



Principal component analysis in RIL population derived from the cross of Improved Samba Mahsuri and Phougak in rice (*Oryza sativa* L.)

Gorige Srihari, Jyothi Badri, B Krishnaveni, V Prakasam and R M Sundaram

Department of Genetics and Plant Breeding, Acharya N G Ranga Agricultural University,
Agricultural College, Bapatla- 522101, Andhra Pradesh, India

ABSTRACT

Rice is one of the principal cereal food crops which fulfills carbohydrate needs and nutritional requirement of half of the world's population. Despite its agricultural significance, rice production faces severe threats from many diseases, notably sheath blight (ShB), caused by *Rhizoctonia solani*. ShB leads to substantial yield losses, ranging from 4-50% based on environmental conditions and disease severity. The present study aimed to evaluate recombinant inbred line (RIL) population, derived from a cross between susceptible elite cultivar ('Improved Samba Mahsuri') and a ShB-tolerant landrace ('Phougak'), for traits associated with ShB tolerance and yield. Using principal component analysis (PCA), nine ShB and yield-related traits were evaluated in two environments, yielding nine and eleven principal components (PCs) for the dry season and wet season respectively. In both environments, the first three PCs explained 50% of observed phenotypic variance, with traits relative lesion height (RLH) and standard evaluation scores (SES) as primary contributors. PCA results revealed that yield contributing traits-tiller number, panicle number, and panicle length are positively correlated with grain yield, emphasizing their breeding relevance. Biplot analysis indicated critical trait interactions for RLH and SES and their strong influence on trait divergence. This research helps rice breeding efforts targeting ShB resistance with improved yield, by identifying promising genotypes for sustainable rice production.

Key words: *Eigen value, Principal components, Sheath blight, Variability.*

Rice is the most essential food crop in the world, serving as primary food source for more people than any other crop. It provides approximately 21% of global human per capita energy and 15% of per capita protein intake (Jeong *et al.*, 2017). In India, rice is cultivated in an area of 45.15 million hectares resulting in a production of 123.81 million tonnes and average productivity of 2742 Kg/ha (www.Indiastat.com, 2023-24). Despite its significance, rice is attacked by a variety of fungal, bacterial and viral diseases, which pose significant challenges to its cultivation and productivity. The rice disease dynamics have evolved over the years, primarily due to reduced varietal diversity, which has resulted in a narrow genetic base and increasing susceptibility to emerging pathogens. Break down of varietal resistance due to appearance of new virulent forms of the pathogen, made the management of diseases more challenging, emphasizing the immediate

necessary for continued research and innovation in rice breeding and disease management practices (Mwang *et al.*, 2023).

Among the various fungal diseases affecting rice, sheath blight (ShB) is considered as an important disease similar to rice blast. It is a soil borne disease caused by the fungus *Rhizoctonia solani* Kuhn AG1-IA (Anastamosis Group-IA). This disease results in both qualitative and quantitative losses and yield losses usually range from 4-50% depending on the crop stage at which the disease appears and level of severity and the environmental conditions (Singh *et al.*, 2004; Bhunkal *et al.*, 2015). The impact of ShB disease significantly hampers productivity, posing a significant risk to rice farming globally affecting the socio-economic status of the rice farmers. Although cultivating resistant varieties is regarded as the most effective strategy for managing ShB, no fully resistant varieties are currently available to farmers. Therefore,

understanding the variability within recombinant inbred line (RIL) populations that exhibit potential resistance to ShB is crucial. In this context, the present study evaluates a bi-parental RIL population against the ShB pathogen, employing principal component analysis (PCA) to better understand genetic resistance and inform future breeding efforts.

Principal Component Analysis (PCA) is a widely recognized and powerful multivariate statistical method that compresses, reduces and transforms data. PCA reduces data with a large number of correlated variables into a much smaller set of new variables through a linear combination of the original variables capturing most of the variation present in the data” (Varanasi *et al.*, 2023). It is used in data analysis to identify the minimum number of components that can explain the maximum variability from the total variability (Anderson, 1972). Keeping this in view, the present investigation was conducted on mapping population derived from the cross between ‘Improved Samba Mahsuri (ISM)’ and ‘Phougak’, to identify the quantitative traits responsible for variation among the lines in the population.

MATERIAL AND METHODS

The experiment was conducted with a bi-parental mapping population consisting of 360 lines developed from a cross between ‘Improved Samba Mahsuri (ISM)’ and ‘Phougak’. ‘ISM’ (ShB susceptible parent) is a popular elite fine grain cultivar with medium slender grains and bacterial blight resistance. ‘Phougak’ is a North Eastern landrace identified with consistent and stable ShB disease tolerance across the years and hot spot locations (Dey *et al.* 2016; Badri *et al.* 2023). This mapping population, along with five check varieties ‘ISM’ (Parent 1-P₁), ‘Phougak’ (Parent 2-P₂), ‘Swarna’ (ShB susceptible cultivar), ‘Tetep’ and ‘Wazuhophek’ (both ShB tolerant cultivars) were planted at research farm of ICAR-Indian Institute of Rice Research (ICAR-IIRR) in Hyderabad during the dry season (DS) of 2022-23 (December-May) and wet season (WS) of 2023 (July-November) with a spacing of 20 x 15 cm. Field experiment was conducted using augmented block design (ABD) with five blocks. The lines of the mapping population were grown only once, while the check entries were replicated in each block.

Data was collected for nine traits, days to 50% flowering (DFF), plant height (PH), tiller number (TN), panicle number (PN), panicle length (PL), grain yield un-inoculated (GY_u), grain yield inoculated (GY_i), relative yield loss (RYL) and SES scores. PCA was used to examine the traits with a maximum contribution to the phenotypic variation in the population. The mean data of ShB disease traits and yield-related traits was used for principal component analysis and the scree plot and biplots were constructed using R Studio (version 4.3.2.) to interpret and analyze the contribution of each trait to overall variability.

Pathogen inoculation to evaluate the phenotypes against ShB disease was carried out by using standard typha method (Bhaktavatsalam *et al.* 1978). RILs and checks were inoculated artificially at maximum tillering stage (40 days after transplanting) under field epiphytic conditions. Inoculum loaded with 3-4 typha bits were placed in the middle of tillers just above the water level. To keep the inoculum intact and ensure favorable microenvironment for uniform disease spread, all the tillers enclosing typha bits were loosely tied with rubber band and kept intact by locking with a small piece of stick. The assessment of the disease was carried out based on length of the lesion (LH) relative to plant height (PH). Disease reaction was recorded at two stages, first observation at 15 days after inoculation (DAI) and second observation at 30 DAI. The total height of lesion spread (cm) was recorded from plant’s base to the tip of the top most lesions on the stem. Relative lesion height (RLH) was computed using the following formula

$$RLH = \frac{\text{Lesion height}}{\text{Plant height}} \times 100$$

RLH was scored by following the Standard Evaluation System (SES) (IRRI, 2002). Rice sheath blight grade chart ranging from 0 to 9 scale was utilized for recording disease reaction and lesion height: 0 indicates no symptoms (immune), 1 indicates 20% symptoms (resistant), 3 indicates 21-30% symptoms (moderately resistant), 5 indicates 31-45% symptoms (moderately susceptible), 7 indicates 46-65% symptoms (susceptible) and 9 indicates 66-100% symptomatic area or collapsed plants (highly susceptible).

RESULTS AND DISCUSSION

Principal Component Analysis (PCA) is a commonly employed method for dimension reduction that can be used to simplify a large set of variables to a smaller set that still contains most of the information in the large set. In this study, PCA effectively captured genetic diversity among the genotypes. For the dry season, nine principal components (PCs) were extracted, with the first four informative PCs (eigen values >1) accounting for 81.19% of the cumulative variance (Table 1). Notably, 50 % of the variance in the population was recorded by the first three major PCs (PC₁ 27.11%; PC₂-22.44%; and PC₃-16.68%). During the WS, eleven (PCs) were extracted, with

the top four being informative PCs (eigen values >1) and accounting for 77.25% of the cumulative variance (Table 2). Here, the first two major PCs alone explained over 50% of the variance, with PC₁ at 29.72% and PC₂ at 22.85%. Similar results have been reported in studies on diverse rice populations, where the first three PCs captured the highest variability in introgression lines from multi-parent marker-assisted forward breeding (Varanasi *et al.*, 2023); in a diverse collection of *indica* and *tropical japonica* breeding lines (Deepika *et al.*, 2021) and among bi-parental lines from an *indica/tropical japonica* inter-sub specific cross (Harnal *et al.*, 2023).

Table 1. Eigen values, Percentage of variation and Cumulative percentage for principal components in mapping population from the cross of ‘Improved Samba Mahsuri’ and ‘Phougak’ during dry season, 2022-23

Principal components	Eigen value	Variance%	Cumulative Variance%
PC ₁	2.44	27.11	27.11
PC ₂	2.01	22.44	49.55
PC ₃	1.5	16.68	66.24
PC ₄	1.34	14.95	81.19
PC ₅	0.73	8.21	89.41
PC ₆	0.59	6.62	96.03
PC ₇	0.22	2.52	98.56
PC ₈	0.06	0.75	99.32
PC ₉	0.06	0.67	100

Table 2. Eigen values, Percentage of variation and Cumulative percentage for principal components in mapping population from the cross of ‘Improved Samba Mahsuri’ and ‘Phougak’ during wet season 2023

Principal components	Eigen value	Variance%	Cumulative Variance%
PC ₁	3.27	29.72	29.72
PC ₂	2.51	22.85	52.58
PC ₃	1.49	13.53	66.11
PC ₄	1.23	11.14	77.25
PC ₅	0.86	7.83	85.08
PC ₆	0.81	7.32	92.4
PC ₇	0.54	4.88	97.28
PC ₈	0.11	1.04	98.32
PC ₉	0.07	0.62	98.95
PC ₁₀	0.06	0.58	99.53
PC ₁₁	0.05	0.47	100

Across these PCs, all the traits exhibited positive loadings in both seasons. In PC₁, relative lesion height (RLH) and standard evaluation score (SES) contributed substantially to the variation, while total tiller number (TN) and panicle number (PN) were prominent in PC₂ and PC₃. These traits play a vital role in the observed variance, capturing a major part of the variability in the dataset. Selecting traits exhibiting high variability will be useful for future breeding programs. In both seasons, the first three PCs cumulatively explained 50% of the total variability

(Tables 3 and 4). Similarly, in a related study, the first three PCs contributed 68.69% to total variability, with the number of productive tillers per plant showing positive loadings in PC₁ (Kumari *et al.*, 2021).

In this investigation, PC₁ accounted for 27.11 % of the variability, with an eigen value of 2.44 during the DS and 29.72 % variability, with an eigen value of 3.27 during the WS. The biplot diagrams in the present study illustrate interactions among genotypes performing well for specific traits (Fig. 1A, 1B, 1C and 1D). The vector length of each trait depicts the

Table 3. Contribution of principal components to variation in population derived from the cross of Improved Samba Mahsuri and Phougak during dry season, 2022-23

Character	PC ₁	PC ₂	PC ₃	PC ₄	PC ₅
DFF	0.01	8.05	5.82	29.11	12.86
LH	7.03	0.64	49.46	0.86	0.01
PH	13.32	2.08	35.88	0.2	0.79
RLH	35.83	0.4	0.47	2.34	0.2
SES	32.16	0.51	1.04	3.24	0.01
TN	0.41	42.41	3.19	3.48	0.12
PN	0.18	43.63	1.17	4.4	0.02
PL	6.44	0.03	2.63	31.44	18.13
GY	4.62	2.25	0.32	24.94	67.86

DFF-Days to 50% flowering, LH-Lesion height (cm), PH-Plant height (cm), RLH-Relative Lesion Height (%), SES-Standard Evaluation Score, TN- Tiller number, PN- Panicle number, PL- Panicle length (cm), GY- Grain yield (g).

Table 4. Contribution of principal components to variation in population derived from the cross of Improved Samba Mahsuri and Phougak during wet season, 2023

Character	PC ₁	PC ₂	PC ₃	PC ₄	PC ₅
DFF	0.35	4.28	3.81	10.97	79.25
LH	11.62	0.66	2.67	33.74	7.59
PH	7.92	11.58	10.19	10.14	4.04
RLH	23.77	4.07	1.84	3.4	0.24
SES	21.67	3.69	2.02	6.75	0.18
TN	0	20.54	29.64	0.22	0.27
PN	0.22	22.9	25.08	0.28	0.02
PL	10.22	2.33	1.93	14.85	6.11
GYu	8.61	9.78	3.1	7.18	0.02
GYi	10.39	14.6	9.47	8.38	0.74
RYL	5.24	5.57	10.25	4.09	1.53

DFF-Days to 50% flowering, LH-Lesion height (cm), PH-Plant height (cm), RLH-Relative Lesion Height (%), SES-Standard Evaluation Score, TN- Tiller number, PN- Panicle number, PL- Panicle length (cm), GYu- Grain yield under un-inoculated conditions (g), GYi- Grain yield under inoculated conditions (g) and RYL-Relative Yield Loss (%).

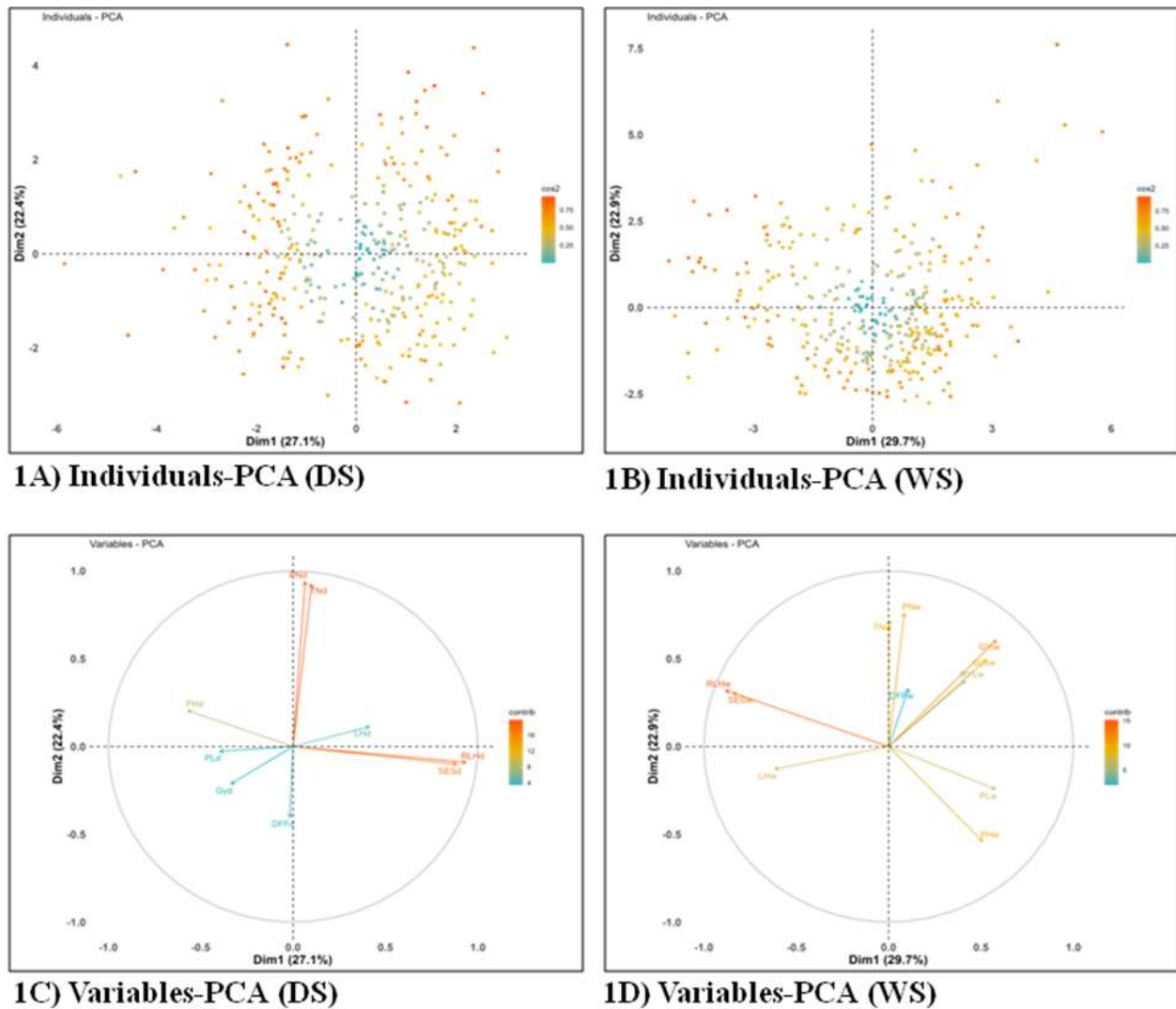


Fig. 1A-1D. Principal Component Analysis for ShB and yield traits among 360 RILs derived from the cross of ISM and Phougak. 1A) PCA biplot depicting contribution of 9 traits to the total variation, 1B) PCA biplot depicting contribution of 11 traits to the total variation, 1C) scree plot for percentage of explained variance of 9 Principal components and 1D) scree plot for percentage of explained variance of 11 PC components. DFF-Days to 50% flowering, LH-lesion height (cm), PH-plant height (cm), RLH-relative lesion height (%), SES-Standard evaluation score, TN- Tiller number, PN- Panicle number, PL- Panicle length (cm), Grain yield (g), Gyu- Grain yield under un-inoculated conditions (g), Gyi-Grain yield under inoculated conditions (g) and RYL-Relative Yield Loss (%).

contribution of the trait to total divergence; longer vector length in general indicates more contribution of concerned trait. The biplot also showed the distribution and diversity of genotypes with respect to sheath blight (ShB) disease and yield related traits. RLH and SES traits exhibited longest vector lengths, indicating that they contribute to total divergence, followed by trait panicle number. The angle between trait vectors in the biplot reflects the nature of association between

traits: an angle $<90^\circ$ implies a positive correlation, 90° indicates no correlation, and $>90^\circ$ suggests a negative correlation. Traits LH, RLH, and SES exhibited negative correlation with grain yield, while other component traits showed a positive correlation with grain yield per plant. Thus, selection with regard to grain yield based on traits like TN, PN, and PL is likely to enhance grain yield. Genotypes positioned close to a trait vector within the same quadrant are

expected to perform well for those traits. The present study identified the key variability contributing parameters that help to identify promising lines that can serve as donors for improving yield traits with ShB tolerance in future breeding programs.

In the previous studies, the percentage of variation for the initial three PCs (PC_1 , PC_2 , PC_3) together accounted for 72.9 % of variability with panicle number (0.1482) showing positive loading (Christina *et al.*, 2021). Similarly, PC_1 and PC_2 together explained 61.2% of the total variation, where PC_2 was associated with variation for total tillers and productive tillers (Worede *et al.*, 2014); and PC_1 with positive loading for panicle number (0.359) (Kumar *et al.*, 2021). These findings explained how different principal components capture distinct aspects of trait variability. The various principal components capture diverse aspects of variability, and the significant eigen values generated from analysis provide valuable insights into the distinct contributions of different variables. This understanding is crucial for developing targeted breeding strategies to improve yield with ShB tolerance. In the present study principal component analysis has demonstrated distinct patterns and relationships among the studied variables.

CONCLUSION

In the present study principal component analysis (PCA) was used to dissect and capture the genetic variability among RIL population of cross between ISM and Phougak, particularly in relation to sheath blight (ShB) disease resistance and yield traits. PCA identified key traits contributing to genetic diversity, with relative lesion height (RLH) and standard evaluation score (SES) showing substantial impacts on variance. The analysis identified that traits like tiller number, panicle number, and panicle length, recorded positive associations with grain yield. The biplot visualization explained the correlations between traits, with RLH and SES as prominent factors contributing to genetic divergence. The results from this PCA will serve as valuable resources for rice breeding, especially for selecting genotypes that exhibit resilience against ShB with enhanced yield potential. By focusing on genotypes within specific quadrants of the PCA biplots, we can identify promising lines that combine favorable yield-related traits with ShB tolerance. These findings confirm similar patterns observed in other studies providing a basis for future breeding strategies that aim to balance high yield with disease resistance.

LITERATURE CITED

- Anderson TW 1972.** An Introduction to Multivariate Analysis. Wiley Eastern Pvt. Ltd. New Delhi.
- Badri J, Prakasam V, Sundaram RM, Prasad M S, Laha G S, Bhadana V P, Subba Rao L V and Priyanka C 2023.** Phougak (D82), a rice (*Oryza sativa* L.) germplasm for tolerance to sheath blight, resistance to neck blast, and resistance to leaf blast. *Indian Journal of Plant Genetic Resources*, 36(2): 319-340.
- Bhaktavatsalam G, Satyanarayana K, Reddy A P K and John V T 1978.** Evaluation of sheath blight resistance in rice. *International Rice Research Newsletter*, 3: 9-10.
- Bhunkal N, Singh R and Mehta N 2015.** Assessment of losses and identification of slow blighting genotypes against sheath blight of rice. *Journal of Mycology and Plant Pathology*, 45: 285-292.
- Christina GR, Thirumurugan T, Jeyaprakash P and Rajanbabu V 2021.** Principal component analysis of yield and yield-related traits in rice (*Oryza sativa* L.) landraces. *Electronic Journal of Plant Breeding*, 12(3): 907-911.
- Deepika K, Lavuri K, Rathod S, Yeshala C M, Jukanti A K, Reddy S N, L V S R and Badri J 2021.** Multivariate analysis of geographically diverse rice germplasm for genetic improvement of yield, dormancy, and shattering-related traits. *Plant Genetic Resources*, 19(2): 144-152.
- Dey S, Badri J, Prakasam V, Bhadana V P, Eswari K B, Laha G S, Priyanka C, Rajkumar A and Ram T 2016.** Identification and agro-morphological characterization of rice genotypes resistant to sheath blight. *Australasian Plant Pathology*. 45: 145-153.
- Haranal A A, Eswari KB, Badri J and Rao P R 2024.** Yield trait association studies through principal component analysis in a bi-parental population derived from indica/tropical japonica inter-subspecific cross in rice (*Oryza sativa* L.). *Journal*

- of Experimental Agriculture International*, 46(8): 651-657.
- International Rice Research Institute (IRRI) 2002.** *Standard Evaluation System for Rice*, IRRI, Philippines.
- Jeong J M, Kim E C, Venkatanagappa S and Lee J S 2017.** Review of rice: Production, trade, consumption, and future demand in Korea and worldwide. *Korean Journal of Crop Science*, 62(3): 157-165.
- Kumari B K, Kumar B R, Jyothula D P and Rao N M 2021.** Diversity analysis in rice breeding lines for yield and its components using principal component analysis. *Journal of Pharmacognosy and Phytochemistry*, 10(1): 905-909.
- Mwangi R W, Mustafa M, Charles K, Wagara I W, Kappel N 2023.** Selected emerging and reemerging plant pathogens affecting the food basket: A threat to food security. *Journal of Agriculture and Food Research*, 14: 100827.
- Singh S K, Shukla V, Singh H and Sinha A P 2004.** Current status and impact of sheath blight in rice (*Oryza sativa* L.): A review. *Agricultural Reviews*, 25(4): 289-297.
- Varanasi Y V, Isetty S R, Revadi P, Balakrishnan, D, Hajira S, Prasad M S, Laha G S, Perraju P, Singh U M, Singh V K and Kumar A 2023.** Molecular and morphological characterization of introgression lines with resistance to bacterial leaf blight and blast in rice. *Plants*, 12(16): 3012.
- Worede F, Sreewongchai T, Phumichai C and Sripichitt P 2014.** Multivariate analysis of genetic diversity among some rice genotypes using morpho-agronomic traits. *Journal of Plant Sciences*, 9(1): 14-24.

Received on 02.01.2024 and Accepted on 13.03.2024