

BAYESIAN ANALYSIS OF THE GENETIC CONTROL OF SURVIVAL IN F₃ FAMILIES OF COMMON BEAN

Freddy Mora^{1*}, Maria C. Gonçalves-Vidigal¹, and Alexandra I. Santos¹

ABSTRACT

The objectives of this study were to examine the genetic control of survival in segregant families F₃ of the common bean (*Phaseolus vulgaris* L.) in southern Brazil during the 2004-2005 growing season, to identify useful genotypes for the breeding program of this crop, and to determine the genetic association between survival and weight of 100 seeds (production trait; P100). A Bayesian approach was used to predict breeding values and to estimate variance components. Survival was recorded as a binary response: dead plant or live plant during harvest. The total population consisted of 11 520 individual plants. The difference in the magnitude between the best and the worst families was as high as 22%, and varied from 57 to 73%. Survival was found to be highly heritable, with an *a posteriori* heritability mean and Bayesian credible interval: H² = 53% (43-65%). The genetic advance by direct selection achieved a value of 18%, considering a selection intensity of 25%. Survival was not correlated with P100 (Pearson = 0.099; Spearman = 0.074), indicating that selection for this trait alone would have little impact on production from a breeding viewpoint. Bayesian analysis, using the Gibbs algorithm, was useful in the genetic evaluation of common bean families based on a binary response variable.

Key words: genetic improvement, breeding, Gibbs algorithm, heritability, *Phaseolus vulgaris*.

INTRODUCTION

The common bean, an important source of protein in human consumption, is grown extensively in countries such as Brazil, China, the United States, India, and Mexico. This species is originally from America and belongs to a crop without a specific center of origin, and with multiple centers of domestication that were located in Central America, Mexico, and South America (Bascur and Tay, 2005). According to FAO (2005), Brazil is the world main producer (with 11% of total production), but is at the same time the main consumer with an average *per capita* of 20 kg per year. The common bean is grown in most of the Brazilian territory due to its extensive adaptability to edaphoclimatic conditions, being an important crop of the productive systems of small and medium-sized farmers who grow it along with other important species or as a single crop. Historically, the common bean has been socially and economically important in Brazil, especially in the State of Paraná, which has the greatest production in the country (Ceolin *et al.*, 2007).

Given the existence of a wide range of variability in the present populations of *P. vulgaris*, the greatest challenge for producers interested in improving productivity is to optimize alternatives to accumulate useful genes in genotypes or specific varieties (Bascur and Tay, 2005; Tay *et al.*, 2006). In Brazil, a key aspect to maximize productivity in bean crops has been to select those genotypes that respond well to specific environmental conditions, considering their productivity and survival

¹ Universidade Estadual de Maringá, Departamentos de Agronomia e Zootecnia, Av. Colombo 5790 Bloco 05, CEP: 87020-900, Maringá, Paraná, Brasil. E-mail: morapoblete@gmail.com; mvidigal@pop.com.br; alexandraines@hotmail.com *Corresponding author.
Received: 3 September 2007. Accepted: 17 January 2008.

per unit of area. However, even though there is varied research related to the study of the genetic control of the productivity components of the common bean (Kelly *et al.*, 1998; Ceolin *et al.*, 2007; Dawo *et al.*, 2007) studies about the genetic control of survival in this crop have been relatively undocumented in Brazil. As with any other agricultural crop, mortality in a common bean crop can have direct implications on the productivity and success of the crop. Varied studies in plant breeding have shown the importance of survival (or the number of plants effectively produced) in commercial production systems (Rojas and Riveros, 2001; Coyle *et al.*, 2006).

Heritability, or the degree of genetic control associated to some interest trait, is one of the most important parameters within the breeding context. Heritability indicates how much of the phenotypic variability has a genetic origin, and gives objective information for the genetic selection process (Falconer and MacKay, 1996). The precision of this parameter depends on an adequate estimation of the associated variance components. In this respect, the mixed linear model theory has been extensively applied in genetic evaluation, which has allowed the evaluation of genotypes in different selection programs (Mora and Scapim, 2007). In general, the prediction of the breeding value of genotypes (families, clones, origins, etc.) and the identification of those that are genetically superior, depend on the correct estimation of the genetic parameters.

In the context of plant breeding, the Bayesian methods can be used efficiently in the estimation of variance components, including heritability and genetic correlations, and in the prediction of breeding values (Mora and Arnold, 2006; Mora *et al.*, 2007). Bayesian analysis is based on the knowledge of the *a posteriori* distribution of the genetic parameters and makes the conformation of confidence intervals (or Bayesian credibility) viable for the estimation of such parameters (Gianola and Fernando, 1986). The methods referred to as Markov Chain Monte Carlo (MCMC) allow the efficient analysis of complex likelihood surfaces and the calculation of the *a posteriori* Bayesian distributions (Cadena-Meneses and Castillo-Morales, 2000; Mora *et al.*, 2007). For this reason, Walsh (2001) mentioned that in the next decades there will be a strong tendency to use such procedures, which will replace the methodology based solely on the likelihood function.

The objectives of this study were to evaluate the genetic control of survival in segregant families (F_3) of common

bean, evaluated in southern Brazil during the 2004-2005 growing season; to identify useful genotypes for the crop breeding program; and to determine the genetic relationship of survival with a production trait, aimed at indirect selection. The results of this study will improve knowledge about the genetic control of survival in bean, and will allow the efficient advancement of the breeding program initiated in southern Brazil.

MATERIALS AND METHODS

Genetic material origin

Twenty-four segregant families of common bean (F_3 generation) were used obtained from a breeding program of the Nucleus of Applied Agricultural Research (NUPAGRI), of the State of Paraná in southern Brazil. This program started in 2002 through diallelic crosses that used the following commercial cultivars: IPR-Juriti, FT-Soberano, IAC-Tybatã, BRSMG-Talismã, BRS-Campeiro and IPR Uirapuru. Subsequently, the best hybrid combinations were selected in greenhouse conditions up to the F_3 generation (Gonçalves-Vidigal *et al.*, 2005).

Field experiment

The 24 families were sown during the 2004-2005 season in the experimental station of the Santa Catarina Agricultural Research and Rural Extension Company Inc. (EPAGRI), located in Chapecó (27°07' S; 52°37' W); 679 m.a.s.l., southern Brazil.

A completely random block design was used, with four-blocks. The experiment consisted in four rows spaced at 0.5 m and 4.0 m long, with a total area of 8.0 m². The usable experimental area was 3.6 m² considering the two central rows. Plantation density was 15 plants per linear meter and was made up of 120 plants per plot.

Measurements

Survival data were collected of individual plants analysis during the 2004-2005 growing season. The total of the plants analyzed was 11 520. This trait was measured as a binary response, considering two possible biological events: dead or live plant in a specified plot. Subsequently, the survival variable response has a Bernoulli distribution, which results in two values: 0 and 1; in which 0 corresponds to the assigned value for a dead plant and 1 for a live plant.

In addition, with the objective of evaluating the association between survival and a production trait,

100 seeds (g) were weighed by randomly collecting 100 seeds from each experimental plot.

Genetic parameters estimations

Broad-sense heritability (H^2) or degree of genetic control for the survival trait was calculated using the following equation:

$$H^2 = \frac{\sigma_g^2}{\sigma_p^2} = \frac{(\sigma_a^2 + \sigma_d^2)}{(\sigma_a^2 + \sigma_d^2 + \sigma_e^2)}$$

where σ_g^2 and σ_p^2 are the genotypic and phenotypic variances, respectively, σ_a^2 and σ_d^2 are the additive and the dominance variances, respectively, and σ_e^2 refers to the residual variance.

Genetic advance or genetic gain was analyzed due to the direct selection, for the survival trait estimated through the breeding values of the family effect: additive genetic effects plus non additive (or the effects of dominance deviation).

Pearson and Spearman correlations were calculated for the family genetic survival values and the 100 seed weight to establish the association for indirect (Spearman) and genetic (Pearson) selection between both traits.

Statistical analysis

A Bayesian approach was used by applying the Gibbs algorithm, a variation of the Markov Chain Monte Carlo (MCMC) methods. In this statistical analysis the threshold model was considered (Van-Tassell *et al.*, 1998), weighing the additive genetic effects and the dominance deviations of the 24 families originating from self-pollination (generation F_3). The implemented model refers to:

$$\eta = X\beta + Za + Wd + \varepsilon$$

where η is the vector of a continuous random variable, normally distributed, not observable (unknown), and refers to the survival variable (interest trait); β is the vector with the block effect (considering a uniform *a priori* distribution); a and d are the vectors of additive and dominant effects of individual plants (*a priori* normal distribution); X , Z and W refer to the incidence matrixes of the respective effects, and ε is the vector of residual effects with a normal *a priori* distribution. It was considered that the variance components follow a prior inverted Chi-square distribution; univariate case (Van-Tassell and Van Vleck, 1996).

The MTGSAM program (Van-Tassell and Van Vleck, 1996) was used to obtain the subsequent variance component densities, broad-sense heritability and genetic parameter values (Van-Tassell *et al.*, 1998) in accordance with the procedure utilized by Mora *et al.* (2007).

The convergence of the chains generated by the Gibbs algorithm were evaluated with the Heidelberger and Welch method (1983) available in the R-CODA statistical program (Cowles and Carlin, 1995).

Bootstrap confidence intervals were calculated (Efron, 1979) with 95% probability for the Pearson and Spearman correlation coefficients, through the generation of 1000 random samples with substitution, using the SURVEYSELECT and CORR procedures of SAS (SAS Institute, 1996), in accordance with Mora and Arnhold (2006).

RESULTS AND DISCUSSION

Survival was moderate, reaching a mean value of 65%. The difference in magnitude between the best and the worst families (FFC-21 and FFC-3) was 22%, varying from 57 to 73%. The estimations of the variance components and broad-sense heritability are presented in Table 1. The *a posteriori* distributions of the interest parameters are shown in Figure 1. Genetic variance mean (and its additive and non additive components) estimations were different from the median and mode, indicating an asymmetry in the distribution.

Heidelberger and Welch (1983) statistic indicated a positive response to the convergence of Gibbs chain for each parameter, considering 20 000 samples with 5 000 ruled out (*burn-in*). Figure 1 shows the result of these iterations, considering three samples as the final sample interval.

Survival was a highly heritable trait, with an *a posteriori* heritability mean of $H^2 = 53\%$, and a credibility interval of 43-65%. Heritability mean value estimations were similar to the median and mode thus indicating a lack of asymmetry in the distribution (Figure 1). The specific values of the additive variance were superior to those of the dominance variance in approximately 27%. However, the interval estimation (95% probability) included the specific estimations of the variance due to the dominance deviations.

Genetic control of survival in breeding populations of common bean, whether due to additive effects or

Table 1. Estimates of the additive ($\hat{\sigma}_a^2$), dominance ($\hat{\sigma}_d^2$), genotypic ($\hat{\sigma}_g^2$), residual ($\hat{\sigma}_e^2$) and phenotypic ($\hat{\sigma}_p^2$) variances, and broad-sense heritability (\hat{H}^2) for survival in 24 segregant families of common bean evaluated in southern Brazil during the 2004-2005 growing season.

| Parameter | Mean | Median | Mode | Credibility interval | |
|--------------------|-------|--------|-------|----------------------|-------|
| | | | | Lower | Upper |
| $\hat{\sigma}_a^2$ | 0.173 | 0.160 | 0.142 | 0.100 | 0.290 |
| $\hat{\sigma}_d^2$ | 0.127 | 0.120 | 0.107 | 0.070 | 0.210 |
| $\hat{\sigma}_g^2$ | 0.300 | 0.290 | 0.268 | 0.190 | 0.460 |
| $\hat{\sigma}_e^2$ | 0.252 | 0.250 | 0.250 | 0.250 | 0.260 |
| $\hat{\sigma}_p^2$ | 0.552 | 0.540 | 0.518 | 0.440 | 0.710 |
| \hat{H}^2 | 0.534 | 0.536 | 0.531 | 0.432 | 0.648 |

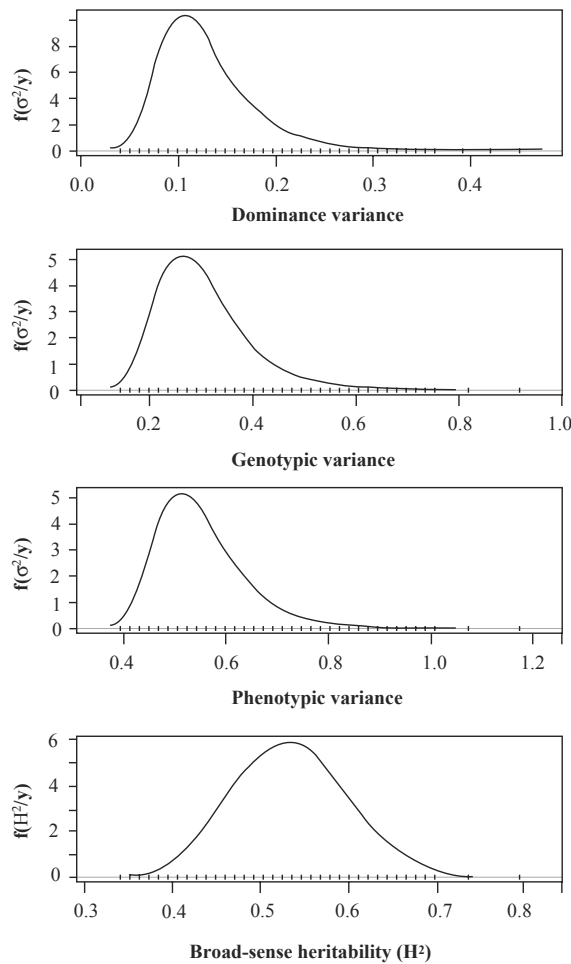


Figure 1. Subsequent distributions of the variance components and broad-sense heritability for the survival trait in common bean families evaluated in southern Brazil, during the 2004-2005 growing season.

$f(\sigma^2/y)$ and $f(H^2/y)$ are the density functions of the variances and heritability, respectively.

dominance deviations, is relatively undocumented in Brazil. Studies have mainly focused their efforts on the knowledge of production traits. For example, Barelli *et al.* (2000) found a greater action of the non additive effects for the traits production of grains and mean number of seeds per plant. A greater action of the additive genetic effects was evidenced in the traits mean weight of seeds and total number of pods per plant. The understanding of the relative contributions of the genetic and environmental components in variance control of interest traits is a crucial aspect of breeding programs (Anbessa *et al.*, 2006).

The significant genetic variability found in this breeding population for the survival trait indicates a favorable situation for the direct selection of endogamous families of the common bean. The quantification of genetic variability is also important in the common

bean breeding programs, since it reveals the genetic structure of the breeding populations (Ceolin *et al.*, 2007). The strategy adopted in the breeding program allows advances to be made in the respective selection cycles, including the survival trait. Ceolin *et al.* (2007) stated that the breeding programs using hybridizations were usually applied in common bean crops, with the objective of obtaining segregant populations that give high production per unit of area and other interest traits of broad genetic variability.

With the results of the genetic evaluation of common bean families, it was possible to select families that responded better in their survival value. This result is shown in Table 2, with the values for the six best families. Genetic advancement for direct selection reached a value of 18% considering a selection intensity of 25%. The standard deviations were similar for each breeding value of the

Table 2. Predicting breeding values for the best six common bean families based on plant survival obtained using a Bayesian approach (Gibbs algorithm).

| Family | Mean | Median | Mode | Standard deviation | Credibility interval | |
|----------|--------|--------|-------|--------------------|----------------------|-------------|
| | | | | | Lower limit | Upper limit |
| FFC-21 | 0.261 | 0.260 | 0.252 | 0.096 | 0.139 | 0.384 |
| FFC-13 | 0.256 | 0.255 | 0.267 | 0.096 | 0.135 | 0.378 |
| FFC-11 | 0.147 | 0.146 | 0.147 | 0.091 | 0.030 | 0.264 |
| FFC-17 | 0.144 | 0.143 | 0.128 | 0.092 | 0.028 | 0.261 |
| FFC-22 | 0.138 | 0.136 | 0.125 | 0.092 | 0.021 | 0.255 |
| FFC-15 | 0.136 | 0.135 | 0.124 | 0.092 | 0.018 | 0.254 |
| GG (25%) | 18.03% | | | | | |

GG (25%): genetic gain considering a selection intensity of 25%.

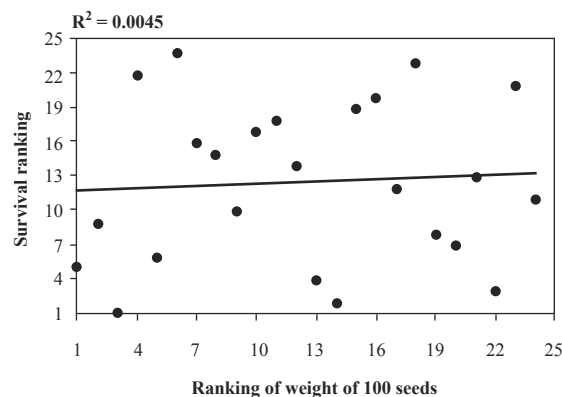


Figure 2. Scatter plot matrix ranking common bean families for survival versus weight of 100 seeds.

families that were evaluated. The Bayesian credibility intervals (95%) did not include zero or negative values in their lower limit, indicating that it is possible to select these genotypes with a significant degree of confidence. According to Wright *et al.* (2000), Bayesian analysis is important in genetic evaluation since it considers the existing variability in the model parameters and the prediction values of the genetic effects. These can be characterized not only with the mode, median or the mean of the *a posteriori* distribution, but also through the credibility intervals (or confidence intervals in the case of traditional evaluation).

Figure 2 shows a scatter plot matrix related to the ranking of the 24 common bean families according to survival and weight of 100 seeds. The line drawn in the graphic gives a measurement of the relation between the two variables with a value of 0.0045 for R^2 . Table 3 shows the results of the Spearman (0.074) and Pearson (0.099) correlation coefficients, including bootstrap confidence intervals. The correlation coefficients were low and not significantly different to zero ($p > 0.05$) in accordance with the Student's t-test. The bootstrap confidence intervals coincided with the significance value since these included the zero value correlation.

The weight of 100 seeds is considered an important characteristic within the common bean breeding programs. Singh *et al.* (1991) showed that the selection for weight of the seeds not only will give a greater weight for each 100 seeds, but also more seeds per planted area. The results in this study indicated a non significant genetic association between the weight of 100 seeds and survival (Pearson correlation between the breeding values), and

the Spearman correlation coefficients indicated that the selection based on plant survival will not have any significant impact on the indirect selection based on the weight of the 100 seeds. This is a relevant aspect within the present program because in a breeding program conducted in the State of Paraná, Ceolin *et al.* (2007) found that the trait number of pods per plant, number of seeds per pod, and weight of 100 seeds are important components in order to obtain high productivity. Dawo *et al.* (2007) have determined that the weight of 100 seeds had a response that was not correlated with other production components, where plants that showed a high productivity also had a smaller seed size.

CONCLUSIONS

The genetic control of survival in this study was high and allowed the identification of useful genotypes for this breeding program, and the selection of segregant families of the common bean by considering its genotypic value with with a reasonable prediction of genetic gain.

From the breeding point of view, selection based on survival can have little impact on some production traits and vice versa. This result is relevant and should be considered within this breeding program conducted in southern Brazil.

Bayesian analysis, through the Gibbs algorithm, could be an important tool in the evaluation of common bean genotypes, based on plant survival, and analyzing the response through a binary type variable.

Table 3. Pearson correlation coefficients (between breeding values) and Spearman ranking correlation coefficients calculated for survival and weight of 100 seeds traits in 24 segregant families of common bean evaluated in southern Brazil.

| Estimated value | Correlation coefficients | |
|---|--------------------------|---------|
| | Spearman | Pearson |
| Mean | 0.074 | 0.099 |
| Standard deviation | 0.230 | 0.221 |
| Lower limit (<i>bootstrap</i>) | -0.332 | -0.314 |
| Upper limit (<i>bootstrap</i>) | 0.442 | 0.413 |
| Statistical significance (Student t-test) | 0.733 | 0.647 |

RESUMEN

Análisis bayesiano del control genético de la supervivencia en familias F₃ de poroto. Freddy Mora^{1*}, Maria C. Gonçalves-Vidigal¹, y Alexandra I. Santos¹. El presente estudio tuvo como objetivos examinar el control genético de la supervivencia en familias segregantes (F₃) de poroto (*Phaseolus vulgaris* L.) en el sur de Brasil, durante la temporada agrícola 2004-2005, identificar genotipos útiles para el programa de mejoramiento del cultivo, y determinar la relación genética de la supervivencia con el peso de 100 semillas (característica de producción; P100). Se utilizó un análisis Bayesiano para la predicción de valores genéticos y la estimación de componentes de varianza. Se midió la supervivencia como una característica binaria (planta viva o muerta durante el período de cosecha). La población total analizada consistió en 11 520 plantas individuales.

La diferencia en magnitud de la supervivencia entre la mejor y la peor familia fue de 22%, y varió de 57 a 73%. La supervivencia fue una característica altamente heredable, con un valor de heredabilidad promedio *a posteriori* e intervalo de credibilidad Bayesiano: H² = 53%: (43-65%). El avance genético por selección directa alcanzó un valor de 18%, considerando una intensidad de selección de 25%. La supervivencia no fue correlacionada con P100 (Pearson = 0,099; Spearman = 0,074), indicando que la selección para la supervivencia tendría poco impacto sobre la producción (desde el punto de vista del mejoramiento genético). El análisis Bayesiano, a través del algoritmo de Gibbs, fue útil en la evaluación genética de familias de poroto, basado en una característica binaria.

Palabras clave: mejoramiento genético, algoritmo de Gibbs, heredabilidad, *Phaseolus vulgaris*.

LITERATURE CITED

- Anbessa, Y., T. Warkentin, A. Vandenberg, and R. Ball. 2006. Inheritance of time to flowering in chickpea in a short-season temperate environment. *J. Heredity* 97:55-61.
- Barelli, M.A.A., M.C. Gonçalves-Vidigal, A.T. Amaral-Júnior, P.S. Vidigal-Filho, C.A. Scapim, and E. Sagrilo. 2000. Diallel analysis for grain yield and yield components in *Phaseolus vulgaris* L. *Acta Sci., Agron.* 22:883-887.
- Bascur, G.B., y J.U. Tay. 2005. Colecta, caracterización y utilización de la variabilidad genética en germoplasma chileno de poroto (*Phaseolus vulgaris* L.). *Agric. Téc. (Chile)* 65:135-146.
- Cadena-Meneses, J.A., and A. Castillo-Morales. 2000. Variance component estimation, a comparison supported by simulation. *Agrociencia* 34:343-352.
- Ceolin, A.C.G., M.C. Gonçalves-Vidigal, and P.S. Vidigal-Filho. 2007. Genetic divergence of the common bean (*Phaseolus vulgaris* L.) group Carioca using morpho-agronomic traits by multivariate analysis. *Hereditas* 144:1-9.
- Cowles, M.K., and B.P. Carlin. 1995. Markov Chain Monte Carlo diagnostics: A comparative review. *J. Am. Stat. Soc.* 91:883-904.
- Coyle, D.R., M.D. Coleman, J.A. Durant, and L.A. Newman. 2006. Survival and growth of 31 *Populus* clones in South Carolina. *Biomass Bioenerg.* 30:750-758.
- Dawo, M.I., F.E. Sanders, and D.J. Pilbeam. 2007. Yield, yield components and plant architecture in the F₃ generation of common bean (*Phaseolus vulgaris* L.) derived from a cross between the determinate cultivar 'Prelude' and an indeterminate landrace. *Euphytica* 156:77-87.
- Efron, B. 1979. Bootstrap methods: another look at the jackknife. *Ann. Stat.* 7:1-26.
- Falconer, D.S., and T.F.C. MacKay. 1996. Introduction to quantitative genetics. 4th ed. Longman, New York, USA.
- FAO. 2005. Base de datos FAOSTAT. Disponible en <http://faostat.fao.org> (Leído 3 septiembre de 2005).
- Gianola, D., and R.L. Fernando. 1986. Bayesian methods in animal breeding theory. *J. Animal Sci.* 63:217-244.
- Gonçalves-Vidigal, M.C., L. Silvério, and P.S. Vidigal Filho. 2005. Diallel analysis of the combining ability of common bean (*Phaseolus vulgaris* L.) cultivars. p. 184-185. *Bean Improvement Cooperative Annual Report 48*. Michigan State University, East Lansing, Michigan, USA.
- Heidelberger, P., and P.D. Welch. 1983. Simulation run length control in the presence of an initial transient. *Oper. Res.* 31:1109-1114.
- Kelly, J., J.M. Kolkman, and K. Schneider. 1998. Breeding for yield in dry bean (*Phaseolus vulgaris* L.). *Euphytica* 102:343-356.
- Mora, F., and E. Arnhold. 2006. Application of the Bayesian inference and mixed linear model method to maize breeding. *Cien. Inv. Agr.* 33:185-190.

- Mora, F., S. Perret, C.A. Scapim, E.N. Martins, and M.P. Molina. 2007. Source-dependent blooming variability of *Eucalyptus cladocalyx* in the Region of Coquimbo, Chile. *Cien. Inv. Agr.* 34:99-106.
- Mora, F., y C.A. Scapim. 2007. Predicción de valores genéticos del efecto de poblaciones de maíz evaluadas en Brasil y Paraguay. *Agric. Téc. (Chile)* 67:139-146.
- Rojas, L.P., y F.B. Riveros. 2001. Efecto del método y edad de las plántulas sobre el prendimiento y desarrollo de injertos en melón (*Cucumis melo*). *Agric. Téc. (Chile)* 61:262-274.
- SAS Institute. 1996. Statistical analysis system: User's guide. 956 p. SAS Institute, Cary, North Carolina, USA.
- Singh, S.P., H. Teran, A. Molina, and J.A. Gutierrez. 1991. Genetics of seed yield and its components in common beans (*Phaseolus vulgaris* L.) of Andean origin. *Plant Breed.* 107:254-257.
- Tay, J., A. France, y A. Pedreros. 2006. Astro-INIA: primera variedad de poroto granado tipo coscorrón de hábito de crecimiento determinado arbustivo. *Agric. Téc. (Chile)* 66:416-419.
- Van Tassell, C.P., and L.D. Van Vleck. 1996. Multiple-trait Gibbs sampler for animal models: flexible programs for Bayesian and likelihood-based (co)variance component inference. *J. Anim. Sci.* 74:2586-2597.
- Van Tassell, C.P., L.D. Van Vleck, and K.E. Gregory. 1998. Bayesian analysis of twinning and ovulation rates using a multiple-trait threshold model and Gibbs sampling. *J. Anim. Sci.* 76:2048-2061.
- Walsh, B. 2001. Quantitative genetics in the age of genomics. *Theor. Popul. Biol.* 59:175-184.
- Wright, D.R., H.S. Stern, and P.J. Berger. 2000. Comparing traditional and Bayesian analyses of selection experiments in animal breeding. *J. Agric. Biol. Environ. Stat.* 5:240-256.