

Scientific Journal. ISSN 2595-9433 Volume 3, Number 2, Article n. 8, July/December D.O.I. <u>http://dx.doi.org/10.35418/2526-4117/v3n2a8</u> Received: 03/15/2021 - Accepted: 05/28/2021



SELECTION OF SOYBEAN F3 SEGREGANTING FAMILIES BY MULTIVARIATE MODELS

Ivan Ricardo Carvalho^{1*}, Danieli Jacoboski Hutra¹, Robison Davi Patias Furlan¹,
 Victor Delino Barasuol Scarton¹, Renan Jardel Rusch Treter¹, Jaqueline Piesanti Sangiovo¹, Christian Szambelam Zimmermann¹, Francine Lautenchleger²,
 Murilo Vieira Loro, D Leonardo Cesar Pradebon

¹ Universidade Regional do Noroeste do Estado do Rio Grande do Sul, Ijuí, RS, Brazil;

2 Universidade do Centro-Oeste, Guarapuava, PR, Brazil;

3 Universidade Federal de Santa Maria, RS, Brazil;

*Corresponding author: Ivan Ricardo Carvalho (carvalho.irc@gmail.com).

Abstract: Breeding programs of soybean preconize the productivity as well as tolerance to biotic and abiotic stresses. In view of the lack of information regarding the multivariate selection of soybean segregating families, the objective of this study was to select superior soybean families and to define segregation patterns through multivariate models. The experiment was conducted in 2017/2018 the experimental design used was augmented blocks with 290 F₃ segregating families, and seven cultivars homozygotes arranged in four repetitions. Was employed the frequency distribution, Euclidean algorithm, Tocher optimized grouping method, relative contribution methodology of Singh's and Artificial Neural Networks. Soybean segregating families express high genetic variability within and between the segregating populations of origin. The heterogeneities in the soybean segregation profile are derived from the genetic complementarity of the contrasting parents used. The multivariate models allow to define patterns for the selection of transgressive families that tend to the agronomic ideotype of the commercial cultivars.

Keywords: *Glycine max* L., genetic breeding, biometric strategies, computer learing, food security, sustainable development.

Introduction

Brazil is the world's largest producer of soybeans (*Glycine max* L.), responsible for 33% of all world production of this crop, surpassing the United States of America which contributes 32% of the amount of production (USDA, 2018). Among the Brazilian states, Mato Grosso, Mato Grosso do Sul, Paraná, Rio Grande do Sul and Goiás represent 81.5% of the national production (Brasil, 2010). The area sown in the 2017/2018 harvest was 35 million hectares with a production of 119 million tons of grains, with the Rio Grande do Sul having an area planted around 6,000 hectares (CONAB, 2018).

Breeding programs of soybean commonly seek to increase morphological, physiological, and agronomic aspects so that these together enhance productivity as well as tolerance to biotic and abiotic stresses. However, selecting superior and promising genotypes becomes a difficult and costly task, since many important traits have low heritability, being controlled by many genes, and suffering large effects from the growing environment, these compiled actions result in difficulties in selection success (Cruz, 2006).

Understanding the characteristics of genotypes, as well as their behavior in face of phenotype changes due to the effects of the environment may make it difficult to define patterns and meanings for genotype selection. Therefore, the researcher should direct their efforts in the selection of superior transgressive families (Filho et al., 2009; Bezerra Neto et al., 2010). Thus, the use of multivariate models obtained by suitable algorithms can compile the trends of all the measured characters and enable the stratification of which are the similar, superior or inferior genotypes.

Among the most efficient multivariate analyzes to discriminate efficiently the genotypes obtained in a breeding program, we highlight the optimized Tocher method that is based on the premise of maximum intra-group homogeneity and intergroup heterogeneity (Cruz e Carneiro, 2006; Cruz et al., 2012), the methodology of Singh (1981), whose purpose is to express which characters are determinant to differentiate the genotypes and to assign specific contributions to each character of interest, and Artificial Neural Networks (ANNs) based on the unsupervised computational learning using the Kohonen algorithm (Cruz et al., 2012; Carvalho, 2018). In view of the lack of information regarding the multivariate selection of soybean segregating families, the objective of this study was to select superior soybean families and to define segregation patterns through multivariate models.

Material and methods

The experiment was conducted in the agricultural crop of 2017/2018 in Campos Borges - RS, located at latitude coordinates 28°52'31 "S and longitude 53°00'55"W, with subtropical humid climate of the Cfa type according to Köppen and soil classified as Dark Red Latosol (Oxisol) (Streck et al., 2008). The F3 segregant families (75% endogamic level

with 25% heterozygosity) were obtained through directed crossings performed in 2014/2015, F1 generation (2015/2016), F2 generation (2016/2017), detailed genealogical information expressed in table 1.

The experimental design used was augmented blocks (Federer, 1956). There were 290 F₃ segregating families, and the cultivars Don Mario 7.0i, Roos Camino RR, BMX Potência RR, NS6700 IPRO, DM5958 RSF IPRO, TMG 7166 RR, Don Mario 5.8i corresponding to the common treatments, and these were arranged in four replicates. The experimental units were composed of a sowing line with five meters in length, spaced by 0.45 meters. The direct and manual sowing was done in the second half of November 2017, using a density of 14 seeds per linear meter, base fertilization of 250 kg ha⁻¹ of N-P-K in the formulation 10-20-20.

Preventive control of weeds, pest insects, and diseases was recommended in order to minimize the biotic effects in the result of the experiment. The characters of agronomic interest measured in 10 random plants were: first pod insertion height (FPI, cm); plant height (PH, cm); number of pods in the main stem (NPMS, units); number of pods in the branches (NPB, units); number of branches (NB, units); number of pods with one seed (NP1, units); number of pods with two seeds (NP2, units); number of pods with three seeds (NP3, units); number of pods with four seeds (NP4, units) and seed mass per plant (SMP, grams).

The data obtained were submitted to the assumptions of the analysis of variance where the normality and homogeneity of the residual variances were verified, after the descriptive analysis was carried out by frequency distribution and definition of the phenotypic classes of interest. After, the Euclidean algorithm was used to make the genetic distances matrix of the studied genotypes using all the measured characters. Through the distance matrix, the Tocher optimized grouping method was applied, with intra-group and inter-group stratification. The relative contribution methodology of Singh's (1981) was used, with the intention of defining which characters are most important. The phenotypic matrix was submitted to the algorithm with unsupervised computational learning where the Artificial Neural Networks were constructed by the Kohonen Map, establishing the centroids of interest and the explanatory synaptic connections (Carvalho, 2018). Statistical analyzes were performed using Genes software (Cruz, 2013) and R (R core Team, 2015).

Table 1. I	Description	of the geneal	logy referi	ring to the	soybean F ₃	segregating	families.
------------	-------------	---------------	-------------	-------------	------------------------	-------------	-----------

Maternal Genitors	Paternal Genitors	F ₂ population of origin	F₃ family		
G1	G ₂	IRC 001	1 to 15		
G ₃	G ₂	IRC 002	1 to 4		
G1	G4	IRC 003	1 to 12		
G ₅	G ₆	IRC 004	1		
G7	Gଃ	IRC 005	1 to 6		
G ₃	G ₆	IRC 006	1 to 3		
G ₃	G_6	IRC 007	1 to 2		
G ₃	G5	IRC 008	1 to 7		
G ₃	Gଃ	IRC 010	8		
G ₁₀	G ₆	IRC 011	1 to 15		
G ₁₁	G3	IRC 012	1 to 18		
G1	G3	IRC 013	1 to 16		
G7	G₅	IRC 016	1 to 3		
G ₁₂	G4	IRC 017	1 to 10		
G ₂	G ₁₃	IRC 019	1 to 6		
G ₁₆	G ₁₇	IRC 021	1 to 4		
G ₁₇	G ₁₈	IRC 022	1 to 6		
G ₂₀	G ₂₀	IRC 024	1 to 8		
G ₂₁	G ₂₁	IRC 025	1 to 6		
G ₂₂	G22	IRC 026	1 to 5		
G ₆	G ₆	IRC 027	1		
G ₂₃	G ₂₃	IRC 028	1 to 8		
G ₂₄	G ₂₄	IRC 029	1 to 6		
G ₂₅	G_6	IRC 030	1 to 14		
G ₂₆	G ₂₁	IRC 031	1 to 4		
G1	G 27	IRC 032	1 to 22		
G ₆	G27	IRC 033	1		
G ₂₆	G ₁₇	IRC 034	1 to 6		
G ₂₅	G_6	IRC 035	1 to 19		
G ₃	G ₂₈	IRC 036	1 to 14		
G ₂₇	G ₆	IRC 037	1 to 11		
G ₁₇	G22	IRC 038	1 to 15		
G ₂₈	G ₂₇	IRC 039	1 to 4		
G ₂₉	G ₂₈	IRC 040	1 to 12		

Results and discussion

The genotypes evaluated for the first pod insertion height (FPI) were distinguished in ten phenotype classes with a range of 4.5 to 31.5 cm (Figure 1a), among these 24.03% of the genotypes present in the prominent class of 10.5 cm. The cultivars comprise classes 10.5 and 19.5, and 65.36% of the segregating genotypes are within agronomic standards. The plant height (PH) shows seven phenotypic classes varying from 30 to 180 cm (Figure 1b), and the classes of 60 and 75 cm are responsible for 57.60% of the segregating genotypes. The number of pods in the main stem (NPMS) revealed twelve phenotype classes (Figure 1c) with variation from 0.4 to 9.2 pods, the class with 2.8 pods being more prominent with 21.91% of the segregating families. The number of pods in the branches (NPB) expressed eleven phenotypic classes with variability (Figure 1d) between 10 and 60 pods. The number of branches (NB) showed eleven phenotype classes (Figure 1e), comprising 5 to 10 branches per plant, where 38.87% of the genotypes show 1 to 3 branches per plant, respectively.



Figure 1. Class frequency distribution for: (a) first pod insertion height (IFV); (b) plant height (PH); (c) number of pods in the main stem (NVMS); (d) number of pods in the branches (NVB); (e) number of branches (NB).

The number of pods with one seed (NP1) presented ten phenotype classes (Figure 2a) with amplitude of 1.25 to 23.75 pods, where 25.44% of the studied genotypes are included in the class with 6.5 pods. The number of pods with two seeds (NP2) revealed ten phenotype classes (Figure 2b) comprising 8 to 80 pods per plant, the class with 24 pods being the most prominent with 25.80% of the segregating genotypes. For the number of pods with three seeds (NP3) twelve phenotypic classes were formed (Figure 2c) with

variation from 3 to 69 pods, where 22.92% of the genotypes were included in the class with 21 pods. In relation to the number of pods with four seeds (NP4), two phenotypic classes were formed (Figure 2d), establishing that 86.57% of the genotypes do not express pods with four seeds. Seed mass per plant (SMP) revealed the formation of nine phenotype classes, ranging from 3 to 63 grams, with a predominance of 32.51% of the genotypes included in the class of 15 grams of seeds per plant.



Figure 2. Class frequency distribution for: (a) number of pods with one seed (NV1); (b) number of pods with two seeds (NV2); (c) number of pods with three seeds (NV3); (d) number of pods with four seeds (NV4); (e) seed mass per plant (SMP).

The relative contribution of Singh (1981) showed that the number of pods in the branches (NPB) and the first pod insertion height (FPI) were responsible for 29.99% of the total variation of the experiment, these being the most polymorphic characters. In contrast, the number of pods with four seeds (NP4) and plant height (PH) expressed variability of less than 5% (Figure 3). The main purpose of the analysis of the relative importance of the characters is to identify the most important variables for the selection and which are more polymorphic (Cruz et al., 2014).

The intra-groups Tocher's grouping method revealed for the 290 genotypes studied, that these can be classified into nine intrinsic groups to their similar phenotypic characteristics (Table 2).

Group I collected the largest fraction of genotypes, in these conditions for the correct interpretation of the genealogy, it is determined that the first terms are referring to the maternal and paternal parents with the subscript term



Figure 3. Relative contribution of characters by Singh (1981). First pod insertion height (IPL); plant height (AP); number of pods in the main stem (NLHP); number of pods in the branches (NLR); number of branches (NR); number of pods with one seed (NL1); number of pods with two seeds (NL2); number of pods with three seeds (NL3); number of pods with four seeds (NL4) and seed mass per plant (MSP).

responsible for the definition of the segregating population F_2 , and the last term related to the identification of the segregating family F_3 . In this way, the genotypes belonging to this group and homogeneous among them were,

G1G4P3F1; G1G3P13F11; G25G6P35F9; G10G6P11F11; G1G4P3F11; G10G6P11F5; G17G22P38F8; G1G4P3F8; G1G27P32F2; G17G18P22F1; G17G22P38F5; G1G3P13F13; G17G19P23F2; G17G22P38F9; G17G22P38F11; G1G27P32F9; G10G6P11F4, G17G22P38F12; G1G27P32F10; G25G6P35F8; G1G27P32F11; G3G5P8F4; G29G28P40F2; G1G27P32F16; G27G6P37F1; G10G6P11F12; G22G22P26F2; G10G6P11F3; G16G17P21F4; G25G6P35F13; G25G6P35F6; G29G28P40F8; G25G6P30F7; G3G28P36F5; G25G6P35F15; G1G27P32F3; G28G27P39F4; $_{G1G27_{P32F}15}; \ _{G3G28_{P36F}13}; \ _{G26G17_{P34F}2}; \ _{G1G27_{P32F}12}; \ _{G27_{G6}7_{P37F}10}; \ _{G25_{G6}7_{P35F}17}; \ _{G1G27_{P32F}6};$ G12G4P17F4; G25G6P30F2; G29G28P40F7; G25G6P30F1; G1G27P32F8; G25G6P35F2; G1G2P1F13; G27G6P37F4; G1G4P3F7; G10G6P11F8; G1G4P3F4; G3G28P36F14; G10G6P11F9; G10G6P11F13; G25G6P35F18; G25G6P35F14; G3G28p36F11; G1G27p32F7; G1G2p1F10; G27G6p37F3; G25G6p35F4; G3G28p36F7; G25G6p35F7; G25G6p35F19; G25G6P30F3; G3G5P8F5; G25G6P30F11; G1G2P1F2; G3G28P36F9; G3G28P36F6; G25G6P30F12; G1G3P13F10; G1G3P13F15; G17G22P38F6; G10G6P11F6; G1G2P1F14; G1G24P17F7; G1G2P1F8; G3G5P8F2; G21G21P25F6; G1G3P13F4; G1G4P3F12; G1G2P1F12; G25G6P30F10; G29G28P40F1; G1G4P3F2; G1G27P32F4; G17G22P38F13; $G_{G_{2}}^{1}G_{G_{2}}^{2}G_{G_{2}}^{2}G_{F_{3}}^{2}F_{5}^{2}; G_{G_{2}}^{1}G_{2}^{2}G_{2}^{2}G_{2}^{2}F_{1}^{2}; G_{G_{2}}^{2}G_{2}^{2}G_{2}^{2}G_{2}^{2}F_{1}^{2}; G_{G_{2}}^{2}G_{2}^{2}G_{2}^{2}G_{2}^{2}F_{1}^{2}; G_{G_{2}}^{2}G_{2}^{2}G_{2}^{2}G_{2}^{2}F_{1}^{2}; G_{G_{2}}^{2}G_{2}^{2}G_{2}^{2}G_{2}^{2}F_{1}^{2}; G_{G_{2}}^{2}G_{2}^{2}G_{2}^{2}G_{2}^{2}F_{1}^{2}; G_{G_{2}}^{2}G_{2}^{2}G_{2}^{2}G_{2}^{2}F_{1}^{2}; G_{G_{2}}^{2}G_{2}^{2}G_{2}^{2}F_{2}^{2}; G_{G_{2}}^{2}G_{2}^{2}G_{2}^{2}F_{2}^{2}; G_{G_{2}}^{2}G_{2}^{2}G_{2}^{2}F_{2}^{2}; G_{G_{2}}^{2}G_{2}^{2}G_{2}^{2}F_{2}^{2}; G_{G_{2}}^{2}G_{2}^{2}F_{2}^{2}; G_{G_{2}}^{2}; G_{G_{2}}^{2}G_{2}^{2}F_{2}^{2}; G_{G_{2}}^{2}; G_{G_{2}}^{2}G_{2}^{2}F_{2}^{2}; G_{G_{2}}^{2}; G_{G_{2}}^{2}$ G29G28P40F3; G3G28P36F4; G25G6P30F5; G1G2P1F15; G3G2P2F1; G3G2P2F2; G25G6P30F4; G7G8P5F5; G17G22P38F15; G22G22P26F3; G3G2P2F4; G25G6P30F14; G29G28P40F4; G1G2P1F11; BMX Potência RR; G25G6P35F11; G3G8P10F8; G21G21P25F5; G11G3P12F18; G29G28P40F6; G22G22P26F5; G26G17P34F3; $_{G25_{G}6_{P35F}16};$ $_{G10_{G}6_{P11F}1};$ $_{G25_{G}6_{P35F}3};$ $_{G17_{G}19_{P23F}3};$ $_{G1_{G}27_{P32F}17};$ $_{G3_{G}28_{P36F}8};$ $_{G25_{G}6_{P35F}10};$ $G_{1G27P32F5}; G_{12G4P17F6}; G_{3G5P8F7}; G_{16G17P21F2}; G_{25G6P30F13}; G_{5G6P4F1}; G_{1G4P3F10}; G_{29G28P40F5}; G_{16G17P21F2}; G_{25G6P30F13}; G_{25G6P4F1}; G_{2$ G11G3P12F10; G3G6P7F1; G3G28P36F12; G7G8P5F6; G3G5P8F3; G10G6P11F14; G25G6P30F9; G3G28P36F1; G17G22p38F14; G1G3p13F1; G1G3p13F9; G11G3p12F15; G1G4p3F9; G1G4p3F6; G23G23p28F8; G28G27p39F1; G1G3P13F6; G1G2P1F1; G26G17P34F6; G27G6P37F11; G29G28P40F10; G11G3P12F1; G21G3P19F5; G7G8P5F4; $_{G1G27_{P32F}13};$ G11G3P12F11; G24G24P29F1; G1G27P32F19; G1G4P3F5; G17G22P38F2; G1G27P32F14; G25G6P30F6;

G1G3P13F16; G7G8P5F1; G17G22P38F1; G17G22P38F10; G12G4P17F5; G20G20P24F2; G12G4P17F2; G1G3P13F2; G17G19P23F4; G7G5P16F3; G12G4P17F1; G7G5P16F2; G12G4P17F10; G11G3P12F8; G24G24P29F3; G1G27P32F21; Don Mario 7.0i; G1G2P1F6; G29G28P40F9; G11G3P12F9; G1G3P13F7; G3G6P6F3; G7G5P16F1; G26G21P31F3; G1G3P13F8; G7G8P5F2; G2G13P19F4; G1G27P32F20; G1G2P1F7; G6G6P27F1; G11G3P12F12; G27G6P37F9; G11G3P12F14; G7G8P5F3; G23G23P28F4; G26G17P34F4; G3G2P2F3; G16G17P21F1; G20G20P24F3; G10G6P11F15; G29G28P40F11; G10G6P11F10; G12G4P17F8; G3G5P8F1; G12G4P17F9; G11G3P12F2; G23G23P28F5; G1G3P13F5; G11G3P12F7; G2G13P19F3; G21G21P25F4; G24G24P29F5; G20G20P24F4; G1G27P32F18; G2G13P19F2; G11G3P12F6; G11G3P12F16; G11G3P12F5; G11G3P12F3; G17G22P38F7; G23G23P28F7; G12G21P25F4; G23G23P28F1; G23G23P28F1; G23G23P28F1; G12G4P17F9; G11G3P12F12; G23G23P28F6; G11G3P12F4; G23G23P28F1; G21G21P25F3; G21G21P25F3; G26G21P31F2; G26G21P31F1; G23G23P28F6; G11G3P12F13; G20G20P24F1; G20G20P24F1; G23G23P28F6; G11G3P12F1; G23G20P24F1; G21G21P25F2; G3G6P7F2; G12G21P31F1; G23G23P28F6; G11G3P12F13; G20G20P24F1; G20G20P24F1; G23G23P28F6; G11G3P12F4; G23G20P24F1; G21G21P25F2; G3G6P7F2; G12G21P31F2; G26G21P31F1; G23G23P28F6; G11G3P12F13; G20G20P24F1; G21G21P25F2; G3G6P7F2; G12G21P32F22; G27G6P37F2; G2G13P19F1.

Table 2. Tocher's optimization method for the 290 soybean genotypes.

Groups Genotypes

I	<pre>e1e4psf1; e1e3prsf11; G25G6P3sF9; G10c6pruF11; e1e4psf11; e10c6pruF5; e17c22pssF8; e1e4psF8; e1e27pszF2; e17c18pszF1; e17c22pssF1; e1c27pszF10; e25c6pssF8; e1e27pszF1; e3c5psF4; e29c28papF2; e10c6pruF4, e17c22PssF12; e1272pszF1; e122pssF2; e17c22pssF4; e29c28papF2; e12c6pruF4, e17c22pssF1; e10c6pruF4; e127eszF1; e25c6pssF13; e25c6pssF6; e29c28papF8; e25c6pssF1; e122pszF2; e17c26pszF13; e25c6pssF13; e25c6pssF12; e3c28pssF5; e25c6pszF13; e25c6pssF12; e25c2pssF4; e1c3pruF12; e3c28pssF2; e127pszF1; e1c2pruF1; e12cPruF3; e15c2pruF4; e25c6pssF12; e25c6pszF13; e27c6pszF14; e3c28pssF1; e1c27pszF1; e1c2pruF1; e1c2PruF2; e25c6pszF12; e25c6pszF13; e27c6pszF4; e1c3cpszF1; e10c6pruF3; e16c4pszF4; e3c28pssF5; e25c6pszF2; e12cPr13; e27c6pszF4; e1c4pszF; e10c6pruF3; e16c4pszF4; e3c28pszF14; e10c6pruF9; e10c6pruF3; e12cPr17; e12cPr13; e25c6pszF4; e1c3cpszF11; e1c27pszF7; e12cPrF10; e27c6pszF1; e12cPrF2; e3c28pszF3; e25c28pszF1; e25c6pszF4; e3c28pszF11; e1c27pszF7; e12cPrF10; e27c6pszF3; e25c6pszF4; e3c28pszF6; e25c6pszF14; e3c28pszF12; e1c3pruF15; e17c22pszF6; e10c6pruF6; e1c2pruF14; e1c24pruF7; e1c2prE8; e3c5pruF2; e21c21pszF6; e13c3pruF4; e27c6pszF3; e25c6psuF1; e1c2pruF2; e3c28psuF3; e22c22psuF1; e1c4pruF2; e1c2PruF3; e12c4pruF3; e29c28psuF3; e32c6psuF1; e1c2pruF2; e3c26psuF1; e22c28psuF1; e1c4psuF3; e1c2PruF3; e12c4pruF3; e29c28psuF3; e32c6psuF1; e1c2pruF2; e3c26psuF3; e12cpr15; e3c2pzF1; e3c2pzF2; e25c6psuF1; e1c3pruF4; e12c4psuF3; e32c2PzeF3; e32c6psuF1; e12c2pruF3; e32c8psuF4; e25c6psuF1; e32c28psuF1; e3c28psuF6; e32c2PzeF3; e32c6psuF1; e1c32pruF1; e32c8psuF4; e32c6psuF1; e32c28psuF1; e3c28psuF6; e32c2pzeF3; e32c6psuF1; e12c2psF3; e12c4pruF3; e32c8psuF3; e32c6psuF1; e32c22pzeF1; e12c5c6psuF1; e32c8psuF3; e32c6psuF1; e12c2psF3; e12c4pruF3; e32c8psF3; e16c17pzuF2; e32c32pseF3; e32c6psF3; e11c3pruF1; e23c6psF3; e11c3pruF1; e32c8psF1; e12c7psF2; e12c4pzF3; e32c6psF3;</pre>
II	TMG 7166 RR; Don Mario 5.8i; g20g20p24F8; g21g21p25F1; g23g23p28F2; g20g20p24F5; g24g24p29F4; g2g13p1F96; g23g23p28F3; g24g24p29F2; g29g28p40F12
III	NS600 IPRO e DM5958 RSF IPRO
IV	G26 _G 17 _{P34F} 1, _G 27 _G 6 _{P37F} 8 e _G 3 _G 6 _{P6F} 1
V	_G 6 _G 27 _{P33F} 1eRoosCamino RR
VI	G10G6P11F7 eG17G22P38F4
VII	G25G6P35F5
VIII	G20G20P24F6
IX	G3G6P6F2

These genotypes show that the maternal G_1 parent was responsible for gathering 23.97% of the progenies in this group, as well as the maternal G₂, G₂₀, G₂₁ and G₂₂ parents responsible for 1.87% of the genotypes intrinsic to this group. The maternal parent G₃ indicated that 10.86% of the progenies had a similar profile, with a lower segregation attributed to progenies from maternal parents G₅ and G₆ with only 0.37%. The G₇ and G_{26} parents indicate that 3.37% of the progenies are in this group. The group II was responsible for gathering eleven genotypes, these being TMG 7166 RR; Don Mario 5.8i; G20G20P24F8; G23G23P28F2; $G_{G20G20P24F}5;$ G21G21P25F1; G24G24P29F4;G2G13P1F96; G23G23P28F3; $_{G}24_{G}24_{P29F}2$; $_{G}29_{G}28_{P40F}12$, the selection of these families could be favorable due to meeting the agronomic pattern of commercial cultivars. The group III defined the presence of the two commercial cultivars NS600 IPRO and DM5958 RSF IPRO, for group IV it is evident that the $G26_{G}17_{P34F}1$, $_{G}27_{G}6_{P37F}8$, genotypes and $_{G}3_{G}6_{P6F}1$ express peculiarities and segregate with the same meaning for the characters of interest. Group V reunited the segregating family $_{\rm G}6_{\rm G}27_{\rm P33F}1$ and the commercial cultivar Roos

Camino RR. Group VI defined that from families ${}_{G10_{G}6_{P11F}7}$ and ${}_{G17_{G}22_{P38F}4}$ present themselves similar even though they come from contrasting parents. Groups VII, VIII and IX express the classification of specific families ${}_{G25_{G}6_{P35F}5}$, ${}_{G20G20P24F6} e_{G3_{G}6_{P6F}2}$, respectively.

In this method, the similar genotypes are grouped in the first hierarchical patterns (Cruz et al., 2014). This analysis made it possible to reveal which genotypes were phenotypically similar to commercial cultivars, and which of these segregating families could be selected considering all the agronomic attributes simultaneously, but those groups that did not reveal the presence of commercial cultivars express the possibility of selections in the next generations due to high variability available. Considering the genetic distinctions between the groups (Table 3), it was observed that the maximum inter group heterogeneity was evidenced between groups V and VI (0.50) and VI and IX (0.46). Therefore, these arch for the maximum genetic variability of the breeding program in soybean can be obtained using the contrast of these groups.

Groups	I	П	Ш	IV	V	VI	VII	VIII	IX
I	0.20	0.34	0.30	0.31	0.36	0.30	0.32	0.32	0.41
II		0.24	0.36	0.31	0.29	0.47	0.44	0.30	0.42
III			0.14	0.38	0.33	0.37	0.42	0.36	0.41
IV				0.23	0.33	0.46	0.34	0.33	0.37
V					0.19	0.50	0.44	0.44	0.43
VI						0.24	0.37	0.44	0.46
VII							0.0	0.44	0.41
VIII								0.0	0.41
IX									0.00

 Table 3. Tocher's optimization method intergroups of 290 soybean genotypes.

The use of Artificial Neural Networks (ANNs) with unsupervised computational learning defined that the variability available

requires 20 centroids to explain and classify the soybean segregating patterns, the classifications are expressed as follows,

 $\begin{array}{l} g1g2p_{1F}15, \ g3g2p_{2F}1, \ g3g2p_{2F}4, \ g5g6p_{4F}1, \ g7g8p_{5F}5, \ g3g6p_{6F}3, \ g3g8p_{10F}8, \ g1g3p_{13F}10, \ g16g17p_{21F}1, \\ g6g6p_{27F}1, \ g25g6p_{30F}9, \ g25g6p_{30F}10, \ g25g6p_{30F}11, \ g25g6p_{35F}11, \ g3g28p_{36F}6, \ g3g28p_{36F}8 \ (Centroid I); \\ g7g8p_{5F}4, \ g7g8p_{5F}6, \ g3g6p_{7F}1, \ g1g3p_{13F}16, \ g20g20p_{24F}2, \ g20g20p_{24F}3, \ g24g24p_{29F}6, \ g26g21p_{31F}4, \\ g28g27p_{39F}2 \ (Centroid II); \\ g3g2p_{2F}3, \ g7g8p_{5F}1, \ g7g8p_{5F}2, \ g7g8p_{5F}3, \ g3g6p_{6F}1, \ g20g20p_{24F}4, \\ g20g20p_{24F}4, \ g20g20p_{24F}6, \ g20g20p_{24F}7, \ g21g21p_{25F}3, \ g21g21p_{25F}4, \ g24g24p_{29F}3, \ g24g24p_{29F}4, \\ g24g24p_{29F}5, \ g26g21p_{31F}2, \ g26g21p_{31F}3, \ g1g27p_{32F}21 \ (Centroid III); \\ g3g6p_{6F}2, \ g20g20p_{24F}5, \\ \end{array}$

G20G20P24F8, G21G21P25F1, G21G21P25F2, G26G21P31F1, G26G17P34F1, Don Mario 5.8i (Centroid IV);_G3_G2_{P2F}2, G10_G6_{P11F}1, G11_G3_{P13F}1, G1_G3_{P13F}9, G7_G5_{P16F}3, G16_G17_{P21F}2, G25_G26_{P30F}8, G25_G6_{P30F}12, G25G6P30F13, G1G27P32F17, G25G6P35F17, G3G28P36F4, G29G28P40F1 (Centroid V);G1G2P1F1, G1G2P1F11, G3G5P8F3, G3G5P8F6, G3G5P8F7, G11G3P12F17, G11G3P12F18, G1G3P13F3, G1G3P13F4, G1G3P13F5, G1G3P13F6, G7G5P16F1, G16G17P21F3, G17G19P23F6, G21G21P25F5, G22G22P26F1, G22G22P26F5, G23G23P28F4, G23G23P28F8, G24G24P29F1, G25G6P30F5, G25G6P30F6, G1G27P32F22, G26G17P34F6, G25G6P35F16, G3G28P36F10, G27G6P37F11, G17G22P38F15, G29G28P40F4, G29G28P40F5 (Centroid VI);G3G5P8F1, G1G3P13F7, G1G3P13F8, G2G13P19F3, G23G23P28F1, G23G23P28F5, G1G27P32F13, G26G17P34F5, G3G28P36F12, G27G6P37F9, G17G22P38F7, G28G27P39F1, Don Mario 7.0i (Centroid VII); G11G3P12F3, G11G3P12F4, $_{G11_G3_{P12F}5}$, $_{G11_G3_{P12F}6}$, $_{G11_G3_{P12F}13}$, $_{G11_G3_{P12F}16}$, $_{G2G13_{P19F}1}$, $_{G23_G23_{P28F}2}$, $_{G23_G23_{P28F}3}$, G23G23P28F6, G6G27P33F1, G29G28P40F12, Roos Camino RR, TMG 7166 RR (Centroid VIII);G3G5P8F5, G10G6P11F14, G1G3P13F12, G22G22P26F3, G25G6P30F1, G1G27P32F1, G1G27P32F4, G1G27P32F5, G1G27P32F15, G25G6P35F1, G25G6P35F2, G25G6P35F7, G25G6P35F15, G25G6P35F18, G28G27P39F3, G29G28P40F6 (Centroid IX);G1G2P1F12, G1G2P1F14, G3G6P7F2, G3G5P8F2, G12G4P17F7, G21G21P25F6, G25G6P30F3, G25G6P30F4, G25G6P30F14, G1G27P32F12, G25G6P35F3, G27G6P37F10, G17G22P38F14, G29G28P40F3, BMX Potência RR (Centroid X);_{G1G2P1F8}, G1G4P3F10, G11G3P12F9, G11G3P12F10, G11G3P12F11, G11G3P12F15, G7G5P16F2, $_{G12G4P17F2}$, $_{G12G4P17F10}$, $_{G2G13P19F4}$, $_{G2G13P19F5}$, $_{G17G19P23F4}$, $_{G12G7P32F19}$, $_{G26G17P34F4}$, G27G6P37F6, G27G6P37F8 (Centroid XI);G1G4P3F5, G1G4P3F6, G11G3P12F2, G11G3P12F7, G11G3P12F8, $_{G11_{G}3_{P12F}12}$, $_{G11_{G}3_{P12F}14}$, $_{G12_{G}4_{P17F}1}$, $_{G12_{G}4_{P17F}9}$, $_{G2_{G}13_{P19F}2}$, $_{G2_{G}13_{P19F}6}$, $_{G24_{G}24_{P29F}2}$, G1G27P32F20, G27G6P37F7, NS6700 IPRO, DM5958 RSF IPRO (Centroid XII);G3G5P8F4, G22G22P26F4, g25g6p30F2, g25g6p30F7, g1g27p32F6, g25g6p35F6, g25g6p35F10, g3g28p36F1, g3g28p36F2, g3g28p36F3, G3G28p36f5. $_{\rm G}3_{\rm G}28_{\rm P36F}7$, $_{\rm G}3_{\rm G}28_{\rm P36F}9$, $_{G28_{G27_{P39F}4}}$, $_{G29_{G28_{P40F}2}}$, $_{G29_{G28_{P40F}7}}$ (Centroid XIII); $G_{1}G_{2}P_{1}F_{9}$, $G_{1}G_{2}P_{1}F_{1}F_{1}$, $G_{1}G_{2}G_{4}P_{1}7F_{6}$, $G_{1}G_{6}G_{1}7_{P_{2}1}F_{4}$, $G_{2}G_{2}G_{2}2_{P_{2}6}F_{2}$, $G_{1}G_{2}7_{P_{3}2}F_{2}$, $G_{1}G_{2}7_{P_{3}2}F_{3}$, $G_{1}G_{1}7_{P_{3}2}F_{3}$, $G_{1}G_{2}7_{P_{3}2}F_{3}$, $G_$ $_{G1G27P32F7}$, $_{G1G27P32F8}$, $_{G1G27P32F16}$, $_{G26G17P34F2}$, $_{G26G17P34F3}$, $_{G25G6P35F12}$, $_{G25G6P35F13}$, $_{G25G6P35}$, $_{G25G6P35}$, G25G6P35F14, G3G28P36F13, G27G6P37F5, G17G22P38F13, G29G28P40F10 (Centroid XIV);G1G2P1F10, G1G4P3F3, G1G4P3F12, G10G6P11F8, G10G6P11F13, G12G4P17F3, G17G18P22F1, G17G19P23F3, G1G27P32F18, G17G22P38F3 (Centroid XV);G1G4P3F2, G1G4P3F7, G11G3P12F1, G1G3P13F2, G1G3P13F14, G1G3P13F15, G1G27P32F14, G17G22P38F6 (Centroid XVI);G1G2P1F2; G1G4P3F8; G10G6P11F4, G1G3P13F11, G1G27P32F9, G1G27P32F10, G1G27P32F11, G25G6P35F8, G25G6P35F9, G3G28P36F11, G3G28P36F14, G27G6P37F1, G17G22P38F1, G17G22P38F9, G17G22P38F11, G29G28P40F8 (Centroid XVII);G1G2P1F5, G1G4P3F1, G1G4P3F4, G10G6P11F3, G10G6P11F11, G10G6P11F12, G1G3P13F13, G17G19P23F2, G25G6P35F19, G27G6P37F4, G17G22P38F8, G29G28P40F11 (Centroid XVIII);G1G2P1F6, G1G2P1F7, G1G4P3F9, G1G4P3F11, G10G6P11F2, G10G6P11F6, G10G6P11F7, G10G6P11F15, G12G4P17F4, G12G4P17F5, G17G19P23F4, G25G6P35F4, G27G6P37F2, G27G6P37F3, G17G22P38F2, G17G22P38F5, G17G22P38F10, G17G22P38F12, G29G28P40F9 (Centroid XIX); G12G4P17F8, G23G23P28F7, G25G6P35F5, G17G22P38F4 (Centroid XX),

thus defining that the 290 genotypes studied are stratified into 20 genetic patterns (Table 4, Figure 4), and the selection can be directed to the centroid level, where when defining the agronomic ideotype of interest, as well as, locate which centroid of reference, so that those genotypes that stand together in this pattern possibly express same sense in these gregation in the next generations.

In this context, Artificial Neural Networks can define and assist the selection strategies

(Figure 5). Therefore, the number of pods in the main stem (NPMS), number of pods with two seeds (NP2) and number of pods in the branches (NPB) determined the formation of Centroid I, the number of pods with three seeds (NP3), and seed mass per plant (SMP) defined centroid II, first pod insertion height (FPI) and number of pods with one seed (NP1) were responsible for the formation of the centroid IV, the number of branches (NB) and the number of pods with four seeds (NP4) define the pattern for centroid VI.

Table 4. Definition of the profiles through Artificial Neural Networks by the Kohonen Map for the 290 soybean genotypes.

Centroids	Neurons
I	G1G2P1F15, G3G2P2F1, G3G2P2F4, G5G6P4F1, G7G8P5F5, G3G6P6F3, G3G8P10F8, G1G3P13F10, G16G17P21F1, G6G6P27F1, G25G6P30F9, G25G6P30F10, G25G6P30F11, G25G6P35F11, G3G28P36F6, G3G28P36F8
II	G7G8p5F4, G7G8p5F6, G3G6p7F1, G1G3p13F16, G20G20p24F2, G20G20p24F3, G24G24p29F6, G26G21p31F4, G28G27p39F2
III	G3G2P2F3, G7G8P5F1, G7G8P5F2, G7G8P5F3, G3G6P6F1, G20G20P24F1, G20G20P24F4, G20G20P24F6, G20G20P24F7, G21G21P25F3, G21G21P25F4, G24G24P29F3, G24G24P29F4, G24G24P29F5, G26G21P31F2, G26G21P31F3, G1G27P32F21
IV	G3G6P6F2, G20G20P24F5, G20G20P24F8, G21G21P25F1, G21G21P25F2, G26G21P31F1, G26G17P34F1, Don Mario 5.8i
V	G3g2p2F2, g10g6p11F1, g11g3p13F1, g1g3p13F9, g7g5p16F3, g16g17p21F2, g25g26p30F8, g25g6p30F12, g25g6p30F13, g1g27p32F17, g25g6p35F17, g3g28p36F4, g29g28p40F1
VI	G1G2P1F1, G1G2P1F11, G3G5P8F3, G3G5P8F6, G3G5P8F7, G11G3P12F17, G11G3P12F18, G1G3P13F3, G1G3P13F4, G1G3P13F5, G1G3P13F6, G7G5P16F1, G16G17P21F3, G17G19P23F6, G21G21P25F5, G22G22P26F1, G22G22P26F5, G23G23P28F4, G23G23P28F8, G24G24P29F1, G25G6P30F5, G25G6P30F6, G1G27P32F22, G26G17P34F6, G25G6P35F16, G3G28P36F10, G27G6P37F11, G17G22P38F15, G29G28P40F4, G29G28P40F5
VII	G3G5P8F1, G1G3P13F7, G1G3P13F8, G2G13P19F3, G23G23P28F1, G23G23P28F5, G1G27P32F13, G26G17P34F5, G3G28P36F12, G27G6P37F9, G17G22P38F7, G28G27P39F1, Don Mario 7.0i
VIII	G11G3P12F3, G11G3P12F4, G11G3P12F5, G11G3P12F6, G11G3P12F13, G11G3P12F16, G2G13P19F1, G23G23P28F2, G23G23P28F3, G23G23P28F6, G6G27P33F1, G29G28P40F12, ROSS Camino RR, TMG 7166 RR
IX	G3G5P8F5, G10G6P11F14, G1G3P13F12, G22G22P26F3, G25G6P30F1, G1G27P32F1, G1G27P32F4, G1G27P32F5, G1G27P32F15, G25G6P35F1, G25G6P35F2, G25G6P35F7, G25G6P35F15, G25G6P35F18, G28G27P39F3, G29G28P40F6
x	G1G2P1F12, G1G2P1F14, G3G6P7F2, G3G5P8F2, G12G4P17F7, G21G21P25F6, G25G6P30F3, G25G6P30F4, G25G6P30F14, G1G27P32F12, G25G6P35F3, G27G6P37F10, G17G22P38F14, G29G28P40F3, BMX Potência RR
XI	g1g2p1f8, g1g4p3f10, g11g3p12f9, g11g3p12f10, g11g3p12f11, g11g3p12f15, g7g5p16f2, g12g4p17f2, g12g4p17f10, g2g13p19f4, g2g13p19f5, g17g19p23f4, g12g7p32f19, g26g17p34f4, g27g6p37f6, g27g6p37f8
XII	G1G4p3f5, G1G4p3f6, G11G3p12f2, G11G3p12f7, G11G3p12f8, G11G3p12f12, G11G3p12f14, G12g4p17f1, G12g4p17f9, G2g13p19f2, G2g13p19f6, G24g24p29f2, G1g27p32f20, G27g6p37f7, NS6700 IPRO, DM5958 RSF IPRO
XIII	G3G5P8F4, G22G22P26F4, G25G6P30F2, G25G6P30F7, G1G27P32F6, G25G6P35F6, G25G6P35F10, G3G28P36F1, G3G28P36F2, G3G28P36F3, G3G28P36F5, G3G28P36F7, G3G28P36F9, G28G27P39F4, G29G28P40F2, G29G28P40F7
XIV	G1G2p1F9, G1G2p1F13, G12G4p17F6, G16G17p21F4, G22G22p26F2, G1G27p32F2, G1G27p32F3, G1G27p32F7, G1G27p32F8, G1G27p32F16, G26G17p34F2, G26G17p34F3, G25G6p35F12, G25G6p35F13, G25G6p35F14, G3G28p36F13, G27G6p37F5, G17G22p38F13, G29G28p40F10
xv	G1G2p1f10, G1G4p3f3, G1G4p3f12, G10G6p11f8, G10G6p11f13, G12G4p17f3, G17G18p22f1, G17G19p23f3, G1G27p32f18, G17G22p38f3
XVI	G1G4p3F2, G1G4p3F7, G11G3p12F1, G1G3p13F2, G1G3p13F14, G1G3p13F15, G1G27p32F14, G17G22p38F6
XVII	G1G2p1F2; G1G4p3F8; G10G6p11F4, G1G3p13F11, G1G27p32F9, G1G27p32F10, G1G27p32F11, G25G6p35F8, G25G6p35F9, G3G28p36F11, G3G28p36F14, G27G6p37F1, G17G22p38F1, G17G22p38F9, G17G22p38F11, G29G28p40F8
XVIII	g1g2p1f5, g1g4p3f1, g1g4p3f4, g10g6p11f3, g10g6p11f11, g10g6p11f12, g1g3p13f13, g17g19p23f2, g25g6p35f19, g27g6p37f4, g17g22p38f8, g29g28p40f11
XIX	g1g2p1f6, g1g2p1f7, g1g4p3f9, g1g4p3f11, g10g6p11f2, g10g6p11f6, g10g6p11f7, g10g6p11f15, g12g4p17f4, g12g4p17f5, g17g19p23f4, g25g6p35f4, g27g6p37f2, g27g6p37f3, g17g22p38f2, g17g22p38f5, g17g22p38f10, g17g22p38f12, g29g28p40f9
XX	g12g4p17F8, g23g23p28F7, g25g6p35F5, g17g22p38F4



Figure 4. Artificial Neural Networks (ANNs) obtained by the *Kohonen* Map defining the centroids (red) as the standard for the 290 genotypes and synaptic connections (dashed lines).



Figure 5. Artificial Neural Networks (ANNs) obtained by the Kohonen Map for the variables: first pod insertion (IFV); plant height (PH); number of pods in the main stem (NVMS); number of pods in the branches (NVB); number of branches (NB); number of pods with one seed (NV1); number of pods with two seeds (NV2); number of pods with three seeds (NV3); number of pods with four seeds (NV4) and seed mass per plant (SMP).

Conclusions

Soybean segregating families express high genetic variability within and between the segregating populations of origin. The heterogeneities in the soybean segregation profile are derived from the genetic complementarity of the contrasting parents used. The multivariate models allow to define patterns for the selection of transgressive families that tend the agronomic ideotype of the commercial cultivars.

References

- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. 2010. Projeções do agronegócio: Brasil 2009/2010 a 2019/2020. Brasília: MAPA.
- CARVALHO, I.R. 2018. **Biometria e modelos mistos aplicados à seleção de genótipos de feijão.** Completion of Course Work (Specialization in Seed Science and Technology), Universidade Federal de Pelotas, Pelotas, 104 p.
- CONAB Companhia Nacional De Abastecimento. 2018. Acompanhamento de safra brasileiro grãos: Décimo segundo levantamento, setembro 2018 safra 2017/2018. 5(12).
- CRUZ, C.D. 2013. GENES a software package for analysis in experimental statistics and quantitative genetics. Acta Scientiarum Agronomy, 35: 271-276.
- CRUZ, C.D.; REGAZZI, A.J.; CARNEIRO, P.C.S. 2012. Modelos biométricos aplicados ao melhoramento genético. UFV, Viçosa, Brasil, 514p.
- CRUZ, C.D.; CARNEIRO, P.C.S.; REGAZZI, A.J. 2014. Modelos biométricos aplicados ao melhoramento de plantas. Editora UFV. Ed.3 Viçosa, Brasil, 668p.
- FEDERER, W.T. 1956. Augmented (hoonuiaku) designs. Hawaiian Planters Record. 55: 191-208.
- FILHO, A.C.; STORCK, L.; RIBEIRO, N.D. 2009. Medidas da precisão experimental em ensaios com genótipos de feijão e de soja. **Pesquisa Agropecuária Brasileira**, Brasília, 44(10): 1225-1231.
- R CORE TEAM. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: <u>http://www.R-project.org</u>. Accessed on: Nov 14th 2018.
- SINGH, D. 1981. The relative importance of characters affecting genetic divergence. Indian Journal of Genetics and Plant Breeding, 41(2): 237-245.
- STRECK, E.V.; KÄMPF, N.; DALMOLIN, R.S.; KLAMT, E.; NASCIMENTO, P.C.; SCHNEIDER, P.; GIASSON, É.; PINTO, L.F.S. 2008. Solos do Rio Grande do Sul. Emater, Ed. 2 Porto Alegre, 222 p.
- USDA United States Department of Agriculture. Available at: <u>http://www.usda.gov</u>. Accessed on: Dec 12th 2018.