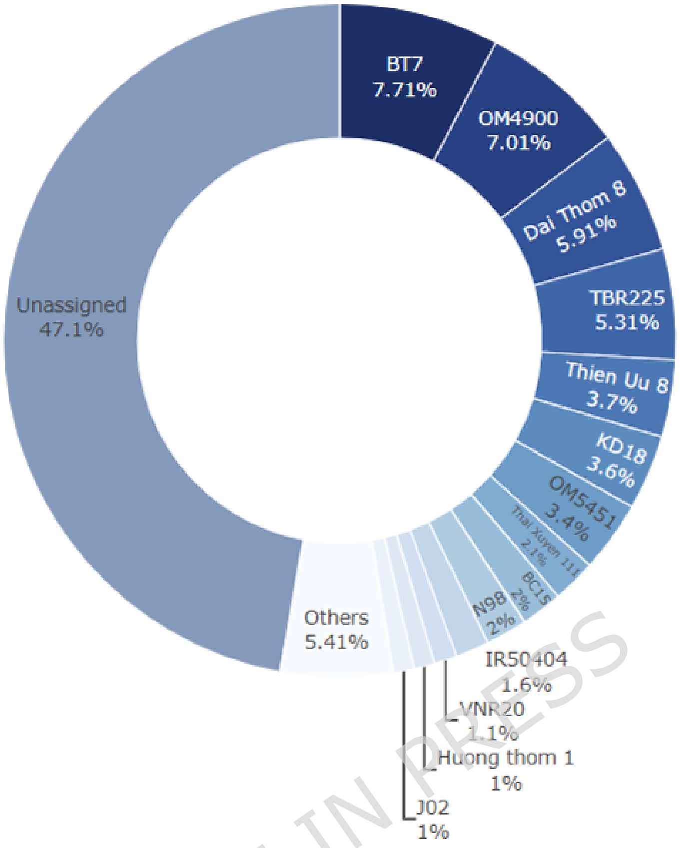


Figure 1. Distribution of Rice Varieties Grown by Farmers in Viet Nam in 2022. Each household in the sample is assigned a rice variety based on a DNA match between the rice sample collected from its plot and the reference library. The figure presents the percentage of households growing each rice variety identified. "Others" includes improved varieties found in less than 1% of households: HUONG UU 98, IR17494, N97, ST24, Q5, Jasmine 85, LH12, LTH31, RVT, VNR10, OM576, OM6162, OM7347, and SH14.

A key question is whether the varieties observed on farmers' fields were developed using marker-assisted selection (MAS or MABC), which has been promoted in Viet Nam since the 2010s [8]. Although some of the most popular varieties, such as OM4900 and its derivative Dai Thom 8, are direct or indirect outcomes of MAS, most widely adopted varieties—including BT7, KD18, TBR225, Tien Uu 8, Thai Xuyen 111, OM5451, and BC15—were either developed through conventional breeding or derived from elite imported lines. These varieties generally originate from genetic backgrounds already rich in high-quality alleles, a fact supported by our QTL screening.

The spatial distribution of improved variety adoption revealed additional patterns. As shown in Figure 2, improved varieties were more likely to have been adopted in the delta regions, characterized by intensive rice cultivation, and in the southeastern region of the country. OM4900 and OM5451 were particularly prevalent in the southern regions, whereas Dai Thom 8 and BC15 were the dominant varieties in the Red River Delta. Additionally, Figure S2b reveals spatial heterogeneity in the QTL density of rice varieties cultivated across Viet Nam. Provinces in the MRD and southeast regions tend to have rice varieties with higher mean numbers of QTLs. In contrast, several northern provinces and upland areas exhibited lower QTL count, consistent with their lower rates of adoption of improved varieties.

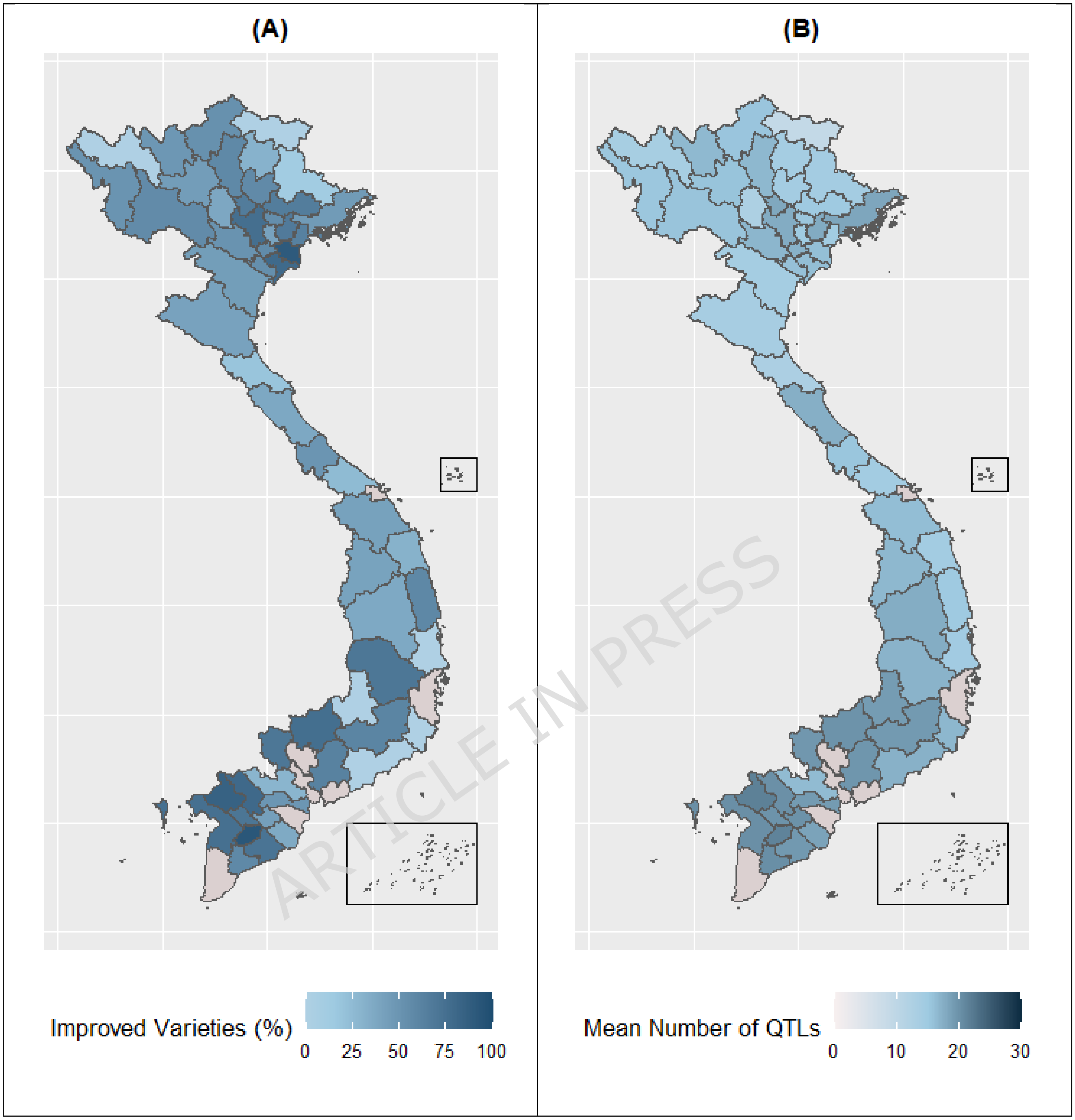


Figure 2. Spatial distribution of (A) improved rice varietal adoption and (B) the mean number of QTL count per rice variety in Viet Nam in 2022. Gray areas indicate provinces where no households grew rice at the time of rice sample collection. The map was generated using R version **4.2.3**. Source: VHLSS 2022.

At the national level, unassigned rice samples and those classified as improved varieties are present in roughly equal numbers, allowing for an effective comparison. In Table S3, we present the differences in QTLs between improved rice varieties and unassigned samples collected from farmers' fields. A t-test comparing the number of QTLs by seed status indicated that the average improved rice variety that farmers grew had a higher number of QTLs (mean = 18.27, range = 9-20) than the unassigned samples (mean = 13.03, range = 10- 17), with the difference significantly different from zero (t = -27.41, df = 673.17, p < 0.0001). These descriptive patterns support the idea that improved rice varieties in Viet Nam possess more advantageous alleles, as expected from a successful breeding program.

## Correlates of improved rice adoption

To understand the correlates of improved rice adoption, we analyzed household-level characteristics associated with adoption. Table 1 reports marginal effects from a multivariate probit regression.

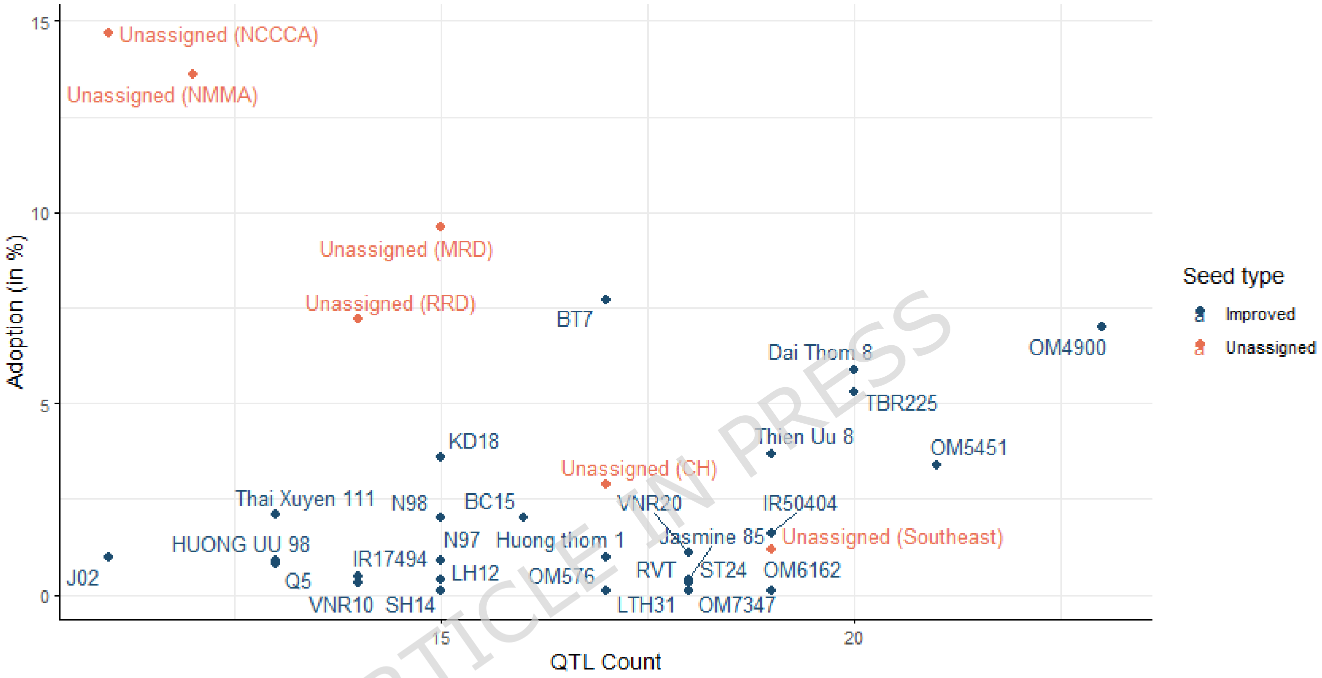
Table 1. MVP estimates of the relationship between improved rice varietal adoption and a set of household's socio-economic characteristics.

| | Marginal effects with 95% CI |
|---|---|
| Household head from an ethnic minority | -0.116 [-0.117, -0.115] |
| Household head is female | 0.020 [0.019, 0.022] |
| Age of household head (years) | 0.007 [0.007, 0.008] |
| Highest education level of household head | 0.008 [0.007, 0.008] |
| Total agricultural land area (hectares) | 0.062 [0.062, 0.062] |
| Commune classified as poor | -0.186 [-0.187, -0.184] |
| Main road surface is asphalt | -0.119 [-0.120, -0.118] |
| Observations | 719 |
| AIF | 9,866,433 |

The marginal effects in **Error! Reference source not found.** indicate that adoption patterns differ across social groups, with all estimates significant at the $p < 0.01$ level. Belonging to an ethnic minority or residing in a classified poor commune substantially reduces the probability of adopting improved varieties. By contrast, higher education and age are positively associated. Although the pooled model suggests female-headed households have slightly higher adoption probabilities, an interaction analysis reveals this effect is spatially heterogeneous. The positive coefficient is primarily driven by the NCCCA and South East regions; however, the latter result warrants caution as it is based on a restricted subsample. Notably, the association of female headship correlates with lower adoption probabilities in the RRD and MRD regions.

## Trait-related QTLs and Improved Rice Adoption

Given the correlation between farmer characteristics and adoption, we next examine whether adopters grew varieties with a broader suite of agronomic traits. To test this, we examine whether a higher density of trait-related QTLs in the rice variety grown is associated with an increased likelihood of adoption. We first analyzed the direct relationship between QTL count and varietal adoption at the national level (Figure 3). We then conduct regression analyses that progressively incorporate controls for potential confounding factors—first adding regional fixed effects and subsequently adjusting for province-level socioeconomic and infrastructure characteristics.



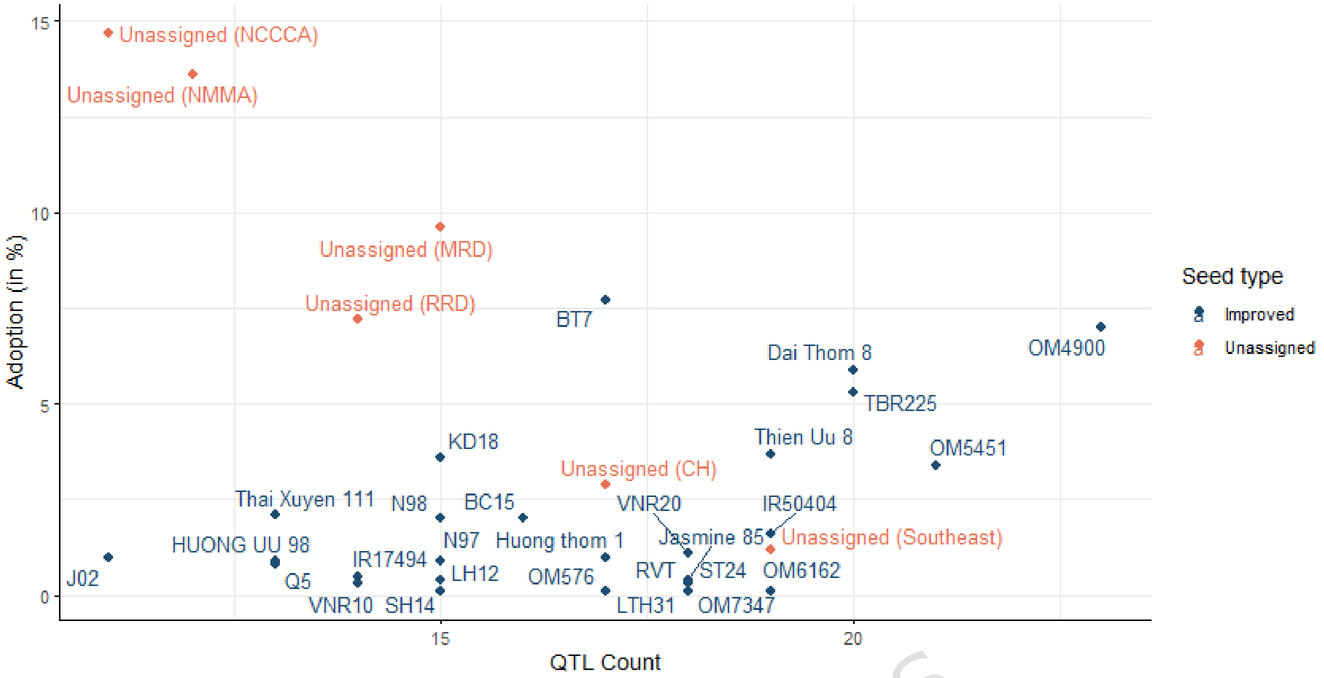


Figure 3. Patterns of Rice Adoption and Number of Trait-Related QTLs at the National Level in Viet Nam, 2022. The x-axis represents the number of QTLs, while the y-axis shows adoption rates at the national level. Unassigned samples were categorized according to their region of origin. NCCCA = North Central Coast and Central Highlands; NMMA = Northern Midlands and Mountainous Area; MRD = Mekong River Delta; RRD = Red River Delta; CH = Central Highlands; Southeast = Southeast Region.

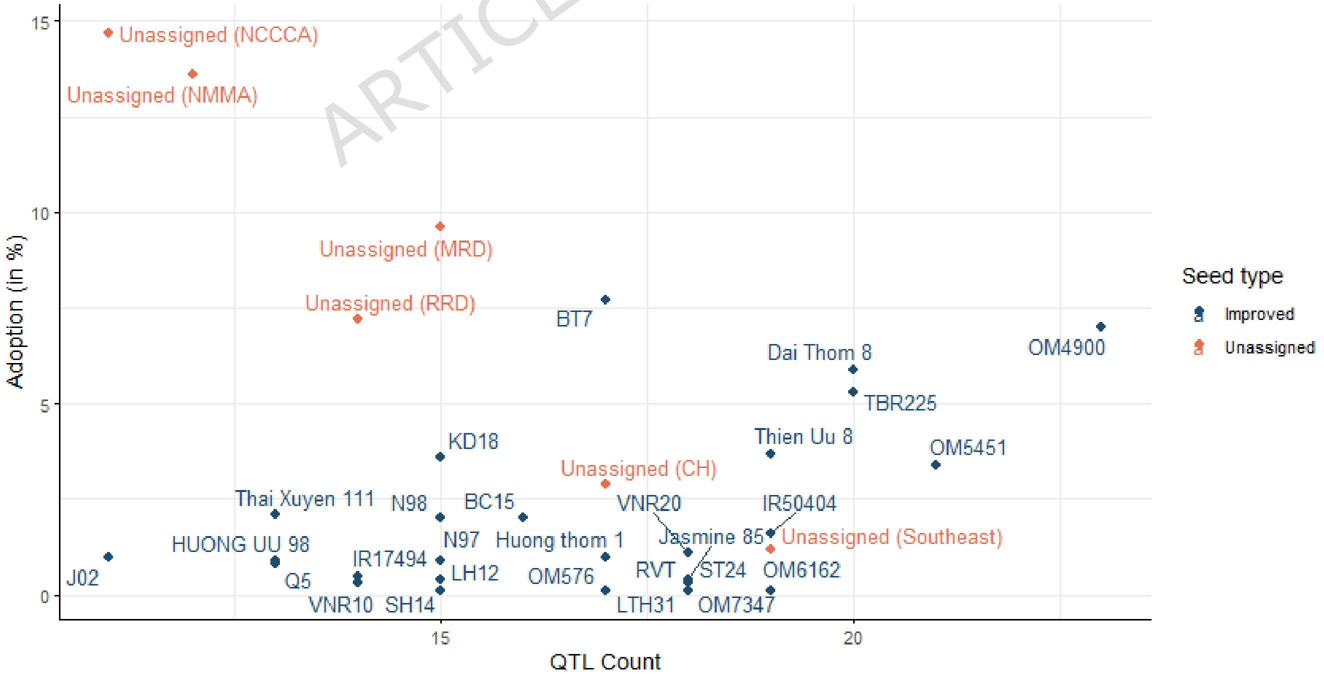


Figure 3 shows two clear patterns in the relationship between varietal adoption and the number of QTLs. In general, the unassigned samples tend to appear toward the left side of the x-axis, indicating that they possess fewer QTLs. Among unassigned samples, adoption rates are inversely correlated with number of QTLs. Despite this, they are found among a large proportion of the households surveyed in the Northern Central Coastal and North Midlands and Mountains regions (NCCCA and NMMA). These genetic entities possess a relatively small number yet diverse QTLs, such as *RFT1, Ehd1, Hd1*, and *Hd2* for flowering time and yield stability; *Chalk5, Alk, GS3, and DTH8* for grain quality; *Xa26* and *STV11* for biotic resistance; and *DRO1, qAG3*, and *HIS1* for abiotic stress tolerance. In contrast, improved varieties generally have a larger number of trait-related QTLs, reflecting greater genetic enrichment, but adoption rates tend to be lower, except for OM4900, BT7, and Dai Thom 8, which as noted earlier, are cultivated widely.

The relationship between province-level adoption rates and QTL count is further explored by aggregating the household-level data to the province level, controlling for regional fixed effects, and subsequently adjusting for province-level socioeconomic and infrastructure factors. The results from the most constrained model are shown in Figure 4. Overall, improved rice varieties with a higher number of QTLs display higher adoption rates at the province level. Each additional QTL in a variety was associated with a 0.9 percentage point increase in the provincial share of households adopting improved rice varieties. Across all three models, these associations are statistically significant (Table S4).

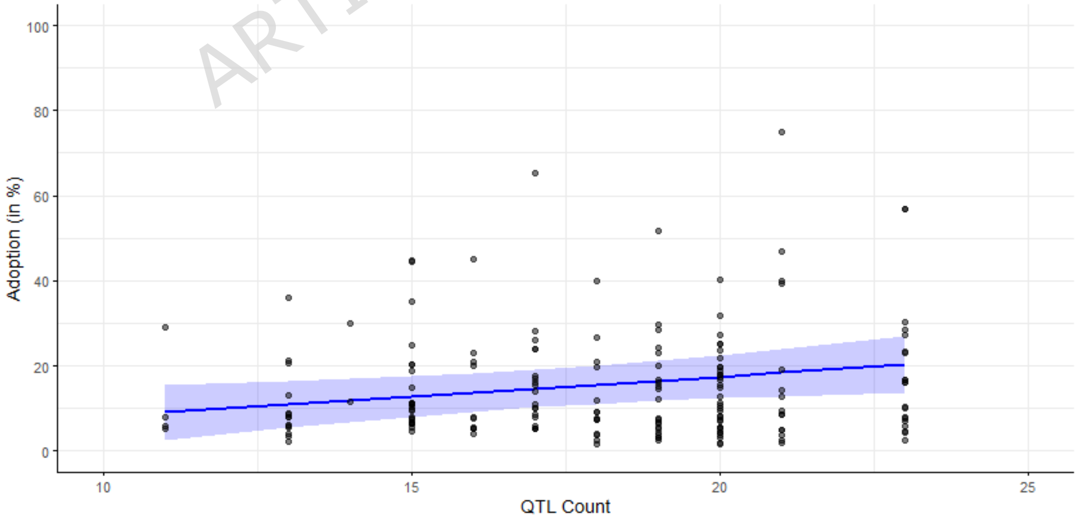


Figure 4. Marginal Effect of QTL count on Improved Varietal Adoption at the Province-Level. The adoption rate of an improved variety is measured as the percentage of farming households in each province that cultivate that variety. The linear regression model estimated the relationship between adoption rates and trait-related QTL count, controlling for

regional fixed effects and province-level covariates such as age of the variety, average education, household income, and infrastructure quality. The model was estimated using ordinary least squares (OLS) at the province level. $R^2 = 0.08$; $p < 0.05$.

## Trait Preferences in Household Adoption

Farmers tend to grow improved rice varieties that exhibit a broader set of agronomic traits. Our final analysis examines whether Vietnamese farmers show preferences for specific traits. In Figure 5 we present the overlap in adoption percentages across specific trait categories in improved rice varieties. The results indicate that multi-trait adoption is primarily driven by yield, with joint adoption of yield and another trait category exceeding 60% in all cases. By contrast, quality traits and abiotic stress tolerance traits appear as complementary choices, often adopted alongside yield but less frequently on their own.

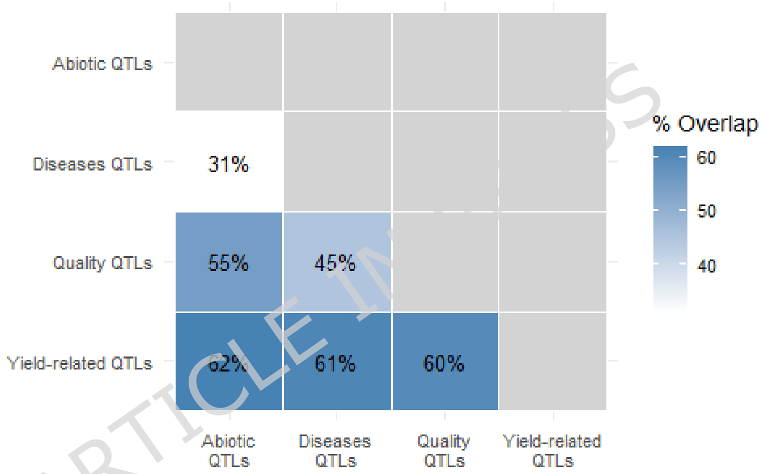


Figure 5. Percentage Overlap in Improved Rice Adoption among Four Key Rice Traits. For each trait, we used the median number of QTLs per variety as a threshold; only varieties with above-median QTL count for a given trait were considered in the corresponding column and cell of the cross-tab.

Building on these descriptive patterns, Table 2 reports results from a multivariate regression analysis assessing the association of different trait categories with improved rice adoption. This approach helps to isolate the relative importance of yield, grain quality, disease resistance, and abiotic stress tolerance traits in shaping varietal adoption, while controlling for household socioeconomic characteristics, infrastructure conditions, and regional fixed effects.

Table 2. Association Between QTL Categories and Adoption of Improved Rice Varieties at the Province Level (2022). Model 1 includes the total QTL Count as the primary independent variable. Model 2 incorporates regional fixed effects. Model 3 adds province-level covariates to the Model 2 specification. Standard errors are reported in parentheses below the

coefficients. NMMA = Northern Midlands and Mountainous Area; NCCCA = North Central Coast and Central Highlands; CH = Central Highlands; Southeast = Southeast Region. MRD = Mekong River Delta. Significance level: *p<0.1.

| | % of Households Adopting | | |
|---|---|---|---|
| **Variable** | **(1)** | **(2)** | **(3)** |
| QTL Count: Yield-related | 0.999<br>(1.244) | 1.327<br>(1.259) | 1.352<br>(1.274) |
| QTL Count: Quality | -0.399<br>(1.853) | -0.528<br>(1.875) | -0.399<br>(2.078) |
| QTL Count: Disease | -0.301<br>(0.767) | 0.118<br>(1.017) | -0.178<br>(1.066) |
| QTL Count: Abiotic | 1.961*<br>(1.155) | 1.980*<br>(1.175) | 1.967*<br>(1.183) |
| Region: NMMA | | 3.223<br>(2.731) | 3.603<br>(3.113) |
| Region: NCCCA | | 0.456<br>(3.007) | 1.685<br>(3.281) |
| Region: Central Highlands | | 4.258<br>(4.217) | 5.268<br>(4.953) |
| Region: South East | | 7.177<br>(5.155) | 4.81<br>(5.478) |
| Region: MRD | | -1.691<br>(3.594) | -3.757<br>(4.153) |
| Age of Variety (in yrs) | | | 0.089<br>(0.133) |
| Highest education level of household head | | | -1.064<br>(2.71) |
| Annual household income (million VND) | | | 0.038<br>(0.028) |
| Main road surface is asphalt | | | -3.3<br>(3.85) |
| Constant | 2.997<br>(7.979) | -0.816<br>(8.263) | -4.702<br>(10.465) |
| Observations | 178 | 178 | 178 |
| R-squared | 0.022 | 0.060 | 0.076 |

Abiotic traits are the only category showing a consistent positive association with adoption across Models (1) and (2) (~1.96, p = .1).

The coefficients remain positive and significant in Model (3), after controlling for socioeconomic and infrastructure variables. This finding suggests that farmers may place greater value on traits that enhance resilience to environmental stress. However, the wider confidence interval for abiotic traits (-2.18; 2.49) likely reflects environmental heterogeneity; these traits may be valued in stress-prone provinces but may be neutral in others. Notably, while the aggregate count of abiotic stress QTLs is a significant predictor, an analysis of individual QTLs (e.g., specific drought or salinity markers) yielded no statistically significant results (Table S7). In contrast, other breeding targets—including yield, grain quality, and disease resistance—showed no significant association with adoption in any model.

## Discussion

[2,29]The observed effect sizes of QTL count in this study align with existing literature, suggesting that while individual loci may have small to moderate impacts on complex traits, their cumulative stacking enhances crop performance[30,31]. Our findings confirm that targeted introgression is successfully reaching farmers' fields in Viet Nam, as evidenced by the presence of improved varieties containing these beneficial genetic markers. While QTL effect sizes can theoretically be subject to heterogeneity, our study operates on the empirical observation that the loci explored are consistently desirable across diverse production contexts. In rice breeding, evidence for significant epistasis or environmental interactions that negate the primary effect of established genes is remarkably limited; generally, these genes maintain stable expression across various genetic backgrounds and ecological conditions [32,33].

When placed in a regional context, Viet Nam's adoption landscape appears more dynamic than that of its neighbors. Previous DNA-based studies in Bangladesh have reported lower rates of improved variety adoption; for instance, [34] found that only 21% of households grew varieties that matched breeders' seed reference samples during the rainfed (Aman) season. More recent data from [35] indicates that while adoption has risen to 39% in the Aman season and 59% in the irrigated Boro season, the average varietal age remains between 21 and 24 years. This is significantly older than the 14-year average observed in Viet Nam. Furthermore, a defining characteristic of the Bangladeshi rice sector is the persistence of mega-varieties such as BRRI dhan29. These varieties, released in the 1990s, continue to dominate the landscape decades later. In contrast, our study found no such dominance of legacy mega-varieties in Viet Nam, suggesting a more fluid varietal turnover.

These findings carry significant policy implications; reducing the average varietal age has become a strategic priority for major agricultural economies seeking to accelerate genetic gain in the field. The Government of India, for instance, has codified this through the National Food Security Mission, which restricts subsidies for cultivars exceeding 10 years of age to incentivize the adoption of modern, climate-resilient alternatives[36]. Ultimately, a seed system's ability to deliver up-todate cultivars—combined with the integration of genetic tracking—is essential to ensuring that farming communities benefit from the latest agricultural innovations.

Although improved varieties with higher QTL densities show higher overall adoption rates, their distribution remains uneven across regions and socioeconomic groups. Our findings reveal that ethnic minority households and smallholders in government-classified poor communes are significantly less likely to access or adopt these cultivars. Climate change exacerbates this inequity, as these populations often reside in areas equally exposed to adverse weather events, yet they lack the genetic tools to mitigate such risks. These results point to an adoption gap possibly caused by restricted seed access, credit constraints, or a potential mismatch between high-performing varieties and the specific needs of marginalized environments. Without targeted interventions, genetic innovations risk bypassing the most vulnerable farmers.

To bridge this gap, a shift toward decentralized, farmer-centric models may prove effective. Participatory Varietal Selection (PVS) has been proposed as a critical solution to ensure that QTL pyramiding aligns with the specific trait preferences demanded by resource-constrained producers[37,38]. By involving farmers in the selection process, breeding programs can validate varietal performance under the marginal conditions typical of poor communes. Furthermore, the establishment of Village-Based Seed Enterprises (VBSE) offers a sustainable mechanism for reaching remote areas where commercial seed companies find operation unprofitable. As highlighted by [39], VBSEs can be effective at introducing stress-tolerant varieties. This strategy not only eliminates the high transportation costs that render improved seeds prohibitively expensive but also creates local income-generating opportunities.

A central question in this study is whether farmers have begun to prioritize environmental resilience over traditional productivity. Our findings provide only suggestive evidence that farmers may favor resilience traits; and this preference appears secondary to entrenched market demands. In Viet Nam, and specifically the Mekong River Delta, yield and grain quality likely remain the well-established pillars of variety selection. Consumer preferences are heavily weighted toward specific sensory attributes—such as slender shape, soft texture, and aroma [20,40]. Consequently, while the literature indicates that farmers value traits like salinity tolerance, they often face a difficult quantity-quality trade-off. Many current stress-tolerant varieties command lower market prices because their grain quality fails to match these established consumer standards[15]. This economic barrier suggests that the weak evidence for resilience preference found in our study may reflect a market constraint rather than a lack of interest: farmers cannot afford to prioritize abiotic stress tolerance unless it is bundled with the high-quality characteristics the market demands.

This study has several limitations that should be considered when interpreting the findings. First, the sampling frame restricted our ability to fully investigate seasonal patterns. In Viet Nam, particularly the Mekong River Delta, regions may experience up to three cropping seasons annually, each with distinct requirements for growth duration and biotic stress resistance. Consequently, our results may not capture the full diversity of trait preferences across different times of the year.

Furthermore, there is a risk of misclassification bias that complicates comparisons with improved varieties. Although our reference library was designed to be comprehensive—encompassing all released cultivars of known importance—it remains possible that some field samples represent improved varieties not currently included in our database. This lack of genomic information could lead to some degree of misclassification. Another limitation is the high level of heterozygosity observed, affecting 30% of field samples, which may reflect the presence of hybrid varieties, accidental admixture within plots, or contamination by weedy rice. The potential impact of heterozygosity on downstream QTL profile analysis was mitigated by aggregating genomic data across field samples assigned to a particular variety. This way, any potential noise from individual plot level admixture was minimized ensuring that the QTL count reflected the representative genetic makeup of the varieties in circulation.

Finally, our results represent a blended reflection of farmer preferences and seed accessibility. Because farmer choices are constrained by the availability of seeds in local markets, it is difficult to isolate inherent trait preferences from the logistical realities of seed supply. Therefore, the observed variety distribution likely indicates what is available and chosen by farmers, rather than a clean expression of farmers' demand.

Future research should explore the dissemination pathways of the on-farm performance of marker-assisted varieties that have not yet gained widespread traction, such as AS996 (released 2002) or HDT8 (released 2011). Longitudinal studies tracking varietal adoption over multiple seasons and environments could clarify seasonal effects. This study is unique in providing a molecular-level perspective on genetic gains in rice, offering a model that can be replicated in other rice growing nations as well as to other crops using established molecular platforms.  However, replicating the approach requires investment in construction of robust reference libraries of locally released varieties based on breeder seed for accurate varietal assignment. Also, QTL profiling should be done with care to ensure they reflect local realities such as quality traits desired by markets. Leveraging these genomic resources could provide broader insights into genetic improvements in smallholder farms.

# Conclusion

This study offers novel insights into the genetic improvements achieved in farmers' fields by examining the presence of trait-related QTLs in rice varieties cultivated across Viet Nam. Our results validate that improved varieties carry significantly more QTLs related to yield, grain quality, and stress resistance than the unassigned samples, supporting the effectiveness and scaling of targeted introgression through breeding programs. However, adoption remains uneven; socioeconomic constraints—primarily ethnic background and geographic location—were negatively associated with improved varietal adoption. Among those who have adopted improved rice, a higher concentration of beneficial QTLs strongly correlates with variety adoption, with suggestive evidence that farmers may favor traits that bolster resilience against environmental stressors. These results illustrate the power of integrating DNA-based varietal identification with socioeconomic data, offering a sound methodology for monitoring the impact of agricultural research and ensuring the equitable distribution of genetic gains.

# Data availability

All data and associated R scripts are stored in the OpenICPSR Repository: https://www.openicpsr.org/openicpsr/project/239565 under a Creative Commons Attribution 4.0 International (CC BY 4.0) License.

# References

1. IPSARD. *Rice Annual Report*. https://thitruongnongsan.gov.vn/images/2013/DANGNHAP/VnSAT_Lua%20gao/2023/N%C4%83m/Bao%20cao%20thuong%20nien%20lua%20gao%202023%20-%20EN.pdf (2023).

2. Collard, B. C. Y. & Mackill, D. J. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**, 557–572 (2008).

3. Geldermann, H. Investigations on inheritance of quantitative characters in animals by gene markers I. Methods. *Theoretical and Applied Genetics* **46**, 319–330 (1975).

4. Broman, K. W. & Sen, S. *A Guide to QTL Mapping with R/Qtl*. (Springer, New York, 2009).

5. Paterson, A. H. What has QTL mapping taught us about plant domestication? *New Phytologist* **154**, 591–608 (2002).

6. Miles, C. & Wayne, M. Quantitative trait locus (QTL) analysis. *Nature Education 1(1):208* (2010).

7. Muthu, V. *et al.* Pyramiding QTLs controlling tolerance against drought, salinity, and submergence in rice through marker assisted breeding. *PLoS One* **15**, (2020).

8. Khanh, T. D., Duong, V. X., Nguyen, P. C., Xuan, T. D. & Trung, N. T. Rice Breeding in Vietnam: Retrospects, Challenges and Prospects. *Agriculture* **11**, 1–21 (2021).

9. Pham, V. P., Huu, M. S. & Vo, C. T. Researching and selecting high quality rice varieties in Mekong Delta. *Can Tho Uni. J. Sci.* **15b**, 97–105 (2010).

10. Khoa, T. *et al.* Pyramiding the Candidate Genes of Rice Bacterial Leaf Blight Resistance xa5, Xa7 and xa13 into the Elite Rice Variety. *Journal of Scientific and Engineering Research* **4**, 92–98 (2017).

11. Nguyen Thi Pha & Nguyen Thi Lang. Marker Assisted Selection in Rice Breeding for Bacterial Leaf Blight. *Omonrice* **12**, 19–26 (2004).

12. Luu M.C. *et al.* Application of Marker Assisted Backcrossing to Introgress the Submergence Tolerance QTL Sub1 into the Vietnam Elite Rice Variety-AS996. *Am. J. Plant Sci.* **03**, 528–536 (2012).

13. Thị, V. *et al. Assessment of Genetic Diversity in Vietnamese Rice Using SSR Markers to Identify Cross Combinations for Developing Drought-Tolerant Cultivars*. (2013).

14. Khanh, T. D., Trung, K. H. & Hoi, P. X. Future perspectives and some approaches to develop rice in Vietnam. in *Biotechnology and Perspectives of Applications in Rice Breeding in Vietnam* (ed. Hanoi National University) 671–701 (Hanoi, 2018).

15. Paik, S. Y., Le, D. T. P., Nhu, L. T. & Mills, B. F. Salt-tolerant rice variety adoption in the Mekong River Delta: Farmer adaptation to sea-level rise. *PLoS One* **15**, 1–23 (2020).

16. Stevenson, J., Macours, K. & Gollin, D. The Rigor Revolution: New Standards of Evidence for Impact Assessment of International Agricultural Research. *Annu. Rev. Resour. Economics* **15**, (2023).

17. Stevenson, J., Gantier, M., Traxler, G. & Kosmowski, F. *The Challenge of Tracking the Reach of Post-Green Revolution Crop Breeding*. Preprint. DOI: 10.21203/rs.3.rs-3028333/v1 (2023).

18. Sandhu, N. *et al.* Positive interactions of major-effect QTLs with genetic background that enhances rice yield under drought. *Sci. Rep.* **8**, (2018).

19. Isnaini, I., Nugraha, Y., Baisakh, N. & Carsono, N. Toward Food Security in 2050: Gene Pyramiding for Climate-Smart Rice. *Sustainability* **15**, (2023).

20. Bairagi, S., Demont, M., Custodio, M. C. & Ynion, J. What drives consumer demand for rice fragrance? Evidence from South and Southeast Asia. *British Food Journal* **122**, 3473–3498 (2020).

21. Euler, M., Krishna, V. V, Jaleta, M. & Hodson, D. Because error has a price: A systematic review of the applications of DNA fingerprinting for crop varietal identification. *Outlook Agric.* **51**, 384–393 (2022).

22. Poets, A., Silverstein, K., Pardey, P., Hearne, S. & Stevenson, J. *DNA Fingerprinting for Crop Varietal Identification : Fit-for-Purpose Protocols and Their Cost and Analytical Implications*. (2020).

23. Arbelaez, J. D. *et al.* 1k-RiCA (1K-Rice Custom Amplicon) a novel genotyping amplicon-based SNP assay for genetics and breeding applications in rice. *Rice* **12**, (2019).

24. Thomson, M. J. *et al.* Large-scale deployment of a rice 6 K SNP array for genetics and breeding applications. *Rice* 10:40 (2017) doi:10.1186/s12284-017-0181-2.

25. Zheng, X. *et al.* A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328 (2012).

26. CGIAR. Rice mid-density genotyping service. *https://excellenceinbreeding.org/toolbox/services/rice-mid-density-genotyping-service* (2025).

27. McCouch, S. R. Gene nomenclature system for rice. *Rice* **1**, 72–84 (2008).

28. R Core Team. R: A Language and Environment for Statistical Computing. Preprint at (2024).

29. Mackay, T. F. C. The Genetic Architecture of Quantitative Traits. *Annu. Rev. Genet.* **35**, 303–339 (2001).

30. Paterson, A. H. *et al.* Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* **335**, 721–726 (1988).

31. Muthu, V. *et al.* Pyramiding QTLs controlling tolerance against drought, salinity, and submergence in rice through marker assisted breeding. *PLoS One* **15**, (2020).

32. Zuo, J. & Li, J. Molecular Genetic Dissection of Quantitative Trait Loci Regulating Rice Grain Size. *Annu. Rev. Genet.* **48**, 99–118 (2014).

33. Wing, R. A., Purugganan, M. D. & Zhang, Q. The rice genome revolution: from an ancient grain to Green Super Rice. *Nat. Rev. Genet.* **19**, 505–517 (2018).

34. Kretzschmar, T. *et al.* DNA fingerprinting at farm level maps rice biodiversity across Bangladesh and reveals regional varietal preferences. *Sci. Rep.* 1–12 (2018) doi:10.1038/s41598-018-33080-z.

35. Singla, S. *et al. SPIA Study 2025: Updating the Green Revolution*. https://iaes.cgiar.org/sites/default/files/pdf/SPIA_Bangladesh_Study_2025.pdf (2025).

36. NNFSN. *Operational Guidelines: National Food Security and Nutrition Mission*. (2024).

37. Sperling, L. *et al.* A framework for analyzing participatory plant breeding approaches and results. *Euphytica* (2001).

38. Nguyen, N. De & Kotaro Ohara. Participatory approaches for improving rice breeding in the Mekong Delta of Vietnam. in *Rice is life: scientific perspectives for the 21st century* (CABI, 2005).

39. MANZANILLA, D. O. *et al.* SOCIAL AND GENDER PERSPECTIVES IN RICE BREEDING FOR SUBMERGENCE TOLERANCE IN SOUTHEAST ASIA. *Exp. Agric.* **50**, 191–215 (2014).

40. Demont, M. & Rutsaert, P. Restructuring the Vietnamese rice sector: Towards increasing sustainability. *Sustainability (Switzerland)* **9**, (2017).

## Acknowledgements

We gratefully acknowledge the support of the General Statistics Office (GSO) of Viet Nam and the Ministry of Agriculture and Rural Development (MARD) during the period of research and data collection. We note that, following government restructuring in 2025, the GSO has since become the National Statistics Office (NSO), under the Ministry of Finance, and MARD has been merged into the Ministry of Environment and Agriculture (MEA). We thank Bùi Chí Bửu for very helpful comments.

## Author contributions

F.K., S.V., J.S., D.G. and J.P. contributed to the conception and design of the work and to the interpretation of data. F.K., D.G. and J.P. performed the analysis. F.K. drafted the manuscript. S.V., J.S., D.G. and J.P. reviewed and revised the manuscript. All authors approved the final version of the manuscript.

# Funding information

The authors gratefully acknowledge funding from the CGIAR System Council for the Viet Nam Country Study.

# Competing interests

The authors declare that they have no competing interests.