

Path analysis under multiple-trait BLUP: application in the study of interrelationships among traits related to cotton (*Gossypium hirsutum*) fiber length

Análisis de ruta bajo BLUP de rasgos múltiples: aplicación en el estudio de las interrelaciones entre los rasgos relacionados con la longitud de la fibra de algodón (*Gossypium hirsutum*)

Rodrigo Silva, Alves ¹, João Romero do Amaral Santos de Carvalho Rocha ², Larissa Pereira Ribeiro Teodoro ³, Paulo Eduardo Teodoro ³, Luiz Paulo de Carvalho ⁴, Francisco José Correia Farias ⁴, Marcos Deon Vilela de Resende ¹, Leonardo Lopes Bhering ²

Originales: Recepción: 10/02/2020 - Aceptación: 23/09/2020

ABSTRACT

Multiple-trait best linear unbiased prediction (BLUP) is generally the most appropriate method to predict genetic values because it considers the genetic and residual correlations among traits and leads to higher selective accuracy. Thus, the aim of the present study was to analyze the interrelationships among traits related to cotton fiber length through path analysis under multiple-trait BLUP. To that end, 36 elite lines were evaluated in three environments and phenotyped for fiber quality and agronomic traits. Variance components were estimated via residual maximum likelihood (REML). The genetic correlation coefficients among traits were obtained through the mixed model output, and a correlation network was constructed to graphically express these results. Subsequently, path analysis was performed considering fiber length as the main dependent variable. Genetic parameters indicate that the phenotypic variance for most traits is mainly composed of residual effects, which reinforces the need for more accurate statistical methods, such as multiple-trait BLUP. The results for genetic correlations and path analysis reveal the difficulty of selection based on important fiber quality traits, especially fiber length, since most traits show a low cause-and-effect relationship, while other important traits have an undesirable cause-and-effect relationship. Path analysis under multiple-trait BLUP proved to be suitable and useful for studying interrelationships among traits related to cotton fiber length.

Keywords

mixed models • genotype × environment interaction • genetic correlation • genetic selection • *Gossypium hirsutum*

1 INCT Café/Federal University of Viçosa. Department of Statistics. University Campus. 36570-000. Viçosa, MG. Brazil.

2 Federal University of Viçosa. Department of General Biology. University Campus. 36570-000. Viçosa. MG. Brazil.

3 Federal University of Mato Grosso do Sul. Department of Plant Science. University Campus. 79560-000. Chapadão do Sul. MS. Brazil. eduteodoro@hotmail.com

4 National Cotton Research Center. Embrapa Cotton. 58428-095. Campina Grande. PB. Brazil.

RESUMEN

La mejor predicción lineal insesgada de rasgos múltiples (BLUP) es, en general, el método más apropiado para predecir valores genéticos porque considera las correlaciones genéticas y residuales entre rasgos y conducta con mayor precisión selectiva. Por lo tanto, el presente estudio tuvo como objetivo estudiar las interrelaciones entre los rasgos relacionados con la longitud de la fibra de algodón a través del análisis de ruta en BLUP de rasgos múltiples. Con este fin, se evaluaron treinta y seis líneas élite en tres ambientes y se fenotipificaron para muchos rasgos relacionados con la calidad de la fibra y los rasgos agronómicos. Los componentes de la varianza se estimaron mediante la máxima verosimilitud residual (REML). Los coeficientes de correlación genética entre rasgos se obtuvieron a través de la salida del modelo mixto, y para expresar gráficamente estos resultados se construyó una red de correlación. Posteriormente, realizamos el análisis de ruta considerando la longitud de la fibra como variable dependiente principal. Los parámetros genéticos indican que la variación fenotípica para la mayoría de los rasgos se compone principalmente de efectos residuales, lo que refuerza la necesidad de utilizar métodos estadísticos más precisos, como BLUP de rasgos múltiples. Los resultados encontrados para las correlaciones genéticas y el análisis de ruta revelan la dificultad de la selección basada en rasgos importantes de la calidad de la fibra, especialmente la longitud de la fibra, ya que la mayoría de los rasgos muestran una relación de causa y efecto muy baja, y otros rasgos importantes presentan una relación de causa y efecto no deseada. El análisis de ruta en BLUP de rasgos múltiples resultó ser adecuado y útil para el estudio de las interrelaciones entre los rasgos relacionados con la longitud de la fibra de algodón.

Palabras clave

modelos mixtos • interacción genotipo x ambiente • correlación genética • selección genética • *Gossypium hirsutum*

INTRODUCTION

Upland cotton (*Gossypium hirsutum* L.r. latifolium Hutch) produces one of the most important natural fibers in the world (4). Due to its wide adaptability and yields (5, 22), it contributes to the world economy as a leading fiber crop in the textile industry (21). To meet industrial demand, cotton breeding programs aim to select genotypes with fibers that are longer and that have greater strength for spinning and baling and greater uniformity of length, as well as genotypes with decreased short fiber content and more consistent micronaire and fiber fineness (2). All these traits are desirable to increase processing speed in the textile industry (3, 20).

To efficiently increase fiber length (FL), it is necessary to know the genetic interrelation of this trait with other agronomic and fiber quality traits, since undesirable changes in other traits may occur when selecting cotton genotypes with higher FL. Some studies have reported a negative relationship between FL and traits such as micronaire (MIC) and fiber percentage (FP) (4, 27). Nevertheless, in a study mapping quantitative trait loci (QTL) for fiber quality traits, Tan *et al.* (2015) found complex significant correlations among fiber quality traits, wherein FL was positively correlated with fiber strength (FS) and MIC, and negatively correlated with fiber elongation (FE). Given this scenario, methods that can make reliable estimates of genetic correlations will help in selection of genotypes (12) with agronomic traits desirable to farmers and that are technological requirements of the textile industry.

In this scenario, the multiple-trait best linear unbiased prediction (BLUP) proposed by Henderson and Quaas (1976) is generally the most appropriate method for genetic evaluation because it considers the genetic and residual correlations among traits, considering pleiotropism and/or gametic phase imbalance. Moreover, it can be fitted to experiments with heterogeneous residual variance and correlated errors within blocks or environments, leading to higher selective accuracy (17).

For genetic selection, it is necessary to identify variables that have a cause-and-effect relationship with the main trait, given that these traits can be used in indirect selection or to compose selection indices. To understand the association of different traits, Wright (1921)

proposed path analysis, which partitions the genetic correlation coefficients into direct and indirect effects on a main trait. Estimates of these effects are obtained by regression equations, which use previously standardized variables. Selection of auxiliary traits should be performed by analyzing the magnitude of their direct effect on the main trait. Traits with a high direct effect (negative or positive) in the same direction as their correlation with the main trait can be used in selection indices aiming to increase the main trait and to keep the other traits within the standards accepted by farmers and the textile industry.

Despite the importance of the multiple-trait BLUP approach, there are still no studies using it for path analysis. Thus, the aim of the present study was to analyze the interrelationships among traits related to cotton fiber length via path analysis under multiple-trait BLUP.

MATERIALS AND METHODS

Genetic material, experimental design, and sites

Phenotypic data was obtained from evaluation of 36 elite lines derived from crosses between two commercial cotton lines that have long fiber. To obtain the elite lines, the plants from these crosses were manually self-pollinated from the F_2 to F_4 generation in a greenhouse. The pedigree selection method was adopted from the F_4 generation on; in the F_4 generation, 271 plants were obtained, of which 51 $F_{4.5}$ plants with FL greater than or equal to 31 mm were selected. These lines were evaluated by Carvalho *et al.* (2015) in a greenhouse, and 36 plants with FL greater than or equal to 32 mm were selected, which were evaluated in the next generation ($F_{4.6}$). Of these, 36 lines with FL greater than or equal to 32 mm were also selected.

These lines were evaluated in three field trials in the municipality of Apodi, RN, in 2013 and 2014 and in Santa Helena, GO, in 2013 in a randomized block design with two replications. The trials in Apodi were set up under irrigation, and the trial in Santa Helena was set up in the rainy season. Plots consisted of two 5-m rows, spaced at 0.80 m, with 60 plants per row. Crop practices followed the recommendations for cotton growing. The herbicides trifloxysulfuron-sodium at 750 g kg^{-1} for dicotyledons and clethodim at 240 g L^{-1} for monocotyledons were used for weed control. Pyraclostrobin and metconazole fungicides were applied at a rate of 65 g ha^{-1} and 80 g ha^{-1} , respectively. Pest control consisted of applying deltamethrin insecticide at a rate of 10 g ha^{-1} and was carried out according to the integrated pest management recommended for the crop in the region.

Evaluated traits

The agronomic traits evaluated were plant height (PH) (cm), cotton boll mass (CBM) (g), and fiber percentage (FP) in five plants from each plot. The PH was measured with a ruler, from soil level to the tip of the plant. CBM was obtained by the average of the bolls, which were harvested manually. The FP was determined by the relationship between the weight of the fiber and the total weight of the sample (fiber plus cottonseed). Twenty bolls were randomly collected from the middle third of the plants from each experimental unit and the following fiber quality traits were evaluated: fiber length (FL, mm), fiber uniformity (FU, %), short fiber index (SFI), fiber strength (FS, gf/tex), fiber elongation (FE, %), and MIC.

Statistical analyses

Single-trait analyses were performed to test the significance of genotype effects and genotype \times environment (G \times E) interaction effects considering the following model:

$$y = Xb + Zg + Wi + e \quad [1]$$

where:

y = the vector of phenotypic data,

b = the vector of replication-environment combinations (assumed to be fixed), which consists of the effects of environment and replication within environment, added to the overall mean,

g = the vector of genotype effects (assumed to be random),

i = the vector of the G \times E interaction effects (random),

e = the vector of residuals (random).

Uppercase letters represent the incidence matrices for these effects. The significance of the random effects was tested by the likelihood ratio test (LRT) (25) using the chi-square statistic with one degree of freedom and significance level equal to 0.05 (16).

According to the LRT, the traits that exhibit genetic variability and the non-significant G×E interaction will be used in subsequent analyses. Variance components were estimated via residual maximum likelihood (REML). Multiple-trait analysis was performed considering the following model:

$$y = Xb + Zg + e \quad [2]$$

In this model, $g \sim N(0, \Sigma_g \otimes I)$ and $e \sim N(0, \Sigma_e \otimes I)$, where Σ_g is the genetic unstructured covariance matrix among the traits, Σ_e is the residual covariance matrix, I is an identity matrix with an appropriate order for the respective random effect, and \otimes denotes the Kronecker product. In addition, Σ_e is composed of diagonal blocks of unstructured covariance structure (Σ_{e1} , Σ_{e2} , and Σ_{e3}) among the traits for each environment, in which

$$\Sigma_e = \begin{bmatrix} \Sigma_{e1} & 0 & 0 \\ 0 & \Sigma_{e2} & 0 \\ 0 & 0 & \Sigma_{e3} \end{bmatrix}$$

Selective accuracies ($r_{\hat{g}g}$) were estimated based on the following expression (17):

$$r_{\hat{g}g} = \sqrt{(\sigma_g^2 - PEV) / \sigma_g^2} \quad [3]$$

where:

σ_g^2 = the genetic variance,

PEV = the prediction error variance, extracted from the diagonal of the generalized inverse of the coefficient matrix of the mixed model equations.

The genetic correlation coefficients between traits were obtained by multiple-trait BLUP output, based on the following expression:

$$\rho_{xy} = \frac{\sigma_{g(xy)}}{\sqrt{\sigma_{gx}^2 \sigma_{gy}^2}} \quad [4]$$

where

$\sigma_{g(xy)}$ = the genetic covariance between x and y traits,

σ_{gx}^2 = the genetic variance of trait x ,

σ_{gy}^2 = the genetic variance of trait y .

A correlation network was used to graphically express the functional relationship between the estimates of genetic correlation coefficients among the traits, in which the proximity between the nodes (traits) is proportional to the absolute value of the correlation between these nodes. The thickness of the edges was controlled by applying a cut-off value of 0.60, which means that only $|\rho_{xy}| \geq 0.60$ have their edges highlighted. Finally, positive correlations were highlighted in green, whereas negative correlations were represented in red.

Subsequently, a multicollinearity diagnosis was performed on the genetic correlation matrix, according to the recommendations of Montgomery and Peck (2001). Thus, a condition number less than 100 indicates the absence of multicollinearity.

Path analysis, considering FL as the main dependent variable, was performed using the following equation:

$$FL = \beta_1 FP + \beta_2 CBM + \dots + \beta_g PH + p_e \quad [5]$$

where:

$\beta_1, \beta_2, \dots, \beta_g$ = the direct effect estimators of the traits FP, CBM, FU, SFI, FS, FE, MIC PH on FL, and p_e = the residual effect of the analysis.

Thus, the normal equation system was used to estimate the direct and indirect effects of each explanatory variable on FL, according to the following expression:

$$\begin{bmatrix} 1.0 & \dots & r_{FP,PH} \\ \vdots & \ddots & \vdots \\ r_{PH,FP} & \dots & 1.0 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_g \end{bmatrix} = \begin{bmatrix} r_{FP,FL} \\ \vdots \\ r_{PH,FL} \end{bmatrix} \quad [6]$$

The coefficient of determination (R^2) of the path analysis was obtained by the following equation:

$$R^2 = \hat{\beta}_1 r_{FP,FL} + \dots + \hat{\beta}_8 r_{PH,FL} \quad [7]$$

The residual effect (\hat{p}_ε) of the path analysis was obtained by the following equation:

$$\hat{p}_\varepsilon = \sqrt{1 - R^2} \quad [8]$$

Software

Statistical analyses were performed using the ASReml 4.1 (7) and R (13) software.

RESULTS

Initially, the LRT was applied to test the random effects of the single-trait model. This test revealed a significant effect of genotypes for all traits evaluated (figure 1), indicating the existence of genetic variability among the cotton lines; and non-significant G×E interaction for all traits evaluated (figure 1).

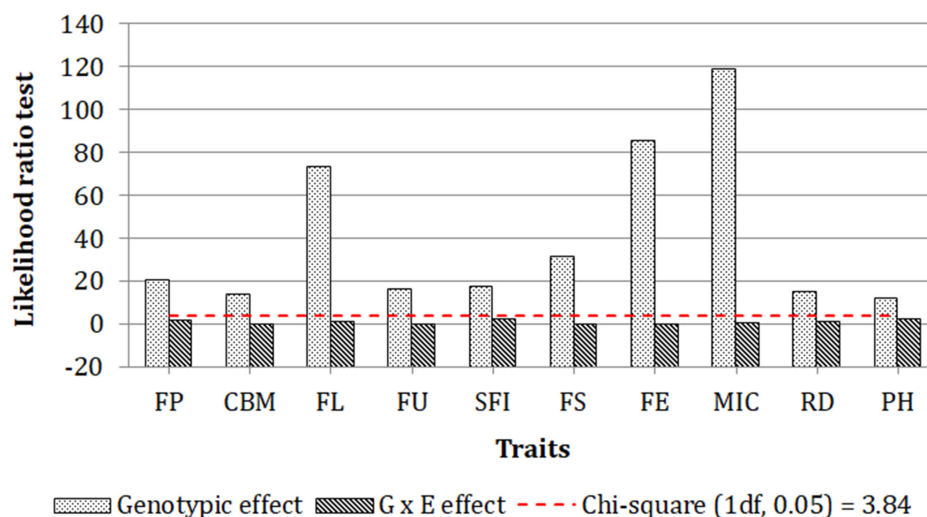


Figure 1. Likelihood ratio test considering single-trait analyses for the fiber percentage (FP), cotton boll mass (CBM), fiber length (FL), fiber uniformity (FU), short fiber index (SFI), fiber strength (FS), fiber elongation (FE), micronaire (MIC), and plant height (PH) traits. All bars above the dashed red line are significant by chi-square test at 5% probability.

Figura 1. Test de razón de probabilidad considerando análisis de un único carácter para porcentaje de fibra (FP), masa de cápsulas de algodón (CBM), longitud de la fibra (FL), uniformidad de la fibra (FU), índice de fibra corta (SFI), fuerza de la fibra (FS), elongación de la fibra (FE), micronaje (MIC) y altura de la planta (PH). Todas las barras por encima de la línea roja discontinua son significativas por la prueba de chi-cuadrado con una probabilidad del 5%.

In general, the estimates of residual variance differed from environment to environment and did not follow any pattern. According to the scale presented by Resende (2015), heritability estimates were of low ($h^2 \leq 0.15$) or moderate ($0.15 < h^2 < 0.50$) magnitude, except for the FL, FE, and MIC traits, which were of high ($h^2 \geq 0.50$) magnitude. The coefficients of experimental variation (CV_ε) were lower than 15% for all the traits evaluated, similar to the values obtained in other studies on the cotton crop (4). The coefficient of genetic variation (CV_g) indicates genetic variability among the cotton lines for all the traits evaluated. In addition, the estimates of average selective accuracy were of high ($0.70 < r_{\hat{g}g} < 0.90$) to very high ($r_{\hat{g}g} \geq 0.90$) magnitude, according to the scale presented by Resende (2015) (table 1, page 6).

Table 1. Estimates of genetic and non-genetic parameters for the fiber percentage (FP), cotton boll mass (CBM), fiber length (FL), fiber uniformity (FU), short fiber index (SFI), fiber strength (FS), fiber elongation (FE), micronaire (MIC), and plant height (PH) traits obtained via REML.

Tabla 1. Estimaciones de los parámetros genéticos y no genéticos para los caracteres porcentaje de fibra (FP), masa de cápsulas de algodón (CBM), longitud de fibra (FL), uniformidad de fibra (FU), índice de fibra corta (SFI), fuerza de fibra (FS), elongación de fibra (FE), micronaje (MIC) y altura de planta (PH) obtenidos a través de REML.

Genetic or non-genetic parameter	FP	CBM	FL	FU	SFI	FS	FE	MIC	PH
σ_g^2	8.67	0.16	1.99	0.42	0.18	1.50	0.44	0.29	28.26
σ_{e1}^2	4.21	0.38	1.53	1.44	0.56	3.32	0.21	0.08	42.08
σ_{e2}^2	4.48	1.05	0.98	1.16	0.28	2.18	0.16	0.06	181.02
σ_{e3}^2	22.60	0.41	0.51	0.79	0.12	1.79	0.29	0.05	61.01
h_{e1}^2	0.67	0.30	0.56	0.23	0.24	0.31	0.67	0.78	0.40
h_{e2}^2	0.66	0.13	0.67	0.27	0.39	0.41	0.74	0.84	0.14
h_{e3}^2	0.28	0.28	0.80	0.35	0.60	0.46	0.60	0.85	0.32
r_{gg}	0.95	0.87	0.96	0.91	0.92	0.95	0.96	0.97	0.88
μ_1	36.53	5.41	28.28	84.42	7.07	30.23	5.95	5.23	87.26
μ_2	35.19	6.98	30.03	85.63	6.64	31.49	5.67	4.93	113.76
μ_3	35.12	6.80	31.22	85.80	6.48	30.48	5.64	4.32	119.74
CV_{e1}	5.62	11.39	4.38	1.42	10.60	6.02	7.76	5.50	7.43
CV_{e2}	6.01	14.71	3.30	1.26	7.95	4.69	7.00	4.85	11.83
CV_{e3}	13.54	9.47	2.28	1.04	5.35	4.39	9.59	5.21	6.52
CV_{ge1}	8.06	7.45	4.99	0.77	5.96	4.05	11.18	10.38	6.09
CV_{ge2}	8.37	5.77	4.70	0.76	6.35	3.88	11.73	11.00	4.67
CV_{ge3}	8.38	5.93	4.52	0.76	6.50	4.01	11.79	12.57	4.44

σ_g^2 : genetic variance; σ_{e1}^2 , σ_{e2}^2 , and σ_{e3}^2 : residual variance for environments 1, 2, and 3, respectively; h_{e1}^2 , h_{e2}^2 , and h_{e3}^2 : individual broad-sense heritability in environments 1, 2, and 3, respectively; r_{gg} : average selective accuracy; μ_1 , μ_2 , and μ_3 : overall mean in environments 1, 2, and 3, respectively; CV_{e1} , CV_{e2} , and CV_{e3} : coefficient of experimental variation in environments 1, 2, and 3, respectively; and CV_{ge1} , CV_{ge2} , and CV_{ge3} : coefficient of genetic variation in environments 1, 2, and 3, respectively.

σ_g^2 : variación genotípica; σ_{e1}^2 , σ_{e2}^2 y σ_{e3}^2 : variación residual para los ambientes 1, 2 y 3, respectivamente, h_{e1}^2 y h_{e2}^2 : heredabilidad amplia individual considerando los ambientes 1, 2 y 3, respectivamente; r_{gg} : precisión de selección media; μ_1 , μ_2 y μ_3 : media general para los entornos 1, 2 y 3, respectivamente; CV_{e1} , CV_{e2} y CV_{e3} : coeficiente de variación experimental para los ambientes 1, 2 y 3, respectivamente; CV_{ge1} , CV_{ge2} y CV_{ge3} : coeficiente de variación genética para los ambientes 1, 2 y 3, respectivamente.

Genetic correlations were expressed graphically by the correlation network (figure 2), a procedure that indicates the interrelations among the traits evaluated. The FL trait was negatively correlated with FP, MIC, and FE, indicating that genotypes with higher FL have lower FP, MIC, and FE. The FP trait was negatively correlated with FS and positively correlated with FE. These results are also partially undesirable because they reveal that genotypes with higher FP have lower strength but higher elongation.

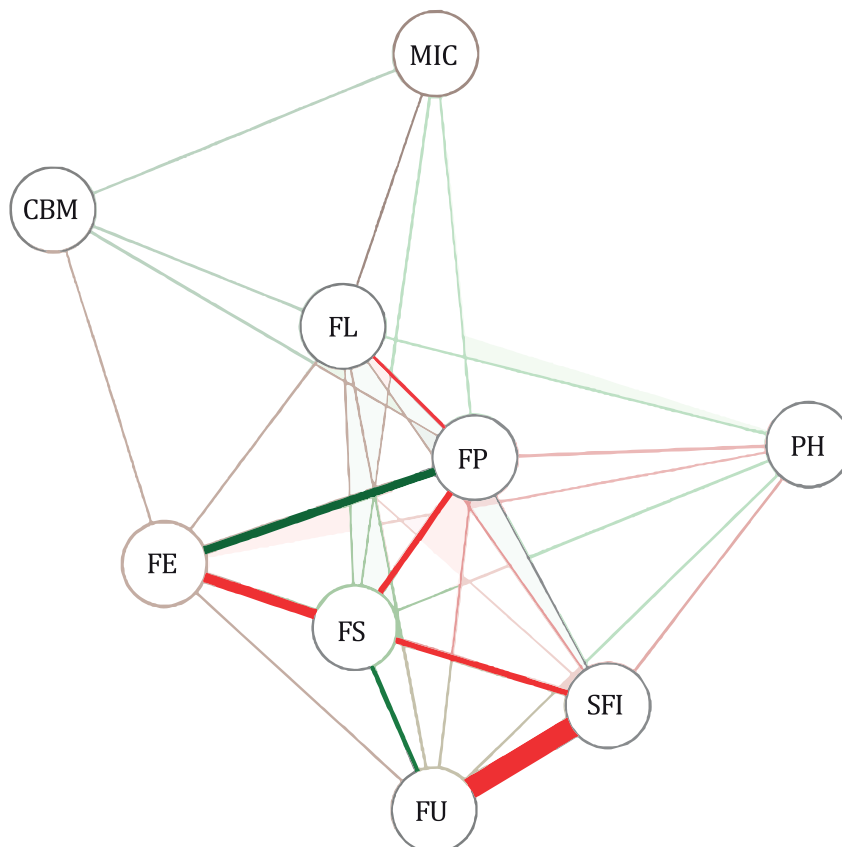


Figure 2. Correlation network among the fiber percentage (FP), cotton boll mass (CBM), fiber length (FL), fiber uniformity (FU), short fiber index (SFI), fiber strength (FS), fiber elongation (FE), micronaire (MIC), and plant height (PH) traits via multiple-trait BLUP.

Figura 2. Red de correlación entre el porcentaje de fibra (FP), masa de cápsulas de algodón (CBM), longitud de la fibra (FL), uniformidad de la fibra (FU), índice de fibra corta (SFI), fuerza de la fibra (FS), elongación de la fibra (FE), micronaje (MIC) y altura de la planta (PH), a través de BLUP multi-caracteres.

The direct and indirect effects of yield and fiber quality on FL are shown in table 2 (page 8). The estimate of the coefficient of determination was of high magnitude ($R^2 = 0.77$) and the residual effect was of low magnitude (0.23). To better understand the magnitudes of the effects and for comparison, effects whose magnitudes were above 0.70 were considered high-magnitude, whereas effects from 0.30 to 0.70 were considered moderate magnitude, and effects lower than 0.30 were considered low-magnitude (6). MIC was the fiber quality trait that obtained the highest direct effect on FL and in the same direction as its correlation. The FE trait had a moderate direct effect in the same direction as its correlations with FL, which is undesirable because an increase in FL is associated with a reduction in FE.

Table 2. Direct and indirect effects of fiber percentage (FP), cotton boll mass (CBM), fiber uniformity (FU), short fiber index (SFI), fiber strength (FS), fiber elongation (FE), micronaire (MIC), and plant height (PH) on fiber length (FL).

Tabla 2. Efectos directos e indirectos del porcentaje de fibra (FP), masa de ramitas de algodón (CBM), uniformidad de la fibra (FU), índice de fibra corta (SFI), resistencia de la fibra (FS), elongación de la fibra (FE), micronaje (MIC), y altura de la planta (PH) en la longitud de la fibra (FL).

Effect	FP	CBM	FU	SFI	FS	FE	MIC	PH
Direct	0.15	-0.18	0.05	-0.09	0.27	-0.35	-0.68	0.14
Indirect via FP	---	0.06	-0.07	0.07	-0.10	0.11	0.03	-0.07
Indirect via CBM	-0.07	---	0.01	0.02	-0.01	-0.06	-0.05	-0.02
Indirect via FU	-0.02	0.00	---	-0.05	0.03	-0.02	0.00	0.01
Indirect via SFI	-0.04	0.01	0.09	---	0.06	-0.03	0.00	0.04
Indirect via FS	-0.18	0.01	0.18	-0.18	---	-0.21	0.10	0.07
Indirect via FE	-0.25	-0.12	0.14	-0.12	0.27	---	0.03	0.07
Indirect via MIC	-0.14	-0.19	0.03	-0.01	-0.24	0.06	---	-0.02
Indirect via PH	-0.07	0.01	0.04	-0.05	0.04	-0.03	0.00	---
Total (genetic correlation)	-0.63	-0.41	0.47	-0.42	0.33	-0.55	-0.60	0.24
R ²	0.77							
Residual effect	0.23							

DISCUSSION

It is important to note that all traits exhibited genetic variability and there was no significant G×E interaction for any of the traits (figure 1, page 5), which indicates similar expression of the genes in controlling these traits in the three environments. Consequently, the estimates of genetic correlation between the environments were equal to 1.00 for all the traits evaluated.

The results for heritability call attention to the MIC, which is often used as an indication of fiber fineness and maturity. Fiber fineness depends on FL, FS, and FU. Therefore, a low estimate of heritability was expected for this trait. However, the residual variance of this trait was very low, that is, the mean values obtained fluctuated little from one replication to another in each genotype. Ulloa (2006) also observed heritability estimates for micronaire higher than the estimates observed for fiber fineness and maturity traits. In general, the results indicate that the phenotypic variance for most traits is mostly composed of residual effects and reinforces the need for using more accurate statistical methods, such as multiple-trait BLUP, to obtain unbiased genetic correlations (1).

Similar results for genetic correlations were also reported by Zeng and Pettigrew (2015) who evaluated genetic correlations among yield traits and fiber properties in upland cotton and found a negative genetic correlation between MIC × FL and MIC × FS. These results raise concerns in plant breeders because FP is related to the amount of product to be marketed, where high values are desirable to farmers. In contrast, the negative correlation between FL × MIC is a favorable result, since genotypes that have greater fiber length and lower micronaire values are desired. Micronaire is an indicator of fiber fineness and maturity, and high micronaire or coarse fibers negatively affect the spinning process and yarn quality (19).

Cotton genotypes with higher FP are desirable because this trait affects the price paid by industry to farmers. FE is a trait related to the strength of cotton fibers and contributes to higher fiber processing speed and yarn and fabric quality (3). Lastly, a high negative correlation was found for SFI × FU, which was expected, since more uniform fibers tend to have a longer length. The other estimates of genetic correlations were of low magnitude.

Studies involving technological traits have been the focus of many cotton breeders (11). FL is among the fiber quality traits that still need to be improved. To that end, knowledge

of the correlations among agronomic and technological traits in breeding programs is essential. This is because when selection is carried out based on a specific trait, changes occur in other traits of agronomic importance.

The results for genetic correlations found here reveal the difficulty of selection based on important fiber quality traits, especially FL, since selection based on this trait could consequently lead to selection of other undesirable traits, such as lower FP and FE. However, correlation estimates can produce misconceptions about the relation between two traits and are not necessarily an accurate cause-and-effect measure (14). In this respect, path analysis is a procedure that can effectively direct simultaneous selection. Teodoro *et al.* (2014) mention that this analysis provides detailed knowledge of the effect of traits and reveals positive and negative correlations of high and low magnitudes among the traits evaluated. For that reason, genetic correlation coefficients of yield and fiber quality traits were partitioned into direct and indirect effects on FL.

The use of path analysis to establish selection criteria in cotton is recent and has been restricted to considering fiber yield as the main dependent variable (14). The R^2 for path analysis indicates that practically all variation in FL was explained by the auxiliary traits considered in the path analysis. Furthermore, it is essential to highlight that path analysis was applied using the genetic values predicted via multiple-trait BLUP, which entails greater selective accuracy (17).

This higher direct effect of MIC on FL is important from the perspective of cotton breeding for FL as it reveals that it is possible to select genotypes with higher FL and lower MIC, which is one of the main interests of cotton breeding programs. The low magnitude of the indirect effects of all traits confirms the difficulty of selecting cotton genotypes superior for both yield and fiber quality traits. Some traits have an undesirable cause-and-effect relationship, or they even show a very low cause-and-effect relationship, making simultaneous selection for traits of interest difficult, thus compromising breeding program efficiency.

Therefore, the findings of this study indicate that there is no genotype carrying all desirable traits, confirming that simultaneous selection for yield and fiber quality traits is a major challenge in cotton breeding. Some studies (9, 18, 19) have already reported that this difficulty in selecting cotton genotypes superior for yield and fiber quality traits may be because there are many genes associated with undesirable traits as a consequence of linkage drag.

CONCLUSION

Path analysis under multiple-trait BLUP proved to be suitable and useful for the study of the interrelationships among traits related to cotton fiber length.

REFERENCES

1. Alves, R. S.; Rocha, J. R. A. S. C.; Teodoro, P. E.; Resende, M. D. V.; Henriques, E. P.; Silva, L. A.; Carneiro, P. C. S.; Bhering, L. L. 2018. Multiple-trait BLUP: a suitable strategy for genetic selection of *Eucalyptus*. *Tree Genet Genomes*. 14: 77.
2. Baxevanos, D.; Goulasb, C.; Rossic, J.; Braojosc, E. 2008. Separation of cotton cultivar testing sites based on representativeness and discriminating ability using GGE biplots. *Crop Science*. 100: 1230-1236.
3. Beyer, B. M.; Smith, C. W.; Percy, R.; Hague, S.; Hequet, E. F. 2014. Test cross evaluation of upland cotton accessions for selected fiber properties. *Crop Science*. 54: 60-67.
4. Carvalho, L. P.; Farias, F. J. C.; Rodrigues, J. I. S. 2015. Selection for increased fiber length in cotton progenies from Acala and Non-Acala types. *Crop Science*. 55: 1-7.
5. Chen, Z. J.; Scheffler, B. E.; Dennis, E.; Triplett, B. A.; Zhang, T. Z.; Guo, W. Z.; Chen, X. Y.; Stelly, D. M.; Rabinowicz, P. D.; Town, C. D.; Arioli, T.; Brubaker, C.; Cantrell, R. G.; Lacape, J. M.; Ulloa, M.; Chee, P.; Gingle, A. R.; Haigler, C. H.; Percy, R.; Saha, S.; Wilkins, T.; Wright, R. J.; Deynze, A. V.; Zhu, Y. X.; Yu, S. X.; Abdurakhmonov, I.; Katageri, I.; Kumar, P. A.; Rahman, M.; Zafar, Y.; Yu, J. Z.; Kohel, R. J.; Wendel, J. F.; Paterson, A. H. 2007. Toward sequencing cotton (*Gossypium*) genomes. *Plant Physiology*. 145: 1303-1310.
6. Cruz, C. D.; Regazzi, A. J.; Carneiro, P. C. S. 2012. Biometric models applied to genetic breeding. UFV. Viçosa.

7. Gilmour, A. R.; Gogel, B. J.; Cullis, B. R.; Welham, S. J.; Thompson, R. 2015. ASReml User Guide Release 4.1 Structural Specification. VSN International. Hemel Hempstead.
8. Henderson, C. R.; Quaas, R. L. 1976. Multiple Trait Evaluation Using Relatives' Records. *Journal of Animal Science*. 43: 11-88.
9. Jenkins, J. N.; Wu, J.; McCarty, J. C.; Saha, S.; Gutierrez, O. A.; Hayes, R.; Stelly, D. M. 2006. Genetic effects of thirteen *Gossypium barbadense* L. chromosome substitution lines with upland cotton cultivars: I. Yield and yield component. *Crop Science*. 46: 1169-1178.
10. Montgomery, D. C.; Peck, E. A. 2001. Introduction to linear regression analysis. New York. John Wiley & Sons.
11. Ng, E. H.; Smith, C. W.; Hequet, E.; Hague, S.; Dever, J. 2015. Generation mean analysis for fiber elongation in upland cotton. *Crop Science*. 54: 1347-1353.
12. Pollak, E. J.; Van der Werf, J.; Quaas, R. L. 1984. Selection bias and multiple trait evaluation. *Journal of Dairy Science*. 67: 1590-1595.
13. R Development Core Team. 2018. R: A language and environment for statistical computing. Vienna. R Foundation for Statistical Computing.
14. Reddy, K. B.; Reddy, V. C.; Ahamed, M. L.; Naidu, T. C. M.; Rao, V. S. 2017. Character association and path coefficient analysis for yield and component traits in upland cotton (*Gossypium hirsutum* L.). *J. Cotton Research Development*. 31: 29-33.
15. Resende, M. D. V. 2015. Quantitative and Population Genetics. Suprema, Visconde do Rio Branco.
16. Resende, M. D. V. 2016. Software Selegen-REML/BLUP: a useful tool for plant breeding. *Crop Breeding and Applied Biotechnology*. 16: 330-339.
17. Resende, M. D. V.; Silva, F. F.; Azevedo, C. F. 2014. Mathematical, Biometric and Computational Statistics. Suprema, Visconde do Rio Branco.
18. Saha, S.; Wu, J.; Jenkins, J. N.; McCarty, J.; Hayes, R.; Stelly, D. M. 2010. Genetic dissection of chromosome substitution lines discovered novel alleles in *Gossypium barbadense* L. with potential for improving agronomic traits including yield. *Theoretical and Applied Genetics*. 120: 1193-1205.
19. Saha, S.; Jenkins, J. N.; McCarty, J. C.; Hayes, R. W.; Stelly, D. M.; Campbell, B. T. 2018. Registration of Two CS-B17-derived upland cotton recombinant inbred lines with improved fiber micronaire. *Journal of Plant Registrations*. 12: 97-100.
20. Smith, C. W.; Hague, S.; Hequet, E.; Thaxton, S.; Brown, N. 2008. Development of extra-long staple upland cotton. *Crop Science*. 48: 1823-1831.
21. Tan, Z.; Fang, X.; Tang, S.; Zhang, J.; Liu, D.; Teng, Z.; Li, L.; Ni, H.; Zheng, F.; Liu, D.; Zhang, T.; Paterson, A. H.; Zhang, Z. 2015. Genetic map and QTL controlling fiber quality traits in upland cotton (*Gossypium hirsutum* L.). *Euphytica*. 203: 615-628.
22. Tang, S.; Teng, Z.; Zhai, T.; Fang, X.; Liu, F.; Liu, D.; Zhang, J.; Liu, D.; Wang, S.; Zhang, K.; Shao, Q.; Tan, Z.; Paterson, A. H.; Zhang, Z. 2015. Construction of genetic map and QTL analysis of fiber quality traits for Upland cotton (*Gossypium hirsutum* L.). *Euphytica*. 201: 195-213.
23. Teodoro, P. E.; Silva Junior, C. A.; Corrêa, C. C.; Ribeiro, L. P.; Oliveira, E. P.; Lima, M. F.; Torres, F. E. 2014. Path analysis and correlation of two genetic classes of maize (*Zea mays* L.). *Journal of Agronomy*. 13: 23-28.
24. Ulloa, M. 2006. Heritability and correlations of agronomic and fiber traits in an okra-leaf upland cotton population. *Crop Science*. 46: 1508-1514.
25. Wilks, S. S. 1938. The large-sample distribution of the likelihood ratio for testing composite Hypotheses. *The Annals of Mathematical Statistics*. 9: 60-62.
26. Wright, S. 1921. Correlation and causation. *Journal of Agricultural Research* 20: 557-585.
27. Zeng, L.; Pettigrew, W. T. 2015. Combining ability, heritability, and genetic correlations for lint yield and fiber quality of Upland cotton in delayed planting. *Field Crops Research*. 171: 176-183

ACKNOWLEDGMENTS

We appreciate the financial support from the Brazilian government through the National Council for Scientific and Technological Development (CNPq) and the Coordination for the Improvement of Higher Education Personnel (CAPES) - Finance Code 001.