Principal Component Analysis for Yield and Yield Attributed Traits in Lowland Rice (*Oryza sativa* L.) Genotypes

Dejen Bekis^{*}, Hussein Mohammed, Bayuh Belay

Department of Agriculture, Hawassa University and Ethiopian Institute of Agricultural Research, Hawassa, Ethiopia

Abstract

Principal component analysis was utilized to determine the variation and to estimate the relative contribution of various characters for total variability. The experiment was laid out using randomized block design with three replications during 2017/2018 main cropping season at Fogera, Ethiopia. The first four principal components axes accounted for 85.3% cumulative variance of the total variability for seventeen agronomic characters. PC1, PC2, PC3 and PC4 explained 44.15%, 19.31%, 14.913% and 6.97% of variation from the total variation, respectively. Thus, maximum variation was found in first PC; therefore, selection for characters under PC1 would be desirable. The variability in PC1 was accounted by flag leaf length, panicle length, days to heading and days to 50% flowering, while PC2 was accounted by harvest index. For future breeding program that employ hybridization, parental material selection should be carried out considering principal components influence to breeders' interest.

Keywords: Genetic variability; Genotype; Principal component analysis; Rice

Introduction

Rice belongs to the family Poaceae and genus *Oryza* (Wang et al). The genus *Oryza* is known to consist of two cultivated species i.e.; Asian rice (*O. sativa*, 2n=24=AA) and African rice (*O. glaberrima*, 2n=24=AA) and 22 wild species (2n=24, 48) [1]. The river valleys of Yangtze, Mekon River area in China could be the primary center of origin of *Oryza sativa* [2]. *Oryza glaberrima* is indigenous to the upper valley of the Niger River and it is cultivated only in western tropical Africa [3]. Cultivated rice (*O. sativa*) is predominantly self-pollinating and has lower out crossing ability than *O.rufipogon*. According to Messeguer et al., the cross-pollination rates of *O. sativa* are less than1% [4]. However, the estimated out crossing rates among wild rice populations ranges from 4.3% to 55.9% [5]. Rice is a highly diverse crop species with wide geographic dispersal from sea level up to 3000 m.a.s.l. in both temperate and tropical climate [5,6].

Rice is the second most-produced cereal in the world after wheat and represents a staple food source for more than half of the world's population [7]. Most of the world's rice is cultivated and consumed in Asia [8]. Although Asia is the main place of rice cultivation, rice is also produced in other continents like Latin America, Europe, USA and some parts of Africa [9]. From the total production, Asia accounts the largest production totaling to about 144.25 million tons whereas Africa produces approximately 11.58 million tons. China, India and Indonesia are the largest rice producing countries in the world, with percentage share of 32.9 %, 24.4% and 11.0 %, respectively. Worldwide the area covered by rice exceeds 163.1 million ha with a production of 748 million tons. The world's average productivity (kg/ha) has doubled during the last 25 years, largely due to the use of improved technologies such as high yielding varieties [10]. In 2014, average yields of rice for countries with high productivity were 6.69, 6.75, 9.52, 5.75, 5.13 and 8.48 tons ha-1 for Japan, China, Egypt, Vietnam, Indonesia and USA, respectively.

Rice was introduced in to Ethiopia during the 1970s and fast distribution of the crop within the country has been achieved [11]. Presence of potential land under irrigation (3.7 million ha) and rain fed (25 million ha), existence of diverse ecosystems such as the uplands, rain fed low lands and flash flood prone areas, long shelf life and acceptance of rice amongst rural population due to the possibility of using rice to a range of traditional food recipes, relatively higher productivity as compared to other main staple crops and the by-products from rice such as straws and husks that shall be fed to livestock and/or used as alternate source, are the main attracting factors for rapid increase in rice production in the country [12].

In Ethiopia, rice covered about 48,418.09 hectares of land and 1,360,007.26 tons of grain was produced per annum in 2016/17 with average productivity of 2.81 tons per hectare [13]. As the demand of rice production is increasing in alarming rate, the area of production almost doubled from 18000 ha in 2006 to 48,418.09 ha in 2016/2017 [13,14]. According to CSA (2017) the number of participant farmers increased from 115,832 in 2012 to 150,041 in 2017 cropping season and the demand for improved rice technologies is also increasing rapidly from time to time.

Rice is a nutritious cereal crop, used mainly for human consumption. It is the main source of energy and protein. In Ethiopia, rice offers a variety of uses. It is used in the preparation of local foods (injera, dabbo, genffo, kinchie, shorba) and local beverages (tella and katikalla/Areki) either alone or mixed with other cereal grains [15].

However, the average rice productivity in Ethiopia is estimated at 2.81 t ha-1 [13], which is much lower than the world's average of 4.6t ha-1. Despite the fact that rice has been recognized by Ethiopian government as "the new millennium crop of Ethiopia" to attain food security, lack of improved varieties, lack of recommended crop management, lack of pre and postharvest management coupled with biotic and abiotic stresses limit the production and productivity of the crop in the country [12,16]. Among these problems, lack of improved varieties for different agro ecologies of the country is the most serious [17]. In many countries, rice is a long established crop and cultivars have been selected that are well adapted to local conditions and the

*Corresponding author: Dejen Bekis, Department of Agriculture, Hawassa University and Ethiopian Institute of Agricultural Research, *Hawassa*, Ethiopia; E-mail: djbks21@gmail.com

Received date: February 02, 2021; Accepted date: February 17, 2021; Published date: February 24, 2021

Citation: Bekis D (2021) Principal Component Analysis for Yield and Yield Attributed Traits in Lowland Rice (*Oryza sativa* L.) Genotypes. Adv Crop Sci Tech 9:461.

Copyright: © 2021 Bekis D, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Page 2 of 5

local market. It is estimated that about 120,000 varieties of rice exist in the world. But in Ethiopia which has diverse agro-ecologies, there are no more than elven lowland rice varieties in the whole country.

Farmers of South Gondar, especially those in Libokemikem, Fogera and Dera districts, largely produce lowland rice under rain-fed condition. Due to swampy nature of the study area, crop production was limited before rice adoption. Fogera and surrounding districts are swampy areas which are ideal for lowland rice cultivation. However, one of the major constraints in the area is the absence of high yielding improved lowland rice varieties resistant to diseases and to terminal water deficit (terminal moisture stress). Hence, as rice is a potential crop in study area, increasing its productivity per unit area and its total production will enable farmers get encouraging returns and improves its role in achieving food self-sufficiency. To increase the productivity of rice in the country, research has been conducted mainly at Fogera National Rice Research and Training Center (FNRRTC). The center introduced a bulk of genotypes from International Rice Research Institute (IRRI) and African Rice Center (WARDA), which are sources of variability for future rice improvement in Ethiopia.

The success of plant breeding research depends on the availability of genetic variation. However, full information is lacking on the genetic variability and character association of grain yield and yield related traits available within recently introduced low land rice genotypes in the study area. Genetic improvement mainly depends on the amount of genetic variability present in the population which is a universal property of all species in nature [18]. Variability in genotypes for yield and yield component traits forms the basic factor to be considered while making selection. The character yield reflects the performance of all plant components and might be considered as the final result of many other traits. i.e. every plant contains an inherent physiological production capacity that operates on energy required for normal plant performance. Not all genotypes have the same inherent physiological capacity to yield.

The knowledge of diversity and genetic distance among groups of genotypes helps to identify parental lines for hybridization programs. Therefore, keeping in view these urgent needs, the present investigation has been undertaken to know the level of genetic divergence among twenty seven rain-fed lowland rice genotypes with three check varieties for yield and yield contributing traits through principal component analysis.

Materials and Methods

Experimental site description

The experiment was conducted in the North-Western part of Ethiopia at Fogera National Rice Research and Training Center (FNRRTC) during the rainy season (June-December) of 2017/18. FNRRTC is located in Amhara Regional state, in the North-Western part of Ethiopia, 607 km far from Addis Ababa. The experimental site is found at Woreta and located 11058' N latitude, 37° 41' E longitude and at an elevation of 1810 m above sea level. Based on ten years' average meteorological data, the annual rainfall, and mean annual minimum, maximum and average air temperatures are 1300 mm, 11.5°C, 27.9°C and 18.3°C, respectively. The soil type is black Vertisol with pH of 5.90. The main water source for rice production in the study area is rain-fall water. Irrigation water from rivers Rib and Gumara was also used in the off season for production of vegetables as the second crop after rice.

Experimental materials

Thirty genotypes consisting of 27 lowland rice genotypes introduced

in 2015 from African Rice Center (Formerly called WARDA), two released lowland varieties (Ediget and Hibir) and one locally available genotype (X-Jigna), obtained from FNRRTC, were used for this study. The majority (20) of the material had its origin in Nepal. Two genotypes each were from Madagascar and Ruawnda. Three genotypes had IRRI coding.

Experimental design

The experiment was laid out in randomized complete block design with three replications. Each plot had six rows each 4m long, with a spacing of 25 cm between rows and 15cm between plants. The plot size was $4 \times 1.5 \text{ m=}6 \text{ m}^2$. Net plot size was $4 \text{ rows} \times 4 \text{ m=}4 \text{ m}^2$. The distance between plots and replications was 0.3 m and 1 m, respectively. Three healthy and uniform sized seeds were drilled per hill on date 28 June 2017 and thinning was conducted after germination to ensure single plant per hill.

Fertilizer in the forms of N and P₂O₅ was applied at a rate of 69/23 Kg/ha, Urea and NPS, respectively. All the NPS was applied at sowing. Urea was applied as split three times, 1/3 at sowing, 1/3 at tillering and the remaining at panicle initiation stage. All other agronomic practices were applied as recommended for rice production in the study area.

Data collection

Based on the standard evaluation system developed by International Rice Research Institute (IRRI, 2002), seventeen quantitative traits were recorded at appropriate growth stage on plot and plant basis. These data were recorded from pre-tagged ten randomly sampled plants in the four central harvestable rows of each experimental unit/plot. However, yield per plot and phenological traits were taken on plot basis. These data along with their details are mentioned below:

Measurements on plot basis

Days to heading (DH): Registered as number of days from seedling emergence up to the date when the tips of the panicle first emerged from the main shoots on 50% of the plants in each plot.

Days to 50% flowering (DF): The number of days was recorded from seedling emergence up to the date of appearance of flowers at the panicle on about 50% of the plants in each plot.

Days to maturity (DM): Number of days from the date of seedling emergence to the date when 85% of the stems, leaves, and floral bracts in a plot changed to light yellow color was recorded.

Thousand-grain weight (GW) (g): Measured by randomly taking 1000 grains from harvested grain and weighed in grams by using a sensitive balance; finally adjusted at 14% moisture content.

Biomass yield (BY) (kg): The total above ground biomass yield produced from the four central harvestable rows of each plot was measured at harvest and converted it to kilograms per hectare.

Grain yield (GY) (kg): Grain yield was measured from the net middle plot area of 4 m2 by using an electronic sensitive balance and converted to kilograms per hectare at 14% moisture content by using rice moisture tester.

Harvest index (HI) (%): The ratio of grain yield per hectare adjusted to 14% moisture to sun-dried biological yield per hectare was calculated and expressed in percent.

Measurements on plant basis

Number of tillers per plant (TP): Number of both ear bearing and non-bearing tillers per plant were counted at the time of harvest.

Number of panicles per plant (PP): Number of panicles per plant was taken by actual counts of the total number of tillers bearing panicles per plant at harvest.

Culm Length (CL) (cm): Culm length was measured from ground level to the base of the central panicle neck node (panicle base node) at maturity stage.

Panicle length (PL) (cm): Actual measurement in centimeters was taken from panicle base to tip of the longest (central) panicle.

Plant height (PH) (cm): The height of plants was measured at maturity stage from the base of the main stem to the tip of the main panicle.

Flag-leaf length (FL) (cm): The flag leaf length was measured from the point of attachment of the ligule to the tip of the blade on the main stem after anthesis.

Flag-leaf width (FW) (cm): Width of flag leaf was measured on the main stem after anthesis.

Number of filled grains per Panicle (FGPP): The number of filled grains per panicle was recorded from the central panicles of sampled plants at harvest.

Number of unfilled grains per panicle (UGPP): The number of unfilled grains was determined by counting only unfilled grains from the central panicles of sampled plants at harvest.

Panicle weight (PW) (g): Panicle weight of main stem was measured after it lost its moisture content by sun drying.

Statistical Analysis

Principal component analysis: Principal Component Analysis (PCA) was computed to find out the characters, which accounted more to the total variation. Principal components analysis was performed using correlation matrix by employing Minitab computer software Released 17. The contribution of each character in PCA is determined by eigenvector that is greater than half divided by the square root of the standard deviation of the Eigen value of the respective PCA. Principal Components (PCs) with Eigen value>1.0 were used as criteria to determine the number of PCs. The bi plot analysis and its graphical output were used to understand and interpret the relationship between the rice genotypes and the observed quantitative trait.

Results and Discussion

Principal component analysis

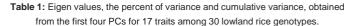
Principal Component Analysis (PCA) was used to examine the variability among the 30 low land rice genotypes. To validate the clustering (grouping) observed by the cluster analysis (Table 1 and Figure 1), an ordination analysis PCA was executed using the 17 quantitative characters. Correlation matrix generated using the genotypic mean values of all traits used as an input and were subjected to the Principal Components Analysis (PCA). From the 17 principal components (equal number to the original traits) extracted, the first four PC's with an Eigen value >1 were significant.

The first four Principal Components (PC's) accounted for 85.34% and the first and the second PC's accounted for 44.15% and 19.31% (total 63.46%) of the variance, respectively. Component loading of the first four principal components is given in Table 1. To aid visualization of the overall variability in the tested genotypes, the first two components scores (PC's) is plotted (Figure 1). Out of the total variation, PC1 and PC4 explained the largest (44.15%) and smallest (6.97%) variation,

respectively, while PC2 and PC3 accounted for 19.31% and 14.913% of the total variation, respectively. In each PC indicated maximum variation was found in first PC, therefore, selection for characters under PC1 may be desirable.

Traits	PC1	PC2	PC3	PC4
DH	0.3	-0.257	0.037	-0.048
DF	0.3	-0.26	0.034	-0.06
DM	0.28	-0.293	-0.04	-0.135
TP	-0.185	-0.16	0.447	0.308
PP	-0.131	-0.167	0.495	0.28
CL	0.229	0.275	-0.096	0.511
PL	0.317	0.076	0.012	0.191
PH	0.256	0.251	-0.08	0.476
FL	0.325	0.026	0.033	-0.007
FW	0.252	0.045	-0.214	0.074
FGPP	0.275	0.194	0.23	-0.233
UGPP	0.071	-0.368	-0.127	-0.137
PW	0.261	0.298	0.006	-0.323
TGW	-0.187	0.272	-0.267	-0.046
BY	0.288	-0.145	0.231	0.013
GY	0.168	0.225	0.447	-0.171
HI	-0.089	0.414	0.308	-0.253
Eigenvalue	7.5051	3.2823	2.5351	1.1852
Variability (%)	44.148	19.308	14.913	6.972
Cumulative (%)	44.148	63.456	78.369	85.341
Variability (%)	44.148	19.308	14.913	6.972

Abbreviations: DH:days to heading, DF:days to Flowering, DM:days to maturity, TP:numberof tillers per plant, PP:number of panicles per plant, CL:culm length, PL:panicle length, PH:plant height, FL:flag-leaf length, FW:flag-leaf width , FGPP:filled grain per panicle, UGPP:unfilled grain per panicle , PW:panicle weight, TGW:thousand grain weight, BY:Biomass yield ha-1 in Kg, HI:harvest index, GY:paddy grain yield ha-1 in Kg



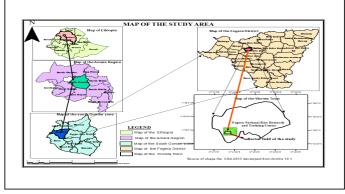


Figure1: Diagrammatic descriptions of experimental site.

Most yield and component characteristics contributed positively to PC1 except tillers per plant, panicles per plant, thousand-grain weight and harvest index. However, the highest variability was related with days to heading, days to 50% flowering, panicle length and flag leaf length. The second principal components (PC2) had positive loading for 10 traits out of the 17 studied traits, but the most contributed trait for the variation was dominated by harvest index. The third principal component (PC3) had positive loading for 11 traits out of the total studied traits, but dominated by panicles per plant, tillers per plant, grain yield and harvest index. The long vectors indicate that, they have a large contribution to the total variation [19].

Characters with lower absolute values, closer to zero, have less influence on clustering than largest absolute values within the principal components. Therefore, the above-mentioned characters, which load high positively or negatively, contributed more to the diversity and they were the ones that most differentiated the principal components. The graph of PC biplot classified the lowland rice genotypes into different group based on the most influential quantitative traits.

As a result clearly showed a breeder can easily pinpoint distances between the genotypes and make decisions based on the principal component simultaneously. In the bi plot, genotypes close to each other are similar while the ones found near the origin are distinct and the ones further out are extremes. PCA analysis classifies the genotypes into groups over the four quadrants based on the concentrations of these seventeen traits. The genotypes were distributed throughout the quadrant demonstrating large genetic variability in these traits.

The ones on the upper left quadrant were related in their harvest index and thousand grain weight while, the upper right quadrant contained genotypes related in their grain yield, panicle weight, plant height, filled grain per panicle, panicle length, flag leaf length and flag leaf width. The selected five rice genotypes (G8, G14, G26, G29 and G30) with grain yield above 6000 kg ha-1 were located in this quadrant, which clearly indicates the positive association of these traits for grain yield of rice genotypes.

The right bottom quadrant contain genotypes related in their biomass yield, days to heading, days to flowering, days to maturity and unfilled grains per panicle plant whereas, the left bottom genotypes related in their number of tillers per plant and number of panicles per plant.

The distance between the locations of any two genotypes on the biplot is directly proportional to the degree of similarity/difference between them as per the traits considered [20,21]. The bi plot showed that genotypes G3, G20, G22 and G29 were the most divergent from the major group. These extreme genotypes are favorable for breeding programs due to their morphological and biochemical differences from the rest which makes unique. Genotypes which nearly overlapped in the principal component axes had similar quantitative characters. Thus, (G14 and G26) at cluster III; (G15 and G24) at cluster I in the upper right quadrant and (G10 and G12) at cluster II in the left bottom quadrant showed similar relationship in the principal component axis (Figure 2).

Conclusion

The present study was an attempt to know the principal components for yield and yield contributed traits in lowland rice genotypes for future utilization in the breeding program. To generate this information a total of 27 lowland rice genotypes with two standard checks and one locally available genotype were evaluated using randomized complete block design with three replications during the 2017/18 main cropping season at FNRRTC. The analysis of variance showed highly significant differences among the tested genotypes for all 17 studied traits, which indicates that there is a considerable genetic variability in the tested rice genotypes.

Paddy grain yield ranged from 2766.7 kg ha-1 for G12 to 7062.0 kg ha-1 for G26 with a mean of 4736.13 kg ha-1. Among the genotypes, G26, G14, G8, G27, G29 and G30 were the six top yielders with corresponding mean grain yield of 7062, 6900, 6583.10, 6486.90, 6400.60 and 6343.10 kg ha-1. About 46.7% and 90% of the genotypes had higher mean grain yield than the standard check G6 (Ediget) and the local check G28 (X-Jigna), respectively.

The maximum cluster distance was found between cluster two and three while the minimum was found between cluster one and three. The first four principal components explained 85.34% of the total variation. PC1, PC2, PC3 and PC4 accounted for 44.15%, 19.31%, 14.913% and 6.97% of the total variation, respectively. Based on the present investigation results, it can be concluded that there is adequate genetic variability for most of quantitative characters evaluated, that the genotypes with high grain yield should be selected from different clusters and crossed so as to improve grain yield. The study also identified the best performing genotype for further evaluation and/or recommended for release for possible commercialization.

For future breeding programs that employ hybridization, parental material selection should be carried out between clusters rather than within clusters. It is recommended to repeat the study at more seasons and locations with more number of genotypes to predict genotypic performance across seasons and locations which helps to validate the obtained current results. Moreover, the future rice research should be supplemented by molecular characterization to further confirm the outcome of current study findings.

Conflict of Interests

The authors have not declared any conflict of interests.

Acknowledgements

The authors acknowledged Ethiopian Institute of Agricultural Research (EIAR) and Fogera National Rice Research and Training Center (FNRRTC) for funding the project. The authors obliged to extend deepest gratitude to Dr. Fisiha Worede, Dr. Tilahun Tadesse, Mr. Desta Abayechaw, Mrs. Helen Asaminew and Mr. Endalew Getu who helped us throughout our research period.

References

- Singh SK, Bhati PK, Sharma A, Sahu V (2015) Super hybrid rice in China and India: Current status and future prospects. Int J Agric. Biol 17:221-232.
- Gross BL, Zhao Z (2014) Archaeological and genetic insights into the origins of domesticated rice. Proceedings of the National Academy of Sciences 111:6190-6197.
- Ansari MU, Shaheen T, Bukhari S, Husnain T (2015) Genetic improvement of rice for biotic and abiotic stress tolerance. Turk J Bot 39:911-919.
- Messeguer J, Fogher C, Guiderdoni E, Marfa V, Catala MM, et al. (2001) Field assessments of gene flow from transgenic to cultivated rice (*Oryza sativa* L.) using a herbicide resistance gene as tracer marker. Theoretic App Gen 103:1151-1159.
- Oka HI (1988) Weedy forms of rice. Origin of cultivated rice. (edn), Elsevier, Amsterdam, pp. 107-114.
- Mickel A, Donini B, Maluszynski M (1990) Induced mutations for crop improvement. Mutuation Breeding Review, Vienna 1-16.
- Luz VK, Silveira SF, Fonseca GM, Groli EL, Figueiredo RG, et al. (2016) Identification of variability for agronomically important traits in rice mutant families. Bragantia 75:41-50.
- Chakravarthi BK, Naravaneni R (2006) SSR marker based DNA fingerprinting and diversity study in rice (*Oryza sativa* L). African J Biotechnol 5.
- 9. Zibaee A (2013) Rice: Importance and Future. J Rice Res 1:2.
- Rahman MM, Syed MA, Adil M, Ahmad H, Rashid MM (2012) Genetic variability, correlation and path coefficient analysis of some physiological traits of transplanted Aman rice (*Oryza sativa* L.). Middle East Journal of Scientific Research 11:563-566.
- Girma BT, Mohammed H, Abegaz K (2016) Physical characteristics and nutritional quality of salt tolerant rice genotypes. J Cereals Oilseeds 7:7-13.
- 12. MoARD (Ministry of Agriculture and Rural Development) (2010) National Rice Research and Development Strategy of Ethiopia. (edn), The Federal Democratic Republic of Ethiopia, Ministry of Agriculture and Rural development, Addis Ababa, Ethiopia. 48pp.

Page 5 of 5

- CSA (Central Statistical Agency) (2017) Agricultural Sample Survey Series: Report on Area and Production for Major Crops (Private Holdings, Main Season). Statistical Bulletin No. 578. Addis Ababa, Ethiopia. 14-16.
- Kirub A, Alemu D, Shiratori K, Assefa K (2011) Challenges and opportunities of rice in Ethiopian agricultural development. Addis Ababa: Ethiopian Institute of Agricultural Research.
- Gebrekidan H, Seyoum M (2006) Effects of mineral N and P fertilizers on yield and yield components of flooded lowland rice on vertisols of Fogera Plain, Ethiopia. J Agricult Rural Develop Tro Subtro 107:161-176.
- 16. Zegeye T, Alemayehu B, Agidie A (2005) Rice production consumption and marketing: The case of Fogera, Dera and Libokemikem districts of Amhara region Paper presented on the national workshop on the status of rice research and promotion in Ethiopia. (edn), Bahir Dar, Ethiopia, pp. 13.
- 17. MoARD (Ministry of Agriculture and Rural Development) (2009) National Rice research and development strategy of Ethiopia. MOARD, Addis Ababa, Ethiopia.
- Dutta P, Dutta PN, Borua PK (2013) Morphological traits as selection indices in rice: A statistical view. Univ J Agri Res 1:85-96.
- 19. Yan W, Kang MS (2002) GGE biplot analysis: A graphical tool for breeders, geneticists and agronomists. CRC press.
- Shegro A, Labuschagne M, Shargie N, Biljon A (2013) Multivariate analysis of nutritional diversity in sorghum land race accessions from Western Ethiopia. J Biol Sci 13: 67-74.
- (2002) Rice Almanac, source book for the most important economic activity on earth. In: Maclean JL, Dawe DC, Hardy B, Hettel GP (eds)(3rdedn), International Rice Research Institute, Manila, Philippines. pp.1-253.