Indian J. Genet. Plant Breed., (2023); 83(3): 343-354
https://doi.org/10.31742/ISGPB.83.3.6

ISSN: 0975-6906
www.isgpb.org

RESEARCH ARTICLE

Assessment of genetic diversity for seed vigor traits in rice (*Oryza sativa* L.) through principal component and cluster analysis

Shruti Kumari, S. K. Chakrabarty* and Debashis Paul

Abstract

Seed vigor, an important characteristic of seed quality attributes, determines the rate of early, rapid, uniform germination, emergence and growth of strong seedling in any environmental condition. In the present set of experiment, 44 rice (*Oryza sativa* L.) varieties were assessed for different seed vigor component traits. For evaluation of these traits, basic statistics, correlation, principal component (PC) and cluster analysis were carried out to obtain genetic diversity for seed vigor traits and early seedling vigor among rice varieties specifically to select the best varieties and important seed vigor traits for its possible use in future rice breeding programme. The estimation of descriptive statistics of 20 quantitative seed vigor-related traits indicated existence of variability among the varieties. The 44 varieties were grouped into five clusters on the basis of average linkage and dendogram. Varieties in Cluster I are considered to possess very high vigor traits and cluster IV with least seed vigor traits. Principal component (PC) analysis showed first 5 PCs having Eigen value >1 explaining 77.01% of the total variation with different seed vigor traits. In the biplot analysis between PCs 1 and 2, the varieties remained scattered in all four quadrants, showing large genetic variation in quantitative traits. Seed vigor index-I, seed vigor index-II, seed dry weight, shoot length, reserve mobilization, germination percentage, dehydrogenase, mean germination rate, germination speed, and mean germination time contributed to high variability among the varieties. The seed vigor traits were positively correlated with germination percentage, shoot length, seedling dry weight, speed of germination, mean germination rate, reserve mobilization, superoxide dismutase, dehydrogenase and alpha-amylase. The positive correlation among these seed vigor-contributing traits suggested that these characteristics are important for direct selection of high seed vigor varieties and early emergence.

Keywords: Seed vigor traits, diversity, PCA, cluster analysis, rice

Introduction

Rice (*Oryza sativa* L.) is one of the important and the second largest food grain crops grown in the world both in terms of area and production. The success of every crop's germination, growth, and final yield depends largely on the quality of the seeds used to grow the crop. Seed quality is defined as a seed's viability and vigor that enables the emergence and establishment of normal seedlings under a wide range of environments (Hilhorst and Koornneef 2007). Seed vigor has been known as a complex characteristic affected by many factors, such as the genetic background and environmental factors during seed development and storage stages (Sun et al. 2007). In rice, seed vigor is a critical factor in determining the success of crop production, especially under adverse environmental conditions. Several factors such as seed size, weight, shape, color, and physiological status influence seed vigor in rice. Therefore, understanding the genetic and environmental factors that influence seed vigor traits can help in developing improved cultivars with high seed vigor. Reddy et al. (1996) studied the relationship between seed vigor, seedling length and initial germination of rice varieties. The results showed that seedling length, early seedling vigor, speed of germination and dry matter production decrease with decreasing initial germination percentage. The cultivars with high seed vigor are desirable for farmers to get optimum stand establishment under sub-

---

Division of Seed Science and Technology, ICAR-Indian Agricultural Research Institute, New Delhi 110 012, India

***Corresponding Author:** S. K. Chakrabarty, Division of Seed Science and Technology, ICAR-Indian Agricultural Research Institute, New Delhi 110 012, India, E-Mail: skchakra_sst@yahoo.com

How to cite this article: Kumari S., Chakrabarty S. K. and Paul D. 2023. Assessment of genetic diversity for seed vigor traits in rice (*Oryza sativa* L.) through principal component and cluster analysis. Indian J. Genet. Plant Breed., **83**(3): 343-354.

Source of support: Nil

Conflict of interest: None.

Received: April 2023 **Revised:** July 2023 **Accepted:** Aug. 2023

optimal field conditions. Seed vigor has been shown to be a varietal attribute and its expression is affected by many factors, such as the genetic background, environmental factors during seed development, and storage stages (Sun et al. 2007). Seed with high vigor may significantly improve the speed and uniformity of seed germination, the final percentage of seedling emergence, good crop performance, and even high yield under different conditions (Foolad et al. 2007). Some biochemical factors, such as the activity of certain enzymes (e.g: α-amylase and glutamine synthetase) and such processes as oxidative phosphorylation, protein, RNA and DNA synthesis and ATP content, show high activity in high-vigor cultivars (Ching 1982). The lower productivity level may be due to a shortage of water, land, labour and other resources to produce higher yields in addition to the stresses created by changing climate scenario (Singh et al. 2013). Various environmental and genetic factors influence the seed biochemical traits, such as CAT, POD and SOD (Shruti et al. 2023). Given these limitations, direct seeding is considered to be a viable option to maintain and increase the productivity level of rice under any stressed situations.

Multiple selection criteria are crucial to identify favourable genotypes with compatible trait combinations. Correlation studies reveal trait relationships guiding strategies for concurrent trait improvement. Genetic diversity assessment among the genotypes facilitates the generation of superior recombinants. (Basavaraj. et al. 2022). On the other hand principal component analysis is to reduce the dimension of data. Watson and Eyzaguirre (2002) reported that PCA of morphological characterization results could identify a few key or minimum descriptors that effectively account for the majority of the diversity observed, saving time and effort for future characterization efforts. Principal Component Analysis (PCA) and Cluster Analysis (CA) are two popular multivariate statistical techniques used to identify the underlying patterns and relationships among variables in a dataset. PCA effectively quantifies genetic divergence among germplasm lines (Beena et al. 2020). On the other hand, CA is a non-parametric method that groups similar observations into clusters based on their distance or similarity measures. For instance, a study by Li et al. (2020) used PCA and CA to investigate the genetic variation and relationship among different seed vigor traits in rice. The study found that seed size, weight, and germination rate were the most important variables that contributed to the variation in seed vigor traits, and that different rice varieties exhibited distinct patterns of seed vigor traits. So, these techniques can help breeders to identify the most important factors that influence seed vigor, and develop effective strategies for improving seed vigor in rice cultivars. Therefore, the present study was carried out to evaluate the genetic diversity among rice varieties specifically for seed vigor traits to select the best genotype that can be exploited in rice breeding.

Material and methods

Seed materials

The seeds of 44 Indian rice varieties were used in this study (Table 1). The seeds of all 44 varieties were produced following the recommended package of practices in the field at ICAR-Indian Agricultural Research Institute, New Delhi during *kharif*, 2020.

Recording of data

Observations were recorded in the laboratory of the Division of Seed Science and Technology for different seed vigor characters which is categorized in three different classes such as seed vigor traits, membrane functions and antioxidant activities during early germination stage. Seed vigor trait comprises of test weight (TW), moisture content (MC)(ISTA 1985), root length (RL), shoot length (SL), seedling dry weight (SDW), germination percentage (GP) (ISTA rules 2021), speed of germination (SOG) **(**Maguire 1962**)**, seed vigor index-I (SVI) (Abdul Baki and Anderson 1973) and seed vigor index-II (SVII) (Abdul Baki and Anderson 1973), mean germination time (MGT) (Nicholas and Heydecker 1968), mean germination rate (MGR), and reserve mobilization (RM) (Soltani et al. 2006). One (1) gram of seed samples were taken to conduct membrane functions traits comprising of electrical conductivity (EC) (AOSA 1983) and water-soluble sugar (WSS) (Dubois et al.1956). Antioxidant enzymes like catalase (CAT) (Aebi 1984), peroxidase (POD) (Castillo et al. 1984) and superoxide dismutase (SOD) (Dhindsa et al.1981), dehydrogenase(Kittock and Law 1968), hydrogen peroxide (H_2O_2) (Mukhejee and Choudhari 1983) and alpha-amylase (Jones and Varner 1967) were also conducted on 1gram of rice seed sample by following the standard protocol.

Statistical analysis

The data were subjected to correlation analysis, principal component analysis (PCA) and cluster analysis using statistical software packages of SAS 9.2. The correlation coefficient was calculated for all seed vigor and its related traits using Pearson's correlation coefficient. Cluster analysis was performed using average linkage clustering while Ward's method developed a tree diagram based on squared Euclidian distances. The first two principal components (PC_1 & PC_2/ Dim1& Dim2/Component 1& Component 2) were plotted against each other to find out the patterns of variability among genotypes and characters using SAS 9.2 software (SAS Institute 1996).

Results and discussion

The estimation of descriptive statistics *viz*., maximum, minimum, mean, Standard Deviation (SD) and Variance (V) for the measured 20 quantitative traits indicated the existence of diversity among the genotypes are shown in (Table 2). Among all the traits investigated, seed vigor index-

Table 1. A list of rice genotypes alongwith their place of origin and pedigree

Sl.no.	Name of variety	Place of origin	Parentage	Sl.no.	Name of variety	Place of origin	Parentage
1.	JD-6	New Delhi	Nagina 22 x Dular	23.	Satabdi	West Bengal	CR-10-114 x CR-10115
2.	VL Dhan 221	Uttarakhand	IR2053 x Ch1039	24.	Vikramarya	Andhra Pradesh	Vikram x PTB-2 (Ponnayan) RP-1125-1548-1-4-3
3.	VL Dhan 81	Uttarakhand	CH-988 x HPU-741	25.	Jaya	Karnataka	T(N)1 x T-141
4.	Remaya	Kerala	Jaya x Ptb 33 (HS)	26.	Shyamala	Madhya Pradesh	R 60-2713 x R 2386
5.	Poornima	Madhya Pradesh	Poorva x IR-8608-298	27.	Mandyavijaya	Karnataka	Sona x Mahsuri
6.	Pant Dhan 11	Uttar Pradesh	VL-206 x Dagi	28.	Dudheswar	West Bengal	Farmers' Traditional Variety
7.	JD-13	New Delhi	Kagalikai x Jaldi Dhan 8	29.	Budiluchai	Madhya Pradesh	Farmers' Traditional Variety
8.	Govind	Chhattisgarh	IR-20 x IR-24	30.	Naveen	West Bengal	Sattari x Jaya
9.	IR 688 97 B	IRRI	-	31.	Vasumati	Odisha	PR 109 X Pakistani Basmati
10.	Rasi	Telengana	T(N)1 x Co.29	32.	CRD-204	Odisha	IRRI 76569- 259-1-2-1 x CT 6510-24-1-2
11.	Annada	West Bengal	MTU-15 x Yaikaku Nantoku (China)	33.	IR-64	IRRI, Philippines	IR-5857-33-2-1 x IR-2061-465-1-5-5
12.	Jyoti	Kerala	Ptb-10 x IR-8 (HS)	34.	Basmati-370	New Delhi	Pure line from a locally adapted landrace
13.	Pusa 834	New Delhi	IR-50 x PP-33-BP	35.	Indira Sugandhit	Chhattisgarh	IET 21842 (R 1536-136-1-17-1)
14.	Pusa 33	New Delhi	Improved Sabarmati x Ratna	36.	Phalguna	Andhra Pradesh	IR-8 x Siam-29
15.	NDR 97	Uttar Pradesh	N-22 x Ratna	37.	MTU1010	Andhra Pradesh	Krishnaveni x IR 64
16.	Vandana	Odisha	C 22 x Kalakeri	38.	Kranti	Tamil Nadu	Cross-16 x IR-8
17.	Ravi	Andhra Pradesh	M-63-83 x (RP-79-5 x Rikutonorin-21)	39.	Swarna	Andhra Pradesh	Vasisa x Mahsuri
18.	Lachit	Assam	CRM13-3241x Kalinga II	40.	Krishnaveni	Andhra Pradesh	Sowbhagya x150 ARC 5984
19.	Satyabhama	Odisha	IR 31238-350- 3-2-1 x IR 41054-102-2- 3-2	41.	Salivahana	Odisha	RP5-32 x Pankaj
20.	Bidhan Gontra-2	West Bengal	-	42.	CRD-300	Odisha	NDR 9370018 x KDML 05 x PSBRC 60
21.	PNR-381	New Delhi	Tainan-3 Mutant X Basmati- 370	43.	Surajone	Chhattisgarh	Farmers' Traditional Variety
22.	Pant Dhan-12	Uttar Pradesh	Govind x UPR-201-1-1	44.	Ranikajar	Chhattisgarh	Farmers' Traditional Variety

Table 2. Estimates of basic statistics for 20 quantitative traits in 44 varieties of rice

Trait (unit)	Mean	Min.-Max.	Range	Variance	SD
Test weight(g)	22.7	12.0-28.3	16.3	0.14	0.37
Moisture content (%)	9.88	8.26-11.28	3.02	0.32	0.56
Root length(cm)	22.39	8.78-29.77	20.99	20.35	4.51
Shoot length (cm)	12.15	6.93-16.78	9.85	5.81	2.41
Seedling dry weight (mg)	0.056	0.019-0.081	0.06	0.00017	0.01
Germination (%)	86.06	23-100	77	523.41	22.88
Seed vigor index-I	3074.06	377.04-4367	3989.96	1147804.48	1071.36
Seed Vigor index-II	5.26	0.45-8.11	7.66	3.74	1.93
Speed of germination	12.75	8.15-15.85	7.71	5.33	2.31
Mean germination time (d)	5.28	2.54-6.88	4.34	1.88	1.36
Mean germination rate	0.25	0.15-0.40	0.25	0.005	0.07
Reserve mobilization	0.08	0.038-0.091	0.054	0.0002	0.02
Electrical conductivity (µS/cm/g)	17.52	10.09-36.29	26.20	48.88	6.99
Water soluble sugars (µg/ml)	43.21	16.5-82	65.50	271.79	16.49
Catalase (µM/min/gFW)	0.019	0.003-0.061	0.058	0.00017	0.0129
Peroxidase (µM/min/gFW)	0.009	0.001-0.037	0.036	0.00009	0.0097
Super oxide dismutase (µM/min/gFW)	128.8	14.35-198.85	184.50	2641.32	51.39
Dehydrogenase ($OD{\cdot}g^{-1}{\cdot}m^{-1}$)	0.07	0.006-0.116	0.110	0.0008	0.03
Hydrogen peroxide (H_2O_2 µm/g)	0.83	0.246-1.773	1.53	0.17	0.41
Alpha amylase (mg/mL/ g)	0.64	0.516-0.964	0.45	0.007	0.08

I, superoxide dismutase, germination (%), root length, water-soluble sugars and electrical conductivity recorded higher variation in mean, range, variance and standard deviation among the varieties. The lowest variation in mean, range, variance and standard deviation was found in seed dry weight, germination time, germination rate, dehydrogenase and alpha-amylase activity. Results of the study revealed that there is a large quantity of variability for seed vigor traits in rice varieties. Thus, genetic diversity analysis aids parent selection, maintains and utilizes desirable variations, and enhances insights into crop evolutionary pattern in breeding programme (Bollinedi et al 2020).

Cluster analysis classifies the varieties into different clusters based on their similarity features on seed vigor traits. Based on average linkage and dendrogram (Fig. 1) 44 varieties were clustered into five groups considering all 20 seed vigor-related parameters. The cluster analysis sequestrates 44 varieties into five clusters that exhibit high similarity within a cluster and high differences between clusters (Table 3).

The distribution pattern of all the varieties into five clusters (Table 3) showed the presence of considerable genetic diversity among the varieties for most of the traits under consideration. The distribution of varieties in the different clusters suggests that there are different levels of performance in seed vigor-related traits among the varieties and can be used to identify and select high-performing varieties. The results of this study are consistent with previous research that has shown the presence of genetic diversity in rice varieties for seed vigor traits (Kumar et al. 2015; Reddy et al. 2017). The use of cluster analysis to identify groups of similar varieties has been widely used in plant breeding and genetics research (Singh et al. 2018; Sahoo et al. 2020). The presence of significant genetic diversity among the varieties suggests the potential of seed vigor traits for genetic improvement.

In cluster analysis, a constellation plot is a graphical representation that helps visualize the clustering structure and relationships among the clusters. It is a scatter plot where each point represents an observation or data point, and different colors or symbols indicate the clusters. In this constellation plot cluster I and II have more similarity and hence present in one side and cluster III, IV and V having more similarity present in another side (Fig. 2). Therefore, the varieties present in cluster I and II are similar and different from those of cluster III, IV, and V for these 20 seed vigor traits.

Principal component analysis for seed vigor traits

It is difficult to understand the reason for variability for seed vigor traits in the large sets of varieties under

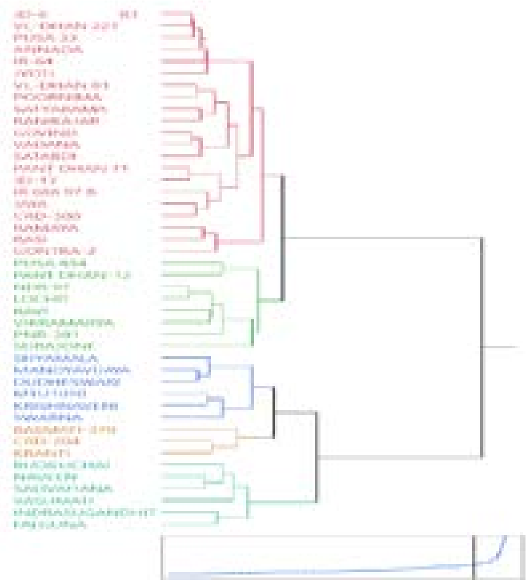


Fig. 1. The dendrogram of rice varieties resulting from cluster analysis using Ward methods based on standardized data of all 20 seed vigor traits

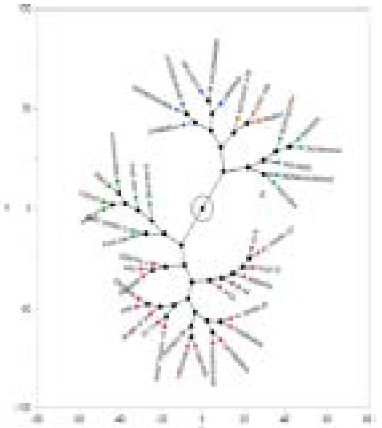


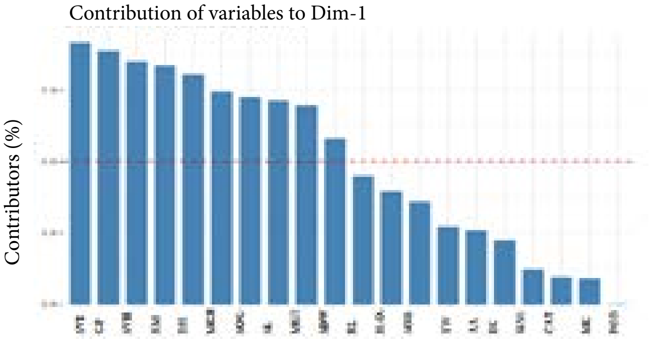
Fig. 2. Constellation plot of 44 varieties of rice

different maturity groups. The principal component analysis is conventionally used for extraction of features or characteristics for better understanding and implementation of cluster analysis. It also measures the importance and contribution of each component to the total variance. PCA can be used for the measurement of the independent impact of a particular trait on the number of components extracted is equal to the number of variables being analyzed. So the PCA grouped the 20 vigor-related parameters into 20 components, accounting for the entire (100%) variability among the varieties. But, according to Chatfield and Collins (1980), components with an eigenvalue of <1 should be eliminated so that only a few key components can be dealt with. Eigenvalue >1 is considered significant, determining a minimum of 10% of the variation. Trait with superior Eigenvalue is considered as best attribute in principal components. In this study, principal factor analysis identified five principal components with eigenvalue of more than one (>1), explaining about 77.01% of the total variability among the 44 varieties by considering all 20 seed vigor-related parameters. The remaining 15 components contributed only 22.99% towards the total variability among the varieties studied (Table 5). PC_1 has an eigenvalue of 9.759 which is the maximum and accounted for 48.79% of the variation. This represents several seed vigor traits *viz.*, seed vigor index-I, seed vigor index-II, germination percentage, shoot length, mean germination time, reserve mobilization, and dehydrogenase activity and indicated that those were important contributing variables for the variation among the varieties. For instance, a study by Yohane et al. (2020) found that the first principal component accounted for the most significant proportion of variation in seed traits of pigeon pea. Similarly, a study by Ghafoor et al. (2001) reported that the first principal component explained the highest proportion of variation in wheat yield traits. Chozin (2007), Mujaju and Chakuya (2008) and Ali et al. (2011) reported the important contribution of the first PC in total variability while studying different traits. The results were found to be in accordance with the studies on seed and seedling vigor traits, including relative stage of seedling development (Namuco et al. 2009), relative seedling growth rate measured as dry weight accumulation, seedling survival (Farooq et al. 2006) and seed response to stress (Rajjou et al. 2012). The remaining traits *viz.*, mean germination rate, electrical conductivity, water-soluble sugars, peroxidase, and hydrogen peroxidase, contributed negatively. PC_2 has an eigenvalue of 1.77, accounting for 8.87% of the variation. The PC_2 is related to variability in the varieties due to test weight having positive factor loading value, root length, and hydrogen peroxide contributed negatively. Similarly, PC_3 expressed loading value for variation among varieties resulting from moisture content, and alpha-amylase, catalase, and peroxidase were negatively contributed. PC_4 was related to variability among rice varieties due to the negative contribution of moisture content, electrical conductivity, water-soluble sugars, peroxidase, and superoxide dismutase, and similarly, PC_5 estimated variation among the varieties mainly for alpha-amylase, electrical conductivity, and water-soluble sugars and negative contribution to moisture content and hydrogen peroxide.

These findings suggest that certain seed vigor traits, such as germination percentage and shoot length, are important contributors to the overall variability in rice varieties. Other variables, such as moisture content and peroxidase, have a negative impact on this variability. Previous studies have shown the importance of seed vigor traits in determining overall crop performance and yield (Finch-Savage et al. 2021; Li et al. 2019). The first component can be expected to account for a large amount of the total variance. PC_1 has a maximum variability of 48.80% of the total variation

Table 3. Distribution of 44 varieties of rice into five clusters

Cluster	Cluster size (No. of variety)	Name of variety
I	21	JD-6, VL Dhan-221, VL Dhan-81, Pusa-33, Vandana, Satyabhama, Bidhan Gontra-2, Satabdi, Jaya, IR-64, CRD-300, Ranikajar, Remaya, Poornima, Pant Dhan 11, JD-13, Govind, IR688 97 B, Rasi, Annada and Jyoti.
II	8	Pusa 834, Surajone, NDR 97, Ravi, Lochit, PNR-381, Pant Dhan-12, and Vikramarya
III	6	MTU1010, Krishnaveni, Swarna, Shyamala, Mandyavijaya, and Dudheswar
IV	3	Kranti, Basmati-370, and CRD-204
V	6	Indira Sugandhit, Falguna, Budiluchai, Naveen, Vasumati, and Salivahana



SVI = Seed vigor index I, GP = Germination percentage, SVII = Seed vigor index II, RM = Reserve mobilization, DH = Dehydrogenase, MGR = Mean germination rate, SOG = Speed of germination, SL = Shoot length, MGT = Mean germination time, SDW = Seedling dry weight, RL = Root length , H_2O_2 = Hydrogen peroxide, SOD = Super oxide dismutase, TW = Test weight, AA = alpha amylase, EC = Electrical conductivity, WSS = Water soluble sugars, CAT = Catalase, MC = Moisture content and POD = Peroxidase

Fig. 3. Contribution (%) of seed vigor related parameters towards variability among the varieties

for seed vigor traits, PC_2 loaded with 8.88% of the total variability and PC_3, PC_4, and PC_5 contributed 7.33, 6.77, and 5.24% of the total variability, respectively. This method has been widely used in various fields, including agriculture, to analyze complex datasets and identify the underlying patterns or factors contributing to variability. PCA has also been found to be effective in identifying the most important traits and help in the selection of desirable traits for crop improvement (Cullis et al. 2006; Langewisch et al. 2017). Hence, varieties and traits selected from this component (PC_1) will be useful in crop improvement for improving the traits (Talekar et. al. 2022).

Contribution (%) of seed vigor-related parameters towards variability among the varieties

Seed vigor is an essential trait that determines the performance of seeds under various environmental conditions. The vigor of seeds depends on various physiological and biochemical factors, and the parameters identified in this study can be used to assess seed vigor accurately. Through PC_1/Dim_1 it was found that out of 20 seed vigor-related parameters, ten parameters contributed a higher proportion of variability among the genotypes, and the remaining contributed low (Fig. 3). These high vigor parameters can be used to assess the seed vigor accurately and select high-vigor genotypes. The findings of this study are similar with previous studies that reported the importance of seed vigor index-I, seed vigor index-II, seed dry weight, germination percentage, and mean germination rate in determining the seed vigor (Baskin and Baskin2014; Farooq et al. 2016). The seed vigor-related parameters that contributed less to the variability were moisture content, water-soluble sugars, root length, test weight, electrical conductivity, catalase, peroxidase, superoxide dismutase, hydrogen peroxide, and alpha-amylase (Fig. 3). Similarly, seed producers can use these parameters to assess the quality of seeds and ensure that only high-quality seeds reach the market. Moreover, the parameters identified in this study can also be used to study seed vigor's genetic and physiological basis and identify the molecular markers associated with this trait. So PCA analysis was conducted to identify a few key traits that accounted for the larger variability.

Ranking of cluster of varieties based on seed vigor traits, membrane functions and antioxidant activities

By considering all of the above important parameters which contributed more variability for seed vigor traits taken and again classified these clusters or genotypes in each cluster in the range of vigor groups *i.e.*, very high, high, medium, low and very low. These seed vigor-related parameters are firstly categorized into seed physiological traits, membrane functions, and antioxidant activities. In each cluster mean data were taken for each parameter and compared. It was found that in cluster 1 consisting of 21 varieties mean data for all important parameters has the highest values (268.86) as compared to another cluster and is considered a very high vigor group and varieties present in this group are also considered as high vigor varieties. Cluster II consisted of eight varieties possessing the second-highest mean (268.83) following cluster I mean and are considered as high vigor group and varieties present in this group are treated as high vigor varieties. Cluster V lies in the medium vigor group (194.72), and varieties in cluster V are considered medium seed vigor varieties. In the low vigor group, cluster III was found and varieties present in cluster III were considered as low vigor varieties (138.21). Cluster IV was considered a very low vigor group (47.30) with the least mean value, and varieties present in this cluster were considered very low

Table 4. Ranking of clusters based on mean value seed vigor traits, membrane function and antioxidant activity

Cluster	Seed vigor trait									Membrane function		Antioxidant activity				Mean	Rank
	SL	SDW	GP	S.V.I	S.V.II	SOG	MGT	MGR	RM	EC	WSS	SOD	DH	HM	AA		
Cluster I	13.531	0.069	98.571	3704.827	6.198	14.430	6.572	0.153	0.085	14.447	32.167	140.563	0.086	0.607	0.722	268.868	1
Cluster II	12.975	0.063	97.000	3643.591	6.654	12.768	6.504	0.154	0.084	15.978	61.975	173.292	0.081	0.738	0.649	268.834	2
Cluster III	9.142	0.048	60.167	1825.463	2.867	10.349	6.837	0.146	0.055	17.113	48.797	90.293	0.037	1.281	0.590	138.212	4
Cluster IV	8.530	0.031	23.667	508.003	0.742	8.592	7.374	0.136	0.043	28.613	61.833	60.037	0.008	1.334	0.541	47.299	5
Cluster V	11.053	0.056	84.833	2638.627	4.783	11.380	6.734	0.148	0.070	25.255	34.962	101.211	0.067	1.011	0.610	194.720	3

SL = Shoot length, SDW = Seedling dry weight, GP = Germination percentage, S.V.I: seed vigor index I, S.V.II = Seed vigor index II, SOG = Speed of germination, MGT = Mean germination time, MGR: mean germination rate, RM = Reserve mobilization, EC = Electrical conductivity, WSS = Water soluble sugars, SOD = Super oxide dismutase, DH: dehydrogenase, AA = alpha amylase and HM = Hydrogen peroxide

vigor varieties (Table 4). This study's findings are consistent with previous research showing seed vigor's importance in crop performance and yield (Finch et al. 2016). Seed vigor is a complex trait that is affected by various factors, including genetics, environmental conditions, and seed treatments (Zhao et al 2021). Therefore, it is essential to understand the underlying mechanisms contributing to seed vigor and identify varieties with high vigor for crop improvement programme. High vigor varieties could lead to increased yield and better crop performance under various environmental conditions.

PCA plots showing relation among various seed vigor traits in rice varieties

In a Scree plot, the eigenvalues of a correlation matrix are plotted in descending order of magnitude to represent the proportion of total variance in the data. When Eigenvalues and principle component numbers were graphed, the Scree plot was applied to explain the percentage of variation associated with each principal component. The ideal pattern is a steep curve followed by a bend and a straight line. PC_1 showed 48.79% variability with an Eigen value of 9.75 (Fig. 4). The Eigenvalues gradually declined from PC_1 to PC_{20} (Fig.4). It is generally recommended to retain principal components with eigenvalues greater than 1 because they explain more variance than a single variable. Based on this recommendation, we would retain only PC_1 and PC_2 as the meaningful principal components in this analysis (Cattell 1966). These two components explain a combined 60.56%

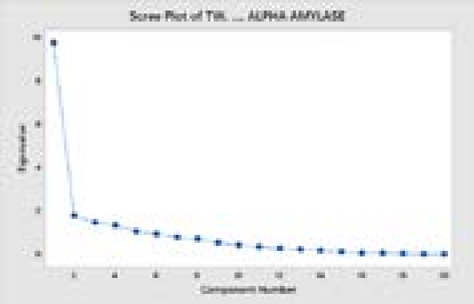


Fig. 4. Scree plot for eigen value and principal component of 44 rice varieties

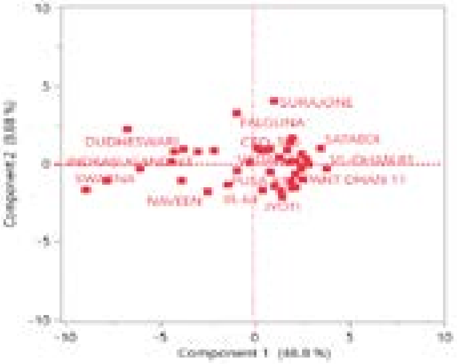


Fig. 5. Score plot for distribution of rice varieties for the first two principal components based on different seed vigor traits

Table 5. Principal component analysis for seed vigor traits in rice varieties

Traits	PC_1	PC_2	PC_3	PC_4	PC_5
Test Weight (g)	2.70	14.58	1.70	2.06	0.011
Moisture Content (%)	-0.90	-7.79	-20.57	-11.81	-7.33
Shoot length (cm)	7.12	0.38	0.27	1.29	5.41
Root length (cm)	4.48	13.12	2.90	3.50	0.17
Seed Dry Weight (mg)	5.80	12.34	2.12	0.15	0.46
Germination (%)	8.86	0.14	0.99	1.56	0.15
Seed vigor index-I	9.16	0.98	1.40	1.08	0.005
Seed Vigor index-II	8.49	3.030	2.25	0.057	0.003
Speed of Germination	7.23	6.053	1.56	1.08	0.028
Mean germination time	-6.94	-10.29	-2.22	-0.0218	-0.129
Mean germination rate	7.44	6.22	1.34	0.001	0.037
Reserve mobilization	8.35	0.088	1.174	0.047	0.231
Electrical conductivity(µS/cm/g)	-2.23	-1.20	-0.015	-.39	-20.11
Water soluble sugars(µg/ml)	-1.221	-0.51	-0.44	-30.26	-25.78
Catalase (µM/min/gFW)	0.96	0.0023	34.38	1.171	1.30
Peroxidase (µM/min/gFW)	9.29E-05	9.90	17.14	15.70	2.57
Superoxide dismutase (µM/min/gFW)	3.59	0.025	0.078	21.006	0.062
Dehydrogenase ($OD{\cdot}g^{-1}{\cdot}m^{-1}$)	8.034	1.39	0.99	0.126	0.00031
Hydrogen peroxide (H_2O_2 µm/g)	-3.94	-11.91	-2.20	-0.142	-12.178
Alpha-Amylase (mg/mL/ g)	2.57	0.056	6.27	0.54	24.049
Eigenvalue	9.76	1.78	1.47	1.35	1.047
Variance (%)	48.80	8.88	7.33	6.77	5.24
Cumulative % of variance	48.80	57.68	65.01	71.78	77.02

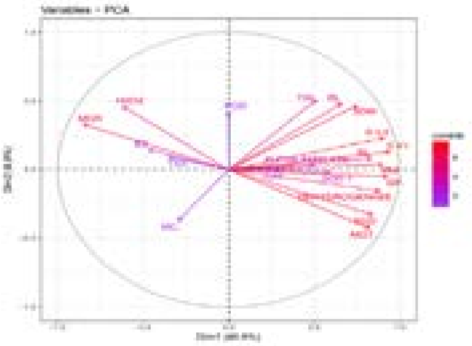


Fig. 6. Loading plot of the first two PCAs showing relation among various seed vigor traits in rice

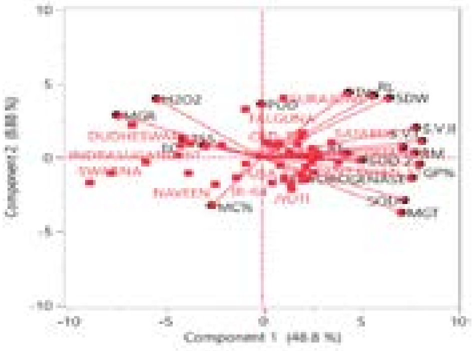


Fig. 7. Biplot of the first two PCAs showing seed vigor traits and varieties of rice

of the total variance in the data.

The genotypes remained scattered in all four quadrants based on all 20 seed vigor traits in the score plot of the first two principal components (Fig. 5). Varieties showed large variability for seed vigor traits. The distance between the locations of any two genotypes on the score plot is directly proportional to the degree of difference or similarity between the vigor in terms of major important seed vigor traits. So, the genotypes in PC_1 have vigor relatively similar for many seed vigor attributes but differ for a few very close to zero attributes. Identifying varieties with high seed vigor traits is essential for ensuring successful crop establishment and optimal plant growth. Principal component analysis (PCA) for seed vigor traits has been previously employed in studies on wheat (Mehra et al. 2021) and soybean (Barbosa et al. 2013), among others, demonstrating its usefulness in

Table 6. Pearson correlation coefficient between 17 quantitative seed vigor traits in rice

	GP	SL	SDW	SVI	SVII	SOG	MGT	MGR	RM	EC	WSS	CAT	POD	SOD	DH	H_2O_2	AA
GP	1																
SL	.716**	1															
SDW	.674**	.590**	1														
SVI	.937**	.809**	.724**	1													
SVII	.894**	.720**	.915**	.906**	1												
SOG	.692**	.712**	.490**	.678**	.634**	1											
MGT	-.639**	-.457**	-.490**	-.606**	-.568**	-.465**	1										
MGR	.626**	.453**	.470**	.596**	.555**	.455**	-.997**	1									
RM	.877**	.769**	.690**	.888**	.863**	.727**	-.562**	.561**	1								
EC	-.443**	-.428**	-.305*	-.422**	-.379*	-.412**	.566**	-.556**	-.398**	1							
WSS	-.335*	-.370*	-0.189	-.356*	-0.266	-.301*	0.012	0.007	-0.264	0.04	1						
CAT	0.267	.416**	0.17	0.296	0.228	.356*	-0.221	0.217	0.267	-0.252	-0.003	1					
POD	0.048	0.123	-0.039	-0.007	-0.02	0.164	0.021	-0.023	-0.019	-0.194	0.009	0.155	1				
SOD	.456**	.571**	.439**	.473**	.492**	.536**	-.425**	.424**	.476**	-.317*	0.039	0.181	0.046	1			
DH	.893**	.736**	.555**	.848**	.782**	.689**	-.635**	.631**	.814**	-.401**	-.354*	0.239	0.018	.527**	1		
H_2O_2	-.584**	-.448**	-0.228	-.494**	-.455**	-.691**	.299*	-0.296	-.551**	0.175	0.059	-0.176	-0.087	-.360*	-.579**	1	
AA	.477**	.334*	.327*	.514**	.438**	.386**	-.396**	.402**	.481**	-0.1	0.025	0.226	0.138	0.295	.486**	-0.278	1

*, **: Correlation is significant at p=0.05 and 0.01, respectively

identifying genotypic variation.

The loading plot shows the distribution of seed vigor traits in the first two principal components (Fig. 6).The loading plot clearly explained the traits which contribute variability in both positive and negative senses. Here the traits come in the right top of PC_1 *viz.*, test weight, seed dry weight, seed vigor index-I, seed vigor index-II, reserve mobilization, germination percent, alpha-amylase, catalase and speed of germination have large positive loading on PC_1 and moisture content has negative loading on PC_2. The maximum loading on the left of the PC_2 is for hydrogen peroxide, mean germination rate, electrical conductivity, and water-soluble sugars, which has less contribution towards variability for seedling vigor traits. The smaller the angle between the two traits more positively correlated to each other and vice versa. Here test weight, seed dry weight, seed vigor index-I, seed vigor index-II, reserve mobilization, germination percent, dehydrogenase, and mean germination rate were positively correlated to each other but mean germination time was negatively correlated with mean germination time, hydrogen peroxide, electrical conductivity, water-soluble sugars and moisture content from first two principal components. It has been shown that seed weight and vigor index are good predictors of seedling vigor and crop establishment (Kandasamy et al. 2020). Similarly, germination percentage and speed of germination have been identified as important traits for evaluating seed quality (Bewley and Black 1994). Electrical conductivity and hydrogen peroxide have also been used as indicators of seed ageing and deterioration (Bailly 2004). The biplot demarcated the varieties with seed vigor traits over the four quadrants explained by the first two principal components (Fig. 7). The varieties in the top right quadrant were closely related to test weight, root length, seed dry weight, seed vigor index-I, seed vigor index-II, reserve mobilization, and catalase. The left top quadrant consists of varieties with related to hydrogen peroxide, mean germination rate, water-soluble sugars, and electrical conductivity. The right bottom quadrant comprised the varieties related to germination percentage, dehydrogenase, speed of germination, superoxide dismutase, and mean germination time, and the left bottom quadrant consisted of varieties related to moisture content. Overall, the biplot analysis provided a useful tool for visualizing the relationships among the varieties and seed vigor traits.

Correlation is a measure of the strength of a linear relationship between the characters. In the present investigation seed vigor was positively correlated with germination %, shoot length, seedling dry weight, speed of germination, mean germination rate, reserve mobilization, superoxide dismutase, dehydrogenase and alpha-amylase and also negatively correlated with mean germination time, electrical conductivity, water-soluble sugars and hydrogen peroxide. (Table 6). Such strong positive correlations recorded among the varieties suggest that they are heritable and genetically controlled traits that could be transmitted into desired varieties. Result of the correlation analysis further showed that characters were differently associated with each other. Adebisi (2008) pointed out that a strong positive correlation suggests that selection for one trait could be used to indirectly select for another character but this can cause difficulties during selection if the association is between desirable and undesirable traits. Seed germination, shoot length, seedling vigor index I and II, speed of germination, seedling dry weight, reserve mobilization, superoxide dismutase, dehydrogenase and alpha-amylase had good correlations with mean germination rate; hence they are good predictors of seed emergence in rice varieties. In general, the seed germination characters have been shown to be an excellent predictor of field emergence and seedling vigor (Egli et al. 1979; Adebisi et al. 2006 and Adebisi 2008) in some crop species. In rice varieties, germination-related characters were reported to be highly significant (Subba Rao 1996). Significant correlation among seed vigor traits was reported by Cui et al. (2002).

Thus, the study's results revealed a large quantity of variability in rice varieties under study. The PC analysis, cluster analysis and correlation coefficient in this present experiment provided facilitation in the classification of genotypes and identification of the subset of varieties having quantitative differences in seed vigor-related parameters. Various useful correlations and aforementioned information extracted from cluster and PC analysis will be helpful in designing crop improvement programmes to obtain high vigor and yielding varieties in rice. PCA is also used for the measurement of the independent impact of a particular trait on the total variance. The dominant characters combining in a certain principal component by contributing to variability has a tendency to combine and offer the possibility for its use in selection.

Authors' contribution

Conceptualization of research (SKC); Designing of the experiments (SKC, SK); Contribution of experimental materials (SKC); Execution of field/lab experiments and data collection (SK,SKC); Analysis of data and interpretation (SK, DP, SKC); Preparation of the manuscript (SK,SKC).

Acknowledgement

The research is part of a Ph.D. thesis of the first author submitted to Post-Graduate School, ICAR-IARI, New Delhi, India. The first author gratefully acknowledges the University Grants Commission (UGC) for providing Senior Research Fellowship during the study period.

References

Abdul-Baki A. A. and Anderson J. D. 1973. Vigor determination in soybean seed by multiple criteria. Crop Sci., **13**: 630-633.

Adebisi, M.A. 2008. Relation of seed testing traits of field performance of sesame genotypes under different environments. ASSET Journal Series A, **8**: 194-204.

Adebisi M.A., Oyekale K.O., Akintobi D.C. and Abdul-Rafiu A.M. 2006. Variation in seed quality of West African rice (*Oryza sativa* (L.). *In.* Proc: 31st Ann. Conf. of Genetic Society of Nigeria: 60-66.

Aebi H. 1984. Catalase *in vitro. In:* Methods in enzymology, vol. 105. Academic Press: United Kingdom: 121-126.

Ali M.A., Jabran K., Awan S.I., Abbas A., Zulkiffal M., Acet T., Farooq J. and Rehman A. 2011. Morpho-physiological diversity and its implications for improving drought tolerance in grain sorghum at different growth stages. Aust. J. Crop Sci., **5**:311-320.

AOSA. 1983. Seed vigor testing handbook: Contribution no. 32 to handbook on seed testing. Association of Official Seed Analyst, Springfield, IL, USA., **33**: 1-93.

Bailly C. 2004. Active oxygen species and antioxidants in seed biology. Seed Sci. Res., **14**: 93-107.

Barbosa R.M., Vieira B.G.T.L., Ferraudo A.S., Corá J.E. and Vieira R.D. 2013. Discrimination of soybean seed lots by multivariate exploratory techniques. J. Seed Sci., **35**: 302-310.

Basavaraj P. S., Gireesh C., Bharamappanavara M., Manoj C. A., Ishwaryalakshmi L. V. G., Senguttuvel P., Sundaram R. M., Subbarao L. V. and Anantha M. S. (2022). Genetic analysis of introgression lines of *Oryza rufipogon* for improvement of low phosphorous tolerance in indica rice. Indian J. Genet. Plant Breed., **82**: 135–142.

Baskin C.C. and Baskin J.M. 2014. Seeds: ecology, biogeography and evolution of dormancy and germination (2nd ed.), Academic Press: San Diego, CA, USA.

Beena R., Veena, V. Jaslam M.P.K., Nithya N. and Adarsh V.S. 2021. Germplasm innovation for high-temperature tolerance from traditional rice accessions of Kerala using genetic variability, genetic advance, path coefficient analysis and principal component analysis. J. Crop Sci. Biotechnol., **24**: 555-566.

Bewley J.D. and Black M.1994. Seeds: germination, structure, and composition. Seeds: physiology of development and germination, Springer, Boston, MA, 1-33.

Bollinedi H., Vinod K. K., Bisht K., Chauhan A., Gopala Krishnan S., Bhowmick Prolay K., Nagarajan D., Rao S., Ellur R. K. and Singh A. K. 2020. Characterising the diversity of grain nutritional and physico-chemical quality in Indian rice landraces by multivariate genetic analyses. Indian J. Genet. Plant Breed., **80**: 26-38.

Castillo F.J., Penel C. and Greppin H. 1984. Peroxidase release induced by ozone in *Sedum album* leaves: involvement of Ca2+. Plant Physiol., **74**: 846-851.

Chatfield C. and A. Collins 1980: Introduction to multivariate analysis. Chapman and Hall.

Ching T.M. 1982. Adenosine triphosphate and seed vigor. The Physiology and biochemistry of seed development, dormancy and germination. Elsevier Biomedical Press, New York, 487-504.

Chozin M. 2007. Characterization of sorghum accessions and choice of parents for hybridization. J Akta Agr Edisi Khusus, **2**:227-232.

Cui K., Peng S., Xing Y., Xu C., Yu S. and Zhang Q. 2002. Molecular dissection of seedling-vigor and associated physiological traits in rice. Theor. Appl. Genet., **105**: 745-753.

Cullis B.R., Smith A.B. and Coombes N.E. 2006. On the design of early generation variety trials with correlated data. J Agr. Biol. Envir. Stat., **11**: 381-393.

Dhindsa R.S., Plumb-Dhindsa P.A. and Thorpe T.A. 1981. Leaf senescence: correlated with increased levels of membrane permeability and lipid peroxidation, and decreased levels of superoxide dismutase and catalase. J. Expt. Bot., **32**: 93-101.

Dubois M., Gilles K. A., Hamilton J. K., Rebers P. T. and Smith F. 1956. Colorimetric method for determination of sugars and related substances. Analytical Chem., **28**: 350-356.

Egli D.B., White G.M. and Tekrony D.M. 1978. Relationship between seed vigor and the storability of soybean seed. J. Seed Technol., **3**: 1-11.

Farooq M., Barsa S.M. and Wahid A., 2006. Priming of field-sown rice seed enhances germination, seedling establishment, allometry and yield. Plant Growth Regul.,**49**: 285-294.

Finch-Savage W.E. and Bassel G.W. 2016. Seed vigor and crop establishment: extending performance beyond adaptation. J. Exp. Bot., **67**: 567-591.

Foolad M.R., Subbiah P. and Zhang L. 2008. Common QTL affect the rate of tomato seed germination under different stress and nonstress conditions. Int. J. Pl. Genomics, vol. 2007. Hindawi Publishing Corporation, doi:10.1155/2007/97386.

Ghafoor A., Sharif A., Ahmad Z., Zahid M.A. and Rabbani M.A. 2001. Genetic diversity in blackgram (*Vigna mungo* L. Hepper). Field Crops Res., **69**: 183-190.

Hilhorst H.W. and Koornneef M., 2007. Dormancy in plants. *In:* Encyclopedia of Life Sciences. Elsevier.doi: 10.1002/9780470015902.a0002045.

International Seed Testing Association. 1985. International rules for seed testing. Rules 1985. Seed Sci. and Tech., **13**: 299-513.

Jones R. and Varner J.E., 1966. The bioassay of gibberellins. Planta, **72**: 155-161.

Kandasamy S., Weerasuriya N., Gritsiouk D., Patterson G., Saldias S., Ali S. and Lazarovits G. 2020. Size variability in seed lot impact seed nutritional balance, seedling vigor, microbial composition and plant performance of common corn Hybrids. MDPI Agronomy, **10**: 157.

Kiran S.Y., Singh V., Mor V.S., Dey S. and Kumar D. 2021. Multivariate analysis of seed vigor parameters in late sown wheat (*Triticum aestivum* L. em. Thell). Int. J. Chem. Stud., **9**: 275-278.

Kittock D.L. and Law A.G., 1968. Relationship of seedling vigor to respiration and tetrazolium chloride reduction by germinating wheat seeds. Agron J., **60**:286-288.

Kumar S., Rajendran K., Kumar J., Hamwieh A. and Baum M. 2015. Current knowledge in lentil genomics and its application for crop improvement. Front. Plant Sci. 6:78. doi: 10.3389/fpls.2015.00078.

Kumari S., Chakrabarty S.K., Paul D., Singh Y., Bhowmick P.K. and Prasad A.S.H. 2023. Variability and heterosis of seed vigor traits in hybrid rice (*Oryza sativa* L.). Indian J. Genet. Plant Breed., **83**: 168-175.

Langewisch T., Lenis J., Jiang G.L., Wang D., Pantalone V. and Bilyeu K., 2017. The development and use of a molecular model for soybean maturity groups. BMC Pl. Biol., **17**: 1-13.

Li W., He X., Chen Y., Jing Y., Shen C., Yang J., Teng W., Zhao X., Hu W., Hu M. and Li H. 2020. A wheat transcription factor positively sets seed vigor by regulating the grain nitrate signal. New Phytol., **225**: 1667-1680.

Maguire J.D. 1962. Speed of germination-aid in selection and evaluation for seedling emergence and vigor. Crop Sci., **2**: 176-177.

Mujaju C. and Chakauya E., 2008. Morphological variation of sorghum landraces from semi-arid areas of Zimbabwe. Int. J. Bot. **4**: 376-382.

Mukherjee S.P. and Choudhuri M.A. 1983. Implications of water stress induced changes in the levels of endogenous ascorbic acid and hydrogen peroxide in *Vigna* seedlings. Physiol. Plant., **58**: 166-170.

Namuco O.S., Cairns J.E. and Johnson D.E. 2009. Investigating early vigor in upland rice (*Oryza sativa* L.): Part I. Seedling growth and grain yield in competition with weeds. Field Crops Res., **113**: 197-206.

Nesselroade J.R. and Cattell R.B. 2013. Handbook of multivariate experimental psychology. Springer Science & Business Media,Springer US,ISBN:9781461308935.

Nichols M.A. 1968. Two approaches to the study of germination data. Proc. Int. Seed Test. Ass., **33**: 531-540.

Nyirenda Yohane E., Shimelis H., Laing M., Mathew I. and Shayanowako A. 2020. Phenotypic divergence analysis in pigeonpea (*Cajanus cajan* (L.) Millspaugh) germplasm accessions. MDPI Agronomy, **10**: 1682.

Qun S., Wang J.H. and Sun B.Q. 2007. Advances on seed vigor physiological and genetic mechanisms. Agri. Sci. China, **6**: 1060-1066.

Rajjou L., Duval M., Gallardo K., Catusse J., Bally J., Job C. and Job D. 2012. Seed germination and vigor. Ann. Rev. Pl. Biol., **63**: 507-533.

Reddy K.R., Brand D., Wijewardana C. and Gao W., 2017. Temperature effects on cotton seedling emergence, growth, and development. Agron. J., **109**: 1379-1387.

Reddy S.K.N., Reddy B.M and Ankaiah R. 1996. Effect of initial germination levels on vigor, growth and yield in rice. Seed Res.,**22**: 22-25.

Sahoo S., Sanghamitra P., Nanda N., Pawar S., Pandit E., Bastia R., Muduli K.C. and Pradhan S.K. 2020. Association of molecular markers with physio-biochemical traits related to seed vigor in rice. Physiol. Mol. Biol. Plants., **26**: 1989-2003.

SAS Institute (1996), SAS/STAT software: Users guide, version 9.2. SAS Inst., Cary, NC.

Singh B., Norvell E., Wijewardana C., Wallace T., Chastain D. and Reddy K.R. 2018. Assessing morphological characteristics of elite cotton lines from different breeding programmes for low temperature and drought tolerance. J. Agron. Crop Sci., **204**: 467-476.

Singh Y.V. 2013. Crop and water productivity as influenced by rice cultivation methods under organic and inorganic sources of nutrient supply. Paddy and Water Environ., **11**: 531-542.

Soltani E., Adeli R., Akbari G.A. and Ramshini H. 2017. Application of hydrotime model to predict early vigor of rapeseed (*Brassica napus* L.) under abiotic stresses. Acta Physiol. Plant., **39**: 1-11.

Subba Rao L.V., Kumar S. and Vanisree G. 1996. Genetic variability for seedling characteristics among rice (*Oryza sativa* L.) cultivars. Seed Res., **24**: 124-128.

Sun Q., Wang J.H. and Sun B.Q. 2007. Advances on seed vigor: physiological and genetic mechanisms. Agri. Sci. China, **6**: 1060–1066.

Talekar S.C., Praveena M.V. and Satish R.G. 2022. Genetic diversity using principal component analysis and hierarchical cluster analysis in rice. Int. J. Pl. Sci., **17**: 191-196.

Watson J.W. 2002. Home Gardens and in Situ Conservation of Plant Genetic Resources in Farming Systems: Proc. 2 Intern. Home Gardens Workshop, 17-19 July 2001, Witzenhausen, Federal Republic of Germany. Bioversity International.

Zhang A., Liu C., Chen G., Hong K., Gao Y., Tian P., Peng Y., Zhang B., Ruan B., Jiang H. and Guo L. 2017. Genetic analysis for rice seedling vigor and fine mapping of a major QTL qSSL1b for seedling shoot length. Breed. Sci., **67**: 307-315.

Zhao J., He Y., Huang S. and Wang Z. 2021. Advances in the identification of quantitative trait loci and genes involved