

WHEAT-BUNT FIELD TRIALS, II

GEORGE A. BAKER AND FRED N. BRIGGS

UNIVERSITY OF CALIFORNIA, DAVIS

1. Introduction

In the study of inheritance of resistance to bunt, *Tilletia tritici*, in wheat it is necessary to grow F_3 progenies from F_2 plants in order to determine the genotype of the latter. This is necessary because all susceptible plants in a population usually do not show evidence of the disease under prevailing conditions. Such susceptible plants merely escape infection and do not differ genetically from those which exhibit bunt. Some trials are further complicated by the fact that some heterozygous and occasionally some homozygous-resistant plants become diseased.

Briggs [2], [3], [4]¹ has shown that there is little difficulty in distinguishing the three genotypes (homozygous-susceptible, heterozygous, and homozygous-resistant) in crosses when the parents differ by a single dominant factor for resistance to bunt if F_3 rows of 50 plants each are grown. Even with 50 plants per row, however, there are usually a few rows which cannot be positively classified. The proportion of infected plants in both heterozygous and homozygous-susceptible rows of 50 plants usually exhibit a range of about 0.50. When parents differ by two or more factors for resistance, as reported by Briggs [2], the overlapping of the proportions of infected plants in rows of 50 belonging to the different genotypes makes it difficult, if not impossible, to recognize the rows belonging to the various genotypes. Thus, genetical analysis under these conditions may be very unsatisfactory.

For these reasons it is important to consider the question of whether the variability in bunt infection among rows of a given genotype can be accounted for on the basis of sampling error, and whether this variability would decrease according to random-sampling expectations as the number of plants per row increased. A decrease in variability ensures that the limits of the three genotypes are more certainly defined when a single-factor difference exists. If the decrease in variability is sufficient, a greater number of genotypes can be recognized when more factors are involved. An alternative approach to analyzing the more complicated crosses on the basis of classifying rows with respect to genotype would be to set up a theoretical distribution, using the necessary factors with the proper mean effects and variabilities, to fit the observed distribution of proportions of infected plants. In order to set up such a theoretical distribution, it is necessary properly to evaluate the expected variability of proportions of infected plants of each genotype for rows of a given length or number of plants.

Salmon [6] has shown that the theoretical binomial sampling error will not account for the variation in bunt infection among wheat varieties. He found it necessary to increase this error by a factor a , which varied for different environmental conditions and levels of infection.

¹ Boldface numbers in brackets refer to references at the end of the paper (see p. 491).

By using rows of about 340 plants for the F_3 population, Stanford [7] was able to make a satisfactory genetical analysis of a cross between Rio and Baart wheats which had not been possible when only 50 plants per row had been grown.

Baker and Briggs [1] indicated briefly the development of a mathematical model of wheat bunt field trials and made detailed application of this model to the interpretation of one cross between a resistant (Banner Berkeley) and a susceptible (Baart) wheat. In the present paper a fuller and modified development of the mathematical model is given which leads to a slight correction in two formulas, (6) and (7), of their earlier paper [1]. Further, the random-sampling variations of some of the estimates are discussed in detail and the sampling errors of other quantities are indicated. Also, additional applications to observed data are made.

2. Homozygous-susceptible plants only can be infected

We shall assume for the moment that a row of n plants is homogeneous with respect to the probability of infection, that the probability, p , of infection of homozygous-susceptible plants does not vary from row to row, and that only homozygous-susceptible plants can be infected. Then the total probability of k infected plants in heterozygous rows of n plants each is

$$(2.1) \quad \sum_{m=k}^{m=n} \frac{n!}{m!(n-m)!} (1/4)^m (3/4)^{n-m} \frac{m!}{k!(m-k)!} p^k (1-p)^{m-k},$$

where m is the number of susceptible plants in a row and the segregation is 3 resistant plants to 1 susceptible. The expression (2.1) reduces to the binomial distribution

$$(2.2) \quad [(1 - p/4) + p/4]^n.$$

If the segregation is not 3:1 but the proportion of susceptible plants is α instead of 1/4, then, corresponding to (2.2), we have

$$(2.2') \quad [(1 - \alpha p) + \alpha p]^n.$$

Denote the mean of the proportions of infected plants in rows of n by M_1' and the second moment of the proportions calculated about zero as an origin by M_2' . Then from (2.2) we have

$$(2.3) \quad \begin{cases} M_1' = p/4 \\ M_2' = \frac{p(1-p)}{4n} + p^2 \left(\frac{3}{16n} + \frac{1}{16} \right). \end{cases}$$

Now consider that we can have a range of values of p 's (probability of susceptible plants being infected) and that the frequency distribution of the p 's is given by

$$(2.4) \quad \varphi(p)dp, \quad 0 \leq p \leq 1,$$

with mean \bar{p} and variance V . Let M and Σ^2 be the mean and variance of the proportions of infected plants under the extended conditions. Then

$$(2.5) \quad \begin{cases} M = \bar{p}/4 \\ \Sigma^2 = \frac{\bar{p}(4 - \bar{p})}{16n} + \frac{(n - 1)V}{16n} \end{cases}$$

The ratio a^2 of Σ^2 to the theoretical binomial variance is

$$(2.6) \quad a^2 = 1 + \frac{(n - 1)V}{\bar{p}(4 - \bar{p})},$$

and hence is always greater than 1. The expression corresponding to (2.6) for homozygous-susceptible rows is

$$(2.6') \quad (a')^2 = 1 + \frac{(n - 1)V}{\bar{p}(1 - \bar{p})}.$$

$(a')^2$ differs much more from 1 than does a^2 .

The value of $(a')^2$ depends on three quantities, n , V , and \bar{p} , and tends to increase with the length of the row, n . However, V probably tends to decrease as the length of the row increases. These remarks are illustrated in the accompanying table.

EFFECT OF LENGTH OF ROW, n , ON $(a')^2$ AND V FOR HOMOZYGOUS-SUSCEPTIBLE BAART PLANTS

n	$(a')^2$	V
50*	1.31	0.003328
100*	1.56	0.003323
150*	1.90	0.004015
342†	2.39	0.001198

* See Baker and Briggs [1].
 † See Stanford [7].

The quantity V (variance of p in rows of n) can be interpreted in terms of a' , the ratio of the observed standard deviation to the theoretical binomial standard deviation. That is,

$$(2.7) \quad \frac{(a')^2 \bar{p}(1 - \bar{p})}{n} = \frac{\bar{p}(1 - \bar{p})}{n} + \frac{n - 1}{n} V,$$

or

$$(2.8) \quad V = \frac{[(a')^2 - 1] \bar{p}(1 - \bar{p})}{n - 1}.$$

Variation of probability of infection within the heterozygous rows may cause a slight decrease in over-all variability, as indicated in a letter from Professor J. Neyman, Statistical Laboratory, University of California, Berkeley.

3. Heterozygous and homozygous-resistant plants can be infected

Suppose that all three types of plants can be infected and let p_1 , p_2 , p_3 be the probabilities of infection for the (rr) homozygous-susceptible, (Rr) heterozygous, and (RR) homozygous-resistant plants respectively. The p 's are assumed to hold uniformly for a row of n plants. Then, corresponding to (2.2), we have

$$(3.1) \quad \left[\left(1 - \frac{p_1 + 2p_2 + p_3}{4} \right) + \frac{p_1 + 2p_2 + p_3}{4} \right]^n.$$

Corresponding to (2.3) we have

$$(3.2) \quad \begin{cases} M_1' = \frac{p_1 + 2p_2 + p_3}{4} \\ M_2' = \frac{p_1 + 2p_2 + p_3}{4n} + \frac{n-1}{16n} (p_1^2 + 4p_2^2 + p_3^2 + 4p_1p_2 + 2p_1p_3 + 4p_2p_3). \end{cases}$$

Let $\varphi(p_1, p_2, p_3)$ be the distribution function of p_1, p_2, p_3 with means $\bar{p}_1, \bar{p}_2, \bar{p}_3$; variances V_1, V_2, V_3 ; and correlations R_{12}, R_{13}, R_{23} . Then, corresponding to (2.5), we have

$$(3.3) \quad \begin{cases} M = \frac{\bar{p}_1 + 2\bar{p}_2 + \bar{p}_3}{4} \\ \Sigma^2 = \frac{1}{16n} [\bar{p}_1(4 - \bar{p}_1) + 4\bar{p}_2(2 - \bar{p}_2) + \bar{p}_3(4 - \bar{p}_3) - 4\bar{p}_1\bar{p}_2 - 2\bar{p}_1\bar{p}_3 - 4\bar{p}_2\bar{p}_3] \\ \quad + \frac{n-1}{16n} [V_1 + 4V_2 + V_3 + 4(V_1V_2)^{1/2}R_{12} + 2(V_1V_3)^{1/2}R_{13} + 4(V_2V_3)^{1/2}R_{23}] \end{cases}$$

It is noted that $M, \Sigma^2, \bar{p}_1, \bar{p}_3, V_1$, and V_3 can be observed or computed from observation and that R_{12}, R_{13} , and R_{23} can be assumed equal to 1; hence \bar{p}_2 and V_2 can be computed. In some cases when only environmental conditions act to increase variability, $V_1 = V_2 = V_3$.

4. Modifying factors present

If modifying factors are present, the situation is quite different as far as the equality of the V 's is concerned. An example is the Rio \times Baart cross reported by Stanford [7]. We have

$$\begin{array}{lll} p_1 = 0.9045 & a_1^2 = 5.73 & V_1 = 0.0011982 \\ p_2 = 0.5294 & a_2^2 = 30.05 & V_2 = 0.0212241 \\ p_3 = 0.0929 & a_3^2 = 10.87 & V_3 = 0.0024391 \\ M = 0.5141 & \Sigma^2 = 0.0095116 & \\ a_1 = 2.39 & a_2 = 5.48 & a_3 = 3.30 \\ n = 342 & & \end{array}$$

These values of a_1, a_2 , and a_3 indicate that the heterozygous plants are relatively more affected by modifiers than are the homozygous-resistant plants.

5. Random-sampling variation for homozygous-susceptible plants

We shall consider the case of the random-sampling variance of \bar{p} determined from N rows of n homozygous-susceptible plants. Under these conditions the variance of the sampled population of proportions is

$$(5.1) \quad \frac{1}{n} [\bar{p}(1 - \bar{p}) + (n - 1)V] ,$$

and hence the variance of our estimate of \bar{p} from N rows is

$$(5.2) \quad \frac{1}{Nn} [\bar{p}(1 - \bar{p}) + (n - 1)V] .$$

Further, the distribution of these estimates of \bar{p} can be assumed to be approximately normal when N is as large as it is in these experiments.

In order to obtain the variance of the standard deviation of the proportions of n homozygous-susceptible plants grown in N rows we make use of several formulas available in Kendall [5], who gives, in particular,

$$(5.3) \quad \frac{M_4 - M_2^2}{4NM_2}$$

for the variance of the standard deviation where the M 's are the indicated moments of the sampled population about its mean but estimated from the sample. We have, under the conditions stated and the further assumption that $\varphi(p)$ is normal,

$$(5.4) \quad \left\{ \begin{array}{l} M_1' = \bar{p} \\ M_2 = \frac{1}{n} [\bar{p}(1 - \bar{p}) + (n - 1)V] \\ M_3 = \frac{1}{n^2} [\bar{p}\{1 - 6(n - 1)V\} - 3\bar{p}^2 + 2\bar{p}^3 + 3(n - 1)V] \\ M_4 = \frac{1}{n^3} [\bar{p}\{1 + 6(n - 1)(n - 6)V\} + \bar{p}^2\{(3n - 7) - 6(n - 1)(n - 6)V\} \\ \quad - 6(n - 2)\bar{p}^3 + 3(n - 2)\bar{p}^4 + 7(n - 1)V + 3(n - 1)(n - 2)(n - 3)V^2] . \end{array} \right.$$

If we use (5.3) and (5.4) to compute standard errors for the standard deviations of the proportions in the homozygous rows of the authors' earlier paper [1], we find that the standard errors are in the neighborhood of 0.009.

In order to discuss the sampling variation in the heterozygous rows we need to replace p by $p/4$ if only susceptible plants can be infected and by $(p_1 + 2p_2 + p_3)/4$ if all three types of plants can be infected in the formulas leading to (5.4). In the latter case we need to consider more fully the probable

characteristics of the distribution function $\varphi(p_1, p_2, p_3)$ and the possibility of expressing the higher moments in terms of the means and variances of the three types of plants. This of course is possible if φ is assumed normal and the correlations assumed equal to 1.

6. Summary

A mathematical model of wheat bunt field trials when only one factor for resistance is involved has been given for the general case in which all three genotypes can become infected. The analysis of variability developed in this connection is also fundamental in interpreting complex hybrids when a row or family may be made up of different genotypes which have different probabilities of infection.

If three or two genotypes can be infected, the mean and spread of the heterozygous plants cannot be observed, but these can be calculated because the means and variances of the proportions of infected plants for the homozygous and segregating rows can be observed and it can be assumed that the probabilities of infection for different types of plants are perfectly correlated as between rows.

In general the variance for the segregating rows is much closer to the binomial variance than is the variance for the homozygous rows. This apparent inconsistency does not result from different a values for the genotypes of the plants present but is merely a result of the mathematics of the situation. If a values are different, then it is likely that the increased variability of some genotypes is due to the presence of modifying factors.

The a values when no modifying factors are present depend on the length of rows, the spread in infectivity levels due to environment, and the general level of infection.

Formulas for the variances of the means and standard deviations of the proportions of infected plants per row of a given length are obtained in some cases and indicated in others.

LITERATURE

1. BAKER, GEORGE A., and FRED N. BRIGGS. "Wheat Bunt Field Trials," *Jour. Amer. Soc. Agron.*, Vol. 37 (1945), pp. 127-133.
2. BRIGGS, FRED N. "Inheritance of Resistance to Bunt, *Tilletia tritici* (Bjerk.) Winter, in Wheat," *Jour. Agr. Res.*, vol. 32 (1926), pp. 973-990.
3. ———. "Inheritance of Resistance to Bunt, *Tilletia tritici*, in Hybrids of White Federation and Banner Berkeley Wheats," *ibid.*, vol. 42 (1931), pp. 307-313.
4. ———. "A Third Genetic Factor for Resistance to Bunt, *Tilletia tritici*, in Wheat Hybrids," *Jour. Gen.*, vol. 27 (1933), pp. 435-441.
5. KENDALL, M. G. *The Advanced Theory of Statistics*, Vol. I, pp. 117 and 224. J. B. Lippincott Co., 1943. [Also, Charles Griffin and Co., Ltd., London, 1943.]
6. SALMON, S. C. "Generalized Standard Errors for Evaluating Bunt Experiments with Wheat," *Jour. Amer. Soc. Agron.*, vol. 30 (1938), pp. 647-663.
7. STANFORD, ERNEST H. "A New Factor for Resistance to Bunt, *Tilletia tritici*, Linked with the Martin and Turkey Factors," *Jour. Agr. Res.*, vol. 33 (1941), pp. 559-568.