

The Genetical Expectations of Doubled Haploid Lines Derived from Different Filial Generations

J.W. Snape and E. Simpson

Plant Breeding Institute, Trumpington, Cambridge (England)

Summary. The genetical expectations of the means, variances and covariances of populations of doubled haploid lines derived from F_1 , F_2 , F_3 and intermated F_2 (S_3) generations are presented. These expectations are identical, regardless of genetical architecture, providing there is no linkage disequilibrium. In the presence of linkage disequilibrium differences will occur whose magnitude and direction will depend on the degree of disequilibrium, recombination frequency and the presence or absence of epistasis.

Data from an experiment to detect linkage disequilibrium in a cross between two spring barley varieties are presented. This involved a comparison of means, variances and covariances of doubled haploid populations derived from the F_1 and F_2 generations using the *H. bulbosum* system. Linkage disequilibrium was detected for important agronomic characters and the effect of this disequilibrium on the choice of generation for doubled haploid production is discussed.

Key words: Doubled haploids – Genetical expectations – Linkage disequilibrium – Barley

Introduction

The production of high frequencies of doubled haploid lines provides the most rapid technique available for the advancement of breeding lines to homozygosity with corresponding advantages in selection efficiency and purity for multiplication (Griffing 1975; Reinbergs et al. 1976). Doubled haploids can also be used to identify the potential frequencies of transgressive segregants in different crosses (Reinbergs et al. 1976; Simpson and Snape 1979) and provide a rapid method for advancing recurrent selection programmes (Griffing 1975; Choo and Reinbergs 1979).

In cereals the most successful system of doubled haploid

production has been the exploitation of chromosome elimination in *Hordeum vulgare* × *Hordeum bulbosum* crosses, and this system forms an integral part of barley breeding programmes in many countries (Kasha and Reinbergs 1980). However, most of these programmes are based on crossing *H. bulbosum* to the F_1 between barley varieties. Thus, the haploid lines produced are products of meiosis in the F_1 so that haploid segregants are the result of only one round of recombination between the parental genomes. If linkage between genes controlling agronomic characters is an important component of the genetical architecture then the complete spectrum of genetic variation potentially available in a cross may not be realised. In this situation it may be more beneficial to produce haploids from more advanced generations which have undergone further rounds of recombination rather than the F_1 .

The present paper examines the theoretical effects of linkage on the means, variances and co-variances between characters, of doubled haploid lines derived from different early filial generations, namely, the F_1 , F_2 , F_3 and intermated F_2 (S_3) generations. Tests for the presence of linkage disequilibrium and epistasis are presented together with results from an experiment designed to detect the effects of linkage disequilibrium on important agronomic characters in a cross of two spring barley varieties.

Theory

In comparing the means and variances of populations of doubled haploid (DH) lines derived from F_1 , F_2 , F_3 and S_3 generations it is assumed that no selection of gametes or genotypes occurs during the production of the haploids. Furthermore, the parental individuals of the F_2 , F_3 and S_3 generations are assumed to be a random sample of genotypes from those generations.

1 Expectations of Means

In the absence of linkage disequilibrium the frequency of any gamete in the F_1 , F_2 , F_3 and S_3 generations is the same, namely, $(\frac{1}{2})^k$ where k is the number of segregating loci. Thus, the frequency of each homozygous genotype in any DH population is $(\frac{1}{2})^k$ and consequently the expected means of the populations are the same. The theoretical expectations of these are defined by Mather and Jinks (1971) whose terminology is used here, as Van der Veen's $F \propto$ metric, m .

The mid-parent mean, however, will differ if epistatic effects are present and its expectation is $m + \{i\}$, where $\{i\}$ is the balanced sum of the homozygous \times homozygous interactions (Mather and Jinks 1971). A test for epistasis in a cross is thus provided by a comparison of the mid-parental mean with that of a DH population.

If linkage between genes is present then the genotype frequencies in the gametes of the different parental generations will differ as a consequence of the number of rounds of recombination in their derivation. This will result in different genotypic frequencies in the different DH populations. However, this will only cause differences between the means of the populations if epistasis between the linked loci is present (Mather and Jinks 1971).

Given a recombination frequency between any two linked loci of p and assuming for simplicity only digenic interactions then the expectations of the means of the DH populations become:

$$\text{Mean } F_1 \text{ DH} = m + \frac{\text{coupling (C)} - \text{repulsion (R)}}{\Sigma (1 - 2p)} i$$

$$\text{Mean } F_2 \text{ DH} = m + \frac{C}{-R} \Sigma (1 - p)(1 - 2p) i$$

$$\text{Mean } F_3 \text{ DH} = m + \frac{C}{-R} \Sigma \frac{1}{2} (1 - 2p)(2 - 3p + 2p^2) i$$

$$\text{Mean } S_3 \text{ DH} = m + \frac{C}{-R} \Sigma (1 - p)^2 (1 - 2p) i$$

where i is the homozygous \times homozygous epistatic effect between any pair of linked loci.

The quantitative effect of linkage on these means will depend on the degree of linkage disequilibrium, the average recombination frequency and the type of epistasis – whether complementary or duplicate (Mather 1967). On the one hand net coupling linkage with complementary epistasis or net repulsion linkage with duplicate epistasis will result in a ranking of means of $F_1 \text{ DH} > F_2 \text{ DH} > F_3 \text{ DH} > S_3 \text{ DH}$ for all recombination frequencies. Alternatively for situations of coupling linkage with duplicate epistasis or repulsion linkage with complementary epistasis, means $F_1 \text{ DH} < F_2 \text{ DH} < F_3 \text{ DH} < S_3 \text{ DH}$. The magnitude of the differences will depend on the recombination frequency

and the biggest differences will be found at intermediate rather than extreme frequencies. A comparison of the means of the $F_1 \text{ DH}$ and $S_3 \text{ DH}$ populations will provide the most sensitive test for linked epistatic effects whilst a comparison of the $F_2 \text{ DH}$ and $F_3 \text{ DH}$ the least sensitive.

With situations of many linked loci both repulsion and coupling linkage may be present and only a net effect can be detected. In certain situations balancing effects may result in epistasis not being detected even when it is present.

2 Expectations of Variances

(a) No Linkage Disequilibrium

In the absence of linkage disequilibrium the expectations of the total variances of the different DH populations, as with the means, are the same, namely $D + I$, where D is the additive variance and I the epistatic variance as defined by Mather and Jinks (1971). However, in the F_2 , F_3 and S_3 derived DH populations a hierarchical family structure can be obtained where more than one doubled haploid line can be derived from each parental individual; and these parental individuals, in the case of the F_3 and S_3 generations can be assigned to families. In this situation the total variance between doubled haploid lines can be partitioned into different effects (Table 1). This partitioning allows a separation and estimation of the additive and epistatic variances.

(b) Linkage Disequilibrium

Linkage disequilibrium will differentially affect the variances of the DH populations irrespective of the contribution of epistasis because of the covariance between individual additive effects caused by the different degrees of linkage disequilibrium. Considering all sources of variation

Table 1. Partitioning of double haploid variances

Statistic	Parental Generation		
	F_2	F_3	S_3
Average variance between DH lines within parental individuals	$\frac{1}{2}D + \frac{3}{4}I$	$\frac{1}{4}D + \frac{7}{16}I$	$\frac{1}{2}D + \frac{3}{4}I$
Average variance between means of parental individuals within families	–	$\frac{1}{4}D + \frac{5}{16}I$	$\frac{1}{4}D + \frac{11}{64}I$
Variance of Means of families	$\frac{1}{2}D + \frac{1}{4}I$	$\frac{1}{2}D + \frac{1}{4}I$	$\frac{1}{4}D + \frac{5}{64}I$
Total variance between DH lines	$D + I$	$D + I$	$D + I$

the theoretical expectations of the total variances of the DH populations are – (see also Jinks and Pooni 1981)

$$VF_1DH = D + \Sigma\{4p(1-p)^2\}i^2 + \frac{C}{-R} \Sigma\{2(1-2p)\}d_a d_b$$

$$VF_2DH = D + \Sigma\{1-2p\}^2(1-p)^2i^2 + \frac{C}{-R} \Sigma\{2(1-2p)(1-p)\}d_a d_b$$

$$VF_3DH = D + \Sigma\{1 - \frac{1}{4}(1-2p)^2(2-3p+2p^2)\}i^2 + \frac{C}{-R} \Sigma\{(1-2p)(2-3p+2p^2)\}d_a d_b$$

$$VS_3DH = D + \Sigma\{1 - (1-2p)^2(1-p)^4\}i^2 + \frac{C}{-R} \Sigma\{2(1-2p)(1-p)^2\}d_a d_b$$

where $d_a d_b$ is the product of the additive effects of two loci.

With no epistatic effects between linked loci a preponderance of coupling linkages will result in $VF_1DH > VF_2DH > VF_3DH > VS_3DH$; and with repulsion linkages of $VF_1DH < VF_2DH < VF_3DH < VS_3DH$. The magnitudes of differences will vary with recombination frequency with bigger differences at intermediate rather than extreme frequencies.

The contribution of epistasis between linked loci to differences between the variances acts in the same direction for both linkage phases resulting in $VF_1DH < VF_2DH < VF_3DH < VS_3DH$. Thus a combination of epistatic and non epistatic linkages will reinforce one another and increase the differences between the generations when linkages are predominantly in repulsion and act in opposition and decrease the differences when in coupling.

A comparison of the total variances of the different generations provides a test of linkage disequilibrium. When family structure is obtained the separate components of variance need to be estimated from the analysis of variance and summed to provide the total variance of that generation. However a more satisfactory test is to grow only one doubled haploid line from each parental family so that the between family mean square is a direct estimate of the total generation variance. The biggest differences are expected between F_1 and S_3 derived DH populations, and a comparison of these provides the most sensitive test of linkage disequilibrium; with a comparison of F_2 and F_3 derived populations the least sensitive.

These differences in variance are caused by changes in the relative proportions of extreme to intermediate genotypes in the DH populations. Any increase in recombination with repulsion linkages will result in a greater proportion of extreme genotypes and vice versa for coupling linkages. Such changes can affect the limits to selection in a cross and the probabilities of transgressive segregants (Jinks and Pooni 1981).

3 Covariances and Correlations between Characters

Genetic correlations between characters are due either to the pleiotropic actions of genes or to the linkage of genes controlling separate characters. If linkages are involved, correlations between characters in the F_1 , F_2 , F_3 and S_3 derived DH populations may differ as a result of different frequencies of recombinant genotypes.

The genetical contribution of linked genes to the total covariance between two characters in the different populations are:

$$F_1DH + \frac{C}{-R} \Sigma (1-2p)d_a d_b$$

$$F_2DH + \frac{C}{-R} \Sigma \{(1-p)(1-2p)\}d_a d_b$$

$$F_3DH + \frac{C}{-R} \Sigma \{\frac{1}{2}(1-2p)(2-3p+2p^2)\}d_a d_b$$

$$S_3DH + \frac{C}{-R} \Sigma \{(1-p)^2(1-2p)\}d_a d_b$$

where d_a , d_b are the respective additive effects of the loci for the two characters.

Coupling linkages generate a positive covariance between characters and repulsion linkages a negative covariance. Differences in recombination will result in differences in covariances over the different DH populations. The direction of the difference will depend on the relative contributions of linkage and pleiotropy to the total covariance although the covariances will always rank in the order of $Cov F_1DH$, $Cov F_2DH$, $Cov F_3DH$, $Cov S_3DH$. If, for example, a positive correlation between characters is due entirely to coupling linkages, then recombination will decrease the genetic covariance differentially in the different DH populations resulting in $\hat{r} F_1DH > \hat{r} F_2DH > \hat{r} F_3DH > \hat{r} S_3DH$.

Tests for the involvement of linkage in the correlation between characters can be carried out by comparing the total covariances and correlations in the different DH populations. However, although statistically it is easier to compare correlation coefficients than covariances, correlation coefficients are ratios of covariances to variances and differences between DH populations could be due to changes in the variances rather than the covariance. Thus when linkages for individual characters are present total covariances rather than correlation coefficients should be tested.

Example

Materials and Methods

The example concerns an experiment containing doubled haploid lines of barley, developed using the *H. bulbosum* system, from the

cross between the spring varieties 'Vada' and 'Sultan'. Two different populations of DH lines were available for a test for linkage disequilibrium. First a population of 54 lines developed from the F_1 and secondly a population of 48 lines derived from a random sample of F_2 individuals where each line was derived from a different F_2 plant.

These two DH populations, together with the parental varieties, were grown in a randomised field experiment consisting of four replications. In each replicate, each DH line was represented by a single plot of eleven plants and each parent by four such plots. During development and after harvest a range of agronomic characters were measured on each individual plot.

2 Results

(a) Means

Genetical variation between DH lines was detected for seven of the characters measured. Thus, the parents differ for alleles controlling all these characters although they only differed in mean performance for three characters, with 'Vada' being significantly earlier and with a larger 250 grain weight and ear weight than 'Sultan' (Table 2). Comparisons between the means of the DH populations and the mid-parental value revealed epistasis for three characters, namely grain number/ear, grain weight/ear and ear number/plant (Table 2). For grain number/ear at least part of the epistasis is attributable to interactions between linked loci

since the F_2 DH mean was significantly greater than the F_1 DH mean. Given the similarity of parental performance this suggests that linkages for this character are of the repulsion type with a duplicate type of epistasis. For the other characters no differences between the DH populations were apparent so that linked epistatic effects appear to be absent, or there is a balance of opposing effects.

(b) Variances

The variances between lines within each DH population, which estimate the additive genetical variance for each character, are presented in Table 3. The significance of differences between the populations were tested using a Bartlett test on the between line mean squares extracted from the analysis of variance for each population.

No significant differences between the populations were detected for any character although quite large percentage changes in the variances were apparent. This reflects the insensitivity of the present experiment because of relatively low numbers of lines in each population, since with the degrees of freedom available at least at 75% change in variance is necessary for 5% significance. Nevertheless the changes in variance are suggestive of linkage disequilibrium particularly for ear emergence time, height, 250 grain weight, grain number/ear and ear number/plant where over

Table 2. Mean performance of parental and doubled haploid populations

	Ear emergence time	Height (cms)	Grain weight/plant (gms)	250 grain weight (gms)	Grain number/ear	Grain weight/ear (gms)	Ear number/plant
'Vada'	4.56	101.3	12.6	14.1	28.3	1.59	7.95
'Sultan'	7.50***	100.8	12.4	13.1***	27.7	1.45***	8.63
Mid Parent	6.03	101.1	12.5	13.6	28.0	1.52	8.28
F_1 DH	5.75	99.8	12.8	13.3	26.8*	1.43	9.04
F_2 DH	6.08	99.9	12.8	13.0	27.5*	1.43	9.05
Significance of difference between mid-parent and DH means.	F_1 DH	N.S.	N.S.	N.S.	*	**	**
	F_2 DH	N.S.	N.S.	N.S.	N.S.	**	**

Probabilities: * = 0.05 - 0.01, ** = 0.01 - 0.001, *** \leq 0.001

Table 3. Variances of doubled haploid populations

	Ear emergence time	Height	Grain weight/plant	250 grain weight	Grain no./ear	Grain weight/ear	Ear no./plant
Variance F_1 DH lines	2.082	14.10	2.431	1.262	1.992	0.0187	0.615
Variance F_2 DH lines	3.023	18.39	1.922	0.831	2.472	0.0166	0.785
Probability of difference	0.27	0.58	0.54	0.25	0.62	0.80	0.44
% change in variance	+45	+30	-20	-34	+24	-11	+28

20% changes in variances were observed. For ear emergence time, height, grain number/ear and ear number/plant the variances of the F₂DH population were larger than the F₁DH population and this reflects higher proportions of extreme genotypes in the F₂DH population at both ends of the phenotypic distribution. This suggests that net repulsion linkages are present in the genetical control of these characters. With grain number/ear this explanation reinforces the conclusion concerning the means of the populations. On the other hand with 250 grain weight and grain weight/plant a decrease in variance is apparent in the F₂DH population, caused by a decrease in the proportion of extreme types, a situation characteristic of the break up of coupling linkages.

(c) Correlations and Covariances

The correlations between characters in the DH populations, estimated on family means, are shown in Table 4. Significant correlations were exhibited for sixteen of the character combinations examined. For twelve of these, similar estimates were obtained in both populations suggesting that, in the present experiment, these relationships are due to the pleiotropic actions of genes. For four of the character combinations, however, significant differences between the DH populations were detected and in all cases these were a consequence of a significant change in the genetic covariance and therefore ascribable to the effects

of linkage disequilibrium. With respect to the relationship between grain weight/plant and grain number/ear a significant positive correlation in the F₁DH population was non-significant in the F₂DH population. This suggests the presence of net coupling linkages in the F₁ which are broken down by further recombination in the F₂.

For the three other combinations, ear emergence time with grain number/ear and with grain weight/ear, and grain number/ear with ear number/plant non-significant correlations were obtained from the F₁DH population but highly significant ones from the F₂DH population. This suggests that linkages between genes in the parents have been built up to oppose effects of pleiotropic genes so that only when the linkage component is broken down by recombination is the pleiotropic relationship phenotypically expressed. For example, genes for early flowering appear to be pleiotropically associated with low grain number/ear. Thus to obtain early flowering with high grain number, characterised by the parent variety 'Vada', has necessitated the build up of linkage relationships which oppose the pleiotropic effect. Such linkages would have been built up by selection during the development of the parental varieties for early flowering, higher yielding genotypes. Similarly the undesirable pleiotropic relationships, as far as breeding objectives are concerned, between early flowering and low grain weight/ear and between high grain number/ear and low ear number/plant appear to have been counteracted by a build up of opposing linkages.

Table 4. Correlation matrix for F₁DH and F₂DH populations

		Height	Grain weight/ plant (gms)	250 grain weight (gms)	Grain number/ear	Grain weight/ ear (gms)	Ear number/ plant
Ear Emergence time (days)	F ₁ DH	-0.05	-0.48***	-0.12	-0.05	-0.16	-0.46***
	F ₂ DH	0.18	-0.35*	-0.12	0.62*** ***	0.30* *	-0.55***
Height (cms)	F ₁ DH	-	0.29*	0.27*	0.42**	0.49***	-0.06
	F ₂ DH		0.28	0.46***	0.40**	0.59***	-0.17
Grain weight/plant (gms)	F ₁ DH		-	0.42**	0.34*	0.58***	0.71***
	F ₂ DH			0.61***	-0.12	0.38**	0.69***
250 grain weight (gms)	F ₁ DH				-0.09	0.79***	-0.22
	F ₂ DH				0.03	0.78***	-0.02
Grain number/ear	F ₁ DH					0.54***	-0.01
	F ₂ DH				-	0.65***	-0.61*** ***
Grain weight/ear (gms)	F ₁ DH					-	-0.16
	F ₂ DH						-0.40**
Ear number/plant	F ₁ DH						-
	F ₂ DH						-

Probabilities: * = 0.05 - 0.01, ** = 0.01 - 0.001, *** < 0.001

Discussion

The importance of linkage in the genetic architecture of agronomic characters can be estimated using appropriate experimentation such as that used here. In the present experiment linkage appears to be a component of the genetic architecture of important agronomic characters of barley both for genes controlling individual characters and for genes affecting the relationship between different characters. Thus it is necessary to examine the effect of this linkage disequilibrium on possible breeding strategies using doubled haploid lines.

For most characters the plant breeder is concerned with transgressing the parental range and this requires the dispersion of increasing genes between the parents. If linkage is important this most likely requires the break up of repulsion linkages. In this situation it is more advantageous to develop doubled haploid lines from the F_2 or from intermated F_2 's rather than the F_1 even though this involves a loss in the time saving element of using a haploid system. In the present example the generation delay has increased variation for ear emergence time, height, grain number/ear and ear number/plant. However, there is also a risk element in that the system can break down desirable coupling linkages, presumably built up by previous recombination and selection. Thus in the present experiment recombination has probably resulted in a decrease in the variance for grain weight/plant and 250 grain weight.

The theoretical expectations for the generation variances of the DH populations show that the greatest difference in the presence of linkage occurs between the F_1 DH and the S_3 DH generations, since the S_3 population allows greater scope for recombination than the F_2 . However, in practical terms any increase in variance in the S_3 DH population is unlikely to compensate for the large amount of extra effort required to develop the S_3 generation which would, in barley, require an extensive hybridisation programme; and also cause a further generation delay. Similarly, a time delay until the F_3 generation will only produce a small change in variance. Thus from a practical and a theoretical viewpoint the F_2 would appear to be the best generation for producing doubled haploid lines if linkage disequilibrium is important. A further advantage of using the F_2 is that some selection can be carried out between individuals before they are used for haploid production, for example, in barley for height, cold tolerance and specific disease resistance genes. This also eliminates a disadvantage of the F_1 system, which by producing a completely random sample of genotypes inevitably produces a high frequency of agronomically undesirable lines, which is wasteful in terms of the resources necessary for a doubled haploid system.

A comparison of DH populations can provide a useful procedure for understanding the genetical relationships

between different characters. This can indicate which is the best breeding strategy to follow in terms of plant ideotype. For example, changes in correlations between generations can indicate whether the frequently found compensatory relationships between yield components can be broken. In the present example, the data for the 'Vada' × 'Sultan' cross indicates that a positive relationship between late flowering and certain yield components are, to a large extent, pleiotropic but can be counteracted by building up appropriate linkages.

The experimental results here are based on only one cross and examination of other crosses is required to evaluate the general extent of linkage disequilibrium for quantitative characters in barley. Nevertheless the results suggest that linkage must be considered in utilising doubled haploid lines in a barley breeding programme. Further, if linkage disequilibrium is important then on theoretical and practical grounds the F_2 generation would appear to be the most suitable parental material for haploid production.

Acknowledgement

Thanks are due to Miss B.B. Parker for technical assistance and to Prof. J.L. Jinks and Dr. A.J. Wright for their helpful comments on the manuscript.

Literature

- Choo, T.M.; Reinbergs, E. (1979): Doubled haploids for estimating genetic variances in presence of linkage and gene association. *Theor. Appl. Genet.* 55, 129-132
- Griffing, B. (1975): Efficiency changes due to use of doubled-haploids in recurrent selection methods. *Theor. Appl. Genet.* 46, 367-386
- Jinks, J.L.; Pooni, H.S. (1981): Properties of pure-breeding lines produced by dihaploidy, single seed descent and pedigree breeding. *Heredity* (in press)
- Kasha, K.J.; Reinbergs, E. (1980): Achievements with haploids in barley research and breeding. In: *The Plant Genome*; (eds. Davies, D.R.; Hopwood, D.A.), pp. 215-230. Norwich, England: The John Innes Charity
- Mather, K. (1967): Complementary and duplicate gene interactions in biometrical genetics. *Heredity* 22, 97-103
- Mather, K.; Jinks, J.L. (1971): *Biometrical Genetics*. London: Chapman and Hall
- Reinbergs, E.; Park, S.J.; Song, L.S.P. (1976): Early identification of superior barley crosses by the doubled haploid technique. *Z. pflanzenzücht.* 76, 215-224
- Simpson, E.; Snape, J.W. (1979): Cross prediction for yield using doubled haploid lines. *Barley Genetic Newsl.* 9, 95-97

Dr. J.W. Snape
Mr. E. Simpson
Plant Breeding Institute
Maris Lane, Trumpington
Cambridge CB2 2LQ (England)

Received February 20, 1981
Communicated by R. Riley