Bragantia
ISSN: 0006-8705
ISSN: 1678-4499
Instituto Agronômico de Campinas

Giordani, Willian; Scapim, Carlos Alberto; Ruas, Paulo Maurício;
Ruas, Claudete de Fátima; Contreras-Soto, Rodrigo; Coan, Marlon;
Fonseca, Inês Cristina de Batista; Gonçalves, Leandro Simões Azeredo

Genetic diversity, population structure and AFLP markers associated with maize reaction to southern rust

Bragantia, vol. 78, no. 2, -June, 2019, pp. 183-196
Instituto Agronômico de Campinas

DOI: 10.1590/1678-4499.20180180

https://doi.org/10.1590/1678-4499.20180180

PLANT BREEDING - Article

Genetic diversity, population structure and AFLP markers associated with maize reaction to southern rust

Willian Giordani[1], Carlos Alberto Scapim[2], Paulo Maurício Ruas[3], Claudete de Fátima Ruas[3], Rodrigo Contreras-Soto[4], Marlon Coan[2], Inês Cristina de Batista Fonseca[1], Leandro Simões Azeredo Gonçalves[1]*

1.Universidade Estadual de Londrina - Departamento de Agronomia - Londrina (PR), Brazil.
2.Universidade Estadual de Maringá - Programa de Pós-graduação em Agronomia - Maringá (PR), Brazil.
3.Universidade Estadual de Londrina - Centro de Ciências Biológicas - Londrina (PR), Brazil.
4.Universidad de O'Higgins - Instituto de Ciencias Agronómicas y Veterinarias - Rancagua, Chile.

ABSTRACT: Maize is one of the species with greater genetic diversity among cereals and possibly the most diverse crop species known. Accessing this variability is essential for maize breeding, allowing breeders to achieve progress of yield increasing, overcome environmental challenges or deal with pests and diseases. Among the maize diseases, southern rust is one of the most important, causing significant losses in yield and presenting severe epidemics worldwide. In the present study, the AFLP technique was applied to analyze population structure and genetic diversity among 145 tropical maize inbred lines, and to test for preliminary evidence of association between AFLP markers and the reaction to southern rust. Disease severity was evaluated in two crop seasons and the accessions were genotyped through AFLP using four primer combinations. The clusters obtained based on the Jaccard genetic distance and Ward's hierarchical clustering and those achieved by structure simulations had high concordance and were capable of establish two big clusters, one predominantly of common maize and another of popcorn. The association analysis was performed using four different statistical models. The more complete model containing both population structure and genetic relatedness narrowed the number of significant associations, demonstrating its importance to control false associations. A total of 19 significant marker associations were identified from which three (EactMctg18, EactMctg205, EactMctg169) are interesting candidates for further investigations.

Key words: *Zea mays*, gwas, association studies, cluster analysis, genetic distance, *Puccinia polysora*.

*Corresponding author: leandrosag@uel.br

Received: May 18, 2018 – **Accepted:** Oct. 8, 2018

INTRODUCTION

Maize (*Zea mays* L.) is one of the species with greater genetic diversity among cereals. The huge diversity is shown by more than 350 races and multiple varieties, which carry many distinct characteristics (Vigouroux et al. 2008), allowing the worldwide crop growing in various environmental conditions (Ranum et al. 2014).

The conservation and identification of the genetic diversity remaining within the germplasm pools of crop species are crucial for the breeding process. Therefore, many efforts have been made over the recent decades in order to assess genetic variability of many species, such as maize, using morphological and molecular markers (Fu 2015). Despite the large genetic diversity of maize, the popcorn germplasm has a narrow genetic base (Ziegler 2001). This lack of diversity is a serious problem in breeding programs and can implicate the capacity of yield increasing, overcome environmental challenges or even in deal with pests or diseases.

One of the most important diseases that affect maize is the Southern rust, a fungal disease caused by *Puccinia polysora* Underw, which is considered the most destructive maize rust, historically presenting worldwide yield losses (Dudienas et al. 2013; Raid et al. 1988; Rhind et al. 1952). The ideal conditions for *P. polysora* development involve temperatures around 23 °C and long leaf wetness (Godoy et al. 1999). Thus the disease has been reported mostly in tropical and subtropical areas (Abadassi 2014).

The high variability and the great spreading of *P. polysora* are confirmed by the documentation of at least ten physiological races. Most of the resistance cases to *P. polysora* are race-specific and controlled by genes nominated *Rpp* (Brewbaker et al. 2011). Although the race-specific genes are more practical to work in breeding programs and usually confer a significant level of resistance, the horizontal resistance commonly is more durable and provides protection against greater range of races. Therefore, in tropical areas as Brazil, considering the high incidence and severity of southern rust, the horizontal resistance should be considered as a long-term strategy (Casela and Ferreira 2002).

Understanding the relationship between DNA polymorphism and variation in phenotypes is important for increasing the speed of selective breeding programs. Genome wide association studies (GWAS) can be used to identify significant association between molecular markers across the genome and phenotypic traits in panel of unrelated genotypes. This approach has been successfully applied to identify QTLs and genomic regions related to maize resistance to many important diseases, including southern rust (Zhou et al. 2018).

Although high-throughput genotyping technologies are becoming increasingly cheaper and can provide a vast number of SNPs, according to Zhang et al. (2014) the traditional markers, such as AFLP, can still act as easy-going approaches for many labs. For some studies, such as genetic diversity analysis, the resolution requirement for distinguishing the individuals can be reached by AFLP, and sequencing all genomes would be unnecessary and inflate the costs. Thus the AFLP technique has still been considerably applied in many investigations as genetic diversity access, population structure analysis and also association studies (Achleitner et al. 2008; Dadras et al. 2014; Ebrahimi et al. 2017; Saeed and Darvishzadeh 2017; Sharma et al. 2016)

Therefore, the aim of the present study was to analyze the population structure and genetic diversity among 145 tropical maize inbred lines from the germplasm collections of Universidade Estadual de Maringá, and test for preliminary evidence of association between AFLP markers and the maize reaction to southern rust.

MATERIAL AND METHODS

Field experiments

Field experiments were conducted in two crop seasons and evaluated in a randomized complete block design with two replicates. We tested 145 maize inbred lines derived from the germplasm collections of Universidade Estadual de Maringá among which 77 were previously classified as common maize and 68 as popcorn. The experiment performed in the first crop season was sown on November 30th 2014, while the second crop season experiment was sown on January 19th 2015, both by manual sown without any fungicide application. The experimental units consisted of two six-meter rows with 0.2 m × 0.9 m spacing between plants and rows, respectively.

Both experiments were performed at Iguatemi Research Station in Maringá, Paraná, Brazil, located at 23°25' S →

and 51°57' W. The region's climate is classified as Cfa following Köppen classification and the soil is categorized as Typical Red Dystrophic Argisol. The crop management and the fertilization were realized following the technical recommendation for maize crop. The climatological data of precipitation and average high and low temperatures were obtained by the weather station of the Iguatemi Research Station.

Regarding the high inoculum pressure of *Puccinia polysora* in the region, there was no necessity of artificial inoculation (Figure 1). Anyway, aiming to guarantee great inoculum pressure and high dispersion of the pathogen, the genotype IAC 112, previously classified as susceptible, was cultivated within and in the boarders of the experimental area. The southern rust severity evaluation was held by the infected leaf tissue area according to the referential scale proposed by Fantin[5] (1997). On average, five plants per treatment were scored when reached the VT stage.

DNA extraction and AFLP analysis

Around 30 days after emergency, leaves from five plants of each accession were collected in bulk and immediately frozen in liquid nitrogen until the storage at –80 °C. For DNA isolation, 300 mg of leaf tissue were powdered into liquid nitrogen and extracted using CTAB buffer according to Doyle and Doyle (1987) with minor modifications. DNA integrity and concentration was confirmed by electrophoresis in a 1% agarose gel and using the spectrophotometer NanoDrop 2000/2000c (Thermo Scientific, USA).

Figure 1. Pustules of *Puccinia polysora* on the upper surface of a maize leaf infected under natural inoculation.

The AFLP reactions were performed following the protocol described by Vos et al. (1995), with some modifications. Approximately 700 ng of DNA from each accession was doubly digested with *Eco*RI and *Mse*I enzymes (5 U each) in a 20 µL volume system for 18 h at 37 °C. The resulting DNA fragments were ligated with T4 DNA ligase (1 U) to *Eco*RI (0.5 µM) and *Mse*I (5 µM) adapters in a mix containing 1X T4 DNA ligase buffer, NaCl (0.05 M), BSA (50 ng/µL), and DTT (0.25 mM) in a final volume of 10 µL. The reaction was carried in a thermocycler at 37 °C for 3 h, 17 °C for 30 min, and 70 °C for 10 min and then diluted (1:4) with ultrapure water.

Subsequently, the fragments were amplified using a pair of primers with one selective base in a step known as pre-amplification. This reaction was performed using 3.5 µL of GoTaq®Green Master Mix (Promega, USA), 0.58 µL of the pre-selective primer (4.75 µM), 3.0 µL of the restriction/binding dilution and ultrapure water for a total volume of 10 µL. The pre-amplification program was 2 min at 72 °C, 20 cycles of 1 s at 94 °C, 30 s at 56 °C, and 2 min at 72 °C, followed by a final stage of 30 min at 60 °C. The pre-amplified product was diluted in ultrapure water (1:4).

Finally, selective amplification was carried out using diluted pre-amplified product and four primer pairs with three selective nucleotides (E-ACG/M-CAG, E-AGC/M-CAG, E-ACT/M-CTG and E-AAG/M-CTG) wherein the primers for *Eco*RI were labeled with fluorophores: NED, PET, VIC and FAM, respectively. In this process, a 2.5 µL aliquot of the previous dilution was mixed with: 0.54 µL of each selective primer (5µM *Mse*I and 1µM *Eco*RI), 3.5 µLGoTaq® Green Master Mix (Promega, USA) and ultrapure water, in a final volume of 10 µL. For the reactions the thermocycler was programmed as follows: one initial cycle of 2 min at 94 °C, 30 s at 64 °C and 2 min at 72 °C, followed by eight cycles of 1 s at 94 °C, 30 s at 64 °C (as touchdown with 1 °C lowering for each cycle) and 2 min at 72 °C, further 23 cycles of 1 s at →

[5]Fantin, G. M. (1997). Avaliação de resistência do milho a ferrugem causada por *Puccinia polysora* Underw. (PhD Thesis). Piracicaba: Escola Superior de Agricultura "Luiz de Queiroz".

94 °C, 30 s at 56 °C and 2 min at 72 °C, and finally the last step of 30 min at 60 °C.

The products of selective reactions (0.2 µL of each) were subjected to capillary electrophoresis using the automated system ABI 3500 xL Genetic Analyzer (Applied Biosystems, USA) in a solution containing 0.2 µL of size standard 600-LIZ (GeneScan v2.0) and 8.8 µL of highly deionized (Hi-Di) formamide. The results of the electrophoresis were combined in a binary matrix, scored for presence (1) or absence (0) of polymorphic bands, using the software GeneMapper® v.4.1 (Applied Biosystems, USA)

Data analysis

When considered the phenotypic data, since the southern rust severity showed heteroscedasticity between classes according to Bartlett's test and non-normality of residuals according to Shapiro–Wilk's test, the data were transformed by the expression: $y = \sqrt{x + 0.5}$.

Analysis of variance (ANOVA) was conducted to determine the main effects and interactions using F test ($p < 0.05$). The following mixed model equation was used (Eq. 1):

$$Y_{ijk} = \mu + G_i + R/E_{jk} + E_j + GE_{ij} + \varepsilon_{ijk,} \quad (1)$$

where Y_{ijk} is the severity measurement for the i^{th} accession, in the j^{th} crop season, on the k^{th} block; µ is the general mean; G_i is the random effect of the i^{th} accession; E_j is the fixed effect of the j^{th} crop season; R/E_{jk} is the random effect of the k^{th} block within the j^{th} crop season; GE_{ij} is the random effect of the genotype by crop season interaction of the i^{th} accession and the j^{th} crop season; and ε_{ijk} is the residual effect. The ANOVA, as well as the estimation of the genetic parameters of: broad-sense heritability (H^2), genotypic variance (σ^2_g), genotype by environment variance (σ^2_{ge}), environmental variance (σ^2_e), intraclass correlation coefficient, genotypic coefficient of variation (CV_g), environmental and genotypic and environmental coefficient of variation ratio (CV_g/CV_e) were computed using GENES software (Cruz 2006).

Analyses of the AFLP data were performed using the Jaccard genetic distance matrix and Ward's cluster analysis. These analyses were performed using the packages *ade4* on R statistical computing environment (R Development Core Team 2008). For the population structure analysis, the program STRUCTURE v2.3.4 was used to identify K discrete subpopulations based on admixture model with correlated allele frequencies. Marker data were coded as 1 or 0 and individuals were treated as haploids to avoid any assumptions about dominance or heterozygotes. Values of K ranging from 1 to 11 were tested, with 11 independent interactions for each grouping. Each numerical solution was optimized setting burn-in of 50,000 repetitions and Markov Chain Monte Carlo (MCMC) simulations of 500,000 iterations. The number of K groups that best fit the data set was determined according to the Δk value by Evanno method (Evanno et al. 2005). The membership in subpopulations was compared for reliability among replications and subsets. A representative solution was selected, and the membership of each accession in the K subpopulations was designated using a numerical index, establishing the Q matrix.

The pairwise relative kinship matrix (K matrix) of the association panel was obtained following Hardy (2003) by running SPAGeDi v1.4. All negative values between two individuals in the output from SPAGeDi were set to 0 and the matrix was formatted to a text file readable by TASSEL software v.5.2.

The association analysis between AFLP markers and maize reaction to *P. polysora* was performed in TASSEL v.5.2 using four different approaches. The first approach consisted in using a simple general linear model (GLM), known as *naïve*, which did not account for population structure or relative kinship as a potential cause of the genotype–phenotype relationship (GLM: G + P). In the second framework, GLM was tested taking population structure matrix (Q) as a cofactor (GLM: G + P+Q). The other two approaches consisted in utilizing mixed linear models (MLM) as suggested by Yu et al. (2006). Both have included the kinship matrix (K) among all accessions as random effect, initially not involving the population structure matrix (MLM: G + P + K) and then containing the kinship matrix plus the structure matrix (MLM: G + P+Q + K). The markers with minor allele frequency (MAF) < 0.05 were not taken into consideration for the analyses.

In this study, two thresholds were considered: Bonferroni test criterion assuming an alpha of 0.05, typically considered very strict; and the Bonferroni-corrected threshold ($p = 1/n$, where n = marker number), considered moderately conservative and thus widely adopted in the literature (Liu et al. 2016; Wang et al. 2012; Yang et al. 2014).

RESULTS AND DISCUSSION

Southern rust severity

Maize southern rust severity was scored in the diversity panel of 145 inbred lines across two crop seasons, with two replications in each environment. Analysis of variance identified significant effect ($p < 0.05$) between accessions, crop seasons and for the interaction between accessions and crop seasons (Table 1). The *P. polysora* severity in the first crop season was lower than in the second (0.96% and 5.11%, respectively). If we consider the frequency distribution of the severity performing class scores as proposed by Fantin (1997), we can see the variability of the association panel for the resistance to southern rust. It is possible to note a great difference in the frequency distribution in each crop season. In the first crop, among the 145 inbred lines evaluated, disease incidence was absent in ten lines which received score 1, and only eight were scored with notes higher than 5. These results showed that in the first crop, despite the high incidence of *P. polysora*, the disease severity was not high (Figs. 2a and 2b). In the second crop only lines A2560-65H23.3-176 and A2560-62H23.2-167 showed no signal or symptom of southern rust, while 86 had scores superior to 5 (Fig. 2b).

Taking into consideration the 145 lines, 13 were classified as highly resistant (A2560-62H23.2-167, A2560-63H23.2-170, W57, A2560-65H23.3-176, CD303-89H4.2-258, 30F33-69H26.1-188, 30F33-70H23.1-191, CML12, AVANT-14H5.5-24, DKB350-78H30.1-219, CML19, DAS2C599-95H34.4-276, DKB747-42-104A2560-62H23.2-167), presenting scores lower than 2 in both crop seasons. Of the 13 genotypes, the line A2560-62H23.2-167 showed no symptom of the disease (score 1) in both crops, revealing the potential of the usage of this material as source of resistance for breeding program, focusing in resistance to *P. polysora*.

Higher severity of southern rust in the second crop was also reported by Kurosawa et al. (2016) in a field experiment evaluating 37 popcorn inbred under natural inoculation. The climatic conditions more favorable to the pathogen development and its previous presence in the experimental area may have contributed to this occurrence. Precipitation and monthly minimum and maximum temperatures in the first crop experiment (Dec-Feb) were higher than those of the second crop season (Jan-Apr) (data not shown). The high temperatures observed in the first crop might be inadequate to the *P. polysora* development, exceeding the 25 °C, described as optimum to the pathogen (Godoy et al. 1999; Raid et al. 1988). Furthermore, according to Godoy et al. (2003) late sowings commonly show higher inoculum presence which, associated to ideal weather conditions, can result in severe epidemics of southern rust.

The genetic parameters associated with the reaction of maize to southern rust were estimated to ascertain the genetic variability in the diversity panel and to predict genetic gains from selection. We detect the prevalence of genotypic variance (σ^2_g) in relation to environmental variance, showing wide genetic differences between the accessions. This vast variability may be result of the →

Table 1. Analysis of variance for the severity of southern rust and estimation of the genetic parameters of 145 maize inbred lines and two crop seasons.

SV	DF	Sum of square	Mean square	F	Sig
Blocks/Crop season	2	0.283	0.141		
Accession (G)	144	234.28	1.627	5.941	**
Crop season (E)	1	164.147	164.147	17.079	**
G*E	144	118.188	0.820	2.997	**
Residual	288	78.866	0.273		
General mean	3.35				
First crop season mean	0.96				
Second crop season mean	5.11				
CV (%)	33.04				
Genotypic variance (σ^2_g)	0.3382				
Genotype by environment variance (σ^2_{ge})	0.1367				
Environmental variance (σ^2_e)	0.2738				
Broad-sense heritability (H^2)	83.168				
Intraclass correlation coefficient	55.263				
Genotypic coefficient of variation (CV_g)	36.726				
Environmental coefficient of variation (CV_e)	33.04				
Ratio CV_g/CV_e	1.111				

** Significant at 1% probability level

existence of resistance genes in some genotypes, allowing the identification of ones with higher levels of southern rust resistance and thus the association studies.

The estimated broad sense heritability (H^2) was 83.17%. This result corroborates with other authors who observed H^2 values from 72% to 93% (Kurosawa et al. 2016; Wanlayaporn et al. 2013) and indicates that the resistance to southern rust can be mostly attributed to genetic rather than environmental factors. The high heritability value suggests the possibility of achieving great genetic gains, enabling the phenotypic selection of genotypes resistant to *P. polysora*.

Considering the genotypic coefficient of variation (CV_g), we observed a value of 33.8%. For southern rust severity, similar values were found by Kurosawa et al. (2016), who asserted that this parameter is important because it allows the comparison of the genetic variability for different traits and helps to more precisely define the breeding strategies. For the best interpretation of the results, CV_g might be analyzed in conjunction with CV_e by means of the ratio CV_g/CV_e, to accurately assess the condition of the trait for genetic breeding. In this study we observed a ratio CV_g/CV_e of 1.11, indicating the presence of wide genetic variability and suggesting good genetic gain with selection.

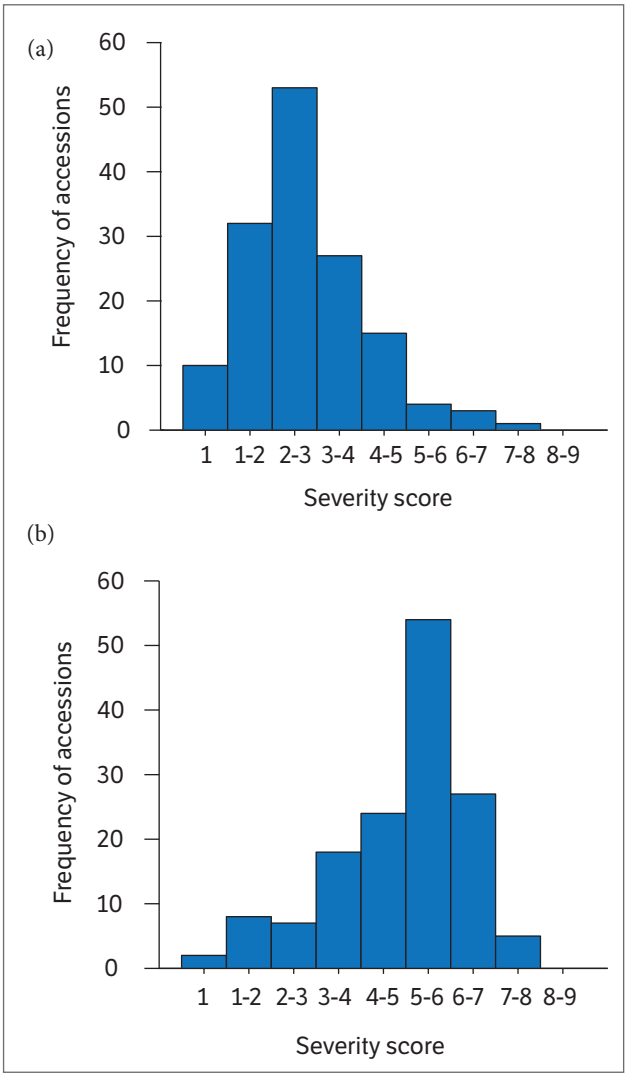

Figure 2. Histogram showing the frequency distribution of southern rust severity score across the maize diversity panel for (a) the first and (b) the second crop seasons.

Genetic diversity and population structure

The four primer pairs used for AFLP analysis resulted in 1008 bands, of which 975 were polymorphic, representing 97% polymorphism. The combinations E-AGC/M-CAG, AAG/M-CTG, E-ACG/M-CAG, E-ACT/M-CTG produced 321, 214, 232, and 241 bands, respectively, what shows the high efficiency of these primers in detecting genetic polymorphisms in maize.

The Jaccard similarity coefficient used to calculate the genetic distance among the 145 maize accessions ranged from 0.51 to 0.84, with a mean distance of 0.74. By Ward's hierarchical clustering analysis, two groups were well defined; clusters I and II, composed predominantly of common maize and popcorn, respectively. Cluster I consisted of 69 accessions of which 68 were previously classified as common maize (Table 2). Cluster II comprised 76 accessions of which 67 were previously defined as popcorn (Table 3). Within the Cluster II was observed the formation of two sub-clusters (Fig. 3). The first sub-cluster involved 47 accessions, all described as popcorn, while the second consisted of 29 accessions, amongst which 20 of popcorn and nine of common maize. This result suggests that the inbred lines belonging to the last sub-cluster share genetic similarities of common maize and popcorn.

The clustering of the popcorn showed that the accessions were not grouped according to their origin. As reported by Kantety et al. (1995) a relatively small number of popcorn inbred lines were developed from flint corn germplasm and selected to popping expansion and quality. Thus, the association pattern of the accessions was possibly a result of the narrow genetic base of the popcorn germplasm (Ziegler 2001).

Based on the simulations performed with the software STRUCTURE and the Δk value methodology proposed →

Table 2. Identification, name and origin of 77 common maize accessions from Universidade Estadual de Maringá germplasm collection.

ID	Accession	Origin	ID	Accession	Origin
L4	AG8080-7H3.1-3	–	L80	AVANT-13H5.4-21	–
L5	AVANT-14H5.5-24	–	L81	W57	–
L6	POP103 -88.1	AS1572	L83	PREMIUM-28H13.2-67	–
L7	30F33-71H26.2-194	–	L84	CML19	CIMMYT
L10	POP101-201-3	GARRA	L86	DKB747-45H17.5-115	–
L11	DKB350-78H30.1-219	–	L88	FORT-85H6.2-242	–
L12	DKB747-50H17.6-130	–	L89	DKB747-29H17.3-95	–
L15	FORT-87H6.4-248	–	L90	AG8080-8H3.2-6	–
L20	STRIKE-67H25.1-182	–	L92	DKB747-37H17.2-89	–
L22	POP201-195.1	P30R50	L95	CD303-88H4.1-255	–
L23	30F33-69H26.1-188	–	L99	30F33-70H23.1-191	–
L28	DKB747-43H17.4-107	–	L105	CD303-90H4.3-261	–
L29	CD303-89H4.2-258	–	L107	TORK-53H20.2-143	–
L34	AVANT-10H5.1-12	–	L108	POP203-56.1	SG6015
L38	DKB350-76H30.1-213	–	L109	POP102-166.5	P30B39
L39	AG6018-23H12.1-55	–	L110	A2560-62H23.2-167	–
L42	AVANT-12H5.3-18	–	L111	POP101-195.2	GARRA
L45	DKB747-38H17.2-92	–	L113	FLASH-22H11.1-52	–
L50	POP202-177.1	AS1570	L116	A2560-63H23.2-170	–
L51	FLASH-20H11.1-46	–	L119	TORK-55H20.3-149	–
L52	DAS422-80H31.2-227	–	L121	FORT-84H6.1-239	–
L56	30-23	P30P70 x Dow8460	L124	DAS2C599-95H34.4-276	–
L57	POP201-198.4	P30R50	L125	CML12	CIMMYT
L58	A2560-66H23.4-179	–	L127	A2560-65H23.3-176	–
L62	POP102-90.1	P30B39	L129	SPEED-81H33.1-230	–
L63	29-14	Penta x P30F53	L132	POP102-91.2	P30B39
L64	POP202-88.2	AS1570	L133	30.11	P30P70 x Dow846
L65	30-15	P30P70 x Dow8460	L134	POP101-197.1	GARRA
L67	POP201-192.1	P30R50	L135	29-154	Penta x P30F53
L68	29-92	Penta x P30F53	L136	DAS422-79H31.1-222	–
L69	POP203-51.2	SG6015	L137	A2560-170	–
L70	CML13	CIMMYT	L138	A2560-164	–
L71	POP103-80.5	AS1572	L139	DKB747-41-101	–
L72	30-29	P30P70 x Dow8460	L140	DKB747-47-121	–
L73	POP103-81.4	AS1572	L141	DKB747-48-124	–
L74	POP202-76.1	AS1570	L142	DAS422-8-222	–
L75	31-33	Penta	L143	DKB747-42-104	–
L77	CML22	CIMMYT	L145	53f-p37	–
L79	DKB350-19H9.1-43	-			

Table 3. Identification, name and origin of 68 popcorn accessions from Universidade Estadual de Maringá germplasm collection.

ID	Accession	Origin	ID	Accession	Origin
L1	GP1	Zélia	L59	P7-4-5	UEM-M2
L2	GP4	CMS43	L60	P9-12-1	IAC112
L3	P9-4-6	IAC112	L61	P9-6-3	IAC112
L8	P7-L7-1	UEM-M2	L66	P9-3-2	IAC112
L9	P16	P1283	L76	P8-2-2-4	Zaeli
L13	P8-1-1	Zaeli	L78	P19	-
L14	P20	-	L82	P7-2-3	UEM-M2
L16	P1-9	Zélia	L85	P1-3	Zélia
L17	P11-1	IAC125	L87	ANGELA-L70	Ângela: Embrapa
L18	P3-3T	CMS42	L91	BEIJAFLOR-L54	Beija-Flor: UFV
L19	P8-2	Zaeli	L93	BEIJAFLOR-L59	Beija-Flor: UFV
L21	P8-1-5-4	Zaeli	L94	GP14	Maradona
L24	P1-12	Zélia	L96	VIÇOSA-L77	Viçosa: UFV
L25	GP13	Jade	L97	P15	Colombiana
L26	P9-1	IAC112	L98	ANGELA-L71	Ângela: Embrapa
L27	P7-2-4	UEM-M2	L100	UFV-L80	Viçosa: UFV
L30	P9-4-5	IAC112	L101	ANGELA-L66	Ângela: Embrapa
L31	VIÇOSA-L75	Viçosa: UFV	L102	P3-1-2	CMS42
L32	P6-1	Catedral	L103	P1-8	Zélia
L33	GP15	Colombiana	L104	BEIJAFLOR-L52	Beija-Flor: UFV
L35	P8-1-5-9	Zaeli	L106	P9-5-3	IAC112
L36	P1-19	Zélia	L112	P9-1-3	IAC112
L37	P18	-	L114	P9-11-1	IAC112
L40	P4-4	CMS43	L115	BEIJAFLOR-L76	Beija-Flor: UFV
L41	P8-2-2-2	Zaeli	L117	P8-1-5-5	Zaeli
L43	P9-7-2	IAC112	L118	P9-8-1	IAC112
L44	P8-2-MULT	Zaeli	L120	P7-2-1	UEM-M2
L46	P11-2	IAC125	L122	P8-1-5-13	Zaeli
L47	GP10	Ângela	L123	GP3	CMS42
L48	P7-4-11	UEM-M2	L126	GP12	IAC125
L49	P9-5-1	IAC112	L128	GP5	UEM-J1
L53	P8-2-2-5	Zaeli	L130	P6-11	Catedral
L54	P8-1-5-10	Zaeli	L131	P1780	-
L55	P9-1-2	IAC112	L144	P9-1-6	IAC112

by Evanno et al. (2005), all 145 analyzed genotypes could be split into two groups (Figs. 4a and 4b). The percentage of membership for the K clusters for each individual was obtained and utilized to assign the accessions to each group (Fig. 4c). A genotype was assigned to a cluster when its percentage of membership was higher than 0.7, otherwise it was defined as admixed (Bitocchi et al. 2012; Dadras et al. 2014; Silva et al. 2015). Figure 4 shows that among the 145 accessions, 76 were assigned to group I, being the vast majority previously classified as common maize (Table 2). The group II was composed by 41 genotypes, most of them classified as popcorn (Table 3). Other 28 genotypes were classified as admixture.

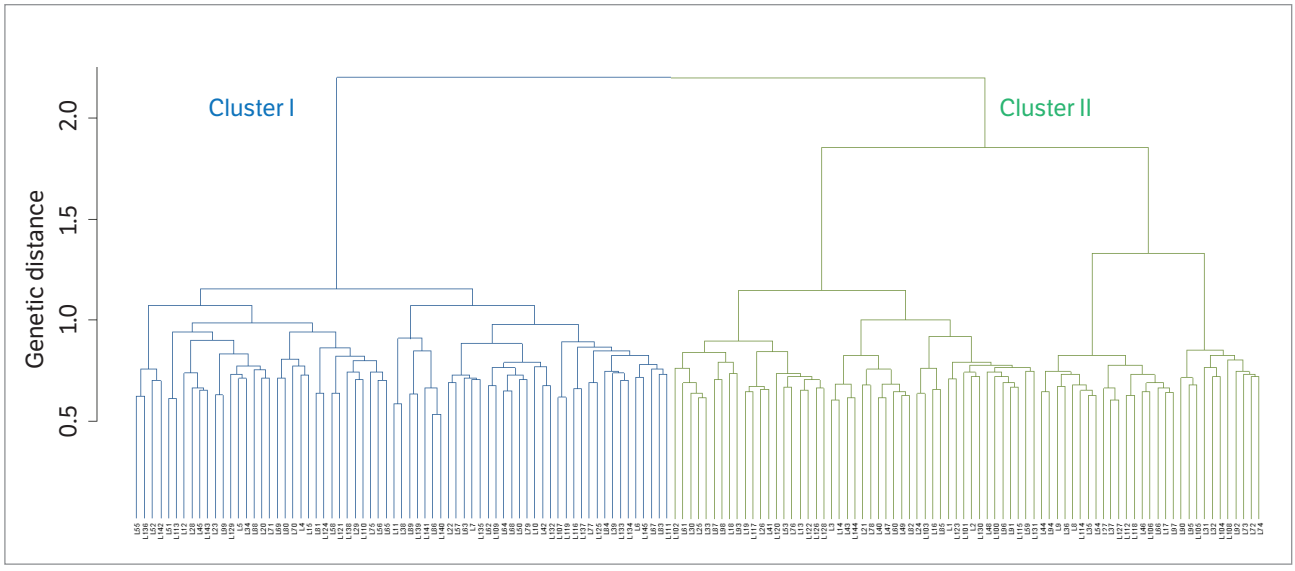

Figure 3. Ward's hierarchical clustering analysis of 145 maize inbred lines, based on Jaccard distances generated by the polymorphism of 975 AFLP markers.

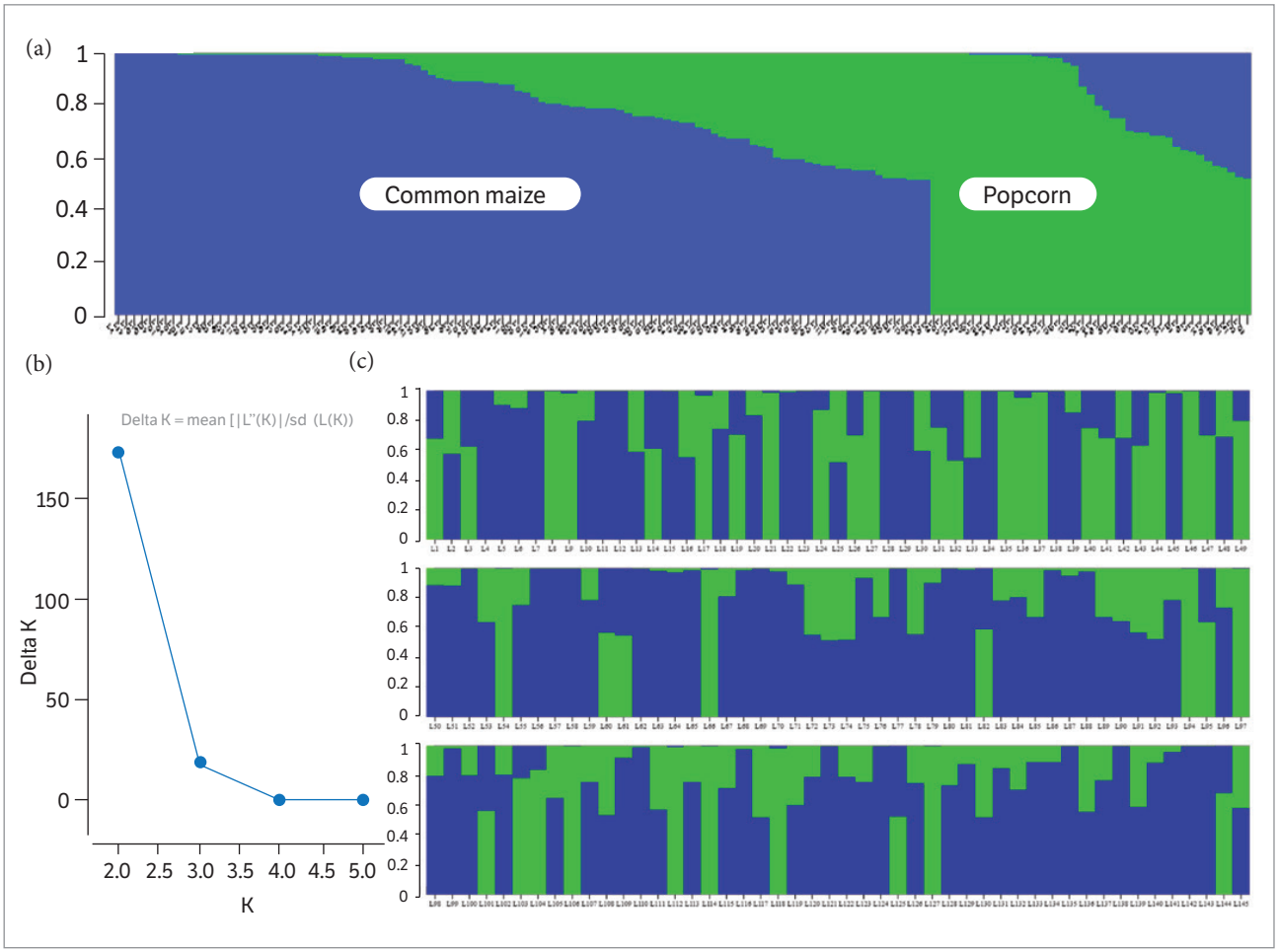

Figure 4. Population structure of 145 maize inbred lines inferred by STRUCTURE software and 975 AFLP markers data. (a) Formation of two groups according to the structure membership; (b) value of Δk and optimal number of K; and (c) membership of each genotype for the two groups.

In this study, the clusters obtained by the Ward's hierarchical clustering analysis and those achieved by structure analysis had high concordance to each other and were capable of establishing two big clusters, one of them predominantly of common maize and another of popcorn. Therefore, the apparent structure might be used to associate with other methods of clustering and showed been fairly consistent. Access the population structure has been frequently applied in many genetic studies as those which aim: to understand the genetic diversity among maize genotypes (Bracco et al. 2016), to determine heterotic groups of maize germplasm lines (Larièpe et al. 2017) and to control false positive associations between marker loci and phenotypic traits (Achleitner et al. 2008; Dadras et al. 2014; Kang et al. 2008; Saeed and Darvishzadeh 2017).

Association Analysis

Significant associations between AFLP markers (MAF > 0.05) and maize response to southern rust for the four tested models are reported in Table 4. The models detected differences in relation to the markers associated to the maize response and to the number of significant associations. Due to the relatively low sample size, which may result in limited statistical power, two thresholds were considered: the usual Bonferroni test ($p < 8.82 \times 10^{-5}$) and a cutoff value with a moderately stringent threshold ($p < 1.76 \times 10^{-3}$).

The independent analysis to each crop season also resulted in different associations. In the second crop season, the number of markers associated to the accessions response to the disease was higher than in the first crop (Table 1). A possible explanation for this phenomenon is the greater levels of severity verified in the second crop, what may be resulted in better representation of the genotypes reaction to southern rust.

Regarding all the models used in this study, the highest rate of significant markers were found by the model GLM: G + P, known as *naïve*, which did not account for population structure or relative kinship, as a potential cause of the genotype–phenotype relationship. When using this model in the first crop season, we identified two AFLP markers (EactMctg187, EaagMctg36), while in the second crop this number was much higher, and 15 markers were found (EaagMctg4, EagcMcag129, EactMctg18, EactMctg186, EaagMctg90, EactMctg207, EagcMcag117, EaagMctg162, EaagMctg131, EactMctg81, EaagMctg201, EactMctg158, EaagMctg94, EaagMctg148, EaagMctg140) (Table 4). Although the *naïve* model do not consider the presence of associations caused by population structure and genetic relatedness, it may lead to a considerable number of false positives (Yu et al. 2006).

In the model GLM: G + P+Q, the structure coefficient membership (Q) was incorporated as a cofactor. This model narrowed the number of significant associations when compared to the *naïve* model, showing the importance of the population structure control in reducing false associations. Using this model, for the first crop season the same two markers (EactMctg187, EaagMctg36) were identified as occurred with the *naïve* model. Although, for the second crop experiment the number of significant associations were cut down to five (EaagMctg117, EactMctg205, EactMctg18, EaagMctg4, EactMctg169), among which two (EactMctg205, EactMctg169) were not observed when applying the *naïve* model.

When considering the mixed linear model MLM: G + P+ K, which did not contemplate the population structure information, but contained the information of genetic relatedness by means of kinship matrix (K), a stronger impact was verified over the number of associations than that produced by GLM: G + P+Q. In the model MLM: G + P+ K, for the first crop experiment there were no significant associations, while for the second crop, three markers were found (EactMctg18, EactMctg205, EactMctg169). This result shows that when the kinship matrix was incorporated to the model, it was capable of reducing the inflation of p-values.

Finally, we tested the model MLM: G + P+Q + K, which incorporated both kinship matrix and population structure. This model is considered the most complete and usually results in better fit with the data (Yu et al. 2006). When applying the MLM: G + P+Q + K to the first crop season data, as well as observed in the MLM: G + P+ K, no significant association was found. For the second crop data, the significant markers (EactMctg18, EactMctg205, EactMctg169) were the same verified by MLM: G + P+ K.

If we consider all the models applied in this study, the number of significant marker associations under the moderately stringent threshold was 19, of which two were significant for the first crop season and 17 were significant for the second crop. The more complete model MLM: G + P+Q + K, which fits both population structure and kinship matrix, narrowed the →

Table 4. Portion of phenotypic variance (r^2), p-values (p) and size of AFLP markers significantly associated with maize reaction to southern rust based on four statistical models in two crop seasons.

First crop season			
Marker	**Size**	**p-value**	**r^2**
GLM: G + P			
EactMctg187	346pb	5.71E-05**	0.107
EaagMctg36	104pb	0.00134*	0.071
GLM: G + P+Q			
EactMctg187	346pb	5.42E-05**	0.109
EaagMctg36	104pb	0.00143*	0.069
Second crop season			
Marker	**Size**	**p-value**	**r^2**
GLM: G + P			
EaagMctg4	54pb	4.89E-05**	0.109
EagcMcag129	197pb	7.48E-05**	0.104
EactMctg18	78pb	1.40E-04*	0.097
EactMctg186	345pb	2.95E-04*	0.088
EaagMctg90	185pb	4.93E-04*	0.082
EactMctg207	393pb	6.97E-04*	0.078
EagcMcag117	183pb	8.08E-04*	0.076
EaagMctg162	327pb	8.32E-04*	0.075
EaagMctg131	247pb	8.74E-04*	0.075
EactMctg81	154pb	9.86E-04*	0.073
EaagMctg201	412pb	0.0012*	0.071
EactMctg158	261pb	0.00146*	0.069
EaagMctg94	194pb	0.0015*	0.068
EaagMctg148	282pb	0.00169*	0.067
EaagMctg140	273pb	0.00171*	0.067
GLM: G + P+Q			
EaagMctg117	227pb	1.24E-04*	0.091
EactMctg205	328pb	2.91E-04*	0.081
EactMctg18	78pb	5.39E-04*	0.074
EaagMctg4	54pb	7.18E-04*	0.071
EactMctg169	293pb	0.00145*	0.063
MLM: G + P+ K			
EactMctg18	78pb	6.72E-04*	0.084
EactMctg205	328pb	9.10E-04*	0.081
EactMctg169	293pb	0.00139*	0.074
MLM: G + P+Q + K			
EactMctg205	328pb	6.98E-04*	0.084
EactMctg18	78pb	7.66E-04*	0.082
EactMctg169	293pb	9.79E-04*	0.078

GLM = General linear model; G = Genotyping data matrix based on AFLP; P = Phenotyping data of southern rust severity; Q = Population structure inferred by STRUCTURE software; MLM = Mixed linear model; K = Kinship matrix estimated according to Hardy (2003). ** p-value significant for Bonferroni test; * p-value significant for the moderately conservative test.

number of significant associations to three, all in the second crop data. The results of association analysis demonstrated that considering both factors, population structure and genetic relatedness, arising from common kinship, is important to control false positive associations between markers and phenotype. This result is in agreement with other association studies using AFLP markers (Achleitner et al. 2008; Dadras et al. 2014; Ebrahimi et al. 2017; Saeed and Darvishzadeh 2017).

Among the 19 significant markers, the majority (13) was associated to maize susceptibility to *P. polysora*. Between these associations, the markers EactMctg205 (328 bp) and EactMctg169 (293 bp) were highlighted, once they were identified by means of the model MLM: G + P+Q + K and presented the highest effects of susceptibility, with a percentage of phenotypic variance (r^2) of 8.4% and 7.8%, respectively (Table 4).

Concerning all the markers associated to maize response to *P. polysora*, only six (EactMctg18, EactMctg186, EagcMcag117, EaagMctg162, EaagMctg201, EactMctg158) were associated to maize resistance. Among these markers, EactMctg18 (78bp) should be highlighted, once it was associated by the most complete model MLM: G + P+Q + K and had r^2 of 8.2% (Table 4).

Although further validation tests are required, the results of this study provided at least three (EactMctg18, EactMctg205, EactMctg169) interesting candidates for further investigations, considering its inclusion in plant breeding strategies as marker assisted selection. Regarding this application, one of the issues that would need to be determined is if the linkage between the markers and QTLs is small enough and permits the employment of these markers for reliably selection for southern rust resistance.

The main reason for the few (GLM) or the absence (MLM) of significant association for the first crop season data is attributed to the lower severity of southern rust, consequence of the weather conditions that were not favorable to the disease aggressiveness. For the second crop season, the weather conditions were more beneficial to the disease expression resulting in a higher number of significant associations.

It is important to emphasize that for the effective use of AFLP markers in assisted selection, these markers should be converted into sequence characterized amplified region (SCAR). SCAR markers, which are detected by single loci, are less laborious, can be identified in agarose gel, and are →

more reproducible, which makes these markers suitable for large-scale screening (Wei et al. 2009).

CONCLUSION

The AFLP markers are efficient to access the genetic diversity of maize accessions from the Universidade Estadual de Maringá germplasm, allowing the identification of two clusters composed of common maize and popcorn. The mixed linear model, MLM: G + P+Q + K, which fits both population structure and kinship, narrowed the number of significant associations, reducing the chance of obtaining false associations. Three AFLP markers: EactMctg18, EactMctg169, and EactMctg205 are associated with the maize response to southern rust and are considered promising for further studies, for genetic validation and for carrying out marker-assisted selection.

AUTHORS' CONTRIBUTION

Conceptualization, Giordani W., Gonçalves L. S. A., Fonseca I. B. and Scapim C. A.; Methodology, Giordani W. and Coan M.; Formal Analysis, Giordani W. and Contreras-Soto, R.; Investigation, Giordani W.; Resources, Gonçalves L. S. A., Scapim C. A., Ruas C. F. and Ruas P. M.; Writing – Original Draft, Giordani W.; Writing – Review and Editing, Giordani W., Gonçalves L. S. A., Ruas C. F. and Ruas P. M.; Supervision, Gonçalves L. S. A.; Project Administration, Gonçalves L. S. A.; Funding Acquisition, Gonçalves, L. S. A.

ORCID IDs

W. Giordani
https://orcid.org/0000-0002-1685-1446
C. A. Scapim
https://orcid.org/0000-0002-7047-9606
P. M. Ruas
https://orcid.org/0000-0001-5462-1900
C. F. Ruas
https://orcid.org/0000-0002-3907-5777
R. Contreras-Soto
https://orcid.org/0000-0001-6468-9394
M. Coan
https://orcid.org/0000-0002-1568-4893
I. C. B. Fonseca
https://orcid.org/0000-0003-0129-8534
L. S. A. Gonçalves
https://orcid.org/0000-0001-9700-9375

REFERENCES

Abadassi, J. (2014). Agronomic evaluation of temperate maize populations in tropical zone. Journal of Applied Biology & Biotechnology, 2, 8-11. https://doi.org/10.7324/JABB.2014.2302

Achleitner, A., Tinker, N. A., Zechner, E. and Buerstmayr, H. (2008). Genetic diversity among oat varieties of worldwide origin and associations of AFLP markers with quantitative traits. Theoretical and Applied Genetics, 117, 1041–1053. https://doi.org/10.1007/s00122-008-0843-y

Bitocchi, E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., Zeuli, P. S., Logozzo, G., Stougaard, J., McClean, P., Attene, G. and Papa, R. (2012). Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. Proceedings of the National Academy of Sciences of the United States of America, 109, 788-796. https://doi.org/10.1073/pnas.1108973109

Bracco, M., Cascales, J., Hernández, J. C., Poggio, L., Gottlieb, A. M. and Lia, V. V. (2016). Dissecting maize diversity in lowland South America: genetic structure and geographic distribution models. BMC Plant Biology, 16, 186. https://doi.org/10.1186/s12870-016-0874-5

Brewbaker, J. L., Kim, S. K., So, Y. S., Logroño, M., Moon, H. G., Ming, R., Lu, X. and Josue, A. D. (2011). General Resistance in Maize to Southern Rust (*Puccinia polysora* Underw.). Crop Science, 51, 1393-1409. https://doi.org/10.2135/cropsci2010.06.0327

Casela, C. R. and Ferreira, A. S. (2002). Variability in isolates of *Puccinia polysora* in Brazil. Fitopatologia Brasileira, 27, 414-417. https://doi.org/10.1590/S0100-41582002000400015

Cruz, C. D. (2006). Programa Genes: estatística experimental e matrizes. Viçosa: UFV.

Dadras, A. R., Sabouri, H., Nejad, G. M., Sabouri, A., and Shoai-Deylami, M. (2014). Association analysis, genetic diversity and structure analysis of tobacco based on AFLP markers. Molecular Biology Reports, 41, 3317-3329. https://doi.org/10.1007/s11033-014-3194-6

Doyle, J. J. and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin, 19, 11-15.

Dudienas, C., Fantin, G. M., Duarte, A. P., Ticelli, M., Bárbaro, I. M., Freitas, R. S., L., Filho, G. C., Bolonhezi, D. and Pântano, A. P. (2013). Severidade de ferrugem polissora em cultivares de milho e seu efeito na produtividade. Summa Phytopathology, 39, 16-23. https://doi.org/10.1590/s0100-54052013000100003

Ebrahimi, F., Majidi, M. M., Arzani, A. and Mohammadi-Nejad, G. (2017). Association analysis of molecular markers with traits under drought stress in safflower. Crop and Pasture Science, 68, 167-175. https://doi.org/10.1071/cp16252

Evanno, G., Regnaut, S. and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology, 14, 2611-2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fu, Y. B. (2015). Understanding crop genetic diversity under modern plant breeding. Theoretical and Applied Genetics, 128, 2131-2142. https://doi.org/10.1007/s00122-015-2585-y

Godoy, C. V, Amorim, L. and Bergamin, A. (1999). Influência da duração do molhamento foliar e da temperatura no desenvolvimento da ferrugem do milho causada por *Puccinia polysora*. Fitopatologia Brasileira, 24, 160-165.

Godoy, C. V, Amorim, L., Bergamin Filho, A., Silva, H. P., Silva, W. J. and Berger, R. D. (2003). Temporal progress of southern rust in maize under different environmental conditions. Fitopatologia Brasileira, 28, 273-278. https://doi.org/10.1590/S0100-41582003000300008

Hardy, O. J. (2003). Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. Molecular Ecology, 12, 1577-1588. https://doi.org/10.1046/j.1365-294X.2003.01835.x

Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J. and Eskin, E. (2008). Efficient control of population structure in model organism association mapping. Genetics, 178, 1709-1723. https://doi.org/10.1534/genetics.107.080101

Kantety, R. V, Zeng, X., Bennetzen, J. L. and Zehr, B. E. (1995). Assessment of genetic diversity in dent and popcorn (*Zea mays* L.) inbred lines using inter-simple sequence repeat (ISSR) amplification. Molecular Breeding, 1, 365-373. https:/doi.org/10.1007/BF01248414

Kurosawa, R. N. F., Vivas, M., Júnior, A. T. A., Santos, A., Mafra, G. S., Guimarães, A. G. and Schwantes, I. A. (2016). Reaction of popcorn germplasm to polysora rust under field conditions and natural inoculation. Tropical Plant Pathology, 41, 415-422. https://doi.org/10.1007/s40858-016-0112-6

Larièpe, A., Moreau, L., Laborde, J., Bauland, C., Mezmouk, S., Décousset, L., Mary-Huard, T., Fiévet, J. B., Gallais, A., Dubreuil, P, and Dubreuil, P. (2017). General and specific combining abilities in a maize (*Zea mays* L.) test-cross hybrid panel: relative importance of population structure and genetic divergence between parents. Theoretical and Applied Genetics, 130, 403-417. https://doi.org/10.1007/s00122-016-2822-z

Liu, S., Fan, C., Li, J., Cai, G., Yang, Q., Wu, J. and Zhou, Y. (2016). A genome-wide association study reveals novel elite allelic variations in seed oil content of *Brassica napus*. Theoretical and Applied Genetics, 129, 1203-1215. https://doi.org/10.1007/s00122-016-2697-z

R Development Core Team (2008). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Raid, R. N., Pennypacker, S. P. and Stevenson, R. E. (1988). Characterization of *Puccinia polysora* epidemics in Pennsylvania and Maryland. Phytopathology, 78, 579-585. https://doi.org/10.1094/Phyto-78-579

Ranum, P., Peña-Rosas, J. P. and Garcia-Casal, M. N. (2014). Global maize production, utilization, and consumption. Annals of the New York Academy of Sciences, 1312, 105-112. https://doi.org/10.1111/nyas.12396

Rhind, D., Waterston, J. M. and Deighton, F. C. (1952). Occurrence of *Puccinia polysora* Underw. in west Africa. Nature, 169, 631. https://doi.org/10.1038/169631a0

Saeed, A. and Darvishzadeh, R. (2017). Association analysis of biotic and abiotic stresses resistance in chickpea (*Cicer* spp.) using AFLP markers. Biotechnology & Biotechnological Equipment, 31, 698-708. https://doi.org/10.1080/13102818.2017.1333455

Sharma, S. S., Islam, M. A., Singh, V. K., Negi, M. S. and Tripathi, S. B. (2016). Genetic diversity, population structure and association study using TE-AFLP markers in *Pongamia pinnata* (L.) Pierre germplasm. Tree Genetics & Genomes, 13, 6-12. https://doi.org/10.1007/s11295-016-1088-6

Silva, T. A., Cantagalli, L. B., Saavedra, J., Lopes, A. D., Mangolin, C. A., Silva, M. F. P. and Scapim, C. A. (2015). Population structure and genetic diversity of Brazilian popcorn germplasm inferred by microsatellite markers. Electronic Journal of Biotechnology, 18, 181-187. https://doi.org/10.1016/j.ejbt.2015.03.005

Vigouroux, Y., Glaubitz, J. C., Matsuoka, Y., Goodman, M. M., Sánchez G., J. and Doebley, J. (2008). Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. American Journal of Botany, 95, 1240-1253. https://doi.org/10.3732/ajb.0800097

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. van de, Hornes, M., Friters, A., Pot, J., Paleman, J. and Kuiper, M. (1995). AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research, 23, 4407-4414. https://doi.org/10.1093/nar/23.21.4407

Wang, M., Yan, J., Zhao, J., Song, W., Zhang, X., Xiao, Y. and Zheng, Y. (2012). Genome-wide association study (GWAS) of resistance to head smut in maize. Plant Science, 196, 125-131. https://doi.org/10.1016/j.plantsci.2012.08.004

Wanlayaporn, K., Authrapun, J., Vanavichit, A. and Tragoonrung, S. (2013). QTL mapping for partial resistance to southern corn rust using RILs of tropical sweet corn. American Journal of Plant Sciences, 4, 878-889. https://doi.org/10.4236/ajps.2013.44108

Wei, P., Feng, H., Piao, Z., Li, C., Liu, Z., Wang, Y., Ji, R., Zou, T. and Ji, S. (2009). Identification of AFLP markers linked to Ms, a genic multiple allele inherited male-sterile gene in Chinese cabbage. Breeding Science, 59, 333-339. https://doi.org/10.1270/jsbbs.59.333

Yang, N., Lu, Y., Yang, X., Huang, J., Zhou, Y., Ali, F., Wen, W., Liu, J. and Yan, J. (2014). Genome Wide Association Studies Using a New Nonparametric Model Reveal the Genetic Architecture of 17 Agronomic Traits in an Enlarged Maize Association Panel. PLOS Genetics, 10, e1004573. https://doi.org/10.1371/journal.pgen.1004573

Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., Buckler, E. S. and Holland, J. B. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature Genetics, 38, 203-208. https://doi.org/10.1038/ng1702

Zhang, Z., Parijs, F. R. D. Van and Xiao, B. (2014). The status of AFLP in the genomics era and a pipeline for converting AFLPs into single-locus markers. Molecular beeding, 35, 1245-1260. https://doi.org/10.1007/s11032-014-0113-4

Zhou, G., Hao, D., Mao, Y., Zhu, Q., Chen, G., Lu, H., Shi, M., Huang, X., Zhang, Z., Zhao, J., and Xue, L. (2018). Identification of genetic loci conferring partial resistance to soutern corn rust through a genome-wide association study. European Journal of Plant Pathology, 150, 1083-1090. https://doi.org/10.1007/s10658-017-1351-1

Ziegler, K. E. (2001). Specialty corns. In A. R. Hallauer (Ed.), Specialty corns (p. 199-214). Ames: CRC Press.