d by ICRISAT Open Access Report

# **J. Hill · W.W. Wagoire · R. Ortiz · O. Stølen** Analysis of a combined  $F_1/F_2$  diallel cross in wheat

Received: 25 May 2000 / Accepted: 31 July 2000

**Abstract** Analysis of a conventional diallel cross offers only perfect-fit estimates of the genetic components of variation, but no test for the goodness of fit of the model based on these estimates. When  $F_2$  progenies are available, however, combining  $F_1$  and  $F_2$  diallels in a single experiment overcomes these problems. Least-squares estimates of these components can be calculated, errors attached to them and the goodness of fit of the resultant model tested. This analysis was applied to data on the severity of yellow rust infection in an  $F_1/F_2$  half-diallel cross among eight bread wheat lines adapted to the East African highlands. After removing two interacting arrays, genetic analysis indicated that an additive/dominance model of gene action satisfactorily explained the variation observed among the remaining six parents and their progenies, in both the individual  $F_1$  and  $F_2$  diallels and the combined  $F_1/F_2$  diallel. Resistance to yellow rust was dominant to susceptibility and genes for increased resistance were more frequent.

**Keywords** Grain yield · Quantitative genetics · Yellow rust

## Introduction

It has been suggested that of all the available experimental mating designs none 'has been used and abused more extensively' than the diallel cross (Hallauer and Miranda

Communicated by P.M.A. Tigerstedt

J. Hill · O. Stølen Department of Agricultural Sciences, The Royal Veterinary and Agricultural University, 40 Thorvaldsensvej, DK-1871 Frederiksberg C, Copenhagen, Denmark

W.W. Wagoire

Namulonge Research Institute, P.O. Box 7084, Kampala, Uganda R. Ortiz  $(\mathbb{Z})$ 

ICRISAT, Patancheru 502 324, Andhra Pradesh, India e-mail: r.Ortiz@cgiar.org

1988). There are four variations on the diallel theme, depending on whether parents and reciprocals are included (Griffing 1956). For the estimation of general and specific combining-ability effects, none of the mating designs supply as much information as the diallel analyses proposed by Griffing (1956) or Gardner and Eberhart (1966). When a complete genetic analysis is required, none of the alternative designs provides as much information as the analyses described by Jinks and Hayman (1953) and Jinks (1954). But this information comes at a price. Not only is the diallel cross labour-intensive, its representation of a reference population is poor. The genetic information supplied, and the conclusions drawn therefrom, relate only to the parents used in that particular cross. Mather and Jinks (1982) stated that many of the benefits of the use of a diallel for a fixed set of inbred lines do not accrue when it is applied to diallel sets of crosses among random samples of individuals from a randomly mating population. A fuller discussion of the issues surrounding the use of a diallel cross can be found in Baker (1978), Mather and Jinks (1982), Wright (1985), Christie and Shattuck (1992) and Hill et al. (1998).

To maximize the genetic information from a diallel cross it should include a set of *n* inbred lines and all possible crosses between them. The ensuing analyses of the data provide tests for non-allelic interactions, and information on the order, level and direction of dominance. If an additive/dominance model of gene action fits the data adequately, non-allelic interactions are absent, and the genetic components of variation may be estimated. But, because there are as many components to be estimated as there are statistics available for their estimation, only perfect-fit estimates of the components can be calculated, and worthwhile estimates of their errors are not available (Mather and Jinks 1982). Mather and Jinks (1982) suggested one way of overcoming this problem by growing the parents,  $F_1s$  and  $F_2s$  of a diallel set of crosses in the same experiment to generate additional statistics. Least-squares estimates of the components can be calculated and the goodness of fit of the resultant

model tested. Our research was conducted to analyse data on the yellow rust (*Puccinia striiformis* Westend) severity obtained from the parents and  $F_1$  and  $F_2$  diallels among bread wheat (*Triticum aestivum* L.) lines adapted to the East African highlands.

## Materials and methods

Eight bread wheat lines from the Uganda Wheat Development Project were chosen for this investigation. These lines were selected at Kalengyere (1° 15' S 29° 45' E), a marginal wheat-growing environment in the south western highlands of Uganda regularly exposed to severe biotic stress caused by yellow rust. The pedigree of the eight lines and their response to yellow rust are listed in Table 1. The eight lines were crossed in a half-diallel mating design to give 28  $F_1$  hybrids, which were subsequently selfed to produce the corresponding  $F_2$  generation.

The trial was conducted at Kalengyere using a randomised complete block design with two replicates. Kalengyere is at 2400 m above sea level, has an Andosol with pH 5.7, and an average temperature of 16°C throughout the year. The high rainfall (750 mm) season lasts from September to March and the relatively low rainfall (480 mm) season from March to August. Parental and  $F_1$  plots included two rows 1.5-m long, and  $F_2$  plots included two rows 5-m long. Inter-row spacing was 0.3 m throughout. Nitrogen was applied at sowing at a rate of 50 kg ha–1.

Yellow rust severity and plot grain yield data were recorded but subsequent analysis of the yield data confirmed the results of previous experiments (Wagoire et al. 1999) for the presence of complex non-allelic interactions. Emphasis, therefore, will focus on the yellow rust data. Yellow rust infection was scored on the flag leaf of individual plants when its severity on the most-susceptible parent was about 100%; i.e. most of the leaf surface was covered with uredinia. The modified Cobb scale (Peterson et al. 1948) was used for scoring the percentage of tissue rusted (disease severity). Host response to infection was scored using  $T (= 0.1)$  for immune plants, R  $(= 0.2)$  for resistant plants showing miniature uredinia, MR  $(= 0.4)$ for moderately resistant plants exhibiting small uredinia,  $MS = 0.8$ ) for moderately susceptible plants with moderate sized uredinia (smaller than the fully susceptible type), and  $S = 1$ ) for fully susceptible plants. Disease severity and host response scores were multiplied together to give the coefficient of infection (CI) for data analysis.

#### Statistical and genetical analyses

Expectations of the  $F_2$  generation and array means, and their  $F_1$ counterparts, are shown in Table 2 for a single gene difference, where  $A^+$  represents the increasing allele, and  $A^-$  the decreasing allele at the A locus. Because of the decrease in heterozygosity that arises from selfing the  $F_1$  to give the  $F_2$ , the dominance contribution to the off-diagonal elements and array means (emboldened) has been halved. Expectations of the various statistics that can be derived from the generation and array means are given in Table 3. For the F<sub>2</sub> statistics, the coefficients of  $d^2$  ( $Vd_{D1}$  and  $Vd_{D2}$ ) are one-quarter, and the coefficients of  $d (Vd_{AD})$  one-half of their corresponding  $F_1$  values (Mather and Jinks 1982). Eight statistics are therefore available for estimating the five components  $Vd_A$ ,  $Vd_{D1}$ ,  $Vd_{D2}$ ,  $Vd_{AD}$  and  $V_{EC}$  (Hill et al. 1998), which correspond to 1/2*D*,  $1/2H_1$ ,  $1/2H_2$ ,  $1/2F$  and *E* in Mather and Jinks' (1982) notation. In their model the genetic components are *D*, which measures only additive effects, while  $1/2H_1$  and  $1/2H_2$  measure only dominance effects. An  $H_1 < H_2$  indicates unequal p and q allele frequencies at the relevant loci. A positive value of *F* suggests that there are more dominant alleles present in the inbred lines than recessive alleles, irrespective of whether these are increasing or decreasing in their effect on the characteristic under investigation. *E* is the environmental parameter.

After equating the expectation of each statistic with its observed value, the resulting eight basic equations are combined in the manner first described by Mather (1949) to give five normal equations, which can be solved by matrix inversion to yield unweighted least squares estimates of the five components (see Mather and Jinks 1982 for details). Expected values for each statistic can then be calculated and the deviation of these values from those actually observed determined. Analysis of the deviation sum of squares enables the goodness of fit of the model to be tested and estimates of the standard errors of the components to be calculated. Other information supplied by the estimates of the components of variation includes the dominance ratio ( $\sqrt{Vd_{\text{D1}}/Vd_{\text{A}}},$  nar-

**Table 1** Designation, parentage, pedigree and yellow rust response of eight bread wheat parents included in diallel crosses



HCWSN Hot Climate Wheat Screening Nursery, CIMMYT, Thailand,

SNACWYT Screening Nursery for African Cooperative Wheat Yield Trial, CIMMYT, East Africa

**Table 2** Contribution of a single gene difference to the family and array means of an  $F_1$  and  $\dot{F}_2$  diallel cross. The lower, emboldened, value in each cell is for an  $F_2$  diallel





**Table 3** Expectations of the statistics for a combined  $F_1/F_2$ diallel; (i) is the coefficient of the environmental component for a complete, and (ii) a halfdiallel, design

row-sense heritability  $[h^2_{\text{n}} = (Vd_A + Vd_{\text{D1}} - Vd_{\text{D2}} - Vd_{\text{AD}})/(Vd_A + Vd_{\text{D1}} 1/2Vd_{D2} - Vd_{AD} + V_{EC}$ ), broad-sense heritability  $[(h_{b}^{2}) = Vd_{A} + Vd_{D1} 1/2Vd_{D2} - Vd_{AD}/(Vd_A + Vd_{D1} - 1/2Vd_{D2} - Vd_{AD} + V_{EC})$  and the mean value of *pq* across all loci ( $1/4Vd_{D2}/Vd_{D1}$ ).

The data were first analysed on a plot mean basis by Griffing's (1956) method 2, model I, after which the techniques described in Mather and Jinks (1982) were used for a genetic analysis.

# Results and discussion

Genetic analysis of the unreduced  $8 \times 8$  F<sub>1</sub> and F<sub>2</sub> half-diallels indicates that an additive/dominance model of gene action was satisfactory for the  $F_2$  data, but was not adequate for the  $F_1$  diallel. Clearly, comparable  $F_1$  and  $F_2$  diallel data sets are a prerequisite before we can proceed with the combined analysis. The standard procedure for removing non-allelic interactions from a diallel cross is to identify and then omit the interacting array(s), followed by reanalysis of the resultant reduced diallel to establish that an additive/dominance model now fits the data (Jinks 1954; Hill et al. 1998). This objective was accomplished by excluding arrays 5 and 7 from both diallels. Further analyses presented here relate to these reduced  $6 \times 6$  half-diallels.

Plot means and their corresponding array means are given in Table 4. From the analysis of these data it is clear that additive (a) and non-additive (b) genetic effects are highly significant for both diallel sets (Table 5). Directional domi-

**Table 4** Coefficient of yellow rust infection plot means, summed over replicates, for the reduced  $6 \times 6$  F<sub>1</sub> and F<sub>2</sub> diallels. F<sub>2</sub> data are the lower, emboldened, values in each cell

Parent	1	2	3	4	6	8	Array mean
1	0.50	1.00	4.30	2.20	6.00	24.00	6.33
2		0.40 0.40	4.75 0.20 0.10	5.53 3.80 4.13	9.81 2.80 16.24	34.87 6.00 20.39	9.31 2.37 6.94
3			0.00	0.40 5.74	2.40 16.62	1.00 13.30	1.38 6.75
4				10.00	18.82 14.01	46.00 74.77	13.54 19.03
6					90.00	100.00 64.84	36.67 35.25
8						140.00	52.83 58.03

nance  $(b_1)$  is also evident, with resistance to yellow rust being dominant to susceptibility in both generations (Wagoire et al. 1998). As expected, both directional dominance and specific combining-ability effects  $(b_3)$  diminish upon selfing. Combining these two analyses reveals that the genetic effects are consistent across generations.

The genetic analyses tests for non-allelic interactions confirm their absence in both diallels. Thus, the joint regression coefficients, calculated from the regression of

**Table 5** Analysis of variance for yellow rust infection of the reduced  $F_1$  and  $F_2$  diallels for six bread wheat lines

Item	df	$MS(F_1)$	$MS(F_2)$	df	$MS(F_1/F_2)$
a	5	2176.13***2081.25***	54 249.35***		
b	15	$304.75***211.06***$	15 471.96***		
$b_1$		1399.45***955.53***	12 333.87***		
	5	$212.23***257.16***$	5461.36***		
$b_2$ <sub>b<sub>3</sub></sub>	9	234.52***102.74***	9270.98***		
Generations (G)					
$G \times a$				5	8.03
$G \times b$				15	43.85
$G \times b_1$					21.11
$G \times b_2$				5	8.03
$G \times b_3$				9	66.28
<b>Block</b> interactions					
(error)	20	40.84	32.69	40	36.77

\*\*\* indicates significant at *P* <0.001 respectively

**Table 6** Tests for non-allelic interactions for yellow rust infection in the reduced  $F_1$  and  $F_2$  diallels of six bread wheat lines

Item	F,	F,
Joint regression coefficient Heterogeneity item Arrays item in $(Wr_i-Vr_i)$ analysis Arrays item in $(Wr_i-Vr_i)$ analysis	$1.017 \pm 0.106$ $0.948 \pm 0.102$ NS. *** NS.	NS * NS.

NS, \* and \*\*\* indicate non-significant or significant at *P* = 0.05– 0.01, and *P* <0.001 respectively



**Fig. 1**  $Wr_i + Vr_i$  for the  $F_1$  and  $F_2$  diallels

**Table 7** Estimates of the available statistics for the combined  $F_1/F_2$  diallel among six bread wheat lines for yellow rust infection

Statistic	Block 1	Block 2	Average	
$V_{\rm Pa}$	1115.618	781.131	948.375	
$r_{F1}$	231.067	269.358	250.213	
$r_{F1}$	328.369	311.510	319.940	
$V r_{\rm F1}$	99.445	128.253	113.849	
$r_{F2}$	243.085	145.958	194.522	
$r_{F2}$	381.465	241.487	311.476	
$\tilde{V}_{F2}$	139.575	77.093	108.334	
$V^{\,}_{\rm EC}$	36.766	36.766	36.766	

array covariances (*Wr*<sub>i</sub>) on array variances (*Vr*<sub>i</sub>), differed significantly from zero but not from their expected value of one, while the regression slopes are homogeneous across replicates (Table 6). The significance of the arrays item in the  $(Wr_i + Vr_i)$  analyses, and its non-significance in the ( $Wr_i$ - $Vr_i$ ) analyses, indicate that the non-additive variation is due to the dominance effects of those genes controlling this character. Two features emerge, both a direct consequence of the reduced heterozygosity of the  $F_2$ , in the  $Wr_i/Vr_i$  graphs for both diallels (Fig. 1). The fitted regression line for the  $F<sub>2</sub>$  diallel lies approximately half-way between the  $F_1$  line and the point of no dominance  $(Vr_i = 1/4V_{Pa}$ ,  $Wr_i = 1/2V_{Pa}$ , Mather and Jinks 1982), and the range of array variances and covariances displayed in the  $F_2$  diallel is less than in the  $F_1$  diallel. On both graphs arrays having a resistant common parent lie close to the origin, while arrays based on increasingly susceptible common parents occupy positions progressively removed from the origin. The order of array points along their fitted regression line is the same for both diallels. In both there is a strong positive correlation between the mean of the common parent  $(P_i)$  and the corresponding ( $Wr_i+Vr_i$ ) value ( $r_{F1} = 0.99$ <sup>\*\*\*</sup>;  $r_{F2} = 0.94$ <sup>\*\*\*</sup>), which confirms that among these six wheat lines resistance to yellow rust is dominant to susceptibility.

Having established that nonallelic interactions do not contribute significantly to the genetic control of CI in the two reduced diallels, the estimation of the genetic components of variation from the combined  $F_1/F_2$  diallel were calculated for each replicate and averaged across replicates (Table 7). The estimate of  $V_{EC}$  was calculated

Component Block 1 Block 2 Overall  $(\pm SE)$  $\begin{array}{cccc} Vd_{\rm A} & 528.244 & 368.538 & 448.390 \pm 25.250^{***} \\ Vd_{\rm D1} & 306.689 & 351.295 & 328.988 \pm 133.185^{**} \end{array}$ *Vd*<sub>D1</sub> 306.689 351.295 328.988 ± 133.185\*\*\*<br> *Vd*<sub>D2</sub> 202.010 258.536 230.270 ± 126.143 NS  $Vd_{\text{D2}}$  202.010 258.536 230.270 ± 126.143 NS<br>  $Vd_{\text{AD}}$  464.820 242.281 353.549 ±84.906\*\*  $V_{\text{4AD}}$  464.820 242.281 353.549 ±84.906\*\*<br>  $V_{\text{FC}}$  48.176 20.135 34.157 ± 4.077\*\*\* *VEC 48.176* 20.135 34.157 ± 4.077\*\*\*<br>0.762 0.976 0.857 Dominance ratio  $0.762$   $0.976$   $0.857$ <br>pq  $0.165$   $0.184$   $0.175$ 0.184  $p$ q<sub>n</sub><sup>2</sup><sub>n</sub> n 0.530 0.594 0.565  $h^2$ <sub>b</sub><sup>2</sup>  $\frac{1}{2}$  0.848 0.945 0.900

calculated from  $F_1$  and  $F_2$  diallels of six bread wheat lines for yellow rust infection

**Table 8** Least squares estimates of the components of variation

NS, \*\* and \*\*\* indicate non-significant or significant at *P*=0.01–0.001 and *P*<0.001, respectively

**Table 9** Testing the adequacy of the model based on the overall components of variation

Item	df	МS
Overall deviation Block heterogeneity	3	5039.957 NS
(remainder)	3	1173.753
Total	h	3106.855

NS indicates non-significant at *P* = 0.05

by pooling the homogeneous error mean squares for the  $F_1$  and  $F_2$  diallels under the combined  $F_1/F_2$  analysis (Table 5). Unweighted least squares estimates of the five components were determined, but of the five components, four –  $Vd_A$ ,  $Vd_{D1}$ ,  $Vd_{D2}$  and  $V_{EC}$  – are quadratic and cannot meaningfully be negative (Table 8). Negative estimates are of no genetic consequence. It seems appropriate, therefore, to apply a one-tailed '*t*' test of significance to these four components, i.e. halving the final probability, since only values significantly greater than zero are of interest. For the remaining component,  $Vd_{AD}$ , a conventional two-tailed test should be used because it can take a sign. Applying these tests to the overall estimates of the components confirms the presence of significant additive and dominance variation. Since  $Vd_{AD}$  is significantly greater than zero, and since dominance is acting towards a decreased CI value (i.e. increased resistance), it can be concluded that, in general, decreasing alleles (*q*) are more frequent than increasing alleles (*p*) at those loci controlling CI in these lines. Moreover, as is 0.175, this would suggest that  $p \approx 0.225$  and  $q \approx 0.775$ . Although dominance is incomplete, the difference between the estimates of narrow- and broad-sense heritability attests to the contribution which dominance makes to the genetic control of this character. Not surprisingly, significant non-heritable variation exists between plots. Because the estimate of  $V_{EC}$  is considerably less than those of the genetic components, an empirical error (Mather and Jinks 1982) was calculated for this component from the variation between the  $F_1$  and  $F_2$  estimates given in Table 5.

The goodness of fit of the model may now be tested. Within each replicate there are eight deviations between observed and expected, one for each statistic. But the resultant deviation sum of squares has only three degrees of freedom in each replicate, as five have been utilized to estimate the components needed to calculate the expected values of the statistics. Of the six degrees of freedom attached to the total deviation sum of squares (three from each of the two replicates), three test the adequacy of the overall estimates of the components in accounting for the variation among the average values of the statistics. Because these values are averaged across two replicates, the resultant deviation sum of squares must be multiplied by two. The remaining three degrees of freedom measure the heterogeneity between replicate estimates of the components. From the analysis it may be concluded that the fit of the overall model is satisfactory since the overall deviation mean square is not significantly greater than the remainder or the heterogeneity mean square (Table 9). A model based on the overall estimates of  $Vd_A$ ,  $Vd_{D1}$ ,  $Vd_{D2}$ ,  $Vd_{AD}$  and  $V_{EC}$  adequately explains the variation recorded here for CI in the combined  $F_1/F_2$  diallel. The overall and the remainder mean squares may therefore be pooled to give a mean square of 3106.855 with six degrees of freedom, from which estimates of the standard errors of the components can be calculated (Mather and Jinks 1982). Thus, in the two individual diallels and in the combined data an additive/dominance model of gene action accounts for the variation in CI observed in these bread wheat lines.

The diallel cross is costly to implement, one problem often associated with this design being the difficulty of obtaining sufficient  $F_1$  seed. This difficulty could be overcome by resorting to an  $F_2$  diallel, but such a course of action is not without its own problems. These stem from the halving of heterozygotes in the  $F<sub>2</sub>$  generation, which in turn reduces the dominance contributions of the genes concerned. This could be compensated for by raising larger  $F_2$  progenies (Mather and Jinks 1982). Tests for non-allelic interactions remain essentially the same as in an  $F_1$  diallel, as does the ordering of array points along the fitted  $Wr_i/Vr_i$  regression line. But the spread of points along this line will be reduced (see Fig. 1), while the point at which this line cuts the  $Wr_i$  axis now underestimates the average level of dominance (Dickinson and Jinks 1956; Hill et al. 1998). Combining  $F_1$  and  $F_2$  diallels in the same experiment is unusual, but it does enable more rigorous tests to be applied to the data, which should put our understanding of the genetic control of continuously varying characters on a firmer footing.

From these results it is apparent that any breeding programme aimed at improving the resistance of this bread wheat germplasm to yellow rust in Uganda must centre on line 3, which has consistently been shown to be the most-resistant line at Kalengyere (Wagoire et al. 1998). Moreover, the position this array point occupies on both the  $F_1$  and  $F_2$  *W*r<sub>i</sub>/*V*r<sub>i</sub> graphs indicates that line 3 carries more dominant genes for yellow rust resistance than any of the other lines included in this trial, and that it transmits this resistance to its progeny. It has been established that a strong negative genetic correlation exists between CI value and yield in this material (Hill et al. 1999). Hence, breeding for increased yellow rust resistance should also increase yield at a yellow rust-prone site such as Kalengyere.

**Acknowledgments** We are grateful to Professor P.D.S. Caligari for reading and commenting upon this manuscript. We also thank the Uganda Grain Milling Company for logistic support and the African Development Bank for financial assistance.

### References

- Baker RJ (1978) Issues in diallel analysis. Crop Sci 18:533–536 Christie BR, Shattuk VJ (1992) The diallel cross: design, analysis
- and use for plant breeders. Plant Breed Rev 9:9–36 Dickinson AG, Jinks JL (1956) A generalised analysis of diallel
- crosses. Genetics 41:65–77 Gardner CO, Eberhart SA (1966)Analysis and interpretation of the variety cross diallel and related populations. Biometrics 22:439–452
- Griffing B (1956) Concept of general and specific combining ability in relation to diallel crossing systems. Aust J Biol Sci 9:463–493
- Hallauer AR, Miranda JB (1988) Quantitative genetics in maize breeding, 2nd edn. Iowa State University Press, Ames, Iowa
- Hill J, Becker HC, Tigerstedt PMA (1998) Quantitative and ecological aspects of plant breeding. Chapman and Hall, London
- Hill J, Ortiz R, Wagoire WW, Stølen O (1999) Effectiveness of indirect selection for wheat yield in a stress environment. Theor Appl Genet 98:305–309
- Jinks JL (1954) The analysis of continuous variation in a diallel cross of *Nicotiana rustica* varieties. Genetics 39:767–788
- Jinks JL, Hayman BI (1953) The analysis of diallel crosses. Maize Coop Newslett 27:48–54
- Mather K (1949) Biometrical genetics. Methuen, London
- Mather K, Jinks JL (1982) Biometrical genetics, 3rd edn. Chapman and Hall, London
- Peterson RF, Campbell AB, Hannah AE (1948) A diagrammatic scale for estimating rust intensity of leaves and stem of cereals. Can J Res Sect C 26:496–500
- Wagoire WW, Stølen O, Hill J, Ortiz R (1998) Inheritance of adult field resistance to yellow rust disease among broad-based hexaploid bread wheat germplasm. Theor Appl Genet 97: 502–506
- Wagoire WW, Hill J, Stølen O, Ortiz R (1999) Impact of genotype-environment interactions on the inheritance of wheat yield in low-yielding environments. Euphytica 105:17–23
- Wright AJ (1985) Diallel designs, analyses and reference populations. Heredity 54:307—311