

Analysis of Diallels

Compilation of Responses (updated 03/25/2004)

Query to Listserv (Forestgen): 'I've had some queries on what diallel programs are available and know of some, but would like to compile a listing of what is available: Program Name (and whether stand alone program or runs in SAS, etc.) and any documentation, articles, reviews; strengths/weaknesses/limitations of the program, availability (and cost).'

NOTE: These responses do not reflect an endorsement of a particular product and are certainly not meant to be an exhaustive listing of the tools available for analysis of diallel data.

Responses:

Respondent	E-mail Address	Program	Respondent's Remarks
Luis Apiolaza	luis.apiolaza@forestrytas.com.au	ASReml	It is my opinion that "diallels programs" per se are not needed anymore. Current general and specialist statistical software packages are more than enough to manage diallels or any more complex mating designs. In particular, ASReml is a generalised linear mixed models software that has been optimised for genetic analyses in such a way that runs orders of magnitude faster than SAS. I put together the ASReml cookbook (http://uncronopio.org/luis/asremml_cookbook.html), which is a collection of explanations for tree breeding analyses using ASReml, including diallel analysis.
Phillippe Baradat	philippe.baradat@club-internet.fr	DIOGENE	Diallel processing using multivariate data is included in the DIOGENE software (in French) which directly runs on Sun stations/servers (Solaris 7, 8 and 9). You can do much more than non-orthogonal diallel analyses (with or without reciprocal crosses, with or without selfings) and integrate it with other integrated programs or routines (Discriminant Analysis, multiple comparison of effects, selection indices, jackknife-bootstrap...). It is free software written in C and Fortran 95 and you can get the compiled codes for Solaris or the source codes. It can run on PC using Linux, after recompilation with Intel compiler (but, in this alternative, it is necessary to make some verification using standard procedures because the last versions of the major part of programs have run only under Solaris).
Greg Dutkowski	g.dutkowski@utas.edu.au	ASReml	General mixed model program that will estimate variance components and solve the mixed model equations. It is very fast and can fit almost any model that you can think of.

Elisabetta Frascaroli	elifrascaroli@hotmail.com	SAS	I wrote a program for my own use (I do research on maize breeding), running in SAS. Now it is prepared for a complete diallel cross scheme, with reciprocals and the analysis as in Hayman. It is quite unrefined (the program computes all the MS needed, but does not assemble the final complete ANOVA). I've never thought of publishing it, but it could be not a bad idea. Of course it should be refined. If you think it could be of more wide interest, I may consider editing it better.
Fikret Isik	fikretisik@yahoo.com	SAS (mixed models)	See Xiang, B. and Li, B. 2001. A new mixed analytical method for genetic analysis of diallel data. Can. J. For. Res. 31:2252-2259.
Manjit S. Kang	MKang@agctr.lsu.edu	SAS	See Zhang, Y., and M. S. Kang. 1997. DIALLEL-SAS: A SAS program for Griffing's diallel analyses. Agron. J. 89:176-182
Colin Matheson	Colin.Matheson@csiro.au	Apart from the SAS programs Harry Wu will have sent you, I also use an updated version of the old Schaffer and Usanis program DIALL.	This program was written in Fortran IV when I came across it in 1972. I spliced Snyder and Nance's program that estimates GCA parental values (and hence breeding values) and SCA values for crosses (by difference) from GCA and SCA effects produced by DIALL. I re-wrote the program in Fortran 77 while in Oxford in 1985-6 with Julie Smith and re-wrote it again in Fortran 90, enlarging it so it will take, I think, 30 variates and 50 parental lines (from memory). It now also calculates heritabilities and standard errors (according to Kempthorne) and genetic correlations and their standard errors (according to Robertson). It thinks it produces heritabilities for the covariance analyses too, but that is only because I've been a bit lazy about telling the program which sort of analysis it is printing out. There is also provision for missing values, but they will be missing for ALL variates. If you have the original NCSU guide that came with it, you will see that it fits a fully random model using least squares including, if required, reciprocal, maternal, GCA x location as well as replicate, GCA & SCA effects. I compiled mine using Lahey LF95 Fortran and you have to run it in a DOS window. The program asks you for the input filename and then for an output file name and then asks for some text that it writes out among the first few lines of output.
Bob Westfall	bwestfall@fs.fed.us		Afraid I can't help you much, wrt diallel software. There is one called DISCLONE, from the Atlantic Forestry Research Center in Canada.
Harry Wu	Harry.Wu@csiro.au	SAS	I published two diallel analyses programs run on SAS: DIAFIXED.SAS (Wu, H.X. and Matheson, A.C. 2000. Analysis of half-diallel mating design with missing crosses: theory and SAS program for testing and estimating GCA and SCA fixed effects. Silvae Genetica 49:130-137.) DIARANDM.SAS (Wu, H.X. and Matheson, A.C. 2001. Analysis of half-diallel mating design with missing crosses: theory and SAS program for testing and estimating GCA and SCA variance

Further Information:

References mentioned by respondents:

Agwanda, C.O., Baradat, P, Eskes, A.B., Cilas, C. and Charrier, A. 2003. Selection for bean and liquor qualities within related hybrids of Arabica coffee in multilocal field trials *Euphytica* **131**: 1–14.

Summary: The success of a new variety of Arabica coffee (*Coffea arabica* L.) depends to an important extent on its liquor and bean qualities. Selection for these traits is however constrained by the prevalence of large genotype-by-environment (G×E) interactions in conjunction with the low genetic variability characteristic of this species. The aim of this study was to assess the extent to which key bean and liquor traits are affected by the environments and could be combined in order to improve the efficiency of selection for coffee quality in a narrow genetic basis set of hybrids obtained from a set of related lines and collectively representing the Ruiru 11 cultivar. Twenty-one full-sib families representative of this hybrid cultivar grown in Kenya in five sites exhibiting strong edaphic and climatic differences were used for the study. Rainfall amounts during various phases of berry development were used to explain the differences observed in the discriminating abilities of the locations for bean and liquor traits. The results showed that these families were best differentiated for bean sizes in the site where moisture supply was optimal throughout berry expansion and filling stages, whereas discrimination on the basis of liquor traits were best observed in the site where moderate moisture stress occurred during bean filling stage. The overall precision in prediction of family values was low for liquor qualities; but a much more efficient selection for large bold beans, optimally based on the AA grade was shown to be possible. Selection indices for family selection were computed to realize a trade-off between genetic gains in bean size and liquor flavour.

Cilas, C., Bouharmont, P., Boccara, M., Eskes, A.B., and Baradat, Ph. 1998. Prediction of Genetic Value for Coffee Production in *Coffea arabica* from a Half-Diallel with Lines and Hybrids *Euphytica* **104**: 49-59.

Summary: Genetic analysis of yield and morphological traits has been carried out in *Coffea arabica* from a half-diallel including the parental lines. The trial was established in west Cameroon with completely randomized single-tree plots. Observations included yield (four years), stem diameter, height and number of primaries. General combining abilities (GCA) and specific combining abilities (SCA) as heritabilities and genetic correlations were estimated. A significant SCA variance was observed for all the traits. Morphological traits, stem diameter, plant height and number of primaries, were genetically correlated to the yield. The hybrids were, on the average, better performing than lines. There was no clear relationship between performance of lines and their general combining ability. Contribution of the seven lines to the SCA sum of squares was shown to be unequal for all the studied traits, one

parent (Java) being far the most interactive. This variation of interactivity seemed to be related to different levels of residual heterozygosity among parental lines. Indirect prediction of yield using morphological traits gave the same value for expected genetic gain as the prediction based on cumulated yield over four years. The multitrait indirect selection for yield operated in a different way as far as GCA and SCA are concerned. The application of these results to coffee breeding strategies involving hybrid selection is discussed.

Xiang, Bin and Li, Bailian. Best linear unbiased prediction of clonal breeding values and genetic values from full-sib mating designs. *Canadian Journal of Forest Research* **33**:2036–2043

Abstract: Full-sib progeny tests with clonal replicates may provide better breeding value estimates and the greatest genetic gain in a tree improvement program. Clonal breeding values (CBV) that combine the family and within-family breeding values due to additive genetic effects can maximize the genetic gain for advanced generation breeding. Clonal genetic values (CGV) that further incorporate full-sib family specific combining ability due to nonadditive genetic effect can maximize gain for a deployment program with clonal propagation techniques. The best linear unbiased prediction (BLUP) is the best statistical method for estimating both CBV and CGV because of its desirable statistical properties compared with the heritability-based gain calculation. A BLUP method for determining both the CBV and CGV for full-sib clonal progeny tests was proposed in this paper. The formulas for CBV and CGV were derived using general BLUP methodology, and formulas were derived for the calculations of their standard errors. An analytical method by using a standard statistical package (SAS PROC MIXED) was presented for CBV and CGV calculations from any fullsib mating designs.

Wu, H. and Matheson, A.C. 2001. Analyses of half-diallel mating designs with missing crosses: Theory and SAS program for testing and estimating GCA and SCA variance components. *Silvae Genetica* **50**(5-6):265-271.

Abstract: The half-diallel mating design, particularly a series of disconnected half-diallel mating design has been widely adopted as a mating design in many commercially important tree species for estimating the variances of general (GCA) and specific (SCA) combining abilities, heritability, and genetic correlation. Standard commercial statistical packages do not allow direct specification of the half-diallel model and therefore are not able to analyse the diallel mating design even for a balanced diallel mating structure (no missing crosses). Published special computer programs for diallel analyses do not provide an adequate solution for GCA and SCA variance components with hypothesis testing for half-diallel mating designs with missing crosses. This paper presents the theory of the method of fitting constants for analysing half-diallel mating designs with or without missing crosses. A SAS computer program (DIARANDM.SAS) for testing hypotheses and estimation of GCA and SCA variance components is developed. The program accommodates multiple variables, multiple environments, environment by GCA and environment by SCA interactions as well as data

imbalance (e.g. unequal number of observations among cells and missing cells). The DIARANDM.SAS output includes (1) hypothesis testing for variance components of GCA, SCA and all interaction effects, (2) estimate of GCA, SCA and all interaction variances, (3) estimate of standard error for GCA and SCA and all interaction variance components, and (4) variance and covariance matrix for computing genetic correlations and other genetic and environmental parameters. Two examples are given to demonstrate the use of the program.

Wu, H.X. and Matheson, A.C. 2000. Analysis of half-diallel mating design with missing crosses: theory and SAS program for testing and estimating GCA and SCA fixed effects. *Silvae Genetica* **49**:130-137.

Abstract: The half-diallel mating design, particularly a series of disconnected half-diallels has been widely adopted as a mating design for estimating genetic parameters and for future selection in many commercially important tree species. Standard commercial statistical packages do not allow direct specification of the linear model associated with the half-diallel design and therefore are not capable of analysing diallel mating designs, even for balanced diallel matings (no missing crosses). Published special computer programs for diallel analyses do not provide an adequate solution for GCA and SCA fixed effects in diallels with missing crosses. This paper presents the least squares theory for analysing half-diallel mating designs with missing crosses, and a SAS computer program (DIAFIXED.SAS), developed to test the significance of GCA and SCA effects and estimate the GCA and SCA fixed effects. The program is flexible enough to accommodate different number of parents, multiple environments and missing individual trees as well as missing whole plots. The DIAFIXED.SAS output includes (1) hypothesis testing for GCA and SCA fixed effects and environmental effects, (2) estimates of GCA and SCA fixed effects, (3) estimates of standard errors of GCA and SCA fixed effects. Results from a 6 by 6 half-diallel for radiata pine planted in two sites are also presented.

Zhang, Y., and Kang, M.S. 1997. DIALLEL-SAS: A SAS program for Griffing's diallel analyses *Agronomy Journal* **89**:176-182.

Abstract: Diallel mating designs are generally used to obtain estimates of genetic effects for a fixed set of parental Lines (fixed effects) or to estimate general combining ability (GCA) and specific combining ability (SCA) variance components from a set of randomly chosen parental lines (random effects) from multiple-environment experiments. Published computer programs for diallel analyses do not provide a partition of F-1 hybrids x environments interaction into GCA x environments, SCA x environments, and reciprocal x environments components for Griffing's Methods 1 and 3, and into GCA x environments and SCA x environments components for Griffing's Methods 2 and 4. Existing programs for Methods 1 and 3 also do not partition reciprocal effects into maternal and nonmaternal effects. This paper reports the development of a comprehensive SAS program (DIALLEL-SAS) for partitioning interactions between environments (years and/or locations) and F-1 hybrids. DIALLEL-SAS can be run on any

microcomputer with SAS, as well as on mainframe computers installed with SAS via UNIX or TSO. The DIALLEL-SAS output includes (i) mean squares for environments, F-1 hybrids, F-1 hybrids x environments, GCA, SCA, reciprocal, maternal, nonmaternal, GCA x environments, SCA x environments, reciprocal x environments, maternal x environments, nonmaternal x environments, (ii) estimates of GCA and maternal effects for each parental line, and (iii) estimates of SCA, nonmaternal, and reciprocal effects for each F-1 hybrid.

Kang, Manjit S. (ed). 2003. Handbook of formulas and software for plant geneticists and tree breeders. Food Products Press, New York. 347 p.

Weblinks:

ASReml:

- Supplier: <http://www.vsn-intl.com/ASReml/index.htm>
- Forestry Examples: http://uncronopio.org/luis/asrem1_cookbook.html
- Forestry Primer: http://members.forestry.crc.org.au/cgi-bin/doc.pl?rm=view_doc&doc_id=405