

Do the six types of Sri Lanka's *Oryza* comprise a domesticated-weed-wild complex?

Asanka Tennakoon¹, Salinda Sandamal², Song Ge², Arthur Melo³, Yao Zhao⁴, Buddhi Marambe⁵, Norman Ellstrand⁶, Beng-Kah Song⁷, and Disna Ratnasekera⁸

¹Eastern University Sri Lanka

²Institute of Botany Chinese Academy of Sciences

³University of New Hampshire

⁴Nanchang University School of Life Sciences

⁵University of Peradeniya Faculty of Agriculture

⁶University of California Riverside

⁷Monash Univ Malaysia

⁸University of Ruhuna

September 27, 2023

Abstract

Genetic studies of Domesticated-Weed-Wild Complexes (DWWC) have typically focused on one-way introgression of crop alleles into wild or weedy populations, with little consideration of the entire natural ecosystem. In Sri Lanka, DWWC is diverse, comprising six evolutionarily discrete groups in the genus *Oryza*. Using 33 neutral simple sequence repeat (SSR) loci, we characterized six *Oryza* groups to understand the genetic background and evolution of DWWC components. Our analysis found that *Oryza* groups have large population sizes and high inter-group long-term gene flows. Asymmetric gene flows were found between wild and weedy rice groups, but the rare alleles shared among DWWC components provide additional evidence for extensive and enduring exchange, highlighting the dynamic nature of this complex genetic admixture among different *Oryza* lineages. We found high genetic diversity at the population and species levels due to mixed DWWC components over the generations. Weedy rice types exhibit genetic incorporation through admixture from both crop and wild species, highlighting the multi-way genetic transfer in the evolution of weedy rice types. Our findings support the idea that the DWWC is an integrated complex in the Sri Lankan rice ecosystem and that its weedy rice has multiple origins, including de-domestication via feralization of cultivated rice, inter-varietal hybridization among distinct cultivated rice types, adaptation, and invasion of rice cultivation areas by wild *Oryza* species, and hybridization events between crop and wild rice populations. Abandoned rice domesticates can also evolve into weedy forms with less intimate human relationships and contaminate the rice ecosystem.

Do the six types of Sri Lanka's *Oryza* comprise a domesticated-weed-wild complex?

Asanka Tennakoon^{a,b,1}, Salinda Sandamal^{c,d,1}, Song Ge^{c,d}, Arthur Melo^e, Yao Zhao^f, Buddhi Marambe^g, Norman C. Ellstrand^h, Beng-Kah Song^{i,2}, Disna Ratnasekera^{a,2}

^aDepartment of Agricultural Biology, Faculty of Agriculture, University of Ruhuna, 81100, Matara, Sri Lanka; ^bDepartment of Biosystems Technology, Faculty of Technology, Eastern University, 30350, Sri Lanka; ^cState Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, 100093, China; ^dUniversity of Chinese Academy of Sciences, Beijing, 100049, China; ^eInova Genética LTDA, Santa Barbara Farm, BR 251, Km18, Cristalina, Goiás, Brazil; ^fJiangxi Province Key Laboratory of Watershed Ecosystem Change and Biodiversity, Center for Watershed Ecology, Institute

of Life Science and School of Life Sciences, Nanchang University, Nanchang, 330031, China; ^gDepartment of Crop Science, Faculty of Agriculture, University of Peradeniya, Kandy, 20400, Sri Lanka; ^hDepartment of Botany and Plant Sciences, University of California, Riverside, CA, USA; ⁱSchool of Science, Monash University Malaysia, 46150, Bandar Sunway, Selangor, Malaysia

¹should be considered joint first author.

²should be considered joint corresponding author.

Corresponding author email:disnaratnasekera@gmail.com

Corresponding author address: Department of Agricultural Biology, Faculty of Agriculture, University of Ruhuna, 81100, Matara, Sri Lanka

Abstract

Genetic studies of Domesticated-Weed-Wild Complexes (DWWC) have typically focused on one-way introgression of crop alleles into wild or weedy populations, with little consideration of the entire natural ecosystem. In Sri Lanka, DWWC is diverse, comprising six evolutionarily discrete groups in the genus *Oryza*. Using 33 neutral simple sequence repeat (SSR) loci, we characterized six *Oryza* groups to understand the genetic background and evolution of DWWC components. Our analysis found that *Oryza* groups have large population sizes and high inter-group long-term gene flows. Asymmetric gene flows were found between wild and weedy rice groups, but the rare alleles shared among DWWC components provide additional evidence for extensive and enduring exchange, highlighting the dynamic nature of this complex genetic admixture among different *Oryza* lineages. We found high genetic diversity at the population and species levels due to mixed DWWC components over the generations. Weedy rice types exhibit genetic incorporation through admixture from both crop and wild species, highlighting the multi-way genetic transfer in the evolution of weedy rice types. Our findings support the idea that the DWWC is an integrated complex in the Sri Lankan rice ecosystem and that its weedy rice has multiple origins, including de-domestication via feralization of cultivated rice, inter-varietal hybridization among distinct cultivated rice types, adaptation, and invasion of rice cultivation areas by wild *Oryza* species, and hybridization events between crop and wild rice populations. Abandoned rice domesticates can also evolve into weedy forms with less intimate human relationships and contaminate the rice ecosystem.

Keywords: Domesticated-weed-wild complexes, Origin and evolution, *Oryza* spp., Weedy rice

1 Introduction

The evolution of crop species may extend beyond the control of a breeder. Perhaps the most dynamic opportunities for spontaneous evolution are in regions rich with reproductively compatible plants in both managed and unmanaged populations. For example, managed populations may consist of both landraces and inbred rice varieties, while unmanaged populations may include both wild rice and weedy rice, descended from the crop or wild species from which the crop has evolved. According to Molina *et al.* (2011), the rice crop has been domesticated from wild *Oryza* species, specifically *O. rufipogon* Griff. in Asia and *O. barthii* A. Shiv. in Africa. It is ecologically logical for the epicenters of rice cultivation to have weedy types that are admixtures of the wild relatives, off-types of the selected and cultivated strains, and various outcrosses (Pusadee *et al.*, 2013; Veltman *et al.*, 2019) as they co-exist in adjoining habitats. In such regions, gene flow between cultivated plants, wild relatives, and their weedy descendants present would generate a variation on which artificial and natural selections could act. Such "domesticated-weed-wild complexes" (hereafter, DWWC) are poised to be genetically rich meta-populations of evolutionary experimentation. While it is easy to tentatively identify such potential complexes by the occurrence of visually apparent hybrids, it is much more difficult to measure whether the gene flow plays a role in the evolution of co-occurring reproductive-compatible relatives (Pusadee *et al.*, 2013; Song *et al.*, 2014). Therefore, the apparent hybrids might be aberrant individuals, or if they are true hybrids, they might be sterile evolutionary dead ends. Genetic analysis is necessary to document whether and how gene exchange occurs (Ellstrand, 2003; Linder *et al.*, 1998; Leak-Garcia *et al.*, 2012). Using a panel of 638 domesticated and wild rice genomes, Wang *et al.*

(2017) examined the genetic ancestry of wild rice and demonstrated that most modern wild rice types are heavily admixed with domesticated rice due to gene flow from both pollen and seeds. This suggests that what has been thought of as wild rice may actually be different stages of feralized domesticated rice. Further, the Southeast Asian wild-like weedy rice may have originated from weedy descendants of coexisting cultivars (Sudoet al ., 2021; Neik et al ., 2019).

Most genetic studies of DWWC have focused on one-way introgression of crop alleles into wild or weedy populations (Ellstrand, 2003), whereas only a few studies have examined multi-way gene flow among all members of the DWWC to evaluate whether it is an integrated "complex". The island nation of Sri Lanka is an ideal system to study the multilateral gene flow of a DWWC due to its geographical isolation from the rest of South Asia, and its highly diverse topography and climate. The DWWC on the island involves six evolutionarily distinct groups in the *Oryza* genus, including landraces, inbred rice varieties, two wild species, feral populations, and weedy rice. Landraces (traditional rice accessions) have grown for more than two and a half millennia across the island (Weerakoon et al., 2011) and inbred rice varieties cultivated since the 1960s (Dhanapala, 2020). Reproductively compatible free-living wild relatives of the rice include two wild species, the annual *O. nivara* and the perennial *O. rufipogon* . Self-sustaining feral populations that have emerged from cultivated fields abandoned (approximately ten years prior) represent an additional category of free-living lineages (Qiu et al., 2020). Nevertheless, their inclusion as part of the DWWC has not yet been recognized. The latter thus presents an unusual opportunity as it represents the early portion of the evolutionary trajectory to de-domestication (Ellstrand et al., 2010). The sixth category is one of the world's worst rice weeds, weedy rice (*Oryza* spp.). Unlike feral populations, which are sustained without human intervention, weedy rice is typically restricted to cultivated rice fields, making it a significant threat to sustainable rice production in Sri Lanka (Ratnasekera, 2015). Given the six categories, there are thirty potential pathways for gene flow among the Sri Lankan *Oryza* types if all of them participate in the DWWC.

Therefore, our study aimed to use SSR loci to characterize the populations of the entire array of Sri Lankan *Oryza* categories and address the following research questions: (1) What is the influence of the genetic background of DWWC components and their participation in an integrated complex on the population? (2) To what extent does multi-way gene flow exist among all members of the DWWC? (3) What role has the DWWC played in the origin and evolution of weedy rice in Sri Lanka?

2 Materials and Methods

2.1 Plant Materials

Six distinct groups within the domesticated-weed-wild complex (DWWC) in the genus *Oryza* , sharing a common ancestry, were sampled from Sri Lanka for this study. These groups include inbred rice, landraces, feral rice, weedy rice, *O. rufipogon* , and *O. nivara* .

Seeds of 42 inbred rice varieties (206 individuals; Table S3) were collected from the Rice Research and Development Institute (RRDI) of Sri Lanka. Seeds of 31 landraces (151 individuals) with 4-5 individuals per variety were collected from the seed bank of the Plant Genetic Resources Centre (PGRC) at Gannoruwa, Peradeniya, Sri Lanka (Table S2). Both inbred and landraces (5-10 seeds of each cultivar) were planted in equal-sized pots (9 cm diameter top, 6.3 cm diameter base, 8.7 cm depth, and capacity: 0.37 liter) filled with sieved paddy soil and grown under greenhouse conditions without added fertilizer. After two weeks, the young tender leaves (3-5 leaves) from randomly selected 3-5 plants were harvested from each cultivar and stored at -80 °C until extraction of DNA.

The other free-living *Oryza* populations, including weedy, feral, and wild rice, were sampled from various locations across Sri Lanka. To prevent repeated samples of the same genets (clones), the distance between sampled individuals was maintained at least 10 m apart. The homogeneity of inbred rice makes it easy to distinguish weedy rice by its characteristic traits, such as differences in seed and panicle features, typically exhibiting taller stature, profuse tillering, a more open or spreading growth habit, and weaker culms. Weedy rice plants were collected from 20 major rice-growing areas (20 populations) in Sri Lanka that had a recent history of heavy weed infestation, as confirmed through personal communication with the Department of

Agriculture, Sri Lanka (see Fig. 1). One population was defined as approximately 25 weedy rice plants randomly sampled from each field, which covered an area of approximately one hectare. Thus, a total of 500 individual weedy rice plants were collected across the entire geographical distribution in Sri Lanka. The sampled rice fields infested with weedy rice were at least 30 km apart from each other. For this study, six *O. nivara* (183 individuals) and five *O. rufipogon* (132 individuals) populations were sampled, comprising 18 to 31 individuals per population and representing the diverse wet, dry, and intermediate zones of Sri Lanka. Feral rice consisted of *O. sativa* populations that emerge spontaneously in paddy fields abandoned for approximately 10 years. These plants are presumed descendants of previously cultivated inbred or landrace varieties. Unlike weedy rice, feral rice is not considered a weed in paddy fields. In this study, we sampled seven feral rice populations, collecting 24 individuals per population (168 individuals in total), from abandoned paddy fields for further analysis. Green leaf samples were collected from wild, weedy and feral rice in the *Maha* season from late February to March 2016. The green leaves were collected individually in the field, placed in zip-lock bags containing silica gel, and stored at 0-4 °C until DNA extraction. Information on population codes, the number of individuals sampled, locations, and GPS readings of the weedy rice populations are given in Table S1.

Fig. 1. Geographical locations of 20 weedy rice (blue circles), five *Oryza rufipogon* (green triangles), six *Oryza nivara* (yellow triangles) and seven feral rice (the red squares) populations sampled in this study. Dark lines separate the climatic zones in this study. Detailed information on these populations is provided in Table S1.

Table S1

Table S2

Table S3

2.2 Molecular Screening

Molecular studies were conducted at the State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China, using the 20 mg silica gel-dried leaves collected from six *Oryza* types from Sri Lanka. Preliminary screening of 150 pairs of simple sequence repeat (SSR) loci across representative individuals sampled from different populations revealed that 33 SSRs produced clear and reproducible products. These 33 SSR primer pairs were designed based on cultivated rice sequences from the RiceGenes Database (<https://gramene.org>), and used to assess the genetic variation of DWWC in this study. The 33 loci were dispersed in all 12 chromosomes in the rice genome, with the polymorphism varying widely among loci. Detailed information on the primer pairs is given in Table S4. Primers were synthesized by ABI DNA Analyzer (Applied Biosystems, Foster City, CA, USA), with the forward primers labelled with blue (FAM), yellow (TAMAR), and green (HEX) fluorophores.

2.3 DNA Extraction and PCR Amplification

The total genomic DNA was extracted from dried leaves of 1340 samples using the Plant genomic DNA kit (Biomed DL114-01, China) following the CTAB protocol (Saghai-Maroo *et al.*, 1984) with minor modifications. The DNA amplification was carried out using a 2,720 thermal cycler (Applied Biosystems, USA) in a 15 µl reaction mixture. Each reaction contained 5 µl buffer (ddH₂O), 7 µl 2 × Taq PCR MasterMix (0.1 U Taq Polymerase/µl, 500 µM dNTP each, 20 mM Tris-HCl, 100 mM KCl, 3 mM MgCl₂) (TIANGEN KT201, TIANGEN Biotech (Beijing) Co., Ltd), 1 µl each primer (forward and reverse) (10 µM), and 1 µl template DNA. The PCR cycling was performed at 94 °C, 3 min; 35 cycles of 94 °C, 30 s; 55 °C, 30 s; and 72 °C, 1 min; and 72 °C, 10 min for the final extension.

The PCR products (2 µl) were diluted with 8 µl of ultrapure water and scoured with 100% alcohol for 15 min. The diluted DNA was then dried and mixed with highly deionized formamide (Applied Biosystems, USA) and submitted to fragment analysis by capillary electrophoresis using an Applied Biosystems 3130 xl DNA Analyzer (Applied Biosystems, USA). The analysis of DNA fragment size and allele calling was performed using GeneScan and GeneMapper software (Applied Biosystems, USA), followed by manual allele binning. To

analyze the FAM-, TAMRA-, and HEX-labelled DNA fragments with the ABI3130 series systems (Applied Biosystems, USA), the matrix standard kit was used to generate the "multi-component matrix". The Data Collection Software (Applied Biosystems, USA) was used to record the data of the multi-component matrix. The instrument automatically analyzed three different coloured fluorescent dye-labelled samples in a single capillary, and reference lanes were further verified by polyacrylamide gel electrophoresis. In the case of non-amplification, the PCR was repeated to exclude technical failure, and a null allele was recorded if both PCRs failed.

Table S4

2.4 Statistical Data Analysis

All 33 SSR loci were evaluated for their adherence to the Hardy-Weinberg equilibrium and the presence of null alleles using the heterozygous deficiency method in Genepop v4.7 (Brookfield, 1996). The presence or absence of linkage disequilibrium between loci was confirmed by testing for genotypic linkage disequilibrium. The genetic parameters for each population were assessed by calculating various parameters, including allelic richness (A), mean observed heterozygosity (H_O), percentage of polymorphic loci (P), and unbiased expected heterozygosity (H_E). All genetic diversity estimates were calculated using GenAlEx 6.5.01 software (Peakall and Smouse, 2006). Further, the fixation index (F_{ST}) values were calculated using FSTAT version 2.9.3 and Arlequin 3.5 (Excoffier and Lischer, 2010) to investigate the species/population differentiation. The AMOVA was performed in Arlequin 3.5 to determine the proportion of genetic variance explained by the differences within and among species/populations. Furthermore, a model-based program, STRUCTURE 2.3.4 (Pritchard *et al.*, 2000), was used to infer the number of distinct genetic clusters and to assign individuals to a specific genetic cluster using default parameters. The program was executed with 10 independent runs for each value of K ranging from 1 to 10, each with 1,000,000 Markov chain Monte Carlo replications, following a 100,000 burn-in period. The admixture ancestry model and a correlated allele frequency model were used for all runs. The STRUCTURE HARVESTER online application was utilized to determine the estimated numbers of genetic components (K values) (Evanno *et al.*, 2005; Verkuil *et al.*, 2012). Clustering patterns and population structure inferences were determined throughout the K using the web tool CLUMPAK (Jakobsson and Rosenberg, 2007; Kopelman *et al.*, 2015). Both inter- and intra-specific genetic structures of the different populations were assessed using multivariate principal component analyses (PCA) via multivariate principal component analyses (PCA) through the `dudi.pca()` function of "ade4" R package (Jombart and Ahmed, 2011). The UPGMA clustering analysis of all populations was performed based on Nei's (1972) unbiased genetic distance using the PowerMarker software (Liu and Muse, 2005), and the resulting tree was visualized with TREEVIEW ver. 1.52. The Venn diagram was tabulated using the number of private alleles identified by genetic analysis for cultivated (inbred, landraces, and feral rice), wild, and weedy types. MIGRATE v. 4.4.4 (Beerli, 2008) was used to estimate effective population size N_e ($\theta/4\mu$) and asymmetric gene flow M (m/μ) between pairs of different *Oryza* groups found in Sri Lanka. The analyses were conducted using Bayesian inference under the structured Coalescent model. First, two shorter runs (10 short chains of 10,000 sampled, 500 records and three final chains of 100,000, 5,000 recorded) were performed. Then, a long final run (10 short chains of 10,000 sampled, 500 recorded, and three final chains of 500,000 sampled and 25,000 recorded) was performed, and results from this final run were reported.

3 Results

All 33 SSR loci exhibited polymorphism among the six *Oryza* types, resulting in the identification of a total of 621 alleles. The genotypic linkage disequilibrium between loci did not reveal significant values ($P > 0.05$), implying the absence of linkage disequilibrium among loci. Among the *Oryza* types, RM426 was the most variable locus with 61 alleles, whereas RM477 displayed only 5 alleles, with an average of 18.18 alleles per locus (data not shown). The highest genetic diversity was found in *O. rufipogon* ($H_E = 0.639$), while *O. nivara* recorded the lowest ($H_E = 0.539$) (Table S6). Moreover, a high level of private alleles was observed in the *O. rufipogon*. Overall, the genetic parameters revealed a considerable genetic diversity among all Sri Lankan *Oryza* lineage (the average H_E of all *Oryza* types was 0.587, was 1.287, Table S6), which was not evenly distributed across the different *Oryza* types (the table of allelic frequencies of each population is

available upon request).

Tables S6

Considerably low genetic differentiation was observed among cultivated *Oryza* types indicating a close genetic relationship (Table S5). Initially, inbred, feral, and landraces were collectively clustered as a cultivated rice group, given the observed genetic similarity among these subgroups. Further, wild rice types, such as *O. nivara* and *O. rufipogon*, were categorized as the wild rice group, based on their distinct genetic differentiation. Additionally, the weedy rice populations were classified as the weedy rice group.

Table S5

To assess the relatedness of weedy rice populations to other *Oryza* categories, the proportion of private alleles shared among the weedy rice group, the cultivated rice group, and the wild rice group was determined. A Venn diagram was employed to graphically illustrate the shared proportion of private alleles among the groups (Fig. S2). A total of 214 private alleles were identified among all rice groups. The cultivated rice group shared 22 alleles, which represented 10.3% of the private allele portion. The three rice groups collectively shared 139 private alleles, which represented 65% of the total. Conversely, the weedy rice populations shared only 1.9% of the private alleles. Specifically, 79% (169 alleles) of private alleles were shared between the weedy and cultivated groups, while 65.9% (141 alleles) were shared between the weedy and wild rice groups. We also observed that approximately 71.1% (152 alleles) of the private alleles were shared between the wild and cultivated *Oryza* groups, indicating the existence of multi-way allele transfer in the DWWC.

Fig. S2

Coalescent-based Bayesian estimates by MIGRATE indicated that the six *Oryza* groups have relatively large effective population sizes, with $N_e > 1,000$ (Table 1, $\mu = 10^{-4}$) (Jamieson and Allendorf, 2012). The long-term gene flows among the six *Oryza* types are shown in Table 1. The gene flow from weedy rice to *O. rufipogon* ($M = 7.299$) was much higher than from *O. rufipogon* to weedy rice ($M = 4.334$). The gene flow from weedy rice to *O. nivara* ($M = 4.2966$) was marginally lower in comparison to the flow from *O. nivara* to weedy rice ($M = 4.9654$). The gene flow from landraces to weedy rice was moderately high ($M = 5.9470$), while gene flow from inbred rice varieties to weedy rice has reported at a high rate ($M = 6.1028$). There were relatively high gene flows between feral rice and two cultivated rice types (inbred rice varieties and landraces). The highest gene flow was from inbred rice to feral rice ($M = 8.1597$). Gene flow between wild species was recorded in moderate values (Table 1).

Table 1

Principle Component Analysis (PCA) revealed three distinct clusters with a slight overlap of individuals from different groups (Fig. 2). The wild species, *O. rufipogon* formed a separate cluster (Cluster I) scattered along with the PCA plot, indicating higher levels of genetic diversity. *Oryza nivara* cluster (Cluster II) was well differentiated from other *Oryza* types, but a considerable portion of landraces, feral and weedy individuals overlapped. The large cluster (Cluster III) reflects the close genetic relationship among inbred cultivars, landraces, feral and weedy types. The feral populations were grouped with inbred rice varieties as they are currently undergoing the process of de-domestication (Ellstrand *et al.*, 2010). A considerable number of feral individuals were grouped with the wild *O. nivara*, potentially a high level of reverse introgression from the wild species. Many inbred rice cultivars overlapped with the feral and landraces. In contrast, weedy rice displayed a diffused distribution in the PCA scatter plot, reflecting a close genetic relationship with inbred cultivars, feral rice, and *O. nivara*. Most of the weedy rice showed a close evolutionary relationship with the inbred and landraces and few displayed close genetic similarities to *O. nivara*.

Fig. 2. Scatter plot of the first and second principal components (PC) based on the variation of 33 SSR loci for 1340 individuals of 20 weedy rice populations, five *O. rufipogon* populations, six *O. nivara* populations, 42 inbred rice varieties, seven feral rice populations and 31 landraces from Sri Lanka showing the considerable overlap of genotypes in different *Oryza* types at 33 SSR loci.

The STRUCTURE analysis of DWWC demonstrated a distinct K peak at $K=2$ levels (Fig. S1), indicating that wild *Oryza* was differentiated from other *Oryza* groups. Within the analysis, two populations of weedy rice (W4 and W6) and three populations of feral rice (F1, F2, and F5) were classified as out-groups (Fig. 3) in relation to their respective *Oryza* groups, leading to the formation of two separate groups for both weedy and feral rice. This observation can likely be attributed to the adaptation of *O. nivara* to the rice ecosystems, a process driven by persistent habitat disturbances, continued selective pressures, or hybridization events between cultivated (inbred and landraces) plants and their reproductively compatible wild counterparts (Cao *et al.*, 2006). Consequently, $K=2$ was determined to be the most biologically realistic population number. Further, K peaks at $K=9$ and $K=5$ (Fig. S1) were considered for further examination of population models. The $K=9$ model is highly intricate, suggesting that the physical classification of *Oryza* groups presents a challenge due to the genetic relationships among populations and individuals based on prevailing patterns of genetic admixture. However, the $K=5$ model is generally consistent with the Principal Component Analysis (PCA) results (Figs. 3 and 4).

The two wild *Oryza* species, *O. nivara* and *O. rufipogon*, differentiated from others. Moreover, these two wild species exhibited genetic differences from one another, i.e. some populations showing admixture, while most individuals can be reliably assigned to a specific population. The *O. rufipogon* group comprised a most distinct group with a more heterogeneous genetic background. Two weedy and three feral populations were highly admixed with the wild *O. nivara*. Widely cultivated types, inbred rice varieties, and landraces shared a common ancestry and were evident as largely close groups to weedy rice with the admixture nature. However, some individuals of landraces were admixed with the wild *Oryza* and weedy rice. The evolution of the weedy type reflects the complex process of genetic incorporation from the crop (inbred or landraces), non-cultivated wild *Oryza*. This illustrates the multi-way genetic transfer to the evolution of weedy types. The STRUCTURE results suggest that there is a complex integration of multi-way gene flow among all members of the DWWC in the rice ecosystem in Sri Lanka. However, when examining K values ranging from 5 to 10, it becomes clear that there is an admixed genetic background for individuals in some populations/cultivars. While most individuals could be assigned to a single population or cultivar, the occurrence of further groupings is uncommon (Fig. 3).

Fig. 3. STRUCTURE graph showing genotype clustering of 20 weedy rice populations, five *O. rufipogon* populations, six *O. nivara* populations, 42 inbred rice varieties, seven feral rice population and 31 landraces by model-based population assignment at K from 2 to 10. Each vertical bar represents an individual, with its assignment probability to genetic clusters represented by different colors. Codes for the weedy rice, *O. rufipogon*, *O. nivara* and feral rice populations are presented in Table S1.

Fig. S1

The UPGMA tree shows similar results, with all *Oryza* types genetically structured into two well-separated major groups (*O. rufipogon* and all other *Oryza* types) and further divided into respective populations and cultivars (Fig. 4). The *O. rufipogon* group forms a separate cluster (Cluster I). Besides, the large cluster was further divided into two sub-clusters (Cluster II and Cluster III). Moreover, all *O. nivara*, three feral rice (F1, F2, and F5) and two weedy (W4 and W6) rice populations, and a few landraces formed a distinct cluster (Cluster III), as shown in the PCA. The large cluster (Cluster III) consisted of inbred varieties, landraces, unmanaged abandoned feral populations, and weedy rice. Most weedy rice populations were grouped with landraces, however, W2 and W3 populations displayed closer relationships with landraces. Furthermore, W1 was distinct from the weedy rice cluster, while the rest of the populations were subdivided into two clusters. Feral rice populations were grouped into a large cluster, but they were also clustered with the inbred rice group. Inbred rice varieties and landraces showed complex clustering patterns due to the relatively high genetic distances among their respective cultivars.

Fig. 4. The UPGMA cladogram is based on Nei (1972) genetic distance. Dendrogram (UPGMA) was constructed based on polymorphisms of 33 SSR loci in six *Oryza* types (20 weedy rice populations, five *O. rufipogon* populations, six *O. nivara* populations, 42 inbred rice varieties, seven feral rice populations, and 31 landraces), using Nei's unbiased genetic distance (Nei, 1972). The bar represents genetic distance, with

the same color sharing the same source of collection. Population or varieties/cultivars codes for the six *Oryza* types of populations are presented in Table S1.

The AMOVA analysis revealed a significant portion of the total genetic diversity present within populations (56%) and among populations (28%), while relatively low (16%) genetic differentiation was observed among *Oryza* types (Table S7; $P < 0.01$). In specific *Oryza* types, the total variation was partitioned into among populations/cultivars and within populations/cultivars. A notably larger variation was observed within populations compared to among populations (Table S7).

Table S7

4 Discussion

The Asian rice gene pool broadly comprises cultivated rice (*Oryza sativa* L.), its wild progenitors (*O. rufipogon* and *O. nivara*), and weedy rice (*Oryza* spp.) with close genetic affinity. In the present study, while the six types of Sri Lankan *Oryza* were grouped together, we observed two major clusters: wild and weedy with cultivated rice. *Oryza rufipogon* showed unique differentiation from *O. nivara* and other *Oryza* types due to its species-specific attributes and relatively high genetic distance (Fig. 2 and 4) (Sandamal *et al.*, 2021; 2022). We find that weedy rice from Sri Lanka had a close genetic relationship to cultivated types and *O. nivara*. Given that both wild rice species are commonly found near paddy fields, recurrent interactions between these DWWC groups in Sri Lanka are possible, suggesting an evolutionarily dynamic rice system with the potential for gene flow and introgression. Below we discuss these findings within the context of evolutionary mechanisms among the DWWC rice groups, particularly exploring how these groups have been influenced by multiple pathways of gene flow.

4.1 Genetic diversity and population structure of DWWC

Genetic diversity (H_E) in *O. rufipogon* and *O. nivara* populations revealed by SSR markers in the present study was relatively higher than those of previously reported, signifying a comparable degree of genetic diversity within the broader South Asian region (Banaticla-Hilario *et al.*, 2013; Kuroda *et al.*, 2007; Pusadee *et al.*, 2013; Samal *et al.*, 2018; Sandamal *et al.*, 2018; Zhou *et al.*, 2003). Among the similar studies, the Chinese *O. rufipogon* population showed an H_E value of 0.413 (Zhou *et al.*, 2013), while in Vientiane Plain of Laos, the H_E values for *O. rufipogon* ranged from 0.37 to 0.77 and for *O. nivara* ranged from 0 to 0.64 (Kuroda *et al.*, 2007). In the Asia Pacific region, reported H_E values were 0.39 for Nepal *O. nivara*, 0.67 for Southeast Asian *O. nivara*, and 0.70 for Southeast Asian *O. rufipogon* (Banaticla-Hilario *et al.*, 2013). In the Indian peninsula, H_E values of 0.63 for *O. rufipogon* and 0.64 for *O. nivara* were reported (Singh *et al.*, 2018). Previous studies revealed that *O. rufipogon* and *O. nivara* have significant but low divergence, implying ongoing speciation processes with potential gene flow between them (Liu *et al.*, 2015; Pusadee *et al.*, 2016; Zheng and Ge, 2010; Sandamal *et al.*, 2018). In the present study, the presence of admixtures in the studied wild populations (Fig. 3), and high genetic diversity along with the PC1 in PCA results supported this hypothesis. Furthermore, *O. rufipogon* showed high within-population variation (63%) and lower among-population variation (37%), while *O. nivara* exhibits 48% within-population and 47% among-population variation. In Sri Lanka, *O. rufipogon* showed lower population differentiation compared to *O. nivara* (Table S7), potentially due to limited gene flow among *O. rufipogon* populations in the wet and intermediate zones.

The present study revealed a moderately high overall genetic diversity ($H_E = 0.566$) in weedy rice populations collected across Sri Lanka, with higher within-population genetic variance (79%) than among-population variance (21%). This suggests the consistently high genetic diversity within weedy populations that can be attributed to the introgression from other DWWC types to weedy lineages. STRUCTURE analysis (Fig. 3) indicated a high admixture of some weedy rice individuals with wild *Oryza*, suggesting bidirectional gene flow between weedy and wild *Oryzas*. Unlike previous studies (Heet *et al.*, 2014), our research included a large number of individuals (1340) from various locations in Sri Lanka, potentially explaining the detection of the grouping of weedy populations with *O. nivara*. The high admixture may result from wild rice adapting to cultivated rice habitats due to ongoing selection and habitat disturbances (De Wet and Harlan, 1975;

Vaughan *et al.* , 2001). Similar studies carried out worldwide have reported low to high genetic diversity in weedy rice populations (Han *et al.* , 2020; Cao *et al.* , 2006; Yu *et al.* , 2005, Neik *et al.* , 2019; Gealy *et al.* , 2003; Song *et al.* , 2014; Prathepha, 2011), suggesting that genetic diversity in weedy rice is variable and influenced by regional and local factors. Our study makes a significant contribution to genetic studies by reporting the genetic structure of South Asian weedy rice populations at a larger scale, highlighting the high genetic diversity and admixture in the DWWC in Sri Lanka.

In Sri Lanka, rice is grown across diverse physical environments with varying altitudes, soils, and hydrological regimes. The rice-growing altitudes range from 0 to 900 m above sea level, with temperatures ranging from 30 °C (± 5 °C) at sea level to about 15 °C at higher elevations (Dhanapala, 2007). The country has a long history of cultivating rice, with numerous endemic cultivars, wild species, and landraces, predominantly long-duration cultivars and photoperiod sensitive (Weerakoon *et al.* , 2011). The development of inbred varieties occurred through artificial hybridization at the Rice Research and Development Institute (RRDI) of Sri Lanka, some of which utilize superior individuals and exotic germplasm imported from Indonesia utilized for developing high-yielding cultivars (Dhanapala, 2020). This led to moderately high genetic diversity in Sri Lankan inbred rice compared to other rice types examined. Landraces and feral populations also displayed high levels of genetic diversity. The genetic differentiation among cultivated *Oryza* types (inbred and landraces), inbred and feral rice, and landraces and feral rice was considerably low (Table S5), indicating a close genetic relationship among rice types. Our findings revealed the dynamic nature of domesticated, wild, and weedy components in the Sri Lankan rice ecosystem, with high admixture and recurrent gene flow observed among different *Oryza* types. The presence of shared private alleles among all DWWC rice groups confirms the high admixture nature of these populations (Fig. S2). This highlights the significant contribution of introgression and multi-way gene flow in shaping the genetic diversity of the DWWC in Sri Lanka.

4.2 Multi-way gene flow among *Oryza* types

Gene flow analysis revealed significant gene flow among all six *Oryza* types (Table 1). Among the rice groups, *O. nivara* showed higher sensitivity to introgression, and asymmetric gene flows were observed between wild and weedy rice groups, particularly between *O. rufipogon* and weedy rice (Table 1). The open panicle structure of *O. rufipogon* makes it more receptive to foreign pollen, resulting in higher gene flow from other *Oryza* groups. The *O. nivara* populations studied were in natural habitats near commercial rice fields. Long-term gene flow from DWWC components to *O. nivara* populations may promote evolution and local adaptation, transitioning their reproductive system from high selfing to slight outcrossing, increasing its sensitivity to introgression.

In Sri Lanka, rice cultivation is carried out in two seasons, *Maha* and *Yala* , coinciding with the inter-monsoon and monsoon rains. Approximately 450,000 hectares of paddy are cultivated during the *Yala* season (March-September), while 850,000 hectares are cultivated in the *Maha* season (October-February) (Punyawardana, 2008). Most inbred rice varieties and landraces have a maturity duration of three to four and a half months, leading to overlap in flowering periods with wild rice, particularly during the *Maha* season. Concurrently, *O. rufipogon* being perennial, can be observed throughout the year, while the annual *O. nivara* seeds germinate at the beginning of the *Maha* season. Both species overlap in flowering time with inbred rice varieties. This increases the possibilities of pollen-mediated gene flow facilitated by monsoon wind currents and insects, contributing to admixtures between wild and weedy rice groups. Further, volunteer weeds may evolve into weedy rice over generations and interact with the DWWC components over time. Long-term introduction and persistence within the DWWC has resulted in complex multi-directional gene flow among its components, shaping the complex rice ecosystem in Sri Lanka (Vaughan *et al.* , 2005).

Our findings support higher gene flow from weedy rice to *O. rufipogon* ($M=7.299$) than vice versa ($M=4.334$), and significant gene flow was observed from cultivated rice to feral rice ($M=8.1597$) (Table 1), implying an early stage of de-domestication (Ellstrand *et al.*, 2010). Pollen-mediated gene exchange occurs between crops and weeds or wild relatives and between weedy and wild relatives. Reverse gene flow from weed to the crop has been well documented (Langevin *et al.*, 1990; Majumder *et al.*, 1997; Song *et al.*, 2002). The presence of wild *Oryza* and greater diversity of crop varieties and landraces in the region (Song *et al.*, 2014) contribute

to the evolutionary dynamics of South Asian weeds. Domesticated rice can mate with wild relatives, leading to gene flow that can influence the genetic structure of wild rice (Wang *et al.*, 2017). Similarly, gene flow between cultivated carrots (*Daucus carota* ssp. *sativus*) and wild relatives (*D. carota* ssp. *carota*) has caused adjacent wild populations to become genetically similar to the cultivars, contributing to the development of more aggressive weeds in carrot fields (Magnussen and Hauser, 2007).

Weedy rice in the U.S.A. is genetically distinct from inbred rice varieties due to de-domestication processes during its evolution and accidental introductions (Reagon *et al.*, 2010; Li *et al.*, 2017). It predominantly descends from wild rice populations in tropical Asia (Londo and Schaal, 2007; Reagon *et al.*, 2010), partly incurred by an extended flowering period and weak post-zygotic reproductive barrier, which promotes gene flow between cultivated and wild rice (Craig *et al.*, 2014). The gene flow patterns elucidated in the present study concerning weedy, wild, and cultivated rice corroborate the scenario (Table 1). Within the context of weedy rice origins and evolution in Sri Lanka, multiple factors exert influence, such as the feralization of cultivated rice, hybridization between landraces and/or inbred rice varieties, the adaptation of wild *O. nivara*, and the hybridization between domesticated and wild rice (Rieseberg *et al.*, 1993; Arnold, 1997; Rosenthal *et al.*, 2008; Ellstrand, 2009; Ellstrand *et al.*, 2010). Wedger *et al.* (2019) discovered that 23% of the studied weedy rice plants in Thailand exhibited introgressed alleles at one or more loci, derived from three domestication gene sequences of DWWC. Similarly, a preceding study done in Thailand (Wedger *et al.*, 2019; Pusadee *et al.*, 2013; Wongtamee *et al.*, 2017) reported that 17.6% of 29 weedy rice plants showed evidence of introgression from wild or cultivated rice groups, indicating gene flow as a significant driving force for the evolutionary dynamics of the *Oryza* complex. Based on pairwise migration data, He *et al.* (2014) reported a significantly high level of gene flow within distinct weedy rice populations in Sri Lanka. Similarly, gene flow from Malaysian weedy rice to *O. rufipogon* substantiates the hypothesis positing potential introgression with other indigenous *Oryzas* within the context of wild rice populations (San Sudo *et al.*, 2021). These studies indicate that weedy rice in the Asian region has undergone a complex evolutionary process involving gene flow with wild and cultivated rice, contributing to its genetic distinctiveness from cultivated rice varieties.

Acknowledgments

We thank the State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences (Beijing, China), for providing laboratory facilities. This work was supported by the National Natural Science Foundation of China (NSFC) (32130008) and Ministry of Science and Technology (2021YFD1200101-02) awarded to G.S., National Research Council, Sri Lanka (NRC 15–108) awarded to D.R., National Science Foundation, Sri Lanka (OSTP/2016/53) awarded to A.T. and (OSTP/2016/54) awarded to S.S., and the MUM-ASEAN Sustainable Development Research Grant (ASEAN-2019-01-SCI) awarded to B.K. We thank Zhe Cai, Mu-Fan Geng, Qing-Lin Meng, Gayani Galhena and M. D. Nirmani for their technical assistances and helpful comments on the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

D. R., S. G., B-K. S., N-C. E., B. M., and A. T. conceived the ideas. A. T., S. S., and D.R. collected the data. A. T., S. S., A. M., and Y. Z. analyzed the data. A. T., S. S., S.G., A.M., Y.Z., B. M., N-C. E., B-K. S., and D.R. wrote the paper.

Data Accessibility

Data available from the Dryad Digital Repository.

Data available in Step 2: File Upload under the title of Supplementary Material section.

References

Arnold, M.L., 1997. *Natural hybridization and evolution*. Oxford University Press on Demand.

- Banaticla-Hilario, M.C.N., van den Berg, R.G., Hamilton, N.R.S. and McNally, K.L., 2013. Local differentiation amidst extensive allele sharing in *Oryza nivara* and *O. rufipogon* . *Ecology and Evolution* , 3(9), pp.3047-3062.
- Beerli, P., 2008. Migrate version 3.0: a maximum likelihood and Bayesian estimator of gene flow using the coalescent. Available at: <http://popgen.scs.edu/migrate.html>.
- Brookfield, J., 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology* , 5, pp.453-455.
- Cao, Q., Lu, B.R., Xia, H.U.I., Rong, J., Sala, F., Spada, A. and Grassi, F., 2006. Genetic diversity and origin of weedy rice (*Oryza sativa* f. *spontanea*) populations found in north-eastern China revealed by simple sequence repeat (SSR) markers. *Annals of Botany* , 98(6), pp.1241-1252.
- Craig, S.M., Reagon, M., Resnick, L.E. and Caicedo, A.L., 2014. Allele distributions at hybrid incompatibility loci facilitate the potential for gene flow between cultivated and weedy rice in the US. *PLoS One* , 9(1), p.e86647.
- De Wet, J.M. and Harlan, J.R., 1975. Weeds and domesticates: evolution in the man-made habitat. *Economic Botany* , 29(2), pp.99-108.
- Dhanapala, M.P., 2007. Bridging the rice yield gap in Sri Lanka.
- Dhanapala, M.P., 2020. Milestones in the History of Rice Improvement in Sri Lanka. *Agricultural Research for Sustainable Food Systems in Sri Lanka : A Historical Perspective* , 1, pp.101-115.
- Ellstrand, N.C., 2003. *Dangerous Liaisons? When Cultivated Plants Mate with Their Wild Relatives* . Johns Hopkins University Press, Baltimore, MD.
- Ellstrand, N.C., 2009. Evolution of invasiveness in plants following hybridization. *Biological Invasions* , 11(5), pp.1089-1091.
- Ellstrand, N.C., Heredia, S.M., Leak-Garcia, J.A., Heraty, J.M., Burger, J.C., Yao, L., Nohzadeh-Malakshah, S. and Ridley, C.E., 2010. Crops gone wild: evolution of weeds and invasives from domesticated ancestors. *Evolutionary Applications* , 3 (5-6), pp.494-504.
- Evanno, G., Regnaut, S. and Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* , 14(8), pp.2611-2620.
- Gealy, D.R., Mitten, D.H. and Rutger, J.N., 2003. Gene flow between red rice (*Oryza sativa*) and herbicide-resistant rice (*O. sativa*): implications for weed management. *Weed Technology* , 17(3), pp.627-645.
- Han, B., Ma, X., Cui, D., Wang, Y., Geng, L., Cao, G., Zhang, H., Koh, H.J. and Han, L., 2020. Analysis of evolutionary relationships provides new clues to the origins of weedy rice. *Ecology and Evolution* , 10(2), pp.891-900.
- He, Z., Jiang, X., Ratnasekera, D., Grassi, F., Perera, U. and Lu, B.R., 2014. Seed-mediated gene flow promotes genetic diversity of weedy rice within populations: implications for weed management. *PLoS One* , 9(12), p.e112778.
- Jakobsson, M. and Rosenberg, N.A., 2007. Clumpp: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* , 23(14), pp.1801-1806.
- Jombart, T. and Ahmed, I., 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* , 27(21), pp.3070-3071.
- Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A. and Mayrose, I., 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* , 15(5), pp.1179-1191.

- Kuroda, Y., Sato, Y.I., Bounphanousay, C., Kono, Y. and Tanaka, K., 2007. Genetic structure of three *Oryza* AA genome species (*O. rufipogon*, *O. nivara* and *O. sativa*) as assessed by SSR analysis on the Vientiane Plain of Laos. *Conservation Genetics*, 8(1), pp.149-158.
- Langevin, S.A., Clay, K. and Grace, J.B., 1990. The incidence and effects of hybridization between cultivated rice and its related weed red rice (*Oryza sativa* L.). *Evolution*, 44(4), pp.1000-1008.
- Leak-Garcia, J.A., Holt, J.S., Kim, S.C., Mu, L., Mejias, J.A. and Ellstrand, N.C., 2012. More than multiple introductions: Multiple taxa contribute to the genesis of the invasive California's wild artichoke thistle. *Journal of Systematics and Evolution* 51, pp.295-307.
- Linder, C.R., Taha, I., Seiler, G.J., Snow, A.A. and Rieseberg, L.H., 1998. Long-term introgression of crop genes into wild sunflower populations. *Theoretical and Applied Genetics*, 96, pp.339-347.
- Li, L.F., Li, Y.L., Jia, Y., Caicedo, A.L. and Olsen, K.M., 2017. Signatures of adaptation in the weedy rice genome. *Nature Genetics*, 49(5), pp.811-814.
- Liu, R., Zheng, X.M., Zhou, L., Zhou, H.F. and Ge, S., 2015. Population genetic structure of *Oryza rufipogon* and *Oryza nivara*: implications for the origin of *O. nivara*. *Molecular Ecology*, 24(20), pp.5211-5228.
- Magnussen, L.S. and Hauser, T.P., 2007. Hybrids between cultivated and wild carrots in natural populations in Denmark. *Heredity*, 99(2), pp.185-192.
- Majumder, N.D., Ram, T. and Sharma, A.C., 1997. Cytological and morphological variation in hybrid swarms and introgressed population of interspecific hybrids (*Oryza rufipogon* Griff.x*Oryza sativa*L.) and its impact on evolution of intermediate types. *Euphytica*, 94(3), pp.295-302.
- Molina, J., Sikora, M., Garud, N., Flowers, J.M., Rubinstein, S., Reynolds, A., Huang, P., Jackson, S., Schaal, B.A., Bustamante, C.D. and Boyko, A.R., 2011. Molecular evidence for a single evolutionary origin of domesticated rice. *Proceedings of the National Academy of Sciences*, 108(20), pp.8351-8356.
- Nei, M., 1972. Genetic distance between populations. *The American Naturalist*, 106(949), pp.283-292.
- Neik, T.X., Chai, J.Y., Tan, S.Y., San Sudo, M.P., Cui, Y., Jayaraj, J., Teo, S.S., Olsen, K.M. and Song, B.K., 2019. When west meets east: the origins and spread of weedy rice between continental and island Southeast Asia. *G3: Genes, Genomes, Genetics*, 9(9), pp.2941-2950.
- Peakall, R.O.D. and Smouse, P.E., 2006. GENALEX 6: Genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), pp.288-295.
- Pritchard, J.K., Stephens, M. and Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155(2), pp.945-959.
- Prathepha, P., 2011. Microsatellite analysis of weedy rice (*Oryza sativa* f. *spontanea*) from Thailand and Lao PDR. *Australian Journal of Crop Science*, 5(1), p.49.
- Punyawardhana, B.V.R., 2008. Rainfall and agroecological zones of Sri Lanka.
- Pusadee, T., Jamjod, S., Rerkasem, B. and Schaal, B.A., 2016. Life-history traits and geographical divergence in wild rice (*Oryza rufipogon*) gene pool in Indochina Peninsula region. *Annals of Applied Biology*, 168(1), pp.52-65.
- Pusadee, T., Schaal, B.A., Rerkasem, B. and Jamjod, S., 2013. Population structure of the primary gene pool of *Oryza sativa* in Thailand. *Genetic Resources and Crop Evolution*, 60(1), pp.335-353.
- Qiu, J., Jia, L., Wu, D., Weng, X., Chen, L., Sun, J., Chen, M., Mao, L., Jiang, B., Ye, C. and Turra, G.M., 2020. Diverse genetic mechanisms underlie worldwide convergent rice feralization. *Genome Biology*, 21, pp.1-11.

- Ratnasekera, D., 2015. Weedy rice: A threat to rice production in Sri Lanka. *Journal of the University of Ruhuna* , 3(1), pp.2-13.
- Reagon, M., Thurber, C.S., Gross, B.L., Olsen, K.M., Jia, Y. and Caicedo, A.L., 2010. Genomic patterns of nucleotide diversity in divergent populations of US weedy rice. *BMC Evolutionary Biology* , 10(1), pp.1-16.
- Rieseberg, L.H., Wendel, J.F. and Harrison, R.G., 1993. Hybrid zones and the evolutionary process. *Oxford university press, New York* , pp.70-109.
- Rosenthal, D.M., Ramakrishnan, A.P. and Cruzan, M.B., 2008. Evidence for multiple sources of invasion and intraspecific hybridization in *Brachypodium sylvaticum* (Hudson) Beauv. in North America. *Molecular Ecology* , 17(21), pp.4657-4669.
- Saghai-Marooif, M.A., Soliman, K.M., Jorgensen, R.A. and Allard, R.W.L., 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences* , 81(24), pp.8014-8018.
- Samal, R., Roy, P.S., Sahoo, A., Kar, M.K., Patra, B.C., Marndi, B.C. and Gundimeda, J.N., R. 2018. Morphological and molecular dissection of wild rice from eastern India suggests distinct speciation between *O. rufipogon* and *O. nivara* populations. *Scientific Reports* , 8(1), p.2773.
- Sandamal, P.S., Tennakoon, A., Amarasekera, D.A.B.N., Marambe, B. and Rathnasekera, D., 2021. Functional trait diversity of wild rice species in Sri Lanka: Implications for field identification and application. *Journal of the National Science Foundation of Sri Lanka* , 49(3), 369–382.
- Sandamal, S., Tennakoon, A., Meng, Q.L., Marambe, B., Ratnasekera, D., Melo, A. and Ge, S., 2018. Population genetics and evolutionary history of the wild rice species *Oryza rufipogon* and *O. nivara* in Sri Lanka. *Ecology and Evolution* , 8(23), pp.12056-12065.
- Sandamal, S., Tennakoon, A., Wijerathna, P., Song, G., Amarasekera, D.A.B.N., Marambe, B., Elwany, S.M., Sorour, S., El Sabagh, A., Hassan, M.M. and Ratnasekera, D., 2022. Phenotypic Characterization of *Oryza nivara* (Sharma et Shastry) collected from different ecological niches of Sri Lanka. *Phyton* , 91(6), pp.1269.
- San Sudo, M.P., Yesudasan, R., Neik, T.X., Masilamany, D., Jayaraj, J., Teo, S.S., Rahman, S. and Song, B.K., 2021. The details are in the genome-wide SNPs: Fine scale evolution of the Malaysian weedy rice. *Plant Science* , 310, p.110985.
- Singh, B., Singh, N., Mishra, S., Tripathi, K., Singh, B.P., Rai, V., Singh, A.K. and 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, p.123.
- Song, B.K., Chuah, T.S., Tam, S.M. and Olsen, K.M., 2014. Malaysian weedy rice shows its true stripes: wild *Oryza* and elite rice cultivars shape agricultural weed evolution in Southeast Asia. *Molecular Ecology* , 23(20), pp.5003-5017.
- Song, Z., Lu, B., Zhu, Y. and Chen, J., 2002. Pollen competition between cultivated and wild rice species (*Oryza sativa* and *O. rufipogon*). *New Phytologist* , 153(2), pp.289-296.
- Vaughan, D.A., Sanchez, P.L., Ushiki, J., Kaga, A., Tomooka, N. and Gressel, J., 2005. Asian rice and weedy rice-evolutionary perspectives. *Crop Fertility and Volunteerism* , pp.257-277.
- Vaughan, L.K., Ottis, B.V., Prazak-Havey, A.M., Bormans, C.A., Sneller, C., Chandler, J.M. and Park, W.D., 2001. Is all red rice found in commercial rice really *Oryza sativa* ?. *Weed Science* , 49(4), pp.468-476.
- Veltman, M.A., Flowers, J.M., van Andel, T.R. and Schranz, M.E., 2019. Origins and geographic diversification of African rice (*Oryza glaberrima*). *PLoS One* , 14(3), p.e0203508.

Wang, H., Vieira, F.G., Crawford, J.E., Chu, C. and Nielsen, R., 2017. Asian wild rice is a hybrid swarm with extensive gene flow and feralization from domesticated rice. *Genome Research* , 27(6), pp.1029-1038.

Weerakoon, W.M.W., Mutunayake, M.M.P., Bandara, C., Rao, A.N., Bhandari, D.C. and Ladha, J.K., 2011. Direct-seeded rice culture in Sri Lanka: lessons from farmers. *Field Crops Research* , 121(1), pp.53-63.

Wongtamee, A., Maneechote, C., Pusadee, T., Rerkasem, B. and Jamjod, S., 2017. The dynamics of spatial and temporal population genetic structure of weedy rice (*Oryza sativa* f. *spontanea* Baker). *Genetic Resources and Crop Evolution* , 64(1), pp.23-39.

Zheng, X.M. and Ge, S., 2010. Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Molecular Ecology* , 19(12), pp.2439-2454.

Zhou, H.F., Xie, Z.W. and Ge, S., 2003. Microsatellite analysis of genetic diversity and population genetic structure of a wild rice (*Oryza rufipogon* Griff.) in China. *Theoretical and Applied Genetics* , 107(2), pp.332-339.

Zhou, W., Wang, Z., Davy, A.J. and Liu, G., 2013. Geographic variation and local adaptation in *Oryza rufipogon* across its climatic range in China. *Journal of Ecology* , 101(6), pp.1498-1508.

Hosted file

Figures.docx available at <https://authorea.com/users/668908/articles/669629-do-the-six-types-of-sri-lanka-s-oryza-comprise-a-domesticated-weed-wild-complex>

Hosted file

Table.docx available at <https://authorea.com/users/668908/articles/669629-do-the-six-types-of-sri-lanka-s-oryza-comprise-a-domesticated-weed-wild-complex>