PLANT SCIENCE TODAY
ISSN 2348-1900 (online)
Vol x(x): xx–xx
https://doi.org/10.14719/pst.6008

HORIZON
e-Publishing Group
HePG

RESEARCH ARTICLE

Multi-trait selection in yellow kernel maize (*Zea mays* L.) genotypes using multi-trait genotype-ideotype distance index

PM Sondarava[1*], MB Patel[2], TH Borkhatariya[1], DJ Parmar[3], KM Akbari[1] & RL Ghetiya[3]

[1]Department of Genetics and Plant Breeding, Anand Agricultural University, Anand 388 110, Gujarat, India

[2]Main Maize Research Station, Anand Agricultural University, Godhra 389 001, Gujarat, India

[3]Department of Agricultural Statistics, Anand Agricultural University, Anand 388 110, Gujarat, India

*Email: prafulsondarva11@gmail.com

ARTICLE HISTORY
Received: 13 November 2024
Accepted: 14 December 2024
Available online
Version 1.0 : 06 February 2025

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

CITE THIS ARTICLE
Sondarava PM, Patel MB, Borkhatariya TH, Parmar DJ, Akbari KM, Ghetiya RL. Multi-trait selection in yellow kernel maize (*Zea mays* L.) genotypes using multi-trait genotype-ideotype distance index. Plant Science Today (Early Access). https://doi.org/10.14719/pst.6008

Abstract

The production of food crops is greatly influenced by maize, which is essential to global food security. Genetic variation and selection are key components in maize breeding that maximize genetic gain and productivity. The present study, 238 maize genotypes were investigated for fourteen quantitative traits to identify diverse and desirable genotypes for future breeding and varietal development programs. Significant genotypic effects were observed for grain yield and its attributes and other agronomic traits, indicating potential for genotype selection. Multivariate PCA analysis revealed that the first four PCs (70.1 % of total variation) effectively captured the considerable diversity within traits. Key traits such as flowering time, plant height, ear height, ear characteristics and grain yield were essential in distinguishing between the genotypes analyzed. A recently introduced multi-trait-ideotype distance index (MGIDI) was used to predict the selection gain and identify the effectively performed genotypes by considering multiple traits. The MGIDI predicts significant desired genetic gains across all characteristics. Strengths and weaknesses of selected genotypes based on MGIDI provided insights into their overall suitability and factor contributions. The genotypes G32, G76, G163, G212 and G169 were identified as performing better using the MGIDI method, considering their strengths and weaknesses for the traits analyzed. MGIDI is a powerful tool that can help breeders effectively select the most desirable genotypes in maize.

Keywords

genetic gain; maize; MGIDI; PCA; strength and weakness; multivariate analysis

Introduction

One of the most adaptable rising crops, maize (*Zea mays* L.), can grow in various agro-climatic circumstances. Maize can play an essential role in crop diversification as it is a dual-purpose crop that produces kernels and fodder for human consumption and livestock feed, respectively (1). Further, it is being exploited for bioethanol production. The production of food crops is greatly influenced by maize, which is essential to global food security. Over 170 countries contribute to global maize production, yielding around 1147.7 million metric tons from an area of 193.7 million ha, with an average productivity of 5.75 tons per ha (2).

Plant breeding targets for the improvement in the genetic makeup of plants, resulting in plant types performing better than older versions. Breeding programs can be more efficiently operated by considering genetic and phenotypic divergence, heritability (narrow sense) and trait association. The kernel yield in maize is the outcome of the yield attributes. Although the primary goal of any crop improvement programme is to increase the economic yield, yield attributes are also important as they directly or indirectly affect yield. Recent studies revealed the importance of key traits, viz., cob length, days to 50 % tasseling, number of cobs per plant, cob weight, kernel rows per cob and cob girth. With the desired correlation coefficients and positive effects on kernel yield (3-4). Breeders can find superior genotypes with desirable qualities by analyzing variance in these variables. Enhancing efficiency by selecting genotypes based on grain yield and agronomic traits is hindered by challenges in effectively incorporating multiple trait data without multicollinearity. Precise statistical approaches reduce and simplify the multivariate data dimensions that operate an efficient selection of desirable traits and genotypes for hybrid breeding (5). Univariate statistical techniques ignore the correlation among the variables and sometimes, the results may not be valid. Multivariate methods like principal component analysis, factor analysis and multivariate clustering consider the inter-dependence and relative importance of various traits involved and display more meaningful genetic information (6).

Principal Component Analysis (PCA) is a data reduction technique that retains essential information while reducing dimensionality. This facilitates easier visualization and interpretation of the data, making it widely used in capturing trait diversity in various crops, including maize (7-10). The Smith-Hazel (SH) index (11-12), a commonly employed multi-trait selection method, seeks to enhance genetic gain for grain yield. However, studies (13-18) point out problems such as biased index coefficients and multicollinearity, which reduce genetic gains and limit their use in plant breeding. The Multi-trait index proposed by (13) is based on factor analysis and ideotype-design principles. This index integrates multiple traits into a single value, allowing for a comprehensive assessment of genotype performance.

A multi-trait genotype-ideotype distance index (MGIDI) was introduced to select genotypes with desirable mean performances, addressing classical index fragility (19). The MGIDI can be employed to select the desired genotypes within the environment. Few earlier studies (20-24) have reported using the MGIDI index for selecting maize inbreds and hybrids based on multiple traits. Hence, an effort was made to unravel the principal components for determining potential secondary traits through indirect selection and (select the superior genotype by considering multiple traits using MGIDI.

Materials and Methods

Climate of the site

The experiment was conducted in the research field of Main Maize Research Station (MMRS), Anand Agricultural University, Godhra, from Rabi 2020-21. Geographically, it is situated at 22° 45' North latitude and 77° 40' East longitude, with an altitude of 119 meters above mean sea level. The soil at the experimental site in Godhra was sandy loam containing 80 % sand and 20 % clay, characterized by medium to low productivity and responsiveness to fertilizers, with a composition of 80 % sand and 20 % clay. It has a very low water-holding capacity and the soil depth is approximately 2.5 feet. Soil nitrogen availability is 0.043-0.046 % (low), medium in phosphorus and high in potash. The climate of the Godhra (middle Gujarat) region is semi-arid and tropical. The temperature of the experimental location is extreme, such as 6.0 ° C in winter and rising to 44.0 °C during summer. The average annual rainfall ranges between 650 and 750 mm over 32 rainy days. The meteorological data were recorded at the Meteorological Observatory, MMRS, Anand Agricultural University, Godhra, during the experimental season of maize.

Selection of the genotypes

The experimental materials were procured from the MMRS, Anand Agricultural University (AAU), Godhra. The list of genotypes is presented in Table 1. A total of 238 yellow kernel maize genotypes were used in the study.

Experimental design and cultural practices

The 238 genotypes (G) were tested in a randomized complete Block Design (RCBD) with three replicates. A border row surrounded the experiment to minimize potential damage and eliminate border effects. Standard agronomic practices and plant protection measures were implemented to ensure successful and healthy crop production.

Data collection and observations

The observations on fourteen quantitative characters were recorded in each row on five randomly selected competitive plants (tagged) from each replication. The mean value was computed for statistical analysis except for days to 50 % tasselling (FT), days to 50 % silking (FS) and days to maturity (DM), which were recorded on a row basis. Plant height (PH) and Ear height (EH) were measured after reproductive maturity. Cob-related traits (CPP: Cobs per plant, KRPC: Kernel rows per cob, KPR: Kernels per row, CL: Cob length, CG: Cob girth, CW: Cob weight, 1000K: 1000 kernel weight, SH: Shelling percent) and Kernel yield per plant (KYPP) were recorded at the physiological maturity of maize.

Statistical analysis

The mean data were primarily subjected to Analysis of Variance (ANOVA). The principal component (PC) analysis was employed to extract the principal components. The research used an MGIDI to select the most optimal genotypes, utilizing data from multiple traits (25). This approach integrates various analytical techniques, including trait scaling, exploratory factor analysis (EFA) and computation of the MGIDI. In the initial step of the analysis, each trait underwent scaling to ensure uniformity in their ranges (25). The scaling equation employed was the following in Equation 1

$$rXij = \frac{(\eta nj - \phi nj)}{(\eta oj - \phi nj)} (\theta nj - \eta nj) + \eta nj \qquad \text{(Eqn. 1)}$$

In this context, φ_{oj} and η_{oj} denote the minimum and maximum original values for the j^{th} trait, respectively, while φ_{nj} and η_{nj} represent the rescaled minimum and maximum values for the j^{th} trait. The original value for the j^{th} trait of the ith genotype is indicated by θ_{ij}. The rescaled values, φ_{nj} and η_{nj}, were chosen based on the desired improvements for each trait: for traits with positive gains (KYPP, 1000K, KPR, SH, KRPC, CL, CG, CW), $\varphi_{nj} = 0$ and $\eta_{nj} = 100$ were assigned, while for traits with negative gains (FT, FS and DM), $\varphi_{nj} = 100$ and $\eta_{nj} = 0$ were used, as recommended by (25). After scaling the traits, Factor Analysis (FA) was performed to reduce the data dimensions and uncover the underlying relationships among the characteristics. The FA model utilized for this analysis was as follows in Equation 2.

Table 1. List of maize genotypes (G) used in the investigation

No.	Genotype	No.	Genotype	No.	Genotype	No.	Genotype
1	CML-269	62	IL-14-61	122	IL-17-04	183	IL-17-78
2	CML-296	63	IL-14-62	123	IL-17-05	184	IL-17-79
4	CML-307	64	IL-14-63	124	IL-17-06	185	IL-17-80
5	CML-482	65	IL-14-64	125	IL-17-09	186	IL-17-81
6	CM-111	66	IL-14-67	126	IL-17-10	187	IGI-19-01
7	CM-135	67	IL-14-68	127	IL-17-12	188	IGI-19-02
8	CM-140	68	IL-14-75	129	IL-17-14	189	IGI-19-03
9	CM-212-2	69	IL-14-77	130	IL-17-15	190	IGI-19-04
10	CM- 500-1	70	IL-14-78	131	IL-17-17	190	IGI-19-04
11	CM-500-2	71	IL-14-80	132	IL-17-19	191	IGI-19-05
12	GYL-1	72	IL-14-81	133	IL-17-20	192	IGI-19-06
13	GYL-2	73	IL-14-85	134	IL-17-21	193	IGI-19-07
14	GYL-4	74	IL-14-86	135	IL-17-22	194	IGI-19-08
15	GYL-5	75	IL-12-06	136	IL-17-24	195	IGI-19-09
16	GYL-6	76	IL-12-11	137	IL-17-25	196	IGI-19-10
17	GYL-7	77	IL-12-12	138	IL-17-26	197	IGI-19-11
18	GYL-8	78	IL-12-16	139	IL-17-27	198	IGI-19-12
19	GYL-9	79	IL-12-17	140	IL-17-28	199	IGI-19-13
20	GYL-10	80	IL-12-19	141	IL-17-29	200	IGI-19-14
21	GYL-11	81	IL-12-23	142	IL-17-30	201	IGI-19-15
22	IGI-19-54	82	IL-12-26	143	IL-17-31	202	IGI-19-16
23	IT-INA-011-2	83	IL-12-29	144	IL-17-32	203	IGI-19-17
24	HKI 3-4-8-5 ER	84	IL-12-38	145	IL-17-33	204	IGI-19-18
25	HKI -163	85	IL-12-40	146	IL-17-34	205	IGI-19-19
26	HKI-193-1	86	IL-12-51	147	IL-17-37	206	IGI-19-20
27	PFSR-S3-5	87	IL-12-52	148	IL-17-38	207	IGI-19-21
28	VL-78	88	IL-12-58	149	IL-17-39	208	IGI-19-22
29	H-07R-04-3	89	IL-12-86	150	IL-17-41	209	IGI-19-23
30	LM-5	90	IL-12-95	151	IL-17-43	210	IGI-19-24
31	H-07-R-01-3	91	IL-12-97	152	IL-17-44	211	IGI-19-25
32	NAI-105-5	92	IL-12-101	153	IL-17-45	212	IGI-19-26
33	LTP 1-1	93	IL-12-109	154	IL-17-46	213	IGI-19-27
34	LM13-2	94	IL-12-123	155	IL-17-47	214	IGI-19-28
35	IL-14-10	95	IL-12-129	156	IL-17-48	215	IGI-19-29
36	CM500	96	IL-12-135	157	IL-17-49	216	IGI-19-30
37	H07R-9-3	97	IL-15-01	158	IL-17-50	217	IGI-19-31
38	H07-49-3	98	IL-15-02	159	IL-17-51	218	IGI-19-32
39	H07R-4-2	99	IL-15-03	160	IL-17-52	219	IGI-19-33
40	H07R-4-3	100	IL-15-13	161	IL-17-53	220	IGI-19-34
41	LM-3	101	IL-15-19	162	IL-17-54	221	IGI-19-35
42	IL-14-01	102	IL-15-20	163	IL-17-55	222	IGI-19-36
43	IL-14-03	103	IL-15-21	164	IL-17-56	223	IGI-19-37
44	IL-14-11	104	IL-15-23	165	IL-17-57	224	IGI-19-38
45	IL-14-15	105	IL-15-24	166	IL-17-58	225	IGI-19-39
46	IL-14-23	106	IL-15-25	167	IL-17-59	226	IGI-19-40
47	IL-14-24	107	IL-15-26	168	IL-17-60	227	IGI-19-41
48	IL-14-26	108	IL-15-27	169	IL-17-61	228	IGI-19-42
49	IL-14-34	109	IL-15-28	170	IL-17-62	229	IGI-19-43
50	IL-14-35	110	IL-15-29	171	IL-17-63	230	IGI-19-44
51	IL-14-38	111	IL-15-30	172	IL-17-64	231	IGI-19-45
52	IL-14-39	112	IL-15-31	173	IL-17-66	232	IGI-19-46
53	IL-14-40	113	IL-15-34	174	IL-17-67	233	IGI-19-47
54	IL-14-47	114	IL-15-35	175	IL-17-68	234	IGI-19-48
55	IL-14-48	115	IGI-19-53	176	IL-17-69	235	IGI-19-49
56	IL-14-51	116	IL-15-38	177	IL-17-70	236	IGI-19-50
57	IL-14-52	117	IL-15-39	178	IL-17-71	237	IGI-19-51
58	IL-14-53	118	IL-15-41	179	IL-17-72	238	IGI-19-52
59	IL-14-56	119	IL-17-01	180	IL-17-73		
60	IL-14-57	120	IL-17-02	181	IL-17-76		
61	IL-14-59	121	IL-17-03	182	IL-17-77		

$$F = Z(A^T R^{-1})^T \qquad \text{(Eqn. 2)}$$

Where F is a $g \times f$ matrix containing the factorial scores, Z is a $g \times p$ matrix representing the rescaled means, A is a $p \times f$ matrix of canonical loadings and R is a $p \times p$ correlation matrix between the measured traits. Here, *g, f* and *p* represent the number of genotypes, the factors retained and the measured characteristics. Afterwards, an ideal genotype (ideotype) was created by assigning the highest rescaled value (100) to all the analyzed traits. The MGIDI was then computed as the Euclidean distance between the genotype scores and the ideotype using the following formula as described in Equation 3:

$$MGIDI = \sum_{j=1}^{f} [(\gamma ji - \gamma j)^2]^{0.5} \qquad \text{(Eqn. 3)}$$

In this context, γ_{ij} represents the score assigned to the i^{th} genotype for the j^{th} factor, where t refers to the total number of genotypes and f indicates the number of factors considered. The score corresponding to the j^{th} trait for the ideal genotype is denoted as γ_j.

The MGIDI was calculated by measuring the Euclidean distance between the genotype scores and those of the ideal genotype. This index quantifies how closely each genotype matches the ideal across multiple traits. The genotype with the lowest MGIDI is considered the closest to the ideal genotype, indicating that it best matches the desired values across all measured traits. In the selection process, all fourteen traits were considered, with a selection intensity of 5 %.

Furthermore, the strengths and weaknesses of each genotype were evaluated by calculating the proportion of the MGIDI index attributed to each factor. This was done by determining how much of the MGIDI score for each genotype (i^{th} row) was explained by each factor (j^{th} factor), as estimated using the following method in Equation 4 (25):

$$\omega_{ij} = \frac{\sqrt{D_{ij}^2}}{\sum_{j=1}^{f}\sqrt{D_{ij}^2}} \qquad \text{(Eqn. 4)}$$

Here, D_{ij} represents the distance between the i^{th} genotype (row) and the MGIDI for the j^{th} factor. A lower contribution from a factor indicates that the traits associated with that factor closely resemble those of the ideal genotype.

The selection differentials were calculated using the MGIDI values, with a selection intensity of about 5%. The selection differential, expressed as a percentage of the population mean (DS %), was then determined for each trait using the following formula in Equation 5.

$$\Delta SI\% = \frac{(Xs - Xo)}{Xo} \times 100 \qquad \text{(Eqn. 5)}$$

Where X_s represents the average performance value of the selected genotypes and X_0 refers to the mean performance of the overall (original) population.

Statistical software

All statistical analyses were performed using R Studio, running R version 4.1.2 software (26). Base package "stats" was used to extract principal components. The "metan" R package was explored within the R Studio environment for MGIDI calculation (27). Additionally, the "ggplot2" package v3.3.3 (28) was utilized for data visualization.

Results

Computation of variance components

The variance components were estimated using the restricted maximum likelihood (REML) approach alongside Best Linear Unbiased Prediction (BLUP) within a mixed-effects model. Genotype was modelled as a random effect, while replication was treated as a fixed effect. The results of the likelihood ratio tests indicated that genotype effects were statistically significant for all traits, with significance levels of $p < 0.05$.

Principal component analysis

The PCA results retained four out of fourteen PCs after varimax rotation with Kaiser normalization, which accumulated 70.1 % variance in the total variance in the dataset (Table 2). These first four PCs had the eigenvalue>1 (29) and were considered for the result interpretation. The first PC explained the highest proportion of variance, *i.e.*, 32.5% and included the three important traits *viz*., days to maturity, days to 50 % silking and days to 50 % tasselling, suggesting their strong correlation and potential contribution to the variability captured by this component. Cob dimensions (cob length, cob girth, cob weight, 1000 kernel weight and shelling per cent) were grouped into the second PC which explained 17.11 % variance of total variation. The third and fourth PC accounted 11.6 % and 08.8 % variance of total variability respectively. Plant height, ear height, kernel rows per cob and kernels per row were grouped into PC3 while cobs per plant and kernel yield per plant were retained into the PC4. Recent studies support the relevance of PCA in understanding trait associations and variability in maize crops. (30) utilized PCA to identify key traits influencing maize kernel yield and highlighted the importance of traits such days to 50 % tasselling and days to 50 % silking which aligns with our PCA results. Similarly, (31) showed that flowering time, plant height, ears traits and yield were the most discriminatory traits emphasizing the utility of PCA in trait characterization and breeding programs.

Table 2. Principal for fourteen quantitative traits showing first four PCs, eigen values and proportion of variance explained

Traits	PC1	PC2	PC3	PC4
Days to 50% tasseling	0.981	-0.112	-0.086	0.006
Days to 50% silking	0.982	-0.112	-0.088	0.000
Days to maturity	0.978	-0.107	-0.087	-0.011
Plant height	-0.055	0.017	0.870	0.090
Ear height	-0.055	0.064	0.858	0.006
Cobs per plant	0.055	-0.100	0.112	0.918
Kernel rows per cob	-0.006	0.070	0.531	0.151
Kernels per row	-0.247	0.386	0.597	0.030
Cob length	-0.008	0.775	0.164	-0.014
Cob girth	-0.166	0.585	0.411	0.115
Cob weight	-0.284	0.754	0.284	0.147
1000 kernel weight	0.107	0.715	-0.081	-0.002
Shelling per cent	-0.191	0.520	-0.001	0.044
Kernel yield per plant	-0.085	0.455	0.207	0.764
Eigen value	4.548	2.396	1.629	1.237
Proportion of variance explained	0.325	0.171	0.116	0.088
Cumulative proportion of Variance (%)	32.5	49.6	61.2	70.1

Multi-trait genotype selection using the MGIDI index

Loadings and factor delineation

Factor analysis is a statistical measure employed to uncover patterns and relationships among a set of variables. To enhance interpretability and data simplicity, the fourteen traits under analysis were grouped into four factors, referred to as FA. The factor loadings and communalities derived from this analysis are shown in Table 3. The communalities of the variables ranged from 0.31 for the shelling (%) and Kernel row per cob to 0.99 for

Table 3. The factorial loadings and communalities resulting from factor analysis

VAR	FA1	FA2	FA3	FA4	Communality	Uniqueness
FT	-0.11	-0.98	-0.09	-0.01	0.98	0.02
FS	-0.11	-0.98	-0.09	0	0.99	0.01
DM	-0.11	-0.98	-0.09	0.01	0.97	0.03
PH	-0.02	-0.05	-0.87	0.09	0.77	0.23
EH	-0.06	-0.05	-0.86	0.01	0.74	0.26
CPP	0.1	0.05	-0.11	0.92	0.87	0.13
KRPC	-0.07	-0.01	-0.53	0.15	0.31	0.69
KPR	-0.39	-0.25	-0.6	0.03	0.57	0.43
CG	-0.77	-0.01	-0.16	-0.01	0.63	0.37
CL	-0.59	-0.17	-0.41	0.12	0.55	0.45
CW	-0.75	-0.28	-0.28	0.15	0.75	0.25
1000K	-0.71	0.11	0.08	0	0.53	0.47
SH	-0.52	-0.19	0	0.04	0.31	0.69
KYPP	-0.46	-0.08	-0.21	0.76	0.84	0.16
Communality Mean				0.700771		

trait Days to 50 % silking. Similar value of communality for SH was reported by (23). FA1 predominantly represents cob traits crucial for maize kernel yield, including cob girth (CG), length (CL), weight (CW), 1000 kernel weight (1000K) and shelling percentage (SH), evidenced by their strong negative loadings (-0.77, -0.59, -0.75, -0.71 and -0.52, respectively). FA2 primarily represents agro-morphological traits, including days to 50 % tasselling (FT), silking (FS) and maturity (DM), evidenced by their strong negative loadings (-0.98). FA3 predominantly represents developmental and yield attributes, including plant height (PH), ear height (EH), kernel row per cob (KRPC) and kernel per row (KPR), as indicated by their strong negative loadings (-0.87, -0.86, -0.53 and -0.6, respectively). FA4 primarily represents kernel yield per plant (KYPP) and cobs per plant (CPP), with positive loadings ranging from 0.76 to 0.92.

Predicted selection gains

The MGIDI index was utilized to assess all measured traits comprehensively. The process involved normalizing the traits using BLUP to estimate the mean performance of each genotype, followed by conducting factor analysis and calculating the genetic distance between the genotypes and the ideal genotype. The predicted genetic gains for the relevant traits based on the MGIDI index are displayed in Table 4. The broad-sense heritability values observed in this study showed substantial variation, ranging from 0.27 for days to maturity to over 0.6 for traits such as cob weight, 1000-kernel weight, plant height, ear height, kernel rows per cob, kernels per row, cobs per plant and kernel yield per plant.

MGIDI-based findings within the FA1 framework indicate a 73.1% increase in cob weight followed by an 8.42 % elevation in 1000 kernel weight. Ear length is expected to extend by 6.42 %, enhancing yield potential. Additionally, a marginal gain of 0.948 % in shelling (%) is projected, contributing to economic yield enhancement. Days to tasselling, days to 50 % silking and days to maturity exhibited a marginal decline of -0.515, -0.486 and -0.19, respectively, indicative of relatively stable performance in this trait within the context of FA2. Within FA3, an anticipated expansion of 26.1 % in kernel per row followed by an increase of 16.9 % in ear height.

Furthermore, the trait of plant height, crucial for overall crop stability, is expected to rise by 9.88 %. Moreover, a significant gain of 5.77 % was observed for kernel row per cob, a key influencer of yield. The highest genetic gain within FA4 was observed in kernel yield per plant, with an impressive increase of 87.7 %. This was closely followed by a substantial gain of 23.7 % for cobs per plant, crucial for enhancing grain yield per plant. Few attempts have been made in maize to select desired genotypes using MGIDI compared to other indices and similar results of high genetic gain in maize using MGIDI have been reported (32).

Selected genotypes

Assuming a controlled selection intensity of 5%, various genotypes were screened using MGIDI. Fig. 1. visually illustrates genotype rankings based on their MGIDI index values, with specific genotypes close to the red cutting point, indicating excellent phenotypes. Genotypes 76, 212, 163, 101, 195, 169, 115, 233, 194, 79, 32 and 104 emerged as distinguished performers.

Table 4. Predicted selection gain using MGIDI index

VAR	Factor	X_o	X_s	SD	SD(%)	h^2	SG	SG(%)	Sense	Goal
CG	FA1	10.9	11.4	0.528	4.86	0.39	0.206	1.9	increase	100
CL	FA1	14.2	15.9	1.65	11.6	0.55	0.914	6.42	increase	100
CW	FA1	55.68	97.6	42.6	77.5	0.94	40.2	73.1	increase	100
1000K	FA1	218.3	241	23.2	10.6	0.79	18.4	8.42	increase	100
SH	FA1	70.4	72.1	1.7	2.41	0.39	0.667	0.948	increase	100
FT	FA2	75.9	74.8	-1.05	-1.38	0.37	-0.391	-0.515	decrease	100
FS	FA2	78.00	77.00	-1.03	-1.32	0.37	-0.379	-0.486	decrease	100
DM	FA2	105	104	-0.745	-0.708	0.27	-0.2	-0.19	decrease	100
PH	FA3	151	167	16	10.6	0.93	14.9	9.88	increase	100
EH	FA3	76.5	90.1	13.6	17.8	0.95	13	16.9	increase	100
KRPC	FA3	12.8	13.8	0.944	7.35	0.79	0.741	5.77	increase	100
KPR	FA3	22.5	29.1	6.54	29	0.9	5.89	26.1	increase	100
CPP	FA4	1.33	1.65	0.319	24	0.99	0.316	23.7	increase	100
KYPP	FA4	67.1	128	61.1	91.2	0.96	58.8	87.7	increase	100

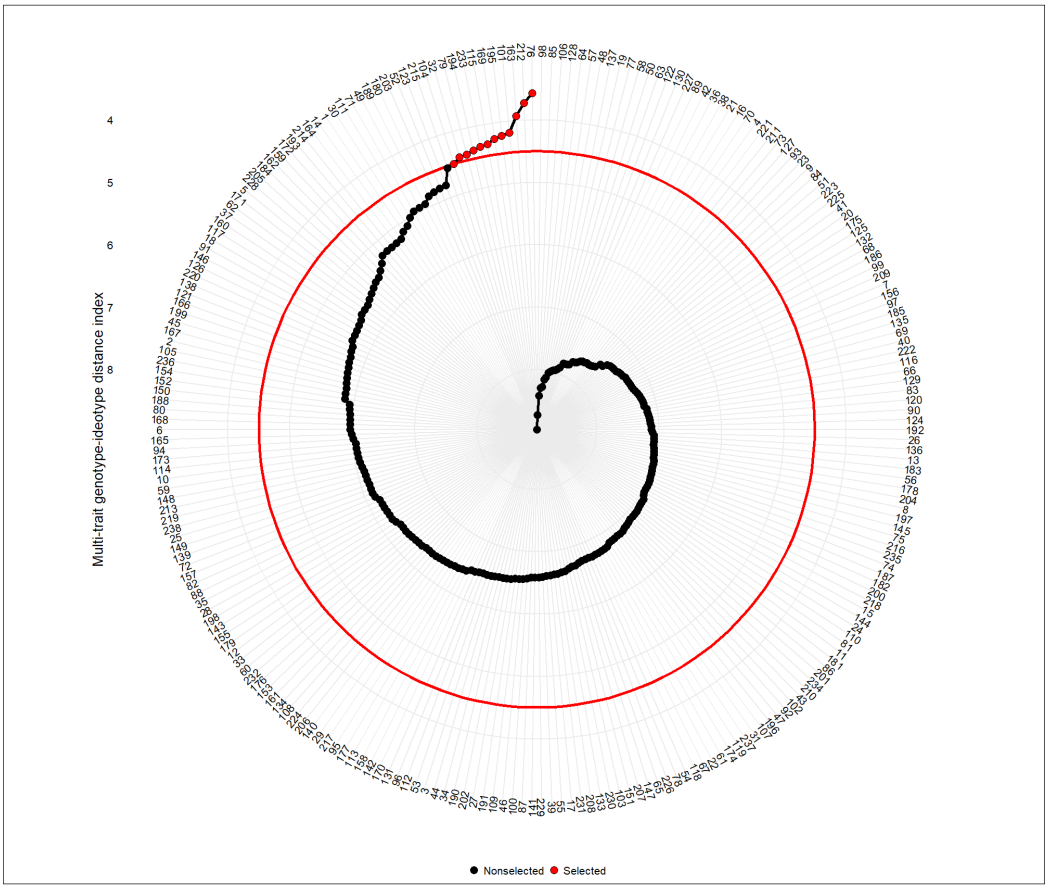

Fig. 1. Maize genotypes ranking based on MGIDI index with highlighted top performers in red.

Moreover, the MGIDI offers crucial insights into the strengths and weaknesses of various genotypes, providing a helpful framework for identifying their advantages and limitations within the complex landscape of multiple traits (25).

Strength and weakness

The radar plot (Fig. 2) illustrates the delineation of genotype attributes, segmenting MGIDI contributors into four key categories. The analysis of strengths and weaknesses indicated that FA1 significantly influenced the MGIDI for all selected genotypes, suggesting relatively poorer performance in key attributes like cob girth, length, weight, 1000 kernel weight and shelling percentage. Genotypes G32, G212, G163 and G195 demonstrated strengths associated with FA2, particularly in flowering characteristics such as days to maturity, days to 50% silking and 50% tasselling. Notably, these traits exhibited negative gains, indicating potential for improvement in these genotypes. FA3 had a relatively lower impact on genotypes G104, G76, G212, G163, G101, G195, G115 and G79, indicating their strong performance in traits associated with FA3, highlighting their suitability for traits *viz*., ear height, plant height, kernel row per cob and kernel per row. FA4 significantly impacted genotypes G32, G104, G76, G212, G169, G233 and G194, particularly influencing traits related to cobs per plant and kernel yield per plant. These genotypes exhibited notable responses to FA4, indicating their potential for enhancing overall yield and productivity traits in maize cultivation. A comprehensive evaluation of multiple traits ranked genotypes G32, G76, G163, G212 and G169 as the top performers among the selected genotypes. Recently, (20) used this index to identify the desired hybrids in maize through their strength and weakness view.

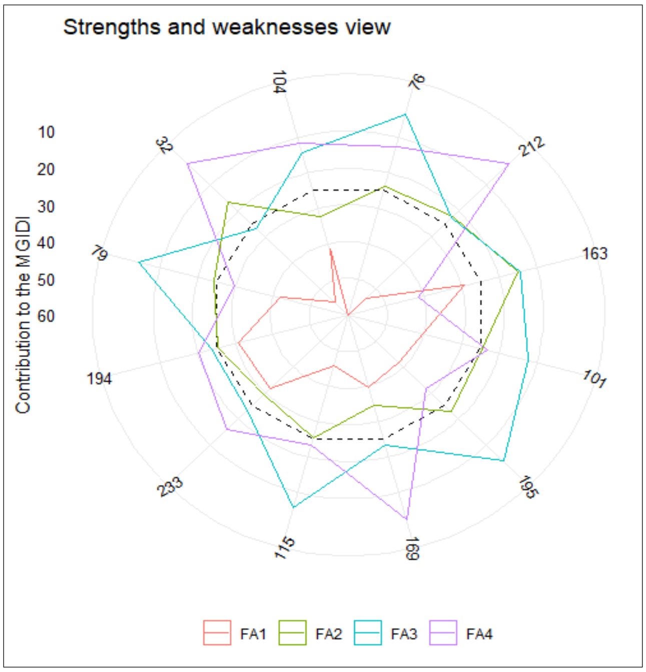

Fig. 2. Comprehensive assessment of selected genotypes, highlighting their strengths and weaknesses weighed by MGIDI.

Discussion

Selection in plant breeding is a fundamental process where breeders choose plants with desirable traits to serve as parents for the next generation. This process aims to enhance yield, disease resistance, drought tolerance and quality. This process necessitates genetic diversity and heritable traits in the population. Maintaining genetic diversity ensures long-term progress and avoids inbreeding depression. Highly heritability traits respond better to selection as the genetic variance is higher than environmental variance. The selection based on a single trait may reduce the selection gain; therefore, selection by considering multiple traits may enhance the genetic gains and success of the breeding programme. The present study investigated 238 maize genotypes based on fourteen quantitative characteristics and we found significant differences in the genotype performance of all evaluated traits. This opens the scope for genotype selection.

Principal component analysis

Principal Component Analysis (PCA) is a statistical technique used to reduce the dimensionality of data while preserving most of the variability present in the data set. Each principal component (PC) is a linear combination of the original variables and the loadings represent the weights assigned to each variable in that linear combination. Interpreting the loadings involves understanding how strongly each original variable contributes to a particular principal component. High loadings (either positive or negative) indicate strong correlations between the original variables and the principal component. Variables that exhibit high loadings on a specific principal component are deemed significant in elucidating the variation represented by that component. PCA was conducted to classify the genotypes based on the most distinguishing traits. Key traits such as flowering time, plant and ear height, ear characteristics and grain yield were essential in distinguishing between the genotypes analyzed and four PCs effectively captured the considerable amount of diversity within traits.

Multi-trait genotype selection using the MGIDI Iindex

The multivariate techniques are widely used for genotypes and trait selections in plant breeding. We have employed the recently developed MGIDI index to select the high-performance genotypes in maize. This process includes exploratory factor analysis and then computation of the MGIDI. The exploratory factor analysis explained fourteen traits into four factors. An average communality of 0.70 suggests that the factors (FA) can account for a significant portion of the variance associated with each variable. This communality reflects the shared characteristics or traits among different genotypes, while the average uniqueness represents the specific features or traits unique to each genotype. High communalities in FA1, FA2, FA3 and FA4 highlight the interconnectedness of the characteristics, emphasizing their shared dependence on underlying factors. Negative loadings in FA1, FA2 and FA3 indicate an inverse relationship between the respective FA traits, suggesting that underlying variables decrease as FA scores increase. A positive correlation between these variables and FA4 demonstrates that as the scores for FA4 increase, both kernel yield per plant (KYPP) and cobs per plant (CPP) also rise. These findings provide valuable insights into the complex architecture of plant traits, informing potential trait groupings and interrelationships. The MGIDI predicts significant genetic gains across all characteristics, including kernel yield per plant, cob weight, kernel per row, cobs per plant and ear height. With its enhanced consistency alongside other genotype ranking methods, such as the SH index (13-15), the MGIDI is poised to find broader applications in future breeding programs for genotype screening.

In the strength and weakness plot, more significant influence factors are positioned proximal to the centre, while lesser contributors align towards the periphery. This aids in identifying suitable parentage for hybridization initiatives. The radar plot for each chosen genotype outlines the MGIDI contributors, showcasing the trait influence hierarchy. This assists in aligning traits with the ideotype, as evidenced by factors nearer the plot's periphery. Through a thorough assessment of multiple traits, genotype ranking identified G32, G76, G163, G212 and G169 as the highest performers among the selected genotypes, indicating their potential for enhancing traits related to overall yield and productivity in maize cultivation. The utilization of the MGIDI index in plant crop research is anticipated to experience rapid growth. (22,23,33) employed this index to identify the best maize genotypes through their strength and weaknesses in stress conditions. The integration of grain yield with other target attributes facilitated genotype ranking, highlighting the significance of assessing multiple traits collectively. This approach, supported by (34), emphasizes evaluating genotypes based on their capacity to harmonize various traits, offering a more comprehensive measure of their suitability in screening or breeding programs.

Conclusion

In conclusion, this study pinpointed the maize genotypes with superior agronomic traits through a comprehensive multi-trait selection model. Key factors like flowering time, plant height, ear height and grain yield emerged as critical in differentiating the genotypes, with the first four principal components capturing significant variation in the trait studied. The use of the MGIDI method provided a nuanced evaluation of genotype strengths and weaknesses, guiding the selection of top-performing genotypes such as G32, G76, G163, G212 and G169. This approach enhances the accuracy of genotype selection and offers breeders a more informed and data-driven method for improving maize breeding programs.

Acknowledgements

We thank Anand Agricultural University for providing research facilities.

Authors' contributions

PM carried out the entire research, writing-original draft preparation, data analysis and visualization. MB conceptualized and supervised the study and reviewed and edited the manuscript. TH and KM participated in research methodology, manuscript preparation and editing. DJ and RL carried out the statistical analysis and visualization. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interest to declare.

Ethical issues: None

References

1. Borkhatariya TH, Gohil DP, Sondarava PM, Patel R, Akbari KM. Character association and path co-efficient analysis among diverse genotypes of forage maize (*Zea mays* L.). Biol Forum Int J. 2022;14 (3):829-33.
2. Food and Agriculture Organization. Statistical Database [Internet]; 2020 [cited:12 Nov, 2024] Rome: FAO. Available from: http://faostat.fao.org
3. Chauhan Z, Shah SN, Chaudhary CS, Rahevar P. Correlation study of yield attributing traits in maize (*Zea mays* L.). Int J Agri Appli Sci. 2022;3(2):9-21. https://doi.org/10.52804/ijaas2022.323
4. Sondarava PM, Patel MB, Borkhatariya TH, Akbari KM, Parmar PK. Correlation and path coefficient analysis of quantitative traits in maize (*Zea mays* L.). Int J Stat Appli Math. 2023;SP-8(6):1140-44. https://doi.org/10.22271/maths.2023.v8.i6So.1516
5. Piepho HP, Möhring J. Computing heritability and selection response from unbalanced plant breeding trials. Genetics. 2007;177 (3):1881-88. https://doi.org/10.1534%2Fgenetics.107.074229
6. Singh M, Sharma SK, Singh TP, Dutta M. Factor analysis of components of yield and some growth parameters in urdbean (*Vigna mungo* (L.) Hepper). Indian J Plant Genet Res. 2011;24(3):346-8
7. Al-Naggar AMM, Shafik MM, Musa RYM. Genetic diversity based on morphological traits of 19 maize genotypes using principal component analysis and GT biplot. Annual Research and Review in Biology. 2020;35(2):68-85. https://doi.org/10.9734/arrb/2020/v35i230191
8. Mengistu S. Maize germplasm characterization using principal component and cluster analysis. Amer J Bio-Sci. 2021;9:122. https://doi.org/10.11648/j.ajbio.20210904.12
9. Tang H, Xu C, Jiang Y, Wang J, Wang Z, Tian L. Evaluation of physical characteristics of typical maize seeds in a cold area of North China based on principal component analysis. Processes. 2021;9(7):1167. https://doi.org/10.3390/pr9071167
10. Sinana HF, Ravikesavan R, Iyanar K, Senthil A. Study of genetic variability and diversity analysis in maize (*Zea mays* L.) by agglomerative hierarchical clustering and principal component analysis. Elect J Plant Breed. 2023;14(1):43-51. https://doi.org/10.37992/2023.1401.015
11. Smith H. A discriminant function for plant selection. Annals of Eugenics. 1936;7:240-50. https://doi.org/10.1111/j.1469-1809.1936.tb02143.x
12. Hazel LN. The genetic basis for constructing selection indexes. Genetics. 1943;28:476-90. https://doi.org/10.1093%2Fgenetics%2F28.6.476
13. Rocha JRDASDC, Machado JC, Carneiro PCS. Multi-trait index based on factor analysis and ideotype-design: Proposal and application on elephant grass breeding for bioenergy. GCB Bioenergy. 2018;10 (1):52-60.
14. Olivoto T, Lúcio ADC, da Silva JAG, Marchioro VS. Mean performance and stability in multi-environment trials I: Combining features of AMMI and BLUP techniques. Agronomy Journal. 2019;111(6):2949-60.
15. Olivoto T, Lúcio ADC, da Silva JAG, Sari BG, Diel MI. Mean performance and stability in multi-environment trials II: selection based on multiple traits. Agronomy Journal. 2019;111:2961-69.
16. Jarquin D, Howard R, Crossa J, Beyene Y, Gowda M, Martini JW, et al. Genomic prediction enhanced sparse testing for multi-environment trials. G3 Genes Genomes Genetics. 2020;10:2725-39. https://doi.org/10.1534/g3.120.401349
17. Woyann LG, Meira D, Matei G, Zdziarski AD, Dallacorte LV, Madella LA. Selection indexes based on linear-bilinear models applied to soybean breeding. Agronomy Journal. 2020;112:175-82.
18. Zuffo AM, Steiner F, Aguilera JG, Teodoro PE, Teodoro LPR, Busch A. Multi-trait stability index: a tool for simultaneous selection of soya bean genotypes in drought and saline stress. J Agro Crop Sci. 2020; 206(6):815-22.
19. Olivoto T, Nardino M, Meira D, Meier C, Follmann DN, Souza VQ, et al. Multi-trait selection for mean performance and stability in maize. Agronomy Journal. 2021;113:3968-74. https://doi.org/10.1002/agj2.20741
20. Azrai M, Aqil M, Efendi R, Andayani NN, Makkulawu AT, Iriany RN, et al. A comparative study on single and multiple trait selections of equatorial grown maize hybrids. Frontiers in Sustainable Food Systems. 2023;7:1185102
21. Palaniyappan S, Ganesan KN, Manivannan N, Ravichandran V, Senthil N. Multi trait genotype-ideotype distance index-A tool for identification of elite parental inbreds for developing heterotic hybrids of fodder maize (*Zea mays* L.). Electronic Journal of Plant Breeding. 2023;14(3):841-49. https://doi.org/10.37992/2023.1403.098
22. Singamsetti A, Zaidi PH, Seetharam K, Vinayan MT, Olivoto T, Mahato A, et al. Genetic gains in tropical maize hybrids across moisture regimes with multi-trait-based index selection. Frontiers in Plant Science. 2023;14:1147424. https://doi.org/10.3389/fpls.2023.1147424
23. Azrai M, Aqil M, Andayani NN, Efendi R, Suarni, Suwardi, et al. Optimizing ensembles machine learning, genetic algorithms and multivariate modeling for enhanced prediction of maize yield and stress tolerance index. Frontiers in Sustainable Food Systems. 2024;8:1334421. https://doi.org/10.3389/fsufs.2024.1334421
24. Trung DN, Tuan PQ, Anh NTN, Van LV. Phenotypic diversity and selection of superior tropical sweetcorn inbred lines by multivariate method and combining ability analysis. Ecological Genetics and Genomics. 2024;30:100215.
25. Olivoto T, Nardino M. MGIDI: toward an effective multivariate selection in biological experiments. Bioinformatics. 2021;37:1383-89. https://doi.org/10.1093/bioinformatics/btaa981
26. R Core Team. R: a language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2021. Available from: https://www.r-project.org/
27. Olivoto T, Lúcio ADC. Metan: An r package for multi-environment trial analysis. Methods in Ecology and Evolution. 2020;11(6):783-89. https://doi.org/10.1111/2041-210X.13384
28. Wickham H. Ggplot2: Elegant graphics for data analysis. 2nd ed. Cham: Springer; 2016
29. Kaiser HF. An index of factorial simplicity. Psychometrika. 1974;39 (1):31-36. https://doi.org/10.1007/BF02291575
30. Shrestha J. Cluster analysis of maize inbred lines. J Nepal Agri Res Council. 2016;2:33-36. https://doi.org/10.3126/jnarc.v2i0.16119
31. Belalia N, Lupini A, Djemel A, Morsli A, Mauceri A, Lotti C, et al. Analysis of genetic diversity and population structure in Saharan maize (*Zea mays* L.) populations using phenotypic traits and SSR markers. Genetic Resources and Crop Evolution. 2019;66:243-57. https://doi.org/10.1007/s10722-018-0709-3
32. Yue H, Wei J, Xie J, Chen S, Peng H, Cao H, et al. A study on genotype -by-environment interaction analysis for agronomic traits of maize genotypes across Huang-Huai-Hai region in China. Phyton. 2022;91 (1):57. https://doi.org/10.32604/phyton.2022.017308
33. Zendrato YM, Suwarno WB, Marwiyah S. Multi-trait selection of tropical maize genotypes under optimum and acidic soil conditions. SABRAO J Breed Genet. 2024;56(1):142-55. http://doi.org/10.54910/sabrao2024.56.1.13
34. Yan W, Frégeau-Reid J. Genotype by yield × trait (GYT) biplot: a novel approach for genotype selection based on multiple traits. Scientific Reports. 2018;8:8242. https://doi.org/10.1038/s41598-018-26688-8