



## Multi-trait BLUP model indicates sorghum hybrids with genetic potential for agronomic and nutritional traits

J.E. Almeida Filho<sup>1</sup>, F.D. Tardin<sup>2</sup>, J.F.R. Guimarães<sup>1</sup>, M.D.V. Resende<sup>3</sup>, F.F. Silva<sup>1</sup>, M.L. Simeone<sup>2</sup>, C.B. Menezes<sup>2</sup> and V.A.V. Queiroz<sup>2</sup>

<sup>1</sup>Departamento de Zootecnia, Universidade Federal de Viçosa, Viçosa, MG, Brasil

<sup>2</sup>Núcleo de Recursos Genéticos e Desenvolvimento de Cultivares, Embrapa Milho e Sorgo, Sete Lagoas, MG, Brasil

<sup>3</sup>Embrapa Floresta, Colombo, PR, Brasil

Corresponding author: J.E. Almeida Filho

E-mail: janeo.eustaquio@ymail.com

Genet. Mol. Res. 15 (1): gmr.15017071

Received June 24, 2015

Accepted October 16, 2015

Published February 26, 2016

DOI <http://dx.doi.org/10.4238/gmr.15017071>

**ABSTRACT.** The breeding of sorghum, *Sorghum bicolor* (L.) Moench, aimed at improving its nutritional quality, is of great interest, since it can be used as a highly nutritive alternative food source and can possibly be cultivated in regions with low rainfall. The objective of the present study was to evaluate the potential and genetic diversity of grain-sorghum hybrids for traits of agronomic and nutritional interest. To this end, the traits grain yield and flowering, and concentrations of protein, potassium, calcium, magnesium, sulfur, iron, manganese, and zinc in the grain were evaluated in 25 grain-sorghum hybrids, comprising 18 experimental hybrids of Embrapa Milho e Sorgo and seven commercial hybrids. The genetic potential was analyzed by a multi-trait best linear unbiased prediction (BLUP) model, and cluster analysis was accomplished by squared Mahalanobis distance using the predicted genotypic values. Hybrids 0306037 and 0306034 stood out in the agronomic evaluation. The hybrids with agronomic prominence, however, did not stand out for the traits related to the nutritional quality of the grain. Three clusters were formed from the dendrogram obtained with

the unweighted pair group method with arithmetic mean method. From the results of the genotypic BLUP and the analysis of the dendrogram, hybrids 0577337, 0441347, 0307651, and 0306037 were identified as having the potential to establish a population that can aggregate alleles for all the evaluated traits of interest.

**Key words:** *Sorghum bicolor*; Biofortification; Mixed models; Mahalanobis distance; Hierarchical cluster

## INTRODUCTION

Thirty five percent of the global production of sorghum [*Sorghum bicolor* (L.) Moench] is estimated to be for human consumption (Awika and Rooney, 2004) with over 500 million people living in developing countries utilizing it as a staple food (Mutisya et al., 2009). This explains the effort aimed at improving the nutritional quality of the grains, which is suitably termed “biofortification” (Olembo et al. 2010).

The biofortification of sorghum is a longstanding practice on the African continent and could substantially supply the nutritional needs of families, helping resolve the issue of food scarcity (Africa Harvest Biotech Foundation International, 2010). Compared to other food species, sorghum is well adapted to water-stress conditions and could be highly valuable and vital in ensuring good-quality food for needy families in areas with low rainfall.

In order to develop a breeding program aimed at biofortification, it is essential to evaluate genotypes for their agronomic performance, as well as, for their nutritional quality. It is also important to distinguish divergent individuals so as to aggregate favorable alleles among these individuals for the traits of interest.

According to Piepho et al. (2008), the standard procedures for the evaluation of genetically superior genotypes are based on the methodology of mixed models, restricted maximum likelihood/best linear unbiased prediction (REML/BLUP), and current interest in improvement is focused on several traits. The multi-trait linear model proposed by Henderson and Quaas (1976) seems promising, as it considers the heritabilities and genetic and environmental correlations among the traits, and thus tends to be more accurate than the univariate model (Schaeffer, 1984; Mrode, 2005).

Measures of distance are much useful for the evaluation of genetic diversity, especially Mahalanobis distance, which additionally considers the residual correlations among the variables (Cruz et al., 2011). Another advantage in obtaining the dissimilarity matrix would be the use of genotypic values instead of phenotypic means to obtain Mahalanobis distance, because it would be a measurement of genetic distance, free of environmental effects.

The objective of the present study was to evaluate the genetic potential and the diversity among the hybrids of grain sorghum aimed at identifying the potential founders in a breeding program for the traits of agronomic and nutritional interest.

## MATERIAL AND METHODS

Data used in the present study were from an experiment conducted in the 2010 growing season, at Sete Lagoas (Brazil: -19°28'S; 44°14'W), with 25 simple hybrids, 18 of which were pre-commercial hybrids from the breeding program of Embrapa (Brazilian Agricultural Research Corporation) Milho e Sorgo and seven were the commercial cultivars, BRS 310, BRS 330, BRS 332, BRS 373, BRS 380,

DKB 599, and DOW 822. The experimental design adopted was a 5 x 5 triple lattice.

The plots were composed of four 5-m-long rows, with a spacing of 0.1 m between the plants within the rows and 0.5 m between the rows. The central 4 m of the two internal rows were considered the useful area of the plot, and plants in the rest of the plot were used as the boundary.

The evaluated agronomic traits were grain yield with moisture corrected to 13% and flowering, which was measured as the number of days from planting to the time at which 50% of the plants showed flowers at the middle third of the panicle in anthesis.

The nutritional traits evaluated were the concentrations of potassium, calcium, magnesium, sulfur, iron, manganese, zinc, and protein in the grain. The plant material collected for these analyses was a sample of grains from each plot, according to the experimental design adopted in the field. These samples were conditioned in paper bags and dried in an oven at 65°C until they attained a constant weight. Subsequently, the bags were ground in a Wiley mill (2-mm-mesh sieve). The nitrogen content was determined by the method of Dumas (AOAC, 1998), and the protein content was calculated by multiplying by the factor 6.25 (Jones, 1941). For other nutrients, a nitric-perchloric digestion and quantification by argon plasma atomic emission spectrometry (ICP-OES; AOAC, 1998) was performed.

For statistical analysis, the multivariate mixed linear model was adopted.

$$y_{ijkl} = \mu_l + R_{jl} + G_{il} + B_k(R_j)_l + \varepsilon_{ijkl} \quad (\text{Equation 1})$$

where:  $\mu_l$  was the overall constant for trait  $l$ ;  $R_{jl}$  was the fixed effect of replicate  $j$  for the trait  $l$ ;  $G_{il}$  was the random effect of genotype  $i$  for trait  $l$ ;  $B_k(R_j)_l$  was the random effect of block  $k$  within repetition  $j$  for the trait  $l$ , and  $\varepsilon_{ijkl}$  was the effect of the random error corresponding to the plot  $ikj$  for trait  $l$ .

This model was expressed in matrix form as follows:

$$y = X\beta + Zg + Wb + \varepsilon \quad (\text{Equation 2})$$

where:  $\beta$  was the vector with the parameters of fixed effect corresponding to the overall constant and the repetition;  $g$  was the vector of random effects of genotypes;  $b$  was the vector of random effects of incomplete blocks within repetitions; and  $\varepsilon$  was the vector of random errors.

After assuming that:

$$\begin{bmatrix} y \\ g \\ b \\ \varepsilon \end{bmatrix} \sim \text{MVN} \left( \begin{bmatrix} X\beta \\ \phi \\ \phi \\ \phi \end{bmatrix}, \begin{bmatrix} V & ZG & WB & R \\ GZ' & G & \phi & \phi \\ BW' & \phi & B & \phi \\ R & \phi & \phi & R \end{bmatrix} \right) \quad (\text{Equation 3})$$

$$V = ZGZ' + WBW' + R; G = G_0 \otimes I; B = B_0 \otimes I; R = R_0 \otimes I$$

$G_0$ ,  $B_0$ , and  $R_0$  were co-variance matrices referring to genotypes, incomplete blocks, and residuals, respectively, and  $\otimes$  was the Kronecker product. These matrices were estimated by the REML method.

The estimate of one of the solutions of the fixed effects and prediction of the random effects was made by maximizing the joint probability density function of the vector of phenotypic observations with the vectors of genotypic values and of incomplete blocks, from which we obtained the mixed models equations by considering the multi-trait model of Henderson and Quaas (1976); these equations were expressed as

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z & X'R^{-1}W \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} & Z'R^{-1}W \\ W'R^{-1}X & W'R^{-1}Z & W'R^{-1}W + B^{-1} \end{bmatrix} \begin{bmatrix} \beta^0 \\ \tilde{g} \\ \tilde{b} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \\ W'R^{-1}y \end{bmatrix} \quad (\text{Equation 4})$$

The genotypic coefficient of determination, which is analogous to the heritability in the broad sense, of a hypothetical population generated by the evaluated hybrids, was given by:

$$\hat{h}^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_b^2 + \hat{\sigma}_e^2} \quad (\text{Equation 5})$$

where  $\sigma_g^2$ ,  $\sigma_b^2$  and  $\sigma_e^2$  were the variance components corresponding to genotypes, blocks, and errors, respectively, for a given trait.

To check the efficiency of the multivariate model, the univariate model

$$y_{ijk} = \mu + R_j + G_i + B_k(R_j) + \varepsilon_{ijk} \quad (\text{Equation 6})$$

was also adjusted for each trait, and the definition of the parameters was analogous to that in the multivariate case; the mixed models equations were also considered in this model, but, in the univariate case, the co-variance matrices were given by:

$$V = ZZ'\sigma_g^2 + WW'\sigma_b^2 + I\sigma_e^2; G = I\sigma_g^2; B = I\sigma_b^2; R = I\sigma_e^2 \quad (\text{Equation 7})$$

In the univariate model, all the variance components were estimated by the REML method independently.

The accuracy of the BLUP of the genotypes in both the multivariate and univariate models, considering the respective variance components, was given by:

$$\hat{r}_{g\tilde{g}} = \sqrt{1 - \frac{\text{Var}(\tilde{g} - g)}{\hat{\sigma}_g^2}} \quad (\text{Equation 8})$$

A measure was used to verify the relative efficiency of the multivariate model in relation to the univariate model. This measure was given by:

$$Ef(\%) = 100 \frac{\hat{r}_{g\tilde{g}(m)} - \hat{r}_{g\tilde{g}(u)}}{\hat{r}_{g\tilde{g}(u)}} \quad (\text{Equation 9})$$

where  $\hat{r}_{g\tilde{g}(m)}$  and  $\hat{r}_{g\tilde{g}(u)}$  were the accuracies in the multivariate and univariate models, respectively.

A diversity analysis was conducted using the unweighted pair group method with arithmetic mean (UPGMA) hierarchical clustering method, based on the squared Mahalanobis distance from the Multi-trait BLUP of the genetic values.

The distance matrix was obtained by:

$$D^2 = (d_{ii}^2) = (\tilde{g}_i - \tilde{g}_{i'})' R_0^{-1} (\tilde{g}_i - \tilde{g}_{i'}) \quad (\text{Equation 10})$$

where  $g_i$  was the vector of the Multi-trait BLUP of the  $i$ -th genotype for all traits, and  $R_0$  was the residual co-variance matrix.

The cutoff point of the dendrogram was determined based on the methodology of Mojena (1977), with the cutoff point given by

$$CP = \hat{\mu}_\alpha + k\hat{\sigma}_\alpha \quad (\text{Equation 11})$$

where,  $\hat{\mu}_\alpha$  and  $\hat{\sigma}_\alpha$  were the mean and standard deviation of the fusion points between two groups of the dendrogram; and  $k$  was a constant. In this study,  $k = 1.25$ , according to Milligan and Cooper (1985).

The ASReml 3.0 software (Gilmour et al., 2009) was used for the analysis of the linear models; for the analysis of genetic diversity, the R software (version 3.0.1; R Core Team, 2013) was adopted, with adaptations of the script generated by the GENES software (Cruz, 2013).

## RESULTS

### Genetic evaluation of the hybrids

For all evaluated traits, the multi-trait model showed accuracy greater than 0.9, and it was more accurate in terms of genotypic evaluation than the univariate model, especially for traits of lower heritability, as was the case for traits influencing the nutritional quality of the grains (Table 1).

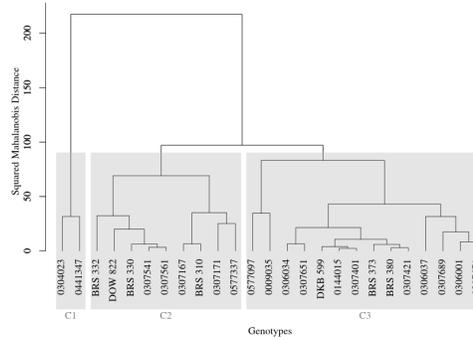
In the agronomic evaluations, hybrids 0306037 and 0306034 were the most productive, especially the hybrid 0306037, which showed the second earliest flowering. The only cultivar that showed genotypic value for grain yield among the seven most productive hybrids was BRS 330. Flowering in the experimental hybrids was, in general, earlier than that in the cultivars.

Hybrids 0577337 and 0441347 showed elevated genotypic values for the nutritional traits except for potassium. For the traits of agronomic interest, these hybrids displayed low potential. Hybrid 0307651 showed fair agronomic performance and elevated genotypic values for potassium and other traits of nutritional interest.

### Genetic diversity

With the squared Mahalanobis distances, it was possible to obtain a consistent dendrogram using the UPGMA clustering method. By the Mantel test ( $P < 0.01$ ) with 1000 permutations, we confirmed the association between the cophenetic matrix and the distance matrix ( $D^2$ ), with the cophenetic correlation coefficient being higher than 70%.

Three clusters were formed according to the cut-off point based on criterion of Mojena (1977) (Figure 1). The first cluster was formed by the experimental hybrids 0304023 and 0441347, both with weak agronomic performance: low genotypic value for grain yield and high genotypic value for flowering. However, regarding the nutritional traits, these hybrids had a great performance for almost all traits; the only exception was potassium, with 0441347 standing out (Table 1).



**Figure 1.** Dendrogram of squared Mahalanobis distance, from the genotypic multi-trait BLUP using the UPGMA method. Cutoff point at the distance of 92.36, with three clusters formed (C1, C2 and C3).

**Table 1.** Genotypic values for the traits: flowering (FLO), grain yield (GY), protein (PRO), potassium (K), calcium (Ca), magnesium (Mg), sulfur (S), copper (Cu), iron (Fe), manganese (Mn), and zinc (Zn) in 25 hybrids of grain sorghum evaluated in Sete Lagoas, MG, Brazil.

Hybrids	Traits										
	FLO (days)	PG (kg/ha)	PRO (%)	K (g/kg)	Ca (g/kg)	Mg (g/kg)	S (g/kg)	Cu (mg/kg)	Fe (mg/kg)	Mn (mg/kg)	Zn (mg/kg)
0009035	64.78	3582.95	13.52	5.20	0.26	1.91	0.99	4.02	37.13	19.02	16.44
0009061	67.56	6437.25	11.06	3.91	0.13	1.37	0.86	2.48	25.13	15.22	13.18
0144015	67.95	5800.15	10.67	4.64	0.15	1.43	0.81	2.75	25.01	15.73	13.5
0304023	75.17	5464.25	15.63	3.63	0.24	1.64	1.05	3.31	38.03	17.28	16.84
0306001	65.18	4446.95	11.08	4.85	0.16	1.58	0.89	2.91	28.14	17.5	13.91
0306034	68.71	7042.95	12.73	4.04	0.16	1.47	0.89	2.73	29.49	15.32	13.67
0306037	63.25	7590.95	11.69	4.20	0.16	1.55	0.83	2.72	28.13	14.48	13.48
0307167	68.40	5784.55	12.22	3.87	0.24	1.56	0.89	3.68	30.46	16.11	16.15
0307171	70.70	3582.95	13.84	4.71	0.31	1.88	1.01	4.35	40.20	19.45	18.70
0307343	68.95	6615.15	11.09	4.00	0.15	1.35	0.83	2.66	24.65	14.97	13.47
0307401	68.14	5843.45	11.11	4.49	0.15	1.45	0.84	2.68	25.44	15.5	13.6
0307421	68.02	6671.95	10.74	3.91	0.13	1.35	0.82	2.52	23.17	14.3	13.41
0307541	71.30	6489.15	11.58	4.14	0.17	1.42	0.85	2.76	28.12	15.68	14.4
0307561	71.41	6252.55	11.3	3.75	0.13	1.29	0.84	2.44	23.31	14.65	13.59
0307651	68.16	6639.75	12.44	4.83	0.20	1.72	0.90	3.04	36.14	16.99	15.38
0307671	65.33	5986.75	11.42	4.90	0.19	1.63	0.87	3.23	32.29	17.5	14.32
0307689	65.86	6467.25	10.96	4.28	0.15	1.48	0.81	2.67	24.24	14.28	13.91
0441347	75.36	4851.05	15.09	3.95	0.38	1.70	1.00	5.22	38.88	18.28	18.97
0577097	62.36	4233.95	10.92	4.49	0.19	1.72	0.92	3.23	29.79	17.03	15.76
0577337	69.23	2742.95	13.12	4.32	0.34	1.77	1.02	5.05	37.25	20.11	18.57
BRS 310	67.97	5180.75	12.43	3.69	0.20	1.54	0.91	3.21	26.82	14.94	15.7
BRS 330	71.32	6910.95	11.57	3.51	0.15	1.33	0.87	2.59	26.51	14.92	14.12
BRS 332	74.51	5709.36	12.31	3.81	0.17	1.36	0.88	2.75	25.5	14.88	14.93
DKB 599	68.88	6467.55	11.13	4.58	0.17	1.46	0.82	2.90	27.38	15.77	13.92
Dow 822	71.16	4703.95	12.49	4.31	0.23	1.56	0.88	3.58	27.69	15.70	16.20
IC <sub>95%</sub> (±) <sup>1/</sup>	1.53	934.92	0.92	0.39	0.05	0.13	0.05	0.57	4.20	1.33	1.30
h <sup>2</sup>	0.92	0.70	0.54	0.45	0.32	0.48	0.45	0.46	0.51	0.44	0.29
r <sub>gg</sub>	0.97	0.93	0.94	0.91	0.94	0.93	0.94	0.93	0.92	0.92	0.93
Ef(%)	0.24	1.42	6.40	7.01	38.25	12.11	4.93	11.39	7.08	14.55	23.51
Mean	68.79	5659.95	12.09	4.24	0.2	1.54	0.89	3.18	29.56	16.22	15.04
Maximum	75.36	7590.95	15.63	5.2	0.38	1.91	1.05	5.22	40.2	20.11	18.97
Minimum	62.36	2742.95	10.67	3.51	0.13	1.29	0.81	2.44	23.17	14.28	13.18

\*Continuous and dashed underlines represent the highest and lowest observed values, respectively. <sup>1/</sup>Asymptotic confidence interval CI = BLUP ± 1.96 SD, where P (t ≥ 1.96) = 0.025 with ∞ degrees of freedom, in which t is a variable that follows the Student t-distribution.

The second cluster contained most of the commercial cultivars. In this group, only BRS 330 and 0307541 showed good grain yield, but these hybrids were late-flowering. With regard to the nutritional traits, almost half of the hybrids presented elevated genotypic values, with emphasis on hybrid 0577337.

The third cluster comprised the majority of hybrids with agronomic performance, among which the hybrid 0306037 stood out. Most of the hybrids from this cluster showed low performance for the nutritional traits; however, the hybrid 0307651, which belonged to this cluster, showed interesting genotypic values for some characteristics of nutritional interest and fair agronomic performance.

## DISCUSSION

### Superiority of the multi-trait model

Based on the observed accuracy values (Table 1), it can be concluded that the multi-trait model was highly efficient in predicting the genotypic value of the individual hybrids. Resende and Duarte (2007) have opined that accuracy in the intermediate phases of a breeding program should be at least 0.7, whereas in the final phases of the program it should be 0.9 or above.

The better performance of the multi-trait model than the univariate models is directly related to the correlations among the traits. As the absolute difference between the genetic and residual correlations among the traits increase, the gain in accuracy of the multivariate model over the univariate model also increases (Schaeffer, 1984; Mrode 2005).

Because the traits evaluated in the present study showed considerable genetic and residual correlations (Almeida Filho JE et al, unpublished results), the superiority of the multi-trait model corroborates theoretical studies on the methodology of mixed models and, thus, the multi-trait model was used for genetic evaluation.

Viana et al. (2010) observed that the multivariate model efficiently addressed the characteristics of grain yield and expansion capacity in popcorn kernels for the selection of families of half-siblings, whereas for selection within families the multivariate model was equivalent to the univariate model.

### Genetic evaluation of hybrids

The results of the genotypic values for grain yield and flowering showed evidence of the efficiency of the grain-sorghum breeding program for agronomic performance, given that among the genotypes with highest grain-yield, the vast majority are experimental hybrids.

Additionally, the experimental hybrids had, in general, early flowering, which is in line with the goal of the program, because, it is preferable that the grain-sorghum hybrids remain on the field for less time to avoid biotic and abiotic stress. Moreover, farmers usually get better prices for their grains with early harvest due to the greater flexibility in negotiations.

In contrast, it is also important to have cultivars requiring different days to flower, because this allows the farmers to distribute the harvest, resulting in fewer losses during this phase.

None of the hybrids with greater agronomic potential showed high performance for nutritional properties. This indicates that in order to obtain a cultivar with agronomic and nutritional potential from the evaluated hybrids, it would be necessary to make crossings among individuals and perform selection in the segregating population; however, parents must be chosen that are genetically divergent and aggregate favorable alleles for all the traits of interest.

## Genetic diversity

The genotypic effects to estimate the dissimilarity matrix were adopted with an aim of obtaining an essentially genotypic measure of divergence. For this purpose, squared Mahalanobis distance was used to obtain an estimate weight by the residual co-variances among the traits (Cruz et al., 2011).

The hierarchical clustering of the distance matrix by the UPGMA method was satisfactory; however, the cophenetic correlation of the present study (0.7) was lower than that obtained by Ritter et al. (2007) using AFLP markers in sorghum accessions, a value considered satisfactory by these authors for the interpretation of the cluster analysis. It should be stressed that higher values of the cophenetic correlation indicate greater consistency of the clustering pattern (Cruz et al., 2011).

Based on the clusters formed and in view of the potential genetics of the hybrids for all the traits, the generation of a genotype with favorable phenotypes for all these traits would be the foundation of a population with hybrids 0577337, 0441347, 0307651, and 0306037. Therefore, the theoretical genetic divergence between the founders would be taken into account, and it would also be expected that the population would have favorable alleles for all the traits.

In a breeding program, a breeder seeks parents that have a wide genetic divergence between themselves also in addition to having genetic merit for the traits of interest. According to Thul et al. (2009), this type of study is of great importance, because the identification and characterization of good combinations between parents provide the basis for selection in segregating generations, as well as for the exploitation of heterosis.

## Conflicts of interest

The authors declare no conflicts of interest.

## ACKNOWLEDGMENTS

The first author acknowledges a graduate scholarship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Research supported by Embrapa.

## REFERENCES

- Africa Harvest Biotech Foundation International (AHBFI) (2010). Africa biofortified sorghum project: five-year progress report. Available at [[http://issuu.com/africaharvest/docs/final\\_document](http://issuu.com/africaharvest/docs/final_document)] Accessed March 3, 2013.
- AOAC (Association of Official Analytical Chemists) (2000). Official Methods of Analysis of AOAC International, 17th edn, AOAC, Arlington.
- Awika JM and Rooney LW (2004). Sorghum phytochemicals and their potential impact on human health. *Phytochemistry* 65: 1199-1221. <http://dx.doi.org/10.1016/j.phytochem.2004.04.001>
- Cruz CD (2013). GENES-A software package for analysis in experimental statistics and quantitative genetics. *Acta Sci. Agron.* 35: 271-276. <http://dx.doi.org/10.4025/actasciagron.v35i3.21251>
- Cruz CD, Ferreira FM and Pessoni LA (2011). *Biometria aplicada ao estudo da diversidade genética*, Suprema Gráfica Editora, Visconde de Rio Branco. 620 p.
- Gilmour AR, Gogel BJ, Cullis BR and Thompson R (2009). ASReml User Guide Release 3.0., VSN International Ltd, Hemel Hempstead, UK.
- Henderson CR and Quaas RL (1976). Multiple trait evaluation using relatives' records. *J. Anim. Sci.* 43: 1188-1197.
- Jones DB (1941). Factors for converting percentages of nitrogen in foods and feeds into percentages of proteins. *USDA-Circ* 183: 1-22. Available at [<http://naldc.nal.usda.gov/download/CAT87213109/PDF>] Accessed June 21, 2012.
- Milligan GW and Cooper MC (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 50: 159-179. <http://dx.doi.org/10.1007/BF02294245>

- Mojena R (1977). Hierarchical grouping method and stopping rules an evaluation. *Comput. J.* 20: 359-363. <http://dx.doi.org/10.1093/comjnl/20.4.359>
- Mrode RA (2005). Linear Models for the Prediction of Animal Breeding Values. 2nd edn. CABI Publishing, Oxfordshire.
- Mutisya J, Sun C, Rosenquist S, Baguma Y, et al. (2009). Diurnal oscillation of SBE expression in sorghum endosperm. *J. Plant Physiol.* 166: 428-434. <http://dx.doi.org/10.1016/j.jplph.2008.06.017>
- Olembo KN, M'mboyi F, Kiplagat S, Sitienny JK, et al. (2010). Sorghum breeding in sub-Saharan Africa: The success stories. The African Biotechnology Stakeholders Forum, Nairobi.
- Piepho HP, Möhring J, Melchinger AE and Büchse A (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161: 209-228. <http://dx.doi.org/10.1007/s10681-007-9449-8>
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available at [<https://www.Rproject.org/>].
- Resende MDV and Duarte JB (2007). Precision and quality control in variety trials. *Pesq. Agropec. Trop.* 37: 182-194.
- Ritter KB, McIntyre CL, Godwin ID, Jordan DR, et al. (2007). An assessment of the genetic relationship between sweet and grain sorghums, within *Sorghum bicolor* spp *bicolor* (L) Moench, using AFLP markers. *Euphytica* 157: 161-176. <http://dx.doi.org/10.1007/s10681-007-9408-4>
- Schaeffer LR (1984). Sire and cow evaluation under multiple traits model. *J. Dairy Sci.* 67: 1567-1580. [http://dx.doi.org/10.3168/jds.S0022-0302\(84\)81479-4](http://dx.doi.org/10.3168/jds.S0022-0302(84)81479-4)
- Thul ST, Lal RK, Shasany AK, Darokar MP, et al. (2009). Estimation of phenotypic divergence in a collection of capsicum species for yield-related traits. *Euphytica* 168: 189-196. <http://dx.doi.org/10.1007/s10681-009-9882-y>
- Viana JMS, Sobreira FM, Resende MDV and Faria VR (2010). Multi-trait BLUP in half-sib selection of annual crops. *Plant Breed.* 129: 599-604. <http://dx.doi.org/10.1111/j.1439-0523.2009.01745.x>