Choosing an analysis

Muhammad Yaseen

2026-11-03

Table of contents

1 Abstract **1**

2 Decision table **1**

3 Which method does a data set fit? **2**

4 What happens when a plot or a cross is missing **3**

5 Column names **4**

6 Summary **4**

7 References **4**

1 Abstract

Which function, method and model to use depends on the crosses grown and on the question asked. This vignette gives a decision table and shows how to check that a data set has the crosses an analysis needs.

2 Decision table

Crosses grown	Question	Function
parents, F_1 and reciprocals	combining ability, reciprocal effects	`Griffing(Method = 1)`
parents and one set of F_1	combining ability	`Griffing(Method = 2)`
F_1 and reciprocals, no parents	combining ability, reciprocal effects	`Griffing(Method = 3)`
one set of F_1 only	combining ability	`Griffing(Method = 4)`

Crosses grown	Question	Function
parents, F_1 and reciprocals	gene action: additive and dominance components, W_r-V_r graph	`Hayman()`
parents, F_1 and reciprocals	maternal and non-maternal reciprocal effects	`Griffing(Method = 1, Reciprocal = "split")`
a published table of cross means	combining ability from the means and the published error	`GriffingMeans()`
any of Griffing's layouts in several environments	stability of gca and sca: interactions with environments	`GriffingEnv()`
parents and crosses (a check variety optional)	heterosis over the mid-parent, better parent, check	`Heterosis()`
parents and one set of F_1	variety effects and average, variety and specific heterosis	`GardnerEberhart()`
two groups of parents crossed with each other	combining ability of each group	`PartialDiallel()`
lines crossed with testers	combining ability, contributions of lines and testers	`LineTester()`
North Carolina Design I, II or III	additive and dominance variance	`NCDesign()`
not yet grown	number of replicates for the tests	`DiallelPower()`, `DiallelSim()`

Model. Choose Model 1 (fixed effects) when the parents are of interest in themselves (the usual case in a breeding programme): it gives effects and standard errors. Choose Model 2 (random effects) when the parents are a random sample of a population and the aim is to estimate variance components (additive and dominance variance).

3 Which method does a data set fit?

```
cells <- function(d) {
  i <- as.numeric(as.character(d$Cross1))
  j <- as.numeric(as.character(d$Cross2))
  u <- unique(data.frame(i = i, j = j))
  c(parents = length(unique(c(i, j))), selfs = sum(u$i == u$j), above =
   ↪  sum(u$i < u$j), below = sum(u$i > u$j))
```

```
}
rbind(
    Data1 = cells(d = GriffingData1)
  , Data2 = cells(d = GriffingData2)
  , Data3 = cells(d = GriffingData3)
  , Data4 = cells(d = GriffingData4)
  )
```

```
      parents selfs above below
Data1       8     8    28    28
Data2       8     8    28     0
Data3       8     0    28    28
Data4       8     0    28     0
```

With P parents: Method 1 needs P selfs and $P(P-1)$ crosses, Method 2 P selfs and $P(P-1)/2$ crosses (either way round), Method 3 $P(P-1)$ crosses and no selfs, Method 4 $P(P-1)/2$ crosses and no selfs.

4 What happens when a plot or a cross is missing

A missing plot is allowed in `Griffing()`, `GriffingEnv()`, `Heterosis()`, `GardnerEberhart()`, `LineTester()` and `NCDesign()`: the cross means are then least-squares means adjusted for blocks, the number of replicates is the harmonic mean of the plot numbers, and the label of the result says so. `Hayman()` and `PartialDiallel()` need every plot. Every cross of the method must be present, though; a missing cross is reported with its name:

```
d <- subset(x = GriffingData1, subset = !(Cross1 == 1 & Cross2 == 2))
tryCatch(
    expr  = Griffing(
        y          = Yield
      , Rep        = Rep
      , Cross1     = Cross1
      , Cross2     = Cross2
      , data       = d
      , Method     = 1
      , Model      = 1
      , Reciprocal = c("combined", "split")[1]
      )
  , error = function(e) conditionMessage(e)
  )
```

```
[1] "Griffing(): 1 cross(es) of Method 1 are missing from `data` (for example 1 x 2). The ar
```

Incomplete diallels need a linear mixed model with the gca, sca and reciprocal effects as random terms.

5 Column names

The arguments take a column by name, as a string, or a vector:

```
d <- GriffingData1
names(d) <- c("Female", "Male", "Block", "GrainYield")
fit <-
  Griffing(
      y          = "GrainYield"
    , Rep        = Block
    , Cross1     = Female
    , Cross2     = Male
    , data       = d
    , Method     = 1
    , Model      = 1
    , Reciprocal = c("combined", "split")[1]
    )
fit$ANOVA
```

```
             Df  Sum Sq Mean Sq F value Pr(>F)
gca           7  3804.6  543.52 18.8868 <2e-16 ***
sca          28 22060.8  787.89 27.3783 <2e-16 ***
reciprocals  28   365.5   13.05  0.4536 0.9923
Error       189  5439.0   28.78
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

6 Summary

The crosses grown fix the analysis and the method; the purpose fixes the model. Check the cells before the analysis; every analysis reports what is missing.

7 References

Griffing, B. (1956). Concept of general and specific combining ability in relation to diallel crossing systems. *Australian Journal of Biological Sciences*, 9(4), 463–493.

Hayman, B. I. (1954). The theory and analysis of diallel crosses. *Genetics*, 39(6), 789–809.

Hayman, B. I. (1954). The analysis of variance of diallel tables. *Biometrics*, 10(2), 235–244.