Regional projections and sampling strategies for the global germplasm resources collection of *Oryza rufipogon* Griff.

Kai Wang, Jinhua Long, Jizhou Lei, Yuntao Liang, Disna Ratnasekera, Qian Qian, Xiaoming Zheng

PII: S2949-9526(25)00037-8

DOI: https://doi.org/10.1016/j.ncrops.2025.100101

Reference: NCROPS100101

To appear in: *New Crops*

Received date: 9 February 2025
Revised date: 13 October 2025
Accepted date: 16 October 2025

Please cite this article as: Kai Wang, Jinhua Long, Jizhou Lei, Yuntao Liang, Disna Ratnasekera, Qian Qian and Xiaoming Zheng, Regional projections and sampling strategies for the global germplasm resources collection of *Oryza rufipogon* Griff., *New Crops,* (2025) doi:https://doi.org/10.1016/j.ncrops.2025.100101

Regional projections and sampling strategies for the global germplasm resources collection of *Oryza rufipogon* Griff.

Kai Wang[a,b,c], Jinhua Long[a,b], Jizhou Lei[b], Yuntao Liang[e], Disna Ratnasekera[a,b,f], Qian Qian[a,b,c,d], Xiaoming Zheng[a,b,g,*]

[a] *State Key Laboratory of Crop Gene Resources and Breeding/Key Laboratory of Grain Crop Genetic Resources Evaluation and Utilization, Ministry of Agriculture and Rural Affairs, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China*

[b] *National Nanfan Research Institute (Sanya), Chinese Academy of Agricultural Sciences, Sanya 572000, China*

[c] *China National Rice Research Institute, Chinese Academy of Agricultural Sciences, Hangzhou 310006, China*

[d] *Yazhouwan National Laboratory, Sanya 572024, China*

[e] *Rice Research Institute, Guangxi Academy of Agricultural Sciences, Nanning 530007, China*

[f] *Department of Agricultural Biology, Faculty of Agriculture, University of Ruhuna, Matara 81000, Sri Lanka*

[g] *International Rice Research Institute, Metro Manila 1301, Philippines*

Corresponding author: Xiaoming Zheng (zhengxiaoming@caas.cn)

Abstract

Oryza rufipogon Griff., the wild progenitor of cultivated rice, contains essential genetic resources for rice improvement and global food security. Preserving the integrity of its germplasm has therefore become a critical conservation priority. In this study, we applied GIS-based spatial analysis and the MaxEnt model, integrating global occurrence data of *O. rufipogon* with 20 environmental variables to predict suitable habitats during the Last Interglacial (LIG, 130 to 115 ka before present [BP]), the Last Glacial Maximum (LGM, 26–19 ka BP), and the present. Model performance was evaluated using the area under the receiver operating characteristic curve (AUC), and spatial clustering was conducted based on the contributions of environmental factors. The results identified three current core ecological zones: the tropical monsoon region of Southeast and East Asia; the tropical grassland–monsoon region of South Asia; and the tropical rainforest region spanning northern Australia, the Philippines, and Indonesia. During the LGM, colder and drier climates led to a more fragmented distribution, whereas the LIG showed a clearer expansion into broader tropical and subtropical zones. Priority regions for *O. rufipogon* germplasm collection were identified across Southeast, East, and South Asia, as well as northern Australia. This study proposes a targeted sampling strategy that incorporates elevation, soil properties, and water availability, providing a solid scientific basis for global *O. rufipogon* surveys and the conservation of high-quality germplasm resources.

Keywords

Oryza rufipogon Griff.; Germplasm resource collection; Genetic diversity; Environmental factor; MaxEnt model

Introduction

Oryza sativa L. is one of the world's most important staple crops (Cheng and Hu, 2008), with *Oryza rufipogon* Griff. (*O. rufipogon*), its wild ancestor, mainly found in subtropical and tropical regions. Over extensive natural selection and adaptation, *O. rufipogon* has developed notable genetic diversity, making it an invaluable resource for rice improvement (Yang and Huang, 2013; He et al., 2024; Zheng et al., 2024). However, recent shifts in climate, natural environment, cropping structure, and land management have seriously affected its natural habitats, leading to its designation as a nationally protected species in China (Fu, 1992). Surveys conducted between 2002 and 2007 in Guangxi revealed that 80% of *O. rufipogon* habitats had disappeared (Huang et al., 2009). Similarly, a 2002–2013 survey across 18 cities in Hainan Province found that only 18 of the 118 *O. rufipogon* distribution sites recorded in the 1980s remained, with 84.7% of sites lost (Yun et al., 2015). In Guangdong Province, a 2005–2016 survey reported a 90.49% reduction in distribution sites (Fan et al., 2017). Thus, the conservation and collection of *O. rufipogon* germplasm resources have become increasingly urgent.

The survey of *O. rufipogon* germplasm resources has evolved from sporadic, small-scale surveys led by individual researchers (Lin and Yuan, 1980; Guarino et al., 1995) to broader, large-scale surveys using documentary, interview, and field survey methodologies (Gao and Yang, 2022). Collection methods have shifted from experience-based random sampling to more systematic approaches, emphasizing clear population identification and strategies that capture the genetic diversity of each population (Fig. 1). Population-specific collection strategies now incorporate molecular markers, setting diversity thresholds and defining sampling standards for spacing and sample sizes (Wang et al., 2010; Gao and Yang, 2022). Technologies like GPS now allow precise location of sample sites, while visual documentation records essential habitat and phenotypic context. These advancements have led to the establishment of numerous *O. rufipogon in situ* conservation sites (Zheng et al., 2019) and the development of *ex situ* conservation programs, preserving a wealth of genetic resources.

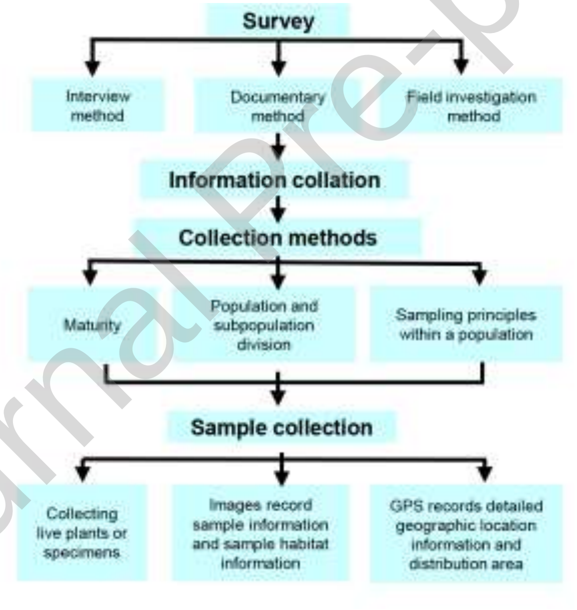


Fig. 1. Roadmap of traditional survey and collection techniques for *O. rufipogon.*

Despite these advancements, traditional collection methods remain time-consuming, labor-intensive, and geographically restricted. Furthermore, over 90% of the *O. rufipogon* germplasm resources conserved in China originate from domestic collections, underscoring the lack of global germplasm representation (Zhang and Yang, 2003; Zheng et al., 2020). Expanding the collection of foreign *O. rufipogon* resources is therefore essential to enrich China’s genetic resource base, yet overseas surveys present unique challenges. Comprehensive field investigations require substantial investment and a high level of expertise, often hindering the efficiency of collection efforts abroad. As a result, the development of accurate and efficient methodologies for global *O. rufipogon* collection has become an urgent priority.

To address these challenges, more advanced survey and collection methodologies are crucial. The integration of GIS technology into germplasm surveys enables researchers to obtain comprehensive landscape information for target populations. As climate change alters species distributions, species distribution models (SDMs), grounded in ecological niche theory—especially when combined with GIS—provide powerful tools for predicting distributional shifts (Xu et al., 2019). Using species occurrence data and environmental factors, SDMs can simulate potentially suitable habitats across time periods, supporting more targeted germplasm collection. Commonly used SDMs include ecological niche factor analysis (ENFA) (Basille et al., 2008), maximum entropy modeling (MaxEnt) (Jaynes, 1957), generalized linear modeling (GLM) (Gotway and Stroup, 1997), and the genetic algorithm for rule-set prediction (GARP) (Stockwell and Noble, 1992). Among these, MaxEnt combined with GIS is particularly robust, providing predictions with high stability and good agreement with observed species distributions (Wang et al., 2007; Wu et al., 2020; Yan et al., 2020; Qin et al., 2020).

The distribution of *O. rufipogon* is influenced by a complex interplay of climatic, ecological, and geographical factors. Therefore, identifying germplasm-rich regions and predicting distributional shifts under various climate scenarios are essential for effective conservation. In this study, we integrated global distribution data of *O. rufipogon* with 20 environmental factors, applying GIS-based spatial analysis and MaxEnt modeling to simulate suitable habitats during the Last Interglacial (LIG, 130–115 ka before present [BP]) (Quiquet et al., 2015), the Last Glacial Maximum (LGM, 26–19 ka BP) (Schaefer et al., 2015), and the present period. We also evaluated the impacts of climate change on *O. rufipogon* distribution, determined the optimal number of sampling groups, and proposed sampling strategies for distinct ecoregions. Together, these efforts provide a scientific foundation for future global exploration, collection, and conservation of *O. rufipogon* germplasm.

Results

Spatial distribution of *O. rufipogon*

O. rufipogon is primarily distributed across Southeast Asia, East Asia, and South Asia, including China, Laos, Vietnam, Thailand, Myanmar, India, Cambodia, the Philippines, and Malaysia (Fig. 2). These regions are characterized by warm, humid climates and abundant water resources, which are conducive to the growth of *O. rufipogon*. In China, *O. rufipogon* occurs mainly in the southern provinces, especially Guangdong, Guangxi, Yunnan, and Hainan, with smaller populations also found in Jiangxi, Fujian, Hunan, and Taiwan. These areas typically experience humid subtropical or tropical climates with ample precipitation, providing favorable conditions for *O. rufipogon* growth. In contrast to its broad distribution in the Northern Hemisphere, *O. rufipogon* has a more restricted presence in the Southern Hemisphere, where it occurs mainly in the humid regions of northern Australia. Although climate conditions in these areas are favorable, its distribution remains sparse, suggesting that the species may have lower adaptability in the Southern Hemisphere. Overall, *O. rufipogon* thrives in warm, humid environments, particularly those with pronounced monsoon patterns and abundant water resources.

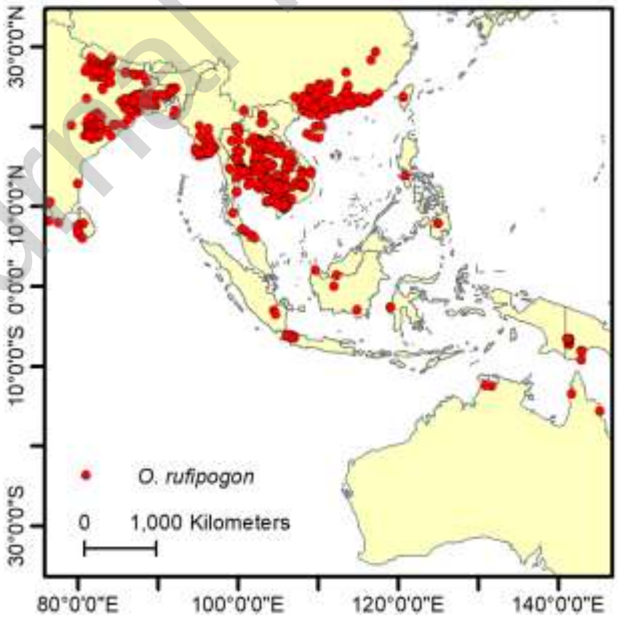


Fig. 2. Global spatial distribution of *O. rufipogon*.

This figure is based on the standard map downloaded from the Standard Map Service System (http://bzdt.ch.mnr.gov.cn/) with the approval number GS(2016)1663. The basemap has not been modified; it has only been color-processed.

Model simulation results

The MaxEnt model was used to simulate and evaluate the potential suitable distribution of *O. rufipogon* across three distinct climatic periods: the LIG, characterized by a relatively warm and wet climate with higher sea levels; the LGM, marked by extremely cold and dry conditions with substantially lower sea levels; and the Current period, defined by rising temperatures, shifting precipitation patterns, and more frequent extreme climate events. The simulations were based on known distribution records combined with 20 environmental variables (Table 1). Model predictive performance for each climatic period is shown in Fig. 3 and was assessed using Receiver Operating Characteristic (ROC) curves. In each ROC curve, the horizontal axis represents "1 - Specificity (Fractional Predicted Area)" and the vertical axis represents "Sensitivity (1-Omission Rate)", which together indicate the classification performance of the model across different thresholds. All three ROC curves are close to the upper left corner, indicating strong predictive performance through a better balance between sensitivity and specificity. The primary evaluation metric, the area under the curve (AUC), measures overall predictive performance, with values closer to 1 indicating stronger discriminative ability. As shown in Fig. 3, the model achieved mean AUC values of 0.952, 0.945, and 0.952 for the three periods, all exceeding 0.9. The consistently high performance demonstrates the model's strong predictive ability and stability under varying climatic conditions. The black diagonal line in Fig. 3 represents a random prediction scenario. Compared with this reference line, the ROC curves for all three models show substantially better performance than random prediction, further confirming the predictive efficiency of the model. In summary, both the ROC curves and AUC values support the high precision of the MaxEnt model in predicting potentially suitable habitats for *O. rufipogon*.

Table 1. Environmental variables for modeling

Variable code	Environmental variables
BIO1	Annual mean temperature
BIO2	Mean diurnal range (Mean of monthly (max temperature-min temperature))
BIO3	Isothermality ((BIO2/BIO7) × 100)
BIO4	Temperature seasonality (standard deviation × 100)
BIO5	Max temperature of warmest month
BIO6	Min temperature of coldest month
BIO7	Temperature annual range (BIO5-BIO6)
BIO8	Mean temperature of wettest quarter
BIO9	Mean temperature of driest quarter
BIO10	Mean temperature of warmest quarter
BIO11	Mean temperature of coldest quarter
BIO12	Annual precipitation
BIO13	Precipitation of wettest month
BIO14	Precipitation of driest month
BIO15	Precipitation seasonality (coefficient of variation)
BIO16	Precipitation of wettest quarter
BIO17	Precipitation of driest quarter
BIO18	Precipitation of warmest quarter
BIO19	Precipitation of coldest quarter
Ele	Elevation

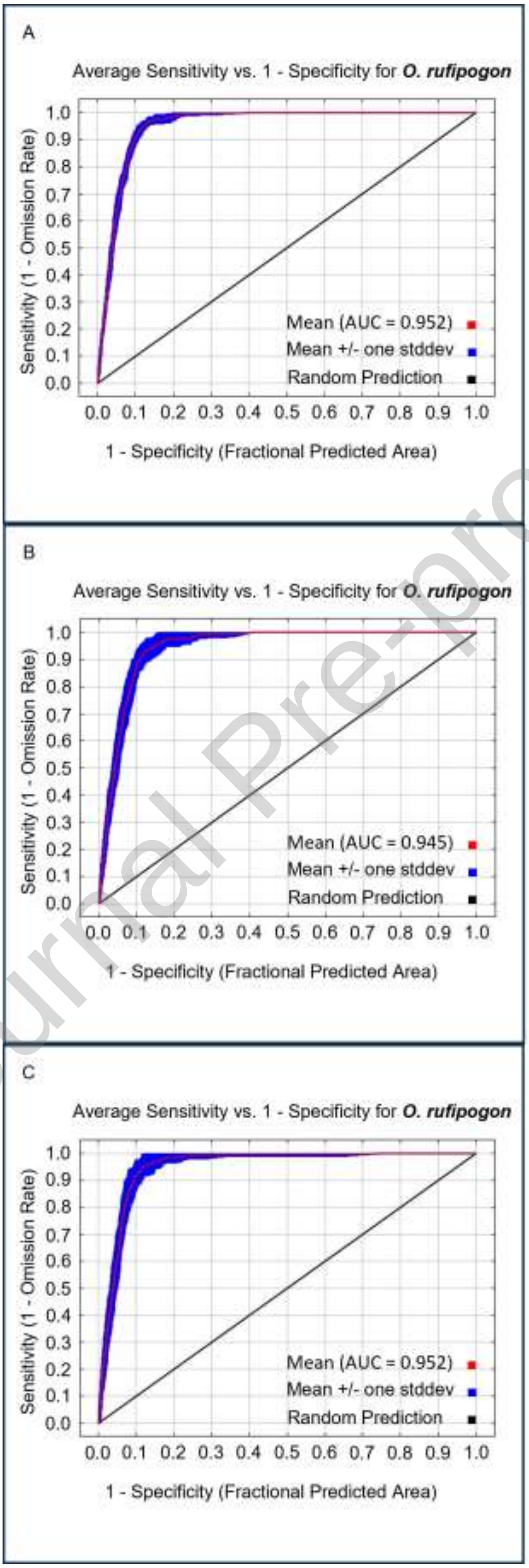

A
Average Sensitivity vs. 1 - Specificity for **O. rufipogon**
Sensitivity (1 - Omission Rate)
1.0 0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0.0
Mean (AUC = 0.952) ■
Mean +/- one stddev ■
Random Prediction ■
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
1 - Specificity (Fractional Predicted Area)
B
Average Sensitivity vs. 1 - Specificity for **O. rufipogon**
Sensitivity (1 - Omission Rate)
1.0 0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0.0
Mean (AUC = 0.945) ■
Mean +/- one stddev ■
Random Prediction ■
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
1 - Specificity (Fractional Predicted Area)
C
Average Sensitivity vs. 1 - Specificity for **O. rufipogon**
Sensitivity (1 - Omission Rate)
1.0 0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1 0.0
Mean (AUC = 0.952) ■
Mean +/- one stddev ■
Random Prediction ■
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
1 - Specificity (Fractional Predicted Area)

Fig. 3. Receiver operating characteristic (ROC) curves illustrating the discriminative performance of the MaxEnt model.

(A) Current; (B) LGM; (C) LIG. AUC, area under the ROC curve.

Key environmental variables influencing the suitable distribution of *O. rufipogon*

The contribution of environmental factors to the distribution of *O. rufipogon* across different periods was assessed using the Jackknife test. This method determines the weight of each variable in shaping the distribution of *O. rufipogon*, and variables with minimal contribution were removed to improve prediction accuracy (Li et al., 2020; Zhang et al., 2021). According to the results (Table 2), although the ranking of environmental factors differs among periods, annual precipitation (BIO12) and mean annual temperature (BIO1) consistently emerged as the two most important variables, each showing the highest contribution rate in every period. These factors exert the strongest influence on the distribution of *O. rufipogon* across the Current, LGM, and LIG periods. Specifically, BIO12 contributed over 60% in each period, whereas BIO1 accounted for more than 10% in all periods.

As shown in Table 2, the consistently high contribution rates of BIO12 and BIO1 across the Current, LGM, and LIG periods highlight their central roles in determining the suitable distribution of *O. rufipogon*. The markedly higher contribution of BIO12 compared with other factors indicates that water availability is a primary limiting factor for the distribution of *O. rufipogon*. Similarly, the substantial contribution of BIO1 suggests that *O. rufipogon* is sensitive to temperature fluctuations. Overall, the Jackknife test demonstrates that water availability and ambient temperature are the two most critical environmental determinants of *O. rufipogon* distribution across climatic periods, with BIO12 playing a particularly dominant role.

Table 2. Contribution rates of major environmental factors to the distribution of *O. rufipogon*

Current		LGM		LIG	
Variable code	Contribution rate (%)	Variable code	Contribution rate (%)	Variable code	Contribution rate (%)
BIO12	66.9	BIO12	64	BIO12	69
BIO1	13.7	BIO1	12.3	BIO1	13.1
BIO15	8.8	BIO15	10.4	BIO14	4.6
Ele	4.8	BIO2	6	BIO8	3.9
BIO3	2.5	BIO14	5.1	BIO5	3.2
BIO14	1.2	BIO19	2.1	BIO15	2.8

Prediction of potential suitable areas for *O. rufipogon* at different periods

The potential distribution of *O. rufipogon* across different periods is illustrated in Fig. 4, with color gradients indicating distribution probability. Red areas represent Low Suitability Zones, where climatic or environmental constraints result in a low probability of occurrence. Green areas represent Medium Suitability Zones, where *O. rufipogon* can survive and reproduce, although certain environmental factors may restrict population size and distribution. Purple areas represent High Suitability Zones, indicating regions with highly favorable conditions for the growth and spread of *O. rufipogon*, and a relatively high probability of its presence. Overall, the potential distributions of *O. rufipogon* during the Current period, the LGM, and the LIG exhibit broadly similar geographic patterns, with primary concentrations in southern China, Southeast Asia, and parts of South Asia. However, the potential distribution of *O. rufipogon* varied across these periods (Table 3). At present, the total potential distribution area is 62,404,680 km², including 58,619,800 km² of low suitability, 2,316,040 km² of medium suitability, and 1,468,840 km² of high suitability. The species is mainly concentrated in Southeast Asia, southern China, and parts of South Asia (Fig. 4A). Overall, the predicted patterns are highly consistent with those reported in previous studies (Liu et al., 2015; Lin et al., 2024). During the LGM, the total potential distribution area expanded to 72,431,160 km², comprising 66,877,030 km² in low suitability areas, 3,035,450 km² in medium suitability areas, and 2,518,675 km² in high suitability zones. This period showed a marked expansion in the range of *O. rufipogon*, with a clear increase in highly suitable areas compared with the other periods (Fig. 4B). In the LIG, the total potential distribution area decreased to 61,586,580 km², consisting of 58,116,230 km² of low suitability, 1,877,250 km² of medium suitability, and 1,593,100 km² of high suitability. Although the total area was slightly smaller than the current range, the extent of high suitability zones exceeded that of the current period (Fig. 4C).

Notably, the distribution of *O. rufipogon* in northern Australia showed significant variation across the three climatic periods. During both the present stage and the LGM, northern Australia included areas of high suitability for *O. rufipogon*, suggesting favorable environmental conditions. However, during the LIG, northern Australia did not contain any high suitability zones, likely because climatic changes rendered the region less suitable for growth. Compared with the LGM, the potential distribution range of *O. rufipogon* has clearly contracted in the present. Areas that were once part of its distribution during the LGM, such as parts of East and Southeast Asia, have experienced a notable decline in suitability. The LIG was characterized by a relatively warm and wet climate with higher sea levels, whereas the LGM was marked by extremely cold and dry conditions and a substantial drop in sea level. In the current period, ongoing climate change has resulted in rising temperatures, altered precipitation patterns, and a growing frequency of extreme climate events. In summary, the potential distribution of *O. rufipogon* varied considerably across climatic periods, with the LGM showing the broadest distribution, while the LIG displayed slightly reduced overall coverage but included more high suitability areas than the present. This pattern underscores the strong influence of climate and environmental conditions on the distribution range of *O. rufipogon* over time.

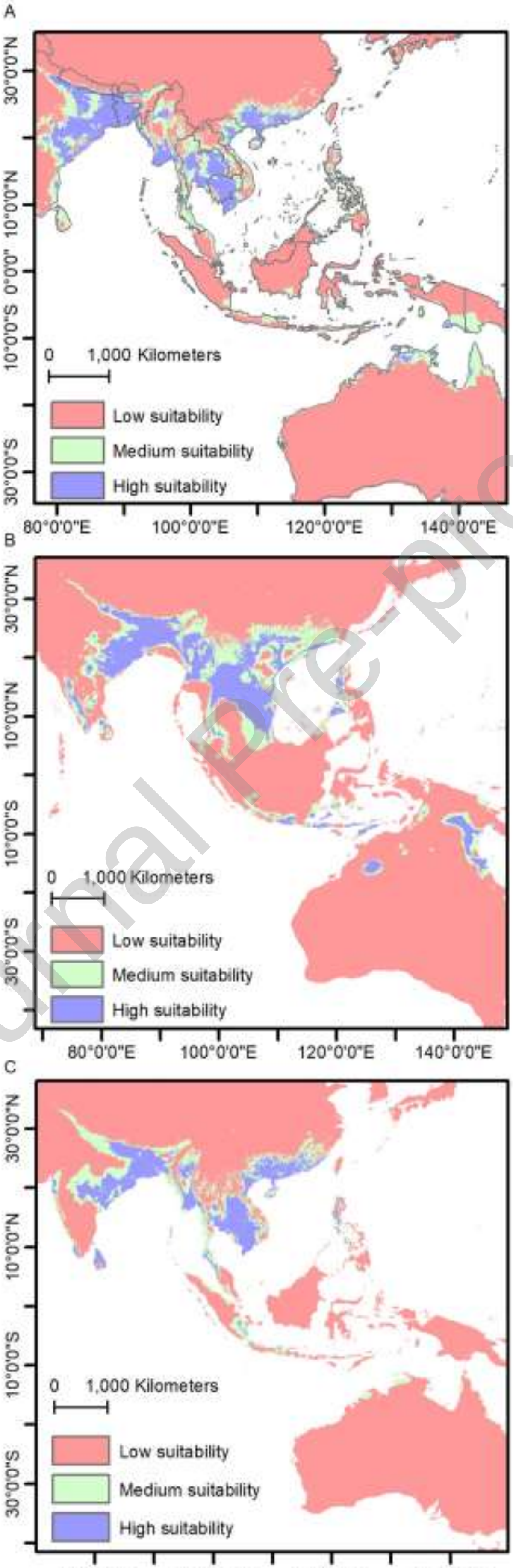

A
30°0'0"N
10°0'0"N
0°0'0"
10°0'0"S
30°0'0"S
80°0'0"E
100°0'0"E
120°0'0"E
140°0'0"E
0
1,000 Kilometers
🟥 Low suitability
🟩 Medium suitability
🟦 High suitability
B
30°0'0"N
10°0'0"N
10°0'0"S
30°0'0"S
80°0'0"E
100°0'0"E
120°0'0"E
140°0'0"E
0
1,000 Kilometers
🟥 Low suitability
🟩 Medium suitability
🟦 High suitability
C
30°0'0"N
10°0'0"N
10°0'0"S
30°0'0"S
80°0'0"E
100°0'0"E
120°0'0"E
140°0'0"E
0
1,000 Kilometers
🟥 Low suitability
🟩 Medium suitability
🟦 High suitability

Fig. 4. Potential suitable habitats of *O. rufipogon* during different periods.

(A) Current; (B) LGM; (C) LIG. Only (A) is based on the standard map downloaded from the Standard Map Service System (http://bzdt.ch.mnr.gov.cn/) with the approval number GS(2016)1663. The basemap in (A) has not been modified; only color processing was applied.

Table 3. Potentially suitable areas of *O. rufipogon* at different periods

Periods	Total suitable region (km^2)	Low Suitability Areas (km^2)	Medium Suitability Areas (km^2)	High Suitability Areas (km^2)
Current	62,404,680	58,619,800	2,316,040	1,468,840
LGM	72,431,160	66,877,030	3,035,450	2,518,675
LIG	61,586,580	58,116,230	1,877,250	1,593,100

Prediction of sampling areas for different periods

The prediction of sampling areas for different periods of *O. rufipogon*, based on the above analysis combined with ecoregion information, provides more precise guidance for the subsequent sampling strategy. By integrating climatic, ecological, and geographical distribution patterns, the ecoregion analysis of *O. rufipogon* offers strong support for selecting sampling areas that cover different climatic zones, topographic features, and ecosystems. In addition, the growth conditions, growth status, and environmental resilience of *O. rufipogon* vary across climatic periods. These responses to environmental differences have led to changes in its genetic diversity. To develop a comprehensive sampling strategy and preserve the genetic diversity and integrity of *O. rufipogon* germplasm resources, we conducted ecological zoning for the LIG, LGM, and the present periods to enable more accurate predictions of suitable sampling areas (Fig. 5 and Table 4).

The potentially suitable areas in the present period can be divided into three ecoregions based on climatic conditions (Fig. 5A). The Southeast and Eastern Asian tropical monsoon ecoregion, including southern China, Vietnam, and Cambodia, has a typical tropical monsoon climate with abundant and strongly seasonal precipitation. The sampling strategy should capture the diversity within these regions, particularly across different altitudes, to reflect the adaptation of *O. rufipogon* to high humidity conditions. The South Asia-dominated grassland and monsoon climate zone, including India and parts of Myanmar and Thailand, is relatively arid but strongly influenced by monsoon systems and contains diverse ecosystems. The sampling strategy should include plains, hills, and areas with variable water availability to represent adaptation to fluctuating soil moisture and precipitation. The tropical rainforest ecoregion in northern Australia, the Philippines, and Indonesia is an important distribution area for *O. rufipogon*, with consistently high temperatures and rainfall. Sampling should be conducted across different forest types and water-rich areas to obtain data on *O. rufipogon* adaptation in tropical rainforest ecosystems.

During the LGM, climatic conditions were relatively cold and dry, and the distribution became fragmented. The ecoregions for this period can be summarized as follows (Fig. 5B). The Fragmented Tropical Monsoon Zone, covering scattered areas in Southeast Asia, requires sampling across many small geographic regions to capture variation in climate and vegetation, including different microclimatic zones such as mountains and river valleys, to explore how *O. rufipogon* survived ice age environments. The arid and mountainous ecoregions of South Asia, particularly in northern and central India, likely provided refuges during the LGM. Sampling should therefore focus on higher elevations and arid mountain regions. Localized humid areas in northern Australia indicate that the region still contained certain ecoregions with relatively humid conditions. Sampling should prioritize water-rich areas to investigate the adaptability of *O. rufipogon* in these relatively isolated ecosystems.

During the LIG, the distribution of *O. rufipogon* expanded considerably, and its ecoregion encompassed extended tropical monsoon and subtropical monsoon climatic zones, especially in Southeast Asia and China, where both distribution range and suitability were broader (Fig. 5C). Sampling should include climatic transition zones across latitudes to ensure coverage of the ecological transition from Southeast Asia to East Asia, which is essential for understanding adaptation during warm and humid periods. Transitional zones between grassland and rainforest in South and Southeast Asia also represent suitable habitats for *O. rufipogon*, and sampling should focus on these ecological transition regions to investigate survival mechanisms under fluctuating climatic conditions. Furthermore, marginal ecoregions in Oceania were suitable for *O. rufipogon*, so sampling should include coastal ecosystems and humid lowlands to clarify distribution patterns within these peripheral environments.

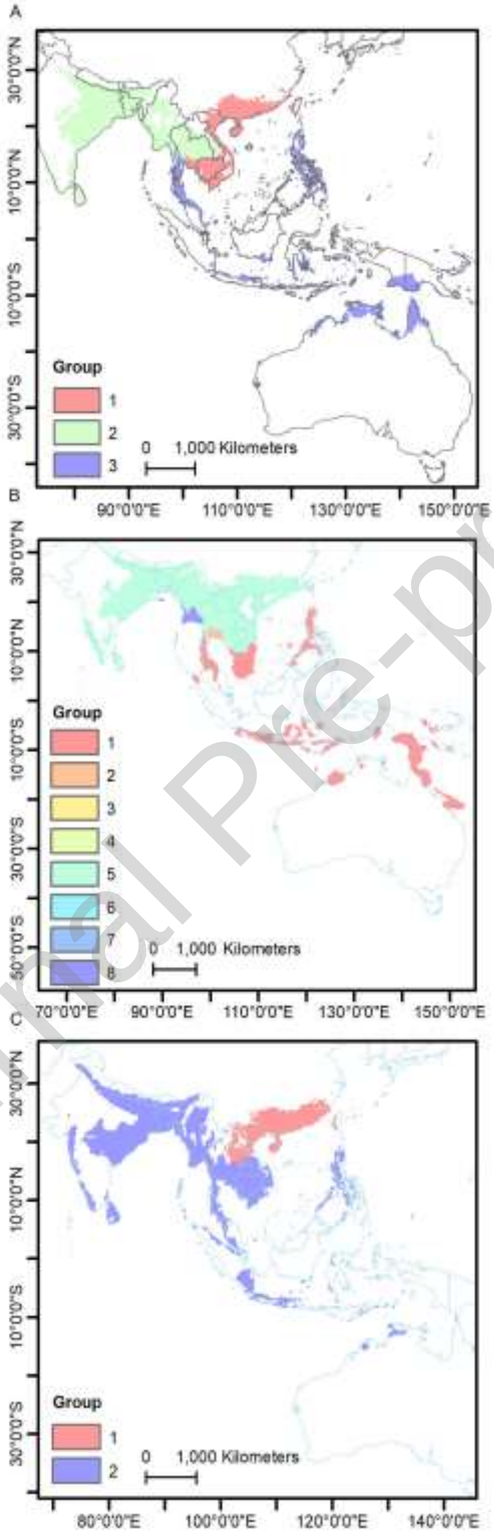


Fig. 5. Division of sampling areas for *O. rufipogon* across different periods.

(A) Current; (B) LGM; (C) LIG. This figure is based on the standard map downloaded from the Standard Map Service System (http://bzdt.ch.mnr.gov.cn/) with the approval number

GS(2016)1663. The basemap has not been modified; only color processing was applied.

Table 4. Areas of *O. rufipogon* sampling zones across different periods

Grouping number	Current (km^2)	LGM (km^2)	LIG (km^2)
1	923,105	1,651,010	895,068
2	2,313,320	57,661	3,384,540
3	1,433,060	5,819	
4		13,225	
5		3,841,070	
6		7,406	
7		3,174	
8		103,155	

Discussion

Impact of climate change on the distribution of *O. rufipogon*

Climate change exerts complex effects on plant distributions, altering both the size and location of suitable habitats. Consistent with previous studies, our analysis of *O. rufipogon* distribution across different climatic periods identified BIO12 as the most important climatic factor determining habitat suitability (Huang and Schaal, 2012; Liu et al., 2015), followed by BIO1. These findings underscore the species' strong dependence on the warm and humid conditions typical of tropical and subtropical regions. Optimal growth of *O. rufipogon* requires adequate light, a warm climate, and sufficient precipitation, factors that collectively support its widespread distribution in these areas.

Changes in potential suitable areas from LIG to present

From the LIG to the present, the size of suitable areas for *O. rufipogon* has fluctuated in a "low–high–medium" pattern. During the LIG, although the climate was relatively warm and humid, which is generally favorable for *O. rufipogon*, the overall suitable area was limited. This may reflect the restricted range of *O. rufipogon* at the time, which may not have fully expanded into all potentially suitable regions, along with geographic or local ecological constraints. By the LGM, the global climate had turned colder and drier, yet the suitable area for *O. rufipogon* unexpectedly increased. This result may be linked to topographic and hydrological changes caused by glacier expansion and retreat, which created new moist environments and soil conditions conducive to growth. Although the general climate during the LGM was harsher, these localized changes allowed the species to establish new microclimates and ecological niches, resulting in an expanded distribution. In modern times, the size of suitable areas for *O. rufipogon* has decreased again, mainly due to habitat destruction or disappearance caused by human-driven land-use change (Huang and Schaal, 2012). Field investigations have provided substantial evidence of this decline. For example, surveys in Guangxi, Hainan, Guangdong, Yunnan, and Thailand have shown significant reductions in *O. rufipogon* habitats due to human activities (Gao et al., 1996; Akimoto et al., 1999; Dai et al., 2004; Huang et al., 2009; Yun et al., 2015; Fan et al., 2017). Moreover, a recent study using Landsat Earth observation data combined with field population surveys also demonstrated a significant decrease in *O. rufipogon* distribution sites linked to human-induced land-use changes (Chen et al., 2022). Factors such as agricultural expansion, urbanization, and habitat loss have substantially reduced the natural habitats of *O. rufipogon*. Additionally, extreme weather events and altered precipitation patterns associated with climate change have further restricted its distribution.

Changes in potential suitable areas in northern Australia

The potential distribution of *O. rufipogon* in northern Australia varied significantly across the three climatic periods, particularly during the LIG when no highly suitable areas were present. This pattern may reflect climatic fluctuations and sea-level changes that altered landscapes and habitats differently across periods. During the LIG, the warmer global climate likely caused sea levels to rise (Dutton and Lambeck, 2012), submerging low-lying areas in northern Australia and reducing suitable habitats for *O. rufipogon*. The distribution and evolution of species are closely linked to environmental changes (Xu et al., 2010). During the LGM, the global sea level was significantly lower than it is today. Studies have shown that sea level around Australia during the LGM was approximately 125 meters lower, resulting in a much larger land area. The Australian continent expanded and became connected to New Guinea, Tasmania, and many smaller islands by land bridges, forming a single landmass known as "Sahul" (Heaney, 1991; Hewitt, 2000; Lewis et al., 2013; Williams et al., 2018; Du et al., 2024). These pronounced geographical changes reshaped local landforms and ecological environments, leading to major shifts in vegetation distribution across Australia and potentially creating new suitable habitats for *O. rufipogon*.

Sampling strategies and conservation of *O. rufipogon* germplasm resources

The basic attributes of crop germplasm resources include genetic diversity, genetic specificity, genetic integrity, and genetic accumulation. Among these, genetic integrity is fundamental to the collection and conservation of germplasm resources (Liu et al., 2023). However, the current survey and collection of wild relatives of crops such as wild soybean, wild maize relatives, and wild wheat relatives still rely predominantly on traditional methods (Fig. 1) (Wang and Li, 2012; Chu et al., 2016; Zeng et al., 2019). These traditional survey and collection approaches have obvious limitations, including time-consuming and labor-intensive procedures and limited geographic coverage. Genetic variation is widespread in *O. rufipogon* both among and within populations. Additionally, *O. rufipogon* shows extensive genetic diversity through adaptive responses to diverse climatic conditions. Therefore, it is necessary to conduct extensive investigations and scientifically structured sampling to preserve the genetic integrity of germplasm resources during collection. To enhance the conservation of *O. rufipogon* germplasm and expand preservation efforts internationally, we integrated GIS and MaxEnt models with global distribution data of *O. rufipogon* and multiple environmental variables to analyze potential habitats across different climate periods (Fig. 6). We propose prioritizing survey and collection efforts in key global regions—particularly Southeast Asia, East Asia, South Asia, and northern Australia (Fig. 7)—where *O. rufipogon* is widely distributed across diverse habitats. These areas may face heightened risks of habitat loss due to climate change and human activities, making sampling efforts especially important for germplasm conservation. Environmental factors such as temperature, topography, and especially precipitation play crucial roles in determining the distribution of *O. rufipogon*, and these factors must be carefully considered in future conservation and management strategies. Understanding the relationship between *O. rufipogon* distribution and its environmental requirements will provide a scientific foundation for developing conservation practices and ensuring the sustainable use of this valuable genetic resource.

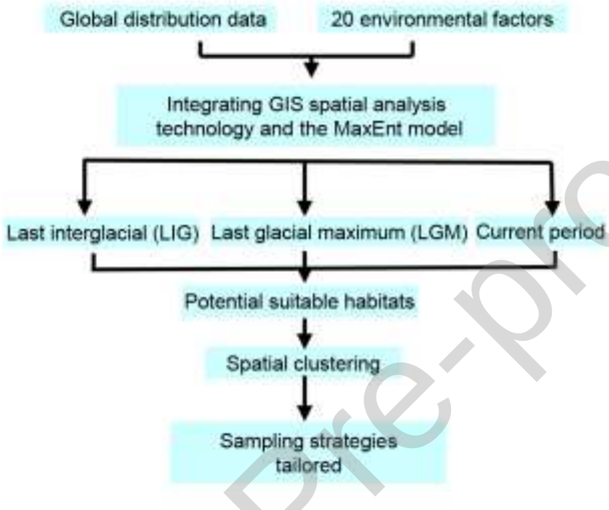


Fig. 6. Technical roadmap outlining the regional prediction and sampling strategy for *O. rufipogon*.

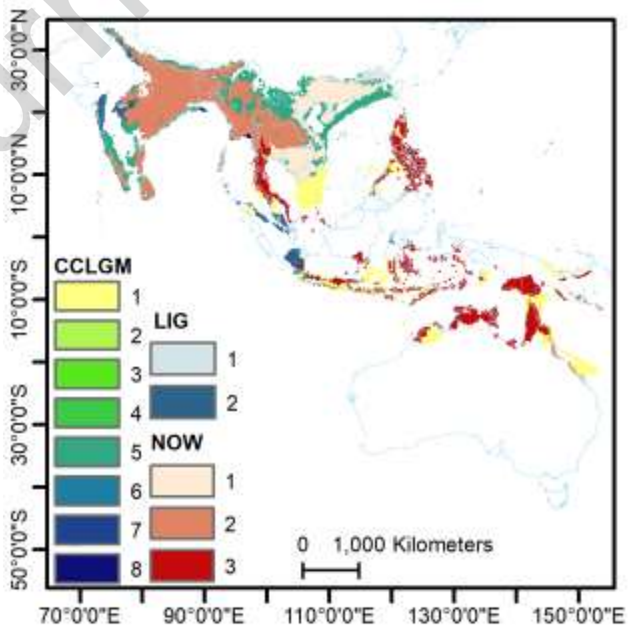


Fig. 7. Diagram of the sampling strategy designed to ensure the genetic integrity of *O. rufipogon* germplasm resources.

This figure is based on the standard map downloaded from the Standard Map Service System (http://bzdt.ch.mnr.gov.cn/) with the approval number GS(2016)1663. The basemap has not been modified; only color processing was applied.

Ecological feasibility and strategic implications of targeted sampling for *O. rufipogon*

This study demonstrates the utility of combining ecological niche modeling with environmental clustering to develop a scientifically grounded and spatially optimized sampling strategy for *O. rufipogon*, a key wild progenitor of cultivated rice. By employing the MaxEnt model alongside GIS-based spatial analysis, we identified major ecological zones currently suitable for the species and revealed their historical dynamics during the LIG and LGM. The model performed strongly, as indicated by high AUC values, capturing known distribution patterns while also identifying additional regions with potential suitability. By incorporating both climatic and topographic variables, the proposed strategy enables targeted germplasm collection that emphasizes ecological diversity, adaptive potential, and representativeness, thereby improving upon traditional uniform or opportunistic sampling approaches.

Despite its demonstrated strength, implementation of this strategy presents logistical and practical challenges, especially in politically sensitive or remote regions such as the Indo-Myanmar frontier, parts of the Philippines, and northern Australia. Field access may be limited by geopolitical factors, weak infrastructure, or environmental hazards. These constraints can be mitigated through collaboration with local agricultural research institutions, non-governmental organizations, and international agencies such as the International Rice Research Institute (IRRI), which can support permitting processes and assist with *in situ* operations. Furthermore, the ecological heterogeneity across the target regions, from lowland wetlands and tropical rainforests to upland grasslands, requires the use of complementary tools such as high-resolution remote sensing, UAV reconnaissance, and landscape-level planning to improve sampling accuracy and efficiency. The risk of capturing hybrid individuals or introgressed populations, particularly near agricultural zones, also highlights the need for early-stage molecular screening using SSR markers or SNP arrays to ensure the genetic purity of collected material.

From a cost–benefit perspective, the adoption of this strategy is expected to significantly lower the financial and logistical burdens of field collection by minimizing redundant efforts and directing resources toward high-probability zones. Preliminary estimates suggest that it could reduce field expenditures by 30–50% while improving the efficiency and genetic utility of collected accessions. Such gains are especially important for global *ex situ* conservation initiatives, where limited resources require strategies that maximize return on investment. More importantly, this strategy enhances the long-term usefulness of germplasm repositories by ensuring the inclusion of alleles relevant to abiotic stress resistance, disease tolerance, and other agronomically important traits that remain underrepresented in cultivated germplasm pools.

Ultimately, the framework established here not only strengthens the conservation of *O. rufipogon* but also provides a model for safeguarding other crop wild relatives whose natural habitats are increasingly threatened by land-use change and climate disruption. By combining predictive modeling with ecological stratification, we offer a replicable and scalable approach to strategic germplasm sampling that balances scientific rigor with practical feasibility. Such integrative conservation planning is essential for sustaining global crop genetic resources and supporting future food security.

Future research prospects

In this study, given the crucial impact of climate on the survival and distribution of species, we used climate and topographic factors to predict the potential suitable habitats of *O. rufipogon*. However, in addition to climate and topography, other ecological factors such as soil type, hydrological conditions, and biotic interactions also have significant effects on the growth and distribution of *O. rufipogon*. As the foundation for plant growth, soil nutrients such as available nitrogen, phosphorus, and potassium are key determinants of the structure of the root-associated bacterial communities of *O. rufipogon*. Differences in these nutrients can change the types and quantities of rhizosphere microorganisms, indirectly affecting plant growth and health, while soluble sugars and proline content in the soil also play important roles in its growth process (Neelam et al., 2017; Tian et al., 2017; Chang et al., 2022). The impact of hydrological conditions on *O. rufipogon* is likewise considerable and cannot be overlooked. Different hydrological regimes strongly influence its tillering dynamics. For instance, *O. rufipogon* grown in cement pots, free from environmental stresses, showed not only an increase in tiller number but also a marked rise in the biomass of both vegetative and reproductive structures. This indicates that stable hydrological conditions in natural environments are crucial for its growth and development, whereas unfavorable changes may limit tillering ability and biomass accumulation, thereby affecting its distribution range (Kariali et al., 2008; Mohapatra et al., 2011). In terms of biotic interactions, *O. rufipogon* competes with surrounding perennial vegetation and is vulnerable to invasive alien species, both of which can strongly influence its distribution patterns (Gao et al., 1996). Although these ecological factors were not included in the model used in this study, we recognize their importance. Future research should integrate soil, biological, and hydrological data to improve predictions of *O. rufipogon*'s potential suitable habitats, enabling a more comprehensive and accurate understanding of its distribution patterns and providing a stronger scientific basis for the conservation and sustainable utilization of its germplasm resources. We hope that future studies will address these aspects through in-depth investigations.

Materials and methods

O. rufipogon distribution data set

The distribution data of *O. rufipogon* used in this study were derived from global genebanks, such as the Genesys platform (Genesys, 2024), which aggregates information on plant genetic resources worldwide. Through collaboration with agricultural research centers and plant resource banks around the globe, this international dataset covers *O. rufipogon* distribution across Asia, Africa, and the Americas. The data document the natural range and ecological environments of *O. rufipogon* in various countries and provide valuable material for international scientific collaboration. Over the years, the Global Wild Rice Germplasm Repository has enriched the diversity data available for *O. rufipogon*, supporting future studies in rice breeding and genetic diversity. This comprehensive dataset provides a strong foundation for the conservation, resource management, and future breeding research of *O. rufipogon*.

Environmental factor data set

MaxEnt operates on the principle of maximum entropy, which improves prediction accuracy by selecting the distribution probability with the fewest assumptions. As outlined previously (Phillips et al., 2006; Phillips and Dudik, 2008), species occurrence records and environmental variables are required to model species distributions effectively. In this study, we selected 19 key climatic variables and one topographic variable known to influence species distributions as input parameters. These variables included mean annual temperature (BIO1), mean diurnal range (BIO2), temperature annual range (BIO7), annual precipitation (BIO12), precipitation of wettest month (BIO13), precipitation of driest month (BIO14), and elevation, among others. Data for the Current, LGM, and LIG periods were sourced from the WorldClim database (https://worldclim.org) (Fick and Hijmans, 2017). The datasets for the Current and LGM periods have a spatial resolution of 2.5 arc-minutes, while the LIG dataset provides a higher resolution of 30 arc-seconds. All environmental layers were standardized to the WGS-1984 geographic coordinate system to ensure consistency and accuracy.

Model analysis

The collected *O. rufipogon* distribution data were screened in ArcGIS 10.7 (Redlands, 2011) to remove duplicate and invalid loci information, ensuring that the distribution points within each raster were uniformly distributed and reducing data bias (Qian et al., 2021). The processed distribution data were then matched with corresponding environmental variables so that each point had both climatic and geographic information, and multivalue extraction was used to obtain climate factor data for each point. The species distribution model (SDM) was generated for *O. rufipogon* using MAXENT 3.4.1 (Phillips et al., 2017), employing maximum entropy modeling to predict its geographical distribution. Twenty environmental variables were selected for model training. Appropriate regularization parameters were chosen to prevent overfitting, and 10-fold cross-validation was used to assess model stability and predictive performance. Seventy-five percent of the processed *O. rufipogon* distribution points were used as a training set, and the remaining 25% served as a test set to evaluate model accuracy. The model iterations were repeated 10 times, and the average output from the 10 runs was taken as the final simulation result.

Model performance

The Receiver Operating Characteristic (ROC) curve was used to evaluate the effectiveness of the MaxEnt model in simulating the potential distribution of *O. rufipogon*. The ROC curve was constructed by comparing the True Positive Rate with the False Positive Rate of the model. The area under the ROC curve (AUC) served as the primary evaluation metric, with values ranging from 0 to 1. The closer the AUC value is to 1, the stronger the discriminative ability of the model, indicating better simulation performance. Conversely, a lower AUC value reflects weaker model performance in predicting the potential distribution of *O. rufipogon*.

Classification of simulation results

Based on the MaxEnt simulations, the potential distribution areas of *O. rufipogon* were classified into three categories: Low Suitability Areas, Medium Suitability Areas, and High Suitability Areas. This classification was derived from the Habitat Suitability Index (HSI), which ranges from 0 to 1 and represents the likelihood of *O. rufipogon* survival in each area.

For ease of interpretation and practical application, we defined the suitability categories as follows: 1. Low Suitability Areas (Unsuitable for Survival): Regions with an HSI between 0 and 0.4 were classified as Low Suitability Areas. Due to climatic or environmental limitations, these areas are generally unfavorable for the growth and distribution of *O. rufipogon* and are considered unsuitable for survival. 2. Medium Suitability Areas (Moderately Suitable Distribution): Regions with an HSI between 0.4 and 0.6 were classified as Medium Suitability Areas. These areas have relatively favorable climatic conditions, allowing *O. rufipogon* to survive and reproduce, although population size and distribution may be limited by certain environmental factors. 3. High Suitability Areas (Optimal Distribution): Regions with an HSI greater than 0.6 were classified as High Suitability Areas. These areas provide highly favorable environmental conditions, making them ideal for the growth and spread of *O. rufipogon*. In these regions, populations are expected to be more stable and widely distributed. This classification provides a practical framework for identifying and managing suitable habitats for *O. rufipogon* conservation and potential breeding programs.

Contribution rates of environmental factors

The contribution of environmental factors to the distribution of *O. rufipogon* over time was analyzed using the Jackknife test. This method is valuable for determining the relative importance of each environmental factor in species distribution by assessing how model predictions change when each variable is removed individually. This approach clarifies the influence of each climate factor on the potential distribution of *O. rufipogon*.

Sampling area division

Based on the MaxEnt modeling results, the classification of suitable habitats, and the contribution of each environmental factor, we spatially divided the sampling region for *O. rufipogon* using a grouping analysis method. In the initial classification stage, we set a default number of groups and applied a function to evaluate the optimal number of groups. To further refine the grouping and determine the most appropriate number of clusters, we calculated the pseudo F statistic following Assunção et al. (2006). This statistic helps assess the quality of grouping and informs the reclassification process to achieve the most reasonable spatial division. A brief explanation of the pseudo F statistic calculation is provided below:

$$F = \frac{(T-P_g)/(g-1)}{P_g/(n-g)} \qquad (1)$$

- T denotes the total sum of squares;
- Pg denotes the intra-group sum of squares when divided into g groups;
- g is the number of groups;
- n is the total number of samples.

When regions are classified into *g* groups, an effective grouping should exhibit minimal variance within each group and maximal variance between groups. The pseudo F statistic evaluates grouping quality by comparing the ratio of within-group and between-group sums of squares. Ideally, the most reasonable classification into *g* groups produces a higher pseudo F value with a smaller number of groups. Based on this principle, we used the pseudo F statistic to identify the optimal number of groupings.

The process involved three steps:

1. **Initial Grouping**: We began with a default number of groups based on the suitability zones of the MaxEnt model and an initial division of regions according to the contributions of environmental factors.
2. **Calculation of Pseudo F Values**: We calculated the pseudo F value for each possible grouping number using the pseudo F statistic. The grouping configuration with the highest pseudo F value and the smallest number of groups was selected as the optimal grouping scheme.
3. **Adjustment and Final Grouping**: Based on the results of the previous steps, we refined the number of subgroups to achieve the most reasonable division. The final classification ensured that each sampling area captured both internal environmental consistency and clear spatial differentiation.

This approach allowed us to reclassify sampling areas effectively while maintaining strong representation of environmental variation across regions. In addition, to ensure the genetic integrity of *O. rufipogon* germplasm resources, we merged the sampling regions predicted for the three periods (Present, LGM, and LIG). The merged area forms the basis of our sampling strategy map, ensuring that all essential genetic information is included in future survey and collection efforts.

Author contributions

Xiaoming Zheng designed the study. **Xiaoming Zheng**, **Kai Wang**, **Jinhua Long,** and **Jizhou Lei** analyzed the data and wrote the manuscript. **Qian Qian**, **Yuntao Liang,** and **Disna Ratnasekera** revised the manuscript. All authors contributed to and approved the final version of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was supported by the National Key Research and Development Program of China (2021YFD1200101); the Project of Sanya Yazhou Bay Science and Technology City (SKJC-2023-02-001 and SCKJ-JYRC-2023-47); the Hainan Province Science and Technology Special Fund (ZDYF2022XDNY260, ZDYF2024KJTPY001); the Nanfan Special Project, CAAS (YBXM2403, YYLH2501, and YBXM2556); the National Natural Science Foundation of China (32261143465, 32350710198); the Hainan Province Science and Technology Innovation Project (KJRC2023A01); and the Hainan Province International Scientific and Technological Cooperation Talent and Exchange Project (Foreign Expert Program) Plan (G20241024007E).

References

Akimoto, M., Shimamoto,Y., Morishima, H., 1999. The extinction of genetic resources of Asian wild rice, *Oryza rufipogon* Griff.: a case study in Thailand. Genet. Resour. Crop Evol. 46, 419-425.

Assunção, R.M., Neves, M.C., Câmara, G., da Costa Freitas, C., 2006. Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. Int. J. Geogr. Inf. Sci. 20 (7), 797-811.

Basille, M., Calenge, C., Marboutin, É., Andersen, R., Gaillard, J.M., 2008. Assessing habitat selection using multivariate statistics: some refinements of the ecological-niche factor analysis. Ecol. Modell. 211 (1), 233-240.

Chang, J., Tian, L., Leite, M.F.A., Sun, Y., Shi, S., Xu, S., Wang, J., Chen, H., Chen, D., Zhang, J., Tian, C., Kuramae, E.E., 2022. Nitrogen, manganese, iron, and carbon resource acquisition are potential functions of the wild rice *Oryza rufipogon* core rhizomicrobiome. Microbiome. 10(1): 196.

Chen, H., Dong, S., He, Z., Chen, Y., Tian, D., Liu, L., Wang, Y., Zhang, W., Li, L., Yan, J., Song, Z., 2022. Effects of land use change on population survival of three wild rice species in China since 2001. Front. Plant Sci. 13, 951903.

Cheng, S., Hu, P., 2008. Development Strategy of Rice Science and Technology in China. Chin. J. Rice. Sci. (03), 223-226.

Chu, L., Tian, X., Wang, W., 2016. Investigation on Wheat Germplasm Resources in Northern

Henan. J. Anhui Agric. Sci. 27, 010.

Dai, L., Wu, L., Wang, L., Yang, Q., Tang, C., Yu, T., 2004. Analysis on the current status of wild rice resource distributed in Yunnan province based on the investigation. Chin. J. Rice Sci. 24, 107-128.

Du, Y., Brown, J.R., Sniderman, J.M.K., 2024. Last Glacial Maximum climate and atmospheric circulation over the Australian region from climate models. Clim. Past. 20(2), 393-413.

Dutton, A., Lambeck, K., 2012. Ice volume and sea level during the last interglacial. Science. 337 (6091): 216-219.

Fan, Z., Pan, D., Chen, Y., Chen, J., Li, C., Sun, B., Zhou, H., Chen, W., Liu, W., 2017. Conservation Suggestions on *Oryza rufipogon* in Guangdong Province Based on Investigation and Collection. J. Plant Genet. Resour. 18 (02), 372-379.

Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37 (12), 4302-4315.

Fu, L., 1992. China Plants Red Date Book. Beijing Science and Technology Press, Beijing, pp. 314-316.

Gao, A., Yang, Q., 2022. Theory and Methods for Survey and Collection of Crop Germplasm Resources. J. Plant Genet. Resour. 23 (01), 21-28.

Gao, L., Zhang, S., Zhou, Y., Ge, S., Hong, D., 1996. A survey of the current status of wild rice in China. Chin. Biodiversity. 4 (3), 160-166.

Genesys. https://www.genesys-pgr.org/a/v2QLWpBmJKW. Accessed on 2024-7-15.

Gotway, C.A., Stroup, W.W., 1997. A generalized linear model approach to spatial data analysis and prediction. Biol. Environ. Stat. 157-178.

Guarino, L., Ramanatha, Rao, V., Reid, R., 1995. Collecting plant genetic diversity: Technical guidelines. Wallingford: CAB International. pp. 1-11.

He, Z., Zhang, P., Jia, H., Zhang, S., Nishawy, E., Sun, X., Dai, M., 2024. Regulatory Mechanisms and Breeding Strategies for Crop Drought Resistance. New Crops.100029.

Heaney, L.R., 1991. A synopsis of climatic and vegetational change in Southeast-Asia. Clim. Change. 19(1), 53-61.

Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. Nature. 405, 907-913.

Huang, J., Yang, Q., Chen, C., Liang, S., Zhang, W., Qiao, W., Wang, J., 2009. Genetic Diversity and the Geographical Characteristics of Wild Rice (*Oryza rufipogon* Griff) in Guangxi. Sci. Agric. Sin. 42 (08), 2633-2642.

Huang, P., Schaal, B.A., 2012. Association between the geographic distribution during the last glacial maximum of Asian wild rice, *Oryza rufipogon* (Poaceae), and its current genetic variation. Am. J. Bot. 99, 1866-1874.

Jaynes, E.T., 1957. Information theory and statistical mechanics. Phys. Rev. 106(4): 620.

Kariali, E., Kuanar, S.R., Mohapatra, P.K., 2008. Individual Tiller Dynamics of Two Wild *Oryza* species in Contrasting Habitats. Plant. Prod. Sci. 11(3), 355-360.

Lewis, S.E., Sloss, C.R., Murray-Wallace, C.V., Woodroffe, C.D., Smithers, S.G., 2013. Post-glacial sea-level changes around the Australian margin: a review. Quat. Sci. Rev. 74, 115-138.

Li, A., Wang, J., Wang, R., Yang, H., Yang, W., Yang, C., Jin, Z., 2020. MaxEnt modeling to predict current and future distributions of *Batocera lineolata* (Coleoptera: Cerambycidae) under climate change in China. Ecoscience. 27(1), 23-31.

Lin, S., Yuan, L., 1980. Hybrid Rice Breeding in China, Innovative Approaches to Rice Breeding. International Rice Research Institute, Manila, Philippines, 35-51.

Lin, Y., Chen, Y., Wang, H., Hong, J., Yang, R., Cao, Y., Yan, S., Fang, W., 2024. Global potential distributions and conservation status of rice wild relatives. Plants, People, Planet. 6 (4), 979-991.

Liu, R., Zheng, X., Zhou, L., Ge, S., 2015. Population genetic structure of *Oryza rufipogon* and *Oryza nivara*: implications for the origin of *O. nivara*. Mol. Ecol. 24 (20): 5211-5228.

Liu, X., Li, Y., Li, L., Jia, J., 2023. Theoretical Framework and Development Strategy for The Science of Crop Germplasm Resources. J. Plant Genet. Resour. 24 (01):1-10.

Mohapatra, P.K., Panda, B.B., Kariali, E., 2011. Plasticity of tiller dynamics in wild rice *Oryza rufipogon* Griff.: A strategy for resilience in suboptimal environments. Int. J. Agron., 2011(1), 543237.

Neelam, K., Thakur, S., Neha., Yadav, I.S., Kumar, K., Dhaliwal, S.S., Singh, K., 2017. Novel alleles of phosphorus-starvation tolerance 1 gene (*PSTOL1*) from *Oryza rufipogon* confers high phosphorus uptake efficiency. Front. Plant Sci. 8, 509.

Phillips, S.J., Anderson, R.P., Schapire, R. E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Modell. 190, 231-259.

Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography. 31 (2): 161-175.

Phillips, S.J., Dudík, M., Schapire, R.E., 2017. Maxent software for modeling species niches and distributions (Version 3.4. 1).

Qian, L., Huang, Z., Yang, S., Cao, W., 2021. Study on spatial conservation priority pattern of key protected plants in Xiamen. Acta. Ecol. Sin. 41 (11), 4367-4378.

Qin, A., Jin, K., Batsaikhan, M.E., Nyamjav, J., Li, G., Li, J., Xue, Y., Sun, G., Wu, L., Indree, T., Shi, Z., Xiao, W., 2020. Predicting the current and future suitable habitats of the main dietary plants of the Gobi Bear using MaxEnt modeling. Glob. Ecol. Conserv. 22, e01032.

Quiquet, A., Archibald, A.T., Friend, A.D., Chappellaz, J., Levine, J.G., Stone, E.J., Telford, P.J., Pyle, J.A., 2015. The relative importance of methane sources and sinks over the Last Interglacial period and into the last glaciation. Quat. Sci. Rev. 112, 1-16.

Redlands, C. E. S. R. I., 2011. ArcGIS Desktop: Release 10.7.

Schaefer, J.M., Putnam, A.E., Denton,G.H., Kaplan, M.R., Birkel,R., Doughty, A.M., Kelley, S., Barrell, D.J.A., Finkel, R.C., Winckler, G., Anderson, R.F., Ninneman, U.S., Barker, S., Schwartz, R., Andersen, B.G., Schluechter, C., 2015. The southern glacial maximum 65,000 years ago and its unfinished termination. Quat. Sci. Rev. 114, 52-60.

Stockwell, D.R.B., Noble, I.R., 1992. Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. Math. Comput. Simul. 33 (5-6), 385-390.

Tian, L., Zhou, X., Ma, L., Xu, S., Nasir, F., Tian, C., 2017. Root-associated bacterial diversities of *Oryza rufipogon* and *Oryza sativa* and their influencing environmental factors. Arch. Microbiol. 199, 563-571.

Wang, K., Li, X., 2012. Fundamental Strategies and Methods for Collection of Wild Soybean Germplasm Resources in China. J. Plant Genet. Resour. 03, 009.

Wang, X., Yang, Q., Yun, Y., Meng, W., 2010. Studies on the Sampling Strategy for Ex-Situ Conservation of *Oryza rufipogon* Griff. Chin. Agric. Sci. Bull. 26 (07), 303-306.

Wang, Y., Xie, B., Wan, F., Xiao, Q., Dai, L., 2007. Application of ROC curve analysis in evaluating the performance of alien species' potential distribution models. Biodivers. Sci. (04), 365-372.

Williams, A.N., Ulm, S., Sapienza,T., Lewis, S., Turney, C.S.M., 2018. Sea-level change and demography during the last glacial termination and early Holocene across the Australian continent. Quat. Sci. Rev. 182, 144-154.

Wu, Z., Zhan, N., Shang, X., Liu, G., 2020. Climatic Characteristics Analysis of *Hibiscus tiliaceus* and Prediction of Its Suitable Range in China. Eucalypt. Sci. Technol. 37 (02), 45-52.

Xu, D., Lu, H., Wu, N., Liu, Z., 2010. 30 000-Year vegetation and climate change around the East China Sea shelf inferred from a high-resolution pollen record. Quat. Int. 227(1), 53-60.

Xu, D., Zhuo, Z., Wang, R., Ye, M., Pu, B., 2019. Modeling the distribution of *Zanthoxylum armatum* in China with MaxEnt modeling. Glob. Ecol. Conserv. 19, e00691.

Yan, H., Feng, L., Zhao, Y., Feng, L., Wu, D., Zhu, C., 2020. Prediction of the spatial distribution of *Alternanthera philoxeroides* in China based on ArcGIS and MaxEnt. Glob. Ecol. Conserv. 21, e00856.

Yang, Q., Huang, J., 2013. Research Progress on Genetic Diversity of *Oryza rufipogon* in China. Acta. Agron. Sin. 39 (04), 580-588.

Yun, Y., Tang, Q., Yan, X., Meng, W., Wang, X., Lin, Y., 2015. Field Survey and Conservation of Wild Rice Resources in Hainan Province. J. Plant Genet. Resour. 16 (04), 715-719.

Zeng, Y., Xie, H., Cheng, W., Jiang, Y., Zhou, J., Xie, X., Tan, X., Zhou, H., Qin, L., 2019. Systematic Field Survey and Collection of Maize Germplasm Resources in Guangxi. J. Plant Genet. Resour. 20(03),654-661.

Zhang, H., Song, J., Zhao, H., Li, M., Han, W., 2021. Predicting the distribution of the invasive species *Leptocybe invasa*: Combining MaxEnt and geodetector models. Insects. 12(2), 92.

Zhang, W., Yang, Q., 2003. Collecting, Evaluation and Conservation of Wild Rice Resources in China. J. Plant Genet. Resour. (04), 369-373.

Zheng, X., Chen, B., Song, Y., Li, F., Wang, J., Qiao, W., Zhang, L., Cheng, Y., Sun, Y., Yang, Q., 2019. *In-situ* Conservation of Wild Relatives of Crops. J. Plant Genet. Resour. 20(05):1103-1109.

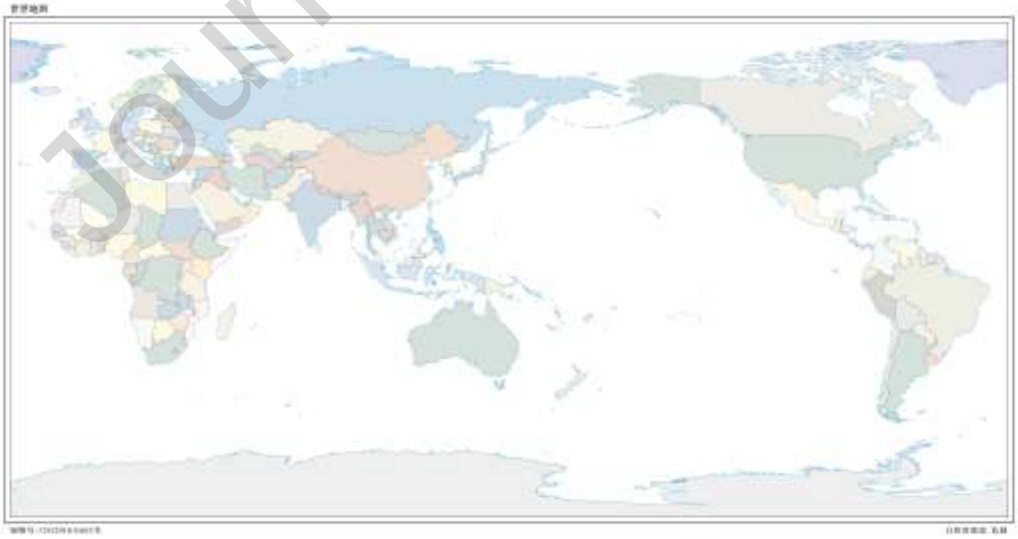
Zheng, X., Peng, Y., Qiao, J., Henry, R., Qian, Q., 2024. Wild rice: unlocking the future of rice breeding. Plant Biotechnol J. 22(11): 3218-3226.

Zheng, X., Zhou, H., Chen, W., Zhou, L., Ren, N., Liu, R., Meng, Q., Geng, M., Liu, K., Li, D., Lu, R., Ma, X., Qiao, W., Zhu, Y., Lu, B., Yang, Q., Ge, S., 2020. Collection and Preliminary Study of Wild Rice Genetic Resources in Southeast and South Asia. J. Plant Genet. Resour. 21(06):1503-1511+1520.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:



Graphical abstract