Published July 28, 2016



# Mixed Modeling of Yield Components and Brown Rust Resistance in Sugarcane Families

Thiago W. A. Balsalobre, Melina C. Mancini, Guilherme da S. Pereira, Carina O. Anoni, Fernanda Z. Barreto, Hermann P. Hoffmann, Anete P. de Souza, Antonio A. F. Garcia, and Monalisa S. Carneiro*

## ABSTRACT

Sugarcane (*Saccharum* spp.) is a complex autopolyploid with high potential for biomass production that can be converted into sugar and ethanol. Genetic improvement is extremely important to generate more productive and resistant cultivars. Populations of improved sugarcane are generally evaluated for several traits simultaneously and in multi-environment trials. In this study, we evaluated two full-sib families of sugarcane (SR1 and SR2) at two locations and 3 yr for stalk diameter, stalk height, stalk number, stalk weight, soluble solid content (Brix), sucrose content of cane, sucrose content of juice, fiber, cane yield, sucrose yield, and resistance to brown rust (*Puccinia melanocephala*). Using a mixed model approach, we included appropriate variance–covariance (VCOV) structures for modeling heterogeneity and correlation of genetic effects and nongenetic residual effects. The genotypic correlations between traits were calculated across the adjusted means as the standard Pearson product-moment coefficient. Through the VCOV structures estimated for each trait, in general, the heritabilities ranged from 0.78 to 0.94. Additionally, we detected 17 and 12 significant genotypic correlations between the evaluated traits for SR1 and SR2, respectively. The analysis of the severity data for brown rust revealed that 66 and 32% of the full-sib genotypes in SR1 and SR2, respectively, had at least 90% probability of being resistant.

## Core Ideas

- A linear mixed model is efficient in production data analysis of sugarcane.
- In general, the broad-sense heritability of the traits were high, ranging from 0.78 to 0.94.
- A generalized linear mixed model can be applied in brown rust analysis of sugarcane.
- Multi-environment trials were applied to the genetic improvement of sugarcane.

Published in Agron. J. 108:1824–1837 (2016)
doi:10.2134/agronj2015.0430
Received 4 Sept. 2015
Accepted 13 Feb. 2016
Supplemental material available online
Available freely online through the author-supported open access option



Sugarcane is a complex autopolyploid plant and one of the most highly consumed crops in the world (FAO, 2014). Sugarcane produces high yields and has been demonstrated to efficiently use resources (i.e., land, water, N, and energy) (de Vries et al., 2010; Eksteen et al., 2014; Gerbens-Leenes et al., 2009; Waclawovsky et al., 2010). A sustainable energy future depends on the increased use of renewable energy. Bioethanol produced from sugarcane through both first- and second-generation technologies is a good example of a renewable energy source that can contribute to reducing the environmental effects of fossil fuels (Goldemberg, 2007). First-generation production is an alternative and economically viable technique that is widely used in Brazil, whereas second-generation production is not as well optimized as first-generation production and is thus less feasible. However, the substantial potential of the generation of bioethanol from lignocellulosic wastes has encouraged studies that seek to improve the process and make it an integral component of the units that already produce first-generation bioethanol (Macrelli et al., 2014; Naik et al., 2010; Saini et al., 2015).

Sugarcane breeding programs have focused on releasing new cultivars with agronomic traits that suit the demand of the sugarcane industry. However, the genetic complexity of sugarcane has hindered the genetic improvement of this crop. Commercial sugarcane cultivars are the result of interspecific crosses between both domesticated *Saccharum officinarum* L. ($2n = 80$) and wild *S. spontaneum* ($2n = 40–120$) species, followed by several

T.W.A. Balsalobre, F.Z. Barreto, H.P. Hoffmann, and M.S. Carneiro, Center for Agricultural Sciences, Dep. of Biotechnology and Plant and Animal Production, Federal Univ. of São Carlos, Highway Anhanguera Km 174, Araras, CEP 13600-970, SP, Brazil; M.C. Mancini and A.P. Souza, Dep. of Plant Biology, Biology Institute, Univ. of Campinas, 255 Monteiro Lobato Ave, Campinas, CEP 13083-862, SP, Brazil, and Center for Molecular Biology and Genetic Engineering, 400 Candido Rondon Ave, Campinas, CEP 13083-875, SP, Brazil; and G.S. Pereira, C.O. Anoni, and A.A.F. Garcia, Dep. of Genetics, Luiz de Queiroz College of Agriculture, Univ. of São Paulo, 11 Pádua Dias Ave, Piracicaba, CEP 13418-900, SP, Brazil. T.W.A. Balsalobre and M.C. Mancini contributed equally to this work.
*Corresponding author (monalisa@cca.ufscar.br).

**Abbreviations:** AIC, Akaike information criterion; BIC, Bayesian information criterion; BLUP, best linear unbiased prediction; FIB, fiber; GEI, genotype × environment interaction; GLMM, generalized linear mixed model; LMM, linear mixed model; METs, multi-environment trials; POL%C, sucrose content of cane in percentage; POL%J, sucrose content of juice in percentage; REML, restricted maximum likelihood; SD, stalk diameter; SH, stalk height; SN, stalk number; SR1, SP80-3280 × RB835486 full-sib family of sugarcane 1; SR2, SP81-3250 × RB925345 full-sib family of sugarcane 2; SW, stalk weight; TCH, tonnes of cane per hectare; TPH, tonnes of sucrose per hectare; VCOV, variance–covariance.

backcrosses with *S. officinarum* (Ha et al., 1999; Irvine, 1999). The polyploid genome complexity of commercial cultivars can be attributed to the following factors: the chromosome number, ranging from 100 to 130 (D'Hont et al., 1998; Irvine, 1999); the genome size of approximately 10 Gb (D'Hont, 2005; D'Hont and Glaszmann, 2001; Piperidis et al., 2010); and the aneuploidy condition, with a variable number of chromosomes in each hom(e)ology group (Grivet and Arruda, 2002).

The primary purpose of a genetic breeding program is to improve yields (Cox et al., 1994), which is possible due to the accumulation of knowledge of plant breeding. Worldwide data indicate that sugarcane yield has increased by 41% during the last 50 yr (Gouy et al., 2013, 2015). Recently, small sugar production increases of approximately 1 to 1.5% per year have been obtained. The average yield of sugarcane in Brazil, the world's largest sugarcane producer, is approximately 74 t $ha^{-1}$; however, the theoretical production potential is approximately 400 t $ha^{-1}$ (Dal-Bianco et al., 2012; Matsuoka et al., 2014; Waclawovsky et al., 2010). Waclawovsky et al. (2010) showed that, in Brazil, the commercial maximum yield (large land areas) was 260 t $ha^{-1}$ and an experimental maximum (individual trials on smaller land areas) was 299 t $ha^{-1}$. These high yields were obtained under irrigation in an area with low precipitation and low cloudiness, hence higher solar radiation than is observed in most sugarcane-producing areas of Brazil. Thus, to achieve high yields, it is necessary to consider the use of agricultural practices that involve additional costs, such as irrigation and fertilization. The genetic components, the environment, and the relationship between the traits of interest are essential to developing breeding strategies. The genetic component of the phenotypic variance in crop traits is most commonly studied, as reflected in the high rate of scientific and technological progress in plant breeding (Edwards et al., 2013; Sadras et al., 2013). In sugarcane breeding programs, many genetic clones are commonly evaluated during several harvests in multi-environment trials (METs). Genotype × environment interaction (GEI) is broadly considered to be the variation in the relative performance of genotypes across environments (Ramburan et al., 2012) and is an important feature when selecting superior cultivars (Jackson et al., 1991; Ramburan, 2014). Furthermore, experiments commonly involve the simultaneous evaluation of several traits because superior cultivars should concentrate favorable alleles for yield, resistance to diseases (e.g., brown rust and smut), pests and abiotic stresses, and agronomic traits, among other factors (Welham et al., 2010).

Brown rust, a disease caused by *Puccinia melanocephala* H. & P. Sydow, affects sugarcane and is present in many production areas worldwide (Asnaghi et al., 2004; Hoy and Hollier, 2009; Ryan and Egan, 1989). Negative impacts of brown rust on sugarcane yields have been reported (McFarlane et al., 2006; Raid and Comstock, 2000). Field losses >50% are associated with brown rust, depending on the cultivar susceptibility and growing conditions (Hoy and Hollier, 2009; Purdy et al., 1983). Therefore, the release of cultivars that are resistant to brown rust is very important and is the most efficient form of disease control. The genetic inheritance of brown rust in sugarcane has been broadly studied, and some researchers have claimed that this resistance trait is controlled by one or a few genes (Asnaghi et al., 2004; Costet et al., 2012; Daugrois et al., 1996; Garsmeur et al., 2011; Glynn et al., 2013; Hogarth et al., 1993; Le Cunff et al., 2008; Parco et al., 2014; Raboin et al., 2006; Racedo et al., 2013; Ramdoyal et al., 2000; Sordi et al., 1988). In contrast to that shown for brown rust, almost every trait that is related to sugarcane production exhibits quantitative variations. For example, sugar yield components depend on a combination of stalk diameter, stalk height, stalk number, stalk weight, and BRIX (Hogarth, 1971). These traits have a complex relationship, which complicates the selection of superior cultivars.

Because of the combination of factors for a sugarcane cultivar ideotype, the efficient estimation of genetic parameters is dependent on the choice of experimental design and statistical models that are appropriate for the response pattern of the evaluated variables (Sadras et al., 2013). In traditional analysis of variance models, all of the effects are considered fixed (except for the residual error), limiting the potential of the analysis. For sugarcane, the data from METs are modeled by assuming variance homogeneity and absence of genetic correlation between the harvest and location for estimating breeding values (Balzarini, 2002). In contrast, linear mixed models (LMMs) (Henderson, 1984) have advantages over fixed linear models for analyzing METs. Specifically, LMMs have the ability to consider variables as random rather than fixed and to use different variance–covariance (VCOV) structures for random effects to investigate the presence of heteroscedasticity and correlations. This approach allows the analysis of unbalanced data (Pastina et al., 2012; Smith et al., 2005) in addition to using more realistic models for residual variation (incomplete blocks and spatial correlation) and assuming sets of effects (e.g., genotypes) as random (Piepho et al., 2008; Smith et al., 2005). The estimation of variance component parameters is obtained preferably by restricted maximum likelihood (REML), and genotype effects may be obtained either by best linear unbiased estimation or best linear unbiased prediction (BLUP), depending on whether genotypes are considered fixed or random factors, respectively (Piepho et al., 2008; Smith et al., 2005). One major property of BLUP is shrinkage toward the mean, which anticipates regression of progeny to the mean and increases the accuracy of prediction of breeding and genotypic values (Piepho et al., 2008). In addition, BLUP maximizes the correlation of true genotypic values and predicted genotypic values, which is the primary aim of breeders (Searle et al., 1992).

The mixed model approach is more realistic, with a higher predictive ability based on modeling the VCOV matrices. This approach also means a great change in the analysis of breeding experiments because genotype observations may be grouped by levels of grouping factors generated from the experimental design, such as the harvest year and location (Pastina et al., 2012). The application of a mixed model approach is becoming increasingly popular in plant breeding, particularly in research involving the prediction of breeding values combined with genomic data (Beaulieu et al., 2014; Bevan and Uauy, 2013; Burgueño et al., 2012; Crossa et al., 2013; Muir, 2007; Wolc et al., 2011; Zhang et al., 2010). For sugarcane, the use of linear mixed models to map quantitative trait loci (Pastina et al., 2012) and genomic selection (Gouy et al., 2013) represents progress in crop improvement. However, the basis of genetic inheritance of all traits of economic interest and the genetic expression of related genes in cultivars that are used as parents by breeding programs should be better understood so that we

can efficiently achieve higher gains with the selection process and apply genomic selection in sugarcane in the future.

The objectives of this study were to: (i) evaluate the morphological and technological traits in full-sib families of sugarcane that were established in two different locations during three harvest years using LMMs, (ii) estimate the heritability and genotypic correlation coefficients using VCOV matrices, and (iii) assess the severity data of brown rust disease using a generalized linear mixed model (GLMM).

## MATERIALS AND METHODS

### Plant Material

The populations were developed by the Genetic Breeding Program of Sugarcane from the Universidade Federal de São Carlos (UFSCar), which is an integral component of the Rede Interuniversitária para o Desenvolvimento do Setor Sucroalcooleiro (RIDESA). Bi-parental crosses between Brazilian commercial cultivars produced two full-sib families from which the phenotypic data were collected. The first family, named SR1, consisted of 153 full-sib genotypes that were derived from a cross between SP80-3280 (female parent) and RB835486 (male parent). The SP80-3280 parent (SP71-1088 × H57-5028) was sequenced by the Sugarcane Expressed Sequence Tags (SUCEST) Project and has higher productivity, sucrose content, fiber content, and resistance to smut (*Sporisoriun scitamineum* Sydow) and brown rust; RB835486 (L60-14 × ?) has a higher sucrose content and is susceptible to smut and brown rust. The second family, named SR2, consisted of 240 full-sib genotypes that were derived from a cross between SP81-3250 (female parent) and RB925345 (male parent). The SP81-3250 parent (CP70-1547 × SP71-1279) is resistant to brown rust, while RB925345 (H59-1966 × ?) is susceptible to brown rust; both parents have high productivity, sucrose content, and fiber content.

### Phenotypic Data

Two independent experiments, one for each family, were planted in 2010 at two locations (Araras and Ipaussu) in the state of São Paulo, Brazil. The Araras site was located at 22°21′25″ S, 47°23′3″ W, at an altitude of 611 m; the soil of the site was a Typic Eutroferric Red Latosol. The Ipaussu site was located at 23°8′44″ S, 49°23′23″ W, at an altitude of 477 m; the soil of the site was a Dark Red Latosol. Historically, Ipaussu is a location with a high natural incidence of brown rust.

At each location, the experimental design consisted of an augmented randomized incomplete block design, which was fully replicated three times. Each incomplete block included 30 genotypes: 27 full-sib genotypes plus three checks (SP80-3280, RB835486, and RB867515 for the SR1 experiment and SP81-3250, RB925345, and RB867515 for the SR2 experiment). The positions of the 30 genotypes in each incomplete block were fully randomized within family. Trial plots consisted of three and two rows in Ipaussu and Araras, respectively. The rows were 3 m long and spaced 1.5 m apart for both locations.

Sugarcane families were evaluated for 10 yield components: soluble solid content (BRIX, in °Brix), sucrose content of the cane (POL%C, in %), sucrose content of the juice (POL%J, in %), fiber (FIB, in %), stalk diameter (SD, in mm), stalk height (SH, in cm), stalk number (SN), stalk weight (SW, in kg), cane yield (TCH, in t ha$^{-1}$), and sucrose yield (TPH, in t ha$^{-1}$). Considering each experiment as multi-harvest-location, trials were harvested when the plants were approximately 12 mo of age during 2011, 2012, and 2013 at both locations. A 10-stalk sample was taken for analysis of the BRIX, POL%C, POL%J, and FIB, and two replicates at each location were evaluated. For analysis of the SD, SH, SN, SW, and TCH, all three replicates at each location were evaluated. A cluster of 10 stalks per plot was weighed and used to measure SH and SD. The weights of the two clusters of 10 stalks from each plot were added to the total weight of the plot (SW) to estimate the TCH, which was calculated as the product of the stalk weight of a linear meter (6667 linear meters compose 1 ha with a spacing of 1.5 m). The number of stalks was estimated by directly counting the tillers in the field. The values of TCH and POL%C were used to estimate TPH from the product of TCH and POL%C divided by 100. The yield component data were evaluated at a plant age of 12 mo according to the methodology described by the State of Sao Paulo Sugarcane, Sugar and Alcohol Growers Council (2006). During the experiments, Family SR1 suffered unforeseen events (the experimental field was attacked by capybaras [*Hydrochoerus hydrochaeris*]) that rendered the collection of yield component data of the second harvest at Araras unviable. Likewise, for both families (SR1 and SR2), the collection of BRIX, POL%C, POL%J, and FIB data for the second harvest at Ipaussu was not possible because there was an accidental fire in the experimental field that prevented the collection of samples for these analyses.

The resistance to brown rust was evaluated in the field through natural infestation. Evaluations of the incidence of brown rust were performed on both full-sib genotypes and checks on 6-mo-old plants in February during 2011, 2012, and 2013. This period is considered the most favorable epidemiological season for the occurrence of brown rust, considering both temperature and humidity conditions favoring the incidence of infection as the plants age. Brown rust resistance was scored in 3+ on each plot on a 1 (most resistant) to 9 (most susceptible) diagrammatic scale according to Tai et al. (1981) and Amorim et al. (1987). This scale is based on a visual assessment of the disease symptoms. A score of 1 indicates the absence of sporulating pustules (uredospores) and resistant plants. A score of 2 indicates very rare sporulating pustules. For grades from 2 to 9, the density of sporulating pustules increases and indicates susceptible plants. Five plants per plot were evaluated.

### Analysis of Yield Components and Brown Rust Resistance Data

A multi-harvest-location mixed model produced the joint-adjusted means to obtain genotypic correlations among traits. The analyses were conducted for each trait for both populations using GenStat 16 (Payne et al., 2009) based on the REML and the following linear model:

$$y_{ijknm} = \mu + l_n + h_m + r_{k(nm)} + b_{j(knm)} + t_{inm} + \varepsilon_{ijknm} \quad [1]$$

where $y_{ijkmn}$ is the phenotype of the $i$th genotype in the $k$th replicate and the $j$th incomplete block at the $n$th location and $m$th harvest; $\mu$ is the overall mean; $l_n$ is the fixed effect of the $n$th location ($n = 1, N = 2$); $h_m$ is the fixed effect of the $m$th harvest ($m = 1, ..., M$; $M = 2$ or 3 depending on the location);

$r_{k(nm)}$ is the fixed effect of the $k$th replicate ($k = 1, ..., K$; $K = 2$ or 3 depending on the trait) at the $n$th location and $m$th harvest; $b_{j(knm)}$ is the random effect of the $j$th block ($j = 1, ..., J$; $J = 9$) in the $k$th replicate at the $n$th location and $m$th harvest; $t_{inm}$ is the random effect of the $i$th genotype ($i = 1, ..., I$; $I = 156$ for SR1 and $I = 243$ for SR2) at the $n$th location and $m$th harvest; and $\varepsilon_{ijknm}$ is the random residual error. The genotypes ($t_{inm}$) were separated into two groups, in which $g_{inm}$ was a random genetic effect of the $i$th full-sib genotype ($i = 1, ..., I_g$; $I_g = 153$ for SR1 or 240 for SR2) at the $n$th location and $m$th harvest, and $c_{inm}$ was the fixed effect of the $i$th check ($i = I_g + 1, ..., I_g + I_c$; $I_c = 3$) at the $n$th location and $m$th harvest. For the genotypes, the vector $\mathbf{g} = (g_{111}, ..., g_{IMN})'$ had a multivariate normal distribution with zero mean vector and genetic VCOV matrix $\mathbf{G} = \mathbf{G}_P \otimes \mathbf{I}_{Ig}$, i.e., $\mathbf{g} \sim N(0,\mathbf{G})$, where $P$ is the number of location–harvest combinations and $\otimes$ represents the Kronecker direct product of both the genetic $\mathbf{G}_P$ and identity $\mathbf{I}_{Ig}$ matrices with the respective dimensions of $P \times P$ and $I_g \times I_g$. Several structures for the $\mathbf{G}_P$ matrix (Table 1) were examined and compared via Akaike (AIC; Akaike, 1974) and Bayesian (BIC; Schwarz, 1978) information criteria (Pastina et al., 2012). Abbreviations of each structure presented in Table 1 will be used hereafter in the text. Models 1 to 6 used a location–harvest factorial combination for different environments (E), i.e., $\mathbf{G}_P = \mathbf{G}^{E}_{P\times P}$, whereas Models 7 to 12 used VCOV matrix direct products for location (L) and harvest (H), i.e., $\mathbf{G}_P = \mathbf{G}^{L}_{N\times N} \otimes \mathbf{G}^{H}_{M\times M}$. For residuals, $\varepsilon \sim N(0,\mathbf{R})$, where $\varepsilon = (\varepsilon_{11111}, ..., \varepsilon_{IJKMN})'$, and $\mathbf{R} = \mathbf{R}_P \otimes \mathbf{R}_K \otimes \mathbf{I}_{I\cdot J}$ is the residual VCOV matrix, whereas matrices $\mathbf{R}_P = \mathbf{R}^{E}_{P\times P}$ (factorial combination) or $\mathbf{R}_P = \mathbf{R}^{L}_{N\times N} \otimes \mathbf{R}^{H}_{M\times M}$ (direct product) and $\mathbf{R}_K = \mathbf{R}^{R}_{K\times K}$ were examined and compared via AIC and BIC for several structures of locations, harvest, and replicates after the selection of $\mathbf{G}_P$. The selected $\mathbf{R}$ matrices were included in the final phenotypic models by considering the existence of non-genetic residual correlations and the heterogeneity of non-genetic residual variances in all harvests and locations. Based on these models, the adjusted means for individual traits in each family and the genetic parameters could be obtained. For each trait, the fixed effects of the interactions between the location, harvest, and checks were tested using the Wald statistics test and were retained in the model if statistically significant ($P < 0.05$). The genotypic correlations between the traits were calculated across the adjusted means as the standard Pearson product-moment coefficient and were tested by assuming a significant global level of $\alpha^* = 0.05$ in R software (http://www.cran.r-project.org) using the package *psych* (Revelle, 2014), which was also used to draw scatterplots between pairs of traits. The broad-sense heritabilities on an individual-plant basis ($\hat{H}^2_{plants}$) were computed based on variance component estimates assuming an identity structure for the $\mathbf{G}_P$ matrix (Model 1) using the ratio $\hat{\sigma}^2_G / \hat{\sigma}^2_P$, where $\hat{\sigma}^2_G$ is the among-genotype variance component and $\hat{\sigma}^2_P$ is the total phenotypic variance for each trait. The ratio $\hat{\sigma}^2_G / \hat{\sigma}^2_{\bar{P}}$ was computed to provide approximate measurements of the broad heritabilities on a genotype-mean basis ($\hat{H}^2_{means}$), where $\hat{\sigma}^2_{\bar{P}}$ is the phenotypic variance among the genotype means for each trait obtained using the harmonic mean of the number of environments as the numerator of the GEI variance estimates and the harmonic mean of the number of full-sib genotypes sampled in each experiment as the numerator of the residual error variance estimates (Holland et al., 2003).

The brown rust severity data did not follow a Gaussian distribution. As a first approach to the problem, a GLMM was used for this analysis (Gianola and Foulley, 1983; Gouy et al., 2013; Thompson, 1979). The average severity of each genotype was transformed into a binary scale, where 0 represents disease resistance (1 on the diagrammatic scale) and 1 represents susceptibility to disease (2–9 on the diagrammatic scale). The transformation of the continuous variable data related to the severity of brown rust in two classes, resistance and susceptibility, was also reported by Asnaghi et al. (2004), Raboin et al. (2006), and Costet et al. (2012).

Table 1. Description and number of parameters ($n_{PAR}$) of the examined models for the genetic variance-covariance matrix $\mathbf{G}_P$ (Models 1–6 used the factorial combination of locations and harvests as different environments [E]; Models 7–12 used the direct product of covariance matrices for locations [L] and harvests [H]; $P = NM$, where $N$ is the number of locations, and $M$ is the number of harvests).

| Model | Parameter | $n_{PAR}$ | Description |
|---|---|---|---|
| | | $\mathbf{G}_P = \mathbf{G}^{E}_{P\times P}$ | |
| 1 | ID | 1 | identity (or homogeneous genetic variances) |
| 2 | UNIF | 2 | uniform |
| 3 | DIAG | P | diagonal (or heterogeneous genetic variances) |
| 4 | $CS_{Het}$ | $P + 1$ | compound symmetry with heterogeneous genetic variation |
| 5 | FA1 | $2P$ | first-order factor analytic |
| 6 | UNST | $P(P + 1)/2$ | unstructured |
| | | $\mathbf{G}_P = \mathbf{G}^{L}_{N\times N} \otimes \mathbf{G}^{H}_{M\times M}$ | |
| 7 | UNST ⊗ ID | $N(N + 1)/2 + 1$ | unstructured and identity for locations and harvest, respectively |
| 8 | UNST ⊗ UNIF | $N(N + 1)/2 + 2$ | unstructured and uniform for locations and harvest, respectively |
| 9 | UNST ⊗ DIAG | $N(N + 1)/2 + M$ | unstructured and diagonal for locations and harvest, respectively |
| 10 | UNST ⊗ AR1 | $[N(N + 1) + 2(M + 1)]/2 - 1$ | unstructured and first-order autoregressive for locations and harvest, respectively |
| 11 | UNST ⊗ $CS_{Het}$ | $N(N + 1)/2 + M + 1$ | unstructured and compound symmetry for locations and harvest, respectively |
| 12 | UNST ⊗ UNST | $[N(N + 1) + M(M + 1)]/2 - 1$ | unstructured for both locations and harvest |

A GLMM with a binomial error distribution and a logit link function was used to model the underlying susceptibility to the disease (observed binary phenotype) (Bolker et al., 2009; De Silva et al., 2014). Location, harvest, and replicate terms were treated as fixed effects, and their significances were assessed by Wald test statistics ($P < 0.05$). A random genotype effect was incorporated into the model and was assumed to be normally distributed, with a zero mean and variance component $\hat{\sigma}^2_G$. The REML was used to estimate the model parameters and variance components using the method of Trust (2014), as implemented in GenStat 16 software (Payne et al., 2009). The approximate broad-sense heritability on a plant-mean basis ($\hat{H}^2_{plants}$) was computed based on $\hat{\sigma}^2_G / (\hat{\sigma}^2_G + \hat{\sigma}^2)$, where $\hat{\sigma}^2_G$ is the genetic variance component estimate and $\hat{\sigma}^2$ is the dispersion parameter estimate.

## RESULTS

### Model Selection for Multi-Harvest-Location Yield Components

Several structures for the $\mathbf{G}_P$ matrix examined and compared via AIC and BIC are summarized in Supplemental Table S1. Four models for the $\mathbf{G}_P$ matrix that consider the heterogeneity of variance were selected according to the data of the evaluated traits (FA1, UNST, UNST ⊗ AR1, and UNST ⊗ UNST) and are provided in Table 2 for each family, SR1 and SR2. The matrix selection for random effects showed that for SR1, the selected models for the traits of SD, SW, BRIX, POL%J, FIB, TCH, and TPH considered the VCOV structure $\mathbf{G}_P = \mathbf{G}^{L}_{N\times N} \otimes \mathbf{G}^{H}_{M\times M}$. The model that was selected for the traits of SN, SH, and POL%C considered the VCOV structure $\mathbf{G}_P = \mathbf{G}^{E}_{P\times P}$. For SR2, the selected models for most traits (SD, SH, SN, BRIX, POL%C, POL%J, FIB, TCH, and TPH) considered a VCOV structure of $\mathbf{G}_P = \mathbf{G}^{E}_{P\times P}$. The selected model considered the VCOV structure of $\mathbf{G}_P = \mathbf{G}^{L}_{N\times N} \otimes \mathbf{G}^{H}_{M\times M}$ only for the SW trait. For each trait, the selection of the mean structure (fixed part of the model) using Wald statistics for SR1 indicated that the interaction effects between harvest and checks were not significant for SH, SN, BRIX, POL%C, POL%J, FIB, TCH, and TPH and that the interaction effects between location and checks were not significant for SD, POL%C, and FIB. In contrast, for SR2, the Wald statistics showed that the interactions between harvest and checks were not significant for SD and FIB and that the interactions between location and checks were not significant for BRIX, POL%C, POL%J, and FIB. On the other hand, the interactions between the location and harvest were not significant for POL%C. The nonsignificant effects were removed from the model, and then the adjusted means were obtained.

For non-genetic residual effects, structures for the **R** matrix examined and compared via AIC and BIC are summarized in Supplemental Table S2 for SR1 and Supplemental Table S3 for SR2. The models that were selected according to the data of the evaluated traits can be viewed in Table 3 for each family, SR1 and SR2. The data of the evaluated traits that fit the factorial combination among locations and harvests in SR1 (SH, SN, POL%C) and SR2 (SD, SH, SN, BRIX, POL%C, POL%J, FIB, TCH, and TPH) showed a pattern of homoscedasticity to nongenetic residual effects between environments, except for POL%C in SR1 and SD, SN, FIB, and TCH in SR2. Furthermore, for these traits that fit the factorial combination among locations and harvests, the covariance between environments was null in SR2 except for the traits SH, SN, TCH, and TPH, whereas in SR1, the covariance was the same between pairs of environments for SH and SN and different for POL%C. For both families, SH and SW showed the same pattern of nongenetic residual effects: SH presented homoscedasticity between environments with the same covariance between pairs of environments, and SW showed heteroscedasticity with null covariance between locations and heteroscedasticity with different covariance between harvests. Every other evaluated trait showed different patterns of behavior for nongenetic residual effects between families.

Table 2. Selected models for the $\mathbf{G}_P$ matrix and number of estimated parameters ($n_{PAR}$) considering each trait separately. The Akaike (AIC) and Bayesian (BIC) information criteria were used to compare the structures of the variance–covariance matrix. The models for the $G_P$ matrix were selected according to the lowest value of the AIC for the stalk diameter (SD) in mm, stalk height (SH) in m, stalk number (SN) by direct counting, stalk weight (SW) in kg, BRIX as °Brix, sucrose content of cane (POL%C) in percentage, sucrose content of juice (POL%J) in percentage, fiber (FIB) as a percentage, cane yield (TCH) in t ha$^{-1}$, and sucrose yield (TPH) in t ha$^{-1}$ for the two families of sugarcane (SR1 and SR2) at two locations (Araras and Ipaussu, Brazil) over three harvest years (2011, 2012, and 2013).

| Trait | Selected model for $\mathbf{G}_P$ matrix† | $n_{PAR}$ | AIC | BIC |
|---|---|---|---|---|
| | SP80–3280 × RB835486 (SR1) | | | |
| SD, mm | 10. UNST ⊗ AR1 | 5 | 18,930.4 | 18,968.1 |
| SH, m | 5. FA1 | 10 | 1274.4 | 1349.8 |
| SN | 5. FA1 | 10 | 37,808.5 | 37,884.0 |
| SW, kg | 10. UNST ⊗ AR1 | 5 | 42,491.7 | 42,529.4 |
| BRIX °Brix | 12. UNST ⊗ UNST | 6 | 6439.5 | 6479.2 |
| POL%C | 6. UNST | 10 | 6407.1 | 6475.0 |
| POL%J | 12. UNST ⊗ UNST | 6 | 7109.1 | 7148.8 |
| FIB, % | 12. UNST ⊗ UNST | 6 | 6050.5 | 6090.1 |
| TCH, t ha$^{-1}$ | 10. UNST ⊗ AR1 | 5 | 36,511.5 | 36,549.2 |
| TPH, t ha$^{-1}$ | 12. UNST ⊗ UNST | 6 | 11,646.2 | 11,685.8 |
| | SP81–3250 × RB925345 (SR2) | | | |
| SD, mm | 6. UNST | 10 | 20,096.9 | 20,245.3 |
| SH, m | 6. UNST | 10 | 728.2 | 876.7 |
| SN | 5. FA1 | 10 | 45,429.0 | 45,519.6 |
| SW, kg | 12. UNST ⊗ UNST | 6 | 47,036.9 | 47,101.5 |
| BRIX, °Brix | 6. UNST | 10 | 7767.2 | 7867.1 |
| POL%C | 6. UNST | 10 | 7653.2 | 7753.0 |
| POL%J | 6. UNST | 10 | 8637.5 | 8737.4 |
| FIB, % | 6. UNST | 10 | 6630.1 | 6730.0 |
| TCH, t ha$^{-1}$ | 6. UNST | 10 | 45,244.1 | 45,392.7 |
| TPH, t ha$^{-1}$ | 6. UNST | 10 | 15,411.7 | 15,546.7 |

† Models selected for the $\mathbf{G}_P$ matrix as described in Table 1.

### Heritability and Components of Variance in Yield Components

The results regarding the ranges, averages of families, averages of parents, estimates of the components of variance, coefficient of variation, and the broad-sense heritability on an individual-plant and genotype-mean basis of the 10 traits evaluated for the two families are summarized in Table 4. In general, the $\hat{H}^2_{means}$ of the traits were high (>0.80) for both families. The $\hat{H}^2_{means}$ ranged from 0.78 (SH) to 0.92 (SD) in SR1 and from 0.79 (POL%C) to

Table 3. Selected models for the R matrix and number of estimated parameters ($n_{PAR}$) considering each trait separately. Akaike (AIC) and Bayesian (BIC) information criteria were used to compare the structures of the variance–covariance matrix. The models for the R matrix were selected according to the lowest value of the BIC for the stalk diameter (SD) in mm, stalk height (SH) in m, stalk number (SN) by direct counting, stalk weight (SW) in kg, BRIX as °Brix, sucrose content of cane (POL%C) in percentage, sucrose content of juice (POL%J) in percentage, fiber (FIB) as a percentage, cane yield (TCH) in t ha$^{-1}$, and sucrose yield (TPH) in t ha$^{-1}$ for the two families of sugarcane (SR1 and SR2) at two locations (Araras and Ipaussu, Brazil) over three harvest years (2011, 2012, and 2013).

| Trait | Selected model for R matrix | $n_{PAR}$ | | | AIC | | | BIC | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SP80–3280 × RB835486 (SR1) | | | | | | | | |
| | $\mathbf{R} = \mathbf{R}^{E}_{P\times P} \otimes \mathbf{R}^{R}_{K\times K}$ | $\mathbf{R}^{E}_{P\times P}$ | $\mathbf{R}^{R}_{K\times K}$ | | $\mathbf{R}^{E}_{P\times P}$ | $\mathbf{R}^{R}_{K\times K}$ | | $\mathbf{R}^{E}_{P\times P}$ | $\mathbf{R}^{R}_{K\times K}$ | |
| SH, m | $\mathbf{R}$ = UNIF ⊗ ID | 2 | 1 | | 1266.1 | 1266.1 | | 1347.8 | 1347.8 | |
| SN | $\mathbf{R}$ = UNIF ⊗ ID | 2 | 1 | | 37,534.3 | 37,534.3 | | 37,616.0 | 37,616.0 | |
| POL%C | $\mathbf{R}$ = UNST ⊗ ID | 10 | 1 | | 6304.7 | 6304.7 | | 6423.6 | 6423.6 | |
| | $\mathbf{R} = \mathbf{R}^{L}_{N\times N} \otimes \mathbf{R}^{H}_{M\times M} \otimes \mathbf{R}^{R}_{K\times K}$ | $\mathbf{R}^{L}_{N\times N}$ | $\mathbf{R}^{H}_{M\times M}$ | $\mathbf{R}^{R}_{K\times K}$ | $\mathbf{R}^{L}_{N\times N}$ | $\mathbf{R}^{H}_{M\times M}$ | $\mathbf{R}^{R}_{K\times K}$ | $\mathbf{R}^{L}_{N\times N}$ | $\mathbf{R}^{H}_{M\times M}$ | $\mathbf{R}^{R}_{K\times K}$ |
| SD, mm | $\mathbf{R}$ = DIAG ⊗ $CS_{Het}$ ⊗ DIAG | 2 | 4 | 3 | 18,906.2 | 18,716.2 | 18,653.8 | 18,950.1 | 18,779.1 | 18,729.2 |
| SW, kg | $\mathbf{R}$ = DIAG ⊗ UNST ⊗ $CS_{Het}$ | 2 | 6 | 4 | 41,985.3 | 41,608.4 | 41,371.4 | 42,029.3 | 41,683.9 | 41,645.7 |
| BRIX, °Brix | $\mathbf{R}$ = ID ⊗ DIAG ⊗ ID | 1 | 3 | 1 | 6439.5 | 6323.4 | 6428.8 | 6479.2 | 6374.4 | 6474.1 |
| POL%J | $\mathbf{R}$ = ID ⊗ UNIF ⊗ ID | 1 | 2 | 1 | 7109.1 | 7002.5 | 7002.5 | 7148.8 | 7047.9 | 7047.9 |
| FIB, % | $\mathbf{R}$ = ID ⊗ UNST ⊗ DIAG | 1 | 3 | 2 | 6050.5 | 5923.5 | 5912.1 | 6090.1 | 5974.5 | 5968.7 |
| TCH, t ha$^{-1}$ | $\mathbf{R}$ = UNST ⊗ UNIF ⊗ DIAG | 3 | 2 | 3 | 36,498.7 | 35,980.8 | 35,961.1 | 36,549.0 | 36,037.4 | 36,030.3 |
| TPH, t ha$^{-1}$ | $\mathbf{R}$ = ID ⊗ UNIF ⊗ DIAG | 1 | 2 | 2 | 11,646.2 | 11,496.5 | 11,490.4 | 11,685.8 | 11,541.8 | 11,541.3 |
| | | SP81–3250 × RB925345 (SR2) | | | | | | | | |
| | $\mathbf{R} = \mathbf{R}^{E}_{P\times P} \otimes \mathbf{R}^{R}_{K\times K}$ | $\mathbf{R}^{E}_{P\times P}$ | $\mathbf{R}^{R}_{K\times K}$ | | $\mathbf{R}^{E}_{P\times P}$ | $\mathbf{R}^{R}_{K\times K}$ | | $\mathbf{R}^{E}_{P\times P}$ | $\mathbf{R}^{R}_{K\times K}$ | |
| SD, mm | $\mathbf{R}$ = DIAG ⊗ UNST | 6 | 6 | | 19,508.9 | 19,335.7 | | 19,689.7 | 19,548.8 | |
| SH, m | $\mathbf{R}$ = UNIF ⊗ ID | 2 | 1 | | 702.6 | 702.6 | | 857.5 | 857.5 | |
| SN | $\mathbf{R}$ = UNST ⊗ $CS_{Het}$ | 21 | 4 | | 43,966.4 | 43,943.4 | | 44,186.2 | 44,186.2 | |
| BRIX, °Brix | $\mathbf{R}$ = ID ⊗ ID | 1 | 1 | | 7767.2 | 7767.2 | | 7867.1 | 7867.1 | |
| POL%C | $\mathbf{R}$ = ID ⊗ ID | 1 | 1 | | 7653.2 | 7653.2 | | 7753.0 | 7753.0 | |
| POL%J | $\mathbf{R}$ = ID ⊗ ID | 1 | 1 | | 8637.5 | 8637.5 | | 8737.4 | 8737.4 | |
| FIB, % | $\mathbf{R}$ = DIAG ⊗ ID | 5 | 1 | | 6489.8 | 6489.8 | | 6613.1 | 6613.1 | |
| TCH, t ha$^{-1}$ | $\mathbf{R}$ = UNST ⊗ UNIF | 21 | 2 | | 44,621.9 | 44,605.1 | | 44,899.7 | 44,889.3 | |
| TPH, t ha$^{-1}$ | $\mathbf{R}$ = UNIF ⊗ ID | 2 | 1 | | 15,319.9 | 15,319.9 | | 15,460.7 | 15,460.7 | |
| | $\mathbf{R} = \mathbf{R}^{L}_{N\times N} \otimes \mathbf{R}^{H}_{M\times M} \otimes \mathbf{R}^{R}_{K\times K}$ | $\mathbf{R}^{L}_{N\times N}$ | $\mathbf{R}^{H}_{M\times M}$ | $\mathbf{R}^{R}_{K\times K}$ | $\mathbf{R}^{L}_{N\times N}$ | $\mathbf{R}^{H}_{M\times M}$ | $\mathbf{R}^{R}_{K\times K}$ | $\mathbf{R}^{L}_{N\times N}$ | $\mathbf{R}^{H}_{M\times M}$ | $\mathbf{R}^{R}_{K\times K}$ |
| SW, kg | $\mathbf{R}$ = DIAG ⊗ UNST ⊗ UNST | 2 | 3 | 6 | 46,497.2 | 45,954.0 | 46,065.3 | 46,568.3 | 46,057.4 | 45,929.5 |

0.94 (SD) in SR2. Estimates for $\hat{H}^2_{plants}$ above 0.30 were found for all of the traits except SH (0.19), SN (0.29), SW (0.26), and TCH (0.27) in SR1 and SH (0.23) in SR2. The $\hat{H}^2_{plants}$ ranged from 0.19 (SH) to 0.45 (SD) in SR1 and from 0.23 (SH) to 0.49 (SD) in SR2. Even considering some indicated exceptions, the values show that much of the observed phenotypic variation can be attributed to differences in the genotypic level.

The genotypic and residual coefficients of variation between SR1 and SR2 for each individual trait were similar, with a few exceptions for the residual coefficient of variation ($CV_R$). The exceptions to the pattern of similarity were as follows: (i) the $CV_R$ for SW was approximately 40% higher in SR1 (25.50) than in SR2 (18.30), and (ii) TCH was approximately 34% higher in SR1 (23.51) than in SR2 (17.54). The values of the estimates of the genetic and phenotypic variances were similar in SR1 and SR2 for each individual trait. An exception was the estimate of the genetic variance component: SN was approximately 77% higher in SR2 (529.90) than in SR1 (298.00). In the estimate of the phenotypic variance component, SW was approximately 70% higher in SR1 (3177.00) than in SR2 (1868.60), TCH was 58% higher in SR1 (2120.00) than in SR2 (1336.00), SD was approximately 53% higher in SR1 (11.10) than in SR2 (7.23), and TPH was 32% higher in SR1 (38.93) than in SR2 (29.46).

The range of variation was different between the SR1 and SR2 families for all of the evaluated traits. Family SR2 showed much higher ranges of variation for the traits of SN, SW, and TCH. The average values of the traits of SH, BRIX, POL%C, POL%J, and FIB showed similar variations between the two evaluated families. However, SD, SN, SW, TPH, and TCH showed differences in the average values between the families. The average values of SD, SW, TCH, and TPH were greater in SR1, whereas the average value of SN was higher in SR2. In SR1, the average of the progeny was higher than the average of both parents for TCH and TPH. In SR2, the average of the progeny was higher than the average of both parents for SD. In addition, in both families and for all traits evaluated, the offspring had higher averages than the parents of the families. These results show the occurrence of transgressive segregation in both families.

Table 4. Ranges, averages, estimates of components of genetic variance ($\hat{\sigma}^2_G$) and phenotype ($\hat{\sigma}^2_P$), coefficients of genotypic variation ($CV_G$) and residual ($CV_R$), and broad heritability on a genotype-mean ($\hat{H}^2_{means}$) and individual-plant basis ($\hat{H}^2_{plants}$), stalk diameter (SD) in mm, stalk height (SH) in m, stalk number (SN) by direct counting, stalk weight (SW) in kg, BRIX as °Brix, sucrose content of cane (POL%C) in percentage, sucrose content of juice (POL%J) in percentage, fiber (FIB) as a percentage, cane yield (TCH) in t ha$^{-1}$, and sucrose yield (TPH) in t ha$^{-1}$ for the two families of sugarcane (SR1 and SR2) at two locations (Araras and Ipaussu, Brazil) over three harvest years (2011, 2012, and 2013).

| Trait | Range | Avg. | SP80-3280 | SP81-3250 | RB835486 | RB925345 | $\hat{\sigma}^2_G$ | $\hat{\sigma}^2_P$ | $CV_G$ | $CV_R$ | $\hat{H}^2_{means}$ | $\hat{H}^2_{plants}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | SP80-3280 × RB835486 (SR1) | | | | | | | | |
| SD, mm | 24.18–34.44 | 28.39 | 29.57 | | 27.16 | | 4.96 | 11.10 | 7.84 | 8.73 | 0.92 | 0.45 |
| SH, m | 1.99–2.60 | 2.31 | 2.37 | | 2.22 | | 0.01 | 0.09 | 5.71 | 11.65 | 0.78 | 0.19 |
| SN | 71.80–138.50 | 105.85 | 108.10 | | 96.20 | | 298.00 | 1012.30 | 16.31 | 25.24 | 0.86 | 0.29 |
| SW, kg | 128.40–254.10 | 184.32 | 203.50 | | 152.3 | | 822.00 | 3177.00 | 15.55 | 25.50 | 0.82 | 0.26 |
| BRIX, °Brix | 18.79–22.88 | 20.85 | 21.19 | | 21.05 | | 0.63 | 1.62 | 3.82 | 4.76 | 0.83 | 0.39 |
| POL%C | 13.46–17.26 | 15.52 | 15.72 | | 15.84 | | 0.64 | 1.65 | 5.15 | 6.45 | 0.83 | 0.39 |
| POL%J | 15.92–20.62 | 18.44 | 18.63 | | 18.85 | | 0.89 | 2.27 | 5.11 | 6.38 | 0.83 | 0.39 |
| FIB, % | 10.78–14.59 | 12.20 | 12.04 | | 12.34 | | 0.59 | 1.46 | 6.30 | 7.65 | 0.84 | 0.40 |
| TCH, t ha$^{-1}$ | 114.50–219.90 | 162.07 | 95.84 | | 85.48 | | 582.00 | 2120.00 | 14.88 | 23.51 | 0.83 | 0.27 |
| TPH, t ha$^{-1}$ | 15.87–31.84 | 23.72 | 16.00 | | 15.00 | | 13.63 | 38.93 | 15.56 | 21.09 | 0.81 | 0.35 |
| | | | | SP81-3250 × RB925345 (SR2) | | | | | | | | |
| SD, mm | 21.16–31.64 | 25.61 | | 25.25 | | 25.18 | 3.55 | 7.23 | 7.35 | 7.15 | 0.94 | 0.49 |
| SH, m | 1.98–2.68 | 2.34 | | 2.25 | | 2.49 | 0.02 | 0.08 | 5.87 | 10.06 | 0.83 | 0.23 |
| SN | 50.50–202.40 | 117.44 | | 127.60 | | 115.20 | 529.90 | 1257.60 | 19.59 | 22.10 | 0.92 | 0.42 |
| SW, kg | 64.40–244.30 | 168.15 | | 172.50 | | 175.70 | 734.20 | 1868.60 | 16.11 | 18.30 | 0.90 | 0.39 |
| BRIX, °Brix | 18.27–22.88 | 20.93 | | 21.02 | | 21.63 | 0.65 | 1.59 | 3.86 | 4.00 | 0.81 | 0.40 |
| POL%C | 13.24–17.35 | 15.54 | | 15.60 | | 16.24 | 0.53 | 1.42 | 4.67 | 5.42 | 0.79 | 0.37 |
| POL%J | 15.62–20.77 | 18.57 | | 18.58 | | 19.47 | 0.81 | 2.11 | 4.84 | 5.42 | 0.80 | 0.38 |
| FIB, % | 11.17–15.13 | 12.64 | | 12.37 | | 12.94 | 0.52 | 1.12 | 5.72 | 5.66 | 0.86 | 0.46 |
| TCH, t ha$^{-1}$ | 84.10–214.10 | 149.04 | | 150.70 | | 157.00 | 543.10 | 1336.00 | 15.63 | 17.54 | 0.90 | 0.40 |
| TPH, t ha$^{-1}$ | 14.02–30.93 | 22.09 | | 22.91 | | 24.26 | 11.53 | 29.46 | 15.36 | 18.26 | 0.82 | 0.39 |

## Genotypic Correlations of Yield Components

Pairwise genotypic correlations among the 10 evaluated traits, considering two locations (Araras and Ipaussu) and three harvests (2011, 2012, and 2013), are shown in Fig. 1. Overall, 17 and 12 significant genotypic correlations ($P < 0.05$) occurred between the evaluated traits in SR1 and SR2, respectively. According to the degree of correlation between the traits, the correlations were grouped into low (≤0.35), moderate (0.36–0.70) and strong (>0.71). Thus, in SR1, seven interactions were classified as low (SD–SN, SD–SW, SD–FIB, SD–TCH, SD–TPH, POL%C–TPH, and POL%J–TPH), three were classified as moderate (SH–SW, SN–SW and SW–TPH), and seven were classified as strong (SN–TCH, SN–TPH, SW–TCH, BRIX–POL%C, BRIX–POL%J, POL%C–POL%J, and TCH–TPH). The correlations SD–SN, SD–FIB, SD–TCH, and SD–TPH were negative (Fig. 1a). In SR2, four interactions were classified as low (SD–SH, SH–FIB, BRIX–FIB, and POL%J–FIB), two were classified as moderate (SD–SN and SD–FIB), and six were classified as strong (SW–TCH, SW–TPH, BRIX–POL%C, BRIX–POL%J, POL%C–POL%J, and TCH–TPH). The correlations SD–SN, SD–FIB, and SH–FIB were negative (Fig. 1b).

Of the total genotypic correlations that were observed, eight were common between SR1 and SR2 (SD–SN, SD–FIB, SW–TCH, SW–TPH, TCH–TPH, BRIX–POL%C, BRIX–POL%J, and POL%C–POL%J). However, three correlations exhibited differences between the two families (SD–SN, SD–FIB, and SW–TPH). The SD–SN and SD–FIB correlations were classified as negative and low in SR1 (−0.31 and −0.29, respectively) and as negative and moderate in SR2 (−0.44 and −0.39, respectively). The SW–TPH correlation was classified as positive and moderate in SR1 (0.62) and positive and strong in SR2 (0.92). The eight exclusive correlations that were present in SR1 were classified as low (SD–SW, SD–TCH, SD–TPH, POL%C–TPH, and POL%J–TPH), moderate (SH–SW and SN–SW) and strong (SN–TCH and SN–TPH), whereas the correlations SD–TCH and SD–TPH were negative (Fig. 1a). In SR2, four exclusive correlations were classified as low (SD–SH, SH–FIB, BRIX–FIB, and POL%J–FIB), whereas the correlation SH–FIB was negative (Fig. 1b).

## Probabilities, Segregations, and Heritability of Resistance to Brown Rust

The GLMM-based analysis revealed that approximately 66% (101) and 32% (74) of full-sib genotypes in SR1 and SR2, respectively, have at most a 10% probability of showing symptoms of the disease (Fig. 2), i.e., at least a 90% probability of being resistant under the evaluated local and environmental conditions. In SR1, the parent SP80-3280 showed a 99.50% probability of being resistant to the disease, while the parent RB835486 showed a 99.80% probability of being susceptible to brown rust. In SR2, the parents SP81-3250 and RB925345 showed 88.90% and 96.01% probabilities of being resistant and susceptible to brown rust, respectively (Fig. 2). The segregation that was observed in SR1 showed a strong displacement of the curve toward the class that was considered resistant (up

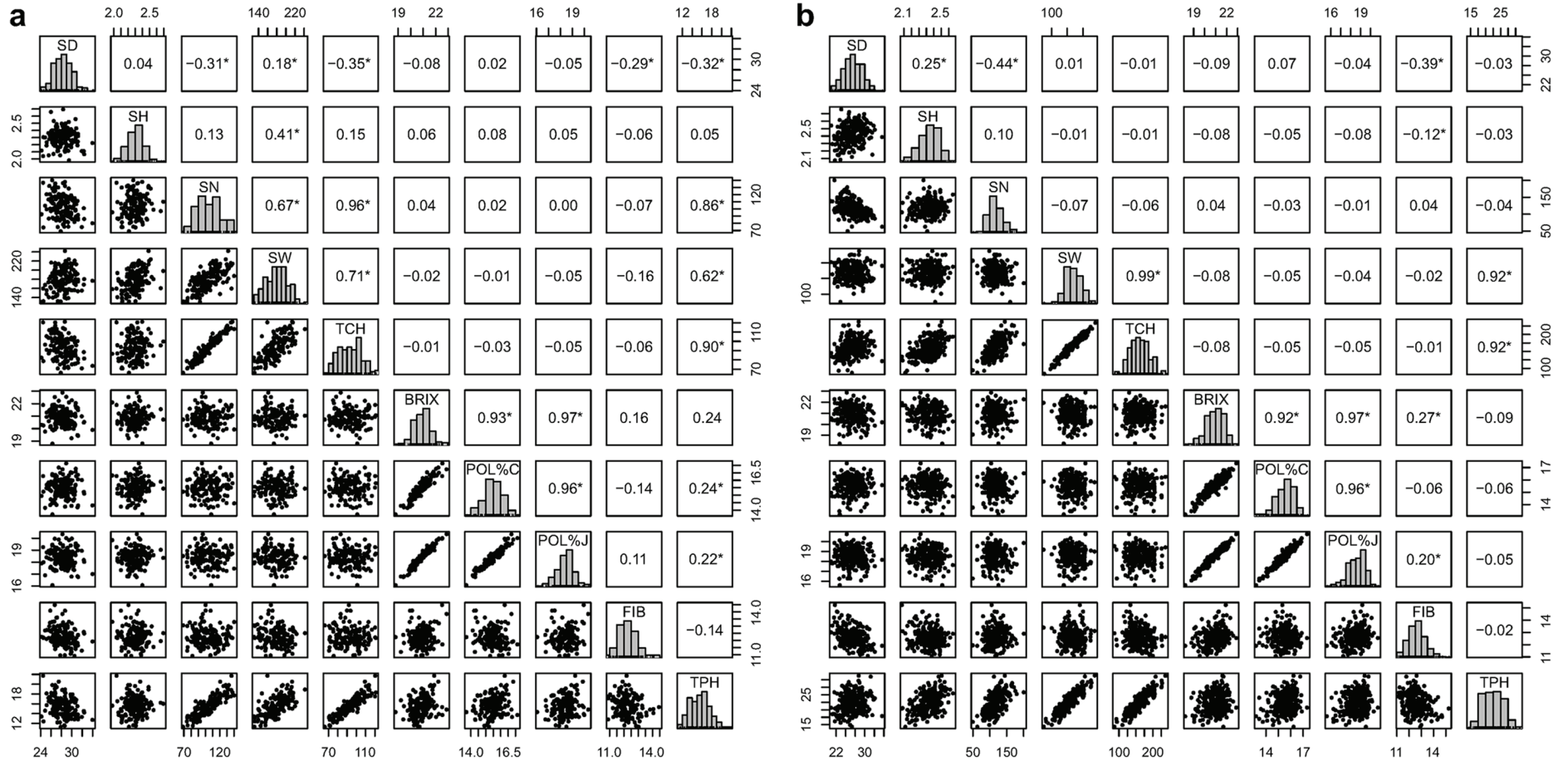


Fig. 1. Estimates of the genotypic correlation of yield components for (a) Family SR1 derived from a cross between SP80-3280 and RB835486, and (b) Family SR2 derived from a cross between SP81-3250 and RB925345 for the stalk diameter (SD) in mm, stalk height (SH) in m, stalk number (SN) by direct counting, stalk weight (SW) in kg, BRIX as °Brix, sucrose content of cane (POL%C) in percentage, sucrose content of juice (POL%J) in percentage, fiber (FIB) as a percentage, cane yield (TCH) in t $ha^{-1}$, and sucrose yield (TPH) in t $ha^{-1}$ at two locations (Araras and Ipaussu, Brazil) over three harvest years (2011, 2012, and 2013). For each trait, the histograms of the adjusted means (diagonal), scatterplots (below diagonal), and values of the genotypic correlation (above diagonal) between pairs of traits are shown. *Significant at the 5% global level ($P < 0.05$).

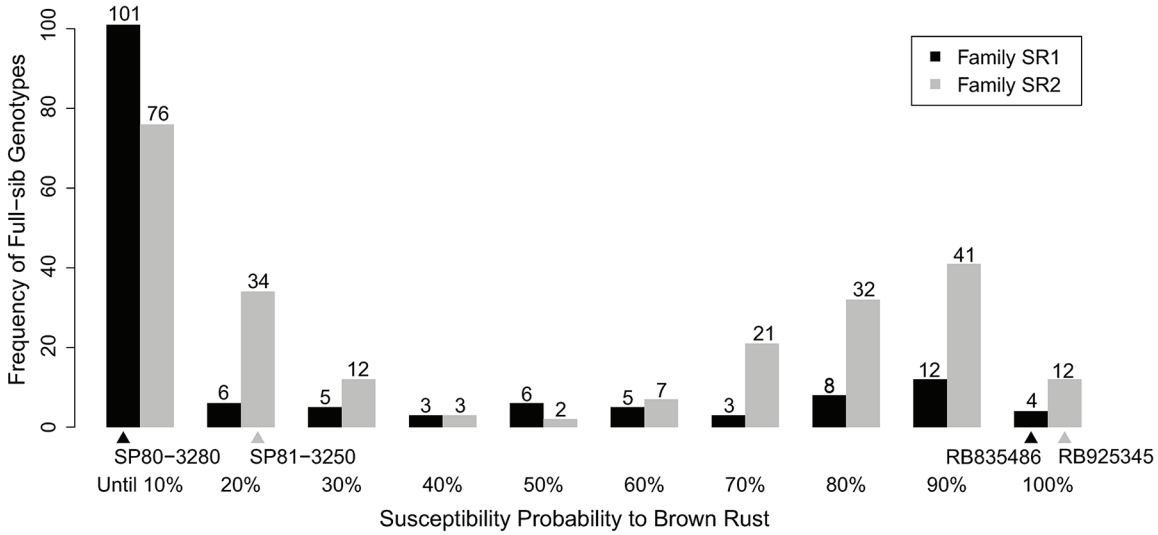


Fig. 2. Frequency of full-sib genotypes in the probability classes of brown rust susceptibility as calculated with a generalized linear mixed model (GLMM) for the two families of sugarcane (SR1, derived from a cross between SP80-3280 and RB835486, and SR2, derived from a cross between SP81-3250 and RB925345) at two locations (Araras and Ipaussu, Brazil) over three harvest years (2011, 2012, and 2013).

to a 10% probability of having the disease). In SR2, a bimodal distribution was observed (Fig. 2). The $\hat{H}^2_{\text{plants}}$ of brown rust resistance were high (0.93 and 0.84 in SR1 and SR2, respectively). The genetic coefficient of variation was 8.76 and 5.31 in SR1 and SR2, respectively, and the residual coefficient of variation was 0.62 and 0.95 in SR1 and SR2, respectively.

## DISCUSSION

Sugarcane is one of the most important crops worldwide, and its importance is mainly attributed to its derivatives, i.e., sugar and ethanol. Brazil is the world's largest producer of sugarcane (FAO, 2014) and is constantly striving for increased production. However, increasing production requires an adequate and sustainable method for the modernization of sugarcane cultivation. Therefore, high-technology agriculture associated with knowledge of the genetic basis of inheritance of all traits that are of economic interest and are directly linked to sugarcane production is very important for increasing productivity without expanding the area planted with sugarcane.

A set of traits should be simultaneously considered because interest lies in the combined selection of traits rather than isolated traits. Several studies that involved yield components have been conducted on sugarcane, albeit using statistical approaches with a series of limitations (Gallacher, 1997; Lin et al., 1993), specifically, without considering the unbalanced data and variance homogeneity assumptions and without assessing whether there were genetic correlations between harvests and locations for estimating breeding values (Balzarini, 2002; Piepho and Möhring, 2007; Smith et al., 2005). Another important limitation is missing phenotypic data, which is very common in experiments with sugarcane. Modeling these limitations would allow more realistic results that should be easier to apply in a sugarcane genetic breeding program. The mixed models approach is suitable for evaluating the heterogeneity of genetic variances and correlations across environments (Malosetti et al., 2013). Therefore, the use of a more sophisticated statistical model would permit data processing that is more appropriate for the experiment and that signifies a major change in the analysis (i.e., the genotype grouping factors, such as the harvest year and location, would be considered) (Pastina et al., 2012). Although the adjustments for VCOV structures were comprehensively studied for the data in the present study, it is important to note that the use of more locations and harvest years would probably permit the adjustment of other variance and covariance structures. For example, in sugarcane data across harvest years, the data from the same individual with time are expected to be correlated. This result is more evident in families from breeding programs with more harvest years and locations because of the long experimentation process until the release of a new cultivar.

The VCOV matrices were primarily constructed by considering the genetic effects matrix (**G** matrix) and then the nongenetic effects matrix (**R** matrix). In principle, the selection of VCOV models for the location effect requires, among other factors, prior knowledge of the soil and climatic conditions of each evaluated location. For harvests, the biological response of the plants across years is important. Locations (i.e., Araras and Ipaussu) have different soil and climatic conditions, which contribute to complex interactions and possible changes in the responses of genotypes. Harvests experience a drop in productivity with time within one cycle. Thus, the individual models for each measured trait are appropriate, and they can reflect the efficiency and reality of the final genotypic response.

Considering the AIC values for the selection of the best models in the sugarcane families, a preferential selection for SR1 was observed in Models 10 (SD, SW, and TCH) and 12 (BRIX, POL%J, FIB, and TPH) with VCOV structure $\mathbf{G}_P = \mathbf{G}^{\text{L}}_{N\times N} \otimes \mathbf{G}^{\text{H}}_{M\times M}$ (Table 2). Model 10 presents an UNST structure for local and AR1 for harvests, indicating a correlation between successive harvests and a systematic explanation of the existing temporal dependence. The productivity of sugarcane tends to decrease over harvests; therefore, we expect a greater correlation among nearby harvests and a lower correlation among distant harvests due to physiological and genetic changes. Model 12 presents a UNST structure for both locations and harvests; this is a complex model that generally captures all of the possible variations, i.e., the traits that exhibit different variances and covariances between locations

and harvests. Model 5 was selected for SH and SN, and Model 6 was selected for POL%C; these models have a VCOV structure of $\mathbf{G}_P = \mathbf{G}^{E}_{P\times P}$. Model 6 assumes a general structure $\mathbf{G}_P$ matrix, which is completely unstructured for different genetic variances in each environment and for different covariances between pairs of environments. Model 5 is an approximated unstructured model that can be interpreted as a linear regression model for genotype effects and GEI of environmental covariates, i.e., it measures the sensitivity of the genotype in relation to the "weight" of each environment. This model has been suggested for MET analysis (Burgueño et al., 2012; Kelly et al., 2007; Piepho, 1998; Thompson et al., 2003) because it captures the genetic variation in genotypes in terms of environment as well as the genetic covariance between environments with more realistic and accurate predictions. In SR2, SW was the only trait with selected VCOV $\mathbf{G}_P = \mathbf{G}^{L}_{N\times N} \otimes \mathbf{G}^{H}_{M\times M}$ structure (Model 12), and a preferential selection by Model 6 with VCOV $\mathbf{G}_P = \mathbf{G}^{E}_{P\times P}$ structure was also observed (SD, SH, BRIX, POL%C, POL%J, FIB, TCH, and TPH) (Table 2). The SN and POL%C were the only traits with the same models selected in SR1 and SR2 (Models 5 and 6, respectively) (Table 2). The genetic complexity of sugarcane is also reflected in the response of genotypes under the conditions to which they were submitted and consequently in the model selection that best fit the response pattern of the data. The biomass production of sugarcane is influenced by several factors (genetic, physiological, and environmental), so Models 6 and 12 are most suitable for estimating the genetic parameters of this measure by considering all possible variations. However, this matrix requires an estimation of the maximum number of parameters. In experiments with many locations and harvests, this analysis can become computationally unfeasible. Alternatively, the FA1 matrix, which considers the genetic effects of location (Boer et al., 2007; Burgueño et al., 2012; Kelly et al., 2007; Meyer, 2009; Smith et al., 2007; Thompson et al., 2003), along with the AR1 matrix adjusted for genetic effects of harvests (Pastina et al., 2012), can accurately predict genetic parameters.

Commonly, sugarcane plant breeders independently assess the results of each experiment. This practice is equivalent to the predictions of the DIAG model, which indicates the heterogeneity of variance but not the correlation of performances of genotypes among the experiments (Kelly et al., 2007). Our results show that none of the analyzed traits adjusted to the DIAG model for the $\mathbf{G}_P$ matrix (Table 2). Thus, the model selection approach that adjusts the natural response pattern for each trait is far superior to that currently practiced by plant breeders because it can capture both the heterogeneity of variance and more complex covariance structures at the genetic level, resulting in a more accurate prediction of individual experiments or multiple environments. The implementation of this data analysis model improves the predictive accuracy directly related to the heritability and genetic gain. Sugarcane breeding programs can increase the efficiency of superior genotype selection in every stage of the selection and in diverse environments.

Defined as the heritable portion transmitted to offspring (Falconer and Mackay, 1996), heritability is an important parameter because it determines the response to selection and because it can help select the optimal strategy for a breeding program (Piepho and Möhring, 2007; Sadras et al., 2013). All of the broad hereditability values presented in this study are high (Bernardo, 2010) ranging from 0.78 to 0.92 for SR1 and from 0.79 to 0.94 for SR2; thus, the observed phenotypic variation is mostly due to the genotypic variation in these populations (Table 4). Comparing the results of SR1 and SR2 with those reported in the literature, Hoarau et al. (2002) found lower $\hat{H}^2_{\text{means}}$ values for SD (0.91), SH (0.83), SN (0.86) and BRIX (0.81) in separate populations of selfing of R570 compared with SR1 and SR2, with the exception of SH, which was lower in SR1. Aitken et al. (2006) found $\hat{H}^2_{\text{means}}$ values that were slightly higher for BRIX (0.88) and POL%J (0.93) in separate populations of a cross between Q165 and IJ76-514. Aitken et al. (2008), using the same parents and a population of 227 individuals, found lower $\hat{H}^2_{\text{means}}$ values for SD (0.88), SN (0.83), and TCH (0.71) and higher values for SH (0.85). Pinto et al. (2010), while working with a separate population of a cross between SP80-180 and SP80-4966, found slightly higher $\hat{H}^2_{\text{means}}$ values for POL%C (0.84) and TPH (0.88) and the lowest value for FIB (0.81). The value for TCH (0.87) was greater than that of SR1 and less than that of SR2. Mancini et al. (2012) found lower $\hat{H}^2_{\text{means}}$ values for SD (0.80), SH (0.72), SN (0.76), SW (0.77), BRIX (0.60), POL%C (0.59), FIB (0.75), and TCH (0.70) when evaluating a separate population of a cross between IACSP95-3018 and IACSP93-3046. These studies did not consider the correlations between harvests. The comparison of results revealed that the broad-sense hereditability values were mostly higher in SR1 and SR2. A good experimental control combined with a statistical model that can integrate data at different locations and multiple harvests generates more accurate estimates of heritability.

The range of the phenotypic values for all of the evaluated traits was greater than the phenotypic range of the parents, i.e., transgressive segregation occurred (Table 4) as also observed by Hoarau et al. (2002) and Mancini et al. (2012). Certainly, the selection of checks is a crucial point for the comprehension of the phenotypic values observed in these experiments. In METs, several environments are tested, and each genotype can develop best in a specific environment compared with others. Comparisons with suitable checks are fundamental for efficient BLUP estimation and the identification of the best genotypes. Furthermore, the selection of checks based on considerations of more than one trait is a challenge because it is desirable that all of the check's phenotypic values are within the range of the phenotypic values of the progeny. In our experiments, we used commercial cultivars as checks, i.e., for each family, the parents and a non-parental cultivar were used as checks (Supplemental Table S4). The check non-parental cultivar RB867515 is currently the most cultivated cultivar in Brazil and exhibits high productivity rates in different production environments, i.e., under different types of soil and climatic conditions. The choice of RB867515 as a non-parental check was primarily based on the substantial knowledge of its production behaviors in different environments. Ensuring the best estimates of the phenotypic values via the appropriate choice of checks is a fundamental step that can contribute to breeding programs and result in the release of cultivars with higher yields.

Genotypes can also be selected through genotypic correlation, which combines more than one desirable trait in the same

plant for indirect selection (Ram et al., 1997). Correlations among traits may reflect biological processes that are of considerable evolutionary interest and are the result of genetic, functional, physiological, or developmental nature (Jamoza et al., 2014; Soomro et al., 2006). The common strong genotypic correlation between SR1 and SR2 (SW–TCH, BRIX–POL%C, BRIX–POL%J, POL%C–POL%J, and TCH–TPH) (Fig. 1) shows that the selection practiced by breeding programs has aimed to increase the amount of sugar and the stalk weight, considering that the parents that originated both families had a different genetic background and still gather favorable alleles for the expression of correlated traits. Using these main traits (SW, BRIX, POL%C, POL%J, TCH, and TPH) throughout the selection period of a breeding program, the cultivars may meet the expectations of highly accumulated sucrose and high production in terms of weight. However, the genetic gain for these traits with the conventional breeding process is nearly stagnant (Dal-Bianco et al., 2012). Several limitations inherent in a breeding program may be noted: (i) a lack of knowledge of the genetic material that is present in germplasm banks, which could be used to perform cross-breeding with the greatest potential to generate superior cultivars; (ii) sparse and mismanaged experimental trials; (iii) failures and a lack of standardization in the collection of phenotypic data; (iv) lack of environmental correlation analysis of the phenotypic data; (v) lack of knowledge of the genetic basis of the traits of interest; (vi) neglect of disease and pest occurrence; and (g) low investment in research and biotechnology (Breseghello and Coelho, 2013; Mahon, 1983; Prohens, 2011).

Among the diseases that affect sugarcane, brown rust is present in almost all of the cultivation areas (Asnaghi et al., 2004; Ryan and Egan, 1989). This disease can hinder the performance of cultivars and exclude them from the breeding stock of producing units. Therefore, sugarcane breeding programs should seek sources of resistance and produce cultivars that are able to resist the pathogen. When evaluating the VCOV structures that are adjusted to residues, we found that DIAG was appropriate for SD and SW in SR1 and SW in SR2 for location. In addition, DIAG was appropriate for SD and FIB for the factorial combination between location and harvest in SR2 (Table 3). Thus, residues, which are causes of variation that are not controlled for in these traits, are possibly different or present different intensities between the two locations. Araras and Ipaussu have different soil and climatic characteristics, as discussed above. Moreover, it is important to highlight that Ipaussu has a great natural incidence of brown rust disease due to the fungus *P. melanocephala*, which could strongly interfere with productivity. The extreme compatibility of the fungus that causes brown rust with sugarcane reduces the life of leaves, which lose their photosynthetic function. The presence of sporulating pustules on the leaves results in the reduced growth of sugarcane and significant productivity losses, compromising the final biomass production, depending on the susceptibility of the cultivar and the environmental conditions (Asnaghi et al., 2000; Hoy and Hollier, 2009; McFarlane et al., 2006; Oloriz et al., 2011, 2012; Purdy et al., 1983; Raid and Comstock, 2000; Taylor et al., 1986). The severity of brown rust is assessed on a particular scale (Amorim et al., 1987; Tai et al., 1981), and the frequency of full-sib genotypes in each severity class is influenced by the genetic basis of the parents of the family. Several researchers have reported that brown rust resistance is controlled by one or a few genes (Asnaghi et al., 2004; Costet et al., 2012; Daugrois et al., 1996; Garsmeur et al., 2011; Glynn et al., 2013; Hogarth et al., 1993; Le Cunff et al., 2008; Parco et al., 2014; Raboin et al., 2006; Racedo et al., 2013; Ramdoyal et al., 2000; Sordi et al., 1988). Costet et al. (2012) showed that resistance to brown rust in modern polyploid sugarcane cultivars essentially depends on the major gene *Bru*1. A GLMM was used to analyze data from SR1 and SR2 because the binomial variable resistance–susceptibility was used to characterize the data of the progenies (Bennewitz et al., 2014; Bolker et al., 2009; De Silva et al., 2014). A strong displacement of the curve toward the class that is considered resistant (up to a 10% probability of having the disease) occurred in SR1, and a bimodal distribution was observed for SR2 (Fig. 2). Our results suggest that one or a few genes originating from the parent SP80-3280 and SP81-3250 may transfer resistance to the full-sib genotypes in SR1 and SR2, respectively. The distribution of SR2 was bimodal and allows us to infer that a combination of a different number of copies of the gene conferring resistance or susceptibility to full-sib genotypes from this family is preferred. In addition, because the broad-sense heritability of brown rust resistance had high values of 0.93 and 0.84 in SR1 and SR2, respectively, *Bru*1 may be present and segregating in the progenies.

Therefore, knowledge of the genotypes, the correct orientation of the cross-breeds, experimental trials with highly accurate measurements, the use of models with VCOV matrices (which consider the heterogeneity of the variance and assess the correlations between locations and harvests), and the commitment to generate truly resistant cultivars are fundamental for obtaining more productive sugarcane cultivars. Our results showed that LMMs and GLMMs for estimating genetic parameters in sugarcane are potentially useful in the investigation of the heterogeneous genetic variances and correlations between environments. These models can be tested in other families and METs of sugarcane for an understanding of the complex relationships among traits and environments.

## ACKNOWLEDGMENTS

We gratefully acknowledge the Raízen group for their support in the field experiments in Ipaussu. We also acknowledge Luiz Plínio Zavaglia and Sandro Augusto Ferrarez for their support in the field experiments and Rodrigo Gazaffi for his comments on the manuscript. This work was supported by grants from INCT-Bioetanol (Instituto Nacional de Ciência e Tecnologia do Bioetanol), FINEP (Finaciadora de Estudos e Projetos), and FAPESP (Fundação de Amparo à Pesquisa de São Paulo, 08/52197-4). Thiago W.A. Balsalobre, Melina C. Mancini, and Guilherme da S. Pereira received doctorate fellowships from FAPESP (10/50091-4, 10/50549-0 and 12/25236-4, respectively). Carina O. Anoni received a doctorate fellowship from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico). Fernanda Z. Barreto received a master's fellowship from CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). Anete P. de Souza and Antonio A. F. Garcia received research fellowships from CNPq.

## REFERENCES


Aitken, K.S., S. Hermann, K. Karno, G.D. Bonnett, L.C. McIntyre, and P.A. Jackson. 2008. Genetic control of yield related stalk traits in sugarcane. Theor. Appl. Genet. 117:1191–1203. doi:10.1007/s00122-008-0856-6

Aitken, K.S., P.A. Jackson, and C.L. McIntyre. 2006. Quantitative trait loci identified for sugar related traits in a sugarcane (*Saccharum* spp.) cultivar × *Saccharum officinarum* population. Theor. Appl. Genet. 112:1306–1317. doi:10.1007/s00122-006-0233-2

Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. 19:716–723. doi:10.1109/TAC.1974.1100705

Amorim, L., A. Bergamin-filho, A. Sanguino, C. Cardoso, V.A. Moraes, and C.R. Fernandes. 1987. Metodologia de avaliação da ferrugem da cana-de-açúcar (*Puccinia melanocephala*). Bol. Tec. COPERSUCAR 39:13–16.

Asnaghi, C., F. Paulet, C. Kaye, L. Grivet, J.C. Glaszmann, and A. D'Hont. 2000. Application of synteny across the Poaceae to determine the map location of a rust resistance gene of sugarcane. Theor. Appl. Genet. 101:962–969. doi:10.1007/s001220051568

Asnaghi, C., D. Roques, S. Ruffel, C. Kaye, J.Y. Hoarau, H. Télismart, et al. 2004. Targeted mapping of a sugarcane rust resistance gene (*Bru*1) using bulked segregant analysis and AFLP markers. Theor. Appl. Genet. 108:759–764. doi:10.1007/s00122-003-1487-6

Balzarini, M. 2002. Applications of mixed models in plant breeding. In: M.S. Kang, editor, Quantitative genetics, genomics and plant breeding. CABI Publ., New York. p. 353–363.

Beaulieu, J., T. Doerksen, S. Clément, J. Mackay, and J. Bousquet. 2014. Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. Heredity 113:343–352. doi:10.1038/hdy.2014.36

Bennewitz, J., S. Bögelein, P. Stratz, M. Rodehutscord, H.P. Piepho, J.B. Kjaer, and W. Bessei. 2014. Genetic parameters for feather pecking and aggressive behavior in a large $F_2$–cross of laying hens using generalized linear mixed models. Poult. Sci. 93:810–817. doi:10.3382/ps.2013-03638

Bernardo, R. 2010. Breeding for quantitative traits in plants. 2nd ed. Stemma Press, Woodbury, MN.

Bevan, M.W., and C. Uauy. 2013. Genomics reveals new landscapes for crop improvement. Genome Biol. 14:206. doi:10.1186/gb-2013-14-6-206

Boer, M.P., D. Wright, L. Feng, D.W. Podlich, L. Luo, M. Cooper, and F.A. van Eeuwijk. 2007. A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. Genetics 177:1801–1813. doi:10.1534/genetics.107.071068

Bolker, B.M., M.E. Brooks, C.J. Clark, S.W. Geange, J.R. Poulsen, M.H.H. Stevens, and J.-S.S. White. 2009. Generalized linear mixed models: A practical guide for ecology and evolution. Trends Ecol. Evol. 24:127–135. doi:10.1016/j.tree.2008.10.008

Breseghello, F., and A.S. Coelho. 2013. Traditional and modern plant breeding methods with examples in rice (*Oryza sativa* L.). J. Agric. Food Chem. 61:8277–8286. doi:10.1021/jf305531j

Burgueño, J., G. de los Campos, K. Weigel, and J. Crossa. 2012. Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. Crop Sci. 52:707–719. doi:10.2135/cropsci2011.06.0299

Costet, L., L. Le Cunff, S. Royaert, L.M. Raboin, C. Hervouet, L. Toubi, et al. 2012. Haplotype structure around *Bru*1 reveals a narrow genetic basis for brown rust resistance in modern sugarcane cultivars. Theor. Appl. Genet. 125:825–836. doi:10.1007/s00122-012-1875-x

Cox, M., D. Hogarth, and P. Hansen. 1994. Breeding and selection for high early season sugar content in a sugarcane (*Saccharum* spp. hybrids) improvement program. Aust. J. Agric. Res. 45:1569–1575. doi:10.1071/AR9941569

Crossa, J., Y. Beyene, S. Kassa, P. Pérez, J.M. Hickey, C. Chen et al. 2013. Genomic prediction in maize breeding populations with genotyping-by-sequencing. G3 3:1903–1926. doi:10.1534/g3.113.008227

Dal-Bianco, M., M.S. Carneiro, C.T. Hotta, R.G. Chapola, H.P. Hoffmann, A.A. Garcia, and G.M. Souza. 2012. Sugarcane improvement: How far can we go? Curr. Opin. Biotechnol. 23:265–270. doi:10.1016/j.copbio.2011.09.002

Daugrois, J.H., L. Grivet, D. Roques, J.Y. Hoarau, H. Lombard, J.C. Glaszmann, and A. D'Hont. 1996. A putative major gene for rust resistance linked with a RFLP marker in sugarcane cultivar 'R570.' Theor. Appl. Genet. 92:1059–1064. doi:10.1007/BF00224049

De Silva, N.H., L. Gea, and R. Lowe. 2014. Genetic analysis of resistance to *Pseudomonas syringae* pv. *actinidiae* (Psa) in a kiwifruit progeny test: An application of generalised linear mixed models (GLMMs). SpringerPlus 3:547. doi:10.1186/2193-1801-3-547

de Vries, S.C., G.W.J. van de Ven, M.K. van Ittersum, and K.E. Giller. 2010. Resource use efficiency and environmental performance of nine major biofuel crops, processed by first-generation conversion techniques. Biomass Bioenergy 34:588–601. doi:10.1016/j.biombioe.2010.01.001

D'Hont, A. 2005. Unravelling the genome structure of polyploids using FISH and GISH: Examples in sugarcane and banana. Cytogenet. Genome Res. 109:27–33. doi:10.1159/000082378

D'Hont, A., and J.C. Glaszmann. 2001. Sugarcane genome analysis with molecular markers: A first decade research. Proc. Int. Soc. Sugar-Cane Technol. 24:556–559.

D'Hont, A., D. Ison, K. Alix, C. Roux, and J.C. Glaszmann. 1998. Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. Genome 41:221–225. doi:10.1139/g98-023

Edwards, D., J. Batley, and R.J. Snowdon. 2013. Accessing complex crop genomes with next-generation sequencing. Theor. Appl. Genet. 126:1–11. doi:10.1007/s00122-012-1964-x

Eksteen, A., A. Singels, and S. Ngxaliwe. 2014. Water relations of two contrasting sugarcane genotypes. Field Crops Res. 168:86–100. doi:10.1016/j.fcr.2014.08.008

Falconer, D.S., and T.F.C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Longman, Harlow, UK.

FAO. 2014. FAOSTAT. [Database.] FAO, Rome. http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#anco (accessed 10 Nov. 2014).

Gallacher, D.J. 1997. Evaluation of sugarcane morphological descriptors using variance component analysis. Aust. J. Agric. Res. 48:769–774. doi:10.1071/A96062

Garsmeur, O., C. Charron, S. Bocs, V. Jouffe, S. Samain, A. Couloux, et al. 2011. High homologous gene conservation despite extreme autopolyploid redundancy in sugarcane. New Phytol. 189:629–642. doi:10.1111/j.1469-8137.2010.03497.x

Gerbens-Leenes, W., A.Y. Hoekstra, and T.H. van der Meer. 2009. The water footprint of bioenergy. Proc. Natl. Acad. Sci. 106:10219–10223. doi:10.1073/pnas.0812619106

Gianola, D., and J. Foulley. 1983. Sire evaluation for ordered categorical data with a threshold model. Genet. Sel. Evol. 15:201–224. doi:10.1186/1297-9686-15-2-201

Glynn, N.C., C. Laborde, R.W. Davidson, M.S. Irey, B. Glaz, A. D'Hont, and J.C. Comstock. 2013. Utilization of a major brown rust resistance gene in sugarcane breeding. Mol. Breed. 31:323–331. doi:10.1007/s11032-012-9792-x

Goldemberg, J. 2007. Ethanol for a sustainable energy future. Science 315:808–810. doi:10.1126/science.1137013

Gouy, M., D. Luquet, L. Rouan, J.-F. Martiné, A. Thong-Chane, L. Costet, et al. 2015. Site and *Saccharum spontaneum* introgression level drive sugarcane yield component traits and their impact on sucrose yield in contrasted radiation and thermal conditions in La Réunion. Field Crops Res. 171:99–108. doi:10.1016/j.fcr.2014.11.002

Gouy, M., Y. Rousselle, D. Bastianelli, P. Lecomte, L. Bonnal, D. Roques, et al. 2013. Experimental assessment of the accuracy of genomic selection in sugarcane. Theor. Appl. Genet. 126:2575–2586. doi:10.1007/s00122-013-2156-z

Grivet, L., and P. Arruda. 2002. Sugarcane genomics: Depicting the complex genome of an important tropical crop. Curr. Opin. Plant Biol. 5:122–127. doi:10.1016/S1369-5266(02)00234-0

Ha, S., P.H. Moore, D. Heinz, S. Kato, N. Ohmido, and K. Fukui. 1999. Quantitative chromosome map of the polyploid *Saccharum spontaneum* by multicolor fluorescence in situ hybridization and imaging methods. Plant Mol. Biol. 39:1165–1173. doi:10.1023/A:1006133804170

Henderson, C.R. 1984. Applications of linear models in animal breeding. Univ. of Guelph, Guelph, ON, Canada.

Hoarau, J.-Y., L. Grivet, B. Offmann, L.M. Raboin, J.P. Diorflar, J. Payet, et al. 2002. Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.): II. Detection of QTLs for yield components. Theor. Appl. Genet. 105:1027–1037. doi:10.1007/s00122-002-1047-5

Hogarth, D. 1971. Quantitative inheritance studies in sugar-cane: II. Correlations and predicted responses to selection. Aust. J. Agric. Res. 22:103–109. doi:10.1071/AR9710103

Hogarth, D.M., C.C. Ryan, and P.W.J. Taylor. 1993. Quantitative inheritance of rust resistance in sugarcane. Field Crops Res. 34:187–193. doi:10.1016/0378-4290(93)90006-9

Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martinez. 2003. Estimating and interpreting heritability for plant breeding: An update. Plant Breed. Rev. 22:9–113.

Hoy, J.W., and C.A. Hollier. 2009. Effect of brown rust on yield of sugarcane in Louisiana. Plant Dis. 93:1171–1174. doi:10.1094/PDIS-93-11-1171

Irvine, J.E. 1999. *Saccharum* species as horticultural classes. Theor. Appl. Genet. 98:186–194. doi:10.1007/s001220051057

Jackson, P.A., D. Horsley, J. Foreman, D.M. Hogarth, and A.W. Wood. 1991. Genotype × environment (GE) interaction in sugarcane variety trials in the Herbert. Proc. Aust. Soc. Sugar Cane Technol. 13:103–109.

Jamoza, J.E., J. Owuoche, O. Kiplagat and W. Opile. 2014. Broad-sense heritability estimation and correlation among sugarcane (*Saccharum* spp. hybrids) yield and some agronomic traits in western Kenya. Int. J. Agric. Policy Res. 2:016–025.

Kelly, A.M., A.B. Smith, J.A. Eccleston, and B.R. Cullis. 2007. The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. Crop Sci. 47:1063–1070. doi:10.2135/cropsci2006.08.0540

Le Cunff, L., O. Garsmeur, L.M. Raboin, J. Pauquet, H. Telismart, A. Selvi, et al. 2008. Diploid/polyploid syntenic shuttle mapping and haplotype-specific chromosome walking toward a rust resistance gene (*Bru*1) in highly polyploid sugarcane ($2n \sim 12x \sim 115$). Genetics 180:649–660. doi:10.1534/genetics.108.091355

Lin, J.F., R.K. Chen, and Y.Q. Lin. 1993. The inheritance of sugar characters in sugarcane. J. Fujian Agric. Coll. 22:392–397.

Macrelli, S., M. Galbe, and O. Wallberg. 2014. Effects of production and market factors on ethanol profitability for an integrated first and second generation ethanol plant using the whole sugarcane as feedstock. Biotechnol. Biofuels 7:26–42. doi:10.1186/1754-6834-7-26

Mahon, J.D. 1983. Limitations to the use of physiological variability in plant breeding. Can. J. Plant Sci. 63:11–21. doi:10.4141/cjps83-002

Malosetti, M., J.M. Ribaut, and F.A. van Eeuwijk. 2013. The statistical analysis of multi-environment data: Modeling genotype-by-environment interaction and its genetic basis. Front. Physiol. 4:44. doi:10.3389/fphys.2013.00044

Mancini, M.C., D.C. Leite, D. Perecin, M.A.P. Bidóia, M.A. Xavier, M.G.A. Landell, and L.R. Pinto. 2012. Characterization of the genetic variability of a sugarcane commercial cross through yield components and quality parameters. Sugar Tech 14:119–125. doi:10.1007/s12355-012-0141-5

Matsuoka, S., A.J. Kennedy, E.G.D. dos Santos, A.L. Tomazela, and L.C.S. Rubio. 2014. Energy cane: Its concept, development, characteristics, and prospects. Adv. Bot. 2014:597275 . doi:10.1155/2014/597275

McFarlane, K., S.A. McFarlane, D. Moodley, and R.S. Rutherford. 2006. Fungicide trials to determine the effect of brown rust on the yield of sugarcane variety N29. Proc. S. Afr. Sugar Technol. Assoc. 79:297–300.

Meyer, K. 2009. Factor-analytic models for genotype × environment type problems and structured covariance matrices. Genet. Sel. Evol. 41:21. doi:10.1186/1297-9686-41-21

Muir, W.M. 2007. Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. J. Anim. Breed. Genet. 124:342–355. doi:10.1111/j.1439-0388.2007.00700.x

Naik, S.N., V.V. Goud, P.K. Rout, and A.K. Dalai. 2010. Production of first and second generation biofuels: A comprehensive review. Renew. Sustain. Energy Rev. 14:578–597. doi:10.1016/j.rser.2009.10.003

Oloriz, M.I., V. Gil, L. Rojas, O. Portal, Y. Izquierdo, E. Jiménez, and M. Hofte. 2012. Sugarcane genes differentially expressed in response to *Puccinia melanocephala* infection: Identification and transcript profiling. Plant Cell Rep. 31:955–969. doi:10.1007/s00299-011-1216-6

Oloriz, M.I., V. Gil, L. Rojas, N. Veitía, M. Höfte, and E. Jiménez. 2011. Selection and characterisation of sugarcane mutants with improved resistance to brown rust obtained by induced mutation. Crop Pasture Sci. 62:1037–1044. doi:10.1071/CP11180

Parco, A.S., M.C. Avellaneda, A.H. Hale, J.W. Hoy, M.J. Pontif, K.A. Gravois, et al. 2014. Frequency and distribution of the brown rust resistance gene *Bru*1 and implications for the Louisiana sugarcane breeding programme. Plant Breed. 133:654–659. doi:10.1111/pbr.12186

Pastina, M.M., M. Malosetti, R. Gazaffi, M. Mollinari, G.R. Margarido, K.M. Oliveira, et al. 2012. A mixed model QTL analysis for sugarcane multiple-harvest-location trial data. Theor. Appl. Genet. 124:835–849. doi:10.1007/s00122-011-1748-8

Payne, R.W., D.A. Murray, S.A. Harding, D.B. Baird, and D.M. Soutar. 2009. GenStat for Windows. 12th ed. VSN Int., Hemel Hempstead, UK.

Piepho, H. 1998. Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance–covariance structures. Theor. Appl. Genet. 97:195–201. doi:10.1007/s001220050885

Piepho, H.P., and J. Möhring. 2007. Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881–1888. doi:10.1534/genetics.107.074229

Piepho, H.P., J. Möhring, A.E. Melchinger, and A. Büchse. 2008. BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161:209–228. doi:10.1007/s10681-007-9449-8

Pinto, L.R., A.A.F. Garcia, M.M. Pastina, L.H.M. Teixeira, J.A. Bressiani, E.C. Ulian, et al. 2010. Analysis of genomic and functional RFLP derived markers associated with sucrose content, fiber and yield QTLs in a sugarcane (*Saccharum* spp.) commercial cross. Euphytica 172:313–327. doi:10.1007/s10681-009-9988-2

Piperidis, G., N. Piperidis, and A. D'Hont. 2010. Molecular cytogenetic investigation of chromosome composition and transmission in sugarcane. Mol. Genet. Genomics 284:65–73. doi:10.1007/s00438-010-0546-3

Prohens, J. 2011. Plant breeding: A success story to be continued thanks to the advances in genomics. Front. Plant Sci. 2:51. doi:10.3389/fpls.2011.00051

Purdy, L.H., L.J. Liu, and J.L. Dean. 1983. Sugarcane rust, a newly important disease. Plant Dis. 67:1292–1296. doi:10.1094/PD-67-1292

Raboin, L.M., K.M. Oliveira, L. Lecunff, H. Telismart, D. Roques, M. Butterfield, et al. 2006. Genetic mapping in sugarcane, a high polyploid, using bi-parental progeny: Identification of a gene controlling stalk colour and a new rust resistance gene. Theor. Appl. Genet. 112:1382–1391. doi:10.1007/s00122-006-0240-3

Racedo, J., M.F. Perera, R. Bertani, C. Funes, V. González, M.I. Cuenya, et al. 2013. *Bru*1 gene and potential alternative sources of resistance to sugarcane brown rust disease. Euphytica 191:429–436. doi:10.1007/s10681-013-0905-3

Raid, R.N., and J.C. Comstock. 2000. Common rust. In: P. Rott et al., editors, A guide to sugarcane diseases. CIRAD and ISSCT, Montpellier, France. p. 85–89.

Ram, B., B.S. Chaudhary, and D.K. Yadav. 1997. General and specific selection indices for single stool stages of selection in sugarcane. Euphytica 95:39–44. doi:10.1023/A:1002965924609

Ramburan, S. 2014. A multivariate illustration and interpretation of non-repeatable genotype × environment interactions in sugarcane. Field Crops Res. 157:57–64. doi:10.1016/j.fcr.2013.12.009

Ramburan, S., M. Zhou, and M.T. Labuschagne. 2012. Investigating test site similarity, trait relations and causes of genotype × environment interactions of sugarcane in the midlands region of South Africa. Field Crops Res. 129:71–80. doi:10.1016/j.fcr.2012.01.017

Ramdoyal, K., S. Sullivan, L.C.Y. Lim Shin Chong, G.H. Badaloo, S. Sautally, and R. Domaingue. 2000. The genetics of rust resistance in sugarcane seedling populations. Theor. Appl. Genet. 100:557–563. doi:10.1007/s001220050073

Revelle, W. 2014. PSYCH: Procedures for psychological, psychometric, and personality research. R Found. Stat. Comput., Vienna. http://cran.r-project.org/package=psych (accessed 12 Nov. 2013).

Ryan, C.C., and B.T. Egan. 1989. Rust. In: C. Ricaud and B.T. Egan, editors, Diseases of sugarcane: Major diseases. Elsevier Science Publ., Amsterdam. p. 189–210.

Sadras, V.O., G.J. Rebetzke, and G.O. Edmeades. 2013. The phenotype and the components of phenotypic variance of crop traits. Field Crops Res. 154:255–259. doi:10.1016/j.fcr.2013.10.001

Saini, J.K., R. Saini, L. Tewari. 2015. Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: Concepts and recent developments. 3 Biotech. 5:337–353. doi:10.1007/s13205-014-0246-5

Schwarz, G. 1978. Estimating the dimension of a model. Ann. Stat. 6:461–464. doi:10.1214/aos/1176344136

Searle, S.R., G. Casella, and C.E. McCulloch. 1992. Variance components. John Wiley & Sons, New York.

Smith, A.B., B.R. Cullis, and R. Thompson. 2005. The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. J. Agric. Sci. 143:449–462. doi:10.1017/S0021859605005587

Smith, A.B., J.K. Stringer, X. Wei, and B.R. Cullis. 2007. Varietal selection for perennial crops where data relate to multiple harvests from a series of field trials. Euphytica 157:253–266. doi:10.1007/s10681-007-9418-2

Soomro, A.F., S. Junejo, A. Ahmed, and M. Aslam. 2006. Evaluation of different promising sugarcane varieties for some quantitative and qualitative attributes under Thatta (Pakistan) conditions. Int. J. Agric. Biol. 8:195–197.

Sordi, R.A., H. Arizono, and S. Matsuoka. 1988. Indicadores de herdabilidade e avaliação da resistência de clones RB à ferrugem da cana-de-açúcar. Bras. Acucareiro 106:18–23.

State of Sao Paulo Sugarcane, Sugar and Alcohol Growers Council. 2006. Manual de instruções. 5th ed. CONSECANA-SP, Piracicaba, SP, Brazil.

Tai, P.Y.P., J.D. Miller, and J.L. Dean. 1981. Inheritance of resistance to rust in sugarcane. Field Crops Res. 4:261–268. doi:10.1016/0378-4290(81)90077-0

Taylor, P.W.J., B.J. Croft, and C.C. Ryan. 1986. Studies into the effect of sugarcane rust (*Puccinia melanocephala*) on yield. Proc. Int. Soc. Sugar-Cane Technol. 9:411–419.

Thompson, R. 1979. Sire evaluation. Biometrics 35:339–353. doi:10.2307/2529955

Thompson, R., B. Cullis, A. Smith, and A. Gilmour. 2003. A sparse implementation of the average information algorithm for factor analytic and reduced rank variance models. Aust. N.Z. J. Stat. 45:445–459. doi:10.1111/1467-842X.00297

Trust, B. 2014. Estimation in generalized linear models with random effects. Biometrika 78:719–727.

Waclawovsky, A.J., P.M. Sato, C.G. Lembke, P.H. Moore, and G.M. Souza. 2010. Sugarcane for bioenergy production: An assessment of yield and regulation of sucrose content. Plant Biotechnol. J. 8:263–276. doi:10.1111/j.1467-7652.2009.00491.x

Welham, S.J., B.J. Gogel, A.B. Smith, R. Thompson, and B.R. Cullis. 2010. A comparison of analysis methods for late-stage variety evaluation trials. Aust. N.Z. J. Stat. 52:125–149. doi:10.1111/j.1467-842X.2010.00570.x

Wolc, A., C. Stricker, J. Arango, P. Settar, J.E. Fulton, N.P. O'Sullivan, et al. 2011. Breeding value prediction for production traits in layer chickens using pedigree or genomic relationships in a reduced animal model. Genet. Sel. Evol. 43:5. doi:10.1186/1297-9686-43-5

Zhang, Z., E. Ersoz, C.Q. Lai, R.J. Todhunter, H.K. Tiwari, M.A. Gore, et al. 2010. Mixed linear model approach adapted for genome-wide association studies. Nat. Genet. 42:355–360. doi:10.1038/ng.546