Breeding strategies for tropical maize targeting *in vivo* haploid inducers

Camila Bastos Ribeiro[1]*, Breno Alvarenga Rezende[1], Júlio Sílvio de Sousa Bueno Filho[2], Éder David Borges da Silva[3], Kaio Olímpio das Graças Dias[1], José Airton Rodrigues Nunes[1], Pedro Laurence Silva Carvalho[1], Suellen Finamor Mota[1], Fernanda de Oliveira Bustamante[1], Rafael Parreira Diniz[1], Carlos Eduardo Caixeta de Castro[1] Renato Barbosa Camargos[1], Josiel Nascimento[1] and João Cândido de Souza[1]

Abstract: *The objective of this study was to compare the selection of plants bred by different pedigree methods using selection among, among and within and only within families. The haploid induction rate of 14 $S_{0:1}$ and seven $S_{2:3}$ families, all crossed with the single-cross hybrid GNZ9501, was evaluated. An experimental area of the Department of Biology of the Federal University of Lavras (UFLA), in Lavras, Minas Gerais, in the growing seasons 2012/2013 and 2014/2015, was used for the experiments. In each growing season, one experiment per was carried out, arranged in a complete randomized design, with one and two replications, respectively. Haploid induction was most effective in the families 2 and 6 in both growing seasons. Selection among and within families resulted in higher genetic gains for haploid induction. The results indicated a high genetic variability for haploid induction rate in plants within families.*

Keywords: *Zea mays, selection among and within family, haploid induction.*

INTRODUCTION

In the past decade, the doubled-haploid (DH) technology based on *in vivo* haploid induction (HI) has become one of the most important tools in maize breeding and has come to replace the conventional method of developing lines by recurrent selfing. The use of DH technology became possible once many maize haploid inducer lines were developed, in particular in temperate climate regions. Maize haploid inducer lines, when used as pollinators, trigger the production of seeds with a haploid embryo at a mean rate of 8% due to a hetero-fertilization together with failed egg-sperm cell fusion (Tian et al. 2018). In spite of the successful development of maize haploid inducers in the tropics (Chaikam 2012, Couto et al. 2020), compared to temperate conditions, data on this DH process are still scarce.

One way to enhance the efficiency of the DH technology is to raise and maintain the stability of the haploid induction rate by using different germplasm sources as females/donor genotypes for haploid induction. Another way is to use the spontaneous doubling in the donor genotype used in the induction

*Corresponding author:
E-mail: camila.cbra2@gmail.com
ORCID: 0000-0002-1948-5234

[1] Universidade Federal de Lavras, Departamento de Biologia, 37.200-900, Lavras, MG, Brazil
[2] Universidade Federal de Lavras, Departamento de Estatística, 37.200-900, Lavras, MG, Brazil
[3] Du Pont Pioneer, 75.904-223, Rio Verde, GO, Brazil

for having more efficiency using DH production (Arshadullah et al. 2018, Chaikam et al. 2019b, Boerman 2020). The *in vivo* method has been very successful in maize in recent growing seasons and has been extensively used in commercial maize breeding programs. Initially, haploid plants occurred naturally in maize fields at a frequency of 0.01% (Chase 1951). The discovery of maize lines Stock 6 and W23 (Coe 1959), and other haploid-inducing lines, such as ZSM, KMS and MHI (Chalyk 1999) revolutionized the application of the DH technology in maize breeding. This revolution is expressed in the rate of 8% already achieved in many breeding programs. There are also other factors that reduce the production costs per DH line as chromosome doubling protocols, research on genetic improvement in spontaneous chromosome doubling (Chaikam et al. 2019b) and more efficiency methodologies to select haploid kernels considering R1-nj expression based in convolutional neural networks for example (Altuntas et al. 2019).

All of the above inducers were developed from temperate germplasm and evaluated for HI mainly under temperate conditions, and some inheritance studies have suggested polygenic control of *in vivo* induction of maternal haploids (Rober et al. 2005). Nair et al. (2020) did a genetic dissection study for HIR in 671 tropical inbred lines. They revealed that the maternal influence of HIR is controled by a few moderate and many small effect QTLs. In a genome-wide association study (GWAS), Hu et al. (2016) analyzed the genetic basis underlying haploid induction in maize. In the GWAS, these authors used 56,110 single nucleotide polymorphism (SNPs) data in 53 maize haploid inducer liners from 29 different breeding programs and 1482 non-inducer inbred lines. This study provided evidence for the hypothesis of Prigge et al. (2012), stating that QTL qhir1 is required for haploid induction. Therefore, it is imperative to evaluate and establish maize haploid inducers for the tropics, with desirable agronomic traits associated with high levels of haploid induction.

To develop maize haploid inducer lines with superior performance, efficient strategies to obtain significant genetic gains for the traits under selection must be applied in a breeding scheme. The efficiency of selection can be maximized by using robust statistical methods, e.g. analysis by a mixed binomial model, especially in cases of experimental evaluations with high data imbalance (Jaeger 2008, Stroup 2013), as often found in the case of haploid induction. This trait is related to an adequate fertilization management and various other environmental factors. Moreover, specifically the use of binomial models has been intensified in other related studies (Wegenast et al. 2010, Batistelli et al. 2013, Couto et al. 2015). Couto et al. (2020) used the multinomial model to evaluate haploid induction rate, diploid seed rate and inhibition seed rate.

In spite of reports of high rates of haploid inducers (up to 8%), information about genetic parameters and on, breeding strategies to develop lines and/or hybrids with high haploid induction rates in tropical maize remain scarce. In this context, this study compared the selection of plants using three different breeding strategies involving the pedigree methods among families, among and within families and within families, using the $S_{0:1}$ and $S_{2:3}$ generations of haploid inducers in tropical maize.

MATERIAL AND METHODS

Experimental area and evaluated genotypes

The experiments were carried out in the experimental area of the department of biology of the Federal University of Lavras (UFLA), in two growing seasons (2012/2013 and 2014/2015). In 2012/2013, the HIR of 14 $S_{0:1}$ families of gynogenetic haploid inducers was evaluated. These families were crossed with the commercial hybrid GNZ9501 in a completely randomized experiment with one replication. The families were selected from a S_0 population, derived from a cross of the Russian inducer Krasnodar Embryo Marker Synthetic or KEMS (Battisteli et al. 2013, Ribeiro et al. 2018) with other tropical lines from diverse origins. This line KEMS is also designated as ZMK 1 and is widely used for selection for haploid maize at the national and international level (Shatskaya 2010). Seed of the families and hybrid was sown on November 23, 2012, in 3-m rows, at a row spacing of 80 cm and plant spacing of 50 cm. Each $S_{0:1}$ family was selfed and crossed with GNZ9501. To ensure synchronous flowering, the seeds of GNZ9501 were planted in weekly intervals on four dates (November 23 and 30 and December 7 and 14). All $S_{0:1}$ plants were tagged in the field according to their family and sowing date.

In the 2013/2014 growing season, the best families were selfed again, establishing the $S_{2:3}$ generation. The generation was only advanced because if the haploid induction test and selfing were performed together, the pollen would have

to be sufficient for the test replications as well as selfing. Consequently, to maintain the advanced generations as well as a reliable analysis of the contribution of additive variance to the trait, selfing was only performed in 2013/2014. In 2014/2015, another completely randomized experiment with two replications evaluated the haploid induction rate (HIR) of the seven $S_{2:3}$ families in crosses with the single-cross hybrid GNZ9501. The replication consisted of two different GNZ9501 plants pollinated with pollen of the same $S_{2:3}$ plant. Each $S_{2:3}$ plant was identified according to its family and selfed. This experiment was initiated on November 28, 2014. In 6-m rows spaced 80 cm apart, two seeds were in 30 holes per row to warrant a high germination percentage. To ensure synchronous flowering, GNZ9501 seeds were sown in weekly intervals on four dates: November 28, December 5, 12 and 19, 2014. In both growing seasons, fertilization at sowing consisted of 500 kg ha^{-1} of 10-30-10 NPK and side dressing of 500 kg ha^{-1} of 20-0-20 NPK fertilizer. Other cultural practices were carried out as recommended for maize.

The seeds resulting from the crosses were evaluated for the purple color of the endosperm and embryo. Seeds with purple and white endosperm/embryo were considered haploid, according to the methodology of Chase and Nanda (1965). The seeds with purple endosperm/embryo and those without purple color were also counted, while seeds with fungus infestation were eliminated. In this way, the total number of seeds per plant was determined. The following parameter was analyzed in this study: Haploid induction (HI) rate = (number of haploid seeds/ total number of seeds) x 100, where: number of haploid seeds = number of seeds with purple endosperm and white embryo; total number of seeds = haploid seeds + purple seeds + non-purple seeds

Statistical analysis

Analyses were performed using generalized linear mixed models (GLMMs). The following binomial GLMMs with the *logit* link function were considered to evaluate the HIR in the 2012/2013 and 2014/2015 growing seasons, respectively:

Logit $(p) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \mu + p_j + e_{ij}$ where: μ is the intercept; p_j the random effect of family j and e_{ij} the random effect of the individual plant i and family j; *Logit* $(p) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \mu + p_j + p(d_i) + r_k + e_{ijk}$

In which: μ is the intercept; p_j the random effect of family j; $p(d_i)$ the random effect of the individual plant i and family j; the random effect of replication k and the random effect of the triple interaction among individual plant i, replication k and family j;

Where $R_{ij}/u_{ij} \sim \frac{Binomial\ (m_{ij},\pi_{ij})}{m_j}$ and: R_{ij}/u_{ij} corresponds to the observed proportion in the plot occupied by the individual plant i and family j; m_{ij} number of haploids; π_{ij}: haploid induction/total number of seeds and m_i: total number of seeds;

Significance of effects of the GLMMs were tested by the chi-square test at 5% probability. For the analysis, the R software package lme4 (R Core Team 2019) was used (Bates et al. 2015).

Comparison of breeding selection strategies

Three pedigree selection strategies were compared using the genotypic value of the haploid induction rate for individual plants and/or families (BLUP), considering different selection indices, according to the following methodologies using Genes software (Cruz 2013):

Selection among families: the best six families for haploid induction rate of each year were selected, to compare the same number of families among seasons. All plants of each family were considered.

Selection among and within families: two selection indices were considered in each year. In 2012/2013, one individual plant was selected within 12 families (12x1) and three individual plants within six families (6x3). In 2014/2015, the best individual plant within the best six families (6x1) and three plants within three families (3x3) were selected.

Selection within families: three selection indices: 50%, 20% and 10% of all plants were considered in both seasons. In this case, the families were irrelevant. In other words, the plants were ranked and selected to estimate the genetic gain, without considering family data.

Estimate of genetic and phenotypic parameters

For the logit function, the environmental variance is an approximation of the variance of the logit distribution. Hence, heritability was computed by the following expressions:

Heritability within families: $h^2_d = \dfrac{\sigma^2_{Gd}}{\sigma^2_d}$, where $\sigma^2_d = \sigma^2_{Gd} + \sigma^2_e$ and $\sigma^2_e = \pi^2/3$

The genetic variance within families (σ^2_{Gd}) could only be calculated for the data of 2014/2015, because of the existence of replications.

Heritability among families: $h^2_{mp} = \dfrac{\sigma^2_{families}}{\sigma^2_{totalF}}$, where $\sigma^2_{totalF} = \sigma^2_{families} + \sigma^2_e$ and $\sigma^2_e = \pi^2/3$

Predicted genetic gain

The predicted genetic gains (GG) were estimated for each selection strategy, according to the following expressions proposed by Furtini et al. (2012):

a) Selection among families: $SG_{among} = ds_{BLUPamong} \times h^2_{mp}$

In which: $ds_{BLUPamong}$: genotypic value of the best families (according to the selection index) – mean BLUP of all families and h^2_{mp}: heritability among families

b) Selection among and within families

In 2012/2013: $SG_{total} = SG_{among} + SG_{within}$, $SG_{among} = ds_{BLUPwithin} \times h^2_{mp}$, where: $SG_{within} = ds_{BLUPwithin}$ and $ds_{BLUPwithin}$ is the genotypic value of the best individual plants – mean genotypic value of all families. Therefore, when individual BLUP is used, the heritability is equal to one.

In 2014/2015: $SG_{total} = SG_{among} + SG_{within}$, $GG_{among} = ds_{BLUPamong} \times h^2_{among}$, $GG_{within} = ds_{BLUPwithin} \times h^2_{within}$ and h^2_{within}: heritability within families.

c) Selection within families

In 2012/2013: $GG_{within} = ds_{BLUPwithin}$ and in 2014/2015: $GG_{within} = ds_{BLUPwithin} \times h^2_{within}$

RESULTS AND DISCUSSION

The mean HIR is one of the main parameters analyzed for decision-making in haploid- inducing maize breeding programs. The mean HIR of each family is shown in Table 1. The mean HIR was different among seasons, in that it decreased to half the value from the first to the second growing season. The HIR among the 14 families in 2012/2013 ranged from 0% to 7.92%. However, in 2014/2015, it was far lower (0 - 2.71%). Lines 7 and 13 had an overall mean of 1.56% and 0.61% in 2012/2013 and zero HIR in 2014/2015, contributing to reduce the overall mean.

Moreover, the discrepancy between HIR in 2012/2013 and 2014/2015 can be explained by natural selection disfavoring the haploidy-inducing gametes during selfing. From an evolutionary point of view, a higher proportion of haploids in the family caused by high HIR of the pollinator would result in reduced plant vigor, because haploid maize plants are less vigorous, often male sterile, and therefore generally less likely to produce progenies than diploid maize plants (Couto et al. 2020), which explains the action of natural selection (Prigge et al. 2012). The latter may also explain the difficulties of maintaining haploid inducers described by maize breeders.

The best haploid inducer was family 6 in both growing seasons (overall mean HIR = 5.32%). Haploid-inducing plants with 6 and 15% HIR were described in other studies under temperate (Rober et al. 2005, Prigge et al. 2012) and tropical climate conditions (Prigge et al. 2011). A partnership of the University of Hohenheim, Germany, with CIMMYT, Mexico, concluded that haploid-inducing lines and hybrids with HIR of 8 to 12%, combining favorable agronomic traits, can be developed in tropical environments (Chaikam et al. 2012). Almeida et al. (2020) demonstrated in their genomic selection study for HIR that HIR can be improved without negatively impacting agronomic performance in temperate germplasm.

The deviance analysis of HIR is shown in Table 2. The sowing date effect (first and second date) in 2012/2013 was not significant at 5% probability and therefore not considered in the statistical model. There was a significant difference

Table 1. Weighted mean and sample size (N) of haploid induction rate (HIR) of 14 and 7 families in 2012/2013 and 2014/2015, respectively

Family name	HIR (%)		
	2012/2013	2014/2015	Family mean
1	0.86		
2	2.3	0.74	1.52
3	2.5	0.34	1.42
4	0.88	0.52	0.7
6	7.92	2.71	5.32
7	1.56	0	0.78
8	1.76	1.53	1.65
9	2.5		
10	0		
13	0.61	0	0.31
15	0		
18	1.16		
20	0.3		
21	0		
Mean	1.60 (1.59-1.65[*])	0.84 (0.74-1.05)	1.67 (1.65-1.75)
Sample Size	92	148	

[*] Mean confidence interval (95% probability).

among the families and plants in both growing seasons, indicating the existence of variability among them, which is an essential condition for selection (Table 2). The heritability among families was greater in 2014/2015 (11.33%) than in 2012/2013 (9.37%). In 2012/2013, the within-family heritability (h^2) could not be estimated, because no replication of HIR of the same plant was available. A high within-family variability was indicated by the h^2 estimates in 2014/2015. The estimates based on plants as selection unit, in other words, within family, were higher in 2014/2015 (22.04%) than the among – family h^2 estimates (11.33%). These h^2 values showed that the phenotypic values of plants are good predictors of genotypic values and that within-family selection is efficient. To the best of our knowledge, there are few reports about HIR among and within family

Table 2. Results of the analysis of deviance of the haploid induction rate (HIR) of 14 $S_{0:1}$ families, evaluated in 2012/2013 and seven $S_{2:3}$ families in 2014/2015

Factors	2012/2013		
	df	Deviance	h^2(%)
Family	13	0.34*	9.37
Family x Plant	89	2.33	
	2014/2015		
Family	6	0.42*	11.33
Plant/Family	196	0.93*	22.04
Replication	1	0.0001	
Family x Plant x Replication	142	6.57	

[*] Significant differences according to the χ² test at 5% probability.

was also estimated by Prigge et al. (2012) and the F₂, F₂:₃ and F₂: HIR heritability estimates in the literature. A h^2 of 46% for HIR evaluated in crosses of the hybrid inductor RWSxUH400 with 45 single-cross hybrids was reported by Prigge et al. (2011). Heritability for ₄ generations were evaluated for HIR in CAUHOI X UH400, F₂ and F₂:₃ generations in 1680 x UH400, and the F₂:₃ generation in CML395 x UH400 and CML495 x UH400 (CML495-F₃). These authors observed higher h^2 among (80%) than within families (approximately 70%).

For ranking purposes and family selection, the Best Linear Unbiased Prediction (BLUP) based on the genotypic value of HIR was considered (Table 3). In both growing seasons, the families 2 and 6 were the most promising while family 8 also had a good performance. The individual BLUP selection method tends to concentrate the highest number of selected plants in the larger families. This is not always realistic because the best plant of each progeny can be the second, third or fourth of other families. It is possible to obtain and exploit genetic variability for HIR in maize breeding programs to raise the HIR and simultaneously, select for other desirable agronomic traits to facilitate cultivation of these plants in the tropics. For breeding purposes, the strategy that results in higher genetic gains (GG) must be used. Since information regarding breeding for HIR in maize is scarce, we compared the strategies

among families, among and within families and within families in this study.

Based on the assumption that the BLUP estimates are closest to the true genotypic value (Bernardo 2010), and that the occurrence of unbalanced data is significant for the estimate of the selection differential (SD), we used the families and individual BLUPs (Table 4). We also used a binomial model which is more recommended for this a binomial trait, as used in ther studies (Couto et al. 2020). In 2014/2015, SD was weighted by heritability to predict the GG, since it was possible to calculate the variance within families, as previously explained. The estimates of predicted GG were highest and most consistent for selection among and within families, in both growing seasons. In addition, various selection indices were evaluated with each selection strategy. By selection among and within families, the possibilities of selecting more families and fewer plants within each family or a greater number of plants in a smaller number of families were compared. According to the predicted GG, the best option is to select a larger number of plants in the best family. In 2012/2013, 6 families and 3 plants (GG = 0.156) was the best option and in 2014/2015, 3 families and 3 plants (GG = 0.172). However, it should be emphasized that, despite the low GG estimates, practicing selection within families based on a high selection index is a good option, particularly in the case of plants with a lower inbreeding level, e.g., in the $S_{0:1}$ generation. This may have been the case since in this generation, the additive variance (V_A) within is 0.5, while in $S_{2:3}$ it is only 0.125 V_A. Thus, selection within $S_{0:1}$ families is justified. Consequently, it is better to use a selection index of 10% than of 20 or 50% for HIR.

In view of the low HIR heritability in both growing seasons, the sample size should be as large as possible to consequently increase the selection index. Another question for low heritability would be the use of an F_1 as donor genotype in the induction, leading to a lower genetic variance values due to a worst exploration of genetic variability. However, Couto

Table 3. Genotypic value (BLUP) of 14 and 7 families evaluated for haploid induction rate (HIR) in the 2012/2013 and 2014/2015 growing seasons, respectively

	Genotypic value	
Family	**2012/2013**	**2014/2015**
1	0.60	
2	1.12	0.15
3	1.12	0.03
4	0.68	0.15
6	1.99	1.39
7	0.47	0.08
8	0.82	0.12
9	0.72	
10	0.67	
13	0.65	0.06
15	0.63	
18	0.52	
20	0.63	
21	0.76	
Mean	0.81	0.28

Table 4. Families ranked from best to worst for haploid induction rate (HIR) and estimates of predicted genetic gains (GG) for different breeding strategies, evaluated in 2012/2013 and 2014/2015

	2012/2013	
Selection strategy	**Selected families**	**GG**
Among families (6 families)	6,2,3,8,21,9	0.005
Among and within families (12x1)[1]	6,2,3,8,21,9,1,15,20,13,10,4	0.078
Among and within families (6x3)[2]	6,2,3,8,21,9	0.156
Within families (50% or 46 plants)		0.036
Within families (20% or 18 plants)		0.132
Within families (10% or 9 plants)		0.244
	2014/2015	
Among families (6 families)	6,4,2,8,7,13	0.0001
Among and within families (6x1)[3]	6,4,2,8,7,13	0.024
Among and within – families (3x3)[4]	6,4,2	0.172
Within families (50% or 74 plants)		0.005
Within families (20% or 28 plants)		0.020
Within families (10% or 14 plants)		0.027

[1] (12x1) – selection of 12 best families and the best individual plant within each family;
[2] (5x3) – selection of 5 best families and the best 3 individual plants within each family;
[3] (6x1) – selection of 6 best families s and the best individual plant within each family;
[4] (3x3) – selection of 3 best families and the best 3 individual plants within each family.

et al. (2020) compared F_1 and F_2 generations in a induction containing 5 tropical commercial hybrids and they observed a higher genetic gain using F_1 generation compared to F_2 generation of the respective F_1s. This can emphasize the importance of sample size in this kind of studies, specially because in DH programs, we evaluate population effects. There are many other factors envolved in a DH production. The maternal influence in the induction crosses (Nair et al. 2020) is also involved in the DH process. The F_1 used in this study was a commercial hybrid with late cycle that can have lower HIR regarding the male used for induction. Other point is the R1-nj inhibition, specially in tropical germplasm. Khulbe et al. (2019) observed that the phenotype expression of R1-nj has complex nature and needs to have further investigation involving larger sets of germplasm.

ACKNOWLEDGMENTS

REFERENCES

Almeida VC, Trentin HU, Frei UK and Lubberstedt T (2020) Genomic prediction of maternal haploid induction rate in maize. **The Plant Genome 1**: 1.

Altuntas Y, Comert Z and Kocamaz AF (2019) Identification of haploid and diploid maize seeds using convolutional neural networks and a transfer learning approach. **Computers and Electronics in Agriculture 163**: 1-11.

Arshad Ullah M, Suhaib M, Baber R, Usama M, Uz-Zaman B, Mahmood IA and Hyder SI (2018) Growth of *Chenopodium quiona* wild under naturally salt affected soils. **Malaysian Journal of Sustainable Agriculture 53:** 258.

Bates D, Machler M, Bolker BM and Walker SC (2015) Fitting linear mixed-effects models using lme4. **Journal of Statistical Software 67**: 1-48.

Batistelli GM, Von Pinho RG, Justus A, Couto EGO and Balestre M (2013) Production and identification of doubled haploids in tropical maize. **Genetics and Molecular Research 12**: 4230-4242.

Bernardo R (2010) **Breeding for quantitative traits in plants**. Woodbury, Stemma, 390p.

Boerman NA, Frei UK and Lubberstedt T (2020) Impact of spontaneous haploid genome doubling in maize breeding. **Plants 9:** 369.

Chaikam V, Gowda M, Nair SK, Melchinger AE and Boddupalli PM (2019a) Genome-wide association study to identify genomic regions influencing spontaneous fertility in maize haploids. **Euphytica 215**: 138.

Chaikam V, Mahuku G and Prasanna BM (2012) Design and implementation of maternal haploid induction. In Prasanna BM, Chaikam V and Mahuku G (eds) **Doubled haploid technology in maize breeding theory and practice**. International Maize and Wheat Improvement Center, Mexico, p. 14-19.

Chaikam V, Molenaar W, Melchinger AE and Boddupalli PM (2019b) Doubled Haploid technology for line development in maize: technical advances and prospects. **Theoretical and Applied Genetics 132**: 3227-3243.

Chalyk ST (1999) Creating new haploid-inducing lines of maize. **Maize Genetics Cooperation Newsletter 73**: 53-54.

Chase SS (1951) Production of monozygous diploids of maize from monoploids. **Agronomy Journal 44**: 263-267.

Chase SS and Nanda DK (1965) Comparison of variability in inbred lines and monoploid-derived lines of maize (*Zea mays L.*). **Crop Science 5**: 275-276.

Coe EH (1959) A line of maize with high haploid frequency. **America Naturalist 93**: 381-382.

Couto EGO, Cury MN, Souza MB, Granato ISC, Vidotti MS, Garbuglio DD, Crossa J, Burgueno J and Fritsche-Neto R (2020) Effect of F_1 and F_2 generations on genetic variability and working steps of doubled haploid production in maize. **Plos One 1**: 1.

Couto EGO, Von Pinho EVR, Von Pinho RG, Veiga AD, Bustamante FO and Dias KOG (2015) In vivo induction and efficiency of two chromosome duplication protocols in tropical maize. **Ciência e Agrotecnologia 39**: 435-442.

Cruz CD (2013) GENES - a software package for analysis in experimental statistics and quantitative genetics. **Acta Scientiarum 35**: 271-276.

Furtini IV, Ramalho MAP, Abad JIM and Aguiar AM (2012) Effect of different progeny test strategies in the performance of eucalypt clones. **Silvae Genetica 61:** 116-120.

Hu H, Schrag TA, Peis R, Unterseer S, Schipprack W, Chen S, Lai J, Yan J, Prasanna BM, Nair SK, Chaikam V, Rotarenco V, Shatskaya OA, Zavalishina A, Scholten S, Schön C-C and Melchinger AE (2016) The genetic basis of haploid induction in maize identified with a novel genome-wide association method. **Genetics 202:** 1267-1276.

Jaeger FT (2008) Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. **Journal of Memory and Language 59**: 434-446.

Khulbe RK, Pattenayak A and Panday V (2019) R1-nj expression in parental inbreds as a predictor of amenability of maize hybrids to R1-nj based doubled haploid development. **Indian Society of Genetics & Plant Breeding 79**: 678-664.

Nair SK, Chaikam V, Gowda M, Hindu V, Melchinger AE and Boddupalli PM (2020) Genetic dissection of maternal influence on in vivo haploid

induction in maize. **The Crop Journal 8**: 287-298.

Prigge V, Sanchez C, Dhillon BS, Schipprack W, Araus JL, Banziger M and Melchinger AE (2011) Doubled haploids in tropical maize: I., effects of inducers and source germplasm on *in vivo* haploid induction rates. **Crop Science 51**: 1498-1506.

Prigge V, Xu X, Li L, Babu R, Chen S, Atlin GN and Melchinger AE (2012) New insights into the genetics of *in vivo* induction of maternal haploids, the backbone of doubled haploid technology in maize. **Genetics 190**: 781-793.

R Core Team (2019) Foundation for statistical computing. Available at <http://www.R-project.org/>. Accessed on November 20, 2019.

Ribeiro CB, Pereira FC, Nóbrega Filho L, Rezende BA, Dias KOG, Braz GT, Ruy MC, Silva MB, Cenzi G, Techio VH and Souza JC (2018) Haploid identification using tropicalized haploid inducer progenies in maize. **Crop Breeding and Applied Biotechnology 18**: 16-23.

Rober FK, Gordillo GA and Geiger HH (2005) In vivo haploid induction in maize-performance of new inducers and significance of doubled haploid lines in hybrid breeding. **Maydica 50**: 275-283.

Shatskaya OA (2010) Haploinductors isolation in maize: three cycles of selection on high frequency of induction of matroclinal haploids. **Russian Journal of Agricultural Biologia 5**: 79-86.

Stroup WW (2013) Rethinking the analysis of non-normal data in plant and soil science. **Agronomy Journal 107:** 811-827.

Tian X, Qin Y, Chen B, Liu C, Wang L, Li X, Dong X, Liu L and Chen S (2018) Hetero-fertilization together with failed egg-sperm cell fusion supports single fertilization involved in *in vivo* haploid induction in maize. **Journal of Experimental Botany 9:** 4689-4701.

Wegenast T, Utz HF, Friedrich HL, Maurer HP, Dhillon BS and Melchinger AE (2010) Hybrid maize breeding with doubled-haploids: V. Selection strategies for testcross performance with variable sizes of crosses and S1 families. **Theoretical and Applied Genetics 120**: 699.