Acta Scientiarum

http://www.periodicos.uem.br/ojs/
ISSN on-line: 1807-8621
Doi: 10.4025/actasciagron.v47i1.69619

PLANT BREEDING

Combination of mixed linear model approach with selection indices in kale breeding programs

Eduardo Alves da Silva, André Boscolo Nogueira da Gama, Valter Carvalho de Andrade Junior*, Orlando Gonçalves Brito, Ariana Lemes da Costa and Ana Izabela Freire

Departamento de Agricultura, Universidade Federal de Lavras, Campus Universitário, Cx. Postal, 3037, 37200-900, Lavras, Minas Gerais, Brazil. *Author for correspondence. Email: valter.andrade@ufla.br

ABSTRACT. Utilizing selection indices is an effective strategy for the simultaneous evaluation of multiple traits in kale breeding programs. This approach allows for the selection of kale genotypes that exhibit enhanced productivity and adaptability by combining desirable attributes for the crop. In this study, we employed a mixed model approach in combination with various selection indices to estimate selection gains and recommend the most suitable index for kale breeding. The experiment was conducted at the Center of Development and Technology Transfer, Federal University of Lavras, Ijaci, MG. Thirty-four experimental genotypes were assessed in a randomized block design with three replicates, featuring four plants per plot. We evaluated several traits, including total leaf yield, number of leaves, average leaf mass, number of sprouts and chlorophyll content. Data analysis was performed at both the plot average level and the average quantity of the five harvests. Statistical analysis of mixed models confirmed the presence of genetic variability among kale genotypes. We examined the Smith and Hazel, Mulamba and Mock, Z-index, and FAI-BLUP indices. Smith and Hazel, Mulamba and Mock, as well as Z-index, were found unsuitable for leafy kale selection in breeding programs. The FAI-BLUP index demonstrated superior performance, aligning with the specific objectives of the kale breeding program and offering desirable gains. Therefore, we recommend the use of the FAI-BLUP index in kale breeding programs.

Keywords: *Brassica oleracea sub acephala* L.; biometry; genetic improvement; multivariate analysis; REML/BLUP.

Received on September 14, 2023.
Accepted on January 4, 2024.

Introduction

In the realm of plant breeding, the development of genotypes boasting high yields and desirable commercial traits holds paramount importance. The ultimate cultivar is anticipated to exhibit superior average performance across the desired traits (Baker, 2020). Consequently, to effectively select a specific genotype or progeny, a comprehensive array of traits must undergo evaluation (Fadhli, Farid, Effendi, Azrai, & Anshori, 2020). This dataset necessitates simultaneous analysis, enabling the identification of the genotype with the most commendable overall average performance (Olivoto, Lúcio, Silva, Sari, & Diel, 2019). Nonetheless, concurrent selection mandates mathematical and statistical resources to predict selection gains (Resende & Alves, 2021).

Among the array of statistical methods, linear mixed models in tandem with restricted maximum likelihood (REML/BLUP) have demonstrated remarkable efficiency in predicting genotypic values for individual traits (Henderson, 1986; Piepho, 2009). Conversely, the use of multivariate models facilitates the analysis of data and the simultaneous selection of superior genotypes from a comprehensive dataset encompassing multiple traits (Piepho & Moring, 2005; Dugard, Todman, & Staines, 2022). Consequently, the amalgamation of these methodologies augments the precision of ideotype selection.

In the context of kale (*Brassica oleracea sub acephala* L.), numerous traits come into play when evaluating and endorsing high-performance genotypes. Recent advancements in kale genetic enhancement have revolved around the identification of genotypes characterized by prolific leaf productivity (Azevedo et al., 2015; Brito et al., 2019), robust plant architecture (Brito et al., 2019), resistance to pests and diseases (Tiwari, Saville, & Wratten, 2019; Hassan, Ahmad, & Hussain, 2022), and exceptional nutritional quality (Reda et al., 2021). Thus, a comprehensive assessment of several traits through a selection index is imperative for pinpointing top-performing genotypes.

Selection indexes employ weighting coefficients, derived from linear combinations of phenotypic data, to facilitate genotype selection and rank the most promising candidates (Olivoto & Nardino, 2021). Consequently, the use of these indexes, tailored to a set of crop-favorable characteristics, enables the identification of more productive and well-adapted genotypes (Cruz, Regazzi, & Carneiro, 2012).

Several selection indexes find application in plant breeding programs, including the classical index (Smith, 1936; Hazel, 1943), the rank sum-based index (Mulamba & Mock, 1978), the Z-index (Mendes, Ramalho, & Abreu, 2009), and the FAI-BLUP index (Rocha, Machado, & Carneiro, 2018). However, these indexes are diverse and not universally applicable to all crops. Moreover, the scientific literature lacks guidance regarding the most suitable index for kale. Therefore, a comprehensive evaluation of the available indexes is essential, intending to validate the one delivering optimal agronomic performance for this crop (Baker, 2020).

Consequently, the objective was to employ the linear mixed model approach in conjunction with selection indexes to estimate selection gains using different indexes and recommend the most suitable one for kale breeding.

Material and methods

Experimental design

The experiment took place between February and April 2019 at the Center of Development and Technology Transfer of the Federal University of Lavras (UFLA) in Ijaci, Minas Gerais State, Brazil (21°10' S, 44°55' W, at an elevation of 832 m). We evaluated thirty-four kale clones sourced from the UFLA germplasm bank. Seedlings were derived from sprouts collected from "mother plants" and placed in 72-cell polystyrene trays, where they remained in the greenhouse for 40 days. The study employed a randomized block design with three replications, featuring four plants per plot. Row spacing was set at 1.0 m, with a 0.5 m gap between plants. Planting, top-dressing fertilizations, and other cultural treatments adhered to crop recommendations (Trani et al., 2015).

Thirty days after planting, agronomic evaluations commenced, spanning five harvests conducted at biweekly intervals. This number of harvests sufficed to assess the average performance of the genotypes (Brito et al., 2019). Each harvest quantified total leaf yield (commercial leaf productivity) (Mg ha^{-1}), number of leaves, average leaf mass (g), and number of sprouts (removed after tallying) at the plot level. Commercial leaves were defined as those exceeding 15 cm in length and devoid of pest damage or signs of senescence (Azevedo et al., 2012). Simultaneously, chlorophyll content was measured using the ChlorofiLOG CFL1030 device (Falker).

Statistical Analysis

The dataset underwent analysis at both the plot average level and the mean quantity of the five harvests. We confirmed the presence of genetic variability among kale genotypes through statistical analysis of mixed models, employing the Likelihood Ratio Test (LRT) for all studied traits, following Henderson (1984):

$Y_{ij} = \mu + B_i + T_j + \varepsilon_{ij}.$

where: Y_{ij} represents the observed data, μ denotes the fixed mean effect (intercept), with $E(\mu) = \mu$. Bi signifies the block effect (fixed), T_j represents the genotype effect (random), and ε_{ij} indicates the error deviation associated with Y_{ij}, with $E(\varepsilon_{ij}) = 0$ and $E(\varepsilon^2_{ij}) = \sigma^2\varepsilon$. We assessed the significance of the mixed model analysis through the LRT for testing random effects, while the approximate F-test was employed for fixed effects.

Direct selection gains were estimated for all traits in the direction of improvement, except for the trait "number of sprouts," where lower values were deemed superior. The gain by direct selection was calculated using the formula:

h^2 = (Genetic variance) / (Genetic variance + Environmental variance)
DS = Xs (Mean of the selected) – Xo (General mean)
GS = (DS × h^2)
GS (%) = GS / Xo (phenotypic mean)

where: GS represents the expected selection gain, DS signifies the selection differential, and h^2 denotes the heritability coefficient.

Additionally, we performed selection via the classical index (Smith, 1936; Hazel, 1943), Z-index (Mendes et al., 2009), the sum of ranks index (Mulamba & Mock, 1978), and FAI-BLUP index (Rocha et al., 2018) for the six kale traits. In all cases, the top ten ranked genotypes were selected to compute selection gains.

Statistical analyses were conducted using R (R Core Team, 2021), employing the *lme4* package (Bates, Mächler, Bolker, &Walker, 2015), and the GENES program (Cruz et al., 2016). Venn diagrams were constructed using the *VennDiagram* package (Chen & Boutros, 2011), and network diagrams were generated using the CYTOSCAPE program (Shannon et al., 2003).

Results and discussion

Assessment of variability

The Akaike Information Criterion (AIC) was employed as a valuable tool to evaluate the model's fitting quality. In Table 1, the significance levels reported are derived from the chi-squared test, determined using the Likelihood Ratio Test (LRT). Notably, significant differences (p-value ⩽ 0.05) were observed among genotypes for all the traits (Table 1). According to Neath and Cavanaugh (2012), when comparing two models, if the difference between their AIC values is less than 10, it can be considered that there is no significant difference between the adjusted models. In such cases, the "reduced model" is assumed, which excludes the treatment effect.

Hence, the full model effectively accounted for the observed variation at the genotype level across traits, including total leaf yield (TLY), number of leaves (NL), average leaf mass (ALM), number of sprouts (NS), and chlorophyll content (CLF).

Table 1. Maximum likelihood ratio test for genetic variability among kale genotypes. Complete model (including treatment factor) and reduced model (no treatment factor).

	npar	logLik	AIC	LRT	Df	*p-value*
			TLY			
Complete Model	5	-247.8	505.6			
Reduced Model	4	-270.41	548.81	45.211	1	0.000001
			NL			
Complete Model	5	-1225.4	2460.7			
Reduced Model	4	-1265.1	2538.3	79.532	1	0.0001
			NS			
Complete Model	5	-305.23	620.45			
Reduced Model	4	-311.46	630.92	12.473	1	0.000001
			ALM			
Complete Model	5	-310.48	630.96			
Reduced Model	4	-319.92	647.84	18.882	1	0.00003
			CLF			
Complete Model	5	-258.9	527.8			
Reduced Model	4	-281.71	571.42	45.62	1	0.00008

*npar = number of parameters in the model; logLik = log-likelihood; AIC = Akaike information criterion; LRT = likelihood ratio test; Df =degrees of freedom; TLY = total leaf yield; NL = number of leaves; NS = number of sprouts; ALM = average leaf mass; CLF = chlorophyll content.

The presence of genetic variability among genotypes is a fundamental prerequisite for achieving genetic progress in breeding programs (Begna, 2021). This genetic diversity is pivotal as it provides the raw material for selection, allowing for the improvement of desirable traits as generations advance.

Selection indexes

Direct selection for TLY yielded a substantial selection gain of 34.40% for the subsequent cycle (Table 2). This result surpasses the findings of Brito et al. (2020), who reported a selection gain of 21% for the same trait. Additionally, direct selection for TLY led to the highest observed indirect gain for ALM at 18.21%. However, it resulted in a reduction in indirect gains for traits related to coloration, such as CFL with a decrease of -1.22%. When conducting direct selection for TLY, it is crucial to consider the indirect effects on other traits. While positive indirect gains are generally desirable, in the case of kale, they are not favorable for NS, where a 4.97% increase is observed. This is because kale breeding programs require a reduction in NS (Azevedo et al., 2015; Brito et al., 2020), making exclusive reliance on TLY for selection inefficient for all traits.

Direct selection based on the best genotypes for NL resulted in a gain of 24.01%, but it also led to an undesired gain of 11.85% for NS. This approach is less advantageous than direct selection based on TLY since the indirect gain for TLY is reduced by 30% (24.18%) compared to TLY-based direct selection (34.40%).

Selecting plants with the lowest NS indicated a significant 19.81% reduction in sprout numbers in the next generation (Table 2). However, this selection also incurred indirect losses for all other traits. Reducing the number of sprouts is particularly beneficial for commercial cultivars as it prevents lateral sprouting, ensuring that producers continue to purchase seeds from the companies developing these genotypes (Azevedo et al., 2012). In a study with kale progenies, Brito et al. (2020) found that direct selection for this trait can reduce the number of sprouts by up to 28%.

Table 2. Average gains from direct selection (in bold) and gain values from indirect selection for the traits in each of the corresponding rows.

Trat	TLY	NL	NS	ALM	CLF
TLY	34.40	12.66	4.97	18.21	-1.22
NL	24.18	24.01	11.85	1.85	1.03
NS	-6.66	-7.15	-19.81	-0.97	-4.27
ALM	29.31	4.85	0.09	20.87	-6.40
CLF	-6.63	2.36	8.27	-9.32	7.66

TLY = total leaf yield; NL = number of leaves; NS = number of sprouts; ALM = average leaf mass; CLF = chlorophyll content. Values on the diagonals represent the gain with direct selection for the trait in the line.

Direct selection of plants with higher values for ALM per plant provided a significant 20.87% increase for the next generation. It also led to an indirect gain for TLY of 29.31% but showed only a minimal increase in NS (0.09%). However, it resulted in a negative gain for CLF at -6.40% due to its association with lighter green leaf color, as chlorophyll content is linked to green color (Taiz et al., 2017). Thus, direct selection for ALM implies an indirect selection of productive genotypes for TLY (29.31%), as genotypes with larger average leaf mass are inherently more productive in terms of total leaf yield. Therefore, evaluating leaf size individually can effectively predict or select the most productive genotypes.

Direct selection based on CLF yielded an average gain of 7.66%. However, this gain was relatively low due to the limited genetic variance of this trait. The reduced variability among genotypes results in lower genetic gains (Kardos et al., 2021). Given its lower agronomic significance, direct selection for CLF is not recommended. Nevertheless, knowledge of CLF remains important, as it serves as a qualitative standard for hue in kale and aids breeders in ideotype selection.

In summary, direct selection for a single trait can lead to substantial gains in a few generations. However, such specific selection may result in "losses" or undesirable gains in other traits of interest (Table 2). As a result, it is advisable to conduct simultaneous selection of traits to ensure that favorable traits are selected in a balanced and homogeneous manner. Selection indexes, with their diverse estimation methodologies, are employed to achieve this balanced selection approach.

Smith and Hazel's index

According to Smith and Hazel's index ranking, the top-performing genotypes were identified as UFLA 32, UFLA 02, UFLA 04, UFLA 20, UFLA 30, UFLA 24, UFLA 13, UFLA 05, UFLA 11, and UFLA 09.

However, it is important to note that Smith and Hazel's index resulted in an undesired gain of 5.45% for the number of sprouts, which was the highest among all the indexes. Evaluating unwanted gains is a critical aspect when assessing the quality of an index (Cruz et al., 2012). Thus, this index was not efficient in selecting genotypes with favorable performance because the selection based on its estimation did not lead to a reduction in the number of sprouts. Additionally, the selected genotypes showed unwanted losses in ALM (-1.87%) and CLF (-0.36%) traits (Table 3).

Smith and Hazel's index is considered one of the simplest and most widely used indexes in breeding programs due to its superior performance (Vieira et al., 2017; Barth et al., 2022; Hidalgo-Contreras, Salinas-Ruiz, Eskridge, & Baenziger, 2021). It is particularly useful for selecting traits with a normal distribution, as it assumes that the traits are independent and follow a normal distribution. However, one of its main disadvantages is that it does not take into account the correlation between traits (Hazel, 1943). This means that an individual can be selected based on its performance in a single trait, even if that trait is unrelated to other important traits for breeding. Moreover, the classic index may result in the selection of individuals with average performance across all traits rather than individuals with high performance in one or more specific traits (Smith, 1936; Hazel, 1943).

Table 3. Selection gains by the Smith and Hazel's index, Mulamba and Mock's index, Z-index, and FAI-BLUP's index.

Chr	Smith and Hazel's index					Mulamba and Mock's index		
	Xo	Xs	h^2 (%)	GS	GS (%)	Xs	GS	GS (%)
TLY	11.03	12.93	87.31	1.66	15.04	14.59	3.11	28.18
NL	467.95	566.60	93.73	92.46	19.76	525.30	53.75	11.49
NS	22.23	24.06	66.17	1.21	5.45	22.66	0.28	1.28
ALM	23.33	22.74	73.81	-0.44	-1.87	27.84	3.33	14.27
CLF	42.81	42.63	86.64	-0.16	-0.36	42.32	-0.42	-0.99
Chr	Z index					FAI-BLUP's index		
	Xo	Xs	h^2 (%)	GS	GS (%)	Xs	GS	GS (%)
TLY	11.03	13.36	87.32	2.03	18.45	13.04	2.01	18.25
NL	468.00	510.83	93.72	40.14	8.58	500.06	32.06	6.85
NS	22.23	23.77	66.17	1.02	4.58	21.19	-1.04	-4.69
ALM	23.33	26.28	73.81	2.18	9.33	25.82	2.49	10.67
CLF	42.81	43.04	86.64	0.20	0.47	39.52	-3.29	-7.69

Chr = characteristic; $h^2_{(\%)}$ = heritability; Xo = average; Xs = average of selected; GS = gain of selection; GS $_{(\%)}$ = percentage of gain; TLY = total leaf yield; NL = number of leaves; NS = number of sprouts; ALM = average leaf mass; CLF = chlorophyll content.

Mulamba and Mock's index

The selected genotypes using Mulamba and Mock's index were: UFLA 02, UFLA 29, UFLA 04, UFLA 20, UFLA 23, UFLA 06, UFLA 09, UFLA 03, UFLA 14, and UFLA 13. These genotypes resulted in a selection gain of 1.28% for the number of sprouts (NS), which was lower than the gain achieved with Smith and Hazel's index. However, this index provided a substantial gain for TLY of approximately 28%. This gain was only 20% lower than the value observed with direct selection and almost twice the gain obtained with Smith and Hazel's index, indicating better performance in selecting for TLY. For CLF, there was a negative gain (-0.99%), and for ALM, there was a positive gain of 4.48%.

Mulamba and Mock's index is considered a simplistic method because it involves minimal mathematical and algebraic resources. It is also a non-parametric method that does not assume a normal distribution of the data (Mulamba & Mock, 1978). This makes it suitable for selecting traits with non-normal distributions and is particularly useful for small sample sizes. One of the main advantages of Mulamba and Mock's index is its robustness to outliers because it relies on the ranks of the traits rather than their raw values, making it widely used in various applications (Beloti, Maciel, Juliatti, Finzi, & Cardoso, 2020; Ramos et al., 2022; Jesus, Passos, & Diniz, 2023).

However, a notable disadvantage of both Mulamba and Mock's index and Smith and Hazel's index is their inability to take into account the correlation between traits (Dalbosco et al., 2018). Several studies have identified significant correlations between commonly evaluated traits in kale breeding programs (Meena, Kale, & Meena, 2014; Azevedo et al., 2015; Brito et al., 2020). This limitation can affect the efficiency of these indexes when used in kale breeding programs, as they may not consider the interdependencies between traits that are important for selection.

Z-index of Mendes et al. (2009)

According to the Z-index of Mendes et al. (2009), the best-performing genotypes were UFLA 02, UFLA 29, UFLA 09, UFLA 13, UFLA 12, UFLA 30, UFLA 14, UFLA 32, UFLA 18, and UFLA 19. This index resulted in a selection gain of 18.45% for TLY, which was 35% lower than the gain achieved with Mulamba and Mock's index. Additionally, it provided a gain of 4.58% for the number of sprouts. However, the gains obtained for ALM and NL were lower, with values of 10.67 and 6.85%, respectively.

The Z-index is a useful method for selecting traits with normal or non-normal distributions and can be used to select individuals with high performance in one or more characteristics (Mendes et al., 2009). This index is based on the standardization of variables, aiming to make them directly comparable by solving the scaling problem.

Despite its utility, the Z-index has some disadvantages. It is more complex to calculate compared to the simpler Smith and Hazel's index. It requires the calculation of the covariance matrix between traits, which can be challenging when dealing with large sample sizes (Mendes et al., 2009). Additionally, the Z-index may not be suitable for the selection of traits with weak correlations because it may lead to the selection of individuals with average performance across all traits. Therefore, the applicability of the Z-index should be considered in the context of the specific breeding program and the traits of interest.

FAI-BLUP's index

The best-performing genotypes identified by the FAI-BLUP's index were UFLA 14, UFLA 28, UFLA 03, UFLA 06, UFLA 09, UFLA 15, UFLA 19, UFLA 13, UFLA 05, and UFLA 04. This index resulted in a selection gain of 18.25% for TLY. However, it also led to a reduction of 5.8% in the number of sprouts (NS). Additionally, the gains obtained for ALM and NL were lower, with values of 10.67 and 6.85%, respectively.

The FAI-BLUP index is considered a more sophisticated and accurate index compared to the Smith and Hazel index and the Z-index. It uses the methodology of Best Linear Unbiased Prediction (BLUP) to calculate the genetic value of individuals (Rocha et al., 2018; Casagrande et al., 2022; Al-Ashkar et al., 2023). This index is useful for selecting traits with normal and non-normal distributions and takes into account the correlation between traits (Rocha et al., 2018). However, the method presents a more complex approach to be calculated than the previous indexes.

This index was the only one to promote simultaneous reduction for NS, demonstrating its efficiency in recommending genotypes. This index also demonstrated satisfactory performance for the recommendation of sorghum (Silva et al., 2018), chickpea (Sellami, Lavini, & Pulvento, 2021), and wheat (Meier, Marchioro, Meira, Olivoto, & Klein, 2021).

Among all the genotypes recommended by the different indexes, only one genotype (UFLA 13) was ranked among the top ten by all indexes (Figure 1A). Generally, the indexes that best approximated direct selection for TLY were the Mulamba and Mock Index (MM) and the FAI-BLUP index. These indexes showed 60% similarity in the selection of genotypes when considered together. However, when examining each method separately, the FAI-BLUP index showed 50% similarity with the genotypes recommended by direct selection for TLY, while the MM index showed 80% similarity.

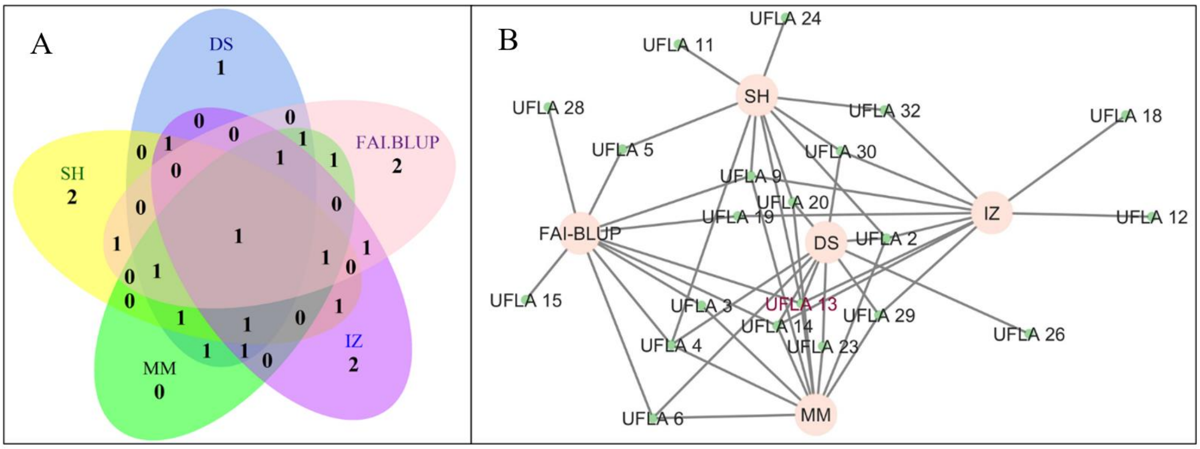

Figure 1. Comparison among selected genotypes: (A) Venn diagram showing similarity between genotypes selected by various indexes and (B) Network diagram illustrating relationships among genotypes selected by four evaluated indexes and direct selection for TLY. DS = Direct selection; FAI.BLUP= FAI-BLUP index; IZ= Z-index; MM= Mulamba and Mock index; SH= Smith and Hazel index.

Notably, the performance of the Smith and Hazel index showed limitations, particularly in terms of TLY and the undesirable increase in NS. This suggests that the Smith and Hazel index may not be the most suitable choice for kale breeding programs.

On the other hand, the Mulamba and Mock index exhibited superior performance for TLY, indicating its potential as a selection tool. However, it was associated with a limited gain in reducing the number of sprouts, which could be a drawback depending on the breeding goals.

The Z-index, although successful in improving other crops like beans and rice (Mendes et al., 2009; Lima, Abreu, Ferreira, & Ramalho, 2015), did not perform satisfactorily for kale, as it resulted in gains for NS and suboptimal performance in terms of TLY.

In contrast, the FAI-BLUP index demonstrated its effectiveness by achieving a reduction in NS (-4.69%) while also providing gains in TLY (Table 3). According to Brito et al. (2020), this is particularly relevant as NS is a crucial consideration in kale breeding programs. The FAI-BLUP index appears to fulfill this demand effectively.

The superiority of the FAI-BLUP index is attributed to its ability to align with specific breeding objectives, which include reducing the number of sprouts. However, the choice of the most suitable index should always be tailored to the requirements of the breeding program and the priorities set by breeders.

It is important to underscore the significance of minimizing the number of sprouts (Azevedo et al., 2014; Brito et al., 2020) in kale breeding programs. High sprout numbers necessitate additional labor for removal, and secondary branches can potentially reduce the availability of metabolites to the canopy, negatively impacting leaf production (Taiz et al., 2017). Consequently, NS is a critical trait to consider in kale breeding programs compared to chlorophyll content.

Conclusion

The FAI-BLUP index demonstrates great promise as a methodology for application in kale breeding programs. Notably, it exhibits the ability to effectively manage and, in some cases, reduce undesired traits, such as the number of sprouts, while simultaneously generating positive improvements in traits of interest. Consequently, the FAI-BLUP index emerges as a valuable and practical tool for enhancing the genetic potential of kale genotypes.

Acknowledgements

To the *Fundação de Amparo à Pesquisa de Minas Gerais* (FAPEMIG), the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq), for providing financial resources and scholarships that supported the execution of this project, and the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES), under financial code 001, for their contributions to this research.

References

Al-Ashkar, I., Sallam, M., Almutairi, K. F., Shady, M., Ibrahim, A., & Alghamdi, S. S. (2023). Detection of high-performance wheat genotypes and genetic stability to determine complex interplay between genotypes and environments. *Agronomy*, *13*(2), 1-24. DOI: https://doi.org/10.3390/agronomy13020585

Azevedo, A. M., Andrade Júnior, V. C. D., Pedrosa, C. E., Fernandes, J. S. C., Valadares, N. R., Ferreira, M. A. M., & Martins, R. A. D. V. (2012). Desempenho agronômico e variabilidade genética em genótipos de couve. *Pesquisa Agropecuária Brasileira*, *47*(12), 1751-1758. DOI: https://doi.org/10.1590/S0100-204X2012001200011

Azevedo, A. M., Andrade Júnior, V. C., Pedrosa, C. E., Valadares, N. R., Fernandes, J. S. C., Ferreira, M. R. A., & Martins, R. A. V. (2014). Divergência genética e importância de caracteres em genótipos de couve. *Horticultura Brasileira*, *32*(1), 51-57. DOI: https://doi.org/10.1590/S0102-05362014000100008

Azevedo, A. M., Andrade Júnior, V. C. D., Figueiredo, J. A., Pedrosa, C. E., Viana, D. J. S., Lemos, V. T., & Neiva, I. P. (2015). Divergência genética e importância de caracteres em genótipos de batata-doce visando a produção de silagem. *Revista Brasileira de Ciências Agrarias*, *10*(3), 479-484. DOI: https://doi.org/10.5039/agraria.v10i3a5165

Baker, R. J. (2020). *Selection indices in plant breeding*. Boca Raton, FL: CRC Press. DOI: https://doi.org/10.1201/9780429280498

Barth, E., Resende, J. T. V., Mariguele, K. H., Resende, M. D. V., Silva, A. L. B. R., & Ru, S. (2022). Multivariate analysis methods improve the selection of strawberry genotypes with low cold requirement. *Scientific Reports*, *12*(1), 1-12. DOI: https://doi.org/10.1038/s41598-022-15688-4

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48. DOI: https://doi.org/10.18637/jss.v067.i01

Begna, T. (2021). Role and economic importance of crop genetic diversity in food security. *International Journal of Agricultural Science and Food Technology*, *7*(1), 164-169. DOI: https://doi.org/10.17352/2455-815X.000104

Beloti, I. F., Maciel, G. M., Juliatti, F. C., Finzi, R. R., & Cardoso, D. B. O. (2020). Genetic parameters and line selection of *Cucurbita pepo* based on selection indices. *Bioscience Journal*, *36*(Suppl. 1), 205-216. DOI: https://doi.org/10.14393/BJ-v36n0a2020-53607

Brito, O. G., Andrade Júnior, V. C. D., Azevedo, A. M., Silva, N. O., Fernandes, J. S. C., & Alves, K. A. (2020). Genetic parameters selection gains and genotypic correlations in kale half-siblings progenies. *Emirates Journal of Food and Agriculture*, *32*(8), 591-599. DOI: https://doi.org/10.9755/ejfa.2020.v32.i8.2136

Brito, O. G., Andrade Júnior, V. C. D., Azevedo, A. M. D., Donato, L. M. S., Silva, L. R., & Ferreira, M. A. M. (2019). Study of repeatability and phenotypical stabilization in kale using frequentist Bayesian and bootstrap resampling approaches. *Acta Scientiarum*. *Agronomy*, *41*(1), 1-11. DOI: https://doi.org/10.4025/actasciagron.v41i1.42606

Casagrande, C. R., Mezzomo, H. C., Silva, C. M., Lima, G. W., Souza, D. J. P., Borém, A., & Nardino, M. (2022). Selection indexes based on genotypic values applied to Brazilian tropical wheat breeding. *Agronomy Science and Biotechnology*, *8*, 1-16. DOI: https://doi.org/10.33158/ASB.r171.v8.2022

Chen, H., & Boutros, P. C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics*, *12*(35), 1-7. DOI: https://doi.org/10.1186/1471-2105-12-35

Cruz, C. D. (2016). Genes software –extended and integrated with the R., Matlab and Selegen. *Acta Scientiarum. Agronomy*, *38*(4), 547-552. DOI: https://doi.org/10.4025/actasciagron.v38i4.32629

Cruz, C. D., Regazzi, A. J., & Carneiro, P. C. S. (2012). *Modelos biométricos aplicados ao melhoramento genético* (1. ed). Viçosa, MG: Editora UFV.

Silva, M. J., Carneiro, P. C. S., Carneiro, J. E. S., Damasceno, C. M. B., Parrella, N. N. L. D., Pastina, M. M., ... Parrella, R. A. C. (2018). Evaluation of the potential of lines and hybrids of biomass sorghum. *Industrial Crops and Products*, *125*, 379-385. DOI: https://doi.org/10.1016/j.indcrop.2018.08.022

Dalbosco, E. Z., Krause, W., Neves, L. G., Araújo, D. V. D., Hiega, K. M. R., & Silva, C. G. D. (2018). Parametric and non-parametric indexes applied in the selection of sour passion fruit progenies. *Revista Brasileira de Fruticultura*, *40*(1), 1-8. DOI: https://doi.org/10.1590/0100-29452018282

Dugard, P., Todman, J., & Staines, H. (2022). *Approaching multivariate analysis: A practical introduction* (2nd ed.). London, UK: Routledge; Taylor & Francis.

Fadhli, N. U. R., Farid, M. U. H., Effendi, R. O. Y., Azrai, M., & Anshori, M. F. (2020). Multivariate analysis to determine secondary characters in selecting adaptive hybrid corn lines under drought stress. *Biodiversitas Journal of Biological Diversity*, *21*(8), 3617-3624. DOI: https://doi.org/10.13057/biodiv/d210826

Hassan, S., Ahmad, T., & Hussain, B. (2022). Food preferences and consumption parameters of Cabbage butterfly, *Pieris brassicae* (Lepidoptera: *Pieridae*) on different kale genotypes. *Journal of Plant Disease and Protection*, *130*, 45-55. DOI: https://doi.org/10.1007/s41348-022-00683-8

Hazel, L. N. (1943). The genetic basis for constructing selection indexes. *Genetics*, *28*(6), 476-490. DOI: https://doi.org/10.1093/genetics/28.6.476

Henderson, C. R. (1984). *Applications of linear models in animal breeding* (v. 462). Guelph, CA: University of Guelph.

Henderson, C. R. (1986). Statistical methods in animal improvement: historical overview. In *Advances in statistical methods for genetic improvement of livestock* (p. 2-14). Berlin, Heidelberg, GE: Springer Berlin Heidelberg.

Hidalgo-Contreras, J. V., Salinas-Ruiz, J., Eskridge, K. M., & Baenziger, S. P. (2021). Incorporating molecular markers and causal structure among traits using a Smith-Hazel index and structural equation models. *Agronomy*, *11*(10), 1953. DOI: https://doi.org/10.3390/agronomy11101953.

Jesus, M. S., Passos, A. R., & Diniz, R. P. (2023). Selection indexes and principal components for agronomic and bromatological traits in forage cactus. *Revista Caatinga, 36*(1), 189-198. DOI: https://doi.org/10.1590/1983-21252023v36n120rc

Kardos, M., Armstrong, E. E., Fitzpatrick, S. W., Hauser, S., Hedrick, P. W., Miller, J. M., ... Funk, W. C. (2021). The crucial role of genome-wide genetic variation in conservation. *Perspective*, *118*(48), 1-10. DOI: https://doi.org/10.1073/pnas.2104642118

Lima, D. C., Abreu, Â. D. F. B., Ferreira, R. A. D. C., & Ramalho, M. A. P. (2015). Breeding common bean populations for traits using selection index. *Scientia Agricola*, *72*(2), 132-137. DOI: https://doi.org/10.1590/0103-9016-2014-0130

Meena, K. Y., Kale, S. V., & Meena, P. O. (2014). Correlation coefficient and path analysis in coriander. *International Journal of Scientific and Research Publications*, *4*(6), 1-4.

Meier, C., Marchioro, V. S., Meira, D., Olivoto, T., & Klein, L. A. (2021). Genetic parameters and multiple-trait selection in wheat genotypes. *Pesquisa Agropecuária Tropical*, *51*, 1-9. DOI: https://doi.org/10.1590/1983-40632021v5167996

Mendes, F. F., Ramalho, M. A. P., & Abreu, A. D. F. B. (2009). Índice de seleção para escolha de populações segregantes de feijoeiro-comum. *Pesquisa Agropecuária Brasileira*, *44*(10), 1312-1318. DOI: https://doi.org/10.1590/S0100-204X2009001000015

Mulamba, N. N., & Mock, J. J. (1978). Improvement of yield potential of the ETO blanco maize (*Zea mays* L.) population by breeding for plant traits. *Egyptian Journal of Genetics Cytology*, *7*, 40-51.

Neath, A. A., & Cavanaugh, J. E. (2012). The Bayesian information criterion: background derivation and applications. *Wiley Interdisciplinary Reviews*: *Computational Statistics*, *4*(2), 199-203. DOI: https://doi.org/10.1002/wics.199

Olivoto, T., Lúcio, A. D., Silva, J. A., Sari, B. G., & Diel, M. I. (2019). Mean performance and stability in multi-environment trials II: Selection based on multiple traits. *Agronomy Journal*, *111*(6), 2961-2969. DOI: https://doi.org/10.2134/agronj2019.03.0221

Olivoto, T., & Nardino, M. (2021). MGIDI: Toward an effective multivariate selection in biological experiments. *Bioinformatics*, *37*(10), 1383-1389. DOI: https://doi.org/10.1093/bioinformatics/btaa981

Piepho, H. P. (2009). Ridge regression and extensions for genomewide selection in maize. *Crop Science*, *49*(4), 1165-1176. DOI: https://doi.org/10.2135/cropsci2008.10.0595

Piepho, H. P., & Möhring, J. (2005). Best linear unbiased prediction for subdivided target regions. *Crop Science*, *45*(3), 1151-1159. DOI: https://doi.org/10.2135/cropsci2004.0398

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, AT: R Foundation for Statistical Computing. Retrieved on Jan. 05, 2023. from https://www.R-project.org/

Ramos, J. P., Cavalcanti, J. J., Freire, R. M., Silva, C. R., Silva, M. D. F., & Santos, R. C. D. (2022). Selection indexes and economic weights applied to runner-peanut breeding. *Revista Brasileira de Engenharia Agrícola e Ambiental, 26*(5), 327-334. DOI: https://doi.org/10.1590/1807-1929/agriambi.v26n5p327-334

Reda, T., Thavarajah, P., Polomski, R., Bridges, W., Shipe, E., & Thavarajah, D. (2021) Reaching the highest shelf: A review of organic production nutritional quality and shelf life of kale (*Brassica oleracea var. acephala*). *Plants People Planet*, *3*, 308-318. DOI: https://doi.org/10.1002/ppp3.10183

Resende, M. D. V., & Alves, R. S. (2021). *Genética: estratégias de melhoramento e métodos de seleção* (3rd ed) (p. 171-202). Brasília, DF: Embrapa Florestas.

Rocha, J. R. A. S. C., Machado, J. C., & Carneiro, P. C. S. (2018). Multitrait index based on factor analysis and ideotype-design: Proposal and application on elephant grass breeding for bioenergy. *GCB-Bioenergy, 10*(1), 52-60. DOI: https://doi.org/10.1111/gcbb.12443

Sellami, M. H., Lavini, A., & Pulvento, C. (2021). Phenotypic and quality traits of chickpea genotypes under rainfed conditions in south Italy. *Agronomy*, *11*(5), 1-15. DOI: https://doi.org/10.3390/agronomy11050962

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T, Ramage, D., Amin, N., … Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*(11), 2498-2504. DOI: https://doi.org/10.1101/gr.1239303

Smith, H. F. (1936). A discriminant function for plant selection. *Annals* of *Eugenics*, *7*, 240-250. DOI: https://doi.org/10.1111/j.1469-1809.1936.tb02143.x

Taiz, L., Zeiger, E., Møller, I. M., & Murphy, A. (2017). *Fisiologia e desenvolvimento vegetal* (6. ed.). Porto Alegre, RS: Editora Artmed.

Tiwari, S., Saville, D. J., & Wratten S. D. (2019). Susceptibility of kale cultivars to the wheat bug *Nysius huttoni* (Hemiptera: *Lygaeidae*) in New Zealand. *New Zealand Journal of Agricultural Research*, *63*(3), 467-477. DOI: https://doi.org/10.1080/00288233.2018.1562480

Trani, P. E., Tivelli, S. W., Blat, S. F., Prela-Pantano, A., Teixeira, E. P., Araújo, H. S., … Novo, M. C. S. S. (2015). *Couve-de-folha: do plantio à pós-colheita* (Série Tecnologia Apta. Boletim Técnico IAC, 214). Campinas, SP: IAC.

Vieira, S. D., Souza, D. C., Martins, I. A., Ribeiro, G. H. M. R., Resende, L. V., Ferras, A. K., … Resende, J. T. V. (2017). Selection of experimental strawberry (*Fragaria x ananassa*) hybrids based on selection indices. *Genetics and Molecular Research*, *16*(1), 1-11. DOI: https://doi.org/10.4238/gmr16019052