GMR

F$_2$ population of snap bean (*Phaseolus vulgaris* L.) potential for use in breeding programs

D.P. Albuquerque[1]*, G.C. Entringer[1], G.A. Gravina[1], D.P. Cruz[1], I.M. Pereira[1], C.L. Leite[1], D.H.L. Teixeira[2], M. Vivas[1] and R.F. Daher[1]

[1]Centro de Ciências e Tecnologias Agropecuárias, Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), 28013602, Campos dos Goytacazes, RJ, Brasil.

[2]Universidade Federal Rural da Amazônia (UFRA), 68650000, Capitão Poço, PA, Brasil.

Corresponding author: D.P. Albuquerque
E-mail: dalcirlei.albuquerque@ufv.br

ABSTRACT. Genetic breeding researches begin with controlled hybridizations to explore the genetic variability and the subsequent best materials selection. The present research aimed to assess the F$_2$ population genetic variability and discharge the minor importance variables. Hybridizations between a yellow and the UENF Goytacá lines were used to obtain the segregating population. The F$_2$ seeds were evaluated at field. Among these, 178 indeterminate habit were the primary research focus. The research obtained data on 16 qualitative and quantitative traits, analyzing them later by descriptive techniques to verify the distribution. The REML/BLUP method obtained the genetic parameters estimates and genetic values prediction. The possibility of discarding redundant characters was verified using Singh's methodologies, principal component analysis and correlation coefficients. Grouping analyses by the Tocher and UPGMA methods assessed the genetic variability in the populations, using the standardized mean Euclidean distance as a dissimilarity measure. In the end, the best thirty individuals were chosen for the characteristic of total pods per plant. All traits displayed higher or lower degrees of phenotypic variation. The selected genotypes were auspicious due to the high values observed for the characteristics. The data analysis suggested to be discarded of the pod width

because it was redundant. Moreover, the data highlighted the formation of divergent groups for both grouping techniques, corroborating the existing variability. The thirty best genotypes selection promoted a 7.62% gain from the pods' number per plant. This experience showed that the genealogical method could improve selected genotypes, continuing the UENF snap bean breeding program.

Key words: Correlations; Multivariate analyses; REML/BLUP; Segregating population

INTRODUCTION

The snap bean (*Phaseolus vulgaris* L.) is global economic relevance horticultural produced and consumed in several countries. The worldwide production of this plant is about 21 million tons (FAO, 2019). Even being the same botanical species as the common bean, its consumption is different, using the immature pods, an excellent food source of fibers, vitamins, minerals, carbohydrates, and proteins (Yuan et al., 2017; Fukuji et al., 2019). Furthermore, the crop has good profitability due to its short production cycle and good productivity (Viçosi et al., 2020). Therefore, the product has a great demand both in the national (Vilela et al., 2009) and international markets (Seif et al., 2016).

Even the relevance of this species in the Rio de Janeiro state agriculture, this species still has a potential to be exploited and constant demand of newly adopted materials, productive and with good pod's quality, as there are still few available cultivars (Sant'anna et al., 2020). These factors reflect the need for research aimed at the snap bean regional genetic breeding, using methods of breeding, evaluation, and analysis in segregating populations.

Genetic breeding programs in self-pollination species commonly include controlled crossbreeding, which must be performed between homozygous parents, contrasting for one or more traits and that have positive agronomic attributes. In this form, it is possible to breed and explore the variability in segregating populations, without which, it would be impossible to make progress in plant breeding. Therefore, parent selection for hybridizations is one of the most relevant decisions to be taken by the breeder, providing a sufficient genetic variability increase to obtain gains by selection (Pereira et al., 2007).

The mixed models REML/BLUP is a primary method in the vegetal genetic breeding context to verify the variability existence (Henderson, 1975). This technique promotes precise estimations and predictions for genetic parameters and values, and contributes to increase the selection candidates genetic assessment efficiency (Alves and Resende, 2008). Mixed models have been used excellently in studies performed on the snap bean (Cruz et al., 2018; Gomes et al., 2018).

After the variability existence is verified, multivariate biometric techniques of genetic diversity can examine and explore. These techniques can unify multiple pieces of information, choosing promising genotypes in the breeding programs. Among the multivariate technologies, the principal components and grouping methods stand out (Cruz et al., 2012) and the Singh (1981) method for the relative variables contribution.

The promising genotypes assessment and choice at the initial steps of a breeding program is crucial to reducing the economic costs, time, and handwork to develop new cultivars. Exploiting an F_2 population can provide sufficient information to succeed in the superior plant selection and

strategies for conducting these materials, leading in a short-medium time to the potential inbred lines development, which may register as new cultivars by the *Ministério da Agricultura, Pecuária e Abastecimento* (MAPA).

This study aimed to 1) assess the genetic variability, estimate genetic parameters in an F_2 population, and identify divergent plant groups as to the morphological traits; 2) identify those traits that mainly contribute to the population variability, and discharge those which contribute few in the observed variability or redundant; 3) forecast the gain due to the selection for the pod number and select the best individuals for the generations advancement.

MATERIAL AND METHODS

Genetic Material

Hybridizations were performed in the Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF) greenhouse to obtain the segregating population. The F_2 population was derived from the crossing between a line identified as *Linhagem Amarela* (LA) and the cultivar *UENF Goytacá* (L7). These parents contrasted by pod color (yellow and green, respectively).

The activity was divided into two steps: in the first step, the F_1 generation was obtained (P_1 x P_2), and in the second step, the segregating generation F_2, derived from the self-fertilization of the F_1 plants. In both steps, the seeds were derived from the planting in 10 liters capacity vases, filled with Basaplant® substrate. The F_2 segregating generation seeds were analyzed in a field-conditions experiment.

Field experiment performance

The experiment was performed at the PESAGRO-Rio experimental area, in the Campos dos Goytacazes city, RJ, Brazil (latitude 21º19'23 "S and longitude 41º19'40" W). The soil is classified as dystrophic argisol, and the regional climate is classified as "Aw" according to the Köppen classification, with a mean year temperature of 22.4 °C, and mean rainfall volume of 1.125 mm. Rains occur mainly between October and March (Xavier et al., 2022).

F_2 generation seeding was performed in March/April 2020. This step used 320 pits divided into ten 16 meters-long lines, spaced 1.0 m x 0.5 m. The central 30 plants of each line were used for the analysis. The plants at both ends were used as borders. Besides this, two lines were planted at the sides, surrounding the experiment completely.

Two seeds per pit were used for sowing at an approximate depth of 2.5 cm. After emergence, plants were thinned to leave just one plant per pit. About 15 days after the emergence, the plants were staked using wire and bamboo so that the wire fixed the bamboo sticks with the crossed points. Cultural and phytosanitary treatments were performed as recommended for the culture (Filgueira, 2008). Plants were irrigated by aspersion during the experiment.

Nine of the initial 300 plants died during the first developing phases, and 78 did not produce enough seeds for the generations' advancements, being discharged from the research and breeding program. Besides this, even if both parents displayed indeterminate growth habits (HI), the F_2 also displayed plants with determinate growth (HD). These latter plants were not used for the present research. One hundred seventy-eight individuals remained in the population after the eliminations derived from the natural selection and the removal of HD individuals.

Evaluated characteristics

For each plant, 16 culture and post-harvest stage evaluated characteristics were assessed: number of days to emergence (NDE) - number of days since seeding to the seedlings emergence; number of days to flowering (NDF) - number of days since the seeding until the first flowers emergence; anthocyanin in the hypocotyl (AH) - 1: presence; 2: absence, assessed at the seedlings initial growth; flower color (FC) - assessed at flowering, in fresh-opened flowers; pod color (PC) - assessed during the physiological ripening; pod transversal section (PTS) - 1 cylindrical; 2 flattened, assessed after the immature pods harvesting, performing a transversal cut of five to ten pods per plant; seed size (SS) - assessed by the weight of 100 seeds, where: 1: small $\leq$ 25g; 2: medium 25 $\leq$ 40g; 3: big > 40g; plant height (PH) - obtained by measuring the height correspondent to the distance between the collar until the end of the main stem, expressed as meters; first pod insertion height (FPIH) - distance between the plant collar and the first pod insertion, in centimeters; total number of pods per plant (TNPP) - quantification of the total pods number after harvest; pod green matter mass (PGMM) - quantification of each plants green pods total mass (g) , obtained by precision scale; green matter mass of one pod (GMMOP) - quantified by the PGMM/TNPP ratio; pod width (PW) - measured as millimeters at the central part of the pod, from the ventral to the dorsal suture in a ten fruits per plant sample; pod diameter (PD) - measured with a precision caliper as millimeters in a ten fruits per plant sample; pod length (PL) - longitudinal length of the pods, including the apical tooth, measured with a plastic rule in a ten pods per plant sample, and expressed as centimeters; mass of one hundred seeds (MOHS) - obtained with a precision scale and expressed as grams.

Genetic and statistical analyses

The descriptive analysis assessed data distribution variables, obtaining column graphs and box plots for the qualitative and quantitative characteristics, respectively.

1. Genetic parameters estimate and superior individual selection

The study performed individual analysis on the per plant values of the studied traits. The software used to estimate the variables and predict the genetic values was the SELEGEN REML/ BLUP (Restricted Maximum Likelihood – Best Linear Unbiased Prediction). The program used the BLUP (Best Linear Unbiased Prediction) procedure to predict the genetic values using the variance estimates obtained by the REML (Restricted Maximum Likelihood) method (Resende, 2016).

The above-cited analysis was performed by the model number 58 of the SELEGEN software, considering the following estimator:

$$y = Xu + Zg + Wl + Tc + e$$

Where: y is the data vector, u is the scalar referent to the general mean, g is the genetic effects vector (assumed to be random), l is the vector of the line effects (assumed to be random), c is the vector of the column effects (random), e is the vector of errors or residues (random). This model applies to experiments without experimental design, suggested for genetic evaluations of equally spaced plants, such as in commercial plantings, where lines and columns are attributed to a plantation rectangle, appropriate for the experimental conditions used in the study.

Based on the predicted genetic values, the thirty best individuals were selected based on the total number of pods per plant (TNPP) for generation advance and breeding program continuity.

2. Discarding of the redundant variables or those less contributing to the variation

The variables elimination that contributed less to the plants discrimination or of the redundant ones was performed according to the following criteria: 1) relative contribution of the variable to the total variation, based on the Sigh method (1981); 2) relevance variable to explain the total variation, by the principal components analysis (PCA), where the disposable variables were those with the highest weighting coefficient, as its absolute value, in the last principal components, which concentrated few variance (Cruz et al., 2012), which eigenvalues are lower than 0.7 (Jolliffe, 1973; Mardia et al., 1979); 3) redundant variables by correlation: where the variables displaying high correlation with another which contributed more to the plant discrimination were deemed as redundant and disposable, as recommended by Cruz et al. (2012). The Pearson's correlation coefficients among the traits were estimated for the BLUP means to obtain genetic correlations approximate estimates.

3. Estimation of the genetic divergence among F_2 individuals

The genetic divergence estimates were obtained by groupings based on the nine quantitative variables data. The standardized Euclidean distance was used as dissimilarity distance. The grouping based on the Toucher optimization method and the unweighted pair group method with arithmetic mean (UPGMA) were performed from the genetic distance matrix (MDG). In the UPGMA, the adjustment between the genetic distance matrix and the grouping matrix was performed by the cophenetic correlation coefficient (CCC) (Sokal and Rohlf, 1962), and the mojena method was used to establish groups (Mojena, 1977).

RESULTS

Descriptive statistics

The column graphs displayed that all the qualitative traits displayed phenotypic variability (Figure 1). The plants emerged from the fourth to the seventh day. Most of the plants emerged on the fifth day (Figure 1a). The number of days to flowering varied widely. The plants flowered between 28 and 44 days (mean 34.5). The highest flowering rate was on the 33rd day (Figure 1b).

Most plants presented anthocyanine in the hypocotyl (Figure 1c). The flower color (FC) segregated the plants into four types displaying white, pink, violet, or bicolor flowers. Pink and violet were predominant (Figure 1d). Pod color (PC) displayed three different segregated varieties: green, yellow, and a mixture of both colors, here named "green & yellow" (Figure 2). Most plants displayed green pods (Figure 1e) and a flattened form (Figure 1f). Besides, most of the plants displayed medium-sized seeds (Figure 1g).

From the quantitative traits behavior evaluation, through the boxplots graphs analysis, it was observed that all of them presented, to a greater or lesser range, phenotypic variation (Figure 3).

The plant height (PH) varied from 1.25 to 3.8 m, and the mean height was 2.6 m (Figure 3a). The first pod insertion height (FPIH) varied from 8 to 48 cm, and the mean height was 19.4 (Figure 3b). As to the green matter mass of one pod (GMMOP), varied from 4.5 to 12.7g, with a mean of 8.1g (Figure 3c). The traits PW, PD, and PL displayed mean values of 10.6 mm, 6.8 mm, and 14.8 cm, respectively (Figures 3d, 3e, and 3f).
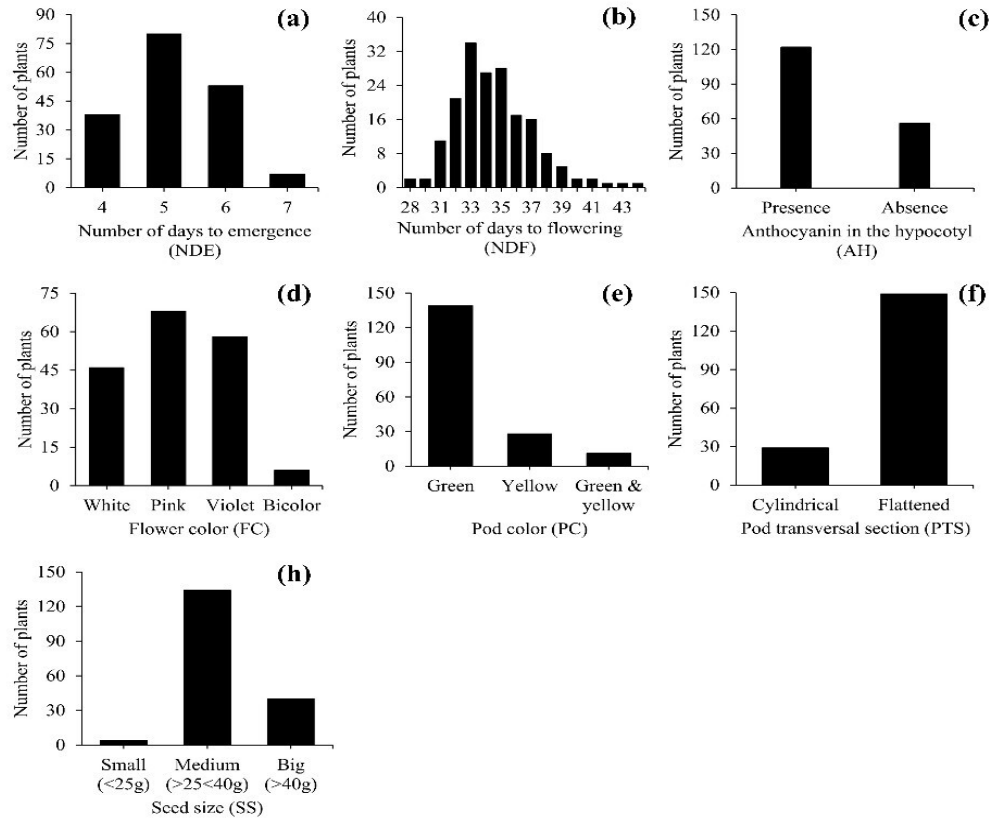
Figure 1. (a) Number days off to emergence (NDE), (b) number days off to flowering (NDF), (c) anthocyanins in the hypocotyl (AH), (d) flower color (FC), (e) pod color (PC), (f) pod transversal section (PTS), and (g) seed size in snap beans of indeterminate growth. Campos dos Goytacazes, RJ, Brazil.

Figure 2. Observed segregation as the pod's color: a- green; b- yellow, and c-green & yellow.
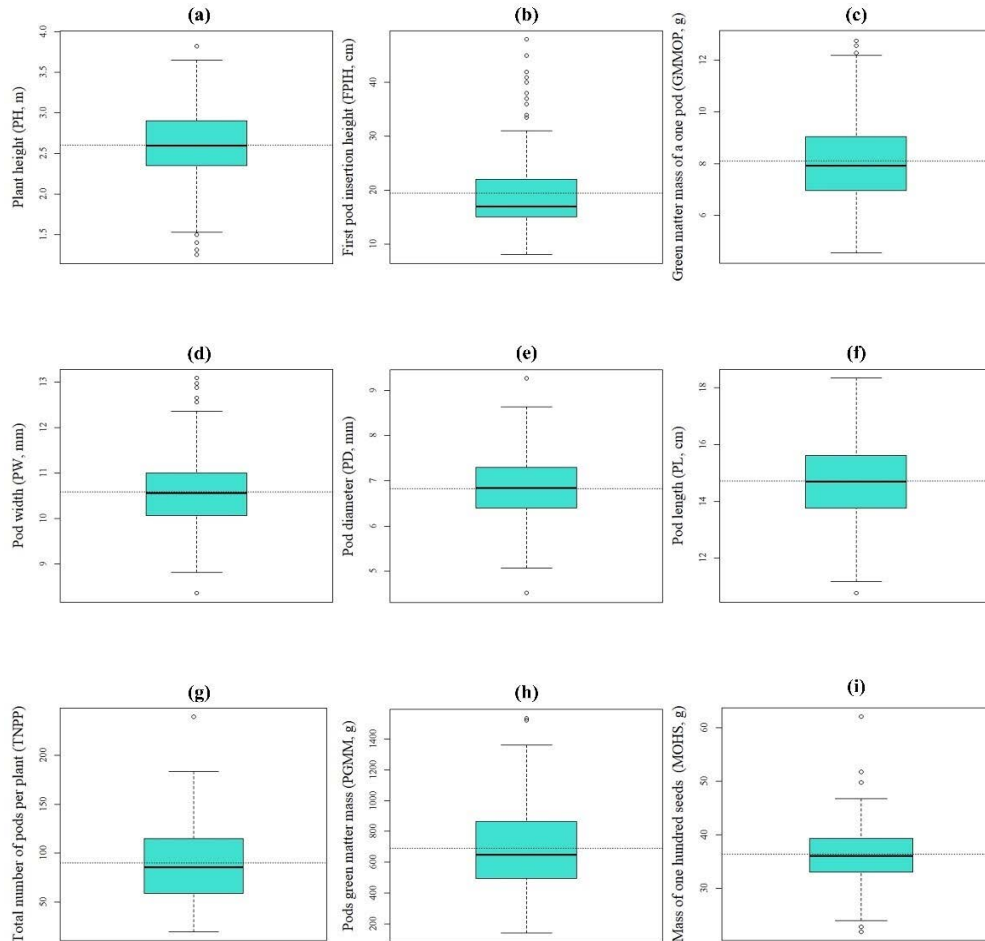
Figure 3. (a) Plant height (PH), (b) first pod insertion height (FPIH), (c) green matter mass of one pod (GMMOP), (d) pod width (PW), (e) pod diameter (PD), (f) pod length (PL), (g) total number of pods per plant, (TNPP), (h) pods green matter mass (PGMM), and (i) mas of one hundred seeds (MOHS) in snap beans of indeterminate growth. Campos dos Goytacazes, RJ, Brazil.

The characters TNPP, PGMM, and MOHS, are intrinsically related to the yield. These traits displayed mean values of 90, 691g, and 3.4g, respectively (Figures 3g, 3h, and 3i).

Estimates of genetic parameters via REML

The estimate of genetic variance (σ_g^2) presented good values for all evaluated traits, to a greater or lesser extent, it ranged from 0.024 to 8264.424 for the traits ALTP and MMVV, respectively (Table 1). In basic terms and without considering the interaction effect, the phenotypic variation is the sum of the genetic and environmental variances, where the highest parcel of the first one plays a key role in the promoting the breeding process. However, the data highlighted that the variation due to environmental factors ($\sigma_{line}^2 + \sigma_{column}^2 + \sigma_r^2$) was higher than the genetic one for all the analyzed traits.

Table 1. Genetic parameters to nine indeterminate growth snap bean characteristics. Campos dos Goytacazes, RJ, Brazil.

Character	Genetic parameters						Mean
	σ^2_g	σ^2_{line}	σ^2_{column}	σ^2_r	σ^2_f	h^2_g	
PH [a]	0.024	0.018	0.007	0.161	0.210	0.112±0.071	2.592
FPIH [b]	6.069	8.623	0.140	39.400	54.232	0.112±0.071	19.289
GMMOP [c]	0.318	0.004	0.035	2.075	2.432	0.131±0.077	8.145
PW [d]	0.084	0.004	0.002	0.537	0.628	0.134±0.078	10.578
PD [e]	0.063	0.020	0.003	0.411	0.497	0.127±0.076	6.828
PL [f]	0.265	0.027	0.004	1.679	1.975	0.134±0.078	14.720
MOHS [g]	3.903	3.197	0.091	25.256	32.448	0.120±0.074	36.312
TNPP [h]	158.107	178.129	170.837	1109.247	1616.320	0.098±0.067	88.787
PGMM [i]	8264.424	12445.374	5840.437	57582.819	84133.054	0.098±0.067	679.911

[a] PH: plant height (m); [b] FPIH: first pod insertion height (cm); [c] GMMOP: green matter mass one pod (g); [d] PW: pod width (mm); [e] PD: pod diameter (mm); [f] PL: pod length (cm); [g] MOHS: mass of one hundred seeds (g); [h] TNPP: total number of pods per plant; [i] PGMM: pods green matter mass (g).

The heritability parameter quantifies the trustworthiness of the phenotypic value as a guide for the genetic value, the greater value, the more favorable for obtaining gains with the advancement of generations. All the traits presented low individual heritability in a broad sense, reflecting the high environmental variance. The individual heritability varied between 9.8% for TNPP and PGMM to 13.4% for PL and PW.

Redundant variables discard

The Singh method (1981) pointed out that the variable with the highest genotypes separation power was PGMM (15.02%), followed by GMMOP (12.90%), PL (12.23%), FPIH (11.95%), and TNPP (11.79%). These variables alone explain 63.89% of the observed variation (Figure 4). On the other side, the variables that contributed less to the genetic variation and could be depicted by the proposed method were PD (7.85%) and MOHS (7.08%).
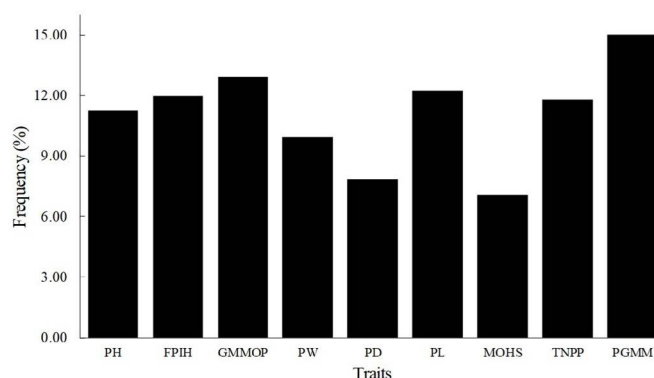
Figure 4. Relative contribution of the intermediate habit snap beans features to the genetic diversity by the Singh criterion (1981). Campos dos Goytacazes, RJ, Brazil. PH: plant height (m); FPIH: first pod insertion height (cm); GMMOP: green matter mass one pod (g); PW: pod width (mm); PD: pod diameter (mm); PL: pod length (cm); MOHS: mass of one hundred seeds (g); TNPP: total number of pods per plant; PGMM: pods green matter mass (g).

PCA method allowed the discard of the seventh Principal Component (PC7), displaying eigenvalues lower than 0.7. The total number of pods per plant (TNPP) was the first trait suggested for discard, as displaying the highest weighting coefficient, in module, in the last principal component (PC9). Next, both the green matter mass of one pod (GMMOP) and the pod width (PW) were suggested for the discard due to the higher weighting coefficients on the PC8, and PC7, respectively (Table 2).

Table 2. Variances (eigenvalues), percentage, and accumulated variances of principal components obtained with the basis of nine characteristics of indeterminate habit snap bean. Campos dos Goytacazes, RJ, Brazil.

Component	Eigenvalue	RH (%)	Accumulated RH (%)	Discharge
1	2.5482908	28.31	28.31	
2	1.8024802	20.03	48.34	
3	1.3400257	14.89	63.23	
4	1.1530102	12.81	76.04	
5	0.8078949	8.98	85.02	
6	0.7450414	8.28	93.30	
7	0.3473008	3.86	97.16	PW [a]
8	0.2397276	2.66	99.82	GMMOP [b]
9	0.0162284	0.18	100.00	TNPP [c]

[a] PW: pod's width (mm); [b] GMMOP: green matter mass one pod (g); [c] TNPP: total number of pods per plant.

The simple correlation estimates displayed that, in general, the correlation values were low and not significant (Table 3). The correlations between TNPP and PGMM (0.89**), GMMOP and PL (0.70**), and PW and PD (0.61**) exhibited highly significant and positive associations. Besides, there was a negative significant correlation between GMMOP and TNPP (-0.39**).

Table 3. Phenotypic correlation among indeterminate habit snap bean characteristics. Campos dos Goytacazes, RJ, Brazil.

	PH	FPIH	GMMOP	PW	PD	PL	MOHS	TNPP	PGMM
PH	1	0.04^{ns}	-0.01^{ns}	0.05^{ns}	-0.10^{ns}	-0.05^{ns}	0.01^{ns}	0.05^{ns}	0.05^{ns}
FPIH		1	0.10^{ns}	0.02^{ns}	-0.13^{ns}	0.02^{ns}	0.15^{*}	-0.17^{*}	-0.11^{ns}
GMMOP			1	0.41^{**}	0.42^{**}	0.70^{**}	0.23^{**}	-0.39^{**}	-0.01^{ns}
PW				1	0.61^{**}	0.25^{**}	-0.16^{*}	-0.12^{ns}	0.03^{ns}
PD					1	0.21^{**}	-0.04^{ns}	-0.11^{ns}	0.04^{ns}
PL						1	0.16^{*}	-0.18^{*}	0.10^{ns}
MOHS							1	-0.10^{ns}	0.01^{ns}
TNPP								1	0.89^{**}
PGMM									1

**, *, and ns: respectively significant at 1%, 5%, and not significant by the t-test.

Genetic divergence by groupings

The grouping by the Toucher method obtained 12 divergent groups. Most of the genotypes were allocated in group I. Some groups contained only one genotype (Table 4).

Table 4. Genetic similarity groups obtained by the Tocher Optimization method, based on nine characteristics assessed on the 178 indeterminate growth snap bean plants. Campos dos Goytacazes, RJ, Brazil.

Group	Genotypes									
I	128	166	205	188	154	136	124	4	114	26
	176	178	135	132	60	183	193	82	21	107
	133	155	1	16	113	144	71	212	137	89
	152	175	10	65	209	190	151	67	160	191
	25	201	210	105	134	104	186	86	148	99
	73	106	11	76	37	177	88	59	3	36
	206	159	96	70	85	173	120	50	20	179
	87	81	58	7	145	149	98	211	172	208
	195	115	143	44	54	45	110	42	92	78
	162	32	5	75	204	9	84	203	197	194
	33	39	35	72	62	27	165	164	157	2
	14	118	29	94	95	69	103	167	40	55
	140	189	129	46	91	52	202	68	153	83
	24	63	192	199	43	77	112	18	23	123
	6	66	127	185	90					
II	150	180	108	131	198	126	80			
III	28	93	15	47	19	13				
IV	56	109	119	121	122					
V	102	117								
VI	171	213	181	207						
VII	184	200	139							
VIII	12	111								
IX	57									
X	17									
XI	101									
XII	146									

A cutting point at 69% distance was established to define dissimilarity among groups by UPGMA. The method disclaimed just six dissimilar groups, half of those observed using the previous method (Figure 5). The cophenetic correlation coefficient (CCC) obtained was 0.64, significant ($P<0.01$), which demonstrates that the genetic distance matrices are well represented in the dendrogram.
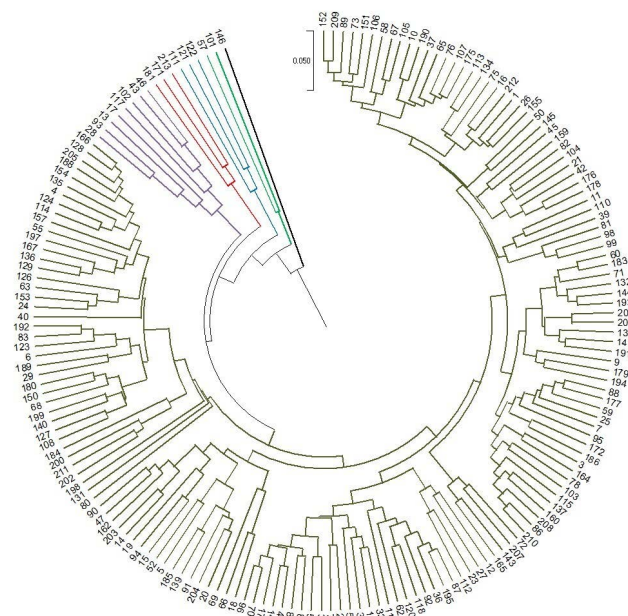
Figure 5. Dendrogram obtained by the unweighted pair group method with arithmetic means (UPGMA), based on 178 indeterminate growth snap bean plants. CCC=0.64**. Campos dos Goytacazes, RJ, Brazil.

Table 5. Predicted genetic value (μ + g) and new selected plants' mean for the characteristic total pods number/plant on the F₂ indeterminate growth snap bean population. Campos dos Goytacazes, RJ, Brazil.

Selected individuals	μ + g	New mean	Selected individuals	μ + g	New mean
121	106.766	106.766	120	94.906	97.025
195	99.316	103.041	78	94.801	96.894
118	97.266	101.116	207	94.699	96.772
23	97.198	100.137	111	94.684	96.662
27	97.049	99.519	2	94.546	96.556
146	96.839	99.072	149	94.214	96.445
119	96.578	98.716	7	94.007	96.334
12	96.454	98.433	112	93.780	96.223
109	96.289	98.195	33	93.590	96.113
122	96.132	97.989	44	93.563	96.011
35	96.058	97.813	36	93.466	95.913
77	95.670	97.635	92	93.381	95.819
56	95.507	97.471	73	93.214	95.726
62	95.242	97.312	151	93.126	95.637
87	95.125	97.166	58	93.071	95.551
General mean			88.787		
Selected plants' mean			95.551		
Predicted gain			6.764		
Percentage gain			7.62%		

DISCUSSION

Natural selection pressure is one of the main factors playing on the variability existing in a population and interfering with the plants' development. As a result, some of these plants suffer due to this stress and may die, not produce seeds, or not produce a sufficient amount of them to produce new generations, as observed in the present research.

Natural selection concentrates, in a specific population, the genetic heritage of the individuals displaying higher adaptive value, reproducing more, and creating a considerable number of viable progenies. In this form, in each new generation, their genes overcome those of the other components of the group, and population genotypes concentrate as adaptative advantages (Lobô and Lobô, 2007).

Descriptive Statistics

The analyzed material displayed higher emergence precocity than Almeida et al. (2014) described using the same culture. The authors' experience described a mean emergence beginning at seven days after seeding. Besides this, F_2 individuals were more precocious than commercial varieties and common beans intermediate habit advanced lines, as Correa et al. (2003) described.

As to NDF, Abreu et al. (2004) observed means varied between 31.25 to 46.75 days and considered precocious the accesses UENF-1430 and UENF-1448, both displaying a mean of 33.75, at the beginning of the breeding program for the indeterminate habit snap bean at the UENF. The genotypes analyzed in the present research displayed lower amplitude, highlighting the higher number of flowering plants on the 33rd day (Figure 1b), which confirms, once again, that most plants are precocious. Previous papers have already described the high variability associated with this traits (Peixoto et al., 2002). The result for NDF shown is relevant to the crop cycle, making it shorter, consequently affecting the time the crop remains in the field. Since the shorter the time in the field, the lower the costs with manual labor and supplies, in addition to reducing the period of exposure to pests and diseases.

The anthocyanins presence or absence in the hypocotyl (AH) is the minimal descriptor for registering to the official department (*Registro Nacional de Cultivares do Brasil*). However, this element is not entirely accepted as a variety descriptor, as sun exposition can influence it, producing variations, mainly of color intensity. The FC is a variable descriptor in the same form, which varies according to the daytime. Therefore, its identification must always be performed at the first daylight, as sunlight rapidly reduces the color tonalities. This protocol was followed in the present research.

As to the trait PC, it can be highlighted that, even if one of the classes is a combination between the remaining two, it is not a codominance or incomplete dominance inheritance, as the low number of either green & yellow pods, which does not fit to the expected proportion in an F_2 population, which is 1:2:1. The minor discrepancies observed in the pods displaying color mixture can be due to differences in the penetrance and expressiveness of the allele responsible for the yellow color, as previously observed in work on the seed color genetic control in beans (Baldoni et al., 2002).

The results point out low diversity between the genotypes as to this trait. It is pretty interesting the matter of market acceptability: green pods, which are the majority still today, are more accepted by the natural product consumers.

It is noteworthy that even the same yellow pods may represent a new production trend. Foreign cuisine, such as the Japanese, is more and more present, highlighting the relevance of bright-colored dishes. Applying different colors to the foods can incorporate visual, taste, olfactory, and evocative aspects. In this context, Minich (2019) proposes the concept of "eating by colors" and points out that it can be a more efficient form to involve a long-term lifestyle change.

Regarding to PTS (Figure 1f), the present research results resemble those of Barbé et al. (2010), as the authors assessed 120 advanced F_{6-7} lines, where most of the plants displayed flattened pods (*manteiga*-type). Furthermore, the authors claimed that consumers from the Rio de Janeiro State prefer *manteiga*-type pods, a product with the highest commercialization volume for the crop.

The SS result is also promising. Peixoto et al. (2002) claimed that larger size is not desirable, conferring to the bud's external surface an undesirable apparency due to the bulges, requiring the harvesting of pods below the desired size for the producer reducing the productivity in this form.

The mean PH value was higher than Almeida et al. (2014), who described 2.41 m. In the same paper, the authors compared their lines with the commercial controls viz., Top Seed Blue Line and Feltrin, obtaining mean values about 2.7 m. This datum points out that most genotypes of the studied population have similar heights or are higher than the known commercial varieties.

The soil conditions, irrigation, culture handling, climate, mild temperatures, typical of the experimental time and favorable for the snap bean, explained the excellent plants development and height.

The plants displayed a much higher first pod insertion than Abreu et al., (2004) and Francelino et al., (2011) described. The breeders find the high first pod insertion genotypes interesting, as the lowest pods may quickly deteriorate due to the contact with the soil, reducing the product quality and increasing harvesting losses.

GMMOP was higher than found by Marinho et al., (2013) in Bom Jesus do Itabapoana in 17 F_8 lines of snap bean (2.36g), and similar to the results observed in the F_9 in the municipalities of Bom Jesus do Itabapoana (8.57g) and Cambuci (7.54g). This result is due to the use of flat pods genotypes, which have higher mass than the round pods, according to Peixoto et al., (2002). This result is significant from the farmers final product quantitative point of view, as it would mean a lower number of pods per kilogram of the product.

The pods were less wide than Almeida et al., (2014) described. This is a prominent characteristic for the product acceptance by the consumers. This work highlighted that most of the genotypes displayed desirable width, similar to other materials in the market.

As to PD, the result resembled Silva et al., (2004), as the authors assessed five parentals lines and their ten mixture snap bean hybrids (7.06 mm). It is well known that consumers prefer smaller diameter pods, such as the Paulista cultivar (7.24 mm). The analyzed genotypes displayed medium to high PD acceptability in the market.

The PL results seemed quite hopeful. A proof of this is that the mean value was similar to that described by Sant'anna et al., (2019) as presenting the cultivar UENF Goytacá (15 cm). This same cultivar was the male parent in the crossings that generated the F_2 population used in the present research. It shall be pointed out that almost 50% of the genotypes analyzed here displayed values higher than 15 cm, highlighted by the equivalence between the median and the general mean (Figure 3f).

The trait TNPP displays a high association with the beans production, which is highly relevant due to the direct contribution to the pod productivity. The mean was considered quite high, higher than the highest mean observed by Francelino et al., (2011) as describing 25 snap bean accesses from the UENF Germoplasm bank, and higher than depicted by Sant'anna et al., (2019) in the cultivars UENF Goytacá (83.93) and Top Seed Blue Line (75.56). Both these cultivars are known to be productive and are available to the farmers.

The high pods mass produced by the genotypes reflected directly on the pods green matter mass per plant (PGMM). So, the genotypes that had higher weights (represented by the outliers) were possibly those which produced the highest pods' amount (Figures 3g and 3h).

The mass of one hundred seeds (MOHS) was higher than the mean observed by Abreu et al., (2004) in 24 out of the 25 accesses studied at the beginning of the UENF snap bean breeding program. Also was higher than the mean the UENF Goytacá cultivar, which display a mean 29.25g value for this trait (Sant'anna et al., 2019).

All traits evaluated in the present research displayed good phenotypic variability values, promising genotypes compared to the means displayed by known performance cultivars and lines, the overall experiment means, and the mean of the UENF Goytacá cultivar. There is, then, a possibility of individual genotypes selection to continue the UENF snap bean breeding program.

Estimates of genetic parameters via REML

Even displaying low values in most cases, the genetic variance results suggest a genetic variability among the genotypes. PH and PGMM estimates reflect a considerable genetic variability or difference among the genotypes, which can be exploited for breeding purposes. These traits connect to the pod productivity, which is required in an intermediate habit breeding program.

The high environmental variance estimates reduced the heritability indexes, becoming a significant hamper for breeding. Even though it is crucial to point out, high environmental variance is mainly observed as dealing with quantitative variables (Cruz et al., 2018), and some factors may have influenced the observed results directly, such as the trait polygenic nature, the lack of an experimental design, and the fact that it was a highly segregating population. Therefore, it is possible to claim that the selection can be performed, but it will be more challenging to obtain gains. This difficulty can be bypassed by a higher control on the environmental variables in the next generations, maximizing the success probability.

The comparison of genetic parameters results to heritability is biased, as the latter may vary according to the analyzed trait, estimation method, population endogamy, genetic diversity, and experimental precision (Falconer and Mackay, 1996). In this form, even if a specific trait may display similar results, heritability values are specific for each population and environment to which the analyzed matters are submitted.

The most snap bean crop relevant traits, such as TNPP, displayed low heritability values. However, still, there is a great success possibility, which justifies the individual plants selection in the population for further generations advancement, most especially as being traits controlled by different genes, which are generally bred gradually. The use of more complex statistics, genetic procedures, and experimental design may help in this aim to select and recommend new cultivars with higher market relevance than those that are available.

Redundant characters discard

According to the Singh method (1981), the most contributing variables are strictly associated with the pods productivity, proofing the presence of genotypic and phenotypic variability in selecting candidates for these relevant traits.

Abreu et al. (2004) found discordant results for MOHS, which is a prominent traits in analyze accesses diversity, modifying the grouping as retrieved from the analyses. However, these authors also observed a great PL contribution, agreeing with the present research.

The Principal Components Analysis suggested discarding TNPP from the last principal component. This is one of the most critical trait of the crop considering the intermediate habit. It is not uncommon in the literature for this method suggests important variables to describe a particular culture, justifying the adoption of more than one methodology.

As to the correlation estimates, it is necessary to point out that the correlation coefficient equal to zero does not implicate a lack of correlation between two variables, reflecting just the lack of linear relationship between these two evaluated variables. The correlations between TNPP and PGMM and GMMOP and PL were considered as strong. That between PW and PD, moderate (Dancey and Reidy, 2018). The strong association between the variables allow an indirect breeding: selecting a unique trait provides simultaneous gains in those strongly associated variables (Costa et al., 2004).

It is essential to point out also that the negative correlation between GMMOP and TNPP suggests that the higher the mean single pod weight, the lower the pods number per plant. This observation may associate with plants' physiological elements: as they direct their metabolites to more robust pods' production, they reduce their number.

Considering the three methods together, the discard of the PW variable not represents a loss of information. This variable displayed a moderated positive correlation with the pod's diameter. The principal component analysis suggested its discard and displayed little contribution according to the Sigh method, which considered it redundant. The dispensable variables in divergence studies display low variability among the studied material or are redundant as they correlate with other variables (Cruz et al., 2012). This justifies the decision to discard this variable.

Genetic divergence by groupings

Four groups had only one genotype, probably due to the existence of the outliers, which is common in segregating individuals. As to the UPGMA method, the CCC significance confirms the good adjustment of the dissimilarity matrix to the dendrogram. There has been a low agreement between the groups formed by the two methods. Genotype 146 is an exception, as both techniques pointed it out as an isolated group.

Amaral Júnior et al. (1996) also highlighted contrasting results between the groupings obtained by using two different grouping methods, as assessing the genetic diversity among pumpkin accesses (*Cucurbita maxima* Duch.). Besides, the disagreement among the results can also occur using a unique method and the same materials, such as described by Teixeira et al. (2004) analyzing snap bean genotypes during two different seasons (autumn/winter and summer).

The groups formation by the presented methods highlights the variability in the population, as it had been previously observed in each trait separately by the columns and boxplots graphics.

The advancement of generations in future works may entail inbred lines, presenting interest traits and superior performances.

Best genotypes selection via BLUP

One of the snap bean breeding programs main aims, as in other crops, is the search for traits appreciated by the consumers' market the release of more productive materials to provide the producer with higher commerce profitability (Andrade et al., 2017). In the literature, there are few studies on early selection in snap bean, but generally, it is forwarded to production traits, such as the pod number and weight per plant (Cruz et al., 2018).

In the present research, the most important traits are the number of pods and the mass of pods green matter per plant (TNPP and PGMM). These traits show a solid and significant correlation, as described previously (Table 3). The selection for one trait produces simultaneous advancements in the other feature too. The priority use of TNPP genotypic values ranking and prediction based on criteria such as the higher importance for the crop of good production performance of green pods per plant aims to potentiate the genetic gains for the best plants selection into the selected population.

The mean value obtained for this traits was 88.787 (Table 1) and, even if this trait has been little used to the selection of individual plants in studies on the snap bean, these values were lower than those described by Cruz et al. (2018). It is essential to point out that the authors performed the selection among and inside the progenies, which may have influenced the higher mean observed (144.266).

Among the 178 F_2 individual used in the selection, 81 displayed predicted genetic values higher than the overall mean, with gains that varied ranged from 17.979 to 3.626. Thus, these data not only reinforce the potential of individuals in the population, but also contribute to indicate the possibility of selecting plants with superior performance for this trait.

The 30 best genotypes from the F_2 population were selected, generating the $F_{2:3}$ generation, applying a selection intensity of 16.9% on the original population.

Table 5 presents the mean of the genetic value of each selected individual and its new mean prediction. The table also highlights the general advances derived from the selective process performed in the present research, including the genetic gains to the original population derived from the best individuals' selection.

As applying the quoted selection pressure, the 30 new individuals displayed an TNPP trait variation of 106.766 (genotype 121) to 95.551 (genotype 58) in their means (Table 5). Selection allowed a 7.62% gain as the number of pods per plant in the further generation compared to the mean of the original population, highlighting the outstanding results in the character improvement. Future generations will probably display better gains by selecting among and within the $F_{2:3}$ progenies, higher environmental controls, and new pressure selections. The selected genotypes will be advanced by the genealogical method, continuing the snap bean breeding program at the UENF, with a mid-term aim to obtain a new superior cultivar.

CONCLUSIONS

All the traits presented degrees of phenotypic variation, showing promising results. Moreover, in most cases, the population mean values were higher than known performance lines.

The tested methods highlighted redundancy. Therefore, the PW variable can be discarded from the future analysis causing no information losses.

The Tocher optimization method and UPGMA showed the studied population genetic variability by forming divergent groups even though there was no complete agreement as considering the estimates of these methods.

The REML/BLUP method selected superior individuals, and the selection of the best 30 genotypes as the total number of pods per plant, providing 7.62% gains for this trait.

Finally, it can be pointed out that the selected individuals can be advanced-generation through the genealogical breeding strategy, continuing the snap bean breeding program at the UENF, with a mid-term aim to obtain a new cultivar for the North-and North-eastern Rio de Janeiro State regions.

ACKNOWLEDGMENTS

CONFLICTS OF INTEREST

The authors declare no conflicts of interest in the development and execution of this research.

REFERENCES

Abreu FB, Leal NR, Rodrigues R, Amaral Júnior AT, et al. (2004). Divergência genética entre acessos de feijão-de-vagem de hábito de crescimento indeterminado. *Hortic Bras*. 22: 547-552.

Almeida SNC, Thiebaut JTL, Gravina GA, Araújo LC, et al. (2014). Avaliação de características morfológicas e agronômicas de linhagens de feijão-de-vagem em Bom Jesus do Itabapoana-RJ, com potencial de recomendação. *Vértices*. 16: 39-50.

Alves FM and Resende MDV (2008). Avaliação genética de indivíduos e progênies de cupuaçuzeiro no estado do Pará e estimativas de parâmetros genéticos. *Rev Bras Frutic*. 30: 696-701.

Amaral Júnior AT, Casali VWD, Cruz CD and Finger FL (1996). Utilização de variáveis canônicas e de análise de agrupamentos na avaliação da divergência genética entre acesso de moranga. *Hortic Bras*. 2: 182–184.

Andrade FA, Gonçalves LSA, Fukujy A, Miglioranza É, et al. (2017) Snap beans for organic farming and evaluation of resistance to the common bacterial blight. *Hortic Bras*. 35: 385–389.

Baldoni AB, Teixeira FF and Santos JB (2002). Controle genético de alguns caracteres relacionados à cor da semente de feijão no cruzamento Rosinha X Esal 693. *Acta Sci Agron*. 24: 1427–1431.

Barbé TC, Amaral Júnior AT, Gonçalves LSA, Rodrigues R, et al. (2010). Association between advanced generation and genealogy inbred lines of snap bean by the Ward-Modified Location Model. *Euphytica*. 173: 337-343.

Correa AM, Gonçalves MC, Destro D, Souza LCF, et al. (2003) Estimates of genetic parameters in common bean genotypes. *Crop Breed Appl Technol*. 3: 223-230.

Costa MM, Mauro AO, Unêda-Trevisoli SH, Arriel NHC, et al. (2004). Ganho genético por diferentes critérios de seleção em populações segregantes de soja. *Pesqui Agropecu Bras*. 39: 1095-1102.

Cruz CD, Regazzi AJ and Carneiro PCS (2012). Modelos biométricos aplicados ao melhoramento genético. 4ª ed. Editora UFV, Viçosa.

Cruz DP, Gravina GA, Oliveira TRA, Silva AB, et al. (2018). Selection of progenies of snap beans using mixed models (REML/BLUP). *Genet Mol Res*. 17: gmr16039914.

Dancey C and Reidy J (2018). Estatística Sem Matemática para Psicologia. 7.ed. Penso Editora, Porto Alegre.

Falconer DS and Mackay TFC (1996). Introduction to quantitative genetics. 4st edition. Longman Group, London.

FAO (Food and Agriculture Organization) (2019). Faostat.

Filgueira FAR (2008). Novo manual de olericultura: agrotecnologia moderna na produção e comercialização de hortaliças. 3ª ed. Editora UFV, Viçosa.

Francelino FMA, Gravina GDA, Manhães CMC, Cardoso PMR, et al. (2011). Avaliação de linhagens de feijão-de-vagem para as regiões Norte e Noroeste Fluminense. *Cienc Agron*. 42: 554-562.

Fukuji AYS, Constantino LV, Zeffa DM, Andrade FAD, et al. (2019). Amino acid concentration, total phenolic compound content and antioxidant activity of snap bean genotypes. *Braz J Food Technol*. 22: e2018069.

Gomes ABS, Oliveira TRA, Cruz DP, Gravina GA, et al. (2018). Genetic gain via REML/BLUP and selection indices in snap bean. *Hortic Bras*. 36: 195-198.

Henderson CR (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics*. 31: 432-449.

Jolliffe IT (1973). Discarding variables in a principal component analysis. II: Real data. *J R Stat Soc Ser C Appl Stat*. 22: 21-31.

Lobô RNB and Lobô ANBO (2007). Melhoramento genético como ferramenta para o crescimento e o desenvolvimento da ovinocultura de corte. *Rev Bras Reprod Anim*. 31: 247-253.

Mardia KL, Kent JT and Bibby JM (1979). Multivariate analysis. 1st edition. Academic Press, London.

Marinho CD, Gravina GA, Araújo LC, Almeida SN, et al. (2013). Snap bean recommendation based on diferente methods of phenotypic stability. *Genet Mol Res*. 12: 248- 255.

Minich DM (2019). A review of the science of colorful, plant-based food and practical strategies for "eating the rainbow". *J Nutr Metab*. 1-19.

Mojena R (1977). Hierarchical grouping methods and stopping rules: an evaluation. *Comput J*. 20: 359-363.

Peixoto N, Braz LT, Banzatto DA, Moraes EA, et al. (2002). Características agronômicas, produtividade, qualidade de vagens e divergência genética em feijão-vagem de crescimento indeterminado. *Hortic Bras*. 20: 447-451.

Pereira HS, Santos JB, Abreu ÂFB and Couto KR (2007). Informações fenotípicas e marcadores microssatélites de QTL na escolha de populações segregantes de feijoeiro. *Pesqui Agropecu Bras*. 42: 707-713.

Resende MDV (2016). Software Selegen-REML/BLUP: a useful tool for plant breeding. *Crop Breed Appl Technol*. 16: 330-339.

Sant'Anna CQSS, Gravina GA, Cruz DP, Oliveira TRA, et al. (2019). Feijão vagem UENF Goytacá: uma nova opção para pequenos produtores. *Hortic Bras*. 37: 239-242.

Sant'Anna CQSS, Oliveira TRA, Gravina GA, Cruz DP, et al. (2020). Distinctness, uniformity and stability and GT biplot tests for the selection of snap bean lines. *Hortic Bras*. 38: 370-377.

Seif YIA, El-Miniawy SEDM, El-Azm NAA and Hegazi AZ (2016). Response of snap bean growth and seed yield to seed size, plant density and foliar application with algae extract. *Ann Agric Sci*. 61: 187-199.

Silva MP, Amaral Júnior AT, Rodrigues R, Daher RF, et al. (2004). Análise dialélica da capacidade combinatória em feijão-de-vagem. *Hortic Bras*. 22: 277-280.

Singh D (1981). The relative importance of characters affecting genetic divergence. *Indian J Genet Plant Breed*. 41: 237-245.

Sokal RR and Rohlf FJ (1962). The comparison of dendrograms by objective methods. *Taxon*. 11: 33-40.

Teixeira AB, Amaral Júnior AT, Rodrigues R, Pereira TNS, et al. (2004) Genetic divergence in snap-bean (*Phaseolus vulgaris* L.) evaluated by diferente methodologies. *Crop Breed Appl Technol*. 4: 57-62.

Viçosi KA, Carvalho ADS, Silva DC, Almeida FP, et al. (2020). Foliar fertilization with boron on the growth, physiology, and yield of snap beans. *J Soil Sci Plant Nutr*. 20: 917-924.

Vilela FO, Amaral Júnior AT, Freitas Júnior SP, Viana AP, et al. (2009). Selection of snap bean recombined lines by using EGT and SSD. *Euphytica*. 165: 21-26.

Xavier TC, Fialho ES, Furtado JB and Oliveira WD (2022). Análise comparativa de índices de conforto térmico em Viçosa-MG, Zona da Mata Mineira. *Geo UERJ*. 41: 54585.

Yuan M, Ruark MD and Bland WL (2017). A simple model for snap bean (*Phaseolus vulgaris* L.) development, growth and yield in response to nitrogen. *Field Crops Res*. 211: 125-136.