

Genetic evaluation of *Theobroma grandiflorum* hybrid families for fruit yield and resistance to witches' broom

Saulo Fabrício da Silva Chaves (✉ saulofabricioagro@gmail.com)

UFV: Universidade Federal de Vicosa <https://orcid.org/0000-0002-0694-1798>

Luiz Antônio dos Santos Dias

Universidade Federal de Vicosa

Rodrigo Silva Alves

instituto nacional de ciencia e tecnologia

Rafael Moysés Alves

Embrapa Amazonia Oriental

Jeniffer Santana Pinto Coelho Evangelista

Universidade Federal de Vicosa

Kaio Olimpio das Graças Dias

Universidade Federal de Vicosa

Research Article

Keywords: quantitative genetics, mixed models, repeated measures, repeatability, fruit tree breeding

Posted Date: August 1st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1848061/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Genetic evaluation of *Theobroma grandiflorum* hybrid families for fruit yield and resistance to witches' broom

Saulo Fabrício da Silva Chaves^{a,*}, Luiz Antônio dos Santos Dias^{a,*}, Rodrigo Silva Alves^{a,b}, Rafael Moysés Alves^c, Jeniffer Santana Pinto Coelho Evangelista^a, Kaio Olimpio das Graças Dias^a

^a Universidade Federal de Viçosa, Viçosa, 36570-900, MG, Brazil

^b Instituto Nacional de Ciência e Tecnologia do Café, Viçosa, 36570-900, MG, Brazil

^c Embrapa Amazônia Oriental, Belém, 66095-903, PA, Brazil

* Corresponding authors:

Saulo Fabrício da Silva Chaves: saulo.chaves@ufv.br

Luiz Antônio dos Santos Dias: lasdias@ufv.br

Statements and Declarations

Acknowledgments

We thank all the field workers who work in the *Theobroma grandiflorum* breeding program of Embrapa Amazônia Oriental. We acknowledge the financial support from National Institute of Science and Technology of Coffee (INCT Café), Minas Gerais State Agency for Research and Development (FAPEMIG), Brazilian National Council for Scientific and Technological Development (CNPq) and Coordination for the Improvement of Higher Education Personnel (CAPES) - Finance Code 001.

Declarations

All authors contributed to the study and conception and desing: **Saulo Chaves**: Formal analysis, Writing – Original draft, Visualization **Luiz Dias**: Supervision, Resources, Writing – Review & Editing **Rodrigo Alves**: Conceptualization, Methodology, Writing – Review & Editing **Rafael Alves**: Data curation, Validation, Writing – Review & Editing **Jeniffer Evangelista**: Conceptualization, Writing – Review & Editing **Kaio Dias**: Validation, Writing – Review & Editing

Funding

This research was supported by National Institute of Science and Technology of Coffee (INCT Café), Minas Gerais State Agency for Research and Development (FAPEMIG), Brazilian National Council for Scientific

31 and Technological Development (CNPq) and Coordination for the Improvement of Higher Education
32 Personnel (CAPES) - Finance Code 001

33 **Conflicts of interests**

34 There are no conflicts of interests in this article

35 **Availability of data and material**

36 The datasets generated during and/or analysed during the current study are available from the corresponding
37 author on reasonable request.

38 **Code availability**

39 The codes are available from the corresponding authors on reasonable request

Genetic evaluation of *Theobroma grandiflorum* hybrid families for fruit yield and resistance to witches' broom

Abstract: *Theobroma grandiflorum* is a perennial fruit-bearing species with a long breeding cycle (~12 years). In multi-harvest data analyses, one frequently observes heteroscedasticity and genetic/statistical imbalance, which requires adequate statistical methods. This study aimed to compare the adequacy of multi-harvest models with homogenous and heterogeneous residual variance and to define the optimal number of harvests for an accurate genetic selection of *T. grandiflorum* families with high fruit yield persistence and resistance to witches' broom. We used a generalized linear mixed model to analyse the resistance to witches' broom. The data refers to an unbalanced trial with 34 triple-cross hybrid families, arranged in a randomized complete block design. We evaluated these families throughout nine consecutive harvests regarding fruit yield and resistance to witches' broom. The variance components, genetic parameters and genotypic values were estimated/predicted through residual maximum likelihood/best linear unbiased prediction methodologies. The model with heterogeneous residuals had a better fit than the homogeneous model for the fruit yield trait, according to Akaike information criterion. The results attest to the genetic variability in the evaluated population. To select the best individual for the next breeding cycle, three and two harvests are enough for selecting accurately for fruit yield and resistance to witches' broom, respectively. For commercial release, nine harvests are necessary for both traits. We identified ten high-performance families regarding fruit yield persistence and resistance to witches' broom.

Keywords: quantitative genetics, mixed models, repeated measures, repeatability, fruit tree breeding

1 Introduction

The biodiversity-rich Brazilian Amazon holds a large variety of fruit-bearing species. About 44% of Brazilian native fruit-bearing diversity is within Amazonian territory (Neves et al., 2012). This is an indication of the enormous potential for the sustainable commercial and scientific exploitation of these species. An example of this statement is the fruit-bearing species *Theobroma grandiflorum* [Willd. (ex. Spreng) Schum.]. Since the 1980s, *T. grandiflorum* is the target species of a breeding program at Embrapa (Brazilian Agricultural Research Corporation). Embrapa's breeding program aims to develop highly productive genotypes, with resistance to *Moniliophthora perniciosa* (Stahel) Aime & Phillips-Mora, the etiological agent of the witches' broom, a disease of great economic importance in the *Theobroma* genus, since it has the potential of making production unfeasible (Alves et al., 2020; Patrocínio et al., 2017).

T. grandiflorum has a long juvenile period (about three years) and it takes about eight years for it to reach full production maturity. For that reason, the species' breeding cycle is long (about 12 years), hindering the release of new cultivars (Souza et al., 2002). Hence, it is of paramount importance to determine the optimal number of harvests for an accurate genetic selection. To reduce the length of the breeding cycle is to increase the genetic gains per time unit (Chaves et al., 2021; Cobb et al., 2019). In perennial crops breeding, there is an additional challenge, once data often originate from repeated measures in the same plants. When using the appropriated model, one can leverage this data to estimate the repeatability coefficient and the genotype-by-harvest interaction (GHI). Accounting for GHI is important to comprehend the temporal variation of the genotypes' performance (Atlin et al., 2017). When facing GHI, the breeder may prioritize genotypes with high persistence, i.e., genotypes that keep a high performance regardless of the environmental changes (Rocha et al., 2018). The concept of persistence was originally developed for perennial grasses (McDonagh et al., 2016; Wilkins & Humphreys, 2003), but can be extended to perennial fruit-bearing trees such as *T. grandiflorum*.

With multi-harvest data, we can also identify the residual variance structure throughout the harvests (Wolfinger, 1996). The variance can be either homogeneous or heterogeneous. When considering homoscedasticity, the breeder assumes that the non-controllable effects (residuals) were similar throughout the harvests. Naturally, heteroscedasticity indicates the opposite (Gezan & Carvalho, 2018). Considering an eventual residual heteroscedasticity may be necessary, given the environmental changes that occur over the years (Faveri et al., 2015; Verbyla et al., 2021). This fact is especially important for perennial species such as *T. grandiflorum*, which can keep producing for several decades.

In this study, our objectives were to compare the adequacy of multi-harvest models with homogenous and heterogeneous residual variance and to define the optimal number of harvests for an accurate genetic selection of *T. grandiflorum* genotypes with high fruit yield persistence and resistance to witches' broom.

2 Material and Methods

2.1 Experimental conditions

The trial was established in a rural property in the Tomé-Açu municipality (latitude 02° 24' 05.6" S; longitude 48° 00' 08.8" W; altitude 45 m above sea level), Pará State, Brazil. The climate of the region is the type B4rA'a', according to Thornthwaite classification (Moraes et al., 2020). The soil is an Oxisol, with medium texture, low natural fertility, and deep with good drainage. The average temperature, humidity and rainfall in the evaluated period were 27 °C, 85.5% and 2.729 mm, respectively. These data were obtained *in situ* by a local weather station.

During the first years in the field, *T. grandiflorum* needs partial shading. For this reason, the plants were intercropped with adult black pepper plants (*Piper nigrum* L.), at 6 x 4 m spacing. The cultural practices followed the recommendation for the species.

2.2 Plant material

The evaluated hybrid families were acquired by triple crossing: at first, we obtained the simple hybrids by crossing clones with complementary features from the *T. grandiflorum* Germplasm Bank at Embrapa Amazônia Oriental. The hybrids were evaluated for 12 consecutive years. After evaluation between and within hybrid families, we cloned the selected individuals and crossed them with four commercial cultivars: BRS Coari (174), BRS Codajás (186), BRS Manacapuru (215) and BRS Belém (286). In sum, the crosses followed the scheme $(C_1 \times C_2) \times C_3$, i.e., three different clones (C). Thus, variability is the result of heterogeneity of the non-endogamous parents.

In total, 34 hybrid families were evaluated in a randomized complete block design, with five to ten replicates and three plants per plot. The number of replicates varied according to the availability of seeds. From the third year after establishment onwards, we evaluated the fruit yield per plant (kg), the product of the number of fruits per plant by the average fruit weight of the same plant. We also evaluated the resistance to witches' broom disease, checking for the presence (or absence) of symptoms. We let the disease infection

occur naturally, so the plants manifest the mechanisms of horizontal resistance. Those evaluations were performed throughout nine consecutive harvests.

2.3 Statistical analyses

We used the REML/BLUP (Henderson, 1975; Patterson & Thompson, 1971) methodology for estimating the variance components and predicting the genotypic values. The linear mixed model associated with the evaluation of hybrid families in a randomized complete block design, three plants per plot and multi-harvest was:

$$\mathbf{y} = \mu \mathbf{1} + \mathbf{X}_1 \mathbf{h} + \mathbf{X}_2 \mathbf{b} + \mathbf{Z}_1 \mathbf{g} + \mathbf{Z}_2 \mathbf{p} + \mathbf{Z}_3 \mathbf{gh} + \mathbf{Z}_4 \mathbf{i} + \mathbf{e} \quad (1);$$

Where $\mathbf{y}^{(\eta \times 1)}$ is the vector of phenotypic data, μ is the intercept, $\mathbf{h}^{(\gamma \times 1)}$ is the vector of the fixed effects of harvests, $\mathbf{b}^{(\theta \times 1)}$ is the vector of the fixed effects of replicates, $\mathbf{g}^{(\tau \times 1)}$ is the vector of the random effects of hybrid families ($\mathbf{g} \sim MVN[0, \sigma_g^2 \mathbf{I}]$), $\mathbf{p}^{(\delta \times 1)}$ is the vector of the random permanent plot effects ($\mathbf{p} \sim MVN[0, \sigma_p^2 \mathbf{I}]$), $\mathbf{gh}^{(\tau \gamma \times 1)}$ is the vector of the random effects of GHI ($\mathbf{gh} \sim MVN[0, \sigma_{gh}^2 \mathbf{I}]$), $\mathbf{i}^{(\eta \times 1)}$ is the vector of the permanent plant effects ($\mathbf{i} \sim MVN[0, \sigma_i^2 \mathbf{I}]$), and $\mathbf{e}^{(\eta \times 1)}$ is the vector of the random residual effects ($\mathbf{e} \sim MVN[0, \mathbf{R}]$). $\mathbf{1}^{(\eta \times 1)}$ is a vector of ones and the capital letters represent the incidence matrices of the respective effects. η , γ , θ , τ and δ are the number of observations, harvests, replicates, hybrid families and plots, respectively. σ_x^2 is the variance of the x^{th} random effect. $\mathbf{I}$ is an identity matrix. $\mathbf{R}$ is the residual covariance matrix, which was modelled as an identity of variances, $\mathbf{R} = \sigma_e^2 \mathbf{I}$, in the homogeneous model; and block diagonal, $\mathbf{R} = \bigoplus_1^{\gamma} \sigma_e^2 \mathbf{I}_{\gamma}$, where $\bigoplus$ is the direct sum, in the heterogeneous model. The plot and plant effects are permanent because there was no new casualization from one year to another, i.e., the genotypes are in the same position throughout all harvests.

From eq. 1, we derived the generalized linear mixed model (GLMM) for the evaluation of witches' broom, a discrete random trait. This model is suitable because the resistance was evaluated through the presence of symptoms, in which we assigned "0" for absence and "1" for presence. Then, we used a binomial distribution with the logit link function [$logit = \mu / (1 - \mu)$, where μ is the mean at the data scale].

We tested the goodness-of-fit of the models with different residual variance structures (Eq.1) using the Akaike Information Criterion (AIC) (Akaike, 1974):

$$AIC = -2 \log L + 2p;$$

where L is the maximum point of the residual likelihood function and p is the number of estimated parameters.

We estimated the 95% confidence intervals for each variance component [$CI(\sigma_x^2)_{0.95}$]:

$$CI(\sigma_x^2)_{0.95} = \sigma_x^2 \pm 1.96 SE_x;$$

where σ_x^2 is the variance component of the x^{th} effect, SE_x is the standard error of σ_x^2 and 1.96 is the tabulated value for the Student's t-test for infinite degrees of freedom. If $CI(\sigma_x^2)_{0.95} \leq 0$, we assumed that the variance component did not differ from zero.

From the variance components, we estimated the individual phenotypic variance (σ_f^2), the generalized heritability (H_g^2) (Cullis et al., 2006), the repeatability coefficient (ρ), the coefficient of determination of the permanent plot (c_p^2), GHI (c_{gh}^2), permanent plant (c_i^2) and residual (c_e^2) effects, the type B genotypic correlations across harvests (r_{gh}) and the accuracy (r), given by the following equations:

$$\sigma_f^2 = \sigma_g^2 + \sigma_{gh}^2 + \sigma_p^2 + \sigma_i^2 + \sigma_e^2;$$

$$H_g^2 = 1 - \frac{V(\overline{\Delta_g})}{2\sigma_g^2}, \text{ where } V(\overline{\Delta_g}) \text{ is the average pairwise prediction error variance};$$

$$\rho = \frac{\sigma_g^2 + \sigma_i^2}{\sigma_f^2};$$

$$c_p^2 = \frac{\sigma_p^2}{\sigma_f^2};$$

$$c_{gh}^2 = \frac{\sigma_{gh}^2}{\sigma_f^2};$$

$$c_i^2 = \frac{\sigma_i^2}{\sigma_f^2};$$

$$c_e^2 = \frac{\sigma_e^2}{\sigma_f^2};$$

$$r_{gh} = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{gh}^2};$$

$$r = \sqrt{1 - \frac{PEV}{\sigma_g^2}}, \text{ where } PEV \text{ is the prediction error variance.}$$

To estimate the optimum number of harvests, we estimated the comparative efficiency (E) and accuracy (r_h) of using h harvests instead of using only one (Resende & Alves, 2020):

$$E = \frac{h}{1 + (h+1)\rho};$$

$$r_h = \sqrt{\frac{h\rho}{h\rho + 1 - \rho}};$$

To select hybrid families with major resilience to the environmental variation that occurs over the years in the field, we estimated their yield persistence (P_t), given by (Rocha et al., 2018):

$$P_t = \frac{1}{\sum_{\gamma=1}^9 [\hat{g}_{t\gamma} - \max(\hat{g}_\gamma)]^2} \bigg/ \sum_{\tau=1}^{34} \left\{ \frac{1}{\sum_{\gamma=1}^9 [\hat{g}_{t\gamma} - \max(\hat{g}_\gamma)]^2} \right\} \quad (2)$$

where $\hat{g}_{\tau\gamma}$ is the genotypic value of the τ -th hybrid family in the γ -th harvest and $\max(\hat{g}_{\gamma})$ is the maximum genotypic value estimated in the γ -th harvest, which represents the ideotype.

At last, we used the additive index (AI) to select hybrids that reunite favourable alleles for resistance to witches' broom and persistent fruit yield:

$$AI = \sum_{\theta=1}^2 CV_{g_{\theta}} \frac{G_{\tau}}{\sigma_{g_{\theta}}},$$

where G_i is the P_{τ} of the τ -th hybrid family for fruit yield and the genotypic value of the τ -th hybrid family for resistance, $\sigma_{g_{\theta}}$ is the genetic standard deviation of the θ -th trait and $CV_{g_{\theta}}$ is the genetic coefficient of variation of the θ -th trait ($CV_{g_{\theta}} = 100[\sigma_{g_{\theta}}/\mu]$) (Ferreira et al., 2021). The equation's numerator differed because we have not estimated P_{τ} for resistance, as the GHI effect was not different from zero (see next section). The direction of selection was positive for fruit yield and negative for resistance (recall that "1" is positive for symptom). All the analyses were done in the R software environment, version 4.2 (R Core Team, 2022), using the ASReml-R package, version 4.1 (Butler et al., 2018).

3 Results

According to AIC, for the yield trait, the model with heterogeneous residual variances between harvests had better fit (Table 1). Therefore, all the results for this trait were acquired from this model.

Table 1. Tested models for the fruit yield trait evaluated in *Theobroma grandiflorum* triple-cross hybrid families and their respective number of parameters, accuracy and Akaike information criterion (AIC) values.

Model	Number of parameters	Accuracy	AIC
Homogeneous residual variances	5	0.79	39847.85
Heterogeneous residual variances	13	0.82	39074.07

The heteroscedastic model particularizes the residual variances for each harvest. Consequently, the phenotypic variance and all the genetic parameters derived from these variance components are also particularized (Table 2).

Table 2. Variance components and genetic parameters for fruit yield (kg plant⁻¹) and resistance to witches' broom evaluated in *Theobroma grandiflorum* triple-cross hybrid families.

Trait	Harvest	Variance component						H_g^2	Coefficient of determination				ρ	r_h
		σ_g^2	σ_{gh}^2	σ_i^2	σ_p^2	σ_e^2	σ_f^2		c_{gh}^2	c_i^2	c_p^2	c_e^2		
Fruit yield (kg plant ⁻¹)	1					70.50	153.98	0.57	0.07	0.32	0.07	0.61	0.28	0.56
	2					64.42	147.90	0.57	0.07	0.33	0.07	0.46		
	3					80.11	163.59	0.53	0.06	0.30	0.07	0.44		
	4					129.21	121.69	0.45	0.05	0.23	0.05	0.49		
	5	13.39*	10.35*	10.83	49.91*	142.69	226.17	0.43	0.05	0.22	0.05	0.61		
	6	± 1.53	± 0.68	± 1.11	± 1.64	263.89	347.37	0.30	0.03	0.14	0.03	0.63		
	7					384.56	467.99	0.23	0.02	0.11	0.02	0.76		
	8					203.60	287.08	0.35	0.04	0.17	0.04	0.82		
	9					206.63	290.11	0.35	0.04	0.17	0.04	0.71		
	Mean	-	-	-	-	-	-	0.42	0.05	0.22	0.05	0.71		
Resistance to witches' broom	1 to 9	0.68* ± 0.29	0.02 ^{ns} ± 0.15	1.13* ± 0.36	0.26 ^{ns} ± 0.30	3.29	5.39	0.59	0.004	0.21	0.05	0.61	0.34	0.97

σ_g^2 is the genotypic variance, σ_{gh}^2 is the genotype-by-harvest interaction (GHI) variance, σ_i^2 is the permanent plant effect variance, σ_p^2 is the permanent plot effect variance,

σ_e^2 is the residual variance, σ_f^2 is the phenotypic variance, H_g^2 is the generalized heritability, c_{gh}^2 , c_i^2 , c_p^2 and c_e^2 are the coefficients of determination of GHI, permanent environment, permanent plot and residual effects; ρ is the repeatability coefficient and r_h is the type B genotypic correlation across harvests.

* and ^{ns}: different and not different from zero, respectively, at the 5% probability level.

All the random effects significantly influenced fruit yield (Table 2). On the other hand, GHI and permanent plot effects did not differ from zero for resistance to witches' broom. This is corroborated by the low values of their respective coefficients of determination. As the GHI was low, the type B genotypic correlation across harvests was high for witches' broom resistance (0.97). The same did not occur for fruit yield (0.56) (Table 2).

In general, the residual variance played an important part in the total phenotypic variation for both traits. This is proved by the high coefficient of determination of the residual effects, 0.67 for witches' broom resistance and from 0.44 to 0.82 for fruit yield. The generalized heritability varied from medium values (0.3 to 0.5) to high values (> 0.5) for fruit yield. In turn, witches' broom resistance had a high heritability (0.59). The resistance's repeatability (0.34) was greater than fruit yield's (0.28).

The differences between the traits' repeatability reflect the divergence between their comparative accuracies and efficiencies (Figure 1). Since the resistance had higher repeatability, it needs fewer harvests than fruit yield to reach an accuracy of 0.7, a satisfactory value when one's goal is selection for cycle advancing, according to Resende and Duarte's (2007) criterion. Indeed, only two harvests are enough for resistance evaluation, whereas fruit yield needs three harvests (Figure 1A). This reflects on the comparative efficiencies, which have decreasing growth rates as more harvests are evaluated (Figure 1B). By comparing the curves of both traits, one can conclude that fruit yield benefits more with increasing evaluated harvests. When one's goal is commercial releasing, Resende and Duarte (2007) suggest 0.9 as the minimum accuracy. Our results indicate that, for both traits, nine evaluations are necessary to reach this value (Figure 1A).

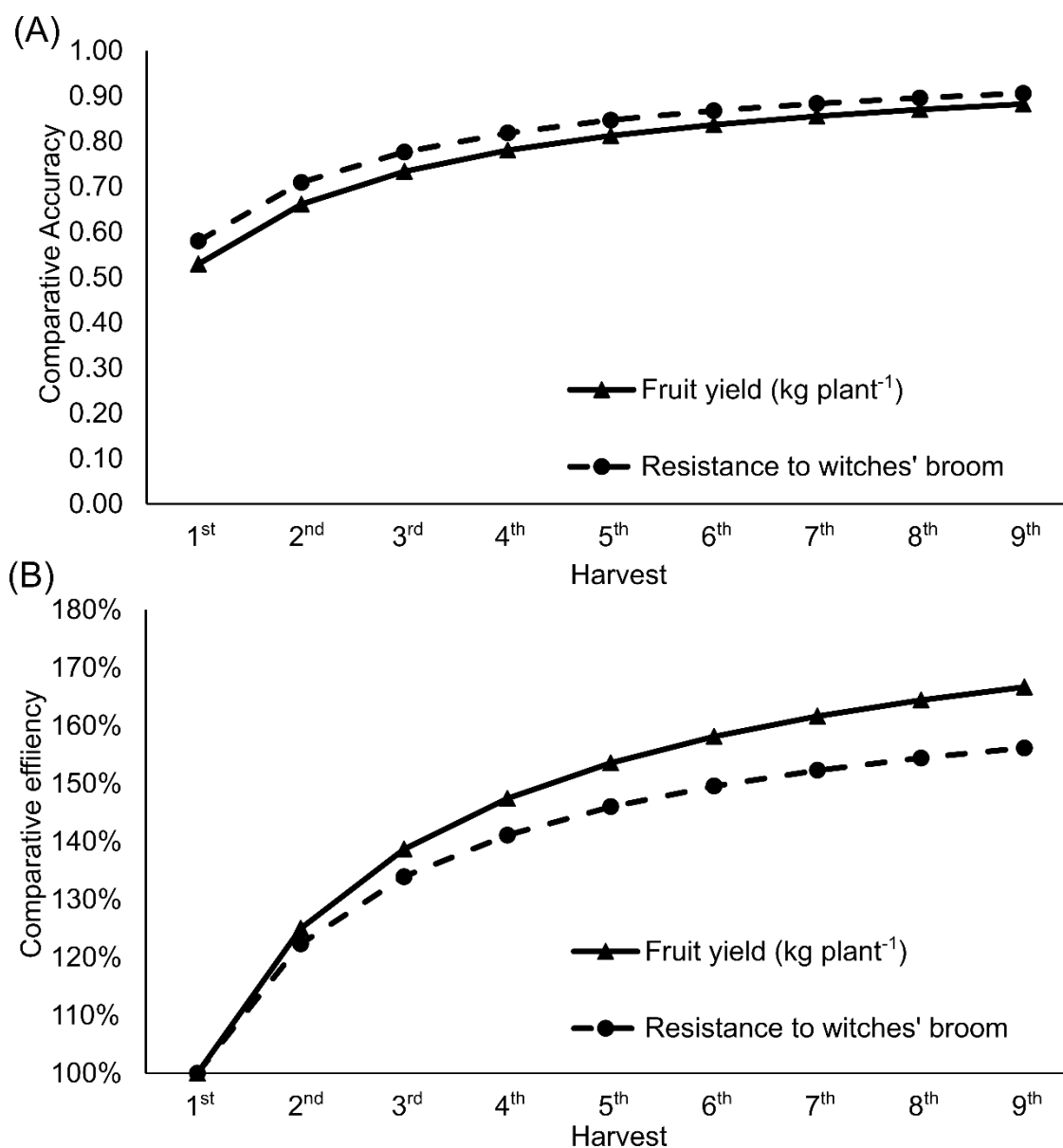


Figure 1. Comparative accuracy (A) and comparative efficiency (B) throughout the harvests for fruit yield and resistance to witches' broom traits evaluated in *Theobroma grandiflorum* triple-cross hybrid families.

The top ten hybrid families were coincident for the fruit yield genotypic value and the persistence (H117, H120, H125, H127, H129, H140, H162, H165 and H169). Selecting these hybrid families will provide a genetic gain of 12.62% in fruit yield (Table 3).

Table 3. Genotypic values ($\hat{g}$), Persistence and Additive Index for fruit yield and resistance to witches' broom evaluated in *Theobroma grandiflorum* triple-cross hybrid families.

Hybrid family	Fruit yield (kg plant ⁻¹)	Resistance to witches' broom	Additive index
---------------	---------------------------------------	------------------------------	----------------

	$\hat{g}$	Persistence	$\hat{g}$	
H117	31.235	3.358	0.0009	11.894
H118	26.510	0.888	0.0005	3.142
H120	30.781	3.074	0.0013	10.883
H121	27.255	1.083	0.0015	3.821
H123	21.040	0.347	0.0009	1.218
H124	27.525	1.125	0.0004	3.985
H125	34.425	33.874	0.0006	120.123
H126	23.912	0.562	0.0012	1.975
H127	30.142	2.438	0.0008	8.634
H128	22.467	0.445	0.0009	1.566
H129	31.113	3.035	0.0008	10.752
H130	28.751	1.522	0.0009	5.387
H131	28.228	1.178	0.0007	4.166
H132	25.902	0.767	0.0008	2.709
H133	25.330	0.714	0.0007	2.523
H134	25.190	0.705	0.0006	2.491
H135	26.902	0.962	0.0004	3.406
H136	24.655	0.590	0.0006	2.084
H137	27.627	1.126	0.0003	3.991
H138	28.158	1.322	0.0013	4.671
H140	31.673	4.541	0.0011	16.088
H143	30.129	2.358	0.0004	8.358
H144	29.631	2.041	0.0010	7.225
H149	27.824	1.274	0.0006	4.511
H150	26.375	0.843	0.0010	2.974
H152	26.358	0.908	0.0010	3.205
H157	27.727	1.053	0.0005	3.727
H161	28.143	1.349	0.0044	4.722
H162	31.899	4.127	0.0043	14.574
H163	28.618	1.554	0.0029	5.471
H165	33.473	10.587	0.0006	37.539
H166	27.118	0.981	0.0008	3.470
H167	29.901	1.978	0.0005	7.008
H169	32.689	7.289	0.0029	25.808

In bold, the top ten hybrid families for both traits, according to the additive index.

Due to the non-significance of the GHI effects for resistance, genotypic values are enough to select the best hybrid families. Since “1” was assigned for symptomatic plants, the lower the hybrid family’s genotypic value, the better. Thus, the hybrid families H118, H124, H135, H136, H137, H143, H149, H157,

H165 and H167 were the top ten in the ranking, and selecting them provides an 11.01% gain for witches' broom resistance.

Using the additive index (AI) to evaluate the hybrid families' performance regarding fruit yield persistence and witches' broom resistance, ten hybrid families stand out: H118, H124, H125, H129, H135, H137, H140, H143, H165 and H167. Five of these hybrid families are within the top ten for persistent fruit yield, whereas seven of them are present among the selected hybrid families for witches' broom resistance. Only two hybrid families (H165 and H143) were selected in all the criteria (fruit yield genotypic value, fruit yield persistence, witches' broom resistance genotypic value and AI) (Figure 2). If one selects the hybrid families indicated by the AI, one's gain in fruit yield will be 6.14% and in witches' broom resistance will be 7.19%.

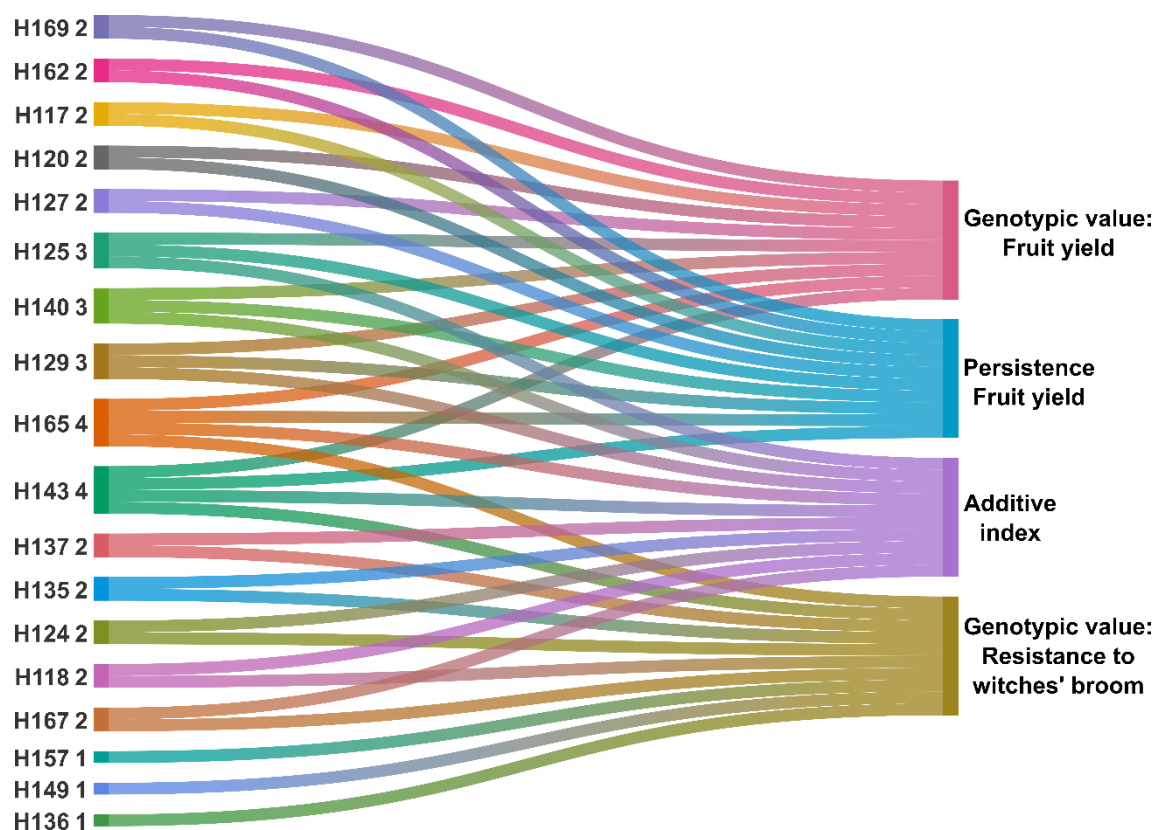


Figure 2. Top ten selected hybrid families (on the left) according to four selection criteria (on the right).

The number beside the hybrid family code represents the number of criteria for which this hybrid family appears in the top ten.

4 Discussion

Phenotyping over time (harvests) or space (locations) implies low probability of homogeneity between residual variances. This is due to the peculiarities of each harvest/location (Edwards & Jannink, 2006; Jaimez et al., 2020). For that reason, we already expected that the heteroscedastic residual variance model would have a better fit. This result has great methodological importance, as it stresses the necessity of considering the heteroscedasticity in multi-harvest data analyses (Faveri et al., 2015; Martínez-García et al., 2017). This is especially relevant in fruit-bearing perennial trees, such as *T. grandiflorum*, for which several measurements are done in the same plant throughout various harvests (Chaves et al., 2021).

The significant genetic variances for both evaluated traits attest to the population's aptitude for selecting the best hybrid families. On the other hand, there is a differential behaviour of each trait regarding GHI. Indeed, the yearly variations have higher effect on fruit yield (significant GHI variance) than witches' broom resistance (non-significant GHI variance), suggesting that fruit yield is more influenced by the environment (Alves et al., 2021).

The non-significance of GHI and the high genotypic correlation across environments indicate that the resistance to witches' broom was stable throughout the evaluated years. From this result, one may presume that i) early selection may be suitable for this trait; and ii) the presence of favourable alleles from different sources contributed to elevating the trait's stability, despite the different pathogen pressures and environmental conditions (Albuquerque et al., 2010; Benjamin et al., 2016; Pereira et al., 2021). Recall that the evaluated hybrid families are a combination of three clones, which supposedly hold resistance alleles. To this day, few studies evaluated in-field witches' broom resistance in *T. grandiflorum* with a suitable genetic/statistical approach. Our study is pioneering in employing a GLMM to analyse this trait.

Witches' broom resistance had a high generalized heritability ($H_g^2 > 0.5$), whereas fruit yield had moderate values for this parameter ($0.15 < H_g^2 < 0.5$) (Resende & Alves, 2020). Fruit yield heritability found herein was somewhat lower than the values usually found in the literature ($H_g^2 > 0.5$) (Alves et al., 2020, 2021). This is probably due to differences in the model: we used a multi-harvest model, whereas the aforementioned authors used the mean across harvests in a single-trait model. Note that from the fourth harvest onwards, generalized heritability values decrease for fruit yield. This is clear evidence of environmental influence. In fact, the different climatic patterns and the increase of witches' broom incidence played a part in this behaviour. The different maturity stages may also be one of the causes.

Like heritability, the repeatability coefficient is also linked with the non-controllable effects and the genetic variability in the population. The lower the repeatability values, the larger the number of

harvests that must be considered for an efficient selection. In our study, we followed the Resende and Duarte's (2007) accuracy classification to propose the optimal number of harvests. Fruit yield must be evaluated for at least three consecutive harvests to provide a reliable selection for cycle advancing (accuracy = 0.7). In *Theobroma* genus, other studies also reached the same conclusion (Carvalho et al., 2002; Mustiga et al., 2018). As it has higher repeatability, the resistance needs fewer harvests (two) to provide the same accuracy. Considering the first three years of the juvenile period, the current selection cycle of *T. grandiflorum* breeding program can be shortened by half. This result has great practical relevance and this study is the first to report this result on resistance to witches' broom.

The evaluation in 12 years (considering the juvenile period) should only be kept in the last stage of the breeding cycle, when the breeder aims to select for commercial releasing and a higher accuracy is demanded (0.9). One must bear in mind that these results are valid for trials that are similar to the one used in our study. In *T. grandiflorum* breeding, recommendation trials are larger, with fewer genotypes evaluated in more replicates and locations. Thus, under these conditions, the number of harvests to be considered may be reduced even further.

Using multi-harvest data allows for the estimation of the hybrid families' fruit yield persistence, regarded herein as a measure of stability. In the perennial forages' context, persistence is the ability of a variety to survive and keep producing dry matter for a long period, which reduces the cost with reseeded (O'Connor et al., 2020; Parsons et al., 2011). In *T. grandiflorum* and any other fruit species' context, persistence is related to the genotypes' capability of maintaining a high fruit yield, regardless of the environmental changes. Note from eq. 2 that both performance and stability are compiled within the estimator (Rocha et al., 2018). That may explain the coincidence between the top ten hybrid families considering both persistence and fruit yield genotypic value. Despite the coincidence, there is a change in the order, suggesting that there are different levels of stability among the hybrid families.

The non-significance of GHI for witches' broom resistance justifies our choice of using only the trait's genotypic value for genetic selection. The Additive Index highlighted the best hybrids for both traits simultaneously. The coincidence between the ten best pointed by the index and the single-trait criteria (Figure 2) proves our point. Selection indices are recurrent in breeding program, since frequently more than one trait has economic value (Jaimez et al., 2020). In the *T. grandiflorum* case, the breeder cannot dissociate fruit yield from witches' broom resistance when selecting families. In other words, an ideal family must

hold favourable alleles for both traits. This is because there is a great incidence of *M. pernicioso* in the Amazon, since the region is its centre of diversity (Artero et al., 2017).

Amongst the selected hybrids, only two are present in the top ten for both yield and witches' broom resistance: H165 and H143. By selecting those genetic materials, one supports the breeding program with genotypes that hold favourable alleles from different sources. This diversifies the program's genetic base and, consequently, keeps the variability with the selection-recombination cycles. In doing so, one goes towards the gene pyramiding, elevating the genetic gains and, at the end of the process, providing the farmers with genotypes that are reliably productive and resistant to witches' broom (Pilet-Nayel et al., 2017; Willocquet et al., 2017).

5 Conclusion

The model with heterogeneous residual variances had the lowest AIC value if compared to the model with homogeneous residual variance.

T. grandiflorum trials with similar conditions as ours should only be evaluated for nine consecutive harvests if the breeder's goal is the recommendation. If this is not the case, the evaluation time can be reduced by half and, consequently, there is a higher genetic gain per time unit.

Considering simultaneously the fruit yield persistence and resistance to witches' broom, we selected the hybrid families H125, H165, H143, H137, H167, H129, H117, H124, H140 and H135. The selection of those families will provide a genetic gain of 6.14% for fruit yield and 7.19% for resistance.

6 References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Albuquerque, P. S. B., Silva, S. D. V. M., Luz, E. D. M. N., Pires, J. L., Vieira, A. M. C., Demétrio, C. G. B., Pascholatti, S. F., & Figueira, A. (2010). Novel sources of witches' broom resistance (causal agent *Moniliophthora pernicioso*) from natural populations of *Theobroma cacao* from the Brazilian Amazon. *Euphytica*, 172(1), 125–138. <https://doi.org/10/dkgsdg>
- Alves, R. M., Chaves, S. F. S., Alves, R. S., Santos, T. G., Araújo, D. G., & Resende, M. D. V. (2021). Cupuaçu tree genotype selection for an agroforestry system environment in the Amazon.

Pesquisa Agropecuária Brasileira, 56, e02139. <https://doi.org/10.1590/S1678-3921.pab2021.v56.02139>

- Alves, R. M., Chaves, S. F. S., Gama, M. A. P., Pedroza Neto, J. L., & Santos, T. G. (2020). Simultaneous selection of cupuassu tree and Brazilian mahogany genotypes in an agroforestry system in Pará state, Brazil. *Acta Amazonica*, 50(3), 183–191. <https://doi.org/10.1590/1809-4392202000711>
- Artero, A. S., Silva, J. Q., Albuquerque, P. S. B., Bressan, E. A., Leal Jr, G. A., Sebbenn, A. M., Griffith, G. W., & Figueira, A. (2017). Spatial genetic structure and dispersal of the cacao pathogen *Moniliophthora perniciosa* in the Brazilian Amazon. *Plant Pathology*, 66(6), 912–923. <https://doi.org/10.1111/ppa.12644>
- Atlin, G. N., Cairns, J. E., & Das, B. (2017). Rapid breeding and varietal replacement are critical to adaptation of cropping systems in the developing world to climate change. *Global Food Security*, 12, 31–37. <https://doi.org/10.1016/j.gfs.2017.01.008>
- Benjamin, C. S., Luz, E. D. M. N., Santos, W. O., & Pires, J. L. (2016). Cacao families and parents selected as resistant to natural infection of *Moniliophthora perniciosa*. *Crop Breeding and Applied Biotechnology*, 16, 141–146. <https://doi.org/10/gj8n5f>
- Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J., & Thompson, R. (2018). *ASReml-R reference manual Version 4*. VSN International. <https://mmade.org/wp-content/uploads/2019/01/asremlRMfinal.pdf>
- Carvalho, C. G. P., Cruz, C. D., Almeida, C. M. V. C., & Machado, P. F. R. (2002). Yield repeatability and evaluation period in hybrid cocoa assessment. *Crop Breeding and Applied Biotechnology*, 2(1), 149–156. <https://doi.org/10.12702/1984-7033.v02n01a19>
- Chaves, S. F. S., Alves, R. M., Alves, R. S., Sebbenn, A. M., Resende, M. D. V., & Dias, L. A. S. (2021). *Theobroma grandiflorum* breeding optimization based on repeatability, stability and adaptability information. *Euphytica*, 217(12), 211. <https://doi.org/10.1007/s10681-021-02944-3>
- Cobb, J. N., Juma, R. U., Biswas, P. S., Arbelaez, J. D., Rutkoski, J., Atlin, G., Hagen, T., Quinn, M., & Ng, E. H. (2019). Enhancing the rate of genetic gain in public-sector plant breeding programs: Lessons from the breeder's equation. *Theoretical and Applied Genetics*, 132(3), 627–645. <https://doi.org/10.1007/s00122-019-03317-0>

- Cullis, B. R., Smith, A. B., & Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological, and Environmental Statistics*, 11(4), 381–393. <https://doi.org/10.1198/108571106X154443>
- Edwards, J. W., & Jannink, J. (2006). Bayesian modeling of heterogeneous error and genotype $\times$ environment interaction variances. *Crop Science*, 46(2), 820–833. <https://doi.org/10.2135/cropsci2005.0164>
- Faveri, J., Verbyla, A. P., Pitchford, W. S., Venkatanagappa, S., & Cullis, B. R. (2015). Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. *Crop and Pasture Science*, 66(9), 947. <https://doi.org/10.1071/CP14312>
- Ferreira, F. M., Rocha, J. R. A. S. C., Bhering, L. L., Fernandes, F. D., Lédo, F. J. S., Rangel, J. H. A., Kopp, M., Câmara, T. M. M., Silva, V. Q. R., & Machado, J. C. (2021). Optimal harvest number and genotypic evaluation of total dry biomass, stability, and adaptability of elephant grass clones for bioenergy purposes. *Biomass and Bioenergy*, 149, 106104. <https://doi.org/10.1016/j.biombioe.2021.106104>
- Gezan, S. A., & Carvalho, M. (2018). Analysis of Repeated Measures for the Biological and Agricultural Sciences. In *Applied Statistics in Agricultural, Biological, and Environmental Sciences* (pp. 279–297). John Wiley & Sons, Ltd. <https://doi.org/10.2134/appliedstatistics.2016.0008.c10>
- Henderson, C. R. (1975). Best Linear Unbiased Estimation and Prediction under a selection model. *Biometrics*, 31(2), 423. <https://doi.org/10.2307/2529430>
- Jaimez, R. E., Vera, D. I., Mora, A., Loor, R. G., & Bailey, B. A. (2020). A disease and production index (DPI) for selection of cacao (*Theobroma cacao*) clones highly productive and tolerant to pod rot diseases. *Plant Pathology*, 69(4), 698–712. <https://doi.org/10/gj7t36>
- Martínez-García, P. J., Famula, R. A., Leslie, C., McGranahan, G. H., Famula, T. R., & Neale, D. B. (2017). Predicting breeding values and genetic components using generalized linear mixed models for categorical and continuous traits in walnut (*Juglans regia*). *Tree Genetics & Genomes*, 13(5), 109. <https://doi.org/10.1007/s11295-017-1187-z>
- McDonagh, J., O'Donovan, M., McEvoy, M., & Gilliland, T. J. (2016). Genetic gain in perennial ryegrass (*Lolium perenne*) varieties 1973 to 2013. *Euphytica*, 212(2), 187–199. <https://doi.org/10.1007/s10681-016-1754-7>

- Moraes, J. R. S. C., Rolim, G. S., Martorano, L. G., Aparecido, L. E. O., Oliveira, M. S. P., & Farias Neto, J. T. (2020). Agrometeorological models to forecast açai (*Euterpe oleracea* Mart.) yield in the Eastern Amazon. *Journal of the Science of Food and Agriculture*, 100(4), 1558–1569. <https://doi.org/10.1002/jsfa.10164>
- Mustiga, G. M., Gezan, S. A., Phillips-Mora, W., Arciniegas-Leal, A., Mata-Quirós, A., & Motamayor, J. C. (2018). Phenotypic description of *Theobroma cacao* L. for yield and vigor traits from 34 hybrid families in Costa Rica based on the genetic basis of the parental population. *Frontiers in Plant Science*, 9, 808. <https://doi.org/10/gdrrd8>
- Neves, L. C., Campos, A. J., Benedette, R. M., Tosin, J. M., & Chagas, E. A. (2012). Characterization of the antioxidant capacity of natives fruits from the Brazilian Amazon Region. *Revista Brasileira de Fruticultura*, 34, 1165–1173. <https://doi.org/10.1590/S0100-29452012000400025>
- O'Connor, J. R., Jahufer, M. Z. Z., & Lyons, T. (2020). Examining perennial ryegrass (*Lolium perenne* L.) persistence through comparative genetic analyses of two cultivars after nine years in the field. *Euphytica*, 216(3), 36. <https://doi.org/10.1007/s10681-020-2568-1>
- Parsons, A. J., Edwards, G. R., Newton, P. C. D., Chapman, D. F., Caradus, J. R., Rasmussen, S., & Rowarth, J. S. (2011). Past lessons and future prospects: Plant breeding for yield and persistence in cool-temperate pastures. *Grass and Forage Science*, 66(2), 153–172. <https://doi.org/10.1111/j.1365-2494.2011.00785.x>
- Patrocínio, N. G. R. B., Ceresini, P. C., Gomes, L. I. S., Resende, M. L. V., Mizubuti, E. S. G., & Gramacho, K. P. (2017). Population structure and migration of the witches' broom pathogen *Moniliophthora perniciosa* from cacao and cultivated and wild solanaceous hosts in southeastern Brazil. *Plant Pathology*, 66(6), 900–911. <https://doi.org/10.1111/ppa.12636>
- Patterson, H. D., & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58(3), 545–554. <https://doi.org/10.1093/biomet/58.3.545>
- Pereira, A. S., Lopes, U. V., Pires, J. L., Barreto, M. A., Santos, L. P., Gonzaga, R. S., Silva, L. S., Andrade, F. J., Bonilla, J. L. S., & Gramacho, K. P. (2021). Quantitative resistance to witches' broom disease in progenies of different sources of resistance. *Crop Protection*, 146, 105678. <https://doi.org/10.1016/j.cropro.2021.105678>
- Pilet-Nayel, M.-L., Moury, B., Caffier, V., Montarry, J., Kerlan, M.-C., Fournet, S., Durel, C.-E., & Delourme, R. (2017). Quantitative resistance to plant pathogens in pyramiding strategies for

- durable crop protection. *Frontiers in Plant Science*, 8, 1838.
<https://doi.org/10.3389/fpls.2017.01838>
- R Core Team. (2022). *R: A Language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Resende, M. D. V., & Alves, R. S. (2020). Linear, generalized, hierarchical, bayesian and random regression mixed models in genetic/genomics in plant breeding. *Functional Plant Breeding Journal*, 2(2), 1–31. <https://doi.org/10.35418/2526-4117/v2n2a1>
- Resende, M. D. V., & Duarte, J. B. (2007). Precisão e controle de qualidade em experimentos de avaliação de cultivares. *Pesquisa Agropecuária Tropical*, 37(3), 182–194.
- Rocha, J. R. A. S. C., Marçal, T. S., Salvador, F. V., Silva, A. C., Machado, J. C., & Carneiro, P. C. S. (2018). Genetic insights into elephantgrass persistence for bioenergy purpose. *PLOS ONE*, 13(9), e0203818. <https://doi.org/10.1371/journal.pone.0203818>
- Souza, A. G. C., Resende, M. D. V., Silva, S. E. L., & Sousa, N. R. (2002). The cupuaçu genetic improvement program at Embrapa Amazônia Ocidental. *Crop Breeding and Applied Biotechnology*, 2(3), 471–478. <https://doi.org/10.12702/1984-7033.v02n03a19>
- Verbyla, A. P., Faveri, J., Deery, D. M., & Rebetzke, G. J. (2021). Modelling temporal genetic and spatio-temporal residual effects for high-throughput phenotyping data. *Australian & New Zealand Journal of Statistics*, 63(2), 284–308. <https://doi.org/10.1111/anzs.12336>
- Villemereuil, P. (2018). Quantitative genetic methods depending on the nature of the phenotypic trait. *Annals of the New York Academy of Sciences*, 1422(1), 29–47.
<https://doi.org/10.1111/nyas.13571>
- Wilkins, P. W., & Humphreys, M. O. (2003). Progress in breeding perennial forage grasses for temperate agriculture. *The Journal of Agricultural Science*, 140(2), 129–150.
<https://doi.org/10.1017/S0021859603003058>
- Willocquet, L., Savary, S., & Yuen, J. (2017). Multiscale phenotyping and decision strategies in breeding for resistance. *Trends in Plant Science*, 22(5), 420–432.
<https://doi.org/10.1016/j.tplants.2017.01.009>
- Wolfinger, R. D. (1996). Heterogeneous variance: Covariance structures for repeated measures. *Journal of Agricultural, Biological, and Environmental Statistics*, 1(2), 205–230.
<https://doi.org/10.2307/1400366>

