

SABRAO Journal of Breeding and Genetics
 56 (2) 708-718, 2024
<http://doi.org/10.54910/sabrao2024.56.2.22>
<http://sabraojournal.org/>
 pISSN 1029-7073; eISSN 2224-8978



SELECTION OF LOWLAND TOMATO ADVANCED LINES USING SELECTION INDICES BASED ON PCA, PATH ANALYSIS, AND THE SMITH-HAZEL INDEX

M. FARID¹, M.F. ANSHORI^{1*}, K. MANTJA¹, I. RIDWAN¹, A. ADNAN², and G. SUBROTO³

¹Department of Agronomy, Faculty of Agriculture, Hasanuddin University, Makassar, Indonesia

²Department of Biology, State University of Makassar, Makassar, Indonesia

³Department of Agronomy, Faculty of Agriculture, Jember University, Jember, Indonesia

*Corresponding author's email: fuad.anshori@unhas.ac.id

Email addresses of co-authors: farid_deni@yahoo.co.id, katriani@yahoo.co.id, ifayanti@unhas.ac.id, adnan@unm.ac.id, gatots.faperta@unej.ac.id

SUMMARY

Developing lowland tomatoes (*Solanum lycopersicum* L.) is a better strategy to resist global warming. Tomato expansion also requires an efficient and organized selection process, including breeding the advanced segregating populations. Using selection indices in combination with biometrical analysis, such as principal component analysis, path analysis, and the Smith-Hazel index, can further improve the effectiveness of the selection concept. Therefore, the presented research aimed to evaluate the lowland F5 tomato lines through various selection indices and select the potential genotypes for further preliminary yield tests. The latest study proceeded in an augmented design, divided into four blocks, with a randomized complete block design as an environmental scheme. For comparison, the five standard tomato cultivars, consisting of Chung, Gustavi, Mawar, Tymoty, and Karina, underwent repeated planting in each block to compare with 45 F5 lines. All the tomato genotypes' analysis used correlation, path, and principal component analyses, and the Smith-Hazel index. Based on the results, the selection indices with a multivariate approach and genetic analysis proved effective in selecting tomato lines in the F5 populations. The PCA, Smith-Hazel, and path analyses were the best approaches for creating selection indices formed by three characteristics, namely, yield, the number of fruits per bunch, and the number of branches. The selection indices recognized and recommended the 23 tomato F5 strains for further studies in the preliminary yield testing.

Keywords: Path analysis, PCA, selection criteria, Smith-Hazel index, *Solanum lycopersicum*

Communicating Editor: Prof. Naqib Ullah Khan

Manuscript received: October 15, 2023; Accepted: December 30, 2023.

© Society for the Advancement of Breeding Research in Asia and Oceania (SABRAO) 2024

Citation: Farid M, Anshori MF, Mantja K, Ridwan I, Adnan A, Subroto G (2024). Selection of lowland tomato advanced lines using selection indices based on PCA, path analysis, and the Smith-Hazel index. *SABRAO J. Breed. Genet.* 56(2): 708-718. <http://doi.org/10.54910/sabrao2024.56.2.22>.

Key findings: The Smith-Hazel, PCA, and path analysis proved better approaches in forming the selection indices in the presented tomato research. Based on these indices, the selection index formed was 0.22, the yield + 0.1 numbers of fruits per bunch + 0.06 number of branches. The selection indices recognized 23 lowland tomato lines, recommended for inclusion in the preliminary yield tests. However, the five promising tomato strains were MC 74.12.8, KM30.5.2, MC 74.12.5, MC 29.4.6, and MC 29.4.5.

INTRODUCTION

Tomato (*Solanum lycopersicum* L.) is a member of the Solanaceae family and has the highest reputation in horticultural crops. Its economic value is reasonable for the community, resulting in its high demand worldwide (Bhandari *et al.*, 2021). However, in the present era, population growth, modernization, and global warming cause the balance of demand and supply of tomatoes to oscillate (Ayankojo and Morgan, 2020; Cărbunar *et al.*, 2022). Therefore, it is necessary to develop tomato genotypes occasionally, including their adaptability to various land elevations and environmental conditions.

Generally, the low and middle plains are unfavorable regions for the best tomato production. In the lowlands, the high temperature is one of the limiting factors that can reduce the productivity and quality of the tomatoes (Ayankojo and Morgan, 2020; Bhandari *et al.*, 2021; Dasgan *et al.*, 2021). Regardless, climate change and competition with other horticultural products at higher altitudes resulted in demand for the plant to ably adapt to comparatively high temperatures (Mawasid *et al.*, 2019; Bhandari *et al.*, 2021). One possible solution is to develop tomato cultivars adaptive at various altitudes, especially in lowlands (Mawasid *et al.*, 2019; Yunandra *et al.*, 2023). The development of lowland tomatoes is one of the themes in the study of the horticultural crops' adaptation to climate change. In developing lowland tomatoes, consistent studies have progressed in F2 to F5 generations (Fadhilah *et al.*, 2022; Farid *et al.*, 2022).

The F5 population is one of the advanced generations in the cultivation process before preliminary yield testing. In addition, several studies also reported decisive levels in

the generation to determine the stages of preliminary yield testing (Melo *et al.* 2017; B *et al.* 2022). Therefore, it is obligatory to conduct preliminary yield testing for suck populations. In addition, systematic selection is desirable to enhance the effectiveness of selection, and a multi-character approach is one valuable selection index related to productivity.

Selection by various traits has also been widely reported in several past studies (Alsabah *et al.*, 2019; Karima *et al.*, 2021; Padjung *et al.*, 2021; Tirtana *et al.*, 2021). This selection is very effective considering that productivity is polygenic and significantly has influences from numerous factors. It indicates that environmental changes will influence the yield, requiring other characteristics to support the productivity potential (Fellahi *et al.*, 2018; Anshori *et al.*, 2019). A multi-characteristic-based selection approach can proceed using the concept of selection index (Batista *et al.*, 2021; Farid *et al.*, 2022). The selection index can accumulate the potential for various qualities within a genotype (Cerón-Rojas and Crossa, 2022). These features will undergo formulation in a regression analysis with weighting values adjusted to the economic value of the selected characteristics (Moreira *et al.*, 2019; Lopez-Cruz *et al.*, 2020; Batista *et al.*, 2021). Therefore, the selection index effectivity closely depends upon the weighting value of each variable (Chung and Liao, 2022).

Determining trait weighting values can sustain evaluation with a principal component analysis (PCA). PCA plays a crucial role in combining the variance of several parameters into several dimensions with a certain level and direction of variance. The method results in a new dimension with no overlap variations (Jolliffe and Cadima, 2016; Anshori *et al.*, 2019; Zafar *et al.*, 2021). This advantage can be the basis for determining the weight of the selection index, and weighting through the PCA

system has also been reported in rice (Alsabah *et al.*, 2019; Karima *et al.*, 2021; Tirtana *et al.*, 2021) and maize (Padjung *et al.*, 2021; Fadhli *et al.*, 2023). Nevertheless, the development of this system continues in optimizing index weighting, which can occur with the path analysis. Path analysis can also run with PCA, optimizing the potential for cultivar priorities based on the direct influence of the attributing traits on the yield. Hence, the yield remains objectively prioritized compared with the secondary qualities supporting the yield (Anshori *et al.*, 2019; Padjung *et al.*, 2021).

Despite this, the combination still does not reach its genetic potential; therefore, the Smith-Hazel index approach is also an option in this development (Cerón-Rojas and Crossa, 2022; Chung and Liao, 2022). This approach emphasizes and considers the genetic potential in forming a selection index (Marulanda *et al.*, 2021; Cerón-Rojas and Crossa, 2022). The Smith-Hazel index also provides a better basis after combining it with the PCA. Therefore, using various combinations of selection indices is critical in enhancing the effectiveness of selection in the F5 populations of lowland tomatoes. The latest research sought to evaluate the diverse combinations of selection indices to select the lowland F5 tomato lines with the genetic potential to be candidate genotypes for further preliminary yield tests.

MATERIALS AND METHODS

Research procedure

The present-day research on tomatoes began in May to September 2023 at the Experimental Garden of the Faculty of Agriculture, Hasanuddin University, Makassar, Indonesia (-5°7'39.0648", 119°28'59.1672"). The experiment had an augmented design that contained four blocks and a randomized complete block as the environmental design. Tomato genotypes served as the main factor incurring division into two groups—repeated genotypes and non-repeated genotypes. Repeated genotypes were the check cultivars that played a role in identifying environmental

variance and served as a comparison with the tested lines. The tested tomato lines were the non-repeated genotypes due to the limited number of seeds. The check cultivars comprised five cultivars, i.e., Chung, Gustavi, Mawar, Tymoty, and Karina, and the tested lines were 52 F5 populations. However, seven strains experienced their death, and 45 were the optimal strains for further study and analysis. Based on this, generally, there were 72 experimental units in this study. In cultivating the tomato, all the field activities have proceeded according to Fadhilah *et al.* (2022) and Farid *et al.* (2022).

Observations made and data analysis

The data recording was on randomly selected plants and adjusted with the tomato descriptor according to IPGRI (1996), Fadhilah *et al.* (2022), and Farid *et al.* (2022). These are plant height (cm), dichotomous height (cm), stem diameter (cm), number of branches (branches), flowering age (days), harvest age (days), number of flowers per bunch (flowers), number of fruits per bunch (fruits), number of flowering bunches, fruit length, thickness, and diameter (cm), fruit weight and yield (g), and the total dissolved solids with refractometer (brix). These observations incur systematic analysis with several concepts. Meanwhile, the series of analyses used were analysis of variance (ANOVA) according to an augmented design, heritability (Acquaah, 2007), and factors analysis (Farid *et al.*, 2022)).

Based on the analysis of variance (ANOVA), the traits that were significant at the source of variance in tomato lines and lines vs. check genotypes underwent a standardization process to equalize the dimensions among the characteristics (Anshori *et al.*, 2019). The standardization results also advanced to a correlation analysis. Subsequently, the traits with a significant positive correlation with fruit yield proceeded to further analysis through the path analysis for the yield-related attributes. Characteristics with high direct influence continued analysis using the principal component analysis (PCA) (Alsabah *et al.*, 2019; Karima *et al.*, 2021; Padjung *et al.*, 2021; Tirtana *et al.*, 2021). The results

obtained through the PCA served as the basis for developing selection indices, both in combination with path analysis (Fadhli *et al.*, 2023) and with the Smith-Hazel index (Marulanda *et al.*, 2021; Cerón-Rojas and Crossa, 2022). The selection index obtained from several approaches sustained testing for its effectiveness using the heritability index (Anshori *et al.*, 2019). The selection index approach with the highest heritability index became the basis for selection in the tomato F5 populations. It is because heritability is the strength of a characteristic that can explain its genetic potential for the emergence of its phenotype; thus, a high heritability index becomes a reference in the genetic role of the accumulation of index formulas in its phenotype.

RESULTS

Based on the analysis of variance, several patterns emerged for grouping from the influence of variance sources on the growth characteristics (Table 1). The significant traits in the three sources of variance (checks, lines, and checks vs. lines) were the number of branches, the number of flowers per bunch, the number of fruits per bunch, fruit length, diameter, weight, yield, and total soluble solids. The second group comprised plant height and fruit thickness, which also have significant influence from the sources of variance of tomato checks and lines. The third group was the dichotomous height characteristic and the number of flowering bunches, which incurred significant influence from check genotypes and the check vs. lines. The fourth group consisted of the flowering age and harvest time, incurring effects from the check genotypes variance source. However, the stem diameter was the only trait with no considerable impact from the three sources of variation. Despite the significance of sources of variance influencing the growth characteristics, Table 1 also showed the heritability for all the observed features. Almost all parameters received groupings with high heritability, except for the stem diameter, flowering time, and harvest time. The attribute with the

highest heritability was the number of flowers per plant (98.09%).

The correlation analysis revealed that fruit yield has a significant positive correlation with the number of branches (0.52), the number of flowers per bunch (0.39), and the number of fruits per bunch (0.47) (Table 2). The number of branches also showed a significant positive association with the number of flowers per bunch (0.54) and the number of fruits per bunch (0.34), and the characteristics of fruit length (-0.38), fruit diameter (-0.34), fruit weight (-0.36), and total dissolved solids (-0.38), indicated a considerable negative correlation. The number of flowers per bunch also showed a substantial positive correlation with the number of fruits per bunch (0.35), and the total soluble solids gave a significant negative correlation (-0.36). The number of fruits per bunch also provided a noteworthy negative linkage with total dissolved solids. Meanwhile, other significant correlations occurred in fruit length – fruit diameter (0.99), fruit length – fruit weight (0.98), and fruit diameter – fruit weight (0.98). The cross-check results showed that the number of branches (0.37) and the number of fruits per bunch (0.32) had a direct effect on the fruit yield (Table 3). On the other hand, the number of flowers per bunch also provided a minimum direct influence (0.08). Similarly, the number of fruits per bunch and the number of branches also have a maximum indirect influence (0.13 and 0.11, respectively).

Based on the principal component analysis results, three PCs had a cumulative value of 80% (Table 4). PC1 has a maximum variance recording (42%). On this PC, the number of branches, the number of flowers per bunch, the number of fruits per bunch, and the fruit yield have the same direction of variance as the eigenvalues, i.e., 0.37, 0.18, 0.22, and 0.21, respectively. The characteristics of fruit length, diameter, weight, and total dissolved solids have the opposite direction of variation with eigenvector values of -0.48, -0.47, -0.48, and -0.24, respectively. On the PC2, productivity (-0.41) showed a higher eigenvector value compared with the PC1, with the same directional variance with the number of branches (-0.30), the number of flowering

Table 1. Analysis of various growth characteristics in tomato F5 populations.

Characteristics	Check (C)	Lines (L)	C vs L	Error	CV	VAR E	VAR G	VAR P	H (%)
PH	3156.37**	623.12**	477.82	105.83	11.66	21.17	103.46	124.62	83.02
DH	243.08*	155.30	703.99**	68.07	25.74	13.61	17.45	31.06	56.17
SD	24.02	20.23	30.55	25.52	47.03	5.10	-1.06	4.05	0.00
NB	1003.93**	58.15**	611.89**	9.10	17.81	1.82	9.81	11.63	84.35
FA	50.23**	6.45	3.86	3.43	5.59	0.69	0.60	1.29	46.76
HA	287.03**	29.08	50.26	20.44	5.75	4.09	1.73	5.82	29.71
NFlpB	18.94**	27.96**	31.89**	0.53	11.59	0.11	5.49	5.59	98.09
NFrpB	71.60**	2.46**	39.19**	0.38	14.06	0.08	0.42	0.49	84.64
NFB	2104.05**	80.45	758.99*	20.29	28.75	4.06	12.03	16.09	74.78
FL	109.14**	34.14**	93.16**	4.52	8.79	0.90	5.92	6.83	86.75
FT	93.85**	14.04*	18.06	4.68	10.17	0.94	1.87	2.81	66.67
FD	116.70**	30.95**	92.92**	4.09	8.47	0.82	5.37	6.19	86.80
FW	116.80**	27.79**	107.90**	2.96	19.59	0.59	4.97	5.56	89.34
TDS	16.00**	1.30**	3.65**	0.25	11.39	0.05	0.21	0.26	80.64
Yield	458127.85**	40179.37**	963790.27**	4532.29	17.92	906.46	7129.42	8035.87	88.72

Notes: CV = coefficient of variance, VAR E = environment variance, VAR G = genetic variance, VAR P = phenotypic variance, H = Heritability, PH = plant height, DH = dichotomous height, SD = stem diameter, NB = number of branches, FA = flowering age, HA = harvest age, NFlpB = number of flowers per bunch, NFrpB = number of fruits per bunch, NFB = number of flowering bunches, FL = fruit length, FT = fruit thickness, FD = fruit diameter, FW = fruit weight, TDS = total dissolved solids.

Table 2. Correlation analysis of selected tomato characteristics based on the analysis of variance.

Characteristics	Statistics	NB	NFlpB	NFrpB	FL	FD	FW	TDS
NFlpB	Coef	0.54**						
	p-value	0.0001						
NFrpB	Coef	0.34**	0.35*					
	p-value	0.0151	0.0117					
FL	Coef	-0.38	-0.01	-0.10				
	p-value	0.0058	0.9609	0.4753				
FD	Coef	-0.34*	0.02	-0.10	0.99**			
	p-value	0.0162	0.8989	0.4953	0.0000			
FW	Coef	-0.36**	0.01	-0.09	0.98**	0.98**		
	p-value	0.0093	0.9682	0.5327	0.0000	0.0000		
TDS	Coef	-0.38**	-0.36**	-0.38**	0.18	0.18	0.17	
	p-value	0.0063	0.0098	0.0059	0.2188	0.2101	0.2427	
Yield	Coef	0.52**	0.39**	0.47**	-0.07	-0.06	-0.08	-0.20
	p-value	0.0001	0.0053	0.0006	0.6158	0.6579	0.5741	0.1623

Notes: numbers followed by signs are significantly different from table $r 0.05 = 0.34 (*)$; $r 0.01 = 0.47 (**)$, NB = number of branches, NFlpB = number of flowers per bunch, NFrpB = number of fruits per bunch, FL = fruit length, FD = fruit diameter, FW = fruit weight, TDS = total dissolved solids.

Table 3. Path analysis of tomato traits that have a significant positive correlation with the fruit yield.

Characteristics	Direct effect	Indirect effect			Correlation
		NB	NFlpB	NFrpB	
NB	0.37		0.04	0.11	0.52
NFlpB	0.08	0.2		0.11	0.39
NFrpB	0.32	0.13	0.03		0.48

Notes: NB = number of branches, NFlpB = number of flowers per bunch, NFrpB = number of fruits per bunch.

Table 4. Principal component analysis of selected tomato traits based on the analysis of variance.

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
NB	0.37	-0.30	0.19	0.38	-0.23	0.73	-0.01	-0.05
NFlpB	0.18	-0.46	-0.04	0.54	0.56	-0.38	0.00	0.00
NFrpB	0.22	-0.39	-0.08	-0.72	0.46	0.26	-0.02	0.00
FL	-0.48	-0.30	-0.02	0.00	-0.07	0.05	-0.56	-0.60
FD	-0.47	-0.32	0.00	0.04	-0.06	0.12	-0.24	0.78
FW	-0.48	-0.31	-0.03	0.01	-0.05	0.10	0.79	-0.19
TDS	-0.24	0.29	0.80	0.03	0.45	0.16	0.00	-0.02
Yield	0.21	-0.41	0.57	-0.23	-0.45	-0.45	0.02	0.01
Standard deviation	1.85	1.49	0.90	0.85	0.69	0.56	0.14	0.11
Proportion of Variance	0.43	0.28	0.10	0.09	0.06	0.04	0.00	0.00
Cumulative Proportion	0.43	0.71	0.81	0.90	0.96	1.00	1.00	1.00
Eigenvalues	3.41	2.23	0.81	0.72	0.47	0.32	0.02	0.01

Notes: PC = principal component, NB = number of branches, NFlpB = number of flowers per bunch, NFrpB = number of fruits per bunch, FL = fruit length, FD = fruit diameter, FW = fruit weight, TDS = total dissolved solids.

Table 5. Testing the effectiveness of several combinations of selection indices based on heritability index.

Parameter	PCA	Path + PCA	SH + PCA	SH + Path + PCA
NB	0.37	0.14	0.22	0.06
NFrpB	0.22	0.07	0.21	0.10
Yield	0.21	0.21	0.27	0.22
H index	86.42	90.10	88.11	90.35

Notes: PCA = principal component analysis, SH = Smith-Hazel, NB = number of branches, NFlpB = number of flowers per bunch, NFrpB = number of fruits per bunch, H index = heritability of index.

bunches (-0.46), and the number of fruits per bunch (-0.39), as well as other characteristics, except total dissolved solids (0.29). For PC3, productivity had the largest eigenvector compared with PC1 and PC2; however, the said trait had a different direction of variation with the number of flowers per bunch (-0.04) and the number of fruits per bunch (-0.08).

Outcomes of the selection index weighting estimation exhibited that the index using the Smith-Hazel approach in the combination of path and PCA analysis had the highest heritability index (90.35%) (Table 5). Nevertheless, this value was not significantly different from the heritability of the index using the combined path approach and PCA without the Smith-Hazel index (90.10%). However, the path analysis revealed that the heritability did not differ much from the autonomous heritability of each selection characteristic.

The results of sorting the index values based on the selection index formulation are available in Table 6. Based on this table, 23

tomato lines had positive index values. Based on the parental comparisons, 34 tomato lines occurred better than Chung's best check. The Karina cultivar was the genotype with the lowest value among the existing check genotypes. However, in the presented study, the five best tomato genotypes were MC 74.12.8 (1.91), KM30.5.2 (1.59), MC 74.12.5 (1.56), MC 29.4.6 (1.32), and MC 29.4.5 (1.25).

DISCUSSION

Based on the analysis of variance, the traits with strong influences from the variance of tomato lines and lines vs. checks were potential selection characteristics for further analysis and exploration. The selection relied on both approaches that have also been reported, especially with augmented designs in past studies of tomatoes and beans (Saba *et al.*, 2017; Fadhilah *et al.*, 2022; Farid *et al.*,

Table 6. Testing selection index for tomato F5 populations based on the Smith-Hazel, path analysis, and PCA selection index.

Rank	Genotypes	Real			Standardization			Index
		NB	NFrpB	Yield	NB	NFrpB	Yield	
1	MC 74.12.8	15.00	5.98	553.49	0.17	12.55	2.92	1.91
2	KM 30.5.2	27.00	4.81	577.20	8.43	3.79	3.22	1.59
3	MC.74.12.5	14.45	5.95	504.45	-0.32	11.88	1.78	1.56
4	MC 29.4.6	19.95	5.76	404.06	4.38	9.45	0.50	1.32
5	MC 29.4.5	17.46	5.90	403.39	1.59	10.74	0.35	1.25
6	MC 10.6.3	14.50	5.42	623.56	-0.53	2.51	4.32	1.17
7	MC 12.3.2	18.31	4.39	519.33	2.45	3.05	3.10	1.13
8	KM 5.3.4	31.56	4.20	508.16	9.03	1.15	1.75	1.04
9	MC 27.12.2	21.38	4.93	405.55	4.46	5.74	0.47	0.95
10	KM 23.3.1	26.75	6.30	384.92	3.99	6.49	0.13	0.92
11	MC 10.7.5	14.69	6.96	502.19	-0.07	6.22	1.14	0.87
12	MC 27.12.1	22.25	4.99	388.60	3.49	3.84	0.17	0.63
13	MC74.12.7	14.17	5.47	368.56	-0.52	6.65	-0.10	0.61
14	MC 12.3.5	17.07	4.38	447.60	1.80	2.39	1.16	0.60
15	MC 8.3.2	13.00	5.52	364.55	-5.88	9.96	-0.20	0.60
16	MC 38.1.4	15.60	5.32	405.56	0.58	4.31	0.39	0.55
17	MC 10.11.6	14.36	7.06	349.94	-0.22	5.37	-0.25	0.47
18	MC 27.7.3	15.05	5.44	429.48	0.29	3.31	0.47	0.45
19	MC 10.4.5	14.18	5.05	457.06	-1.04	1.94	1.20	0.40
20	MC 8.3.7	14.13	4.82	434.04	-0.46	1.44	0.83	0.30
21	MC 10.4.6	14.18	4.90	406.73	-0.70	2.01	0.53	0.27
22	MC 10.7.8	13.61	6.63	271.26	-0.48	5.48	-1.49	0.19
23	MC79.2.7	18.25	4.00	368.95	2.30	0.01	-0.10	0.12
24	KM 69.4.3	14.54	2.70	437.12	-0.21	-2.64	1.10	-0.03
25	MC 9.2.4	10.62	3.98	411.90	-2.85	-0.04	0.54	-0.06
26	KM 69.4.8	15.00	2.93	367.37	0.07	-1.20	-0.12	-0.14
27	MC 38.7.5	11.84	3.63	419.86	-4.08	-1.35	0.75	-0.21
28	KM 69.4.8	15.75	3.18	329.15	0.58	-1.42	-0.64	-0.25
29	MC 38.7.3	11.57	3.64	379.03	-2.93	-0.85	0.03	-0.25
30	KM 69.4.1	15.39	3.03	345.03	0.38	-1.77	-0.46	-0.25
31	MC 9.2.2	8.90	3.85	389.89	-4.43	-0.42	0.19	-0.26
32	KM 69.5.6	15.65	2.58	363.55	0.31	-2.66	-0.19	-0.29
33	MC 35.7.5	13.40	3.39	351.84	-1.49	-1.49	-0.35	-0.32
34	MC 8.3.3	10.95	5.01	340.13	-7.18	2.99	-0.96	-0.34
35	Chung	22.74	5.68	255.38	2.04	2.29	-3.27	-0.37
36	KM 69.5.1	15.00	2.34	338.86	0.07	-2.98	-0.43	-0.39
37	MC 35.7.3	9.33	1.80	431.52	-2.28	-3.72	0.49	-0.40
38	KM 69.6.2	13.69	1.65	438.51	-0.57	-5.90	0.89	-0.43
39	KM 69.5.3	14.39	2.26	368.30	-0.21	-4.16	-0.15	-0.46
40	MC 38.8.1	9.00	3.87	341.32	-4.95	-0.49	-0.53	-0.46
41	MC 38.8.3	10.00	3.96	332.26	-4.71	-0.13	-0.91	-0.50
42	KM 69.5.5	12.28	1.54	342.78	-1.91	-6.80	-0.65	-0.94
43	MC. 30.10.7	17.65	1.87	359.75	0.48	-13.39	-0.22	-1.36
44	Mawar	10.40	2.45	197.34	-12.44	-2.54	-1.82	-1.40
45	KM 71.10.3	14.33	0.91	305.92	-0.13	-15.89	-0.84	-1.78
46	KM 71.10.3	8.25	0.83	346.29	-3.13	-19.71	-0.52	-2.27
47	Timoty	6.04	2.41	99.32	-4.70	-4.81	-7.65	-2.45
48	Gustavi	9.70	2.38	138.00	-4.67	-6.59	-8.99	-2.92
49	KM 69.5.1	9.13	1.58	263.34	-8.22	-7.51	-12.94	-4.09
50	Karina	8.50	1.74	86.94	-6.12	-19.15	-8.38	-4.13

Notes: NB = number of branches, NFpB = number of flowers per bunch, NFrpB = number of fruits per bunch.

2022; Yaseen *et al.*, 2022). Therefore, the characteristics of the number of branches, number of flowers per bunch, the number of fruits per bunch, fruit length, diameter, weight, total soluble solids, and fruit yield showed considerable potential to be the best candidate selection traits in the process of cultivating F5 tomato populations.

Combining correlation and path analyses served in recognizing and selecting secondary criteria for selection. Based on these two analyses, the number of branches and total fruits per bunch were selection characteristics that support fruit productivity. The selection concept also has a report on the tomato (*Lycopersicon esculentum*) genotypes (Anshori *et al.*, 2019; Tsagaye *et al.*, 2022). Both analyses were considerably efficient because the reduction approach proceeded systematically. The correlation was highly efficient in initially identifying the selection criteria related to the fruit yield; hence, the cross-tracing process will be more focused on partitioning the direct effect of a secondary characteristic on the variance of the fruit yield (Fadhilah *et al.*, 2022). Therefore, combining both analyses is a solution to selecting secondary attributes. In addition, in the presented research, the number of branches, the number of fruits per bunch, and the fruit yield can come together as selection criteria.

Determination of the number of branches and fruits per bunch also had support from various aspects. The number of branches is identical to indeterminate tomato plants, which constitute most of this population (Amare and Gebremedhin, 2020). Therefore, this was one of the reasons why the number of branches is essential as a selection trait. Apart from that, several other studies also reported the number of tomato branches as an indication of an efficient selection quality (Adewale and Adebo, 2018; Tsagaye *et al.*, 2022); therefore, the use of the number of branches is relevant as a selection characteristic supporting the fruit yield. Another characteristic crucial to this research was the number of fruits per bunch. The said trait also depends on the parents' potential, one of which has the potential for many fruits in one bunch (Fadhilah *et al.*, 2022). Several

studies also explained the potential number of fruits per bunch as an influential factor in determining rice (Rasheed *et al.*, 2020) and tomato yields (Anisa *et al.*, 2022).

The trait fruit weight is a commonly used selection criterion (Farid *et al.*, 2022) and not the chief selection characteristic in this population. However, in this study, fruit weight showed a significant value in the source variance of tomato lines and checks vs. lines. It also depended upon the dominance of the number of fruits per bunch, supporting the yield potential. In addition, most tomato populations have relatively similar fruit appearance and weight; thus, fruit weight was not a prime feature in this population. Therefore, the number of branches and the number of fruits per bunch were the superior traits that supported the fruit yield and were relevant to benefit as a selection trait with productivity.

Based on the PCA analysis, PC1 was the best in developing selection index weights. PC1 can describe the directional pattern of variance between characteristics that were relevant to correlation analysis and also has the highest proportion of variance and eigenvalue compared with other PCs (Jolliffe and Cadima, 2016; Tirtana *et al.*, 2021; Zafar *et al.*, 2021). Meanwhile, in several studies, PC2 and PC3 can also be a basis for weighting (Anshori *et al.*, 2019; Karima *et al.*, 2021). However, in the current research, the two PCs insufficiently explained the pattern of variance among the traits like PC1, and consequently, PC1 attained more priority over the other PCs. Therefore, in PC1, the eigenvectors can be a basis for weighting the selection index and combined with several other selection indices, such as path analysis and the Smith-Hazel index. The results of the index formation underwent comparison with each other to assess the effectiveness of the indices through heritability.

Based on the validation through the heritability index, the combination of the Smith-Hazel approach to PCA + path analysis has the best heritability index compared with other approaches. This heritability also outperformed the sovereign heritability for each selection characteristic. However, the

effectiveness of the Smith-Hazel approach was like the PCA + path analysis approach without Smith-Hazel. It also indicates that the PCA + path analysis approach is considerably efficient in forming selection indices. The effectiveness of this approach also acquired support from high heritability for each independent characteristic and high genetic covariance among the yield traits and supporting yield components. It provides the effectiveness of the Smith-Hazel approach that without it, not much difference occurs. However, in terms of evaluation, the Smith-Hazel approach enhances the effectiveness of the index, based on PCA alone and in a combination of PCA and path analysis. The efficacy of using Smith-Hazel has also come from Marulanda *et al.* (2021) and Cerón-Rojas and Crossa (2022). Therefore, the index formation still depended upon the Smith-Hazel approach, PCA, and path analysis. Based on this approach, the selection index formed was:

Selection index = 0.22 the yield + 0.1 number of fruits per bunch + 0.06 number of branches.

Based on the selection indices, the 23 tomato advanced lines showed considerable effectiveness for further testing in the preliminary yield test. These findings had a basis on the positive values of these 23 lines. A positive value on a standardized index indicates that the potential of the strain was better than the general population on average (Peternelli *et al.*, 2017; Anshori *et al.*, 2019; Karima *et al.*, 2021; Farid *et al.*, 2022). Therefore, the positive index values approach was the basis for the line selection process for continuation at the preliminary yield testing.

CONCLUSIONS

The results showed that using selection indices with a multivariate and genetic analysis proved effective in selecting promising tomato lines in the F5 populations. The Smith-Hazel, PCA, and path analysis approaches effectively formed selection indices. Based on this approach, the selection index formed was 0.22 the yield + 0.1 number of fruits per bunch + 0.06 number of branches. However, the PCA analysis represents Smith-Hazel's potential, provided

the heritability and correlation between the constituent characteristics have high values. In this study's tomato populations, the index selection showed 23 lowland tomato lines recommended for inclusion in the preliminary yield testing. In this selection, the five best and leading tomato strains were MC 74.12.8, KM30.5.2, MC 74.12.5, MC 29.4.6, and MC 29.4.5.

ACKNOWLEDGMENTS

The authors are grateful to the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia through the scheme of BIMA Fundamental Research with grant number: 124/E5/PG.02.00PL/2023.

REFERENCES

- Acquaah G (2007). Principles of Plant Genetics and Breeding. Blackwell Publishing, Oxford, United Kingdom, pp. 130-134.
- Adewale BD, Adebo UG (2018). Phenotypic identity, similarity and stability for selection of tomatoes (*Lycopersicon esculentum* L.) cultivars in South-western Nigeria. *Sci. Hortic.* 235: 264-269.
- Alsabah R, Purwoko BS, Dewi IS, Wahyu Y (2019). Selection index for selecting promising doubled haploid lines of black rice. *SABRAO J. Breed. Genet.* 51(4): 430-441.
- Amare G, Gebremedhin H (2020). Erratum: Effect of plant spacing on yield and yield components of tomato (*Solanum lycopersicum* L.) in Shewarobit, Central Ethiopia. *Scientifica* 2020: 8357237.
- Anisa WN, Afifah EN, Murti RH (2022). Selection of tomato breeding lines based on morphological traits associated with high yield potential in double-cross population. *Biodiversitas* 23: 2973-2980.
- Anshori MF, Purwoko BS, Dewi IS, Ardie SW, Suwarno WB (2019). Selection index based on multivariate analysis for selecting doubled-haploid rice lines in lowland saline prone area. *SABRAO J. Breed. Genet.* 51(2): 161-174.
- Ayankojo IT, Morgan KT (2020). Increasing air temperatures and its effects on growth and productivity of tomato in South Florida. *Plants* 9: 1-16.
- B LP, K SD, C HDR, R J, G R, Sinha P, Aleena D, Harika G, Mastanbee SK, Ravindra KS,

- Kousik MBVN, Swamy H KM, Anila M, Kale RRRR, Kumar TD, Punniakotti E, Hajira SK, Bhadana VP, Sundaram RM (2022). Marker assisted pedigree breeding based improvement of the Indian mega variety of rice MTU1010 for resistance against bacterial blight and blast and tolerance to low soil phosphorus. *PLoS ONE* 17(1): e0260535.
- Batista LG, Gaynor RC, Margarido GRA, Byrne T, Amer P, Gorjanc G, Hickey JM (2021). Long-term comparison between index selection and optimal independent culling in plant breeding programs with genomic prediction. *PLoS One* 16: 1-5.
- Bhandari R, Neupane N, Adhikari DP (2021). Climatic change and its impact on tomato (*Lycopersicon esculentum* L.) yield in plain area of Nepal. *Environ. Challenges* 4: 100129.
- Cărbunar M, Mîntaş O, Sabău NC, Borza I, Stanciu A, Pereş A, Venig A, Curilă M, Cărbunar ML, Vidican T, Oneţ C (2022). Effectiveness of measures to reduce the influence of global climate change on tomato cultivation in solariums — A case study: Crişurilor Plain, Bihor, Romania. *Agric.* 12: 1-26.
- Cerón-Rojas JJ, Crossa J (2022). The statistical theory of linear selection indices from phenotypic to genomic selection. *Crop Sci.* 62: 537-563.
- Chung PY, Liao CT (2022). Selection of parental lines for plant breeding via genomic prediction. *Front. Plant Sci.* 13: 934767.
- Dasgan HY, Dere S, Akhoundnejad Y, Arpacı BBI (2021). Effects of high-temperature stress during plant cultivation on tomato (*Solanum lycopersicum* L.) fruit nutrient content. *J. Food Qual.* 2021: 7994417.
- Fadhilah AN, Farid M, Ridwan I, Anshori MF, Yassi A (2022). Genetic parameters and selection index of high-yielding tomato F2 population. *SABRAO J. Breed. Genet.* 54: 1026-1036.
- Fadhli N, Farid M, Azrai M, Nur A, Efendi R, Priyanto SB, Nasruddin AD, Novianti F (2023). Morphological parameters, heritability, yield component correlation, and multivariate analysis to determine secondary characters in selecting hybrid maize. *Biodiversitas* 24: 3750-3757.
- Farid M, Anshori MF, Ridwan I (2022). Tomato F3 lines development and its selection index based on narrow-sense heritability and factor analysis. *Biodiversitas* 23: 5790-5797.
- Fellahi ZEA, Hannachi A, Bouzerzour H (2018). Analysis of direct and indirect selection and indices in bread wheat (*Triticum aestivum* L.) segregating progeny. *Int. J. Agron.* 2018: 1-11.
- International Plant Genetic Resources Institute (IPGRI) (1996). Descriptors for tomato (*Lycopersicon spp.*). International Plant Genetic Resources Institute, Roma, p. 44.
- Jolliffe IT, Cadima J (2016). Principal component analysis: A review and recent developments. *Phil. Trans. R. Soc. A.* 374: 20150202.
- Karima AW, Putri RK, Purwoko BS, Dewi IS, Suwarno WB, Kurniawati A (2021). Selection of doubled haploid black rice lines in advanced yield trial based on multivariate analysis. *Biodiversitas* 22: 5425-5431.
- Lopez-Cruz M, Olson E, Rovere G, Crossa J, Dreisigacker S, Mondal S, Singh R, Campos G dL (2020). Regularized selection indices for breeding value prediction using hyperspectral image data. *Sci. Rep.* 10: 1-12.
- Marulanda JJ, Mi X, Utz HF, Melchinger AE, Würschum T, Longin CFH (2021). Optimum breeding strategies using genomic and phenotypic selection for the simultaneous improvement of two traits. *Theor. Appl. Genet.* 134: 4025-4042.
- Mawasid FP, Syukur M, Trikoesoemaningtyas (2019). Epistatic gene control on the yield of tomato at medium elevation in the tropical agroecosystem. *Biodiversitas* 20(7): 1880-1886.
- Melo RCd, Trevisani N, Pereira TCV, Guidolin AF, Coimbra JLM (2017). Heterozygosity level and its relationship with genetic variability mechanisms in beans. *Rev. Cienc. Agron.* 48(3): 480-486.
- Moreira SO, Kuhlcamp KT, Barros FLdS, Zucoloto M, Godinho TdO (2019). Selection index based on phenotypic and genotypic values predicted by reml/blup in papaya. *Rev. Bras. Frutic.* 41: 1-7.
- Padjung R, Farid M, Musa Y, Anshori MF, Nur A, Masnenong A (2021). Drought-adapted maize line based on morpho-physiological selection index. *Biodiversitas* 22(9): 4028-4035.
- Paternelli LA, Moreira ÉFA, Nascimento M, Cruz CD (2017). Artificial neural networks and linear discriminant analysis in early selection among sugarcane families. *Crop Breed. Appl. Biotechnol.* 17: 299-305.
- Rasheed A, Hassan MU, Aamer M, Batool M, Fang S, Wu Z, Li H (2020). A critical review on the improvement of drought stress tolerance in rice (*Oryza sativa* L.). *Not. Bot. Horti Agrobot. Cluj-Napoca* 48: 1756-1788.
- Saba I, Sofi PA, Zeerak NA, Mir RR, Gull M (2017). Using augmented design for evaluation of

- common bean (*Phaseolus vulgaris* L.) germplasm. *Int. J. Curr. Microbiol. Appl. Sci.* 6: 246-254.
- Tirtana A, Purwoko BS, Dewi IS, Trikoesoemaningtyas (2021). Selection of upland rice lines in advanced yield trials and response to abiotic stress. *Biodiversitas* 22: 4694-4703.
- Tsagaye D, Gadebo A, Aklilu S (2022). Correlation and path analysis in tomato (*Lycopersicon esculentum* MILL) genotypes. *Ecol. Evol. Biol.* 7: 46-53.
- Yaseen M, Kashif M, Nazish HT, Munir R, Iqbal J, Usman M, Rabbani G (2022). Effect of rain-fed conditions on yield of mash bean genepool by using augmented design. *J. Stat. Theory Appl.* 21: 186-199.
- Yunandra Y, Deviona D, Zuhry E, Syukur M, Ardian A, Effendi A, Nurbaiti N, Yoseva S, Auliyanda FN (2023). Selection criteria for lowland tomatoes (*Solanum lycopersicum* L.). *E3S Web Conf.* 373: 03005.
- Zafar MM, Manan A, Razzaq A, Zulfqar M, Saeed A, Kashif M, Khan AI, Sarfraz Z, Mo H, Iqbal MS, Shakeel A, Ren M (2021). Exploiting agronomic and biochemical traits to develop heat resilient cotton cultivars under climate change scenarios. *Agronomy* 11(9): 1885.