

# Cluster Analysis and Evaluating the Genetic Diversity of Different Upland Rice Genotypes Based on their Quantitative Traits

Gedifew Gebrie\* and Desta Abebe

Department of Agriculture, Ethiopian Institute of Agricultural Research Center, Pawe, Ethiopia

\*Corresponding author: Gedifew Gebrie, Department of Agriculture, Ethiopian Institute of Agricultural Research Center, Pawe, Ethiopia, Tel: 0946454437; E-mail: gebriegedifew1976@gmail.com

**Received date:** June 22, 2022, Manuscript No. IPPBA-22-14168; **Editor assigned date:** June 27, 2022, PreQC No. IPPBA-22-14168 (PQ); **Reviewed date:** July 11, 2022, QC No. IPPBA-22-14168; **Revised date:** October 13, 2022, Manuscript No. IPPBA-22-14168 (R); **Published date:** October 20, 2022, DOI: 10.36648/IPPBA.6.5.001

**Citation:** Gebrie G, Abebe D (2022) Cluster Analysis and Evaluating the Genetic Diversity of Different Upland Rice Genotypes Based on their Quantitative Traits. J Plant Bio Agric Sci Vol:6 No:5

## Abstract

Determining the level of germplasm diversity and genetic relationships among breeding materials is an invaluable aid in crop improvement strategies with an understanding that genetic variability is the base for crop improvement providing opportunity for plant breeders to develop new and improved cultivars with desirable characteristics and it is a key to reliable and sustainable production of the food crops through breeding. It has been also confirmed that measuring the available diversity of crops is important for effective evaluation and utilization of germplasms to explore their variability so as to identify desirable agronomic attributes. For eradicating the problem of national rice production, the national rice breeding and genetics research program of the country is introducing and evaluating different rice germplasms for their environmental adaptability and agronomic performance with increasing the crops' genetic diversity. With the same approach, 100 upland rice genotypes were introduced and evaluated using Augmented RCBD experimental design so as to assess and determine the extent and pattern of their genetic and phenotypic variability using cluster analysis bringing them into similarity groups based on important agronomic traits. Each genotype was planted on a plot area of 2.5 m<sup>2</sup> involving 4 rows per plot with 0.25 m spacing between each row. The seeds were drilled in rows with a seed rate of 60 kg ha<sup>-1</sup>. Nitrogen Phosphorus Sulfur (NPS) and Urea fertilizers were applied in the amount of 124 kg ha<sup>-1</sup> and 100 kg ha<sup>-1</sup> respectively. The quantitative traits such as days to 50% heading, days to 85% maturity, plant height, panicle length, number of filled grains per a panicle and number of unfilled grains per a panicle, grain yield and 1000 seed weight were measured and subjected to clustering analysis using XLSTAT 5.03 statistical software so as to determine the extent and pattern of the genetic and phenotypic variability of 97 upland rice genotypes. During clustering analysis, the genotypes were grouped into five clusters with different Euclidian distances confirming the presence of genetic and phenotypic variability among the evaluated upland rice genotypes. The genotype with the highest grain (6298 kg ha<sup>-1</sup>) yield was obtained under cluster-III.

**Keywords:** Cluster analysis; Upland rice genotypes; Variability

## Introduction

The national rice breeding and genetics research program of Ethiopian Institute of Agricultural Research (EIAR) is continually introducing and evaluating different rice germplasms targeting for their environmental adaptability and agronomic performance (high grain yield, tolerant for biotic and abiotic stresses and for other quality traits) aiming at enhancing the rice national productivity with increasing its annual production and area coverage [1].

Determining the extent and degree of germplasm diversity and genetic relationships among breeding materials is a vital aid in crop improvement strategies with an understanding that genetic variability is the base for crop improvement providing the opportunity for plant breeders to develop new and improved cultivars with desirable characteristics and it is a key to reliable and sustainable production of the food crops through breeding [2]. It has been also confirmed that measuring the available diversity of crops is important for effective evaluation and utilization of germplasms to explore their variability so as to identify desirable agronomic attributes.

Bringing improvements over the existing crop varieties is a continuous process in plant breeding there identifying diverse parents having high genetic variability for combining desirable characters and similarly the effectiveness of any rice improvement program depends on the utilization of different germplasm from around the world. Therefore, knowledge of sound genetic diversity is important for implementing any breeding program targeting crop improvement with exploring the variability present in their germplasm for identification of desirable agronomic attributes. Similarly, to bring improvements over the rice varieties, introduction and evaluation of different rice germplasms for different ecosystems to increase the national genetic diversity of the crop is one of the usual researches works of Ethiopian institute of agricultural research and a total of 3,336 germplasms for rainfed upland, rainfed lowland and irrigated rice ecosystems have been introduced

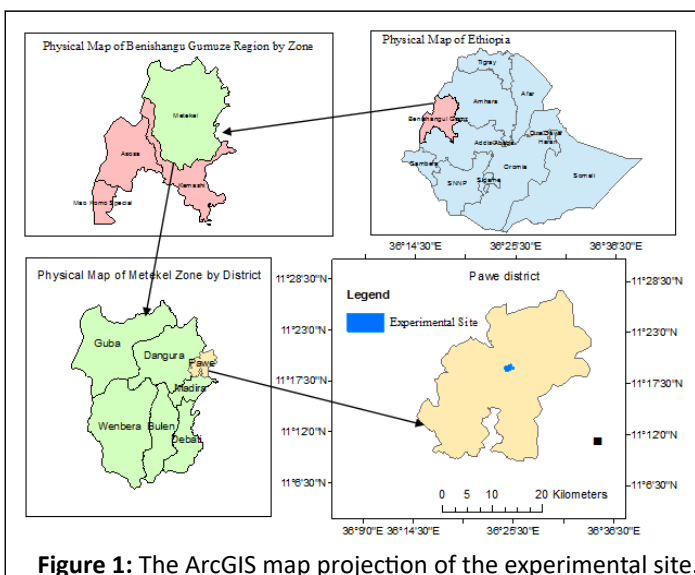
with the major objectives of high yielding, medium maturing, high biomass, white caryopsis, abiotic stress tolerance (cold and salinity), biotic resistance (blast, brown spot and sheath rot) rice varieties where the introduced germplasms pass through a serious of evaluation stages for variety release and a total of almost 39 improved rice varieties has been released.

Still, the national demand for improved rice technologies is increasing from time to time that is why the national rice research program of the country collected and introduced a set of rice germplasm from different external sources with the objective of characterizing and evaluating their environmental adaptability in lined with their genetic diversity and agronomic potential so as to identifying the promising rice genotypes for further rice improvement research program. With the same trend, 100 upland rice genotypes have been introduced to be involved in the rice germplasm enhancement research sub component of the national rice research program of the country for being evaluated under the observation nursery trial at Pawe agricultural research center to assess their environmental adaptability and agronomic performances. In addition, this activity was also designed to determine the extent and pattern of the genetic and phenotypic variability of the upland rice genotypes with using cluster analysis bringing them into similarity groups based on important agronomic traits.

## Materials and Methods

### Description of the experimental site

The trial was implemented at Pawe agricultural research center for one year during 2019/20 main cropping season under rain-fed condition. The experimental area is located in the north western regions of Ethiopia at about 575 km north west of Addis Ababa at a latitude of 11°19'N and longitude of 36°24'E and at an altitude of 1120 mas which is characterized by hot to warm moist conditions with mean minimum and maximum temperature of 16°C and 32°C, respectively and has an average annual rainfall of 1587 mm for five to seven months duration (Figure 1).



**Figure 1:** The ArcGIS map projection of the experimental site.

### Plant materials

Hundred upland rice genotypes with three standard checks were used. The genotypes were collected and introduced from China and Japan in collaboration with IRRI and Africa-Rice research center [3].

### Experimental design and procedures

An Augmented-RCBD experimental design was used with a plot area of 2.5 m<sup>2</sup> with 4 rows per plot, 0.5 m and 1 m spacing between plots and blocks respectively, an inter-row spacing of 0.25 m and a seed rate of 60 kg ha<sup>-1</sup> was used. Inorganic fertilizers of NPS (Nitrogen-phosphorus-sulfur containing fertilizer) and Urea with an amount of 124 kg ha<sup>-1</sup> and 100 kg ha<sup>-1</sup> respectively were applied. The total NPS was applied at planting, whereas the urea was used with three splits (1/3 at planting, 1/3 at tillering after weeding and 1/3 at the crop's panicle initiation stage).

### Data collection and statistical analysis

A plot and plant based quantitative characteristics of rice were collected following the appropriate growth stage of the rice crop from each experimental unit. Six newly tested genotypes failed to germinate and hence no data collection was done from the plots where those missed genotypes were planted. Thus, the quantitative data were collected from 97 upland rice genotypes. The traits such as Days to 50% Heading (DH), Days to 85% Maturity (DM), Plant Height in cm (PH), Panicle Length in cm (PL), Number of Filled Grains per a central main panicle (NFG), Number of Unfilled Grains per a central main panicle (NUFG), grain yield in kg ha<sup>-1</sup> and 1000 Seed Weight in gm (TSW) were considered and subjected to average linkage hierarchical clustering analysis using XLSTAT 5.03 statistical software.

Clustering analysis is very important to see through complex relationships among populations of diverse origins in a more simplified manner. Thus, the Average linkage Hierarchical Clustering (AHC) analysis was computed to determine the extent and pattern of genetic and phenotypic diversity among the tested upland rice genotypes based on their corresponding traits [4]. The closeness between two upland rice genotypes was measured by calculating at what point they are dissimilar using Euclidian distance (Ward's method) while grouping 97 upland rice genotypes.

## Results and Discussion

### The descriptive statistical estimation of the quantitative traits

From the descriptive statistical analysis result of the measured quantitative traits of the upland rice genotypes (Table 1), the average value of 131 days for 85% Days to Maturity (DM) was obtained indicating that there is a probability of getting a medium maturing upland rice genotypes which is compatible with the rainfed upland rice cultivation season of the testing location and other agroecologically similar areas of Ethiopia.

As it has been noted by IRRI, standards of classifying the rice genotypes based on their time of physiological maturity, the genotypes could be classified as extra early maturing (up to 95 days), very early maturing (96-110 days), early maturing (111-125 days), medium maturing (126-140 days) and late maturing (141 days and above) [5]. Thus, 16 of the tested 97 upland rice genotypes could be classified as an early maturing

genotype. 73 of the upland rice genotypes with days to maturity ranging from 126 to 140 days to 85% were identified as a medium maturing and 8 of the genotypes found to be late maturing with a number of 85% days to maturity ranging from 143 to 158 numbers of days to 85% maturity.

**Table 1:** The descriptive statistical analysis result of 8 quantitative traits collected from 97 upland rice genotypes tested at the research station of Pawe agricultural research center.

S/N	Agronomic traits	Min. value	Max. value	Mean value	Std. dev
1	DH	79	118	95	9.93
2	DM	117	158	131	6.75
3	PH	55.8	114.4	85.64	13.68
4	PL	16.2	23	20.96	1.94
5	NFG	67	172	105	20.73
6	NUFG	1	27	6	3.39
7	TSW	18.38	36.1	26.13	4.09
8	GY	1493.48	6298	3950.48	231.96

**Where:** DH: Days to 50% Heading; DM: Days to 85% Maturity PH: Plant Height in cm; PL: a Panicle Length of the main central tiller of a sampled upland rice plant in cm; NFG: Number of Filled Grains per a central main panicle of the sampled upland rice plant; NUGF: Number of Unfilled Grains per a central main panicle of the sampled upland rice plant; GY: Grain Yield in kg per hectare and TSW: 1000 Seed Weight in gram

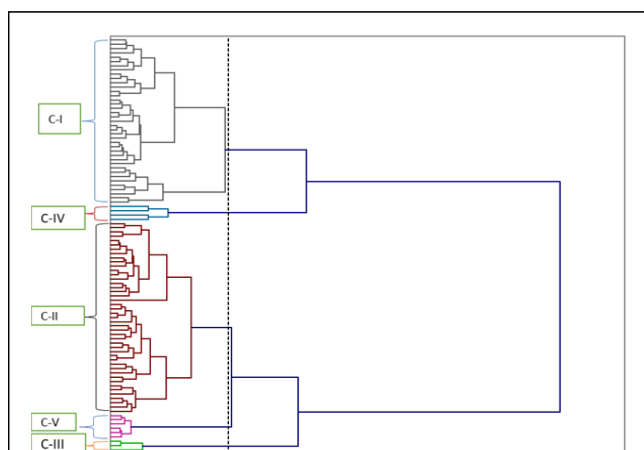
While considering their Plant Height (PH), the genotypes obtained values ranging from 55.80 to 114.40 cm with an average Panicle Length (PL) of 20.96 cm. For Number of Filled Grain per a panicle (NFG), the average value of 105 grains (ranged from 67 to 172 grains) was obtained from which 29 upland rice genotypes scored more than the average value of NFG scored by the standard check called 'Adet-1' which scored a better value of 113 NFG ensuring the higher probability of getting the upland rice genotypes with a better expected grain yield. It was also found that almost 61% of the upland rice genotypes had in determinant number of unfilled grains per a panicle which was less than or equal to the average value of NUGF (6 grains per a panicle). These two results for number of filled and unfilled grains per a panicle indicated that these genotypes had a higher predicted yield potential when compared to the best performing genotypes among the checks directing that they will have a greater chance to be subjected to further rice improvement program focusing on grain yield as the trait of interest.

The 1000 Seed Weight (TSW) of the tested upland rice genotypes were ranged from 18.38 g to 36.10 g with an average value of 26.13 g where 15 of the genotypes scored a better TSW when compared to a standard check called 'Nerica-4' with the

higher TSW value of 30.27 g [6]. A high TSW will increase germination percentage, seedling emergence, tillering efficiency, plant population density and grain yield. Therefore, the higher TSW for 15 upland rice genotypes is an indication for the presence of most promising rice genotypes with a better genetic advantage over the best performing standard checks for their population stand, seedling vigor, tillering efficiency and yield potential.

### Clustering the upland rice genotypes using eight quantitative traits

From the clustering result, the upland rice genotypes were grouped into five main clusters (Figure 2) approving that there was a genetic variation among the tested rice genotypes with 86.39% of between cluster variance decomposition and 13.61% of a variance decomposition with in each cluster (cluster).



**Figure 2:** Dendrogram showing the clustering patterns of 97 upland rice genotypes.

**Where:** C-stands for Cluster

Among the reported clusters, cluster II contained the highest number of genotypes (45 genotypes) and cluster I contained the second large number of genotypes (36 genotypes) while, cluster III, cluster IV, cluster V contained 3, 4 and 6 genotypes, respectively (Table 2).

**Table 2:** Variance decomposition and extent of within and between-cluster variance

Variance decomposition	Percent of variance
Within-cluster	13.61%
Between-clusters	86.39%

The cluster mean value of each measured trait in each cluster was computed for the tested upland rice genotypes (Table 3). From the analysis result, an average cluster mean value of 85% Days to Maturity (DM) ranging from 129 to 134 days was obtained for genotypes grouped under cluster III and V respectively while clustering the 97 upland rice genotypes [7].

This indicated that there is a greater chance of obtaining medium maturing genotypes which were grouped under all the five clusters.

**Table 3:** List of the upland rice genotypes grouped under each cluster.

Cluster	Number of genotypes	List of genotypes grouped under the assigned cluster number
I	39	G-1, G-2, G-8, G-14, G-17, G-18, G-19, G-20, G-22, G-24, G-25, G-26, G-28, G-31, G-32, G-38, G-40, G-41, G-42, G-44, G-55, G-56, G-57, G-58, G-59, G-60, G-61, G-65, G-67, G-69, G-78, G-79, G-81, G-86, G-89, G-91, G-93, G-94 and G-103
II	45	G-3, G-4, G-5, G-9, G-11, G-12, G-13, G-15, G-16, G-27, G-30, G-33, G-35, G-36, G-37, G-39, G-43, G-45, G-47, G-48, G-49, G-50, G-51, G-52, G-54, G-62, G-63, G-64, G-66, G-68, G-70, G-71, G-72, G-75, G-76, G-77, G-80, G-82, G-84, G-85, G-87, G-88, G-90, G-101 and G-102
III	3	G-6, G-7 and G-10
IV	4	G-21, G-23, G-53 and G-92
V	6	G-29, G-34, G-46, G-73, G-74 and G-83

**Where:** G-stands for Genotype

The highest cluster mean values of plant height (105 cm) was recorded among the genotype grouped under cluster III and the shortest cluster mean value of the trait (83.05 cm) was obtained among the genotypes grouped under cluster IV (Table 4).

**Table 4:** The mean values (Cluster centroids) of eight quantitative traits in each cluster of 97 upland rice genotypes.

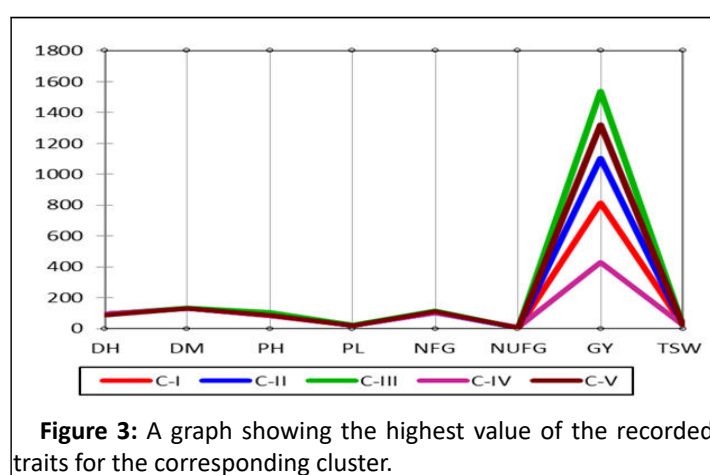
Cluster/ Traits	DH	DM	PH	PL	NFG	NUFG	TSW	GY
I	94	132	84.55	20.97	105	6	25.43	3270.32
II	96	131	85.57	20.84	103	6	26.75	4413.4
III	95	134	105	22.8	115	5	26.02	6152.92
IV	99	131	83.05	20.23	104	9	25.16	1718.04
V	90	129	85.23	21.3	113	6	26.69	5286.72

**Where:** DH: Days to 50% Heading, DM: Days to 85% Maturity, PH: Plant Height in cm; PL: Panicle Length in cm, NFG: Number of Filled Grains per a panicle; NUFG: Number of Unfilled Grains per a panicle; GY: Grain Yield in kg per hectare and TSW: 1000 seed weight in gram

Tall varieties (104 cm and taller) are suitable for flood-prone and unlevelled fields where lodging may be a problem [8]. The rice varieties with a medium height (100 to 102 cm) are not susceptible to lodging and are suitable for most areas when fertilizer is used. Short varieties (less than 100 cm in height) are best suited to level fields especially in Irrigated areas and are responsive to fertilizers. Thus, most genotypes included under all clusters except cluster III had a Plant Height (PH) less than 100 cm indicating that there is indication for obtaining short upland rice genotypes which might be best suited to level fields especially in irrigated areas and will be more responsive to fertilizers.

The highest cluster mean value of plant height (105 cm) for the upland rice genotypes grouped under III indicates the probability of obtaining a genotype which is suitable for flood-prone and unlevelled fields where lodging may be a problem. A highest cluster mean value of panicle length (21.30 cm) and number of filled grains per panicle (113 grains per panicle) with a higher cluster mean value for a grain yield (5286.72 kg $ha^{-1}$ ) was obtained for cluster V directing that those genotype with a better expected grain yield potential could be obtained from those tested upland rice genotypes clustered under this cluster.

In rice, yield could be determined by indirect traits like panicle length and grains per panicle as well as direct traits like filled grains per panicle and 1000 grain weight. Thus, the above results of the higher cluster mean values of Panicle Length (PL) and Number of Filled Grains per panicle (NFG) for genotypes grouped under cluster V indicated that the respective rice genotypes could have a better expected grain yield potential when compared to those genotypes in other clusters though the highest cluster mean value of grain yield was recorded by those genotypes grouped under cluster III (6152.92 kg $ha^{-1}$ ). Therefore, a better attention could be given for those genotypes included under cluster V and cluster III while considering enhanced grain yield as the primary breeding objective of the newly designed rice breeding activity [9]. The upland rice genotype with the highest grain yield value was grouped under cluster III (Figure 3)

**Figure 3:** A graph showing the highest value of the recorded traits for the corresponding cluster.

**Where:** DH: Days to 50% Heading; DM: Days to 85% Maturity; PH: Plant Height in cm; PL: Panicle Length in cm; NFG: Number of Filled Grains per a panicle; NUFG: Number of Unfilled Grains per a panicle; GY: Grain Yield in kg per hectare and TSW: 1000 Seed Weight in gram.

**Note:** The unit of GY (grain yield) in this case is in gram per plot (the grain yield harvested from a plot size of 2.5 m<sup>2</sup>). This has been done to make the graph well-structured and visible.

While computing the inter-cluster distances (the distance between cluster centroid of the quantitative trait dissimilarity) using Euclidean distance calculation, the maximum inter-cluster distance was obtained between cluster III and IV (1109.03) while clustering the tested upland rice genotypes (Table 5) [10]. Inter cluster distance is the main criterion for selection of the genotypes implying that the genotypes belonging to the cluster with maximum inter cluster distance are genetically more divergent suggesting that selection of genotypes should be based upon large inter cluster distances, which may lead to broad spectrum of favorable genetic variability. Therefore, it was confirmed that selection of genotypes for the improvement of traits of interest among the eight quantitative traits of the tested upland rice genotypes would be focused on those rice genotypes grouped under cluster III and IV.



**Table 5:** Inter-cluster distances (Euclidean distances) and intra-cluster distances for the clustering pattern of 97 upland rice genotypes and their variances with in a cluster.

Clusters	I	II	III	IV	V
I	0	285.79	721.03	388.12	504.19
II		0	435.52	673.86	218.67
III			0	1109.03	217.6
IV				0	892.28
V					0
Within-class variance	9809.18	6668.28	1136.66	4284.55	482.26
Average intra-cluster distances	86.85	74.68	25.54	54.93	19.36

The higher average intra cluster distance was observed for cluster I, II, and IV intra-cluster distances (86.85, 74.68 and 54.93 respectively) indicating that there was a high genetic diversity among the tested genotypes grouped with in these clusters. These clusters were also identified with a higher with in cluster variance of 9809.18, 6668.28 and 4284.55 for cluster I, II, and IV respectively confirming that there is a higher genetic diversity among the upland rice genotypes of the respective clusters. The highest inter cluster distance was observed between cluster. The lowest inter clusters distance was found between cluster III and V (217.60) and the highest distance was found between cluster III and IV (1109.03).

The genotypes grouped into the same cluster, as well as, the lowest values of inter cluster distance indicating a close relationship between them and exhibited the lowest divergence's degree from one another giving a hint that no need of making hybridization between genotypes belonging to the same cluster or between genotypes from closely related clusters because no transgressed segregants are predicted from these combinations [11]. In addition, the highest values of inter cluster distance suggests the maximum variability among the clusters or between different genotypes from two different clusters with the highest inter cluster distances from which hybridization programs could be designed in such a way that the parental genotypes relating to different clusters with high amount of divergence could be used to get appropriate transgressive segregants. Thus, hybridization program could be designed by considering genotypes from cluster III and IV or from cluster I.

## Conclusion

The clustering analysis where 97 upland rice genotypes group into five different clusters directs the presence of genetic diversity among the tested upland rice genotypes from which cluster-I has the highest intra-cluster distance (86.85) confirming the presence of a highest genetic diversity among the genotypes included in this cluster. The highest value of inter cluster

distance between cluster III and IV also suggests that there is the maximum variability between the upland rice genotypes grouped under these cluster. Thus, hybridization or other rice improvement programs could consider the genotypes from cluster III and IV or genotypes grouped under cluster I so as to get appropriate transgressive segregants of the traits of interest.

The obtained cluster mean values of 85% days to maturity ranging from 129 for cluster-V to 134 for cluster-III indicates that there is a greater chance of obtaining medium maturing upland rice genotypes clustered under all the five clusters. The finding of upland rice genotypes with higher number of filled grains and lower number of unfilled grains per a central panicle, higher grain yield potential and with the higher 1000 seed weight values when compared to the best performing standard checks indicates that there is a greater chance of getting genotypes with a better genetic advantage over the best performing standard checks used during the experiment for their population stand, seedling vigor, tillering efficiency and yield potential. Thus, an attention could be given for those genotypes with a greater value of these quantitative traits while selecting genotypes to be improved for the desired traits of the genotypes to be implemented in the next breeding program and a value will be added on the rice research and development program of Ethiopia.

## Acknowledgement

Gratitude will be given for the rice breeding researchers and technical staffs working at Pawe agricultural research center and Fogera national rice research and training center for their help in collecting the phenotypic data and managing the overall activities of this trial. My gratitude will also go to Ethiopian institute of agricultural research for its full financial support.

## References

1. Abd El-Hadi AH, El-Maksoud MM, El-Abd AB, El-Leithy SA (2017) Genetic Variability Study of some Rice Genotypes (*Oryza sativa* L.) for Morphological, Yield and Grain Quality Traits. J Agric Chem Biotech 8:131–138
2. Abebaw D (2020) Rice Breeding Achievements, Potential and Challenges in Ethiopia. Int J Res Stud Agric Sci (IJRSAS) 6:35-42
3. Bhattarai U, Subudhi PK (2019) Genetic Diversity, Population Structure, and Marker-Trait Association for Drought Tolerance in US Rice Germplasm. Plants 8:530
4. Gedifew G, Tsige G (2019) Morphological Characterization and Evaluation of Sorghum (*Sorghum bicolor* (L.) Moench) Landraces in Benishangul Gumuz, North-western Ethiopia. Greener J Agric Sci 9:37-56
5. Govindaraj M, Vetriventhan M, Srinivasan M (2014) Importance of Genetic Diversity Assessment in Crop Plants and Its Recent Advances: An Overview of Its Analytical Perspectives. Genet Res Int 1-14
6. Moshatati A, Gharineh M (2012) Effect of grain weight on germination and seed vigor of wheat. Int J Agric Crop Sci 4:458-460
7. Ronghua L, Meijuan L, Umair A, Shiwei L, Jiaen Z (2019) Exploring the Relationships Between Yield and Yield-Related Traits for Rice Varieties Released in China From 1978 to 2017. Front Plant Sci 10:1-12
8. Shaibu AA, Uguru MI (2017) Application of Multivariate Analysis to Access Selected Rice Germplasm Phenotypic Diversity. J Exp Agric Int 16:1-8
9. Singh RK, Dubey BK, Gupta RP (2013) Intra and Inter Cluster Studies for Quantitative Traits in Garlic (*Allium sativum* L.). SAARC J Agri 11:61-67
10. Soe I, Alex T, Maxwell A, Daniel DA, Nyadanu, Richard (2019) Genetic diversity analyses of rice germplasm using morphological traits. J Plant Breed Crop Sci 11:128-136
11. Swarup S, Cargill EJ, Crosby K, Flagel L, Kniskern J, et al. (2021) Genetic diversity is indispensable for plant breeding to improvecrops. Crop Science 839–852