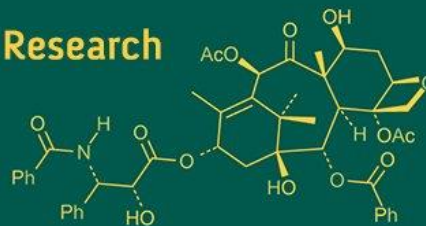
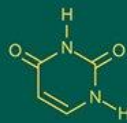
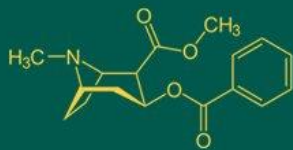


International Journal of Advanced Biochemistry Research



ISSN Print: 2617-4693
 ISSN Online: 2617-4707
 IJABR 2024; SP-8(6): 463-467
www.biochemjournal.com
 Received: 13-04-2024
 Accepted: 21-05-2024

Rohit Kundu

Department of Mathematics
 and Statistics, College of Basic
 Sciences and Humanities, CCS
 Haryana Agricultural
 University, Hisar, Haryana,
 India

OP Sheoran

Department of Mathematics
 and Statistics, College of Basic
 Sciences and Humanities, CCS
 Haryana Agricultural
 University, Hisar, Haryana,
 India

Vinay Kumar

Department of Mathematics
 and Statistics, College of Basic
 Sciences and Humanities, CCS
 Haryana Agricultural
 University, Hisar, Haryana,
 India

Corresponding Author:**OP Sheoran**

Department of Mathematics
 and Statistics, College of Basic
 Sciences and Humanities, CCS
 Haryana Agricultural
 University, Hisar, Haryana,
 India

Empowering plant breeding research: A comprehensive web-based analysis tool for partial diallel crosses

Rohit Kundu, OP Sheoran and Vinay Kumar

DOI: <https://doi.org/10.33545/26174693.2024.v8.i6Sf.1363>

Abstract

The OPSTAT "Partial Diallel Analysis" module represents a pivotal advancement in plant breeding research, offering a sophisticated yet user-friendly platform for analyzing partial diallel cross data. This online tool employs a robust algorithm to automate complex statistical analyses, facilitating the exploration of genetic architecture and trait inheritance in breeding experiments. Through Randomized Block Design (RBD) and combining ability analysis, the tool enables researchers to assess treatment effects, estimate genetic parameters such as general combining ability (GCA) and specific combining ability (SCA), and identify superior parental lines and effective parent combinations for breeding programs. The outputs generated, including ANOVA tables, estimated genetic effects, and variance component estimates, provide comprehensive insights into experimental data, empowering breeders to make informed decisions and accelerate the development of improved crop varieties with desirable traits.

Keywords: Partial diallel crosses, plant breeding, genetic analysis, OPSTAT, statistical algorithm, combining ability, trait inheritance, genetic parameters, breeding decisions, crop improvement, sustainability

Introduction

Partial diallel crosses analysis is a powerful statistical technique widely used in agriculture and animal breeding programs to assess the genetic potential of different parental lines and understand the inheritance of traits. This method involves crossing a subset of parents in a planned manner, enabling researchers to estimate various genetic parameters such as general combining ability (GCA), specific combining ability (SCA), and heritability. These estimates are crucial for selecting superior parental lines for future breeding programs, leading to the development of improved crop varieties and livestock breeds.

The concept of diallel crosses was first introduced by G.H. Shull in 1936, who studied the inheritance of quantitative traits in maize. J.W. Green further developed this technique in 1966 by introducing the concept of partial diallel crosses, which involves crossing only a subset of parents. This significantly reduces the number of crosses needed, making the analysis more practical and cost-effective.

Partial diallel crosses help researchers assess genetic diversity within a population by evaluating the performance of different crosses. This information is critical for identifying superior parental lines and developing new varieties or breeds with desirable traits. The analysis helps breeders predict the performance of hybrid offspring by estimating the combining ability of different parental lines, crucial for selecting the best parents for future breeding programs. Additionally, partial diallel crosses allow researchers to estimate the heritability of traits, providing valuable insights into the genetic control of traits and informing breeding strategies.

Notable contributions to the theory and application of partial diallel crosses have been made by researchers such as Kempthorne and Curnow (1961)^[9], Kaushik and Puri (1985)^[8], and Arya and Narain (1986)^[11]. Further advancements have been driven by researchers like Hinkelmann (2009)^[6], who focused on the statistical design and analysis of these crosses.

These developments have led to more robust and sophisticated methods for analyzing data from partial diallel crosses, providing researchers with greater insights into the genetic architecture of traits and enabling more precise predictions of hybrid performance.

These techniques have been widely used in various agricultural and animal breeding programs. For instance, they are used to evaluate the combining ability of different parental lines to develop improved crop varieties with traits such as higher yield, disease resistance, and improved quality (Singh et al., 2020; Gupta et al., 2019) ^[13, 5]. In animal breeding, partial diallel crosses help improve productivity traits such as milk production in dairy cows, meat quality in beef cattle, and egg production in poultry (Kumar et al., 2021; Wang et al., 2018) ^[11, 15]. This technique is also utilized to study the inheritance of complex traits, aiding researchers in understanding the genetic basis of these traits and developing more effective breeding strategies (Zhang et al., 2022; Liu et al., 2017) ^[16, 12].

Recent advancements in molecular markers and genomic technologies have significantly revolutionized partial diallel crosses analysis. Researchers now use these tools to identify and select desirable genetic combinations with greater accuracy and efficiency, leading to faster and more efficient breeding programs.

Despite the advancements, the analysis of partial diallel crosses can be computationally complex, especially for large datasets. While sophisticated statistical software packages like SAS, R, and specialized breeding software have streamlined the process, there is still a need for accessible and user-friendly online computational tools. These tools would allow researchers to perform analyses without installing specialized software, making the process more accessible and providing a platform for collaboration and data sharing among researchers.

Developing an online tool for partial diallel crosses analysis aims to automate the process of data analysis, eliminating the need for manual calculations and minimizing errors. The tool would be freely available, making the analysis accessible to researchers with limited resources. Additionally, the tool would offer interactive visualizations of the results, making them more intuitive and easier to understand. It would also facilitate data sharing and collaboration among researchers, fostering advancements in breeding research. By achieving these objectives, an online tool for partial diallel crosses analysis would empower researchers to conduct more efficient and collaborative studies, ultimately leading to the development of improved crop varieties and livestock breeds, contributing to greater agricultural productivity and sustainability.

Methods and Materials

The OPSTAT "Partial Diallel Analysis" module is a freely accessible online tool designed to facilitate the analysis of partial diallel cross data, a fundamental technique in plant and animal breeding. This module, constructed using a combination of HTML, ASP (Active Server Pages), and JavaScript, provides a user-friendly interface and robust computational capabilities. It can be accessed at the following URL: [<http://14.139.232.166/opstat/partial.htm>]. The module operates on a server-client architecture, where the user interacts with the client-side interface to input data and select analysis options, while the server-side handles the computational tasks and data processing.

a) Data Arrangements and Entry

Partial diallel crosses are used when the number of parents involved in a breeding program becomes too large to manage all possible crosses. In these cases, a sample of crosses is selected, ensuring that each parent is involved in a similar number of crosses. This approach, known as the circulant sample (Kempthorne and Curnow, 1961) ^[9], is employed in the OPSTAT "Partial Diallel Analysis" module.

To construct a partial diallel cross based on a circulant sample, two key parameters are defined firstly Sample size (s) i.e. This represents the number of times a parent is involved in crosses and secondly constant (k) i.e. This constant is calculated as $k = (v + 1 - s)/2$, where 'v' is the number of parental lines. The selection of 's' and 'v' must ensure that they are not both odd or both even. For example, with 8 parents ($v = 8$) and each parent involved in 3 crosses ($s = 3$), then $k = (8 + 1 - 3)/2 = 3$. In this scenario, the sampled crosses would be 1 x 4, 1 x 5, 1 x 6, 2 x 5, 2 x 6, 2 x 7, 3 x 6, 3 x 7, 3 x 8, 4 x 7, 4 x 8, and 5 x 8.

The data for these crosses is arranged text area of the web page in a specific manner. The replications of the first cross (1 x 4) are entered on the first line, separated by spaces or tabs. The data for cross 1 x 5 is entered in a similar manner on the next line, and so on. If you are analyzing multiple traits, the data for each additional trait is entered immediately after the first trait data, following the same pattern for all crosses.

If you are including the parents in the analysis (i.e., self-pollinations), you would enter the data for the parents after the cross data. In the example above, the data would be arranged as 1 x 4, 1 x 5, 1 x 6, 2 x 5, 2 x 6, 2 x 7, 3 x 6, 3 x 7, 3 x 8, 4 x 7, 4 x 8, 5 x 8, 1 x 1, 2 x 2, 3 x 3, 4 x 4, 5 x 5, 6 x 6, 7 x 7, and 8 x 8. The data arrangements and their entry in the text area has been depicted in Fig 1 for partial diallel crosses when parents are excluded.

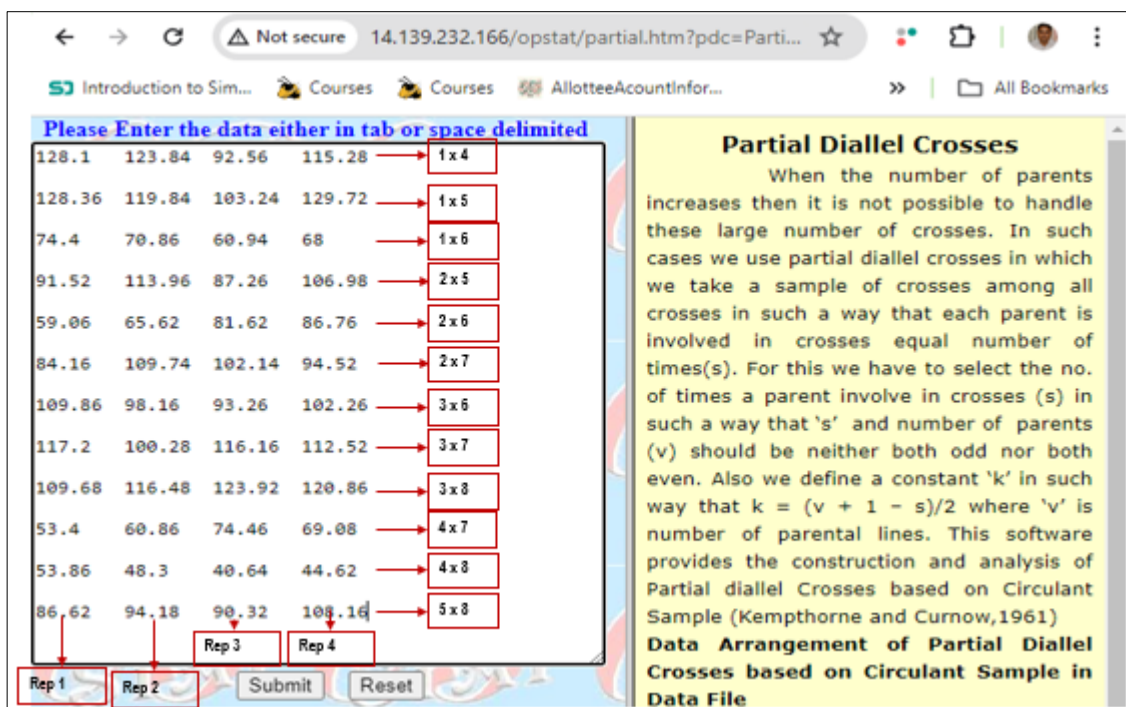


Fig 1: Data arrangements and entry in interface

The user interface (Fig 2) guides researchers through a series of inputs to define the experimental design and data characteristics of a partial diallel cross.

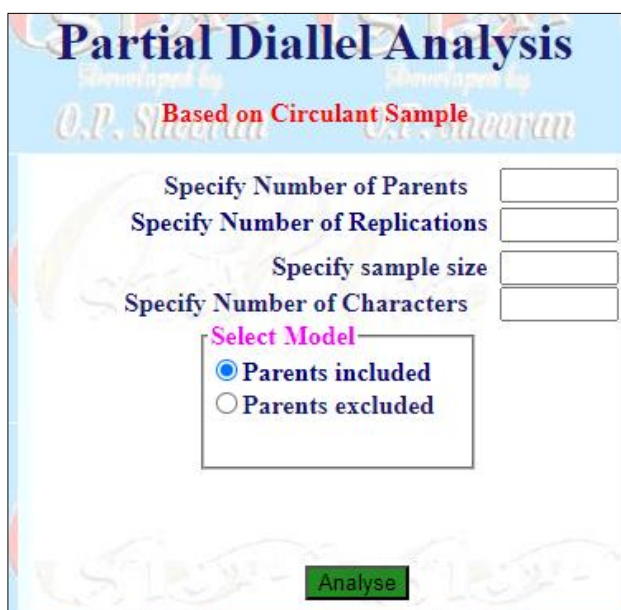


Fig 2: Parameters input interface

First, the number of parental lines used in the experiment is specified. This information is crucial for determining the appropriate design matrix and calculating the degrees of freedom for statistical analysis. Next, the number of times each cross is repeated within the experiment, known as the number of replications, is entered. This parameter is essential for estimating the error variance, which is crucial for determining the significance of differences between treatment means. The user then defines the sample size, which refers to the number of times each parent is involved in crosses within the partial diallel. This parameter plays a key role in determining the size of the subset of crosses selected for analysis.

Finally, the number of traits or variables being measured in the study is specified. The software will then perform the analysis for each trait, generating separate results. Once these parameters are entered, the user has the option to select the appropriate model, which determines whether parental lines are included in the analysis. The inclusion or exclusion of parents directly influences the statistical analysis performed and the specific genetic parameters that are estimated.

These parameters are collected and transmitted to the server-side script for further processing and analysis. The user-friendly interface allows researchers to easily specify the experimental details required for a comprehensive and informative analysis. This step ensures that the analysis is conducted accurately, taking into account the specific characteristics of the experiment.

b) The Algorithm

The algorithm for partial diallel cross analysis involves a series of systematic steps to elucidate the genetic basis of traits in plant breeding experiments.

- Initially, user-submitted data is read and stored in a multidimensional array. The algorithm then calculates the degrees of freedom for various sources of variation, including total degrees of freedom, degrees of freedom for replications, treatments, and error. These calculations depend on the number of parents, replications, sample size, and the chosen model, whether parents are included or excluded.
- Next, the algorithm performs a Randomized Block Design (RBD) analysis of variance (ANOVA) for each character being analyzed. It calculates the sums of squares and mean squares for each source of variation and computes the F-statistic to assess the significance of the treatment effect.
- The core of the analysis lies in the combining ability analysis, where treatment means for each cross are calculated. The algorithm constructs incidence matrices to represent the structure of the partial diallel crosses

based on the chosen model. Using these matrices, the algorithm calculates the general combining ability (GCA) effects for each parent and the specific combining ability (SCA) effects for each cross. It then calculates sums of squares and mean squares for GCA, SCA, and error, followed by F-tests to determine the significance of GCA and SCA effects.

- If a random effects model is chosen, the algorithm proceeds to estimate variance components for GCA, SCA, and error. These components represent the relative contributions of each genetic source to the total phenotypic variance.
- Finally, the algorithm presents a comprehensive set of results, including ANOVA tables for both the RBD and combining ability analysis, tables displaying estimated GCA effects for each parent, estimated SCA effects for each cross, variances, standard errors, and critical differences among effects. If applicable, variance component estimates are also provided.
- The algorithm relies heavily on user-defined parameters and the specific model selected. Calculations for GCA, SCA, and variance component estimation are implemented using VBScript. The data is assumed to be arranged in a specific format, as described in the help documentation.

c) **Outputs of the partial diallel analysis algorithm & their interpretations**

The partial diallel cross analysis algorithm generates a comprehensive set of results, offering detailed insights into the genetic architecture of traits and aiding breeding decisions.

First, the algorithm produces two ANOVA tables: one for the Randomized Block Design (RBD) and another for the combining ability analysis. The RBD ANOVA table partitions the total variation into sources such as replications, treatments (crosses), and error, providing sums of squares (SS), mean squares (MS), F-statistics, and p-values for each source. This table highlights the significance of differences among the treatments (crosses) and the overall variability within the experiment.

The combining ability ANOVA table, on the other hand, partitions the treatment variation into general combining ability (GCA), specific combining ability (SCA), and error (if applicable). This table also includes SS, MS, F-statistics, and p-values for each source of variation, revealing the contributions of GCA and SCA effects to the observed trait variation.

Next, the algorithm generates tables summarizing the estimated genetic effects. The GCA effects table presents the estimated GCA effects for each parent, indicating the overall contribution of each parent to the trait in the analyzed crosses. Parents with significant and positive GCA effects are good general combiners and are beneficial for trait improvement in their offspring, while negative GCA effects suggest a negative influence on the trait.

The SCA effects table shows the estimated SCA effects for each cross, indicating specific interactions between parents that influence the trait. Significant SCA effects imply that certain parent combinations are particularly effective in expressing the desired trait, making them candidates for future breeding programs.

If a random effects model is chosen, the algorithm provides variance component estimates for GCA, SCA, and error.

These estimates quantify the relative contributions of these genetic sources to the total phenotypic variation. High variance components for GCA indicate a significant influence of genetic differences among parents on the trait, while high variance components for SCA suggest that specific parent combinations play a major role in determining trait expression.

The results from this analysis offer valuable information for breeders. Parents with significant and positive GCA effects for desirable traits should be prioritized for breeding due to their good general combining ability. Crosses with significant SCA effects can be identified for developing hybrid varieties with desirable trait combinations. Understanding the relative contributions of GCA and SCA to trait variation guides strategies for selecting parents and crosses, ultimately aiding in the development of improved crop varieties or livestock breeds.

Summary and Conclusions

Partial diallel crosses provide a powerful framework for dissecting the genetic architecture of traits in plant breeding. This technique allows researchers to efficiently evaluate combining abilities of parental lines and identify promising combinations for developing improved varieties. Recognizing the need for a streamlined and accessible tool to analyze partial diallel data, the OPSTAT "Partial Diallel Analysis" module was developed. This freely accessible online tool, with a user-friendly interface, facilitates data entry, analysis option selection, and generates comprehensive results.

The algorithm implemented in the tool enables researchers to explore the genetic basis of traits by examining combining abilities of parental lines and identifying promising combinations for breeding programs. Through Randomized Block Design (RBD) analysis and combining ability analysis, the tool partitions variation, assesses treatment effects, and estimates genetic parameters such as general combining ability (GCA) and specific combining ability (SCA). These analyses provide valuable insights into the relative contributions of different genetic sources to trait variation, guiding breeding decisions and strategies.

The outputs generated by the tool, including ANOVA tables, estimated genetic effects, and variance component estimates, offer researchers a detailed understanding of the experimental data. By highlighting significant treatment effects, identifying superior parental lines, and pinpointing effective parent combinations, the results empower breeders to make informed decisions in selecting parents and designing breeding programs. This data-driven approach facilitates the development of improved crop varieties and livestock breeds with desirable traits, ultimately contributing to enhanced agricultural productivity and sustainability.

The OPSTAT "Partial Diallel Analysis" module represents a significant advancement in plant breeding and genetic analysis. This user-friendly online tool streamlines the analysis process, making it accessible to a wider audience of researchers and breeders. Through its efficient analysis capabilities, the tool empowers breeders to make informed decisions, contributing to the development of improved crop varieties and livestock breeds, ultimately supporting agricultural productivity and sustainability.

References

1. Arya AS, Narain P. Construction and analysis of partial diallel crosses based on circulant sample. *J Indian Soc. Agric. Stat.* 1986;38(1):01-10.

2. Comstock RE, Robinson HF. The components of genetic variance in populations of biparental progenies and their use in estimating the genetic variances and covariance of the parents. *Biometrics*. 1952;8(1):01-10.
3. Crossa J. Genomic selection and prediction: A review. *Euphytica*. 2014;199(1):01-16.
4. Green JW. The use of diallel crosses in plant breeding. *Res. Bull. Ohio. Agric. Res. Dev. Cent*, 1966, (981).
5. Gupta PK, Kumar A, Singh VK. Combining ability analysis and heterosis for yield and quality traits in rice (*Oryza sativa* L.). *J Crop Improv*. 2019;33(2):251-268.
6. Hinkelmann K. Design and analysis of experiments. *Wiley Series in Probability and Statistics*; c2009.
7. Holland JB. Genetic architecture of complex traits. *Oxford University Press*; c2012.
8. Kaushik LS, Puri PD. Augmented partial diallel crosses based on circulant sample. *Biorn. 1*. 1985;27(4):397-403.
9. Kempthorne O, Curnow RN. The genetic basis of heterosis. *Iowa State University Press*; c1961.
10. Kumar R, Kumar D, Kumar S. Molecular marker assisted selection for crop improvement. *Springer International Publishing*; c2016.
11. Kumar S, Singh AK, Singh VK. Combining ability analysis and heterosis for grain yield and its attributes in barley (*Hordeum vulgare* L.). *J Appl Nat Sci*. 2021;13(1):34-41.
12. Liu J, Zhang X, Zhou Y, Liu Y, Zhang J, Gao C. Genetic architecture of grain yield and its component traits in rice under different water regimes. *PLoS One*. 2017;12(4).
13. Singh VK, Kumar A, Singh R. Combining ability analysis and heterosis for yield and yield-related traits in chickpea (*Cicer arietinum* L.). *J Appl. Nat. Sci*. 2020;12(2):322-330.
14. Shull GH. The origin of the maize plant. *J Hered*. 1936;27(4):121-128.
15. Wang Y, Chen S, Zhang Q, Gao Y, Zhang Z. Combining ability analysis and heterosis for growth traits in Chinese Holstein dairy cattle. *Anim. Sci. J*. 2018;89(11):1038-1046.
16. Zhang B, Zhu J, Yan Q, Chen X, Zhang J, Wu Y. Genetic architecture of grain yield and its component traits in wheat under contrasting water regimes. *Front Plant Sci*. 2022;12:763241.