

NOTE

<http://dx.doi.org/10.1590/1984-70332014v14n4n41>

Genetic evaluation of popcorn families using a Bayesian approach via the independence chain algorithm

Marcos Rodovalho¹, Freddy Mora^{2*}, Osvin Arriagada², Carlos Maldonado², Emmanuel Arnhold³ and Carlos Alberto Scapim⁴

Received 26 April 2013

Accepted 2 September 2014

Abstract – *The objective of this study was to examine genetic parameters of popping expansion and grain yield in a trial of 169 half-sib families using a Bayesian approach. The independence chain algorithm with informative priors for the components of residual and family variance (inverse-gamma prior distribution) was used. Popping expansion was found to be moderately heritable, with a posterior mode of h^2 of 0.34, and 90% Bayesian confidence interval of 0.22 to 0.44. The heritability of grain yield (family level) was moderate ($h^2 = 0.4$) with Bayesian confidence interval of 0.28 to 0.49. The target population contains sufficient genetic variability for subsequent breeding cycles, and the Bayesian approach is a useful alternative for scientific inference in the genetic evaluation of popcorn.*

Key words: *Zea mays L., popping expansion, grain yield, heritability.*

INTRODUCTION

In Brazil, there is an annually increasing demand for popcorn, with an estimated consumption of approximately 70,000 t yr⁻¹, of which 71% is being imported from Argentina and the United States, mainly (Moterle et al. 2012). The low number of cultivars is a limiting factor for an expansion of the crop in Brazil. Popcorn represents a good alternative for small farmers due to the high price obtained per production unit and to the widespread consumption in Brazil (Pena et al. 2012).

In popcorn breeding, not only yield and agronomic traits, but also aspects related to the quality, e.g., popcorn texture and softness, should be taken into consideration. Farmers are interested in high yields and in other features of a normal healthy maize population, whereas consumers care about a high expansion volume (EC), which improves the popcorn texture and softness (Miranda et al. 2008).

In maize breeding, intrapopulation improvement methods are highly important (Van Inghelandt et al. 2010). The procedure aims to increase the frequency of favorable alleles for a desirable trait, while maintaining the genetic variability of the improved population (Doerksen et al. 2003). Selec-

tion among and within half-sib families is one of the most commonly used methods in Brazilian breeding programs (Rangel et al. 2008). The main goal of the intrapopulation improvement is to obtain highest gains with the new selected population, but uncareful selection could develop generations with high inbreeding levels, leading to a reduction in population vigor (Dhillon and Malhi 2005, Peiris and Hallauer 2005). The half-sib selection methods provide estimates of genetic parameters, selection gains, selection differentials, among others (Peiris and Hallauer 2005).

In maize breeding, the estimation of variance components and prediction of breeding values are key aspects in the genetic evaluation of traits of economic importance, which can be computed by means of Bayesian methods, incorporating the different variants of MCMC procedures (Markov chain Monte Carlo). These procedures determine the marginal posterior distribution of estimates of parameters. The Bayesian approach is considered an important tool in genetic evaluation, since the existing uncertainty about all model parameters (including variance components) is taken into consideration (Arriagada et al. 2012, Mora and Serra 2014). In the context of Bayesian inference, all model parameters are considered random variables according to the

¹ Syngenta Seeds, Paraná 1241, 85.812-010, Cascavel, PR, Brazil

² Universidad de Talca, Instituto de Ciencias Biológicas, 2 Norte 685, Talca, Chile. *E-mail: morapoblete@gmail.com

³ Universidade Federal de Goiás, Campus Samambaia, C.P. 131, 74.001-970, Goiânia, GO, Brazil

⁴ Universidade Estadual de Maringá, Departamento de Agronomia, Av. Colombo, 5790, 87.020-900, Maringá, PR, Brazil

concept of subjective probability. Using a generalization of Bayes' theorem, *a priori* information about the parameters was used in combination with sample data (represented by the likelihood function theorem), underlying a posterior inference on the data (Corander et al. 2003).

The purpose of this study was to estimate genetic parameters of popping expansion and grain yield in a trial of 169 half-sib families, using Bayesian inference as an alternative of scientific inference in the genetic evaluation of maize.

MATERIAL AND METHODS

The popcorn population used in this study consisted of a white-seeded EMU-C1. 169 half-sib families were evaluated in a square lattice design, with three replications. The experiment was conducted on the Iguatemi experimental farm of the State University of Maringá, in the 2008/2009 growing season.

The measured traits were grain yield (kg ha⁻¹) and expansion volume (g mL⁻¹). The grain yield was adjusted to 13% moisture and popping procedures were performed according to Pacheco et al. (2002).

Phenotypic data were analyzed according to the following linear mixed model (Bueno Filho and Vencovsky 2000):

$$y = X\beta + Zf + \varepsilon$$

Where y is the vector of the phenotypic observations (popping expansion and yield), X is the design matrix, β is the vector of replications and block within replications, Z is the incidence matrix for the family effects, f is the vector of family effects and ε is the vector of residuals or error vector. Variance (f) = $G = I \sigma_f^2$, and Variance (ε) = $R = I \sigma_e^2$, where I is an identity matrix with dimension equal to number of records; σ_f^2 and σ_e^2 are the family and residual variances, respectively. Each *a priori* distribution of the elements of the Bayesian model is given by

$f(\beta) \propto U$ (Flat prior; uniform)

$f(f | \sigma_f^2) \sim N(0, G)$ (Normal prior distribution)

$f(\sigma_f^2) \sim IG(a, b)$ (Inverse Gamma; shape parameter a and scale parameter b)

$f(\sigma_e^2) \sim IG(c, d)$ (Inverse Gamma; shape parameter c and scale parameter d)

Phenotypic observations are assumed to be normally distributed:

$f(y | \beta, f, \sigma_e^2, \sigma_f^2) \sim N(X\beta + Zf, R)$

The joint posterior distribution is given by

$$f(\beta, f, \sigma_e^2, \sigma_f^2 | y) \propto f(y | \beta, f, \sigma_e^2, \sigma_f^2) \cdot f(\beta) \cdot f(f | \sigma_f^2) \cdot f(\sigma_e^2) \cdot f(\sigma_f^2)$$

Variance components, heritability and prediction of breeding values were obtained using the *mixed* procedure of SAS (SAS Institute Inc. 2007), with the *prior* option and posterior sample size of 50,000, generated by the independence chain algorithm, which is a variant of the Markov Chain Monte Carlo (MCMC) methods. The following two steps were fit for Bayesian analysis in SAS: 1) In the first run, Jeffreys' prior distribution (a non-informative prior distribution) was considered to obtain the base density values of the variance component distributions; the base density is taken as a product of approximate inverse gamma distributions for the variance parameters, and 2) A second run was performed using the base density values obtained in step 1, as *a priori* information for the variance component distributions. In *mixed* procedure the base density for the variance components is obtained by deriving approximate inverse gamma distributions for each of the variance components, and taking the product of these distributions as the base density (Brown and Prescott 2006).

The significance of the family effect, and thus the significance of the family variance, from zero for each trait, was tested using the Schwarz Bayesian information criterion (BIC: Bayesian Information Criterion) and a 'one-tailed' likelihood ratio test (LRT). Response to selection for both traits were estimated at two selection rates (20.12% and 10.06%), according to Cané-Retamales et al. (2011). The Pearson correlation coefficient between the predicted breeding values of popping expansion and of grain yield was calculated to determine the genetic relationship between the traits, using the *corr* procedure of SAS (SAS Institute Inc. 2007).

RESULTS AND DISCUSSION

Significant differences were found among families for both traits, according to the Bayesian information criterion and the likelihood ratio test (Table 1). According to BIC there was positive evidence against the reduced model where the family effect is assumed to be zero (i.e. $\hat{\sigma}_f^2=0$) for popping expansion (ΔBIC 2 to 6) and strong evidence against the reduced model for grain yield (ΔBIC 7 to 10); significance differences for ΔBIC are based on Neath and Cavanaugh (2012).

The estimates of variance components and heritability using Bayesian inference via independence chain algorithm, showed moderate inheritance for popping expansion and grain yield (Table 2) with h^2 of 0.34 (posterior mode) and confidence intervals from 0.22 to 0.44 (90% probability), for

popping expansion, and h^2 of 0.4 with confidence interval from 0.28 to 0.49, for grain yield.

Analyzing the heritability values, it was noted (Table 2) that these values for popping expansion were similar to those of grain yield; the confidence interval of heritability of popping expansion includes the point estimates of yield heritability. The results of the heritability of yield and popping expansion are lower than those reported elsewhere, for example, by Arnhold et al. (2009), Daros et al. (2002) and Santos et al. (2008). In a trial with S2 families of the Beija-Flor population, Santos et al. (2004) estimated a heritability of 0.72 for popping expansion. Popping expansion has positive, negative or zero dominance (Lu et al. 2003, Miranda et al. 2008, Moterle et al. 2012). This lack of agreement hampers the exploitation of heterosis. According to several authors, the heritability of popping expansion ranges generally from 70 to 90% (Pereira and Amaral Júnior 2001, Arnhold et al. 2009) and is a trait which has additive

effects as main component of genetic variance (Lu et al. 2009, Moterle et al. 2012). Therefore, in general, according to Lu et al. (2003), popping expansion is less influenced by the environment and dominance deviations than yield, which may result in a greater association between *per se* performance and crosses.

Table 3 shows the estimates of genetic gain (family selection) considering two selection rates (20.12 and 10.06%) calculated according to the predicted Bayesian values of popcorn families (posterior mode). Genetic gains were of 4% - 5.2%, for popping expansion, and of 7.1 - 9.2% for grain yield. The significant genetic variability among popcorn families of the EMU-C1 population allowed obtaining the genetic gains for both traits.

The point estimates of mean, median and mode of the posterior distributions of the variance components and heritability were similar, showing a lack of asymmetry.

Table 1. Bayesian information criterion of Schwarz (BIC), and the restricted likelihood ratio test (RLRT) for the significant effect of family in popcorn

Indicators	Popping expansion		Grain yield	
	FM	RM	FM	RM
-2 LLR	2870.7	2879.6	6930.2	6940.9
BIC	2881.0	2885.7	6940.3	6947.0
Δ BIC (Model NM – Model FM)		4.7		6.7
RLRT	8.9*		10.7*	

FM: full model (with family effect) RM: reduced model (no family effect) * statistically significant, according to the test χ^2 ($p < 0.05$); -2 LLR: -2 multiplied by log likelihood (restricted).

Table 2. Bayesian estimation of variance components and heritability (family level) for grain yield and expansion volume in popcorn

Popping expansion	Bayesian estimates*					
	Mean	Median	Mode	SD	LL	UL
Family variance	3.249	3.218	3.056	0.865	1.889	4.721
Residual variance	19.064	19.024	18.980	1.094	17.347	20.938
Additive variance	12.994	12.870	12.224	3.459	7.555	18.886
h^2	0.334	0.337	0.344	0.066	0.220	0.438
Grain yield						
Family variance	71702	71049	71962	17453	44219	101528
Residual variance	329333	328582	328852	20084	297954	363756
Additive variance	286806	284195	287848	69811	176876	406112
h^2	0.390	0.394	0.402	0.066	0.277	0.493

* SD: standard deviation; LL and UL: lower and upper limits of the confidence intervals (P=90%)

Table 3. Estimates of genetic gain (family selection) considering two selection intensities (20.12 and 10.06%) calculated according to the predicted Bayesian values for popcorn families

	Popping expansion		Grain yield	
	20.12%	10.06%	20.12%	10.06%
MBSF	1.30	1.67	149.8	195.7
NSF	34	17	34	17
Experimental mean	32.15	32.15	2121.4	2121.4
GG%	4.0	5.2	7.1	9.2

MBSF: Mean breeding values of the selected families; NSF: Number of selected families; % GG: genetic gain in percentage

Chains with up to 50 000 random samples with a discarding period (burn-in) of 5 000 were considered. This result is in agreement with Wolfinger and Kass (2000). In the IC algorithm, the base distribution for the variance components is considered accurate and, therefore, the MCMC samples are only rejected if they are not within the parameter space, which could occur, for example, if a variance component is estimated with a negative value. Thus, the IC algorithm is a special case of rejection sampling (Wolfinger and Kass 2000).

One advantage of the Bayesian procedures is the possibility of using the confidence limits (or credibility regions in the Bayesian context) as an alternative estimation of genetic parameters, which are obtained directly from posterior distribution (Wright et al. 2000, Mora and Serra 2014). According to Zech (2002), if there is *a priori* information, Bayesian inference can determine narrower confidence intervals than the conventional ones, and more powerful tests. However, it is possible to work with so-called non-informative prior distributions when appropriate. The great advantage of Bayesian inference in the context of mixed models is the relative ease of obtaining standard deviations for the elements of the main genotypic effect.

Figure 1 shows a graph with the breeding values of families, obtained via Bayesian prediction for popping expansion and grain yield, which were compared with each other. The straight line indicates the degree of association between breeding values of the 169 half-sib families. The Pearson correlation coefficient (calculated between breeding values of families for each trait) was negative, however statistically not different from zero ($r=-0.103$, $p=0.2112$), confirming a weak association between the predicted values of family effect of both characteristics. A genetic correlation not significantly different from zero would indicate that selecting families based on popping expansion would not have a significant effect on grain yield. This result is widely

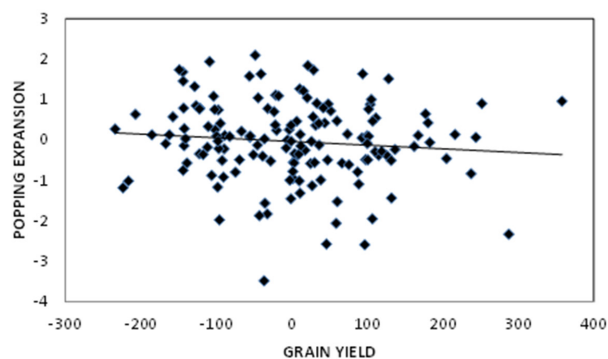


Figure 1. Scatterplot with breeding values of families, obtained by Bayesian prediction for popping expansion and grain yield, evaluated for 169 half-sib popcorn families (population EMU-C1).

reported in the literature, for example, Arnhold et al. (2009) and Broccoli and Burak (2004) confirmed that grain yield is inversely correlated (negative genetic association) with popping expansion. The Spearman correlation coefficient estimated by Munhoz et al. (2009) was not significant for 96 S3 popcorn families. Broccoli and Burak (2004) mentioned that the negative association between popping expansion and yield can affect the yield gains in simultaneous selection of a breeding population. From the genetic point of view, this phenomenon indicates that the two traits could be controlled by distinct genes.

There was significant variability among popcorn families of the EMU-C1 population, which allowed genetic gains for yield and expansion volume, however, the low genetic correlation between both traits would restrict genetic gains by simultaneous selection.

The genetic evaluation using the Bayesian approach allowed scientific inferences for the breeding population and was found to be a useful tool for the selection of popcorn families, corresponding to an equivalent procedure to the classical linear mixed models.

Abordagem Bayesiana na avaliação genética de famílias de meios-irmãos de milho pipoca via algoritmo de Cadeias de Independência

Resumo – O objetivo deste trabalho foi examinar parâmetros genéticos da capacidade de expansão e rendimento de milho pipoca em um ensaio de 169 famílias de meios-irmãos, utilizando a metodologia Bayesiana como alternativa de inferência científica. Foi utilizado o algoritmo de Cadeias de Independência com informações *a priori* informativas para os componentes de variância familiar e residual (distribuição *a priori* Inversa Gama). A capacidade de expansão mostrou ser moderadamente herdável com moda a posteriori de $h^2=0,34$ e intervalo de credibilidade de 0,22–0,44 (90% de probabilidade). O rendimento de grãos apresentou uma herdabilidade moderada (nível de família) de $h^2=0,4$ com intervalo de credibilidade de 0,28–0,49. A população em estudo apresentou variabilidade genética suficiente para os ciclos subsequentes de melhoramento, e a metodologia Bayesiana pode auxiliar no processo de seleção de famílias de milho pipoca.

Palavras-chave: Zea mays L., capacidade de expansão, rendimento de grãos, herdabilidade.

REFERENCES

- Arnhold E, Mora F, Silva RG, Good-God PIV and Rodovalho MA (2009) Evaluation of top-cross popcorn hybrids using mixed linear model methodology. **Chilean Journal of Agricultural Research** **69**: 46-53.
- Arriagada O, Mora F, Dellarossa JC, Ferreira MFS, Cervigni GDL and Schuster I (2012) Bayesian mapping of Quantitative Trait Loci (QTL) controlling soybean cyst nematode resistant. **Euphytica** **186**: 907-917.
- Broccoli AM and Burak R (2004) Effect of genotype x environment interactions in popcorn maize yield and grain quality. **Spanish Journal of Agricultural Research** **2**: 85-91.
- Brown H and Prescott R (2006) **Applied mixed models in medicine**. John Wiley & Sons, Chichester, 455p.
- Bueno Filho JS and Vencovsky R (2000) Alternativas de análise de ensaios em látice no melhoramento vegetal. **Pesquisa Agropecuária Brasileira** **35**: 259-269.
- Cané-Retamales C, Mora F, Vargas-Reeve F, Perret S and Contreras-Soto R (2011) Bayesian threshold analysis of breeding values, genetic correlation and heritability of flowering intensity in *Eucalyptus cladocalyx* under arid conditions. **Euphytica** **178**: 177-183.
- Corander J, Waldmann P and Sillanpaa MJ (2003) Bayesian analysis of genetic differentiation between populations. **Genetics** **163**: 367-374.
- Daros M, Amaral Júnior AT and Pereira MG (2002) Genetic gain for grain yield and popping expansion in full-sib recurrent selection in popcorn. **Crop Breeding and Applied Biotechnology** **2**: 339-344.
- Dhillon BS and Malhi NS (2005) Recurrent selection for intrapopulation improvement and cultivar development in maize (*Zea mays* L.) - an analysis. **Indian Journal of Genetics and Plant Breeding** **65**: 84-87.
- Doerksen TK, Kannenberg LW and Lee EA (2003) Effect of recurrent selection on combining ability in maize breeding populations. **Crop Science** **43**: 1652-1658.
- Lu HJ, Bernardo R and Ohm HW (2003) Mapping QTL for popping expansion volume in popcorn with simple sequence repeat markers. **Theoretical and applied genetics** **106**: 423-427.
- Miranda GV, Souza LV, Galvão JCC, Guimarães LJM, Melo AV and Dos Santos IC (2008) Genetic variability and heterotic groups of Brazilian popcorn populations. **Euphytica** **162**: 431-440.
- Mora F and Serra N (2014) Bayesian estimation of genetic parameters for growth, stem straightness, and survival in *Eucalyptus globulus* on an Andean Foothill site. **Tree Genetics and Genomes** **10**: 711-719.
- Moterle LM, Braccini AL, Scapim CA, Pinto RJB, Gonçalves LSA, Rodrigues R and Amaral Júnior AT (2012) Combining ability of popcorn lines for seed quality and agronomic traits. **Euphytica** **185**: 337-347.
- Munhoz REF, Prioli AJ, Amaral Júnior AT, Scapim CA and Simon GA (2009) Genetic distances between popcorn populations based on molecular markers and correlations with heterosis estimates made by diallel analysis of hybrids. **Genetics and Molecular Research** **8**: 951-962.
- Neath AA and Cavanaugh JE (2012) The Bayesian information criterion: background, derivation, and applications. **Computational Statistics** **4**: 199-203.
- Pacheco CAP, Santos MX, Cruz CD, Parentoni SN, Guimarães PEO, Gomes e Gama EE, Silva AE, Carvalho HWL and Vieira Júnior PA (2002) Inbreeding depression of 28 maize elite open pollinated varieties. **Genetics and Molecular Biology** **25**: 441-448.
- Peiris BL and Hallauer AR (2005) Comparison of half-sib and full-sib reciprocal recurrent selection and their modifications in simulated populations. **Maydica** **50**: 25-37.
- Pena GF, Amaral Júnior AT, Gonçalves LSA, Candido LS, Vittorazzi C, Ribeiro RM and Freitas Júnior SP (2012) Stability and adaptability of popcorn genotypes in the State of Rio de Janeiro, Brazil. **Genetics and Molecular Research** **11**: 3042-3050.
- Pereira MG and Amaral Júnior AT (2001) Estimation of genetic components in popcorn base on the nested design. **Crop Breeding and Applied Biotechnology** **1**: 3-10.
- Rangel RM, Amaral Júnior AT, Scapim CA, Freitas Júnior SP and Pereira MG (2008) Genetic parameters in parents and hybrids of circulant diallel in popcorn. **Genetics and Molecular Research** **7**: 1020-1030.
- Santos FS, Amaral Júnior AT, Freitas Júnior SP, Rangel RM, Scapim CA and Mora F (2008) Genetic gain prediction of the third recurrent selection cycle in a popcorn population. **Acta Scientiarum Agronomy** **30**: 651-655.
- Santos JF, Viana JMS, Vilarinho AA and Câmara TMM (2004) Efficiency of S2 progeny selection strategies in popcorn. **Crop Breeding and Applied Biotechnology** **4**: 183-191.
- SAS Institute Inc. (2007) **SAS 9.1.3 for Windows Microsoft**. SAS Institute Inc, Cary, 212p.
- Van Inghelandt D, Melchinger AE, Lebreton C and Stich B (2010) Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. **Theoretical and Applied Genetics** **120**: 1289-1299.
- Wolfinger RD and Kass RE (2000) Non-conjugate Bayesian analysis of variance component models. **Biometrics** **56**: 768-774.
- Wright DR, Stern HS and Berger PJ (2000) Comparing traditional and Bayesian analyses of selection experiments in animal breeding. **Journal of Agricultural, Biological, and Environmental Statistics** **5**: 240-256.
- Zech G (2002) Frequentist and Bayesian confidence intervals. **European Physical Journal** **12**: 1-81.