PAB
Pesquisa Agropecuária Brasileira

Genetics/ Original Article

Identification of core locations for soybean breeding in Southern Brazil

Abstract – The objective of this work was to identify core locations in two soybean macroregions in Brazil for the evaluation and selection of soybean (*Glycine max*) lineages. Twenty-two cultivars were tested in 23 locations in these two macroregions (MR1 and MR2), during four years, from 2012 to 2015. Trials were conducted in a randomized complete block design with three replicates. All analyses were performed using the GGEbiplot software. The genotype main effects plus genotype × location interaction + genotype main effects plus genotype × environment interaction and genotypic effect vs. genotype x environment effect analyses were used to identify core locations, i.e., locations with high representativeness and consistency of results. Chapada, in the state of Rio Grande do Sul, and Maracaju, in the state of Mato Grosso do Sul, were the core locations in MR1 and MR2, respectively. These locations were the most representative and consistent over the years, and the genotypic effect explained a high proportion of phenotypic variance.

Index terms: *Glycine max*, GGL + GGE, soybean macroregion, test locations.

Identificação de locais de teste para melhoramento de soja no Sul do Brasil

Resumo – O objetivo deste trabalho foi identificar locais-chave de teste em duas macrorregiões sojícolas no Brasil, para avaliação e seleção de linhagens de soja (*Glycine max*). Foram testadas 22 cultivares em 23 locais, nessas duas macrorregiões (MR1 e MR2), durante quatro anos, de 2012 a 2015. Os ensaios foram conduzidos em delineamento de blocos ao acaso, com três repetições. Todas as análises foram realizadas com uso do programa GGEbiplot. As análises efeito principal de genótipo mais interação genótipo × local + efeito principal de genótipo mais interação genótipo × ambiente e efeito genotípico vs. efeito do genótipo x ambiente foram utilizadas para identificar locais de teste, isto é, locais com alta representatividade e consistência de resultados. Chapada, no Rio Grande do Sul, e Maracaju, no Mato Grosso do Sul, foram os principais locais em MR1 e MR2, respectivamente. Estes locais foram os mais representativos e consistentes ao longo dos anos, e o efeito genotípico explicou uma alta proporção da variância fenotípica.

Termos para indexação: *Glycine max*, GGL + GGE, macrorregião sojícola, locais de teste.

Gilvani Matei[1] ,
Andrei Daniel Zdziarski[1] ,
Leomar Guilherme Woyann[1] ,
Rodrigo Zanella[1] ,
Vinícius de Bitencourt Bez Batti[1] and
Giovani Benin[1 ✉]

[1] Universidade Tecnológica Federal do Paraná, Campus Pato Branco, Via do Conhecimento, Km 1, CEP 85503-390 Pato Branco, PR, Brazil.
E-mail: gilvani.matei@syngenta.com, dz_andrei@hotmail.com, leowoyann@gmail.com, zanella_rodrigo@hotmail.com, viniciusbezbatti95@gmail.com, benin@utfpr.edu.br

✉ Corresponding author

Introduction

In plant breeding programs, it is essential to conduct trials at locations that are representative of the target environment. This representativeness must be consistent over the years and present a sufficiently high genetic

correlation with the target environment (Dia et al., 2016; Yan, 2016). Furthermore, locations and selection must be efficient in differentiating superior genotypes (Qin et al., 2015; Krishnamurthy et al., 2017). When a location combines both representativeness and the ability to discriminate genotypes, it is called a core location (Yan, 2014). The identification of a core location is fundamental in generations with low seed availability, such as in segregating populations and recently selected lines, i.e., in progenies from a single plant and in preliminary trials (Wu et al., 2013). Some of the difficulties faced by soybean breeders when choosing the best location for the selection of superior genotypes include both limited seed availability in early generations (Yan, 2014) and resources (Mi et al., 2011); because of the latter, trials are usually conducted in a few locations – usually in only one –, which has direct implications on the effectiveness of the breeding program.

Macroregions for soybean adaptation are widely used to identify core locations for the selection and evaluation of new lineages (Kaster & Farias, 2012). Within these regions, breeders identify representative locations to conduct their trials (Yan et al., 2010). It should be noted that the task of identifying a single location is an important factor (Das et al., 2019), requiring a theoretical and practical foundation to avoid mistakes in selection. Misconceptions in selecting the best lineages in the preliminary testing phase may result in the elimination of highly desirable lineages and in a reduction in the efficiency of the breeding program. Up to date, no known work has been performed to identify the key test locations for soybean breeding in the South of Brazil.

The objective of this work was to identify core locations in two soybean macroregions in Brazil for the evaluation and selection of soybean lineages.

Materials and Methods

Data from soybean [*Glycine max* (L.) Merr.] value for cultivation and use (VCU) trials carried out in the 2012–2013, 2013–2014, 2014–2015, and 2015–2016 crop seasons were used. Trials were conducted in 23 locations, 11 in the first and 12 in the second evaluated soybean macroregion (MR1 and MR2, respectively) (Figure 1 and Table 1). MR1, located in the South of Brazil, is characterized by a higher altitude

and a mesothermal climate without dry seasons, classified as Cfa or Cfb, according to Köppen, with temperatures in the coldest month ranging from -3 to 18ºC (Alvares et al., 2013); this macroregion covers the states of Rio Grande do Sul and Santa Catarina, as well as the mid-southern and southeastern regions of the state of Paraná and the south of the state of São Paulo. In contrast, MR2, which consists of western and northern Paraná, southern and western São Paulo, and southern Mato Grosso do Sul, is considered a transitional region between Cfa and Cwa climates (Alvares et al., 2013), presenting higher predominant temperatures in summer and dry winters (Kaster & Farias, 2012).

A total of 22 genotypes were tested, including 8 cultivars and 14 lineages (Table 2). It should be pointed out that, as locations and genotypes were not the same in all studied years, the dataset is highly imbalanced.

Experiments were conducted in a randomized complete block design with three replicates. Plots consisted of four rows with a length of 5 m, and row and plot spacing were of 0.5 m. Sowing density was
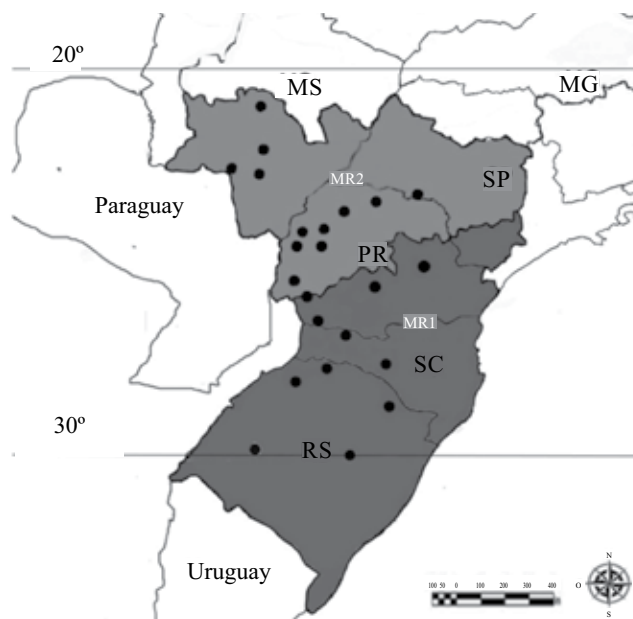

Figure 1. Brazilian map with trial locations (black dots) in the two evaluated macroregions (MR1 in dark gray and MR2 in light gray) of soybean (*Glycine max*) adaptation in Southern Brazil. States: MS, Mato Grosso do Sul; MG, Minas Gerais; SP, São Paulo; PR, Paraná; SC, Santa Catarina; and RS, Rio Grande do Sul.

standardized for all genotypes at 30 seeds per square meter. Basic fertilization consisted of a mineral fertilizer with 7 kg ha^{-1} N, 70 kg ha^{-1} P$_2$O$_5$, and 70 kg ha^{-1} K$_2$O. Harvest was performed using a plot harvester when plants reached harvest maturity at the R8 phenological stage. Both central rows of each plot (5 m^2) were harvested, and seed moisture content was routinely adjusted to 13% to obtain grain yield (kg ha^{-1}).

The statistical analysis for the identification of core locations was performed using the GGEbiplot software (Yan, 2001). The genotype main effects plus genotype × location interaction (GGL) + genotype main effects plus genotype × environment interaction effect (GGE) analysis was used to identify core locations in each of the two previously defined macroregions. In this analysis,

the location on the biplot was defined by the mean of both PC1 and PC2 over the tested years. The data were scaled using a heritability-adjusted environment interaction effect (HA-GGE) biplot, in which data are scaled by standard deviation and adjusted heritability (Yan & Holland, 2010). The cosine of the angle between the vector of the location and the average environment axis (AEA) line corresponds to the genetic correlation (r$_g$) between them (Yan & Holland, 2010; Yan, 2014). Therefore, the smaller the angle between location and average environment, the more representative is the test location. Vector length indicates the consistency of the results over the years, i.e., the representativeness of the location; moreover, when the biplot explanation is

Table 1. Locations (municipalities) of the value for cultivation and use trials in the two evaluated macroregions (MR1 and MR2) of soybean (*Glycine max*) adaptation in Brazil.

Location	Code	Region	2012	2013	2014	2015
Abelardo Luz	ABL	MR1	X	X	X	X
Brasilândia do Sul	BRA	MR2			X	X
Cascavel	CCV	MR2	X	X	X	X
Chapada	CHA	MR1	X	X	X	X
Cândido Mota	CMO	MR2	X	X	X	
Campos Novos	CPN	MR1	X		X	X
Dourados	DOU	MR2	X	X	X	X
Erechim	ERC	MR1	X	X		
Guarapuava	GUA	MR1		X	X	X
Londrina	LON	MR2	X	X	X	X
Maracaju	MCJ	MR2	X	X		X
Palotina	PLT	MR2	X	X	X	X
Ponta Porã	PPO	MR2	X		X	X
Perobal	PRB	MR2	X		X	X
Palma Sola	PSO	MR1	X	X	X	
Realeza	RLZ	MR1	X	X	X	X
São Francisco de Assis	SFA	MR1			X	X
Sidrolândia	SID	MR2	X		X	X
São Jorge do Ivaí	SJI	MR2			X	X
Santa Cruz do Sul	STC	MR1	X	X	X	X
Santo Augusto	STO	MR1	X	X		X
Ubiratã	UBI	MR2		X	X	X
Vacaria	VAC	MR1	X	X		X

Table 2. Soybean (*Glycine max*) genotypes and lineages tested in the value for cultivation and use trials in the 2012/2013–2015/2016 crop seasons in the two evaluated macroregions (MR1 and MR2) of soybean adaptation in Brazil.

Genotype	Specification[1]	2012/2013	2013/2014	2014/2015	2015/2016
A 4724RG	RG	X	X		
BMX ENERGIA RR	RG	X	X		
DMario 58i	RG	X	X	X	X
NA 5909 RG	RG	X	X	X	X
BMX TURBO RR	RG	X	X	X	X
NS 6262	RG	X	X	X	X
BMX Potência RR	RG	X	X	X	X
NK 7059 RR	RG	X	X	X	X
NS L01	L	X	X		
NS L02	L	X	X		
NS L06	L	X	X		
NS L07	L	X	X	X	X
NS L08	L	X	X		
NS L11	L	X	X		
NS L12	L	X	X	X	X
NS L13	L	X	X		
NS L14	L	X	X	X	X
NS L23	L			X	X
NS L24	L			X	X
NS L25	L			X	X
NS L26	L			X	X
NS L27	L			X	X

[1]RG, released genotype; and L, line.

high, vector length is proportional to the squared root of heritability (h) (Yan, 2014).

A linear map was built to facilitate the identification of similarity among environments. This way, it was possible to identify the patterns of the genotype × location interaction and genotype x environment (GE) interaction. GE was defined as follows: GE = GL + GY + GLY, where GL is the genotype × location interaction, GY is the genotype × year interaction, and GLY is the genotype × location × year interaction (Yan, 2016). Therefore, the GE interaction increases when the year effect is high. The genotypic effect (G) vs. GE analysis allows for inferences about the capability of the environment (considered a combination of location + year) for the selection of superior genotypes (Yan, 2014). The vector analysis of the environments, which also allows for the construction of a linear map, shows the distance between environments; the proximity among environments indicates a positive genetic correlation among them (Yan, 2014). In addition, the position of environments on the linear map makes it possible to identify patterns in relation to GL and GE. Therefore, if environments are mainly grouped by location and not year on the map, GL effects will surpass those of GE and vice versa (Yan, 2014).

For all analyses, the following parameters of the GGEbiplot software (Yan, 2001, 2014) were used: "Transform = 0", meaning data were not transformed; "Scaling = 2", data scaled by standard deviation (SD-scaled) and by adjusted heritability (h-weighted); and "Data centering = 2", data centering by genotype + genotype x environment interaction (G + GE) and singular value partitioning ("SVP = 2"), focusing on the environment.

Results and Discussion

According to the GGL + GGE analysis, Chapada is the most representative test location because it was highly correlated with the average environment, indicating that the obtained results must be consistent over the years (long vector) (Figure 2). Furthermore, it was observed that this test location discriminates genotypes mainly through G (Yan, 2015).

In the G vs. GE analysis based on several environments, the interpretation of the biplot may be limited by location overlapping (Figure 3). However,

more intelligible results for MR1 and MR2 are presented in Tables 3 and 4, respectively. The average environment coordinate index on the x-axis (AEC_X) is integrated for environmental evaluations and is useful even in cases of a low explanation of the biplot (Yan, 2014). Therefore, the higher the AEC_X value, the higher is the potential of the environment to select superior genotypes. The AEC y-axis (AEC_Y) column refers to the potential of a location for the selection of genotypes through G or GE. The higher the modular AEC_Y value, the lower the potential of the location to select genotypes by G. Therefore, environments characterized by AEC_Y values between 0.30 and -0.30 were considered adequate for selecting genotypes through G, whereas environments with AEC_Y above these values were considered adequate for selecting genotypes based on GE. In summary, the most desirable environment will be characterized by a high AEC_X value and a close-to-zero AEC_Y value.

The GGL + GGE analysis also indicated that Chapada and Palma Sola are the most representative locations in MR1, since both had a high r_g with an elevated average environment and a genetic correlation between them (Figure 2). The G vs. GE analysis revealed that Campos Novos, in 2014, presented the highest potential for selection of superior genotypes, i.e., the highest $r_g h$ (Figure 3 A and Table 3); however, this environment is less representative of MR1, compared with Chapada and Palma Sola (Figure 2). The Chapada environment was placed to the right of the AEC line across years, which indicates its effectiveness as a site for selecting superior genotypes. This environment showed the best performance among these locations in 2015, but it exceeded optimal vector length in 2014, which indicates that it would not be an effective site for selecting genotypes by G. Also in 2014, Palma Solo had the third highest potential for selecting superior genotypes. The linear correlation analysis (linear mapping) showed that Chapada, in 2012 and 2015, was highly correlated and close on the linear map, indicating its similarity over the years (Figure 3 B). However, most environments were mainly grouped by year and not by location, which indicates that GE dominates over GL.

GE consists mainly in the GL, GY, and GLY interactions (Yan, 2016). In this way, the identification of locations that have a part of GL that can be predictable, such as Chapada, is important to confirm
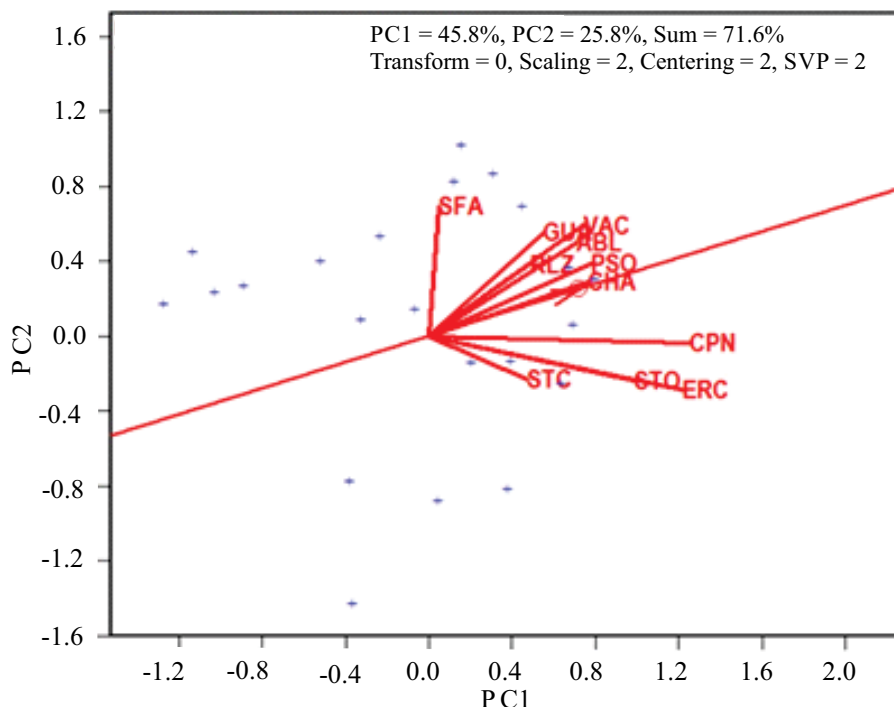
Figure 2. Biplot for genotype main effects plus genotype × location interaction (GGL) + genotype main effects plus genotype × environment interaction (GGE), using data from the value for cultivation and use trials in the 2012/2013–2015/2016 crop seasons in macroregion 1 of soybean (*Glycine max*) adaption in Brazil. Genotypes are represented by +. Environments consist of locations (municipalities) in the Southern region of the country: SFA, São Francisco de Assis; GUA, Guarapuava; VAC, Vacaria; ABL, Abelardo Luz; RLZ, Realeza; PSO, Palma Sola; CHA, Chapada; CPN, Campos Novos; STO, Santo Augusto; ERC, Erechim; and STC, Santa Cruz do Sul.
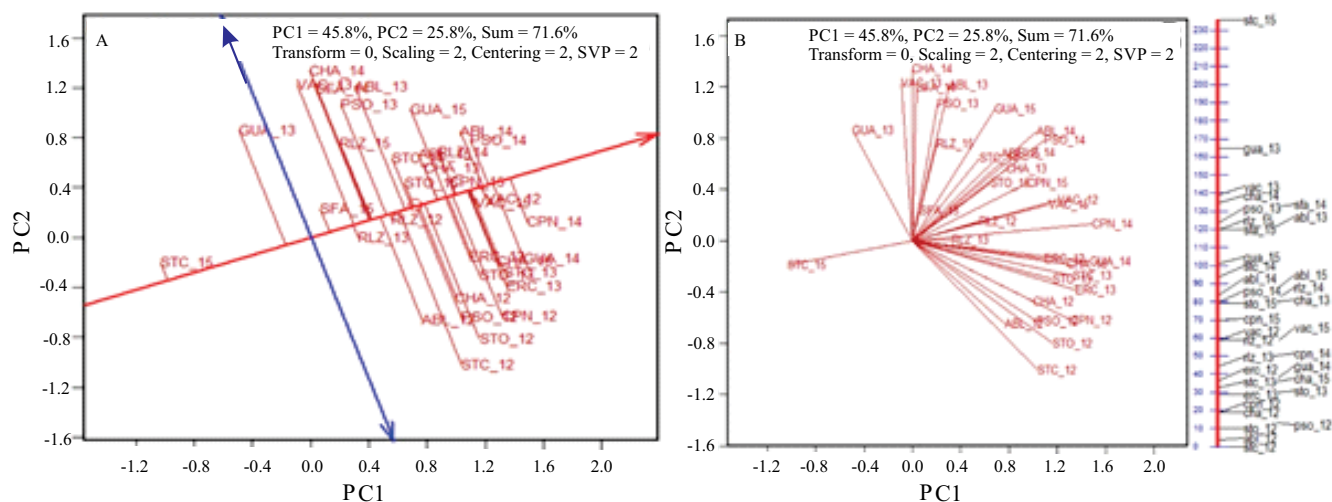

Figure 3. Biplot displaying the ability of each environment to select according to genotypic effect (G) vs. genotype x environment effects (GE) (A) and environment vector view of the GGE biplot in a linear map (B), using data from the value for cultivation and use trials in the 2012/2013–2015/2016 crop seasons in macroregion 1 of soybean (*Glycine max*) adaptation in Brazil. Environments consist of location (Brazilian municipality) combined with different experimental years: STC, Santa Cruz do Sul; GUA, Guarapuava; SFA, São Francisco de Assis; RLZ, Realeza; VAC, Vacaria; CHA, Chapada; ABL, Abelardo Luz; PSO, Palma Sola; STO, Santo Augusto; CPN, Campos Novos; and ERC, Erechim.

Table 3. Numerical values of the genotypic effect (G) vs. genotype x environment effect (GE) analysis for locations with the ability to select superior genotypes (AEC_X), stable by G and unstable by GE (AEC_Y), as well as vector length and representativeness of the environment correlated with the average environment axis (AEA), for macroregion 1 (MR1) of soybean (*Glycine max*) adaptation in Brazil.

Environment[1]	AEC_X	AEC_Y	Vector length	Correlation with AEA
CPN_2014	1.449	0.373	1.496	0.968
GUA_2014	1.327	0.644	1.475	0.9
PSO_2014	1.289	-0.37	1.341	0.961
ABL_2014	1.252	-0.459	1.333	0.939
VAC_2012	1.24	0.105	1.244	0.996
STC_2013	1.16	0.703	1.356	0.855
VAC_2015	1.155	0.112	1.16	0.995
CHA_2015	1.145	0.602	1.294	0.885
ERC_2013	1.143	0.816	1.404	0.814
RLZ_2014	1.049	-0.352	1.106	0.948
CPN_2015	1.047	-0.106	1.052	0.995
CPN_2012	1.033	1.029	1.458	0.708
STO_2013	0.995	0.673	1.201	0.828
GUA_2015	0.981	-0.74	1.229	0.798
ERC_2012	0.98	0.501	1.101	0.89
CHA_2013	0.919	-0.264	0.956	0.961
ABL_2015	0.917	-0.387	0.995	0.921
STO_2012	0.83	1.147	1.416	0.586
CHA_2012	0.783	0.788	1.111	0.705
PSO_2012	0.766	0.94	1.212	0.632
STO_2015	0.756	-0.201	0.782	0.966
STC_2014	0.734	-0.415	0.843	0.87
ABL_2013	0.684	-1.037	1.243	0.551
STC_2012	0.641	1.305	1.454	0.441
RLZ_2012	0.57	0.05	0.573	0.996
PSO_2013	0.545	-0.935	1.082	0.503
ABL_2012	0.507	0.877	1.013	0.501
RLZ_2015	0.437	-0.642	0.776	0.563
CHA_2014	0.435	-1.258	1.331	0.327
SFA_2014	0.422	-1.112	1.189	0.355
VAC_2013	0.317	-1.186	1.227	0.258
RLZ_2013	0.308	0.109	0.327	0.943
SFA_2015	0.134	-0.194	0.236	0.568
GUA_2013	-0.184	-0.969	0.986	-0.187
STC_2015	-1.04	-0.149	1.051	-0.99

[1]Environment consists of location (Brazilian municipality) combined with different experimental years: CPN, Campos Novos; GUA, Guarapuava; PSO, Palma Sola; ABL, Abelardo Luz; VAC, Vacaria; STC, Santa Cruz do Sul; CHA, Chapada; ERC, Erechim; RLZ, Realeza; STO, Santo Augusto; and SFA, São Francisco de Assis. Data are r_gh values of AEC_X, which is an index used to evaluate test environments and that is defined by the ratio between r_g (close to the cosine of the angle between environment and AEA) and h (environment vector length).

their efficiency for the selection of superior genotypes. The predictable GL mainly occurs when the location factor is dominated by stable, predictable, or controllable physical factors. The controllable physical factors consist of characteristics as altitude, latitude, soil properties, and established routine managements; the latter includes planting dates, planting density, irrigation, fertilization, and herbicide and fungicide applications (Yan, 2016).

The São Francisco de Assis and Santa Cruz do Sul locations are the less representative of MR1, with a lower genetic association with the average environment and short vectors in the GGL + GGE analysis (Figure 2). Furthermore, Santa Cruz do Sul, in 2015, and Guarapuava, in 2013, showed negative AEC_X values, which indicate a negligible potential for selecting superior genotypes. Therefore, Chapada, Palma Sola, and Campos Novos are the best candidates for a core location within MR1, but Chapada stands out, indicating that it can be considered the core location for this macroregion. According to Yan et al. (2011), test locations as Chapada are of type I, with a high representativity (small acute angles with the AEA) and repeatability (small acute angles between years within a location), indicating that a high selection intensity can be applied. Locations such as Palma Sola and Campos Novos, with a low-to-moderate representativity and a high repeatability, represent type II(a), i.e., they are useful for culling unstable genotypes, and a low selection intensity should be adopted during plant selection. Therefore, due to these characteristics, only Chapada is appropriate for plant selection in preliminary trials.

In MR2, the GGL + GGE analysis indicated that Maracaju is the most representative location, i.e., it had the highest r_g (Figure 4). In 2013 and 2012, the environment showed the highest potential for identifying superior genotypes, since it was located to the right of the AEA line (higher r_gh). In addition, selection in 2013 seems to be mostly dependent on G, because the environmental vector is shorter towards AEC than towards AEA (Figure 5 A and Table 4). The linear map showed that this environment was closely related in 2012, 2013, and 2015, representing a high r_g, and that the GL effect partially dominates over that of GE (Figure 5 B). Similarly to Chapada, the predictable GL in Maracaju is important for choosing this core
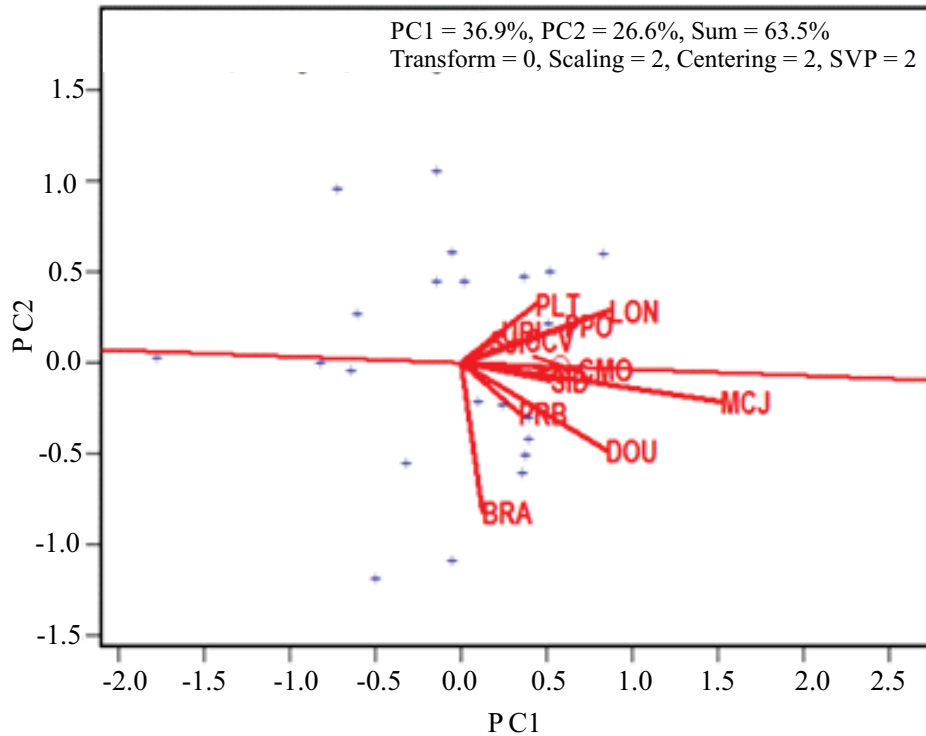
Figure 4. Biplot for genotype main effects plus genotype × location interaction (GGL) + genotype main effects plus genotype × environment interaction (GGE), using data from the value for cultivation and use trials in the 2012/2013–2015/2016 crop seasons for macroregion 2 of soybean (*Glycine max*) adaption in Brazil. Genotypes are represented by +. Environments consist of locations (municipalities) in the Southern region of the country: PLT, Palotina; UBI, Ubiratã; SJI, São Jorge do Ivaí; CCV, Cascavel; PPO, Ponta Porã; LON, Londrina; CMO, Cândido Mota; SID, Sidrolândia; MCJ, Maracaju; DOU, Dourados; PRB, Perobal; BRA, Brasilândia do Sul.
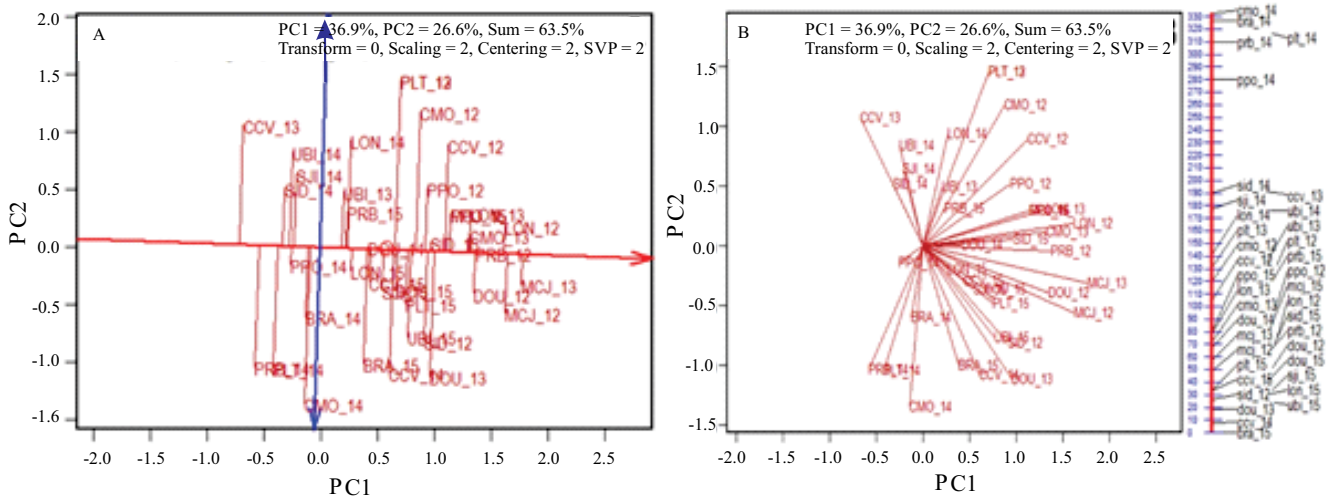

Figure 5. Biplot displaying the ability of the environment to select for genotypic effect (G) vs. genotype x environment effects (GE) (A) and environment vector view of the GGE biplot in a linear map (B), using data from the value for cultivation and use trials in the 2012/2013–2015/2016 crop seasons in microregion 2 of soybean (*Glycine max*) adaptation in Brazil. Environments consist of location (Brazilian municipality) combined with different experimental years: CCV, Cascavel; PRB, Perobal; SID, Sidrolândia; PLT, Palotina; UBI, Ubiratã; SJI, São Jorge do Ivaí; PPO, Ponta Porã; BRA, Brasilândia do Sul; CMO, Cândido Mota; LON, Londrina; DOU, Dourados; and MCJ, Maracaju.

Table 4. Numerical values of the genotypic effect (G) vs. genotype x environment effect (GE) analysis for locations with the ability to select superior (AEC_X) genotypes, stable by G and unstable by GE (AEC_Y), as well as vector length and representativeness of the environment correlated with the average environment axis (AEA), for macroregion 2 (MR2) of soybean (*Glycine max*) adaptation in Brazil.

Environment[1]	AEC_X	AEC_Y	Vector length	Correlation with AEA
MCJ_2013	1.779	0.263	1.799	0.989
MCJ_2012	1.643	0.517	1.722	0.954
LON_2012	1.611	-0.234	1.628	0.99
PRB_2012	1.373	0.006	1.373	1
DOU_2012	1.36	0.359	1.407	0.967
CMO_2013	1.316	-0.151	1.325	0.993
LON_2013	1.307	-0.335	1.349	0.969
MCJ_2015	1.133	-0.326	1.179	0.961
PPO_2015	1.133	-0.326	1.179	0.961
CCV_2012	1.091	-0.924	1.429	0.763
DOU_2013	0.993	1.09	1.474	0.674
SID_2015	0.973	-0.076	0.976	0.997
SID_2012	0.937	0.79	1.225	0.765
PPO_2012	0.918	-0.542	1.066	0.861
CMO_2012	0.835	-1.197	1.46	0.572
UBI_2015	0.793	0.752	1.093	0.726
PLT_2015	0.762	0.469	0.895	0.852
DOU_2015	0.678	0.349	0.763	0.889
PLT_2012	0.655	-1.466	1.606	0.408
PLT_2013	0.655	-1.466	1.606	0.408
CCV_2014	0.642	1.07	1.248	0.514
SJI_2015	0.56	0.36	0.665	0.841
CCV_2015	0.446	0.29	0.532	0.838
DOU_2014	0.42	-0.014	0.421	0.999
BRA_2015	0.417	0.996	1.079	0.386
LON_2015	0.272	0.198	0.336	0.809
LON_2014	0.237	-0.936	0.966	0.246
PRB_2015	0.224	-0.313	0.384	0.582
UBI_2013	0.186	-0.487	0.522	0.356
CMO_2014	-0.098	1.358	1.361	-0.072
BRA_2014	-0.108	0.609	0.619	-0.174
SJI_2014	-0.234	-0.624	0.666	-0.351
PPO_2014	-0.253	0.163	0.301	-0.84
UBI_2014	-0.274	-0.816	0.861	-0.319
SID_2014	-0.335	-0.495	0.598	-0.56
PLT_2014	-0.379	1.065	1.13	-0.335
PRB_2014	-0.542	1.061	1.191	-0.455
CCV_2013	-0.707	-1.036	1.254	-0.564

[1]Environment consists of location (Brazilian municipality) combined with different experimental years: MCJ, Maracaju; LON, Londrina; PRB, Perobal; DOU, Dourados; CMO, Cândido Mota; PPO, Ponta Porã; CCV, Cascavel; SID, Sidrolândia; UBI, Ubiratã; PLT, Palotina; SJI, São Jorge do Ivaí; and BRA, Brasilândia do Sul. Data are r_gh values of AEC_X, which is an index used to evaluate test environments and that is defined by the ratio between r_g (close to the cosine of the angle between environment and AEA) and h (environment vector length)

location, where a repeatable GE occurs across years (Yan, 2016).

According to the G vs. GE analysis, the following environments had negative AEC_X values and were located on the left of the AEC line: Cascavel in 2013, Perobal in 2014, Palotina in 2014, Sidrolândia in 2014, Ubiratã in 2014, Ponta Porã in 2014, São Jorge do Ivaí in 2014, and Brasilândia do Sul in 2014 (Figure 5 A). Moreover, Palotina, Ubiratã, Cascavel, São Jorge do Ivaí, Sidrolândia, and Perobal showed short vectors in the GGL + GGE analysis, which indicates inconsistency of results over the years and a low heritability. It was also observed that Brasilândia do Sul and Palotina are scarcely representative environments, and that Palotina, Ubiratã, Cascavel, São Jorge do Ivaí, Sidrolândia, and Perobal present short vectors with inconsistent results throughout the tested years. Therefore, the obtained results are indicative that Maracaju is the most suitable to be designated as a core location in MR2 for evaluating segregating populations. This means that this environment is representative of the macroregion, i.e., has a high r_g with the average environment and is characterized by a long vector, which indicates consistent results over the years. In addition, in 2012 and 2013, Maracaju exhibited the highest potential as a site for the efficient selection of superior genotypes. In this way, this location is suitable to conduct trials with no-repeatable evaluation in locations or years, as done for progeny lines, commonly in one or two locations in a single year, with no or few replicates (Yan et al., 2011).

In summary, Chapada and Maracaju were found to be the closest to a core location for MR1 and MR2, respectively. Both are desirable environments in most of the studied years, according to the G vs. GE analysis (Figures 2 and 4). Environments with short vectors towards AEC are desirable for selecting superior genotypes through G (Yan, 2014), as observed in Chapada in 2015 and 2013 and in Maracaju in 2012 and 2013. However, locations to the left side of the AEC line present negative r_gh and are inadequate for selecting superior genotypes (Yan, 2014). In practice, the identification of core locations, such as Chapada and Maracaju, can improve selection accuracy and provide a better allocation of resources. This is important due to the limited availability of resources in plant breeding programs, which requires breeders to choose carefully the correct number of replicates

and environments for each genotype (González-Barrios et al., 2019). Moreover, the repeatability of the genotype x environment interaction between years, which was verified mainly in Maracaju, is essential to increase the predictive ability of the environment (Lado et al., 2016).

Another important use of GGE biplots, besides identifying the best test locations, is the possibility of also identifying those that are inappropriate for trials (Krishnamurthy et al., 2017). In this sense, environments on the left side of the AEC line on the G vs. GE analysis are considered ineffective for selection and, therefore, undesirable, and may be replaced by other environments with a higher potential for selecting superior genotypes (Yan, 2014). This was the case of Santa Cruz do Sul in 2015 and Guarapava in 2013 in MR1, and of Cascavel in 2013, Perobal in 2014, Ponta Porão in 2014, Brasilândia do Sul in 2014, Ubiratã in 2014, and São Jorge do Ivaí in 2014 in MR2 (Figures 3 and 5). Nonrepresentative locations with short vectors (inconsistent results) in the GGL + GGE analysis may also be replaced. São Francisco de Assis and Santa Cruz do Sul, in MR1, and Brasilândia do Sul, Palotina, Ubiratã, Cascavel, São Jorge do Ivaí, Sidrolândia, and Perobal, in MR2, had short vectors and an elevated angle (low r_g with the average environment); therefore, they do not represent the target environment by the GGL + GGE analysis. Moreover, all these environments showed inconsistent results across years, indicating that they cannot be designated as core locations (Yan, 2015).

In both macroregions, the environment effects are mainly ranked by year and not by location, leading to a higher correlation due to GE. This occurs because the year effect is naturally randomized and unpredictable; therefore, GY and GLY are not reproducible (Yan, 2016). Minimizing environmental effects is essential for plant breeding, because environmental control decreases the effect of GE over GL, besides enabling a higher heritability and experimental precision. This was observed for Maracaju, which kept a similar position on the linear map, especially in 2012, 2013, and 2015. A portion of GL may repeat with a decreased year effect, which is related to location effects such as sowing date, planting density, irrigation, soil fertility, and other crop management factors. These characteristics minimize the effects of the interactions involving the year factor (Yan, 2016). Therefore, the

success of any breeding effort greatly depends on the high representativeness and potential for the selection of superior genotypes of each core location, as well as on the homogeneity of the experimental area and on an efficient crop management practice (Lucio & Sari, 2017).

The obtained results are indicative that the identification of core locations for the evaluation and selection of lineages in preliminary trials, with a large number of lines and a low seed availability, is of great importance for plant breeding.

Conclusion

The municipalities of Chapada, in the state of Rio Grande do Sul, and of Maracaju, in the state of Mato Grosso do Sul, Brazil, are the best core locations for macroregions 1 and 2 of soybean (*Glycine max*) adaptation in the country, respectively.

References

ALVARES, C.A.; STAPE, J.L.; SENTELHAS, P.C.; GONÇALVES, J.L. de M.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v.22, p.711-728, 2013. DOI: https://doi.org/10.1127/0941-2948/2013/0507.

DAS, A.; PARIHAR, A.K.; SAXENA, D.; SINGH, D.; SINGHA, K.D.; KUSHWAHA, K.P.S.; CHAND, R.; BAL, R.S.; CHANDRA S.; GUPTA, S. Deciphering genotype-by-environment interaction for targeting test environments and rust resistant genotypes in field pea (*Pisum sativum* L.). **Frontiers in Plant Science**, v.10, art.825, 2019. DOI: https://doi.org/10.3389/fpls.2019.00825.

DIA, M.; WEHNER, T.C.; HASSELL, R.; PRICE, D.S.; BOYHAN, G.E.; OLSON, S.; KING, S.; DAVIS, A.R.; TOLLA, G.E.; BERNIER, J.; JUAREZ, B. Value of locations for representing mega-environments and for discriminating yield of watermelon in the U.S. **Crop Science**, v.56, p.1726-1735, 2016. DOI: https://doi.org/10.2135/cropsci2015.11.0698.

GONZÁLEZ-BARRIOS, P.; DÍAZ-GARCÍA, L.; GUTIÉRREZ, L. Mega-environmental design: using genotype × environment interaction to optimize resources for cultivar testing. **Crop Science**, v.59, p.1899-1915, 2019. DOI: https://doi.org/10.2135/cropsci2018.11.0692.

KASTER, M.; FARIAS, J.R.B. **Regionalização dos testes de valor de cultivo e uso e da indicação de cultivares de soja**: terceira aproximação. Londrina: Embrapa Soja, 2012. Available at: <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/54939/1/Doc-330-OL1.pdf>. Accessed on: Feb. 14 2018.

KRISHNAMURTHY, S.L.; SHARMA, P.C.; SHARMA, D.K.; RAVIKIRAN, K.T.; SINGH, Y.P.; MISHRA, V.K.; BURMAN, D.; MAJI, B.; MANDAL, S.; SARANGI, S.K.;GAUTAM, R.K.; SINGH, P.K.; MANOHARA, K.K.; MARANDI, B.C.;

PADMAVATHI, G.; VANVE, P.B.; PATIL, K.D.; THIRUMENI, S.; VERMA, O.P.; KHAN, A.H.; TIWARI, S.; GEETHA, S.; SHAKILA, M.; GILL, R.; YADAV, V.K.; ROY, S.K.B.; PRAKASH, M.; BONIFACIO, J.; ISMAIL, A.; GREGORIO, G.B.; SINGH, R.K. Identification of mega-environments and rice genotypes for general and specific adaptation to saline and alkaline stresses in India. **Scientific Reports**, v.7, art.7968, 2017. DOI: https://doi.org/10.1038/s41598-017-08532-7.

LADO, B.; GONZÁLES BARRIOS, P.; QUINCKE, M.; SILVA, P.; GUTIÉRREZ, L. Modeling genotype × environment interaction for genomic selection with unbalanced data from a wheat breeding program. **Crop Science**, v.56, p.2165-2179, 2016. DOI: https://doi.org/10.2135/cropsci2015.04.0207.

LUCIO, A.D.; SARI, B.G. Planning and implementing experiments and analyzing experimental data in vegetable crops: problems and solutions. **Horticultura Brasileira**, v.35, p.316-327, 2017. DOI: https://doi.org/10.1590/s0102-053620170302.

MI, X.; WEGENAST, T.; UTZ, H.F.; DHILLON, B.S.; MELCHINGER, A.E. Best linear unbiased prediction and optimum allocation of test resources in maize breeding with doubled haploids. **Theoretical and Applied Genetics**, v.123, p.1-10, 2011. DOI: https://doi.org/10.1007/s00122-011-1561-4.

QIN, J.; XU, R.; LI, H.; YANG, C.; LIU, D.; LIU, Z.; ZHANG, L.; LU, W.; FRETT, T.; CHEN, P.; ZHANG, M.; QIU, L. Evaluation of productivity and stability of elite summer soybean cultivars in multi-environment trials. **Euphytica**, v.206, p.759-773, 2015. DOI: https://doi.org/10.1007/s10681-015-1513-1.

WU, J.; BONDALAPATI, K.; GLOVER, K.; BERZONSKY, W.; JENKINS, J.N.; MCCARTY, J.C. Genetic analysis without replications: model evaluation and application in spring wheat. **Euphytica**, v.190, p.447-458, 2013. DOI: https://doi.org/10.1007/s10681-012-0835-5.

YAN, W. Analysis and handling of G × E in a practical breeding program. **Crop Science**, v.56, p.2106-2118, 2016. DOI: https://doi.org/1770.2135/cropsci2015.06.0336.

YAN, W. **Crop variety trials**: data management and analysis. Chichester: Wiley-Blackwell, 2014. 360p.

YAN, W. GGEbiplot – a Windows application for graphical analysis of multienvironment trial data and other types of two-way data. **Agronomy Journal**, v.93, p.1111-1118, 2001. DOI: https://doi.org/10.2134/agronj2001.9351111x.

YAN, W. Mega-environment analysis and test location evaluation based on unbalanced multiyear data. **Crop Science**, v.55, p.113-122, 2015. DOI: https://doi.org/10.2135/cropsci2014.03.0203.

YAN, W.; FRÉGEAU-REID, J.; PAGEAU, D.; MARTIN, R.; MITCHELL-FETCH, J.; ETIENNE, M.; ROWSELL, J.; SCOTT, P.; PRICE, M.; HAAN, B. de; CUMMISKEY, A. LAJEUNESSE, J.; DURAND, J.; SPARRY, E. Identifying essential test locations for oat breeding in eastern Canada. **Crop Science**, v.50, p.504-515, 2010. DOI: https://doi.org/10.2135/cropsci2009.03.0133.

YAN, W.; HOLLAND, J.B. A heritability-adjusted GGE biplot for test environment evaluation. **Euphytica**, v.171, p.355-369, 2010. DOI: https://doi.org/10.1007/s10681-009-0030-5.

YAN, W.; PAGEAU, D.; FRÉGEAU-REID, J.; DURAND, J. Assessing the representativeness and repeatability of test locations for genotype evaluation. **Crop Science**, v.51, p.1603-1610, 2011. DOI: https://doi.org/10.2135/cropsci2011.01.0016..