Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Agro-Morphological Characterization and Diversity analysis of Wainuchara (Oryza rufipogon Griff.) and Murshi (Oryza nivara Sharma & Shastry) Collected from Two Valley Districts of Manipur, Northeastern India

Hussain SYED

ssaddamhussain2013@gmail.com

Dhanamanjuri University

Medhabati KANGABAM

Dhanamanjuri University

Research Article

Keywords: Wild rice, Wainuchara (Oryza rufipogon Griff.), Murshi (Oryza nivara Sharma & Shastry), agro-morphological characterization, germplasm conservation, rice improvement programmes

Posted Date: August 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10687355/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Agro-Morphological Characterization and Diversity analysis of Wainuchara (*Oryza rufipogon* Griff.) and Murshi (*Oryza nivara* Sharma & Shastry*)* Collected from Two Valley Districts of Manipur, Northeastern India.

SYED Saddam Hussain, Medhabati KANGABAM
(Department of Botany, Dhanamanjuri University, Imphal 795001, Manipur, India)
Corresponding author: SYED SADDAM HUSSAIN (ssaddamhussain2013@gmail.com, 9366483450)
ORCID iD: 0009-0004-9643-9735

ABSTRACT

Wild relatives of rice (*Oryza sativa L.*) constitute an invaluable reservoir of novel alleles for broadening the genetic base of cultivated rice and enhancing resilience to biotic and abiotic stresses. The present study evaluated the agro-morphological diversity of ten wild rice accessions comprising Wainuchara (*Oryza rufipogon* Griff.) and Murshi (*Oryza nivara* Sharma & Shastry) collected from two valley districts of Manipur, India. The accessions were characterized using 102 agro-morphological descriptors (65 qualitative and 37 quantitative) following the DUS and Bioversity International guidelines. Qualitative traits were analyzed using polymorphism, Jaccard similarity, hierarchical clustering, and Principal Coordinate Analysis(PCoA), whereas quantitative traits were evaluated using descriptive statistics, coefficient of variation(CV), Pearson's correlation-based hierarchical clustering, and Principal Component Analysis(PCA). *O. rufipogon* showed higher qualitative polymorphism (53.85%) than in *O. nivara* (27.69%), whereas non-polymorphic traits were higher in *O. nivara* (72.31%). Quantitative analysis showed a wide CV(3.30–72.73%), with *O. rufipogon* while *O. nivara* was characterized by earlier flowering, lower grain shattering, and heavier panicles. Hierarchical clustering of qualitative traits based on the Jaccard similarity coefficients(0.205–0.776) classified agro-morphological characters into six major clusters, with Clusters I–III exhibiting strong similarity and Clusters IV–VI showing comparatively greater divergence. Unweighted Pair Group Method with Arithmetic Mean(UPGMA) analysis further separated the ten accessions into two distinct species-specific clusters indicating moderate phenotypic diversity and pronounced interspecific differentiation. Pearson's correlation-based clustering grouped quantitative traits into three major clusters representing (I) vegetative growth and yield, (II) plant architecture and phenology, and (III) reproductive and panicle-related traits, while the accessions were again partitioned into two major species-specific clusters, confirming strong within-group similarity and substantial between-group divergence. PCA based on Jaccard's distance explained 80.48% of total morphological variation, while PCA based on the Pearson correlation matrix explained 99% of quantitative variation, distinctly separating *O. rufipogon* and *O. nivara* and corroborating UPGMA clustering results. The concordance among similarity analysis, hierarchical clustering, PCoA, and PCA demonstrates clear phenotypic differentiation between the two wild accessions evaluated in Manipur and highlights *O. rufipogon* as a rich source of morphological diversity, while *O. nivara* contributes valuable agronomic attributes. These findings provide a robust scientific basis for germplasm conservation, parental selection, and the utilization of wild rice genetic resources in future rice improvement programmes.

Keywords: Wild rice, Wainuchara (*Oryza rufipogon* Griff.), Murshi (*Oryza nivara* Sharma & Shastry*)*, agro-morphological characterization, germplasm conservation, rice improvement programmes

INTRODUCTION

Rice (*Oryza sativa* L.) ranks among the world's most essential crops, nourishing over 50% of the global population. More than 90% of the planet's rice production and consumption occurs in Asia, home to 60% of humanity. As the demand for rice continues to rise with population growth, increasing production through sustainable genetic improvement has become a major priority (Khush, 2005). Wild *Oryza* species represent an invaluable reservoir of phenotypic diversity, providing beneficial alleles for enhancing yield, resistance to biotic and abiotic stresses, and environmental adaptability. In addition to traditional breeding methods, wide hybridization with wild species and leveraging gene bank resources—characterized via advanced genomic technologies—offer promising strategies (McCouch et al., 2012; Khush 2013). *O. rufipogon* and *O. nivara* enhances the understanding of the roles of phenotypic plasticity and environmental factors in the processes of adaptation and speciation (Sandamal et al. 2024). The comparatively greater phenotypic diversity detected among the *O. rufipogon* accessions in the present study is useful in light of recent genomic evidence showing that *O. rufipogon* possesses exceptionally high nucleotide diversity among AA-genome wild rice species (Tao et al., 2026).

Numerous wild relatives of cultivated rice have been gathered from diverse ecological zones across India (Patra *et al*. 2008; Singh *et al*. 2013, 2018), offer a rich reservoir of resistance genes against biotic stresses like blast, brown planthopper, and grassy stunt virus, aiding the genetic enhancement of *O. sativa*. Throughout history, wild *Oryza* species have contributed key traits, including resistance to diseases and pests (Brar & Khush, 1997). A recent case involves developing and molecularly confirming a cytoplasmic male sterility source from wild *O. rufipogon* in China (Shen *et al*. 2013; Yan *et al*. 2017). For instance, the grassy stunt virus resistance gene was isolated and transferred from an *O. nivara* accession (IRRI AC 101508) sourced from eastern Uttar Pradesh's Basti region in India (Khush 1977, 1989, 1997). Numerous genes conferring resistance to diseases and insects, as well as genes associated with high yield and tolerance to abiotic stresses, have been identified in wild *Oryza* species (Khush *et al*., 1990; Brar *et al*., 1996; Xiao *et al*., 1996). Wild rice is also a source of novel alleles, stress tolerance, disease resistance, yield improvement and pre-breeding (Zheng et al. 2024). Several of these beneficial genes have been successfully introgressed into cultivated rice varieties. Therefore, the conservation of wild *Oryza* biodiversity is crucial for ensuring a sustainable global food supply and for maintaining the continued availability and effective utilization of these valuable genetic resources. Researchers have frequently employed morphological characteristics for taxonomic classification and diversity studies in wild rice germplasm (Patra *et al*. 2008; Singh *et al*. 2013, 2018; Samal *et al*. 2018). Also, morphological traits fall short in defining genetic identities within the *O. sativa–nivara–rufipogon* complex and their hybrid intermediates (Singh et al. 2018).

Agro-morphological characterization of germplasm varieties is essential for generating useful information for plant breeding programmes (Lin, 1991). Recent pangenomic analyses have further demonstrated the breadth of genomic diversity retained in wild rice, identifying thousands of wild-rice-specific genes and greater diversity of resistance-gene analogues than in cultivated rice, denoting the importance of conserving and characterizing wild germplasm for future crop improvement (Guo et al., 2025). Although the Green Revolution greatly enhanced food grain production and helped achieve self-sufficiency, the widespread adoption of high-yielding varieties has indirectly led to the rapid loss of traditional landraces and wild rice varieties (Fowler and Moony 1990). Nevertheless, landraces remain critically important to agricultural systems, as the improvement of modern varieties depends on desirable genes that are often found only in traditional and wild varieties (Shiva 1991, Holden *et al.,* 1993). Wild rice habitats

in North Eastern India are shrinking rapidly due to climate change, shifting land uses for development, and expanding cultivation in recent years. This arises the needs to survey and gather the diverse wild relatives of rice from varied ecological niches, characterize the genetic variation and agro-morphology in these wild accessions, and preserve them for future applications in rice genetics and breeding. The primary objective of the present study was to characterize ten wild rice accessions from the Thoubal and Bishnupur district of Manipur using 102 agro-morphological characters, in order to generate useful information for selecting suitable genitors in rice breeding programmes.

MATERIALS AND METHODS

Extensive field surveys and germplasm collection missions were conducted between 2024 and 2025 across two valley districts of Manipur, resulting in the collection of 10 wild rice accessions (Supplementary Table 1; Fig. 1). Seeds were collected whenever possible; however, because seed collection from lake- and swamp-dwelling populations was often difficult, panicle-bearing plants were collected in such cases. Field observations included husk colour, seed coat colour, grain type, awn characteristics, growth habit, habitat, geographic coordinates, altitude, and other ecological information, all of which were recorded as passport data. The collected accessions were stored in labelled plastic bags before being transferred for long-term conservation. For ex situ characterization, the wild rice accessions were established at Lilong, Thoubal district, Manipur (24°42′ N, 93°55′ E) during the 2024 Kharif season. Seeds were soaked in water for 48 h to enhance germination and sown in soil-filled plastic containers. Two-week-old seedlings were transplanted into Styrofoam boxes containing fungicide-treated puddled soil for preliminary growth and seed multiplication. Adjacent accessions were maintained at a distance of 1 m to minimize cross-pollination and seed contamination caused by grain shattering. Initial observations included germination percentage, days to leaf emergence, and leaf sheath pigmentation. During the 2025 Kharif season, the accessions were evaluated in the field using an Augmented Block Design with three replicated cultivated rice checks per block. Each accession was planted in a single 1-m row with a spacing of 25 cm between rows and 20 cm between plants. Standard agronomic practices, including irrigation, weeding, and plant protection measures, were followed throughout the experiment.

A total of 65 qualitative and 37 quantitative agro-morphological traits were evaluated. Qualitative observations were recorded at specific growth stages according to the National Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability (DUS) in rice (*Oryza sativa* L.) under the Protection of Plant Varieties and Farmers' Rights (PPV&FR) Act, 2001, using standard morphological (non-metric) DUS descriptors. Quantitative traits were recorded following the Standard Evaluation System for Rice described by Bioversity International, IRRI and WARDA (2007). Descriptive statistics, including the mean, standard deviation (SD), and coefficient of variation (CV%), were computed for quantitative traits. Qualitative traits were analyzed using Jaccard's similarity coefficient, followed by hierarchical clustering using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) and Principal Coordinate Analysis (PCoA) based on the Jaccard distance matrix to assess morphological relationships among the wild rice accessions. Quantitative traits were also analyzed using Pearson's correlation coefficients, Principal Component Analysis (PCA) based on the Pearson correlation matrix, and UPGMA hierarchical clustering to evaluate phenotypic variation and relationships among accessions and traits. The cophenetic correlation coefficient (CCC) was calculated to assess the goodness-of-fit between the original Jaccard similarity matrix and the corresponding UPGMA dendrogram. All statistical analyses were performed using NCSS software (Version 26.0.3).

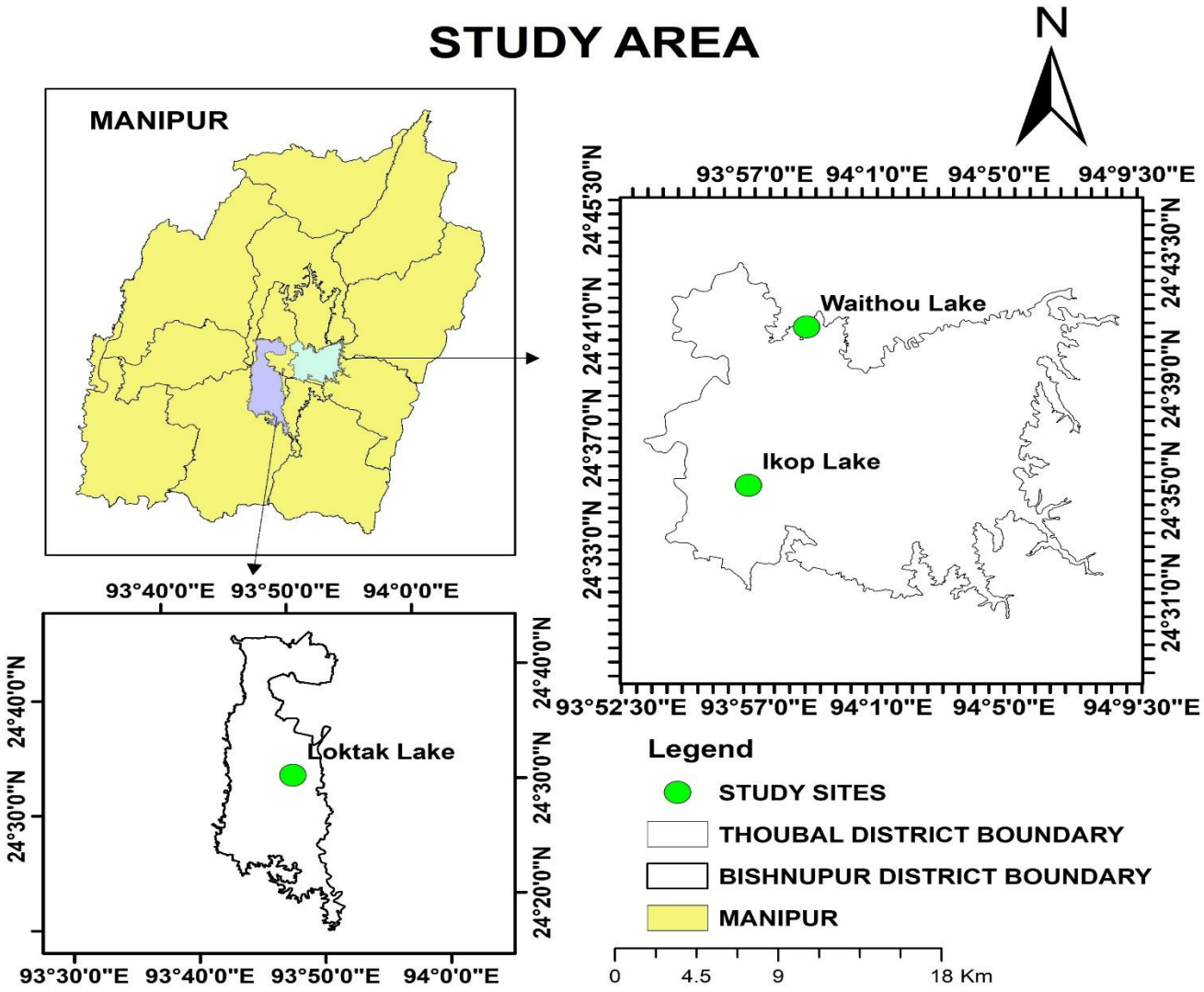


Fig. 1 Geo-referencing and mapping of wild rice accessions collected from different ecological regions located in two districts of Manipur namely Thoubal (08) and Bishnupur (02) using GIS tools

RESULTS AND DISCUSSION:

Collection and morphological characterization of wild rices

10 wild rice accessions were gathered— four Murshi(*O. nivara)* and six Wainuchara(*O. rufipogon* Griff.)—from diverse agro-ecological zones across two Manipur districts during various collection expeditions (Supplementary Table 1). These wild relative germplasm samples came from seasonal pools, shallow lake waters, swamps, deep-water zones, wetlands adjacent to cultivated fields and ex-situ (Fig. 2&3). Distinct wild rice features include seed shattering, long fully awned structures, black hulls, prominent auricles, and red awns. Across all 10 accessions regardless of species, basal leaf sheath and blade colors were predominantly green; ligules showed 2-cleft shapes in white; culms stood erect; awns were long and fully developed; auricles appeared mostly pale green; apiculi were mainly red; and panicles were typically intermediate and well-exserted (Fig. 2&3).

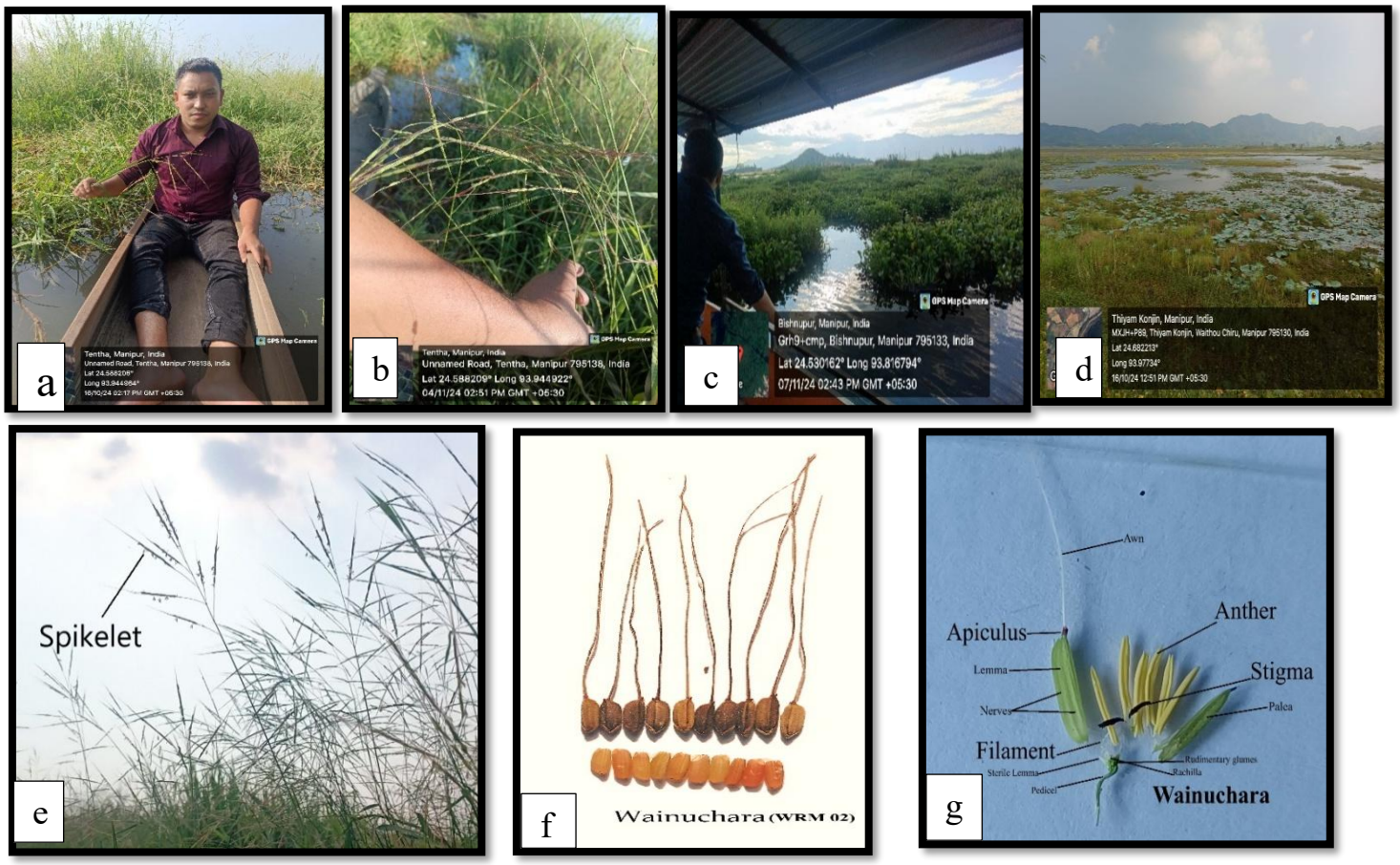


Fig. 2 Ecological niche sites of wild rice collecting Wainuchara (O. rufipogon Griff.) accessions were found growing in seasonal lake (a) in stagnant water of swampy Ikop Paat in Thoubal District, (b) collecting mature Wainuchara, (c) in and around Phumdis of Loktak lake, Thanga Area of Bishnupur District, (d) Waithou Pat, (e) Panicle and spikelet of Wainuchara,(f) Rice grain and caryopsis of Wainuchara and (g) flower of Wainuchara

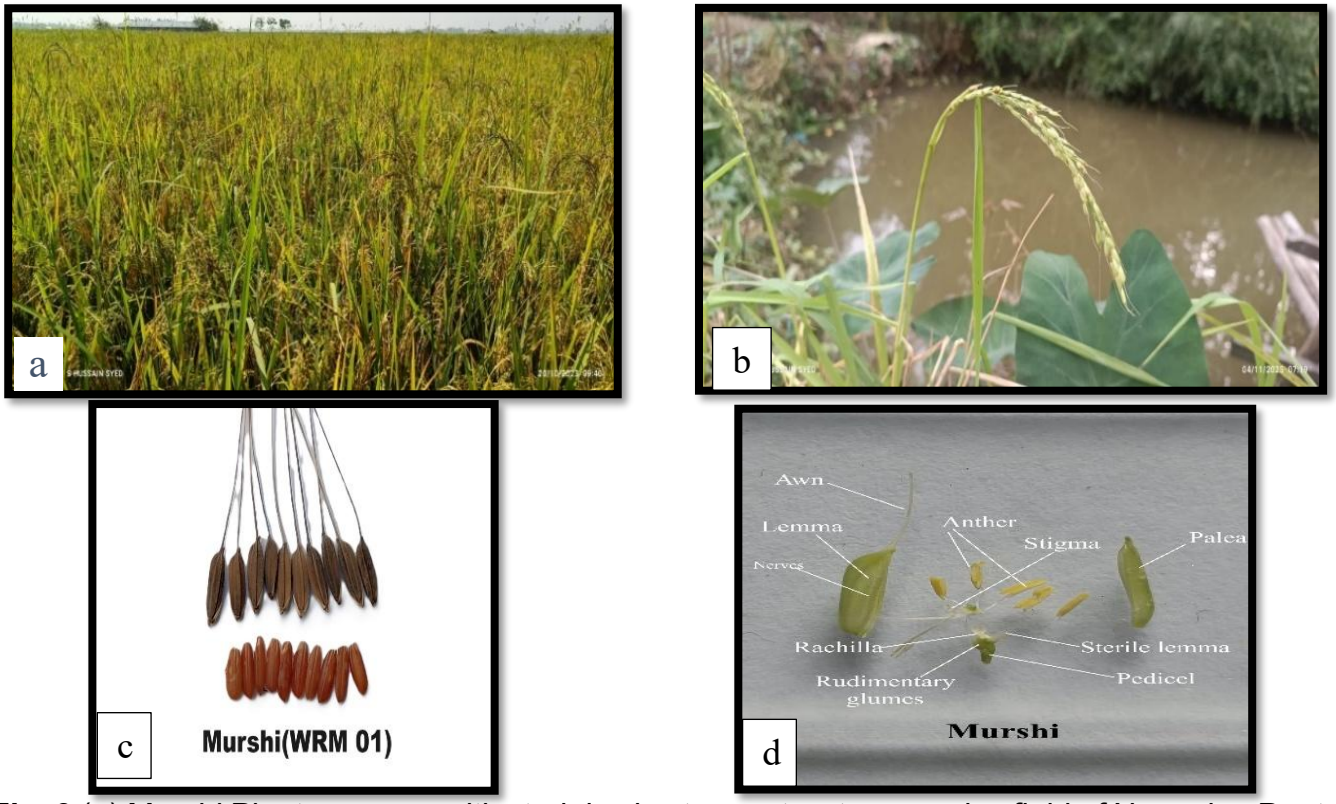


Fig. 3 (a) Murshi Plants among cultivated rice in stagnant water near rice field of Nganglou Paat, (b) early seed shattering in Murshi (O. nivara), (c) Rice grain and caryopsis of Wainuchara, and (d) flower of Murshi

Qualitative Characters

Table 1: Frequency Distribution of 65 Qualitative traits recorded in 10 accessions of Wainuchara and Murshi according to the National Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability (DUS) in rice (*Oryza sativa* L.)

Sl. No.	Qualitative traits	WCM001	WCM002	WCM003	WCM004	WCM005	WCM006	MUR001	MUR002	MUR003	MUR004
1.	CLC	5	5	5	5	5	4	1	1	1	1
2.	SH	3	3	3	5	3	5	3	3	3	3
3.	BLC	3	3	3	4	3	4	1	1	1	1
4.	LSC	5	5	5	5	5	5	3	0	3	0
5.	LBC	1	1	1	1	1	0	0	0	0	0
6.	LDC	1	1	2	1	2	1	0	0	0	0
7.	LBI	3	5	5	5	5	3	3	3	3	5
8.	LBA	5	1	5	5	7	5	1	1	1	1
9.	LBP	1	1	1	1	1	1	2	1	2	1
10.	LBL	5	5	3	5	5	3	3	5	3	3
11.	LBW	5	5	5	5	5	5	5	5	5	5
12.	LMP	2	2	2	2	2	2	1	1	1	1
13.	AUC	2	1	1	2	1	1	2	2	2	2
14.	CC	2	2	2	2	2	2	1	1	1	1
15.	LS	7	6	7	7	7	6	1	1	1	1
16.	LC	6	6	6	6	6	2	0	0	0	0
17.	LMS	1	1	1	1	1	1	0	0	0	0
18.	LMH	0	1	0	1	0	0	0	0	0	0
19.	LP	1	1	1	1	1	1	1	1	1	1
20.	CH	9	7	9	9	9	7	1	1	1	1
21.	CK	1	1	1	1	1	1	0	0	0	0
22.	CL	5	4	4	3	5	3	3	4	4	3
23.	CN	5	3	3	3	3	3	3	3	3	3
24.	CD	1	1	1	1	1	1	1	1	1	1
25.	CAC	1	2	1	2	1	1	0	2	0	0
26.	CNC	0	0	0	0	0	0	2	2	2	2
27.	CIA	1	1	1	1	1	1	0	0	0	0
28.	CIC	0	2	0	0	0	2	2	2	2	2
29.	CR	5	5	5	5	5	5	3	3	3	3
30.	CS	5	5	7	7	5	5	5	5	5	5
31.	FLA	5	5	3	5	5	3	3	1	3	1
32.	LS	7	7	7	7	7	7	7	7	7	7
33.	RS	4	4	4	4	4	3	0	0	0	0
34.	MS	1	2	1	1	2	1	1	1	1	1
35.	SGC	5	5	5	5	5	5	2	2	2	2
36.	ANC	1	1	1	1	1	1	1	1	1	1
37.	ELPC	8	9	9	9	9	8	9	9	9	9
38.	ELAP	8	8	8	8	8	8	4	4	4	4
39.	ELAC	1	1	3	3	1	3	0	0	0	0
40.	AWP	2	2	2	2	2	2	1	1	1	1
41.	EAWC	4	4	4	4	4	4	2	2	2	2
42.	AWL	9	9	9	9	9	9	7	7	9	9
43.	PAPB	2	2	2	2	2	2	2	2	2	2
44.	PTM	2	2	2	2	2	2	2	2	2	2
45.	PNP	5	5	5	5	5	3		7	7	7

46.	PAM	1	1	1	1	1	1	2	3	3	3
47.	PAB	5	5	5	3	5	3	1	1	1	1
48.	PSB	1	1	2	1	1	2	2	2	2	3
49.	PE	9	9	9	9	9	9	5	3	5	3
50.	PS	9	7	9	9	9	7	3	1	1	1
51.	PL	7	5	5	5	5	3	5	5	5	3
52.	PT	3	3	3	3	3	3	2	2	2	2
53.	LAWC	4	3	4	4	4	3	1	1	1	3
54.	LPP	2	2	2	2	2	2	2	2	2	2
55.	LLPC	4	4	4	4	4	4	3	4	4	3
56.	LAC	7	7	5	7	7	5	5	5	3	5
57.	LACA	7	7	5	7	7	5	5	5	3	5
58.	LLAP	3	3	2	3	3	3	3	3	2	3
59.	SLS	2	2	2	2	2	2	2	2	2	2
60.	SLC	1	1	1	1	1	1	1	1	1	1
61.	SF	3	3	4	4	3	3	5	5	5	5
62.	CS	4	4	4	4	4	4	4	4	4	4
63.	CPC	4	2	2	2	4	4	5	5	5	5
64.	ET	2	1	2	1	1	1	1	1	1	1
65.	GA	1	1	1	1	1	1	0	0	0	0

Coleoptile: anthocyanin colouration(CLC), Seedling Height(SH), Basal leaf sheath: colour (BLC), Leaf sheath: anthocyanin colouration(LSC), Leaf blade: anthocyanin colouration(LBC), Leaf blade: distribution of anthocyanin colouration(LDC), Leaf blade: intensity of green colour(LBI), Leaf blade: attitude(LBA), Leaf blade: pubescence(LBP), Leaf blade: length(LBL), Leaf blade: width(LBW), Leaf margin: pubescence(LMP), Auricle Colour(AUC), Collar: colour (CC), Ligule Shape (LS), Ligule Colour(LC), Ligule margin shape(LMS), Ligule margin hairiness(LMH), Ligule pubescence(LP), Culm: habit (CH), Culm: kneeing(CK), Culm: length(CL), Culm: number(CN), Culm: diameter at basal internode(CD), Culm: anthocyanin colouration on node(CAC), Culm: underlying node colour(CNC), Culm: internode anthocyanin(CIA), Culm: underlying internode colouration(CIC), Culm: lodging resistance(CR), Culm: strength(CS), Flag leaf attitude (FLA), Leaf: senescence(LS), Rhizome and stolon: formation(RS), Reproductive traits before harvesting, Male sterility(MS), Stigma colour (SGC), Anther colour(ANC), Lemma and palea colour (ELPC)- early observation, Lemma: apiculus colour(ELAP)- early observation, Lemma: anthocyanin colouration of area below apiculus(ELAC)- early observation, Awn: presence(AWP), Awn colour (EAWC)- early observation, Awn: length(AWL), Panicle arrangement of primary branches(PAPB), Panicle: texture to main axis(PTM) , Panicle: number per plant(PNP), Panicle: attitude to main axis(PAM), Panicle attitude to branches(PAB), Panicle: secondary branching (PSB), Panicle: exsertion (PE), Panicle: shattering(PS), Panicle: length(PL), Panicle: threshability(PT), Awn colour (LAWC)- late observation, Lemma and palea pubescence(LPP), Lemma and palea colour (LLPC)- late observation, Lemma-Anthocyanin colouration of keel (LAC), Lemma: anthocyanin colouration of area below apiculus(LACA)- late observation, Lemma: apiculus colour(LLAP)- late observation, Sterile lemma shape(SLS), Sterile lemma colour (SLC), Spikelet fertility (SF), Caryopsis: shape(CS), Caryopsis: pericarp colour(CPC), Endosperm type(ET) Grain aroma (GA)

Morphological characterization

Qualitative agro-morphological traits are widely recognized as reliable descriptors for varietal identification and phenotypic differentiation, particularly in rice germplasm characterization studies (IRRI, 2013; Bioversity International, 2007). Although qualitative traits are generally governed by major genes and exhibit discrete expression patterns, their phenotypic manifestation may still be influenced to some extent by environmental conditions, especially under diverse agro-ecological settings (Tanksley & McCouch, 1997). The comparative analysis of qualitative descriptors revealed that Wainuchara (*O. rufipogon*) possessed 53.85% polymorphic traits, whereas Murshi (*O. nivara*) exhibited only 27.69% polymorphism among the evaluated characters. This finding is consistent with earlier reports that *O. rufipogon*, being a perennial and widely distributed wild progenitor of cultivated rice, harbors broader phenotypic diversity than *O. nivara*, which is typically annual and ecologically more restricted (Vaughan et al., 2003; Khush, 1997). Conversely, the proportion of non-polymorphic traits was found to be higher in Murshi (72.31%) than in Wainuchara (46.15%). This clearly indicates a greater extent

of morphological diversity within the Wainuchara population, which may be effectively exploited in future rice improvement and breeding programmes. The comparative assessment of qualitative morphological traits between Wainuchara (*O. rufipogon*) and Murshi (*O. nivara*) phenotypes revealed significant variation in the expression pattern of several descriptors, indicating the presence of considerable phenotypic diversity between the two wild rice populations. The varietal similarity analysis based on qualitative morphological traits revealed that the highest similarity (86.15%) was observed between the genotype pairs WCM001 and WCM005, MUR001 and MUR003, and MUR002 and MUR004, indicating very low phenotypic divergence and close phenotypic relationship within the Wainuchara and Murshi populations. Such high phenotypic resemblance suggests possible shared ancestry or gene flow within each population. Ren et al., 2003; Thomson et al., 2007 reported genotypes of wild rice accessions collected from similar ecological niches tend to cluster together due to comparable selective pressures. Other closely related combinations included MUR002 and MUR003 (84.62%) and WCM002 and WCM005 (83.08%). In contrast, the lowest similarity values were recorded between WCM001 and MUR001 (27.69%), followed by WCM001 and MUR004 (29.23%), WCM004 and MUR001 (29.23%), WCM004 and MUR003 (29.23%) and WCM005 and MUR001 (29.23%), reflect substantial inter-population divergence, reinforcing the distinct evolutionary pathways of *O. rufipogon* and *O. nivara*.

Trait-wise similarity analysis further revealed that descriptors such as Culm diameter (CD), Sterile lemma colour (SLC), Leaf senescence (LS), Anther colour (ANC), and Ligule pubescence (LP) exhibited complete uniformity (100% similarity), indicating limited discriminatory power for varietal differentiation. Conversely, traits such as Panicle shattering (PS), Culm anthocyanin colouration (CAC), Late awn colour (LAWC), Early lemma anthocyanin colour (ELAC), and Caryopsis pericarp colour (CPC) showed low similarity values (22.22–26.67%), highlighting their substantial contribution to phenotypic divergence and cluster formation. Traits related to pigmentation and shattering are often under strong genetic control and subject to natural and artificial selection, particularly in the context of domestication and adaptation (Sweeney & McCouch, 2007; Fuller et al., 2009). Therefore, their variability among the studied genotypes may reflect adaptive differentiation and evolutionary significance. Overall, the pronounced phenotypic differentiation observed between *O. rufipogon* and *O. nivara* underscores the importance of wild rice populations as reservoirs of genetic variability. The greater polymorphism within Wainuchara suggests its potential utility as a valuable donor in rice improvement programmes, particularly for broadening the genetic base of cultivated rice. The identification of highly divergent genotypes between the two populations provides a strategic foundation for selecting genetically diverse parental lines to enhance heterosis and introgress desirable traits in future breeding efforts.

Jaccard Similarity Matrix

Pairwise genetic similarity among genotypes was calculated using Jaccard's similarity coefficient.

Key Results:

Comparison	Jaccard Similarity
WCM001 – WCM005	**0.766** (highest within WCM)
MUR001 – MUR003	**0.776** (highest within MUR)
WCM001 – MUR001	**0.205** (lowest between groups)
WCM002 – MUR002	**0.283**
WCM006 – MUR004	**0.308**

Table 2. Jaccard's similarity coefficient of 10 wild rice accessions of Wainuchara and Murshi

	WCM001	WCM002	WCM003	WCM004	WCM005	WCM006	MUR001	MUR002	MUR003	MUR004
WCM001		0.254717	0.366337	0.424242	0.345794	0.288288	0.31068	0.342857	0.333333	0.378641
WCM002			0.266667	0.346535	0.362745	0.266055	0.368421	0.36	0.293478	0.330097
WCM003				0.317757	0.358491	0.311927	0.377551	0.368932	0.305263	0.339623
WCM004					0.324324	0.327273	0.339806	0.358491	0.322917	0.394231
WCM005						0.330357	0.355769	0.4	0.287129	0.333333
WCM006							0.333333	0.351852	0.30303	0.348624
MUR001								0.4375	0.273684	0.275229
MUR002									0.309278	0.330275
MUR003										0.242718
MUR004										

Original 65 qualitative multi-state DUS traits were converted into a binary presence (1) / absence (0) matrix using multistate one-hot encoding as recommended by the International Rice Research Institute descriptor guidelines for morphological diversity studies. The Jaccard similarity coefficients computed from the binary data matrix of 143 qualitative trait states revealed varying degrees of morphological similarity among the ten wild rice genotypes. The similarity values ranged from 0.205 to 0.776, indicating the presence of both high intra-population resemblance and marked inter-population divergence. Higher similarity coefficients were observed among genotypes belonging to the same population group. The Wainuchara accessions showed similarity values ranging from 0.478 to 0.766, whereas the Murshi accessions exhibited comparatively higher similarity ranging from 0.679 to 0.776. In contrast, the similarity coefficients between Wainuchara and Murshi accessions were consistently low and ranged from 0.205 to 0.308. The low inter-group similarity values clearly indicate considerable morphological divergence between the Wainuchara (*Oryza rufipogon* Griff.) and Murshi (*Oryza nivara* Sharma & Shastry) wild rice populations.

UPGMA Cluster Analysis (Dendrogram)

The hierarchical clustering dendrogram based on the Jaccard similarity coefficient reveals the presence of six major clusters among the qualitative traits, separated at an approximate distance threshold of 2.0–2.5. Cluster I (red) comprise closely related traits such as Early lemma apiculus colour(ELAP) and Panicle exsertion(PE), which merge at relatively low distance values (< 1.0), indicating high similarity among these characters in all varieties. Cluster II (blue) forms a comparatively large group including traits such as Leaf blade anthocyanin colour(LBC), leaf blade anthocyanin colour distribution(LDC), Leaf blade pubescence(LBP), Leaf margin pubescence(LMP), Auricle colour(AUC), Collar colour(CC), Leaf margin shape(LMS), Leaf margin hairiness(LMH), Ligule pubescence(LP), Culm kneeing(CK), Culm diameter at basal node(CD), Culm anthocyanin colouration on node(CAC), Culm underlying node colour(CNC), Culm internode anthocyanin(CIA), Culm underlying internode colouration(CIC), Male sterility(MS), Anther colour(ANC), Early lemma anthocyanin colouration below apiculus(ELAC), Awn presence(AWP), Panicle arrangement on primary branches(PAPB),Panicle texture to main axis(PTM), panicle attitude to main axis(PAM), Panicle secondary branches(PSB), Lemma and palea pubescence(LPP), Sterile lemma shape(SLS), Sterile lemma colour(SLC), Embryo type(ET) and Grain aroma(GA), most of which coalesce below a distance of 1.5, suggesting moderate to strong association among the accessions. Cluster III (green), represented by Ligule shape(LIS), Culm habit(CH) and Panicle shattering(PS), merges at a very small distance (< 0.5), demonstrating a particularly strong relationship between these two characters. Cluster IV

(purple) includes traits like Leaf senescence(LS), Early lemma and Palea colour(ELPC) and Awn length(AWL), joins with other clusters at a higher linkage distance (≈2.0), indicating greater divergence from the preceding groups. Cluster V (orange), including traits such as Leaf blade width(LBW), Culm strength(CS), Panicle number per plant(PNP), Panicle length(PL), Lemma anthocyanin colouration of keel(LAC) and Late lemma anthocyanin colouration below apiculus(LACA), clusters at distances around 0.8–1.2, reflecting moderate similarity. Cluster VI (yellow) constitutes the largest aggregation, incorporating characters such as Coleoptile Colouration(CLC), Seedling height (SH), Basal leaf colour(BLC), Leaf sheath anthocyanin colouration(LSC), Leaf blade green colour intensity(LBI), Leaf blade attitude(LBA), Leaf blade length(LBL), Ligule colour(LC), Culm length(CL), Culm number(CN), Culm lodging resistance(CR), Flag leaf attitude(FLA), Rhizome and stolon formation(RS), Stigma colour(SGC), Early awn colour(EAWC), Panicle attitude to branches(PAB), Panicle thresshability(PT), Late awn colour(LAWC), Late lemma and palea colouration(LLPC), Late lemma apiculus colour(LLAP), Spikelet fertility(SF), Caryopsis shape(CS) and Caryopsis pericarp colour(CPC), which generally merge below 1.5, signifying substantial homogeneity within this group of accessions. The relatively large fusion distance (≈2.5–3.0) separating the upper major branches indicates considerable dissimilarity between the principal clusters. Overall, the dendrogram demonstrates clear structural grouping among qualitative traits, with clusters 1–3 exhibiting closer affinity, while clusters 4–6 display comparatively greater divergence.

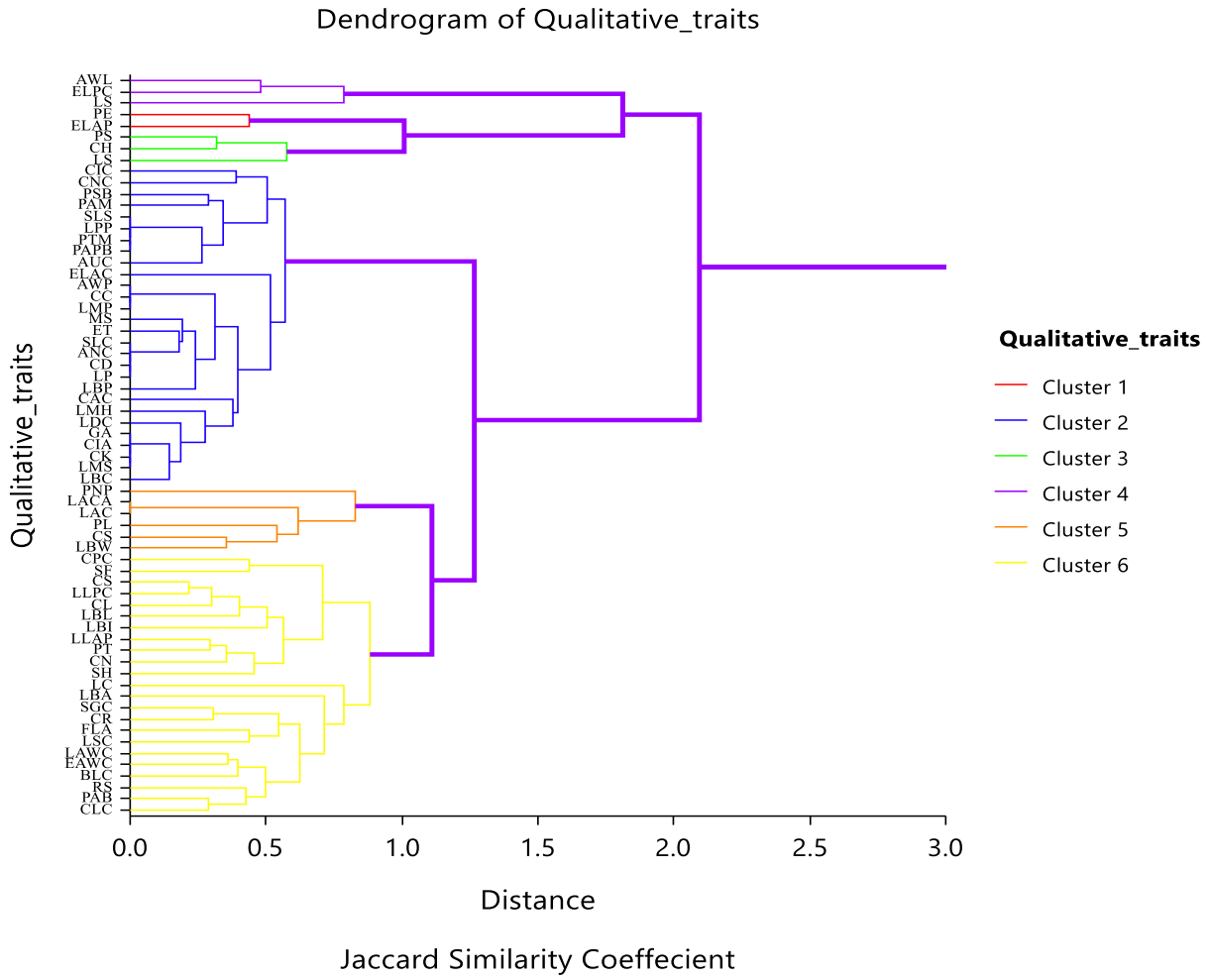


Fig. 4 Dendrogram of similarity patterns by the UPGMA methods estimated from 65 qualitative agro-morphological traits in 10 wild rice accessions of Wainuchara and Murshi

The UPGMA-based hierarchical clustering using the Jaccard similarity coefficient revealed a well-defined structure among the qualitative descriptors, indicating non-random association patterns among morphological characters. The presence of two major clusters merging at a relatively high linkage distance (≈2.20–2.30) reflects substantial dissimilarity among the principal trait groups, suggesting that the qualitative descriptors represent distinct morphological domains. The use of Jaccard's coefficient, which is particularly appropriate for binary (presence–absence) data, enhances the reliability of trait association analysis by excluding joint absences from similarity estimation (Jaccard, 1908; Sneath & Sokal, 1973). Clusters I and III, which merged at very low distance values (<1.0 and <0.5, respectively), indicate strong co-expression or possible genetic linkage among traits such as ELAP, PE, LIS, CH, and PS. Traits related to pigmentation and shattering are often governed by major genes and may exhibit coordinated inheritance patterns, particularly in wild and semi-domesticated rice populations (Sweeney & McCouch, 2007; Fuller et al., 2009). In contrast, Cluster IV, comprising LS, ELPC, and AWL, joined the main dendrogram at a higher distance (≈2.0), reflecting greater divergence and suggesting that these descriptors contribute substantially to phenotypic differentiation. The large and heterogeneous Clusters II and VI, which incorporated numerous vegetative and reproductive characters merging mostly below 1.5, indicate moderate to strong association within these trait complexes. Such grouping of vegetative (leaf and culm) and reproductive (panicle and grain) traits has been observed in rice diversity studies and may reflect functional integration or shared developmental pathways (IRRI, 2013; Thomson et al., 2007). The relatively high fusion distance (≈2.5–3.0) separating the uppermost branches further confirms the existence of distinct qualitative trait assemblages, underscoring the broad morphological variability present within the evaluated material. Hence, the clustering pattern supports the concept that qualitative morphological descriptors are effective tools for discriminating among genetically diverse accessions and for identifying key traits contributing to population differentiation. Such information is valuable for germplasm management, conservation planning, and strategic parent selection in breeding programmes aimed at broadening the genetic base of rice improvement efforts.

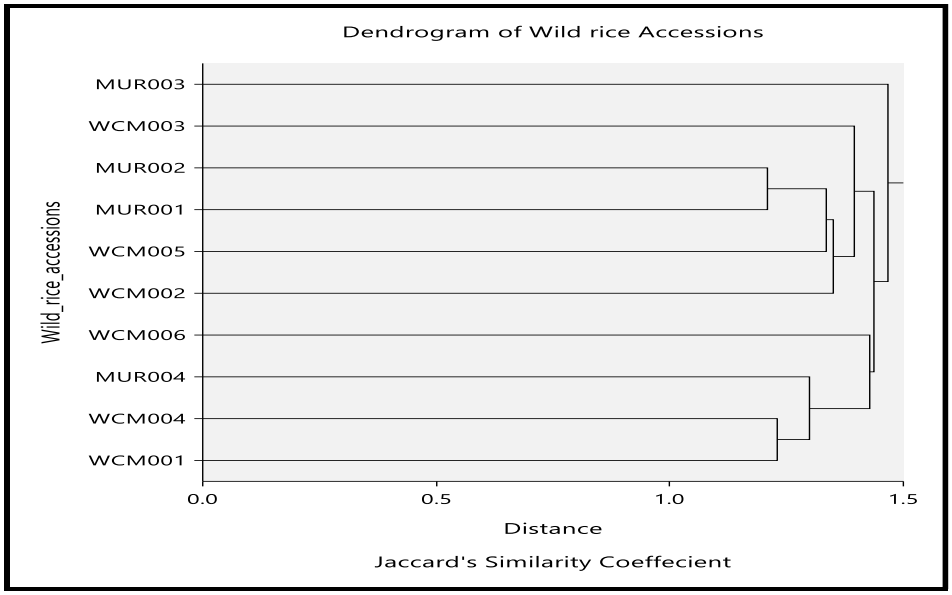


Fig. 5 Dendrogram of 10 wild rice accessions of Wainuchara and Murshi by the UPGMA methods based on Jaccard Similarity coefficient estimated from 65 qualitative agro-morphological traits.

The hierarchical clustering pattern derived from the Jaccard similarity coefficient (Jaccard, 1908) clearly demonstrates the presence of structured genetic variability among the ten varieties. The separation into two major clusters at a linkage distance of approximately 1.47 indicates appreciable inter-group divergence, suggesting that the studied genotypes are not genetically homogeneous. Similar clustering patterns have been widely reported in diversity studies employing similarity coefficients and hierarchical methods, where distinct clusters reflect differences in allelic composition and evolutionary history (Sneath & Sokal, 1973; Mohammadi & Prasanna, 2003). The distinct grouping of MUR003 and WCM003 into an independent cluster at a relatively high distance (≈1.42) highlights their genetic uniqueness, which may be attributed to differential selection pressure, adaptation to specific agro-ecological conditions, or unique pedigree backgrounds. Such divergence is particularly valuable in breeding programs, as genetically distant parents are often associated with greater heterotic responses and broader recombination potential (Falconer & Mackay, 1996; Hallauer et al., 2010). Within the second major cluster, the formation of subclusters at comparatively lower linkage distances (1.20–1.38) indicates higher intra-group similarity, suggesting shared ancestry or parallel selection histories among these varieties. The closer association of MUR002, MUR001, WCM005, and WCM002, as well as the grouping of WCM004, WCM001, MUR004, and WCM006, reflects moderate genetic affinity, which may result from similar breeding objectives or gene flow among related lines. Shorter linkage distances within clusters denote closer genetic relationships, while higher fusion levels correspond to broader genetic divergence (Nei & Li, 1979; Mohammadi & Prasanna, 2003). Therefore, the dendrogram confirms moderate phenotypic diversity among the evaluated varieties and identifies distinct genetic groupings that could be strategically exploited in hybridization programs. Crosses between genotypes from the {MUR003–WCM003} cluster and those in the second major cluster may enhance the probability of obtaining superior recombinants due to increased genetic distance. These findings shows the importance of the utility of Jaccard-based hierarchical clustering as an effective tool for assessing genetic relationships and guiding germplasm management and crop improvement strategies.

Principal Coordinate Analysis (PCoA)

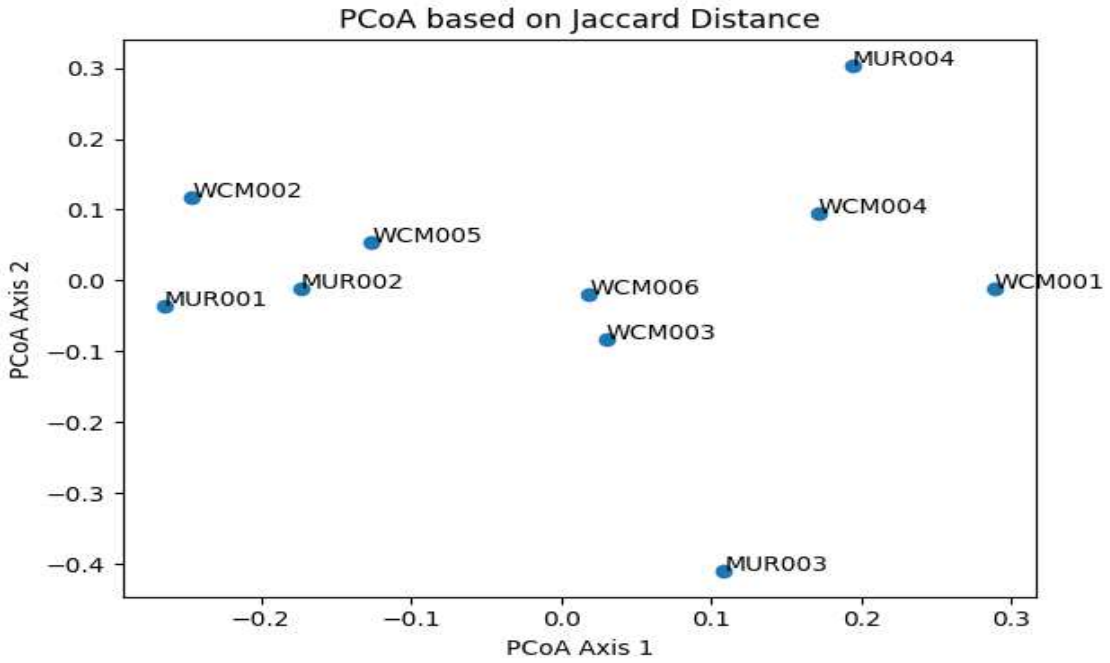


Fig. 6 2D scatters diagram of Principal Coordinate analysis(PCoA) of 10 wild rice accessions of Wainuchara and Murshi used in experiment from 65 qualitative traits.

Principal Coordinate Analysis (PCoA) based on Jaccard's distance matrix was performed to further elucidate the morphological relationship among the studied wild rice accessions. The first two principal coordinate axes accounted for 70.53% and 9.95% of the total variation respectively, cumulatively explaining 80.48% of the total morphological variability. Such a high cumulative variance suggests strong differentiation between groups and validates the suitability of Jaccard's coefficient for analyzing binary qualitative data (Jaccard, 1908; Legendre & Legendre, 2012). Ordination methods like PCoA are particularly valuable in germplasm characterization because they provide a visual representation of phenotypic relationships without imposing a strictly hierarchical structure, thereby complementing cluster analysis (Mohammadi & Prasanna, 2003). The two-dimensional PCoA plot distinctly separated the genotypes into two well-defined groups corresponding to Wainuchara (*Oryza rufipogon* Griff.) and Murshi (*Oryza nivara* Sharma & Shastry) accessions. All Wainuchara genotypes were clustered on one side of the coordinate axis, whereas the Murshi genotypes formed a separate and compact group on the opposite side, indicating substantial inter-population divergence. The clear spatial segregation of the two groups in the ordination space corroborated the clustering pattern obtained from UPGMA dendrogram analysis, thereby confirming the existence of significant morphological diversity between the Wainuchara and Murshi wild rice populations. Phylogeographic and molecular studies show significant divergence between *O. rufipogon* and *O. nivara*, attributing this differentiation to ecological adaptation, life-history differences (perennial vs. annual habit), and independent evolutionary trajectories (Londo et al., 2006; Vaughan et al., 2003). The compact clustering of genotypes within each group further indicates relatively higher intra-population similarity compared to inter-population variation, which is consistent with patterns expected in geographically or ecologically structured wild populations (Hamrick & Godt, 1996). The concordance between PCoA and UPGMA dendrogram results strengthens the reliability of the observed grouping pattern, as agreement between clustering and ordination approaches is generally considered strong evidence of true underlying genetic structure (Sneath & Sokal, 1973; Mohammadi & Prasanna, 2003). Hence, the pronounced separation between Wainuchara and Murshi accessions highlights the substantial morphological diversity present in these wild rice populations and underscores their importance as valuable genetic reservoirs for broadening the genetic base of cultivated rice improvement programmes.

Principal Component Analysis (PCA)

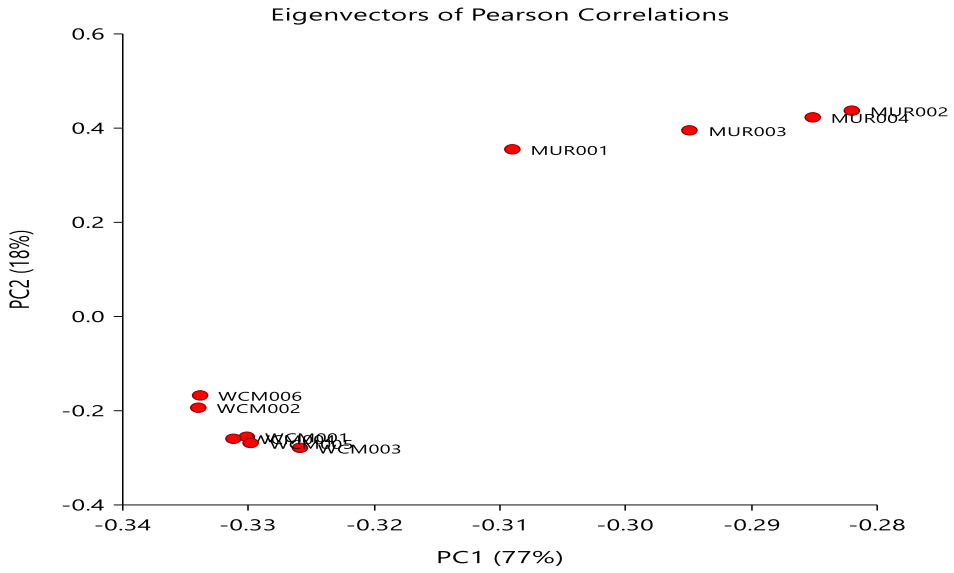


Fig. 7 Eigenvector scatters plot of Principal Component analysis(PCA) based on Pearson correlation matrix of 10 wild rice accessions of Wainuchara and Murshi used in experiment from 65 qualitative traits.

Table 3 Principal component analysis of 10 wild rice accessions - Eigen values, percentage variation and cumulative percentage of ten principal components derived from 65 qualitative traits.

Component	Eigenvalues		
	Total	% of Variance	Cumulative %
1	7.720	77.199	77.199
2	1.788	17.875	95.074
3	.136	1.356	96.430
4	.101	1.008	97.438
5	.074	.735	98.174
6	.061	.607	98.780
7	.044	.440	99.220
8	.038	.385	99.605
9	.021	.209	99.814
10	.019	.186	100.000

Principal Component Analysis (PCA) based on the Pearson correlation matrix was performed to assess the pattern of association and variation among the studied rice genotypes using qualitative morphological traits. The eigenvector scatter plot revealed that the first two principal components, namely PC1 and PC2, accounted for a major proportion of the total variation observed among the genotypes. The first principal component (PC1) explained 77% of the total variation, whereas the second principal component (PC2) contributed an additional 18% of the total variation. Together, these two components accounted for 95% of the cumulative variability, indicating that the majority of phenotypic variation among the genotypes could be effectively represented in a two-dimensional space. In multivariate studies, components explaining more than 70% of total variation are generally considered sufficient to reliably interpret genetic divergence patterns (Jolliffe, 2002; Mohammadi & Prasanna, 2003). Based on the PCA scatter plot, the genotypes were clearly separated into two distinct groups corresponding to the WCM (Wainuchara) and MUR (Murshi) populations. Such distinct grouping patterns in PCA ordination space are indicative of underlying agro-morphological differentiation and have been widely reported in rice diversity studies employing morphological and molecular data (Vaughan *et al*., 2003; Thomson *et al*., 2007). The Murshi genotypes were positioned in the positive direction of PC2 with coordinate values ranging approximately from +0.35 to +0.45, while their PC1 values ranged between –0.31 and –0.28. Among these, MUR004 exhibited the highest PC2 value of nearly +0.45, followed by MUR003 (+0.40) and MUR001 (+0.35), indicating relatively greater contribution of qualitative traits towards variability in this group. In contrast, the Wainuchara genotypes were located in the negative quadrant of PC2 with values ranging from approximately –0.15 to –0.30, and PC1 values between –0.34 and –0.32. For instance, WCM006 and WCM002 were positioned at PC2 values of nearly –0.18 and –0.20, respectively, whereas WCM001 and WCM003 exhibited lower PC2 values of approximately –0.25 to –0.28, reflecting relatively lesser contribution of these traits towards variability within this group. The clear separation of WCM and MUR genotypes along PC2 suggests significant phenotypic divergence between the two populations based on qualitative descriptors. The higher dispersion of Murshi genotypes in the positive PC2 direction further indicates the presence of greater variability among them. Therefore, the PCA results corroborate the findings of cluster analysis and confirm the existence of substantial diversity among the studied rice genotypes, which may be exploited for selection of genetically diverse parents in breeding programmes. These findings align with ecological and evolutionary distinctions between the two taxa, particularly their contrasting life-history strategies (annual vs. perennial forms), which are known to influence

patterns of genetic structure and diversity (Hamrick & Godt, 1996). Importantly, the concordance between PCA results and the previously obtained UPGMA clustering and PCoA analyses strengthens the reliability of the observed grouping pattern. Agreement among multiple multivariate approaches is generally considered strong evidence of true genetic structure rather than analytical artefacts (Sneath & Sokal, 1973; Mohammadi & Prasanna, 2003). Therefore, the PCA findings confirm excellent phenotypic divergence between *O. rufipogon* and *O. nivara* populations and highlight the potential utility of these wild genotypes as valuable sources of genetic variability for broadening the genetic base in rice breeding programmes. The consistent separation of *O. rufipogon* and *O. nivara* accessions across complementary multivariate analyses indicates substantial phenotypic differentiation between the two species. This phenotypic divergence is biologically plausible given the extensive genetic and genomic diversity reported across wild *Oryza* species (Long et al., 2024). However, the correspondence between phenotypic clustering and genomic differentiation remains to be established for the Manipur accessions through molecular marker or genome-wide sequencing approaches.

Cophenetic Correlation Coefficient (CCC)

The cophenetic correlation coefficient (CCC) is widely used to evaluate the fidelity of a dendrogram in preserving the original pairwise similarity or distance relationships among operational taxonomic units (Sneath & Sokal, 1973). In the present study, the CCC value of r = 0.767 indicates a good level of concordance between the original Jaccard similarity matrix and the UPGMA-derived clustering pattern. Values exceeding 0.70 are generally interpreted as indicating that the dendrogram provides a reliable representation of the underlying similarity structure, whereas lower values suggest distortion in the hierarchical grouping (Rohlf, 2000; Saraçli et al., 2013). The obtained coefficient therefore confirms that the clustering topology accurately reflects the morphological relationships among the studied wild rice genotypes. The satisfactory CCC value further supports the appropriateness of using the Jaccard coefficient for binary qualitative data, as it considers shared presences while excluding joint absences, thereby reducing bias in similarity estimation (Jaccard, 1908). The relatively high cophenetic correlation observed in this study suggests minimal distortion between the similarity matrix and dendrogram structure, thereby strengthening confidence in the inferred separation between *Oryza rufipogon* Griff. (Wainuchara) and *Oryza nivara* Sharma & Shastry (Murshi) accessions. Overall, the CCC result validates the robustness of the hierarchical clustering and confirms that the observed grouping pattern is statistically dependable. Such validation is critical in diversity studies, as reliable dendrogram representation ensures accurate interpretation of morphological differentiation and supports informed decision-making in germplasm conservation and breeding programmes.

Quantitative Character

Using 33 morphological descriptors, Zamora et al. (2003) documented substantial morphological diversity among three wild rice species: *Oryza grandiglumis* (Doell) Prod., *Oryza latifolia* Desv., and *O. glumaepatula*. Likewise, Arrieta-Espinoza *et al*. (2005) analyzed 27 morphological characters through multivariate techniques to determine morphometric relationships among wild *Oryza* species, commercial cultivars, and traditional landraces of cultivated rice. Wild rice accessions exhibited a wide array of unique morphological features, highlighting their potential value as donor parents for the introgression of desirable traits in rice breeding programs. So, an investigation on 37 quantitative traits was carried out on 10 wild rice accessions of Wainuchara and Murshi, collected from different ecological niches of Thoubal and Bishnupur districts of Manipur and recorded frequency, Mean, SD and coefficient of variation (CV) for each 10 accessions (analyse 5 plants for each accession) as on Table 4.

Table 4 Frequency, mean, SD and coefficient of variation (CV) values for 37 quantitative characters recorded in 10 accessions of Wainuchara and Murshi according to the National Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability (DUS) in rice (*Oryza sativa* L.)

Characters	WCM 001	WCM 002	WCM 003	WCM 004	WCM 005	WCM 006	MUR 001	MUR 002	MUR 003	MUR 004	Mean	SD	CV (%)
SH	15	11	18	14	14.5	15	10	8	8.5	8.4	12.24	3.49	28.55
PH	150	115	124	118	116	112	115	110	140	106	120.6	13.91	11.54
ML	2.5	2	2.4	4	3.9	4	1.5	1.2	1	4	2.65	1.23	46.53
COL	2.2	4.5	2.5	2	2.8	5	4.5	4.2	2	2.9	3.26	1.16	35.72
CUL	120	87	99.5	96	88.5	92	90	87	116.5	90.8	96.73	12.02	12.43
CUD	1	1.6	1.2	0.9	1.5	1	1.5	1.48	1.52	1.2	1.29	0.26	20.18
BIL	21	25	15	22.5	23.2	31	10.5	12	11.5	15	18.67	6.85	36.69
IEP	0.17	0.28	0.15	0.23	0.26	0.33	0.11	0.13	0.09	0.16	0.19	0.08	41.7
TPP	35	20	24	33	42	8	5	11	6	8	19.2	13.65	71.11
RL	15	16.2	18	12.5	10.25	8.5	5.2	8.5	9	5.5	10.86	4.41	40.59
PAL	30.6	28	24.5	22	27.5	20	25	23	23.25	15.2	23.9	4.37	18.29
PAW	3.5	3.57	4	3.2	3.9	4.5	10.6	11.5	11.5	11	6.73	3.83	56.94
DSP	65	61	72	75	70	55	50	48	55	46	59.7	10.46	17.52
PPBN	7	8	8	9	9	7	12	15	13	12	10	2.79	27.89
PSI	40	35	30	25.5	33.5	21	5	4.5	4.3	3	20.18	14.68	72.73
ANL	6.25	6.5	6	6.6	6.2	6	4.5	4.3	4.4	4.8	5.55	0.94	16.84
GSP	84	80	75	60	85	72	20	10	15	12	51.3	32.73	63.8
PEL	60.6	55	50	45	55.5	44	15	10.2	13	13.2	36.15	20.66	57.16
LLB	35	30	31	29.5	32.05	30	29	29.5	28.2	30.5	30.48	1.91	6.28
WLB	1.37	1.25	1.55	1.1	1.3	1.26	1.5	1.29	1.37	1.55	1.35	0.15	10.73
ALB	47.95	37.5	48.05	32.45	41.66	37.8	43.5	38.05	38.63	47.27	41.29	5.3	12.84
FLL	20	17	16.5	16.8	15.5	15	22	21.5	22	20	18.63	2.75	14.78
FLW	1.32	1.54	1.22	1.45	1.35	1.25	1.4	1.35	1.33	1.2	1.34	0.1	7.8
FLA	90.5	85	90	95	89	85	85	83	84	90.5	87.7	3.85	4.39
LIL	14.2	10.5	11	6.5	7	4	3.5	3.2	3.33	2	6.52	4.11	63.05
DTF	150	125	124	126	110	90	112	95	110	100	114.2	17.73	15.53
AWL	6	6.8	5.5	5.9	5.85	4.4	2.5	2.9	4.1	4.2	4.82	1.42	29.49
AWT	0.021	0.022	0.02	0.02	0.02	0.023	0.041	0.045	0.035	0.04	0.03	0.01	35.76
SLL	2.1	2.2	2	1.9	3	3.1	2	2.5	1.9	3.1	2.38	0.51	21.22
GRL	8.25	8.5	8	8.6	8.2	8	8.7	8.5	8.7	7.8	8.32	0.32	3.85
GRW	3.2	3.1	3.15	3	3.1	2.8	3.4	3.5	3.1	3.5	3.18	0.22	7.02

GLW	2.57	2.74	2.58	2.86	2.64	2.85	2.55	2.42	2.8	2.22	2.62	0.2	7.69
GRT	1.75	1.85	2	2	1.9	2	2.1	2.1	2.05	2.6	2.04	0.23	11.17
CRL	6.25	6.8	6.5	6.6	6.3	6.2	6.5	6.4	6.65	6.15	6.44	0.21	3.3
CRW	2.2	2.1	2.15	2	2.1	1.8	2.2	2.25	2	2.3	2.11	0.15	6.96
GW	2.11	2.1	2.1	2.05	2.01	1.99	1.9	1.85	2	2.05	2.02	0.09	4.29
HLP	18.33	20.3	15.6	16.33	18.33	16.7	20.2	21.22	23.22	20.9	19.11	2.45	12.84

Seedling Height (cm), Plant height (cm), Mesocotyl length (cm), Coleoptile length (cm), Culm length(cm), Culm diameter(mm), Basal Internode length(cm), Internode elongation pattern(internode length/culm lenth), Culm/Tiller number per plant, Root length, Panicle length (cm), Panicle weight(gm), Density (number of spikelet per panicle), Panicle: number of basal primary branches, Panicle: distance from base to lowest spikelet insertion(mm), Anther length (mm), Grain Shattering (Percentage), Peduncle length (cm), Length of Leaf blade(cm), Width of Leaf blade(cm), Area of Leaf Blade(sq. cm), Flag-leaf length (cm), Flag-leaf width (cm), Flag Leaf Angle(in degree), Ligule length (cm), Days to 50% flowering(DTF), Awn length (cm), Awn thickness(mm), Sterile lemma length(mm), Grain length (mm), Grain width (mm), Grain L/W ratio, Grain Thickness(mm), Caryopsis length(mm), Caryopsis width(mm), 100-grain weight (g), Hull Percentage(10 seeds)

Coefficient of Variation (CV%)

The Coefficient of Variation (CV%) was used to assess the extent of variability present among the studied wild rice genotypes for different quantitative traits. The CV% values obtained ranged from 3.30% to 72.73%, indicating the presence of both low and high levels of variability among the characters under study. The highest CV% was recorded for Panicle distance from base to lowest spikelet insertion (PSI) (72.73%), followed by Culm or tillers per Plant (TPP) (71.11%), Percentage of grain shattering (GSP) (63.80%), Ligule Length (LIL) (63.05%), Peduncle Length (PEL) (57.16%), and Panicle Weight (PAW) (56.94%), which reflects a high degree of phenotypic variability for these traits among the genotypes. In plant breeding, traits exhibiting high CV% are generally considered more amenable to selection due to the greater dispersion of phenotypic values, which may reflect underlying genetic variability (Burton & DeVane, 1953; Johnson *et al.,* 1955). Similarly, Mesocotyl length (ML) (46.53%), Internode elongation Pattern (IEP) (41.70%), Root Length (RL) (40.59%), Basal Internode Length (BIL) (36.69%), Awn Thickness (AWT) (35.76%), and Coleoptile Length (COL) (35.72%) also exhibited comparatively high CV% values, suggesting substantial variation and scope for selection in breeding programmes. Moderate variability was observed for traits such as Seedling Height (SH) (28.55%), Number of basal primary branches per panicle (PPBN) (27.89%), Awn Length (AWL) (29.49%), Sterile Lemma Length (SLL) (21.22%), and Culm Diameter (CUD) (20.18%). On the other hand, relatively low CV% values were recorded for Plant Height (PH) (11.54%), Culm Length (CUL) (12.43%), Density- number of spikelet per panicle (DSP) (17.52%), Anther Length (ANL) (16.84%), Days to 50% Flowering (DTF) (15.53%), and Flag Leaf Length (FLL) (14.78%), indicating lesser variability and comparatively stable expression of these traits. Very low CV% values were noted for grain-related traits such as Grain Length (GRL) (3.85%), Grain Width (GRW) (7.02%), Grain Thickness (GRT) (11.17%), Caryopsis Length (CRL) (3.30%), Caryopsis Width (CRW) (6.96%), Grain Weight (GW) (4.29%), and Flag Leaf Angle (FLA) (4.39%), suggesting minimal variation among the genotypes for these parameters. Conversely, the relatively low CV% observed for grain size parameters (e.g., grain length, width, thickness, and grain weight) and certain developmental traits such as days to 50% flowering (DTF) indicates comparatively stable expression and possible stabilizing selection acting on these characters, particularly those directly linked to reproductive success and domestication (Falconer & Mackay, 1996).

The presence of high CV% in several yield-contributing traits indicates considerable phenotypic variability, which can be effectively exploited for selection and genetic improvement in rice breeding programmes. The comparative analysis of morphological and yield-related traits

indicates that *O. rufipogon* (Wainuchara) exhibits greater vegetative vigour, characterized by taller plants, stronger seedling growth, longer mesocotyl and basal internodes, higher tiller number, longer roots, longer panicles, and greater grain shattering percentage, along with delayed flowering. Such traits are typical of perennial wild rice forms adapted to competitive natural environments and are often associated with higher genetic diversity and plasticity (Vaughan et al., 2003; Khush, 1997). In contrast, *Oryza nivara* (Murshi) demonstrates relatively compact growth, earlier flowering, reduced shattering, and heavier panicles—attributes that are consistent with annual life history strategies and traits associated with partial domestication or adaptation to seasonal habitats (Londo et al., 2006; Sweeney & McCouch, 2007). While both species show similarities in culm length, leaf area, seed length, and flag leaf width, the overall pattern suggests that *O. rufipogon* is more vigorous and structurally expansive, whereas *O. nivara* displays characteristics linked to earlier maturity and improved grain robustness. Therefore, the observed quantitative variability underscores the rich genetic reservoir present within these wild rice populations. The high CV% values for several yield-contributing and vegetative traits highlight the potential of *O. rufipogon* as a valuable donor for broadening the genetic base of cultivated rice, whereas *O. nivara* may contribute traits related to earliness and grain quality. These findings emphasize the importance of conserving and systematically evaluating wild rice germplasm for effective utilization in crop improvement programmes. The clear agro-morphological differentiation observed between *Oryza rufipogon* and *O. nivara* accessions in the present study is consistent with the findings of Sandamal et al. (2024), who identified distinct morphological groups corresponding to these two species in Sri Lanka. Their study reported that *O. rufipogon* generally exhibited larger values for several vegetative and reproductive traits, whereas *O. nivara* was characterized by shorter culms, earlier flowering, shorter and less-exserted panicles, and greater panicle compactness.

Table 5. Pearson's correlation coefficient of 10 wild rice accessions of Wainuchara and Murshi

	WCM001	WCM002	WCM003	WCM004	WCM005	WCM006	MUR001	MUR002	MUR003	MUR004
WCM001		0.99145	0.992833	0.985569	0.984158	0.974996	0.935577	0.915233	0.923226	0.914271
WCM002			0.992332	0.98508	0.988602	0.980387	0.920436	0.895744	0.89675	0.897614
WCM003				0.991402	0.988757	0.982781	0.941347	0.921804	0.92371	0.92353
WCM004					0.984068	0.973506	0.949402	0.935938	0.93266	0.935702
WCM005						0.977038	0.901227	0.882945	0.880969	0.883814
WCM006							0.929893	0.913751	0.919635	0.918986
MUR001								0.99491	0.98925	0.993612
MUR002									0.992529	0.995179
MUR003										0.98524
MUR004										

Hierarchical Cluster Analysis

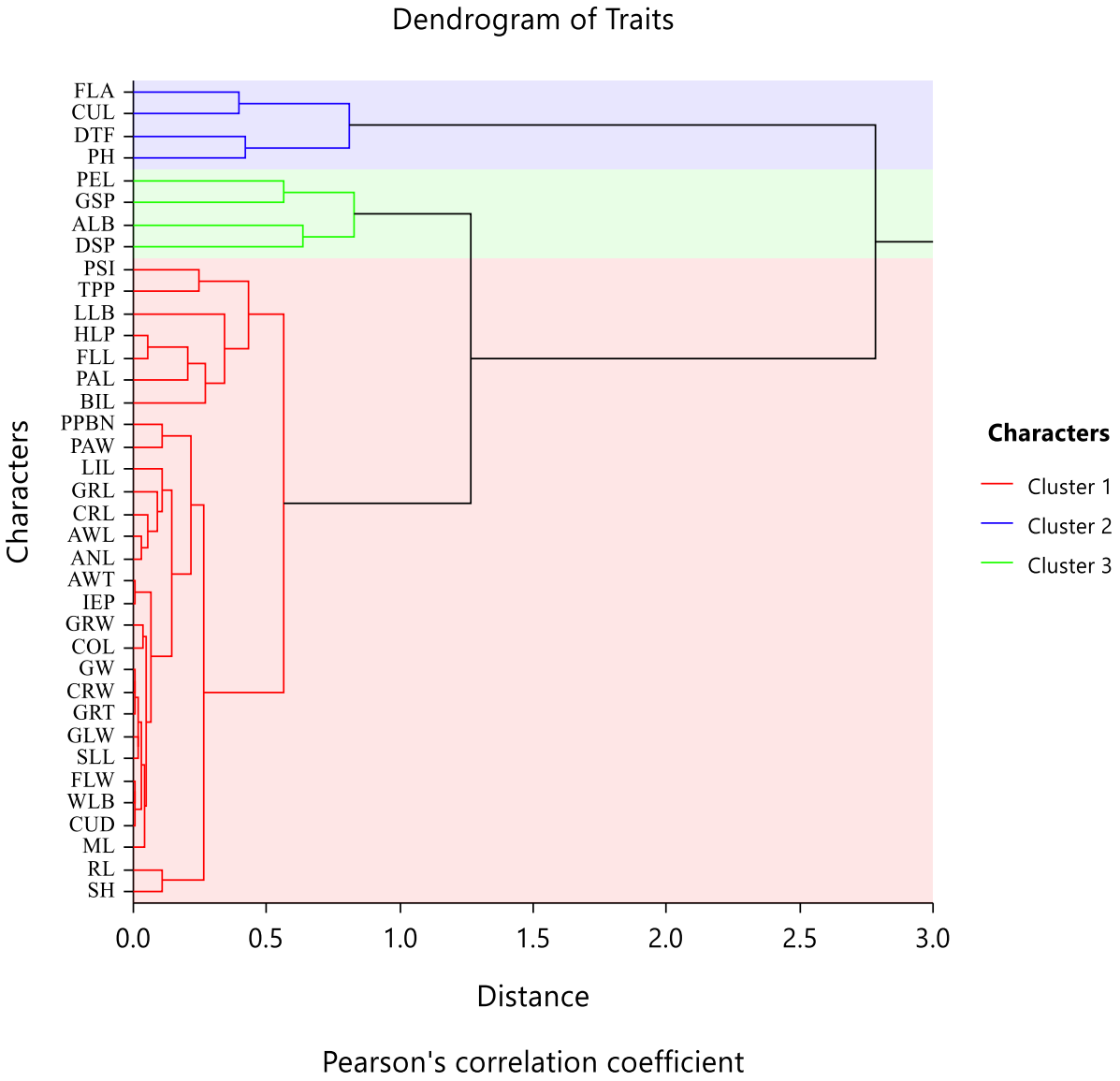


Fig. 8 Dendrogram of 10 wild rice accessions of Wainuchara and Murshi of Thoubal and Bishnupur District derived by UMPGA based on Pearson's correlation coefficient from 37 Quantitative traits

The clustering of quantitative traits based on Pearson's correlation coefficient provides insight into the structural and functional relationships among agronomic characters. The association of plant height, culm length, and flowering time within a single cluster suggests coordinated genetic control and possible pleiotropic effects influencing plant architecture and phenology, as reported in quantitative genetic studies (Falconer & Mackay, 1996). The dendrogram of quantitative traits generated using Pearson's correlation coefficient grouped the characters into three major clusters (figure 7), reflecting distinct patterns of linear association among the studied traits. Cluster II comprised phenological and architectural traits such as Flag Leaf Angle (FLA), Culm Length (CUL), Days to 50% Flowering (DTF), and Plant Height (PH), which merged at relatively low distance values, indicating strong positive correlations among growth duration and plant stature parameters. Cluster III included reproductive and yield-associated traits such as Peduncle Length (PEL), Grain Shattering Percentage (GSP), Area of Leaf blade (ALB), and Density of Spikelets per Panicle (DSP), suggesting coordinated expression of panicle architecture and reproductive behavior. Cluster I, the largest cluster, encompassed a wide range of vegetative vigor and grain-related traits including Seedling Height (SH), Mesocotyl Length (ML), Coleoptile length(COL), Culm diameter(CUD), Basal internode

length(BIL), Internode elongation pattern(IEP), Tillers per Plant (TPP), Root Length (RL), Panicle length(PAL), Panicle Weight (PAW), Panicle- number of basal primary branches(PPBN), Panicle- distance from base to lowest spikelet insertion(PSI), Anther length(ANL), Length of leaf blade(LLB), Width of leaf blade(WLB), Flag leaf length(FLL), Flag leaf width(FLW), Ligule Length (LIL), Awn length(AWL), Awn thickness(AWT), Sterile lemma length(SLL), Grain Length (GRL), Grain Width (GRW), Grain L/W ratio(GLW), Grain thickness(GRT), Caryopsis Length (CRL), Caryopsis Width (CRW), 100-Grain Weight (GW) and Hull percentage(HLP) among others. The largest cluster, comprising yield components and vegetative vigor traits, reflects the interdependence among tillering ability, panicle weight, grain dimensions, and early growth parameters. Such associations have been widely observed in multivariate analyses of rice, where yield is influenced by complex interactions among component traits (Mohammadi & Prasanna, 2003; Jolliffe, 2002). The presence of distinct clusters separated at higher linkage distances further indicates that these trait groups contribute independently to total phenotypic variability. From a breeding perspective, traits grouped within the same cluster may respond simultaneously to selection due to their correlated nature, whereas traits in separate clusters could be manipulated independently to achieve targeted genetic improvement. Therefore, the dendrogram not only elucidates the correlation structure among quantitative traits but also provides a practical framework for designing effective selection strategies in rice improvement programmes.

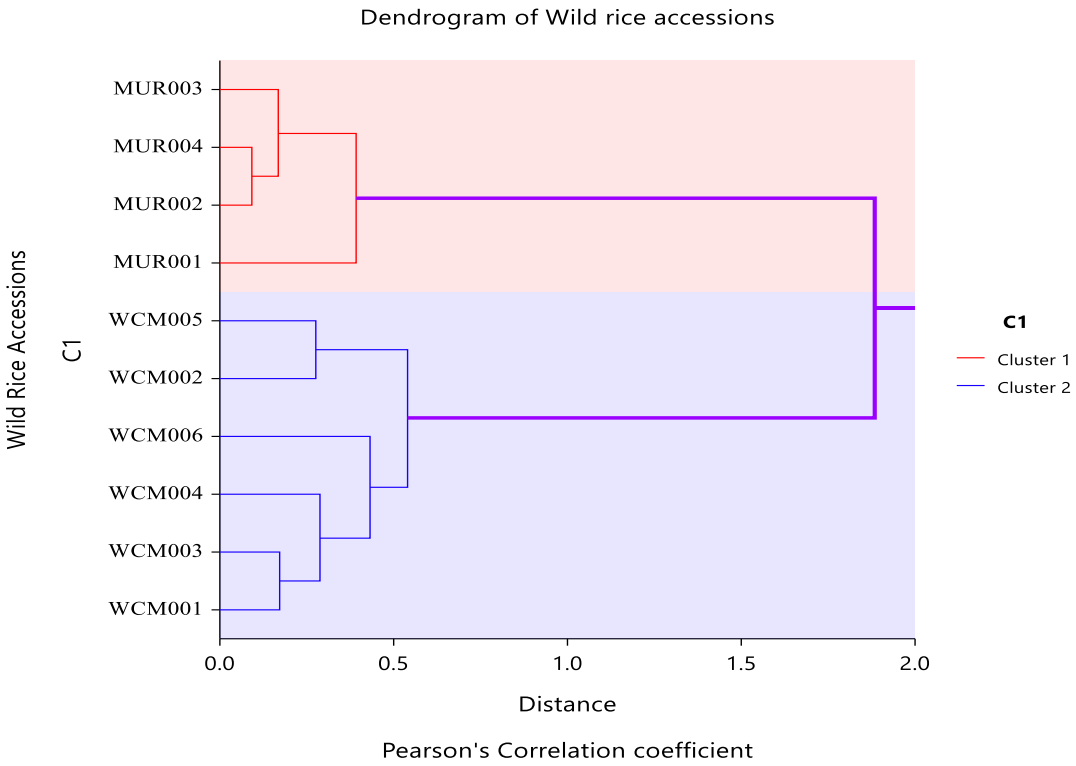


Fig. 9 Hierarchical Clustering of Wild Rice Accessions Based on Pearson's correlation coefficient (UPGMA Method)

The hierarchical cluster analysis based on Pearson's correlation coefficient and UPGMA method grouped the ten wild rice accessions into two well-defined major clusters at a linkage distance of approximately 1.9–2.0 (Figure 9). Cluster analysis using UPGMA is widely applied in phenotypic diversity studies to infer relationships among accessions based on multivariate similarity matrices (Sneath & Sokal, 1973). The clear bifurcation of Murshi and Wainuchara accessions into two major clusters suggests distinct genetic backgrounds, possibly shaped by geographical isolation, local adaptation, or restricted gene flow. Such clustering patterns have

been reported in wild and cultivated rice populations, where eco-geographical origin significantly influences genetic structure (Khush, 1997; Thomson et al., 2007). Cluster I consisted of four Murshi accessions (MUR003, MUR004, MUR002, and MUR001), while Cluster II comprised six Wainuchara accessions (WCM005, WCM002, WCM006, WCM004, WCM003, and WCM001). The large fusion distance separating these two clusters indicates substantial multivariate divergence between the Murshi and Wainuchara groups. Within Cluster I, MUR004 and MUR002 exhibited the closest similarity, merging at a very low distance (~0.15–0.20), followed by MUR003 (~0.25), whereas MUR001 joined the cluster at a comparatively higher distance (~0.45–0.50), suggesting moderate intra-group variability. In Cluster II, WCM003 and WCM001 formed the most similar pair (~0.20), followed by WCM004 (~0.30–0.35) and WCM006 (~0.50–0.55). Additionally, WCM005 and WCM002 clustered closely (~0.30) and merged with the larger WCM subgroup at ~0.60. The relatively low within-cluster distances (<0.60) indicate strong phenotypic similarity and potential shared ancestry. However, the slightly greater internal divergence observed within the WCM group suggests higher variability, which may reflect broader adaptation or greater genetic heterogeneity within that population. The pronounced inter-cluster divergence (~1.9) indicates substantial differentiation, implying that these two groups represent distinct genetic pools. From a breeding perspective, crosses between accessions from different clusters are expected to generate higher heterosis and broader recombination potential due to increased genetic distance (Falconer & Mackay, 1996). Moreover, the identification of two discrete genetic groups provides valuable information for germplasm conservation, as maintaining representatives from both clusters would maximize captured diversity. Overall, the dendrogram confirms strong within-group homogeneity and marked between-group divergence, highlighting the importance of geographical origin in structuring genetic variation among wild rice accessions. These findings provide a scientific basis for strategic parental selection and conservation planning in wild rice improvement programs.

Principal Component Analysis (PCA)

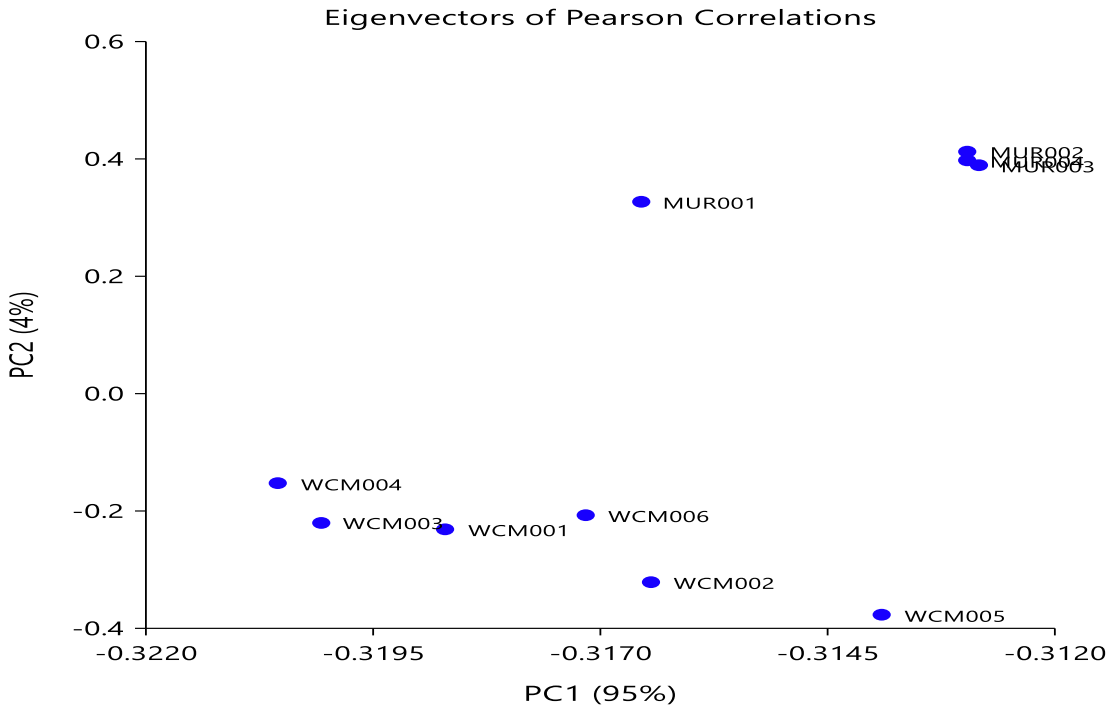


Fig. 10 Principal Component Analysis (PCA) Based on Pearson Correlation Matrix of Wild Rice Varieties

Table 6 Principal component analysis of 10 wild rice accessions - Eigen values, percentage variation and cumulative percentage of ten principal components derived from 37 quantitative characters

Component	Eigenvalues		
	Total	% of Variance	Cumulative %
1	9.548	95.477	95.477
2	.366	3.665	99.142
3	.031	.314	99.456
4	.022	.223	99.679
5	.014	.140	99.819
6	.007	.069	99.887
7	.005	.049	99.936
8	.004	.044	99.980
9	.001	.014	99.994
10	.001	.006	100.000

The principal component analysis (PCA) derived from the Pearson correlation matrix indicates that the first two principal components collectively explain 99% of the total variation, with PC1 contributing 95% and PC2 contributing 4% of the variability. Despite the narrow numerical range of PC1 scores (approximately −0.322 to −0.312), clear separation between varietal groups is evident along PC2. The MUR genotypes (MUR001, MUR002, MUR003, and MUR004) are positioned in the positive region of PC2 (≈0.33 to 0.41), with MUR002 and MUR004 showing the highest values (≈0.40–0.41), indicating close similarity among them. In contrast, the WCM genotypes (WCM001–WCM006) are located in the negative region of PC2 (≈−0.15 to −0.38), demonstrating distinct divergence from the Murshi group. Among the Wainuchara varieties, WCM005 exhibits the lowest PC2 value (≈−0.38), suggesting it is the most differentiated within its group, while WCM003 and WCM004 show relatively higher PC2 scores (≈−0.15 to −0.20), indicating comparatively closer similarity. The strong clustering of Murshi genotypes in the positive quadrant and Wainuchara genotypes in the negative quadrant reflects clear genetic or phenotypic structuring. Although PC1 accounted for the largest proportion of total variance, the separation between the two accession groups was visually more evident along PC2. Such pronounced group separation along principal components is consistent with multivariate differentiation patterns commonly observed in genetic diversity studies (Jolliffe & Cadima, 2016).

Cophenetic Correlation Coefficient (CCC)

The Cophenetic Correlation Coefficient (CCC) was calculated to assess how faithfully the dendrogram preserves the original pairwise distances and its value is found to be **0.935**. A CCC value greater than 0.90 indicates an excellent representation of the original distance matrix. Therefore, the hierarchical clustering model provides a highly reliable structural representation of the dataset.

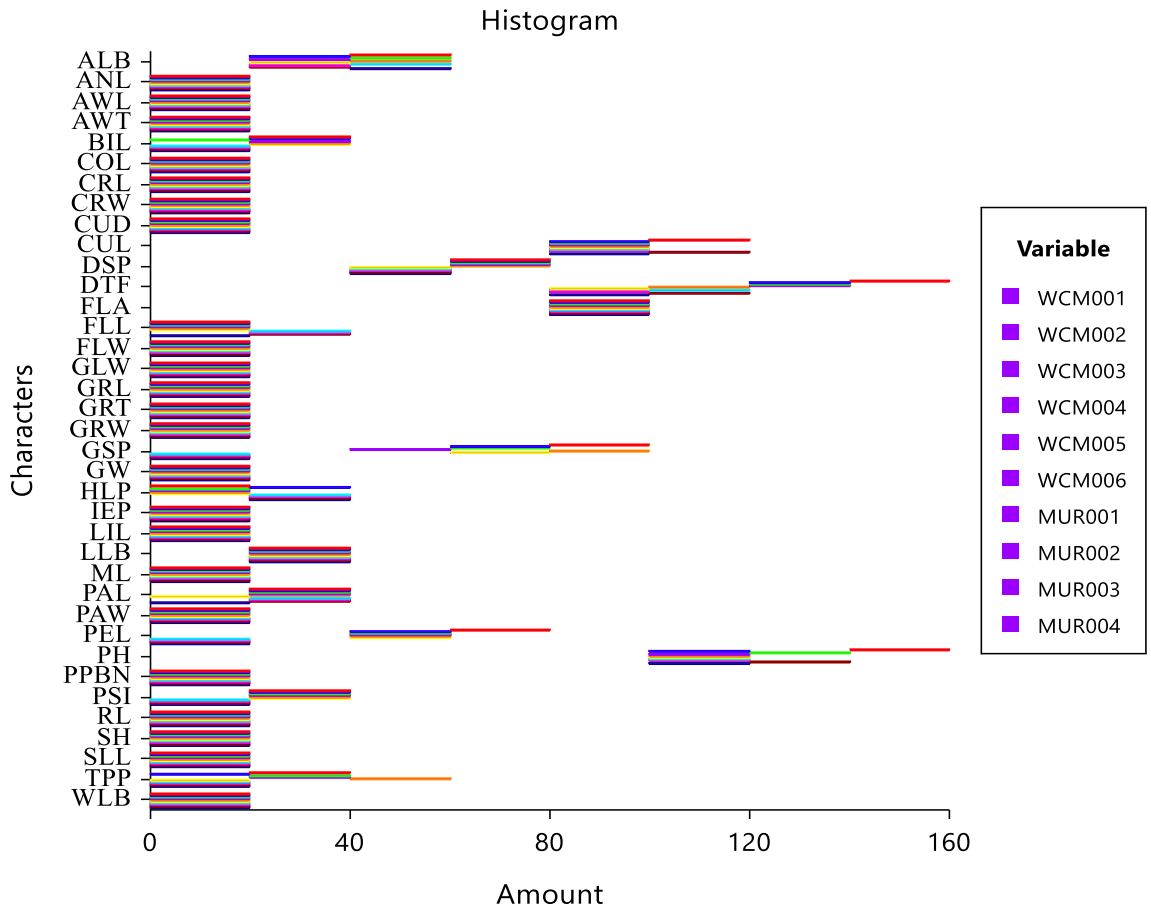


Fig. 11 Frequency-histograms of 37 qualitative traits of 10 wild rice accessions of Wainuchara and Murshi collected from Thoubal and Bishnupur district of Manipur

CONCLUSIONS

Breeding programmes depend on the availability of high level of phenotypic diversity for achieving progress through wide-hybridization. Wild genetic resources are the rich source of genes for genetic improvement of the cultivated rice varieties. The phenotyping studies in wild genetic resources are essential to provide information for their utilization in hybridization, wild relatives management and conservation programmes. In this study, we analyzed 10 wild rice accessions collected from different ecological niches and ex-situ of Thoubal and Bishnupur districts of Manipur. The study shows that a high level of phenotypic diversity exists in Indian wild rice collections from different ecological niches. This set of wild rice germplasm with sufficient variability provides opportunity of their possible use in allele mining and rice breeding.

The integrated qualitative and quantitative analyses collectively demonstrate substantial and structured phenotypic diversity between Wainuchara (*Oryza rufipogon* Griff.*)* and Murshi (*Oryza nivara* Sharma & Shastry*)*, confirming the presence of distinct genetic groupings within the evaluated wild rice populations. Qualitative morphological descriptors effectively discriminated Murshi and Wainuchara accessions, while quantitative traits revealed high coefficients of variation for several yield-contributing and vegetative characters, indicating a rich and exploitable genetic reservoir. The consistent separation of populations in dendrogram and principal component analyses highlights strong within-group homogeneity and marked between-group divergence, largely influenced by geographical origin and dominant sources of variation. The identification of correlated trait complexes and distinct cluster groupings suggests that certain characters may respond simultaneously to selection, whereas others can be independently manipulated for targeted improvement. Notably, *O. rufipogon* shows promise as

a donor for yield enhancement and genetic base broadening, while *O. nivara* may contribute traits related to earliness and grain quality. Overall, the combined multivariate approaches provide a robust and cost-effective framework for germplasm characterization, strategic parental selection, heterosis exploitation, and conservation planning in wild rice–based crop improvement programs.

REFERENCES

Arrieta-Espinoza, G., Sanchez, E., Vargas, S., Lobo, J., Quesada, T., & Espinoza, A. M. (2005). The weedy rice complex in Costa Rica: I. Morphological study of relationships between commercial rice varieties, wild *Oryza* relatives and weedy types. *Genetic Resources and Crop Evolution*, *52*(5), 575–587

Bioversity International, International Rice Research Institute, & WARDA. (2007). *Descriptors for wild and cultivated rice (Oryza spp.)*. Bioversity International

Brar, D. S., & Khush, G. S. (1997). Alien introgression in rice. *Plant Molecular Biology*, *35*(1–2), 35–47

Brar, D. S., & Khush, G. S. (2003). Utilization of wild species of genus Oryza in rice improvement. In J. S. Nanda & S. D. Sharma (Eds.), *Monograph on genus Oryza* (pp. 283–309). Oxford & IBH Publishing. [1, 2]

Burton, G. W., & DeVane, E. H. (1953). Estimating heritability in tall fescue from replicated clonal material. *Agronomy Journal*, *45*(10), 478–481

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics* (4th ed.). Longman, Essex. DOI:10.1093/genetics/167.4.1529

Fowler, C., Mooney, P. 1990. Shattering: Food, Politics and Loss of Genetic diversity. University of Arizona Press. Tucson.

Fuller, D. Q., Sato, Y. I., Castillo, C., et al. (2009). Consilience of genetics and archaeobotany in the entangled history of rice. *Archaeological and Anthropological Sciences*, 1, 115–131.

Guo, D., Li, Y., Lu, H., Zhao, Y., Kurata, N., Wei, X., Wang, A., Wang, Y., Zhan, Q., Fan, D., Zhou, C., Lu, Y., Tian, Q., Weng, Q., Feng, Q., Huang, T., Zhang, L., Gu, Z., Wang, C., … Han, B. (2025). A pangenome reference of wild and cultivated rice. *Nature*. https://doi.org/10.1038/s41586-025-08883-6

Hallauer, A. R., Carena, M. J., & Miranda Filho, J. B. (2010). *Quantitative Genetics in Maize Breeding* (3rd ed.). Springer, New York. DOI:10.1007/978-1-4419-0766-0

Hamrick, J. L., & Godt, M. J. W. (1996). Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1345), 1291–1298. doi:10.1098/rstb.1996.0112

Holden, J. H. W., Peacock, W. J., & Williams, J. T. (1993). *Genes, crops and the environment*. Cambridge University Press. [1] DOI:10.1017/S0014479700024546

Hore, D. K., & Sharma, B. D. (1993). Wild rice genetic resources of North-East India. *Indian Journal of Plant Genetic Resources*, 6(1), 27–32. [1]

IRRI (International Rice Research Institute). (2013). *Standard Evaluation System for Rice (SES)* (5th ed.). IRRI, Los Baños, Philippines.

Jaccard, P. (1908). Nouvelles recherches sur la distribution florale [New research on floral distribution]. *Bulletin de la Société Vaudoise des Sciences Naturelles*, *44*(163), 223–270

Johnson, H. W., Robinson, H. F., & Comstock, R. E. (1955). Estimates of genetic and environmental variability in soybeans. *Agronomy Journal*, 47, 314–318. DOI:10.2134/AGRONJ1955.00021962004700070009X

Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.). Springer. https://doi.org/10.1007/b98835 [1]

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A*, 374(2065), 20150202. doi: 10.1098/rsta.2015.0202

Khush, G. S. (1977). Disease and insect resistance in rice. Adv Agron 29:265–341

Khush, G. S. (1989). Multiple disease and insect resistance for increased yield stability in rice. In *Progress in irrigated rice research* (pp. 79–92). International Rice Research Institute. [1, 2]Khush GS (1997) Origin, dispersal, cultivation and variation of rice. Plant Mol Biol 35:25–34

Khush, G. S. (2005). What it will take to feed 5.0 billion rice consumers in 2030. *Plant Molecular Biology*, *59*(1), 1–6. http://dx.doi.org/10.1007/s11103-005-2159-5

Khush, G. S. (2013). Strategies for increasing the yield potential of cereals: Case of rice as an example. *Plant Breeding*, *132*(5), 433–436. **https://doi.org/10.1111/pbr.1991**

Legendre, P., & Legendre, L. (2012). *Numerical Ecology* (3rd ed.). Elsevier, Amsterdam

Lin, M. S. (1991). Genetic base of japonica rice varieties released in Taiwan. *Euphytica*, *56*(1), 43–46. https://doi.org/10.1007/BF00041742 [1, 2]

Londo, J. P., Chiang, Y. C., Hung, K. H., Chiang, T. Y., & Schaal, B. A. (2006). Phylogeography of Asian wild rice (*Oryza rufipogon*) reveals multiple independent domestications of cultivated rice (*Oryza sativa*). *Proceedings of the National Academy of Sciences USA*, 103, 9578–9583. doi: 10.1073/pnas.0603152103

Long, W., He, Q., Wang, Y., Wang, Y., Wang, J., Yuan, Z., Wang, M., Chen, W., Luo, L., Luo, L., Xu, W., Li, Y., Li, W., Yan, L., Cai, Y., Du, H., & Xie, H. (2024). Genome evolution and diversity of wild and cultivated rice species. *Nature Communications, 15*, 9994. https://doi.org/10.1038/s41467-024-54427-3

McCouch, S. R., McNally, K. L., Wang, W., & Hamilton, R. S. (2012). Genomics of gene banks: A case study in rice. *American Journal of Botany*, *99*(2), 407–423.[1, 2, 3] DOI: 10.3732/ajb.1100385

McCouch, S. R., Baute, G. J., Bradeen, J., Bramel, P., Bretting, P. K., Buckler, E., Burke, J. M., Charest, D., Cloutier, S., Cole, G., Dempewolf, H., Dingkuhn, M., Feuillet, C., Gepts, P., Grattapaglia, D., Guarino, L., Jackson, S., Knapp, S., Langridge, P., … Zamir, D. (2013). Feeding the future. *Nature, 499*(7456), 23–24. https://doi.org/10.1038/499023a

Mohammadi, S. A., & Prasanna, B. M. (2003). Analysis of genetic diversity in crop plants—Salient statistical tools and considerations. *Crop Science*, *43*(4), 1235–1248. DOI:10.2135/CROPSCI2003.1235

Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences USA*, 76, 5269–5273. https://doi.org/10.1073/pnas.76.10.5269

Patra, B. C., Dhua, S. R., Marndi, B. C., Nayak, P. K., Swain, P., Kumar, G. A. K., & Singh, K. (2008). Exploration, collection, characterization, evaluation and conservation of wild rice germplasm of east India. *Oryza*, *45*(2), 98–102. [1]

Ren, F., Lu, B. R., Li, S., et al. (2003). A comparative study of genetic relationships among the AA-genome species of the genus *Oryza* using RAPD and SSR markers. *Theoretical and Applied Genetics*, 108, 113–120. DOI: 10.1007/s00122-003-1414-x

Rohlf, F. J. (2000). *NTSYSpc: Numerical taxonomy and multivariate analysis system* (Version 2.1) [Computer software]. Exeter Software. [1]

Sandamal, S., Tennakoon, A., Wijerathna, P., Zhang, H.-X., Yu, W.-H., Qiang, C.-G., Han, J.-D., Zhang, F.-M., Ratnasekera, D., & Ge, S. (2024). Phenological and morphological variations of *Oryza rufipogon* and *O. nivara* in Sri Lanka and their evolutionary implications. *Scientific Reports, 14*, 31126. https://doi.org/10.1038/s41598-024-82383-x

Samal, R., Roy, P. S., Sahoo, A., Kar, M. K., Patra, B. C., Marndi, B. C., & Gundimeda, J. N. R. (2018). Morphological and molecular dissection of wild rices from eastern India suggests distinct speciation between *O. rufipogon* and *O. nivara* populations. *Scientific Reports, 8*, 2773. https://doi.org/10.1038/s41598-018-20693-7

Saraçli, S., Doğan, N., & Doğan, İ. (2013). Comparison of hierarchical cluster analysis methods by cophenetic correlation. *Journal of Inequalities and Applications*, *2013*(1), Article 203. DOI**:**10.1186/1029-242X-2013-203

Shiva, V. (1991). *The violence of the green revolution: Third World agriculture, ecology, and politics*. Third World Network. [1, 2]

Shen, X. H., Yan, S., Huang, R. L., Zhu, S., Xiong, H. L., & Shen, L. J. (2013). Development of novel cytoplasmic male sterile source from Dongxiang wild rice (*Oryza rufipogon*). *Rice Science, 20*(5), 379–382. https://doi.org/10.1016/S167

Singh, A., Singh, B., Panda, K., Rai, V. P., Singh, A. K., Singh, S. P., Chouhan, S. K., Rai, V., Singh, P. K., & Singh, N. K. (2013). Wild rices of eastern Indo-Gangetic plains of India constitute two sub-populations harbouring rich genetic diversity. *Plant Omics Journal*, *6*(2), 121–127. semanticscholar.org

Singh, B., Singh, N., Mishra, S., Tripathi, K., Singh, B. P., Rai, V., Singh, A. K., & Singh, N. K. (2018). Morphological and molecular data reveal three distinct populations of Indian wild rice *Oryza rufipogon* Griff. species complex. *Frontiers in Plant Science*, *9*, Article 123. https://doi.org/10.3389/fpls.2018.00123

Sneath, P. H. A., & Sokal, R. R. (1973). *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. W. H. Freeman, San Francisco.

Sweeney, M., & McCouch, S. (2007). The complex history of the domestication of rice. *Annals of Botany*, 100, 951–957. https://doi.org/10.1093/aob/mcm128

Tanksley, S. D., & McCouch, S. R. (1997). Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science*, 277, 1063–1066. DOI:10.1126/SCIENCE.277.5329.1063

Tao, H., Wu, S., Wu, G., Chen, X., Sun, Y., Zhu, X., Chen, L., Feng, C., Kabore, M. A. F., Woldegiorgis, S. T., Zhang, L., Ai, Y., Liu, W., & He, H. (2026). Pan-genome of wild and cultivated rice uncovers genetic diversity, lost and selected sequences during rice domestication. *Rice, 19*, 21. https://doi.org/10.1186/s12284-026-00893-w

Thomson, M. J., Septiningsih, E. M., Suwardjo, F., Santoso, T. J., Silitonga, T. S., & McCouch, S. R. (2007). Genetic diversity analysis of traditional and improved Indonesian rice (*Oryza sativa* L.) germplasm using microsatellite markers. *Theoretical and Applied Genetics*, *114*(4), 559–568.

Vaughan, D. A., & Morishima, H. (2003). Biosystematics of the genus *Oryza*. In C. W. Smith & R. H. Dilday (Eds.), *Rice: Origin, history, technology, and production* (pp. 27–65)

Xiao, J., Grandillo, S., Ahn, S. N., McCouch, S. R., Tanksley, S. D., Li, J., & Yuan, L. (1996). Genes from wild rice improve yield. *Nature, 384*(6606), 223–224. https://doi.org/10.1038/384223a0

Yan, S., Zhu, S., Mao, L.-H., Huang, R.-L., Xiong, H.-L., Shen, L.-J., & Shen, X.-H. (2017). Molecular identification of the cytoplasmic male sterile source from Dongxiang wild rice (*Oryza rufipogon* Griff.). *Journal of Integrative Agriculture, 16*(8), 1669–1675. https://doi.org/10.1016/S2095-3119(16)61505-8

Zamora, A., Barboza, C., Lobo, J., & Espinoza, A. M. (2003). Diversity of native rice (*Oryza*, Poaceae) species of Costa Rica. *Genetic Resources and Crop Evolution*, *50*(8), 855–870. DOI:10.1023/A:1025998801637

Zhao, Q., Han, B., & colleagues. (2024). Genome evolution and diversity of wild and cultivated rice species**.** *Nature Communications, 15*, Article 10358. https://doi.org/10.1038/s41467-024-54427-3

Zheng, X., Peng, Y., Qiao, J., Henry, R. J., & Qian, Q. (2024). Wild rice: Unlocking the future of rice breeding. *Plant Biotechnology Journal, 22*(11), 3218–3226. https://doi.org/10.1111/pbi.14443

ACKNOWLEDGMENTS

The authors are thankful to Department of Botany, Dhanamanjuri University, Manipur, India for providing necessary facilities to conduct the research. Authors are also thankful to Dr. Medhabati Kangabam, Dhanamanjuri University, Imphal and Dr. Athokpam Pinokiyo, HOD, Department of Botany, Dhanamanjuri University, Imphal to give constructive comments for improvement of the manuscript.

FUNDING

This work has been done without any financial support.

AUTHOR INFORMATION

Authors and Affiliations

1. Department of Botany, Dhanamanjuri University, Manipur, India

   Syed Saddam Hussain

2. Department of Botany, Dhanamanjuri University, Manipur, India

Dr. Medhabati Kangabam

CONTRIBUTIONS

Syed S H conceptualized designed the article, data collection, designed the methodology, wrote the manuscript, and approved the final version.

DATA AVAILABILITY

The datasets analysed during the current study are not publicly available due to ongoing research and data ownership considerations but are available from the corresponding author on reasonable request.”

CORRESPONDING AUTHOR

Correspondence to SYED SADDAM HUSSAIN.

ETHICS DECLARATIONS

Consent for publication
Not applicable.
Conflict of interest
The authors declare that they have no any financial or personal relationship that could affect this research work.
Ethical approval
Not Applicable

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable1.docx
- Supplementary2.docx
- Supplementary3.docx
- BinaryDataMatrix.xlsx