RESEARCH ARTICLE Indian Journal of Agricultural Research, Volume 60 Issue 7: 976-982 (July 2026)

Genetic Classification of Traditional Rice Varieties Collected from Vietnam using *matK* and *rbcL* DNA Barcode Markers

Nguyen Thanh Nhung[1], Le Thi Thu Trang[2], La Tuan Nghia[2], La Hoang Nhat Minh[2], Nguyen Nhu Toan[3], Nguyen Thi Phuong Doai[4], Khuat Huu Trung[4]

10.18805/IJARe.AF-1063

ABSTRACT

Background: Vietnam is home to a rich collection of traditional rice (*Oryza sativa* L.) landraces, many of which are increasingly threatened by modern agricultural practices and genetic erosion. This study aimed to genetically classify 90 traditional Vietnamese rice varieties collected from the Northern, Central and Southern regions of Vietnam using the chloroplast DNA barcode markers *matK* and *rbcL.*

Methods: Genomic DNA from 90 rice accessions was amplified using PCR with *matK* and *rbcL* primers, followed by sequencing and polymorphism analysis. Sequence variation, including single nucleotide polymorphisms (SNPs), transitions, transversions and indels, were identified. Phylogenetic relationships were analyzed using the Tamura-Nei model and UPGMA clustering to evaluate genetic diversity and relationships among the rice varieties.

Result: PCR amplification success rates reached 85.6% for *matK* and 100% for *rbcL*. Sequence analysis identified 11 SNPs in *matK* and 7 polymorphic sites in *rbcL*. Phylogenetic analysis grouped the rice varieties into two major clades for each marker. The *matK* marker provided higher resolution of intraspecific variation, distinguishing unique landraces such as Beo but bua, Nep bo and Ba la, whereas *rbcL* revealed generally low genetic differentiation but identified a divergent subgroup of eight genetically distinct accessions. Genetic distances ranged from 0.000-0.007 for *matK* and 0.000-0.004 for *rbcL*. These findings demonstrate that *matK* and *rbcL* are complementary molecular markers for genetic classification and varietal authentication of traditional Vietnamese rice, providing a valuable molecular basis for germplasm conservation and future rice breeding programs.

Key words: DNA barcode, *matK*, *rbcL*, Rice landraces.

INTRODUCTION

Rice (*Oryza sativa* L.) ranks as one of the most widely cultivated crops worldwide, serving as a staple food for over half of the global population (Mursyidin *et al*., 2021). Vietnam, a prominent rice-producing nation, is home to an extensive diversity of traditional rice varieties that have been cultivated for generations and exhibit remarkable adaptability to diverse local environmental conditions (Han *et al*., 2025). These landraces possess invaluable genetic traits, including pest resistance, tolerance to abiotic stresses and distinct nutritional characteristics, making them indispensable for rice breeding programs and genetic conservation initiatives (Jekkaral *et al.,* 2025). However, the increasing reliance on high-yielding hybrid varieties and modern agricultural practices has contributed to a decline in the cultivation of traditional landraces, thereby placing their genetic diversity at significant risk (Bonnin *et al.,* 2014). Accurate identification and classification of these traditional varieties are vital for their conservation, utilization in breeding programs and safeguarding under intellectual property laws.

Historically, rice varieties have been identified based on morphological and agronomic traits, such as grain shape, plant height, flowering time and resistance to biotic and abiotic stresses (Akter *et al.,* 2025). However, these traits are frequently influenced by environmental factors, resulting in inconsistencies in classification (Patindol *et al.,* 2015). Biochemical markers, including isozymes and protein profiles, have also been employed for varietal discrimination but generally lack the resolution required to distinguish closely related genotypes (Eevera and Vanangamudi, 2009). In contrast, DNA-based methodologies offer a precise and objective approach to varietal identification, enabling the detection of genetic differences at the molecular level (Azizi *et al.,* 2021). Among these, DNA barcoding has emerged as an efficient tool for species authentication and genetic diversity assessment. This technique employs short, standardized DNA sequences to differentiate species and infer phylogenetic relationships, providing a rapid and reliable method for plant identification (Hebert *et al.,* 2003).

[1]Vo Truong Toan University, Can Tho, Vietnam.
[2]Plant Resource Center, Hanoi, Vietnam.
[3]Hanoi Metropolitan University, Hanoi, Vietnam.
[4]Agricultural Genetics Institute, Hanoi, Vietnam.

Corresponding Author: Khuat Huu Trung, Agricultural Genetics Institute, Hanoi, Vietnam. Email: khuathuutrung@gmail.com

How to cite this article: Nhung, N.T., Trang, L.T.T., Nghia, L.T., Minh, L.H.N., Toan, N.N., Doai, N.T.P. and Trung, K.H. (2026). Genetic Classification of Traditional Rice Varieties Collected from Vietnam using *matK* and *rbcL* DNA Barcode Markers. *Indian Journal of Agricultural Research*. **60(7):** 976-982. doi: 10.18805/IJARe.AF-1063.

Submitted: 09-05-2026 **Accepted:** 14-06-2026 **Online:** 26-06-2026

In plants, the *matK* (maturase K) and *rbcL* (ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit) genes have been recommended by the Consortium for the Barcode of Life (CBOL) as core DNA barcode markers for land plants (CBOL Plant Working Group, 2009). While the *rbcL* gene is relatively conserved and suitable for broad taxonomic classification, the *matK* gene demonstrates higher sequence variability, offering superior resolution for distinguishing closely related taxa (Mursyidin *et al*., 2021). These markers have been successfully applied in various plant barcoding studies, including research on economically significant crops such as leguminosae, grass forages and rice (Li *et al*., 2022; Singh *et al*., 2017).

Despite the growing application of DNA barcoding in plant systematics and crop research, limited efforts have systematically utilized this technique to characterize the genetic diversity of traditional rice varieties in Vietnam. For giving the extensive genetic wealth of Vietnamese rice landraces, establishing a barcode reference dataset based on *matK* and *rbcL* sequences would provide critical support for species authentication, germplasm management and biodiversity conservation.

This study analyzed 90 traditional rice accessions collected from Northern, Central and Southern regions of Vietnam using *matK* and *rbcL* barcode markers. The key objectives were to (1) Evaluate the effectiveness of these chloroplast genes for species-level discrimination, (2) Assess genetic diversity and phylogenetic relationships among the accessions and (3) Develop a molecular framework for classifying and conserving traditional Vietnamese rice varieties. By employing DNA barcoding for varietal authentication, this research contributes to the conservation of rice genetic resources, enhances understanding of genetic diversity within traditional Vietnamese landraces and supports breeding programs aimed at developing improved rice varieties with desirable traits. Furthermore, this study holds broader implications for agricultural biodiversity conservation, food security and safeguarding indigenous genetic resources under international intellectual property frameworks.

MATERIALS AND METHODS

Plant materials

A total of 90 traditional rice accessions from Vietnam were analyzed in this study. These accessions were collected in 1994 and 2015, representing three major geographic regions: Northern, Central and Southern Vietnam. All seed samples were conserved at the plant resource center, where each accession was assigned a unique voucher code for reference in data presentation (Table 1).

For germination, seeds were wrapped in cotton cloth, soaked in water overnight and incubated at 37°C for two days under moist conditions to promote sprouting. Germinated seeds were subsequently sown in soil trays. Young leaves were harvested from seedlings at a height of 15-20 cm for DNA extraction. Molecular experiments were conducted at the Agricultural Genetics Institute in 2024.

DNA extraction and amplification

Genomic DNA was extracted from fresh leaf tissue using the TopPURE® Food DNA Extraction Kit (ABT, Vietnam), following the manufacturer's instructions. DNA quantity and quality were assessed using a Nanodrop 2000 UV-VIS Spectrophotometer (Thermo Scientific, USA). The *matK* and *rbcL* regions of the chloroplast genome were amplified using universal primers (Table 2). PCR was performed in a 25 µL reaction volume containing 20 ng of genomic DNA, 0.2 µmol of each primer and MyTaq HS Red Mix (Bioline, UK). The PCR cycling conditions were as follows: Initial denaturation at 94°C for 3 minutes, followed by 35 cycles of denaturation at 94°C for 30 seconds, annealing at 58°C for 40 seconds, extension at 72°C for 45 seconds, final extension at 72°C for 7 minutes. PCR products were separated by electrophoresis on a 2% agarose gel in 1X TBE buffer, stained with GelRed (ABT, Vietnam) and visualized under UV transillumination. Successfully amplified products were purified and sequenced bidirectionally using the Sanger method. DNA sequence data were obtained in ABI, FASTA and chromatogram formats.

Data analysis

Forward and reverse reads were processed with trimming by using BioEdit software. Species identification was performed utilizing the BLAST tool on the NCBI platform. Multiple sequence alignments were conducted with CLUSTAL W. Phylogenetic analysis and genetic diversity evaluation were executed using MEGA 11, employing the UPGMA method with the Tamura-Nei model. The reliability of tree topology was assessed through bootstrap analysis with 1,000 replicates (Tamura *et al*., 2021).

RESULTS AND DISCUSSION

Analysis of *MatK* sequences in 90 local rice varieties collected from Vietnam

PCR amplification of the *matK* gene was successful in 77 out of 90 rice varieties, achieving a success rate of 85.56%. Each amplified sample exhibited a distinct, single band of approximately 900 bp, consistent with the expected amplicon size of the *matK* region, with no evidence of non-specific amplification. This high amplification efficiency underscores the reliability of the selected primers for analyzing Vietnamese local rice germplasm.

The nucleotide composition of the *matK* sequences averaged 29.4% adenine (A), 35.8% thymine (T), 16.2% guanine (G) and 18.6% cytosine (C).

Multiple sequence alignment of *matK* gene sequences from 77 Vietnamese local rice varieties, performed using MEGA software, identified 11 single nucleotide polymorphisms (SNPs) at positions 218, 231, 792, 843, 844, 845, 847, 848, 849, 867 and 875 (data not shown).

Table 1: List of samples were used in this study.

Sample code	Genbank code	Name of accession	Origin
1	19	Khau mua venh Cao Bang	Cao Bang, North Vietnam
4	1846	Lua mo do	Ha Giang, North Vietnam
5	2151	Nep bo	Quang Nam, Central Vietnam
6	2166	Vang nghe	Phu Yen, Central Vietnam
7	2289	Nang chet do	An Giang, South Vietnam
8	2293	Nang muon	An Giang, South Vietnam
9	2304	Nep than	An Giang, South Vietnam
10	2372	Nep cai hoa vang	Nam Dinh, North Vietnam
11	2375	Tam thom ap be	Nam Dinh, North Vietnam
12	2598	Hong cu	Lai Chau, North Vietnam
13	3323	Lua nep 3 thang	Quang Nam, Central Vietnam
14	3412	Gie trang	Quang Binh, Central Vietnam
15	3482	Nep duoi nai	Quang Ngai, Central Vietnam
16	3484	Lua tri do	Binh Dinh, Central Vietnam
17	3495	Lua cang	Binh Dinh, Central Vietnam
18	3521	Ba du	Binh Dinh, Central Vietnam
19	3536	Nep ong gia	Quang Ngai, Central Vietnam
20	3537	Lua lup	Quang Ngai, Central Vietnam
21	3542	Ankhe	Quang Ngai, Central Vietnam
22	4124	Khau lay deng	Cao Bang, North Vietnam
23	4166	Clang pe	Tuyen Quang, North Vietnam
24	4746	Cham duoc	Thanh Hoa, Central Vietnam
25	4762	Lo cang	Thanh Hoa, Central Vietnam
26	5199	Tar rang	Thua Thien Hue, Central Vietnam
27	5801	Them dia	Tra Vinh, South Vietnam
28	5815	Chim lun	Tra Vinh, South Vietnam
29	5836	Nang thom muon	Ho Chi Minh city, South Vietnam
30	5839	Lua thuoc trang	Ho Chi Minh city, South Vietnam
32	5862	Hai bong	Tien Giang, South Vietnam
33	5867	Nang leo	Long An, South Vietnam
34	5875	Nang thom do	Ho Chi Minh City, South Vietnam
35	6180	Nep rau Thai Binh	Thai Binh, North Vietnam
36	6249	Tam xoan	Nam Dinh, North Vietnam
37	6254	Nep cai	Thai Binh, North Vietnam
38	6373	Nep trang	Hai Duong, North Vietnam
40	6467	Hai hoanh lun	Soc Trang, South Vietnam
41	6476	Nep do	Ca Mau, South Vietnam
42	7052	Te di huong	Hai Phong, North Vietnam
43	7116	Te hat dai	Kon Tum, Central Vietnam
44	7117	Xe dang	Kon Tum, Central Vietnam
45	7219	Nep luong	Thanh Hoa, Central Vietnam
46	7297	Vang me	Thanh Hoa, Central Vietnam
47	7332	Pom trang	Quang Binh, Central Vietnam
48	7340	I hom	Quang Binh, Central Vietnam
49	7342	Ca rang	Quang Binh, Central Vietnam
50	7346	Lua troi	Quang Binh, Central Vietnam
52	7811	Cam	Quang Tri, Central Vietnam
53	7814	A muon	Thua Thien Hue, Central Vietnam
54	7815	A mo	Thua Thien Hue, Central Vietnam
55	7823	Ran trang	Binh Thuan, Central Vietnam

Table 1: Continue...

Table 1: Continue...

56	7824	Nep ray	Binh Thuan, Central Vietnam
57	7826	Nep lua	Ba Ria Vung Tau, South Vietnam
58	7827	Nang thiet	Ba Ria Vung Tau, South Vietnam
59	8219	Bao thai hong	Son La, North Vietnam
60	8789	Lua de	Hai Phong, North Vietnam
66	9221	Nep ngoi	Ca Mau, South Vietnam
67	9222	Lua tram bong	Ca Mau, South Vietnam
68	9226	Lun den	Bac Lieu, South Vietnam
69	9227	Nep gui giang	Bac Lieu, South Vietnam
70	9228	Lun nhan	Bac Lieu, South Vietnam
72	9244	Nep ba thang	Binh Thuan, Central Vietnam
73	9246	Ca ro	Tay Ninh, South Vietnam
74	9247	Xuong ga	Tay Ninh, South Vietnam
78	9510	Ba la	Khanh Hoa, Central Vietnam
79	9511	Coi xoang	Khanh Hoa, Central Vietnam
80	9537	Viet Huong chiem	Khanh Hoa, Central Vietnam
86	9592	Lun can	Ben Tre, South Vietnam
87	9594	Nep nho	Ben Tre, South Vietnam
88	9598	Ba xuyen	Ben Tre, South Vietnam
90	12046	Va tai deo vinh	Ninh Thuan, Central Vietnam
91	12048	Va tai ca pla	Ninh Thuan, Central Vietnam
92	12057	Tay bong sen	An Giang, South Vietnam
93	12065	Chong ba la	An Giang, South Vietnam
94	12078	Than nong cao	Kien Giang, South Vietnam
95	12079	Ba bui	Kien Giang, South Vietnam
96	12087	Trang nho	Kien Giang, South Vietnam
97	12316	Chieng Thay	Ha Giang, North Vietnam
99	12377	Khau cai ca	Dien Bien, North Vietnam
101	12563	Cham hon	Hoa Binh, North Vietnam
102	12572	Khau cham pom	Hoa Binh, North Vietnam
103	12581	Blaopulau	Hoa Binh, North Vietnam
104	12584	Bloa co coong	Hoa Binh, North Vietnam
105	13443	Mua chua sung	Lai Chau, North Vietnam
106	14216	Khau pom	Son La, North Vietnam
107	14225	Khau cu giang	Son La, North Vietnam
108	14348	San cu	Lai Chau, North Vietnam
109	14462	Ple mang mu	Yen Bai, North Vietnam
110	14509	Khau xe	Lai Chau, North Vietnam
111	14548	Khau on mi pom	Lai Chau, North Vietnam
112	14602	Beo but bua	Lao Cai, North Vietnam

Most of the varieties shared conserved nucleotides at these positions; however, a few accessions showed unique SNP profiles. Hai bong (sample 32) exhibited a unique T substitution at position 218, where all other samples had C. Nep bo (sample 5) showed two SNPs at positions 231 (A) and 792 (T), in contrast to the conserved G allele. Ran trang (sample 55) presented a deletion (gap) at position 875. Ba la (sample 78) had substitutions at positions 843 (T) and 847 (T). Beo but bua (sample 112) displayed the most divergence, with unique alleles at four consecutive positions: 844 (G), 845 (A), 848 (C) and 849 (T). These SNP patterns provide clear molecular distinctions among certain rice accessions.

A phylogenetic tree was constructed based on *matK* sequences using the UPGMA method under the Tamura-Nei model, supported by 1,000 bootstrap replicates (Fig 1). The resulting tree had a total branch length of 0.009 and pairwise genetic distances ranged from 0.000 to 0.007, with an average of 0.0033. The phylogenetic tree revealed two major clades. Clade I consisted exclusively of Beo but bua (sample 112), indicating its strong genetic divergence from the remaining accessions. Clade II included the

remaining 76 samples, which were further divided into smaller subclusters. Notably, Nep bo (5) formed a distinct branch. Hai bong (32) and Ba la (78) also branched separately from the core group.

Analysis of rbcL sequences in 90 local rice varieties from Vietnam

All 90 local rice accessions produced successful PCR amplification of the *rbcL* gene, with 100% amplification efficiency. Amplicons were approximately 600 base pairs in length and showed sharp, single bands on agarose gels, confirming both high DNA quality and primer specificity.

Nucleotide composition analysis of the aligned *rbcL* sequences revealed an average of 26.7% adenine (A), 29.1% thymine (T), 23.2% guanine (G) and 20.9% cytosine (C). The sequence lengths varied from 552 bp (sample 83-Tam thom ap be) to 571 bp (sample 20-Nep do).

Alignment and polymorphism analysis using MEGA software identified seven SNP positions at nucleotide sites 27, 28, 29, 30, 31, 42 and 543 (data not shown). The majority of accessions shared the consensus nucleotide pattern A - (–) - (–)-C-T - (–) - C, but eight samples displayed clear sequence divergence. Samples 6, 11, 43, 44, 45, 52, 93 and 99 exhibited unique combinations of substitutions at positions 27 (T), 28 (A), 29 (C) and 30 (T). Several of these also had distinct alleles at positions 31 and 42, particularly showing the presence of an “A” nucleotide at position 42 instead of a deletion found in most accessions.

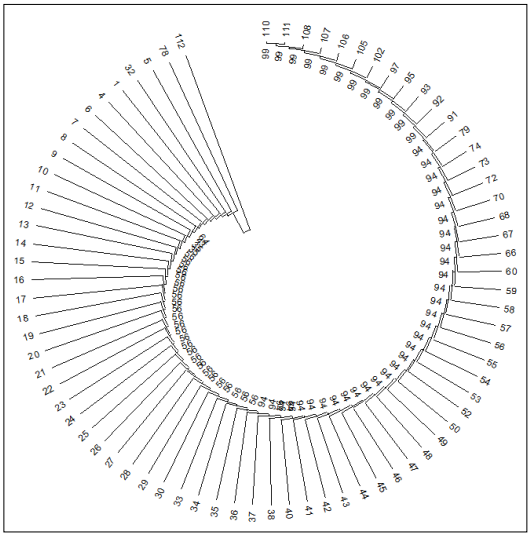


Fig 1: Phylogenetic tree of local rice varieties based on *MatK* sequences.

These variants indicate possible evolutionary divergence or environmental adaptation among these accessions.

A phylogenetic tree was constructed using the UPGMA method and the Tamura-Nei model, with 1,000 bootstrap replicates (Fig 2). The tree revealed a total branch length of 0.003 and pairwise genetic distances among accessions ranged from 0.000 to 0.004, with an average of 0.00053, indicating relatively low overall genetic diversity. The tree clearly separated the 90 rice accessions into two major clades. Group I included the eight genetically divergent samples listed above. These varieties occupied peripheral branches and showed multiple unique SNPs. Group II comprised the remaining 82 accessions, which were tightly clustered and displayed minimal sequence divergence, reflecting high genetic similarity.

This study provides the first comprehensive evaluation of genetic diversity among 90 local Vietnamese rice varieties using two chloroplast DNA barcoding markers, *matK* and *rbcL*. By integrating both markers, we were able to assess sequence variation and infer phylogenetic relationships across landraces collected from northern, central and southern Vietnam.

The amplification efficiency for *matK* reached 85.56%, higher than previously reported by Singh *et al.* (2017), who achieved 66.2% amplification across diverse rice genotypes using multiple primer sets. A consistent about 900 bp amplicon was observed in 77 accessions, confirming the reliability of the selected *matK* primers for Vietnamese rice germplasm. Meanwhile, *rbcL* demonstrated a 100% amplification success rate, producing clean 600 bp bands in all 90 accessions, reaffirming its high conservation and primer robustness. These findings support the utility of both *matK* and *rbcL* as effective tools for initial genetic diversity assessments in rice.

The *matK* region showed moderate variability, with eleven informative SNPs across 77 accessions, while *rbcL* revealed only seven SNPs across 90 accessions-consistent with the low mutation rate typically observed in *rbcL* for angiosperms (Clegg, 1993). Despite its low variability, *rbcL* analysis revealed meaningful divergence in a small subset of accessions, highlighting its complementary value in diversity studies.

Notably, the *rbcL* gene sequence analysis revealed a generally low level of genetic diversity among Vietnamese local rice varieties, with an average genetic distance of just 0.00053. This narrow genetic base likely stems from traditional farming practices, prolonged local adaptation and limited seed exchange, all of which have

Table 2: Primers used in the study.

Region	Sequence (5′-3′)	Target (bp)	Reference
MatK	F: 5′- CGATCTATTCATTCAATATTTC-3′	900	Vijayan and Tsou (2010)
	R: 5′- GTTCTAGCACAAGAAAGTCG-3′		
rbcL	F: 5′- ATGTCACCACAAACAGAGACTAAAGC-3′	600	Mursyidin *et al.* (2021)
	R: 5′- GTAAAATCAAGTCCACCRCG-3′		

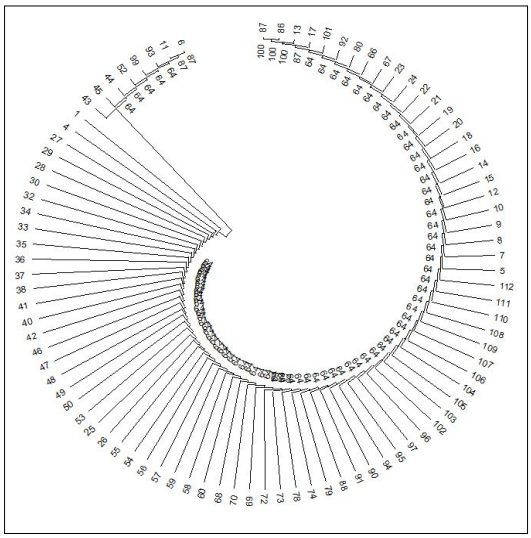


Fig 2: Phylogenetic tree of local rice varieties based on *rbcL* sequences.

contributed to the high genetic homogeneity observed across most accessions.

However, eight accessions-Vang nghe (sample 6), Tam thom ap be (sample 11), Te hat dai (sample 43), Xe dang (sample 44), Nep luong (sample 45), Cam (sample 52), Chong ba la (sample 93) and Khau cai ca (sample 99) -were clearly differentiated from the remaining samples, forming a distinct phylogenetic clade (Group I) in the *rbcL*-based tree. These accessions were genetically distinct and carried unique allelic combinations at multiple SNP positions, most notably at positions 27, 28, 29, 30 and 42. These polymorphisms included base substitutions (*e.g.*, T replacing A or C) and an insertion at position 42. Such SNPs serve as potential diagnostic markers for varietal grouping and molecular identification, underscoring their utility in rice germplasm characterization and genetic resource management.

The identification of this genetically distinct Group I suggests that these accessions may harbor valuable agronomic traits or unique ecological adaptations. As such, they should be prioritized in conservation efforts-both in situ and *ex situ*-and considered as important genetic materials for future breeding programs aimed at improving stress tolerance, yield stability, or local adaptation.

Phylogenetic reconstruction using the UPGMA method and Tamura-Nei model revealed clear and consistent clustering patterns. The *matK*-based tree grouped 76 accessions into a large, homogeneous cluster (Group II), while Beo but bua (sample 112) formed a separate clade (Group I), supported by its unique SNP profile and longer branch length. Similarly, the *rbcL* tree divided accessions into two main groups: Group I, composed of the eight genetically distinct accessions mentioned above and Group II, comprising the remaining 82 samples with highly conserved sequences. Although the total branch length in the *rbcL* tree was relatively short (0.003), compared to *matK* (0.009), it still captured meaningful variation, particularly among outlier accessions. These findings align with previous studies by Mursyidin *et al.* (2021) and Dang *et al*. (2021) which reported that *matK* generally offers greater resolution in phylogenetic analyses, though *rbcL* can still detect key divergences, particularly when combined with other markers.

Despite these observations of divergent lineages, overall nucleotide diversity remained low across both markers-0.0033 for *matK* and 0.00053 for *rbcL*. This genetic uniformity likely reflects the effects of domestication bottlenecks, local seed-saving traditions and limited gene flow across geographical regions. Such trends have also been observed in other crops subjected to similar cultural and ecological pressures (Johnston-Monje and Raizada, 2011).

Nonetheless, accessions identified as outliers in both SNP and phylogenetic analyses -particularly Beo but bua (sample 112), Chong ba la (sample 93) and Khau cai ca (sample 99) -warrant further attention. These varieties may harbor rare alleles associated with adaptive traits and should be prioritized for in-depth molecular characterization, agronomic evaluation and conservation. As noted by Wilberg (2015), peripheral clustering in phylogenies often corresponds to ancestral or underutilized lineages, which may hold untapped potential for crop improvement.

In summary, this study demonstrates that while *rbcL* exhibits a lower level of polymorphism compared to *matK*, it remains a valuable complementary marker for detecting genetic differentiation among closely related rice varieties. The identification of eight distinct accessions within Group I and their unique SNP profiles-especially at positions 27-30 and 42-highlights the marker's applicability in varietal classification and conservation. These findings reinforce the role of chloroplast DNA markers in supporting sustainable rice breeding and the preservation of traditional germplasm in Vietnam.

CONCLUSION

This study presents the first comprehensive assessment of chloroplast DNA diversity in 90 Vietnamese rice landraces using *matK* and *rbcL* markers. A total of 18 informative polymorphic sites were identified, revealing both conserved and divergent maternal lineages across agro-ecological regions. The *matK* marker demonstrated higher resolution for detecting intraspecific variation, while *rbcL* offered complementary support for broader phylogenetic relationships. Notably, genetically distinct landraces such as Beo but bua, Chong ba la and Khau cai ca were consistently differentiated, representing valuable genetic resources for future breeding and conservation efforts. The narrow range of chloroplast variation observed among most accessions highlights the risk of genetic erosion and the need for urgent conservation strategies. This DNA barcode dataset, combined with SNP profiling, provides an effective and cost-efficient tool for varietal identification,

germplasm management and monitoring of genetic integrity. Integration of these molecular findings into national rice conservation strategies is recommended to enhance sustainable utilization and long-term preservation of Vietnam's traditional rice germplasm.

ACKNOWLEDGEMENT

The present study was supported by the program: The Project for the Development of Biotechnology in the Agricultural Sector by 2030, Ministry of Agriculture and Environment of Vietnam.

Disclaimers

The views and conclusions expressed in this article are solely those of the authors and do not necessarily represent the views of their affiliated institutions. The authors are responsible for the accuracy and completeness of the information provided, but do not accept any liability for any direct or indirect losses resulting from the use of this content.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this article. No funding or sponsorship influenced the design of the study, data collection, analysis, decision to publish, or preparation of the manuscript.

REFERENCES

Akter, M.B., Rana, M.S., Nahar, N., Sharif, M.A.R., Hasan, M.N., Sultana, A. and Al Mamun, M.A. (2025). Assessing morphological traits of high yielding aus rice varieties in Bangladesh. *Tropical Agriculture*. **102(1):** 25-35.

Azizi, M.M.F., Lau, H.Y. and Abu-Bakar, N. (2021). Integration of advanced technologies for plant variety and cultivar identification. *Journal of Biosciences*. **46:** 1-20.

Bonnin, I., Bonneuil, C., Goffaux, R., Montalent, P. and Goldringer, I. (2014). Explaining the decrease in the genetic diversity of wheat in France over the 20th century. *Agriculture, Ecosystems and Environment*. **195:** 183-192.

CBOL Plant Working Group, Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S. and Little, D.P. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences*. **106(31):** 12794-12797.

Clegg, MT. (1993). Chloroplast gene sequences and the study of plant evolution. *Proceedings of the National Academy of Sciences*. **90(2):** 363-367.

Dang, T.L., Hoang, T.K.H., Le, L.T.T. and Nguyen, T.Q.T. (2021). Evaluation of genetic diversity by DNA Barcoding of local lotus populations from Thua Thien Hue Province. *Indian Journal of Agricultural Research*. **55(2):** 121-128. doi: 10.18805/IJARe.A-564.

Eevera, T. and Vanangamudi, K. (2009). Morphological, biochemical and molecular characterization of 26 rice cultivar seed and seedlings for cultivar discrimination. *Seed Science and Biotechnology*. **3(1):** 27-34.

Han, N.H., Cuong, H.H., Cong, H.S., Nhan, P.V., Mai, T.T., Quynh, N.X.T. and Nhi, P.T.P. (2025). Biological characteristics and nutrition of new colored rice lines in central Vietnam. *Indian Journal of Agricultural Research*. **59(11):** 1705-1710. doi: 10.18805/IJARe.AF-949.

Hebert, P.D., Cywinska, A., Ball, S.L. and DeWaard, J.R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. **270(1512):** 313-321.

Jekkaral, S.A., Kumar, B.D., Gangaprasad, S. and Halingali, B.I. (2025). Assessment of Morphological, genetical and diversity studies in landraces of rice (*Oryza sativa* L.). *Indian Journal of Agricultural Research*. **59(12):** 1841-1845. doi: 10.18805/IJARe.A-5726.

Johnston-Monje, D. and Raizada, M.N. (2011). Conservation and diversity of seed associated endophytes in Zea across boundaries of evolution, ethnography and ecology. *Plos One*. **6(6):** e20396.

Li, Y., La, B., Cao, L. and Zhao, S. (2022). Establishment of a DNA barcoding database for legume and grass species identification. *Legume Research: An International Journal*. **45(6):** 659-668. doi: 10.18805/AF-727.

Mursyidin, D.H., Nazari, Y.A., Badruzsaufari, B. and Masmitra, M.R.D. (2021). DNA barcoding of the tidal swamp rice (*Oryza sativa*) landraces from South Kalimantan, Indonesia. *Biodiversitas Journal of Biological Diversity*. **22(4):** 1593-1599.

Patindol, J.A., Siebenmorgen, T.J. and Wang, Y.J. (2015). Impact of environmental factors on rice starch structure: A review. *Starch Stärke*. **67(1-2):** 42-54.

Singh, J., Kakade, D.P., Wallalwar, M.R., Raghuvanshi, R., Kongbrailatpam, M., Verulkar, S.B. and Banerjee, S. (2017). Evaluation of potential DNA barcoding loci from plastid genome: Intraspecies discrimination in rice (*Oryza* species). *International Journal of Current Microbiology and Applied Sciences*. **6(5):** 2746-2756.

Tamura, K., Stecher, G. and Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*. **38(7):** 3022-3027.

Vijayan, K. and Tsou, C.H. (2010). DNA barcoding in plants: Taxonomy in a new perspective. *Current Science*. pp 1530-1541.

Wilberg, E. (2015). What's in an outgroup? The impact of outgroup choice on the phylogenetic position of *Thalattosuchia* (*Crocodylomorpha*) and the origin of *Crocodyliformes*. *Systematic Biology*. **64(4):** 621-637.