Article

SAS code to analyze a complete dialelic and heterosis. An environment

Delfina de Jesús Pérez López
Claudia Saavedra Guevara
Martín Rubí Arriaga
J. Ramón Pascual Franco Martínez
Francisco Gutiérrez Rodríguez
Andrés González Huerta[§]

Center for Research and Advanced Studies in Phytomeroration-Faculty of Agricultural Sciences-Autonomous University of the State of Mexico-University *Campus* 'El Cerrillo'. El Cerrillo Piedras Blancas, Toluca, State of Mexico, Mexico. AP. 435. Tel. 722 2965531, ext. 148. (djperezl@uaemex.mx; csg1003@yahoo.com; m-rubi65@yahoo.com.mx; jrfrancom@uaemex.mx; fgrfca@hotmail.com.

[§]Corresponding author: agonzalezh@uaemex.mx.

Abstract

The development of statistical analysis system (SAS) programs and their validation with freely available software is essential when there are no financial resources to acquire the license of an appropriate statistical package. In this study a code for SAS is presented and its validation is performed with the program proposed by Zhang and Kang (1997), modified by Saavedra (2019). The code generates an analysis of variance with partition of the effects of treatments on parents (P), direct crosses (CD), reciprocal crosses (CR), P *vs* crosses, and CD *vs* CR. In addition to generating the comparison of treatment means with the Tukey test, the genetic effects for parents or for their crosses are estimated (Gi, Sij, Rij, Mi); as well as, those of heterosis with the average of both parents or with the best of them. Since both codes only coincide in the calculation of the previously indicated genetic effects, their simultaneous application is suggested to carry out a complete analysis of Griffing's method 1 (1956a, b). The code that has been proposed will be especially useful for plant breeders and geneticists and especially for undergraduate and graduate level biological and agricultural science students with little training in the programming language at SAS.

Keywords: Griffing method 1, model 1, randomized complete blocks, Tukey test.

Reception date: February 2020
Acceptance date: April 2020

Introduction

Dialectic crosses were designed before the 1950's, but soon became a powerful tool for plant and animal breeders, who to recognize the merit of various parents evaluated their progenies through the effects and variances of general combinatorial aptitude (ACG) and specific (ACE) (Sprague and Tatum, 1942; Griffing, 1956a, b; González *et al*., 2007a, b). These define new heterotic patterns or a segregating population from which it is possible to isolate again outstanding plants, predict the response to the selection or the behavior of hybrids or synthetics formed with new lines (Hallauer and Miranda, 1988; Christie and Shattuck, 1992; González *et al*., 2007a, b).

Analysis of a complete diallel crosses experiment without a personal computer (PC) is laborious and to save time, there are several statistical packages such as SAS (https://www.sas.com/store/index.ep), Excel (Microsoft Office), Indostat (https://www.indostat.org), AGD-R (https://data.cimmyt.org/dataset.xhtml?persistentld= hdl:11529/10202), Agrobase II, generation (http://www.agronomix.com), PB Tools (https://pbtools.software.informer.com/2.0/), TNAUSTAT (https:// sites. google.com/site/tnaustat) and GSCA (https://bioseqdata.com/gsca/gsca. htm), among others; of these, only Agrobase II generation and Indostat must be purchased under license with a cost higher than $1 000.00 USD, because at least three modules are required to properly operate both softwares. Although SAS is the best statistical package, it is common for breeders and geneticists to use various software to analyze data from experiments designed in the agricultural and biological sciences (Padilla *et al*., 2019a; Padilla *et al*., 2019b; Saavedra, 2019).

Also, for many users it is difficult to download free software because there is an incompatibility problem between it and their PCs, there are technical problems during the downloads, the necessary permission is not obtained, the researchers do not respond to the requests or the program does not work in versions old or recent Windows. In this context, it would be desirable to elaborate and validate some codes for SAS, for versions 6.01 or higher (SAS, 1989), that allow complementing the genetic-statistical analysis for experiments of complete dialectic crosses.

Materials and methods

Full dialectic

In methodology 1, described in Saavedra (2019), the analysis of variance (ANOVA) for a single environment contains repetitions (R), treatments (Trat) and experimental error, its statistical model corresponds to a randomized complete block design. In ANOVA, the effects of Trat are divided into progenitors (P), direct crosses (CD), reciprocal crosses (CR), P *vs* crosses and CD *vs* CR, as suggested by González *et al*. (2007b), both contrasts estimate average heterosis and maternal and non-maternal effects.

The program calculates the differences between Trat with the Tukey test (SAS, 1989). This code can be easily modified if the user requires other means comparison tests, or various regression and correlation analyzes, these analyzes can be extended to series of experiments in time and space (Saavedra, 2019).

In methodology 2, which corresponds to method 1 of Griffing (1956a, b), the ANOVA for a single trial has repetitions (R), general combinatorial aptitude (ACG), specific combinatorial aptitude (ACE), maternal effects (EM) and reciprocal effects (ER); in the series of experiments the interactions of these with only two environments could be estimated.

Also, in both cases, the effects of gi for each parent or of sij for each cross, the reciprocal and maternal effects would be estimated (Zhang and Kang, 1997). The variance and heritability components, and the prediction of hybrids and synthetics could be estimated with other programs for SAS (Martínez, 1983; González *et al*., 2007a, b; Montesinos *et al*., 2007).

Defining variables in code

In the database called "diallel" female, male, YH, YP, YM, X, Y, A, B, C, D and M are defined, in the female and male variables the combinations of each female with each male, YH, YP and YM correspond to the cross, female and male means, respectively. In X, Y the totals for each pair of CD and CR are captured. After the sum over repetitions has been done. In A, B, C, D, each line of the CD or the CR appears twice, as female and as male (Yi. or Y.i.; Yj.. or Y.j.)M is the great arithmetic mean, GI, SIJ, RIJ and MI are the same genetic effects that are estimated with the formulas proposed in method 1 of Griffing (1956a, b).

Values used in the code

In this study, 96 data were used, corresponding to four parents, their six direct crosses and their six reciprocal crosses, registered in six repetitions (Saavedra, 2019).

Results and discussion

Since its creation in 1972, SAS programs for the analysis of diallelic cross experiments have been implemented on personal computers PC's by several researchers. The great achievements that have been obtained for PC's are attributed to Schaffer and Usanis (1989); Burow and Coors (1994); Magari and Kang (1994); Zhang and Kang (1997); Martínez (1983, 1991), among others. More recently, Mastache and Martínez (1998a, 1998b, 1999a, 1999b), they refined their algorithms to obtain the best empirical linear and unbiased predictors (MPLI) of the effects of the parents, to help users with little training in programming, when using completely random designs (DCA) and complete random blocks (BCA).

Also, Mastache and Martínez (2003) obtained an integrated algorithm for its simultaneous analysis in balanced experiments for fixed or random effects models. These and other programs could also be used to validate and complement the outputs that were obtained with the code proposed in the present study (Zhang *et al*., 2005; Montesinos *et al*., 2007).

Zhang *et al*. (2005) modified the codes of Zhang and Kang (1997); in Diallel-SAS05, they discussed a more efficient program for the genetic-statistical analysis four methods of Griffing's (1956a, b), including those corresponding to designs II and III of Gardner and Eberhart (1966). This program is friendlier and easier to modify than Diallel-SAS, when parents vary from 4 to 12,

when there is no restriction on the number of environments, and when the effects and variances of ACG and ACE for parents and crosses are estimated, as well as their interactions with environments. As with other statistical packages, there are problems in deploying on personal computers with recent versions of Windows (Padilla *et al*., 2019a, b).

With program 1a the ANOVA and the comparison of means (Tukey, p= 0.01) are calculated. Since Trat and its components are considered as fixed effects, the F tests are tested with the mean square of the experimental or residual error of the model. In your code, Data, SET, IF-THEN, ANOVA, and GLM are used to define subsets of data. The user will be careful to respect the correct order in the database: P, CD and CR, the signs and the coefficients of the contrasts, as for other statistical packages, must be captured within the program. If there is any doubt to design this type of contrasts, it is suggested to consult Padilla *et al*. (2019a).

At the SAS output, if R= 6 and Trat= 4, the ANOVA corresponds to parents; your hypothesis test is not correct because it was constructed as a subset and its mean square of the residual is a fraction of the 96 data. In this context, a table of F should be consulted at this stage there are no restrictions regarding the number of variables to analyze. The code can be modified to include tests for the least significant difference (DMS or LSD), Dunnett, or mutually orthogonal contrasts, among others. With two or more variables, it is possible to modify the program to perform regression and correlation, estimate simple statistics and apply multivariate methodologies, among others.

The code corresponding to program 1a is presented below:

```
Data corn; Input rep trat PVG;Cards;
1 01 758
1 02 761
6 15 768
6 16 758;
DATA PARENTS;SET CORN;IF TRAT>4 THEN DELETE;*parents only;
DATA CD;SET CORN;IF TRAT<5 OR TRAT>10 THEN DELETE;* only direct crosses;
DATA CR; SET CORN; IF TRAT<11 THEN DELETE; *only reciprocal crosses;
PROC ANOVA DATA= CORN; CLASS REP TRAT; MODEL PVG=REP TRAT; MEANS
TRAT/TUKEY LINES ALPHA=0.01;* Analysis with 96 data;
PROC GLM DATA= CORN; CLASS REP TRAT; MODEL PVG=REP TRAT;
CONTRAST "P VS CROSSES"TRAT 12 12 12 12 -4 -4 -4 -4 -4 -4 -4 -4 -4 -4 -4 -4;
CONTRAST "CD VS CR" TRAT 0 0 0 0 1 1 1 1 1 1 -1 -1 -1 -1 -1 -1;
PROC ANOVA DATA= PARENTS; CLASS REP TRAT; MODEL PVG=REP
TRAT;*Analysis of variance for parents;
PROC ANOVA DATA=CD; CLASS REP TRAT; MODEL PVG=REP TRAT;* Analysis of
variance for direct crosses;
PROC ANOVA DATA=CR; CLASS REP TRAT; MODEL PVG=REP TRAT;*Analysis of
variance for reciprocal crosses; RUN;
```

With program 1b, estimates of genetic effects (Gi, Sij, Rij, MI) and heterosis (%) are obtained. The definition of variables before CARDS must be correctly indicated using the data in Tables 1 and 2 (Saavedra, 2019), but the denominator values of the GI, SIJ, RIJ and MI formulas should be

corrected, if R, P or both of them. At this stage it is essential to resort to the devices that Martínez (1983) used establishing a logical way of relating formulas Griffing (1956a, b) with the programming language in SAS (SAS Institute, 1989).

Table 1. Volumetric weight of the grain (g L^{-1}) of 16 crosses formed with four lines.

Crosses	R1	R2	R3	R4	R5	R6	Total	Mean
1) 1x1	758	734	750	790	758	765	4 555	759.1
2) 2x2	761	762	737	779	763	773	4 575	762.5
3) 3x3	802	812	802	838	793	782	4 829	804.8
4) 4x4	790	768	780	772	783	775	4 668	778
5) 1x2	814	792	770	781	775	755	4 687	781.1
6) 1x3	805	803	806	832	813	824	4 883	813.8
7) 1x4	791	775	777	791	795	780	4 709	784.8
8) 2x3	819	816	793	814	818	786	4 846	807.6
9) 2x4	779	778	758	798	783	755	4 651	775.1
10) 3x4	830	830	850	853	828	806	4 997	832.8
11) 2x1	774	772	786	750	794	769	4 645	774.16
12) 3x1	789	808	816	808	824	806	4 851	808.5
13) 4x1	787	815	815	825	802	796	4 840	806.6
14) 3x2	817	832	808	775	790	797	4 819	803.1
15) 4x2	756	768	756	754	753	768	4 555	759.1
16) 4x3	850	820	840	850	805	758	4 923	820.5
Total	12 722	12 865	12 644	12 810	12 677	12 495	76 033	792.01

Table 2. Values used to estimate genetic effects and heterosis.

Lines	1	2	3	4	Total
1	4 555	4 687	4 883	4 709	18 834
2	4 645	4 575	4 846	4 651	18 717
3	4 851	4 819	4 829	4 997	19 496
4	4 840	4 555	4 923	4 668	18 986
Total	18 891	18 636	19 481	19 025	76 033

Note: added over repetitions and row or column totals are the contribution of each female or male line, respectively.

In some columns, such as for the large arithmetic mean (M), which is a constant for the 12 mattings, there are duplicate values, but it is easy to establish which parent or cross they correspond to because the data is shown in descending order (González *et al*., 2007a, b; Saavedra, 2019).

The code for program 1b is presented below:

```
DATA HETERO; INPUT FEMALE MALE YH YP YM X Y A B C D M;
MP= YP+YM)/2;*to calculate the mean of the parents (MP);
BP= MAX (YP,YM);*to choose the best parent (BP);
DMP= YH-MP;* to estimate the numerator of the heterosis formula with MP;
HMP= (DMP/MP)*100;*to estimate heterosis with the mean of the parents, in %;
```

```
DBP= YH-BP;*calculate the numerator of the heterosis formula with BP;
HBP= (DBP/BP)*100;*calculates heterosis with the best father, in %;
GI= (A+B)/48 - M;*estimate the effects of gi;
SIJ= (X+Y)/12 -(A+B+C+D)/48 + M;*calculate the effects of Sij;
RIJ= (X-Y)/12;*determines the effects of rij;
MI= (A-B)/48;* calculate the effects mi;
CARDS;
1 2 781.1 759.1 762.5 4687 4645 18834 18891 18717 18636 792.01
1 3 813.8 759.1 804.8 4883 4851 18834 18891 19496 19481 792.01
1 4 784.8 759.1 778.0 4709 4840 18834 18891 18986 19025 792.01
2 3 807.6 762.5 804.8 4846 4819 18717 18636 19496 19481 792.01
2 4 775.1 762.5 778.0 4651 4555 18717 18636 18986 19025 792.01
3 4 832.8 804.8 778.0 4997 4923 19496 19481 18986 19025 792.01
2 1 774.1 762.5 759.1 4645 4687 18636 18717 18891 18834 792.01
3 1 808.5 804.8 759.1 4851 4883 19481 19496 18891 18834 792.01
4 1 806.6 778.0 759.1 4840 4709 19025 18986 18891 18834 792.01
3 2 803.1 804.8 762.5 4819 4846 19481 19496 18636 18717 792.01
4 2 759.1 778.0 762.5 4555 4651 19025 18986 18636 18717 792.01
4 3 820.5 778.0 804.8 4923 4997 19025 18986 19481 19496 792.01

TITLE 'Effects of gi, sij, rij, mi and heterosis for the general dialectic';
DATA DOS; SET HETERO; PROC PRINT; RUN;
```

ANOVA procedure

Dependent variable: PVG

Source	DF	Sum of square	Square of the mean	F-Value	Pr > F
Model	20	51191.16667	2559.55833	9.69	<.0001
Error	75	19801.82292	264.02431		
Total correct	95	70992.98958			

R-square	Coef Var	Root MSE	PVG Mean
0.721074	2.051592	16.24882	792.0104

Source	DF	ANOVA SS	Square of the mean	F-Value	Pr > F
rep	5	3365.67708	673.13542	2.55	0.0347
trat	15	47825.48958	3188.36597	12.08	<.0001

Tukey studentized range test (HSD) for PVG

Note: This test controls the rate of the probability of making an experimentwise Type I error, but usually has a higher Type II error rate than REGWQ.

Alpha	0.01
Degrees of freedom error	75
Mean square error	264.0243
Critical value of the studentized range	5.76634
Minimal significant difference	38.251

Means with the same letter are not significantly different.

Tukey	Grouping	Mean	N	trat
	A	832.833	6	10
B	A	820.500	6	16
B	A C	813.833	6	6
B D	A C	808.500	6	12
B D	A C	807.667	6	8
B D	A C	806.667	6	13
B D	A C	804.833	6	3
B D	A C	803.167	6	14
B D	E C	784.833	6	7
D	E C	781.167	6	5
D	E C	778.000	6	4
D	E	775.167	6	9
D	E	774.167	6	11
	E	762.500	6	2
	E	759.167	6	1
	E	759.167	6	15

Contrast	DF	Contrast SS	Square of the mean	F-Value	Pr > F
P VS CRUZAS	1	8075.086806	8075.086806	30.58	<.0001
CD VS CR	1	272.222222	272.222222	1.03	0.3132

Dependent variable: PVG

Source	DF	ANOVA SS	Square of the mean	F-Value	Pr > F
rep	5	1952.875000	390.575000	2.00	0.1376
trat	3	7805.458333	2601.819444	13.30	0.0002

Source	DF	ANOVA SS	Square of the mean	F-Value	Pr > F
rep	5	2869.25000	573.85000	3.57	0.0143
trat	5	15157.25000	3031.45000	18.84	<.0001

Source	DF	ANOVA SS	Square of the mean	F-Valor	Pr > F
rep	5	1741.80556	348.36111	0.90	0.4948
trat	5	16515.47222	3303.09444	8.56	<.000

Effects of Gi, Sij, Rij, Mi and heterosis for methodology 1

Obs	FEMALE	MALE	YH	YP	YM	X	Y	A	B	C	D	M	MP
1	1	2	781.1	759.1	762.5	4687	4645	18834	18891	18717	18636	792.01	760.80
2	1	3	813.8	759.1	804.8	4883	4851	18834	18891	19496	19481	792.01	781.95
3	1	4	784.8	759.1	778.0	4709	4840	18834	18891	18986	19025	792.01	768.55
4	2	3	807.6	762.5	804.8	4846	4819	18717	18636	19496	19481	792.01	783.65
5	2	4	775.1	762.5	778.0	4651	4555	18717	18636	18986	19025	792.01	770.25
6	3	4	832.8	804.8	778.0	4997	4923	19496	19481	18986	19025	792.01	791.40
7	2	1	774.1	762.5	759.1	4645	4687	18636	18717	18891	18834	792.01	760.80
8	3	1	808.5	804.8	759.1	4851	4883	19481	19496	18891	18834	792.01	781.95
9	4	1	806.6	778.0	759.1	4840	4709	19025	18986	18891	18834	792.01	768.55
10	3	2	803.1	804.8	762.5	4819	4846	19481	19496	18636	18717	792.01	783.65
11	4	2	759.1	778.0	762.5	4555	4651	19025	18986	18636	18717	792.01	770.25
12	4	3	820.5	778.0	804.8	4923	4997	19025	18986	19481	19496	792.01	791.40

Obs	BP	DMP	HMP	DBP	HBP	GI	SIJ	RIJ	MI
1	762.5	20.30	2.66824	18.6	2.43934	-6.0725	5.5517	3.5000	-1.1875
2	804.8	31.85	4.07315	9.0	1.11829	-6.0725	5.2183	2.6667	-1.1875
3	778.0	16.25	2.11437	6.8	0.87404	-6.0725	9.9267	-10.9167	-1.1875
4	804.8	23.95	3.05621	2.8	0.34791	-13.8225	7.2183	2.2500	1.6875
5	778.0	4.85	0.62967	-2.9	-0.37275	-13.8225	-10.9067	8.0000	1.6875
6	804.8	41.40	5.23124	28.0	3.47913	20.0108	14.7600	6.1667	0.3125
7	762.5	13.30	1.74816	11.6	1.52131	-13.8225	5.5517	-3.5000	1.6875
8	804.8	26.55	3.39536	3.7	0.45974	20.0108	5.2183	-2.6667	0.3125
9	778.0	38.05	4.95088	28.6	3.67609	-0.1142	9.9267	10.9167	-0.8125
10	804.8	19.45	2.48198	-1.7	-0.21123	20.0108	7.2183	-2.2500	0.3125
11	778.0	-11.15	-1.44758	-18.9	-2.42931	-0.1142	-10.9067	-8.0000	-0.8125
12	804.8	29.10	3.67703	15.7	1.95080	-0.1142	14.7600	-6.1667	-0.8125

The previous results were validated with the program developed by Zhang and Kang (1997). In ANOVA the code allows the partitioning of possible cross effects in ACG, ACE, ER and EM, when SORT, BY, GLM, IF-THEN, DROP, ARRAY, ELSE, GLM, CONTRAST, ESTIMATE and some MACROS were implemented. In Martínez (1991) these and other components are presented to elaborate the reference code.

The program of Zhang and Kang (1997) applies to the four methods of Griffing (1956a, b), for method 1 m variables are analyzed in two environments. In the present study, this was adjusted to a single environment by implementing the restriction I IF ENV> 1 THEN DELETE or IF ENV< 2 THEN DELETE, captured before DROP and after INPUT. Modifying it is more laborious for users with little training in programming and, especially when the analysis extends to series of experiments (Singh, 1973; Mastache and Martínez, 2003; Zhang *et al*., 2005).

TNAUSTAT software, in addition to calculating the genetic effects related to parents and their crosses in method 1 of Griffing (1956a, b) also, simultaneously, allows the calculation of hybrid vigor with the mean of both parents, with the best of them and additionally, based on commercial heterosis. This has the additional advantage of estimating the genetic parameters corresponding to the mating design I, proposed by Hayman (1954). However, this software was designed to work properly on a platform with MS Dos, so DOSBox software must be downloaded in advance.

The Zhang and Kang program (1997), modified by Saavedra (2019), is presented below:

```
OPTIONS PS=56 LS=78; TITLE 'METHOD 1'; DATA METHOD1;
INPUT I J REP HYBRID YIELD ENV; IF ENV>1 THEN DELETE;DROP N NI NJ P;
P=4;*NUMBER OF PARENTAL LINES? ; ARRAY GCA (N) G1 G2 G3;DO N=1 TO (P-1);
GCA= ((I=N)-(I=P)) + ((J=N)-(J=P)); END;ARRAY SCA(N) S11 S12 S13 S22 S23 S33;
N=0; DO NI=1 TO (P-1); DO NJ=NI TO (P-1); N+1; IF NI=NJ THEN DO;
SCA=(I=NI)*((J=NJ)-(J=P))+(I=P)*((J=P)-(J=NI));END;ELSE DO;
SCA=(I=NI)*(J=NJ)-(J=P)*((I=NI)+(I=NJ)-(I=P)*2)+(I=NJ)*(J=NI)
-(I=P)*((J=NI)+(J=NJ));END;END;END;
ARRAY REC (N) R12 R13 R14 R23 R24 R34; N=0; DO NI=1 TO (P-1);
DO NJ= (NI+1) TO P; N+1; REC= (I=NI)*(J=NJ)-(j=NI)*(I=NJ); END;END;
ARRAY MAT (N) M1 M2 M3; DO N=1 TO (P-1); MAT= (I=N) + (J=P)-(J=N)-(I=P);
```

```
END;ARRAY NONM (N) N12 N13 N23;N=0;DO NI=1 TO (P-2);DO NJ=(NI+1) TO (P-
1);N+1;NONM=((I=NI)*(J=NJ))-(I=NJ)*(J=NI)-((I=NI)*(J=P))+(I=NJ)*(J=P)
+ ((I=P)*((J=NI)-(J=NJ))); END;END;CARDS;
1 1 1 01 758 1
1 2 1 02 814 1
1 3 1 03 805 1
4 3 6 15 758 1
4 4 6 16 775 1;
PROC SORT; BY REP ENV I J; PROC GLM;CLASS REP ENV HYBRID;MODEL
YIELD=ENV REP(ENV) HYBRID HYBRID*ENV;TEST H=HYBRID
E=HYBRID*ENV;LSMEANS HYBRID;
RUN; TITLE 'DIALLEL-SAS 1'; PROC GLM; CLASS REP ENV HYBRID;
MODEL YIELD= ENV REP (ENV) G1 G2 G3 S11 S12 S13 S22 S23 S33 R12 R13 R14 R23
R24 R34 G1*ENV G2*ENV G3*ENV S11*ENV S12*ENV S13*ENV S22*ENV S23*ENV
S33*ENV R12*ENV R13*ENV R14*ENV R23*ENV R24*ENV R34*ENV;
%MACRO GCASCA; CONTRAST 'GCA' G1 1, G2 1, G3 1;
CONTRAST 'SCA' S11 1, S12 1, S13 1, S22 1, S23 1, S33 1;
ESTIMATE 'G1' G1 1; ESTIMATE 'G2' G2 1; ESTIMATE 'G3' G3 1;
Estimate 'G4' G1 -1 G2 -1 G3 -1;
ESTIMATE 'S11' S11 1; ESTIMATE 'S12' S12 1; ESTIMATE 'S13' S13 1;
ESTIMATE 'S22' S22 1; ESTIMATE 'S23' S23 1; ESTIMATE 'S33' S33 1;
Estimate 'S14' S11 -1 S12 -1 S13 -1;
Estimate 'S24' S12 -1 S22 -1 S23 -1;
Estimate 'S34' S13 -1 S23 -1 S33 -1;
Estimate 'S44' S11 1 S12 2 S13 2 S22 1 S23 2 S33 1;
%MEND GCASCA; %GCASCA %MACRO INTERACT;
CONTRAST 'GCA*ENV' G1*ENV 1 -1, G2*ENV 1 -1, G3*ENV 1 -1;
CONTRAST 'SCA*ENV' S11*ENV 1 -1, S12*ENV 1 -1, S13*ENV 1 -1, S22*ENV 1 -1,
S23*ENV 1 -1, S33*ENV 1 -1; %MEND INTERACT; %INTERACT
CONTRAST 'REC' R12 1, R13 1, R14 1, R23 1, R24 1, R34 1;
ESTIMATE 'R12' R12 1; ESTIMATE 'R13' R13 1; ESTIMATE 'R14' R14 1;
ESTIMATE 'R23' R23 1; Estimate 'R24' R24 1; ESTIMATE 'R34' R34 1;
CONTRAST 'REC*ENV' R12*ENV 1 -1,R13*ENV 1 -1,R14*ENV 1 -1,R23*ENV 1 -
1,R24*ENV 1 -1,R34*ENV 1 -1;
CONTRAST 'MAT SS' R12 1 R13 1 R14 1, R12 -1 R23 1 R24 1, R13 -1 R23 -1 R34 1, R14 -1
R24 -1 R34 -1; ESTIMATE 'MAT1' R12 1 R13 1 R14 1/DIVISOR=3;
ESTIMATE 'MAT2' R12 -1 R23 1 R24 1/DIVISOR=3;
ESTIMATE 'MAT3' R13 -1 R23 -1 R34 1/DIVISOR=3;
ESTIMATE 'MAT4' R14 -1 R24 -1 R34 -1/DIVISOR=3; RUN;
TITLE 'DIALLEL-SAS 2'; PROC GLM; CLASS REP ENV HYBRID;
MODEL YIELD= ENV REP (ENV) G1 G2 G3 S11 S12 S13 S22 S23 S33
M1 M2 M3 N12 N13 N23 G1*ENV G2*ENV G3*ENV
S11*ENV S12*ENV S13*ENV S22*ENV S23*ENV S33*ENV
M1*ENV M2*ENV M3*ENV N12*ENV N13*ENV N23*ENV;
%GCASCA %INTERACT
CONTRAST 'MAT SS' M1 1, M2 1, M3 1;
CONTRAST 'NONM SS' N12 1, N13 1, N23 1;
```

```
CONTRAST 'MAT*ENV' M1*ENV 1 -1, M2*ENV 1 -1, M3*ENV 1 -1;
CONTRAST 'NONM*ENV' N12*ENV 1 -1, N13*ENV 1 -1, N23*ENV 1- 1;
ESTIMATE 'M1' M1 1; ESTIMATE 'M2' M2 1; ESTIMATE 'M3' M3 1;
Estimate 'M4' M1 -1 M2 -1 M3 -1;
ESTIMATE 'N12' N12 1; ESTIMATE 'N13' N13 1; ESTIMATE 'N23' N23 1;
Estimate 'N14' N12 -1 N13 -1;
Estimate 'N24' N12 1 N23 -1;
Estimate 'N34' N13 1 N23 1; RUN;
```

To validate the code presented in program 1a, some results are shown that are generated by the program of Zhang and Kang (1997).

GLM procedure

Dependent variable: YIELD

Source	DF	Sum of Square	Square of The mean	F-Value	Pr > F
Model	20	51191.16667	2559.55833	9.69	<.0001
Error	75	19801.82292	264.02431		
Total correct	95	70992.98958			

R-square	Coef Var	Root MSE	YIELD mean
0.721074	2.051592	16.24882	792.0104

Source	DF	Type I SS	Square of The mean	F-Value	Pr > F
ENV	0	0.00000	.	.	.
REP(ENV)	5	3365.67708	673.13542	2.55	0.0347
HYBRID	15	47825.48958	3188.36597	12.08	<.0001
ENV*HYBRID	0	0.00000	.	.	.

Contrast	DF	Contrast SS	Square of The mean	F-Value	Pr > F
GCA	3	30162.39583	10054.13194	38.08	<.0001
SCA	6	14715.59375	2452.59896	9.29	<.0001
REC	6	2947.50000	491.25000	1.86	0.0988
(MAT SS)	(3)	(240.75000)	80.25000	0.30	0.8224
(NONM SS)	(3)	(2706.75000)	902.25000	3.42	0.0216

Parameter	Estimate	Error Standard	Value t	Pr > |t|
G1	-6.0729167	2.03110309	-2.99	0.0038
G2	-13.8229167	2.03110309	-6.81	<.0001
G3	20.0104167	2.03110309	9.85	<.0001
G4	-0.1145833	2.03110309	-0.06	0.9552
S12	5.5520833	3.70826994	1.50	0.1385
S13	5.2187500	3.70826994	1.41	0.1635
S23	7.2187500	3.70826994	1.95	0.0553
S14	9.9270833	3.70826994	2.68	0.0091
S24	-10.9062500	3.70826994	-2.94	0.0043
S34	14.7604167	3.70826994	3.98	0.0002
M1	-1.1875000	2.03110309	-0.58	0.5605
M2	1.6875000	2.03110309	0.83	0.4087

M3	0.3125000	2.03110309	0.15	0.8781
M4	-0.8125000	2.03110309	-0.40	0.6903
R12	3.5000000	4.69063167	0.75	0.4579
R13	2.6666667	4.69063167	0.57	0.5714
R14	-10.9166667	4.69063167	-2.33	0.0226
R23	2.2500000	4.69063167	0.48	0.6329
R24	8.0000000	4.69063167	1.71	0.0922
R34	6.1666667	4.69063167	1.31	0.1926

Conclusions

The programs '1a' and '1b' are easy to use and modify to carry out an analysis of variance in a single environment, with the subdivision of the effects of the treatments in parents (P), direct crosses (CD), reciprocal crosses (CR), P versus crosses and CD versus CR. It is also useful for comparing treatment means (Tukey, p= 0.01) and for estimating heterosis with the parents' mean and with the best of these when analyzing a variable.

Zhang and Kang's (1997) program was designed to analyze 'm' variables, but it is more difficult to manipulate when parents and environments are different from 5 and 2, respectively. Due to this restriction, it was necessary to modify the code with P= 4. The code proposed by Zhang and Kang estimates general and specific combinatorial aptitude, reciprocal and maternal effects, but does not include the Tukey test or the estimation of heterosis.

All three codes for SAS run on versions over 10 years from their commercial release and on the most recent academic proofs. The program of Zhang and Kang (1997) allowed the reliable validation of the code proposed in the present study when the genetic effects were estimated, but the three codes must be used to carry out a more complete diallelic analysis.

Cited literature

Burow, M. D. and Coors, J. G. 1994. DIALLEL: a microcomputer program for the simulation and analysis of diallel crosses. Agron. J. 86(1):154-158.

Christie, B. R. and Shattuck, V. I. 1992. The diallel cross: design, analysis and use for plant breeders. Plant Breeding Reviews. 9(1):9-36.

Gardner, C. O. and Eberhart, S. A. 1966. Analysis and interpretation of the variety cross diallel and related populations. Biometrics. 22(3): 439-452.

González, H. A.; Sahagún, C. J. y Pérez, L. D. J. 2007a. Estudio de ocho líneas de maíz en un experimento dialélico incompleto. Ciencias Agrícolas Informa. 16(1):3-9.

González, H. A.; Pérez, L. D.; Sahagún, C. J.; Norman, M. T. H.; Balbuena, M. A. and Gutiérrez, R. F. 2007b. Análisis de una cruza dialélica completa de líneas endogámicas de maíz. Ciencias Agrícolas Informa. 16(1):10-17.

Griffing, B. 1956a. A generalized treatment of the use of diallel crosses in quantitative inheritance. Heredity. 10(1):31-50.

Griffing, B. 1956b. Concept of general and specific combining ability in relation to diallel crossing systems. Austr. J. Biol. Sci. 9(4):463-491.

Hallauer, A. R.; Miranda, F. O. J. B.1988. Quantitative Genetics in Maize Breeding. Iowa State University Press, Ames. Second Edition. USA. 468 p.

Hayman, B. I. 1954. The theory and analysis of the diallel crosses. Genetics. 39(1):798-809.

Magari, R. and Kang, M. S. 1994. Interactive BASIC program for Griffing's diallel analysis. Journal of Heredity. 85(4):336.

Martínez, G. A. 1983. Diseño y análisis de los experimentos de cruzas dialélicas. Centro de Estadística y Cálculo. Colegio de Postgraduados. Chapingo, Estado de México. 252 p.

Martínez, G. A. 1991. Análisis de los experimentos dialélicos a través del procedimiento IML del SAS. Comunicaciones en Estadística y Cómputo. 10(2):1-36.

Mastache, L. A. A.; Martínez, G. A.; Castillo, M. A. y González, C. F. V. 1998a. Los mejores predictores lineales e insesgados (MPLI) en experimentos dialélicos parciales sin efectos maternos. Revista Fitotecnia Mexicana. 21(1):49-60.

Mastache, L. A. A.; Martínez, G. A. y Castillo, M. A. 1998b. Los mejores predictores lineales e insesgados (MPLI) en experimentos dialélicos parciales con efectos maternos. Revista Fitotecnia Mexicana. 21(2):171-184.

Mastache, L. A. A.; Martínez, G. A. y Castillo, M. A. 1999a. Los mejores predictores lineales e insesgados (MPLI) en los diseños dos y cuatro de Griffing. Agrociencia. 33(1):81-90.

Mastache, L. A. A.; Martínez, G. A. y Castillo, M. A. 1999b. Los mejores predictores lineales e insesgados (MPLI) en los diseños uno y tres de Griffing. Agrociencia. 33(3):349-359.

Mastache, L. A. A. y Martínez, G. A. 2003. Un algoritmo para el análisis, estimación y predicción en experimentos dialélicos balanceados. Revista Fitotecnia Mexicana. 26(3):191-200.

Montesinos, L. O. A.; Mastache, L. A. A.; Luna, E. I. y Hidalgo, C. J. V. 2007. Mejor predictor lineal e insesgado combinado para aptitud combinatoria general y análisis combinado de los diseños uno y tres de Griffing. Técnica Pecuaria en México. 45(2):131-146.

Padilla, L. A.; González, H. A.; Pérez, L. D. J.; Rubí, A. M.; Gutiérrez, R. F. y Franco, M. J. R. P. 2019a. InfoStat, InfoGen y SAS para contrastes mutuamente ortogonales en experimentos en bloques completos al azar en parcelas subdivididas. Rev. Mex. Cienc. Agríc. 10(6):1417-1431.

Padilla, L. A.; González, H. A.; Pérez, L. D. J.; Rubí, A. M.; Gutiérrez, R. F.; Ramírez, D. J. F.; Franco, M. J. R. P. y Serrato, C. R. 2019b. Programas para SAS e InfoStat para analizar una serie de experimentos en parcelas subdivididas. En: temas selectos en la innovación de las ciencias agropecuarias. Alfaomega Grupo Editor SA. de CV. Primera edición (Salgado y otros, eds.). México, DF. 724 p. ISBN: 9786075384115.

SAS Institute, Inc. 1989. SAS/IML software: Usage and reference. Version 6. First edition. Cary, N. C.

Saavedra, G. C. 2019. Estimación de parámetros genéticos en maíz con dos metodologías usando datos de una cruza dialélica completa. I. Un ambiente. Tesis de Maestro en Fitomejoramiento. Facultad de Ciencias Agrícolas, Universidad Autónoma del Estado de México. Toluca, Estado de México. 96 p.

Singh, D. 1973. Diallel analysis for combining ability over several environments-II. Indian Journal of Genetics and Plant Breeding. 33(1):469-481.

Schaffer, H. E. y Usanis, R. A. 1989. General least squares analysis of diallel experiments: A computer program. Genetics Dep. Res. Rep. 1. North Carolina State University, Raleigh. 61 p.

Sprague, G. F. and Tatum, L. A. 1942. General *vs* specific combining ability in single crosses of corn. J. Amer. Soc. Agron. 34(1):923-932.

Zhang, Y. and Kang, M. S. 1997. DIALLEL-SAS: A SAS program for Griffing's Diallel Analyses. Agronomy Journal. 89(2):176-182.

Zhang, Y.; Kang, M. S. and Lamkey, K. R. 2005. DIALLEL-SAS05: A comprehensive program for Griffing's and Gardner-Eberhart analysis. Agron. J. 97(4):1097-1106.