OnLine Journal of Biological Sciences 1 (8): 704-707, 2001

Diallel Analysis for Inbred Lines Involving Genotype × Environment Interaction Effects on Additive-Dominance Genetic Model

[1]Abderrahmane Achouch and [2]Zhu Jun
[1]Faculty of Agronomy, Blida University, Algeria
[2]Department of Agronomy, Zhejiang University, Hangzhou 310029, China

Abstract: Additive-dominance genetic model was employed for eight inbred lines by mating design of modified diallel crosses with F_1's and reciprocal F_1's, F_2's and reciprocal F_2's. Monte Carlo simulations were conducted to compare specific effects and their interactions in single environment and in multiple environments when GE exits, when Genotype Environment (GE) is ignored and when it is null. Estimation of variance components and genetic prediction were computed by using respectively minimum norm quadratic unbiased estimation (MINQUE (1) and adjusted unbiased estimation (AUP). The genetic models in single and multiple environments were robust and gave unbiased estimates with high efficiencies. The variance residual $\sigma^2\varepsilon$ was significantly smaller and more efficient when GE interaction effects were considered in the full model. Whereas σ^2_A, σ^2_D and $\sigma^2\varepsilon$ were overestimated when σ^2_{AE} and σ_{2DE} were ignored in the full model. If there are no genotype by environment interaction (GE) effects ($\sigma^2AE = \sigma^2DE = 0$), σ^2_A and were estimated with similar bias, MSE and power value as when GE interaction effects are present. AUP method gave extremely low biases for mean of predicted genetic effects with absolute value between $8.10^{-5} \sim 12.10^{-4}$. When GE is not ignored, variance of prediction is significantly similar to the true value.

Key words: Additive-dominance genetic model, diallel crosses, Inbred lines, genotype × environment interaction

Introduction

Many species of plants and animals continue to be degraded or disappeared for many reasons essentially involved by genetic effects related to the environment. In plant and animal breeding, few research projects were done in this field in spite of the importance of the subject. Most certainly, it is known that phenotypic variation of quantitative traits is attributed to segregation of some loci and for another part due to the minor effects of the undetermined number genes (Falconer, 1996). However phenotypic variation of quantitative traits cannot only due to the genetic effect side. Environment effects and genotype by environment interaction effects can be other important causes of this variation. The statistical analysis of both effects was practically impossible for some genetic models with mixed effects. Cockerham, 1980, resolved this problem and proposed methodology for constructing general genetic models and set up fundamental principles in developing many kinds of complicated genetic models. Nowadays, the development of statistical methods applied to biometrics (Hartley and Rao, 1967; Rao, 1971 a, b; Henderson, 1988; Weir, 1996) merged progressively to the computational methods in data analysis (William *et al.*, 1992) handle AD genetic model with more complicated effects. Applications of AD models were also reinforced by using experimental designs and some mating designs appropriately conceivable for genetic issues. Nested design, factorial design and diallel designs (Griffing, 1956, Gardner and Eberhart, 1966) are the most employed mating designs. Many kinds of genetic models were used to study the differences and relationships of phenotypic traits (Zhu, 1997). Cockerham and Weir (1977) developed a bio-model including additive, dominance, maternal and paternal effects for F_1's from diallel mating. Zhu and Weir (1996) proposed an animal model, which is a modification of Eisen's model. Simple additive-dominance (AD) genetic models, without environment effects, were used from 1950's for the study of quantitative traits economically important such as yield, seed quality and resistance to disease, rate growth of the animals, etc.

For the analysis of the phenotypic variation of quantitative traits, the purpose of this research is centered particularly on the influence of the environment and on the interaction between genetic and environment effects. An understanding of the inheritance of these differences is of fundamental significance in the study of evolution and in the application of genetics to animals and plant breeding.

Materials and Methods

AD genetic model was employed for genetic entries I x j for (I = 1,2...,8; j = 1,2,...,8). In single environment the genetic model of diallel analysis for parent (I = j) F_1 and F_2 with (I ≠ j) in the *l*th block of the *k*th mating type of genetic entry can be translated as:

$$y_{ijkl} = \mu + G_{ijk} + B_1 + \varepsilon_{ijkl}$$

The expression of the phenotypic values for P's, F_1's, F_2's generations can be written as,

$(P_i \times P_i)$: $y_{ijkol} = \mu + 2A_i + D_{ij} + B_l + \varepsilon_{ijol}$
$(F_{1ij}) = (P_i \times P_j)$: $y_{ij1l} = \mu + A_i + A_j + D_{ij} + B_l + \varepsilon_{ij1l}$
$(F_{2ij}) = (F_{1ij} \times F_{1ij})$: $y_{ij2l} = \mu + A_i + A_j + \frac{1}{4}D_{ii} + \frac{1}{4}D_{jj} + \frac{1}{2}D_{ij} + B_l + \varepsilon_{ij2l}$

For the different mating types, the general genetic AD models in single environment can be written as mixed linear model.

$$y = 1\mu + U_A e_A + U_D e_D + U_B e_B + e_\varepsilon$$

$$=$$

where $1\mu + U_A e_A + \sum_{i=1}^{3} e_{ij} + U_B e_B$ and e_ε are the random effects corresponding to the additive, dominance, block and residual effects, U_A, U_D, U_B, U_ε are the known coefficient matrix relating to the random vector e_i.

In multiple environment, the general genetic model for parent (I = j) and F or F_2 with (I ≠ j) in the *l*th block within the *h*th environment is:

$$y_{hijkl} = \mu + E_h + G_{ijk} + GE_{hijk} + B_{l(h)} + e_{hijkl}$$

The phenotypic values of the three following generations can be written as,

$(P_i \times P_i)$: $y_{hiiol} = \mu + E_h + 2A_i + D_{ii} + 2AE_{hi} + DE_{hii} + B\mu_{(h)} + \varepsilon_{hiiol}$
$(F_{1ij}) = (P_i \times P_j)$: $y_{hij1} + \mu + E_h + A_i + A_j + D_{ij} + AE_{hi} + AE_{hj} + DE_{hij} + B_{l(h)} + \varepsilon_{hij1l}$

$(F_{2ij}) = (F_{1ij} \times F_{1ij})$: $y_{hij2l} = \mu + E_h + A_i + \frac{1}{4}D_{ii} + \frac{1}{4}D_{jj} + \frac{1}{2}D_{ij} + AE_{hi} + AE_{hj} + \frac{1}{4}DE_{hii} + \frac{1}{4}DE_{hjj} + \frac{1}{2}DE_{hij} + B_{l(h)} + \varepsilon_{hij2l}$

In multiple environments, these general genetic AD models can be expressed as matrix form of mixed linear model,

$$y = 1\mu + U_A e_A + U_D e_D + U_{AE} e_{AE} + U_{DE} e_{DE} + U_B e_B + e_c$$

$$= \mu + \sum_{i=1}^{n} U_i e_i$$

where U_i is the known coefficient matrix relating to the random vector e_i assumed that the random effects are not correlated in (1) and (2), so the random vector y has multi variate distribution with mean μ and variance-covariance

$$\gamma: \quad \gamma(I, j) = \sum_{i=1}^{n} \sigma_i^2 U_i U_j'$$

AD genetic model was employed for genetic entries I x j for (I = 1, 2...,8; j = 1, 2,...8) by using mating design of modified diallel crosses with F_1's and reciprocal F_1's, F_2's and reciprocal F_2's. Randomized complete block design of three blocks was used with a total of 120 genetics entries per block in single environment, and a total of 360 genetics entries per block in multiple environments. The environmental effects were assumed fixed effects with value of 100 in single environment. The three environments were assumed fixed with values of 50, 100, 150 respectively for the first, second and third environments. Additive and dominance genetic effects were assumed independent and random. The genetic entries were assigned at random within each block. Variance components were estimated by Minimum Norm Quadratic Unbiased Estimation method (MINQUE) (Rao, 1971 a, b; 1972) and predicted of genetic effects by AUP method. Jackknifing over block method (Miller 1974; Efron, 1982) conducted for estimating standard error. Pseudo-random normal deviates with zero mean and unit variance (0,1) were generated by the method of Kinderman and Monathan (1977). Given the true values of the specific effects of the variance components, balanced data were generated for additive dominance genetic model in single and multiple environments. 500 Simulations was used for computing sample mean of estimate, bias, Mean Squared Error (MSE) and power for genetic variance component of each specific effect of the studied model. Predicted mean, predicted variance and power were also computed for the prediction of the genetic effects for each specific effect of the AD model. MSE is calculated by Var, which $\mathrm{Var}(\hat{\theta}) + (\mathrm{bias})^2$ is usually used as a main criterion for comparing efficiency of estimation methods. Bias is calculated as $\hat{\theta} - \theta$. If $\mathrm{Bias}/\theta \leq 5\%$, the estimate $\hat{\theta}$ is considered as unbiased (Graybill and Wortham, 1956). Sampling variance of estimates is calculated by $Var(\hat{\theta}) = \frac{1}{n-1}\sum(\hat{\theta} - \bar{\hat{\theta}})^2$. The coefficient of efficiency (Zhu and Weir, 1994 b) is defined as: $C.E. = \frac{\sqrt{MSE}}{|\theta| + |Bias|}$.

Results

Simulation results of bias, C.E, MSE and power value for one trait variance are listed in Table 1. In single environment with 8 inbred lines in modified diallel crosses, variance components σ^2_A, σ^2_D and σ^2_ε, were unbiased with respective value of bias equal to 4.98, 1 and 0.4 % of the parameter value. 92% of the mean square error is belonging to additive effects, more than 7% to dominance and less than 1% to residual effects. Higher is the efficiency, less is the MSE. It implies that efficiency of additive variance in single environment is appreciable but less good than that of dominance and particularly that for residual variance. The significance of non-zero σ^2_A, σ^2_D and σ^2_ε can be detected with a probability of over 99%. Robustness of estimating one-trait variance is tested by simulation under the conditions of no specific variation. When the true value is zero for a specific parameter, the conclusion of non-significance can be drawn with a probability around 95% by the t-test. If the null hypothesis of zero variance is true, the probability of rejecting the null hypothesis of no variation is 5%. Non-significance of the additive and dominance can be respectively detected with probabilities around 95 and 99%. Probability of non-significance of the residual variance can be over 99%.

Table 1: Estimation of variance components for AD model in single environment

Variance component	True Value	Estimate	Bias	MSE	C.E.	Power
σ^2_A,	80	83.98	3.98	2370.82	0.58	1.00
σ^2_D	50	50.67	0.67	198.55	0.28	1.00
σ^2_ε	10	9.96	−0.04	0.61	0.08	1.00

However in multiple environments when GE interaction effects are considered, σ^2_A became highly unbiased with bias around 1% of the parameter value and with power value more than 97%. The changing of dominance variance is relatively small. Comparing to the results obtained in single environment bias of additive variance component decreased. MSE and C.E of additive and dominance variances increased but that of residual variance decreased significantly. σ^2_{AE} and σ^2_{DE} were highly unbiased with small biases and their significance can be detected with probability of over 99%. It can concluded that the effects of GE interaction effects in multiple environments tend to provide better estimation of additive variance component and particularly involve unbiased estimation of residual variance component which presents very small bias and MSE, and high efficiency and power value.

Under the assumption of no genotype by environment interaction effects (GE=0), all estimated variance components present minimum bias and MSE by the model of eight parent modified diallel crosses (Table 3). Bias of σ^2_A was around 0.15% of the parameter value and 0.01% for that of σ^2_D. Bias, MSE and power value of $\sigma^2\varepsilon$ (Table 3) were exactly similar to that of genetic full model with significant GE interaction effects (Table 2). σ^2_A, σ^2_D and σ^2_ε were significant with high power value and efficiency.

Table 2: Estimation of variance components for full model including GE interaction effects

Variance component	True Value	Estimate	Bias	MSE	C.E.	Power
σ^2_A	80	81.19	1.19	2839.40	0.65	0.98
σ^2_D	50	49.03	−0.97	259.14	0.31	1.00
σ^2_A	40	39.24	−0.76	267.05	0.40	1.00
σ^2_{DE}	25	24.68	−0.32	25.67	0.20	1.00
σ^2_ε	10	10.01	0.01	0.21	0.04	1.00

Table 3: Estimation of variance components when GE=0 in the full model

Variance component	True Value	Estimate	Bias	MSE	C.E.	Power
σ^2_A	80	80.12	0.12	2143.28	0.58	1.00
σ^2_D	50	49.75	−0.25	193.28	0.27	1.00
σ^2_{AE}	0	0.01	0.01	0.01	10	0.03
σ^2_{DE}	0	0.02	0.02	0.16	20	0.01
σ^2_ε	10	10.01	0.01	0.21	0.04	1.00

Table 4: Estimation of variance components for full model ignoring GE interaction effects

Variance component	True Value	Estimate	Bias	MSE	C.E.	Power
σ^2_a	80	91.01	11.01	2949.69	0.59	0.98
σ^2_D	50	54.87	4.87	280.80	0.30	1.00
σ^2_{AE}	—	—	—	—	—	—
σ^2_{DE}	—	—	—	—	—	—
σ^2_ε	10	71.41	61.41	4186.08	0.90	1.00

Table 5: Evaluation of genetic predictor in AD models

Predictor	Variance	Mean of Predictor	Variance of Predictor	Power
Single Env.				
$\hat{e}_A$	80	−0.0012	84.10	1.00
$\hat{e}_D$	50	−0.0001	52.20	1.00
Multiple Env.				
$\hat{e}_A$	80	−0.0027	81.30	0.98
$\hat{e}_D$	50	−0.0003	49.63	1.00
$\hat{e}_{AE}$	40	−0.0009	39.37	1.00
$\hat{e}_{DE}$	25	−0.0001	26.42	1.00
Multiple Env.				
$\hat{e}_A$	80	−0.0007	94.34	1.00
$\hat{e}_D$	50	−0.0000	857.33	1.00
Multiple Env.				
$\hat{e}_A$	80	−0.0030	80.17	1.00
$\hat{e}_D$	50	−0.0003	50.39	1.00

Robustness of estimating one-trait variance components was also tested by simulation under the conditions of no specific variation. The non-significance of GE interaction variance components was detected with a probability of 97% for σ^2_{AE} and 99% for σ^2_{DE}, (Table 3). It indicated that no matter whether or not genotype by environment interaction effects exit correct conclusions can be obtained for σ^2_A, σ^2_D and σ^2_ε.

In the opposite, when GE interaction effects were ignored, the additive and dominance variances were respectively overestimated with values of bias 14 and 10% of the corresponding parameter values. Residual variance component was particularly overestimated with the highest value (61.41) of bias (Table 4).

Significance of σ^2_A, σ^2_D and σ^2_ε can be detected with probability of over 97%. Ignored effects of GE interaction affected residual variance $\sigma 2\varepsilon$ in particular, then σ^2_A. Bias of σ^2_ε increased from 1 to 614% of the parameter value. Note that MSE of σ^2_ε also increased from 0.006 to 56.4% of the total value. However power value was still the same for all the components.

Simulation were conducted to predict the different specific effects in single and in multiple environment when GE exit or ignored and also under the assumption GE = 0. Mean and variance of predictor and power value were calculated. Extremely low biases for mean prediction of genetic effects were obtained with absolute values around $8.10^{-5} \sim 12.10^{-4}$ (Table 5). Both in single and multiple environments, additive and dominance variances of the genetic predictors were closed to the true value. The probability for detecting significance of additive and dominance variance of predictors was more than 99%. Additive and dominance variances of the predicted genetic interaction effects were also near the true value and were represented with high power value. When GE interaction effects were ignored, the values of additive and dominance genetic variances of prediction increased. However under the assumption GE=0, that two variances decreased and the results were better to those obtained before.

Discussion

Results obtained in single environment (Table 1) and that of multiple environments (Table 2 and 3) gave unbiased estimates for σ^2_A, σ^2_D, σ^2_{AE}, σ^2_{DE}, σ^2_ε, . According to the robustness of estimating one trait variance, it is indicated that, no matter whether or not GE interaction effects exit, correct conclusion can be obtained for variance components. In multiple environments when σ^2_{AE} and σ^2_{DE} were highly significant with power value of over. 99%, σ^2_A and σ^2_ε became highly unbiased with bias around 1 and 0.15% of the parameter value, respectively. Compared to the results obtained in single environment bias of σ^2_A and σ^2_ε decreased significantly. It is concluded that the effects of GE interaction effects in multiple environments tend to provide better estimation for additive variance component and particularly for residual variance component which involve very small bias and MSE. If there are no interaction effects when GE effects is ignored, values of σ^2_{AE} and σ^2_{DE} should be around zero. In this condition, genotype and environment effects should not be interdependent. The equation form of the genetic model of the full model in multiple environments became identical to that in single environment. The results can be similar to those in multiple environment with GE=0 (Table 3), where all variance components were more unbiased and MSE smaller than those in single environment. When the significance of GE interaction was ignored in the model giving the results in Table 4, effects were distributed particularly to the residual then to the additive and dominance variances of the reduced model. Consequently, variances of the specific effects were overestimated. When the full model was used with eight parents modified diallel crosses, the adjusted unbiased prediction with prior value 1 for all variances gave unbiased predicted effects with extremely small biases of mean prediction. The model in single environment is preferred to that of multiple environments when GE interaction effects are known as non-significant for the studied trait. Therefore, before the starting of experiment, previous information is necessary about the specific effects and GE interaction effects. By default, the appropriate model must be the full genetic model.

References

Cockerham, C. C., 1980. Random and fixed effects in plant genetics. Theor. Appl. Genet., 56:119-131.

Cockerham, C. C. and B. S. Weir, 1977. Quadratic analysis of reciprocal crosses. Biometrics, 33: 187-203.

Efron, B., 1982. The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia. Pa.

Falconer, D. S., 1996. Introduction to Quantitative Genetics, Ed. 4. Longman Press, New York.

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

Graybill, F. A. and A. W. Wortham, 1956. A note on uniformly best unbiased estimators for variance components. J. Am. Stat. Assoc., 51: 266-268.

Griffing, B., 1956. Concept of general and specific combining ability in relation to diallel crossing systems. Aust. J. Biol. Sci., 9: 463-493.

Hartley, H. D. and J. N. K. Rao, 1967. Maximum-likelihood estimation for the mixed analysis of variance model. Biometrika, 54: 93-108.

Henderson, C. R., 1988. Progress in statistical methods applied to quantitative genetics since 1976. Ch. 8 in Weir, B. S., Eisen, E. J., Goodman, M. M. and Namkoong, G. (eds), Proc. 2nd Internat. Conf. On Quantitative genetics. Sinauer, Sunderland, Mass., USA.

Kinderman, A. J. and J. F. Monathan, 1977. Computer generation of random variables using the ratio of normal deviates. Assoc, Comput. Math. Trans. Math. Soft., 3: 257-260.

Miller, R. G., 1974. The jackknife-a review. Biometrika, 61: 1-15.

Rao, C.R., 1971 a. Estimation of variance and covariance components MINQUE theory. J. Multi variate Analysis, 1: 257-275.

Rao, C.R., 1971 b. Minimum variance quadratic unbiased estimation of variance components. J. Multi variate Analysis, 1: 445-456.

Rao, C.R., 1972. Estimation of variance and covariance components in linear models. J. Multi variate Analysis, 67: 112-115.

Weir, B. S., 1996. Genetic data analysis II. Sinauer Associates, Inc. Publishers Sunderland, Massachusetts.

William, H. P., A. Saul, T. William and P. Brian, 1992. Numerical recipes in C: the art of scientific computing. Press Syndicate of the Univ. of Cambridge. USA.

Zhu, J., 1997. Analysis methods for genetic models (Chinese), pp: 295. Beijing: Agricultural Publishing House of China.

Zhu, J. and B.S. Weir, 1994 b. Analysis of cytoplasmic and maternal effects: II. genetic model for triploid endosperm. Theor. Appl. Genet., 89:160-166.

Zhu, J. and B.S. Weir, 1996. Diallel analysis for sex-linked and maternal effects. Theor. Appl. Genet., 92: 1-9.