Theor Appl Genet (1988) 76:341–351



# Gene differences in heading date, height, seed weight and seed yield between two pure line varieties of *Triticum aestivum* L.

**C. F. Wehrhahn [1] and G. C. C. Tai [2]**

[1] Resource Ecology, Resource Management Science and Department of Zoology, University of British Columbia, Vancouver, BC V6T 1W5, Canada
[2] Agriculture Canada Research Station, P.O. Box 20280, Fredericton, New Brunswick E3B 4Z7, Canada

Received March 13, 1987; Accepted March 10, 1988
Communicated by A. L. Kahler

**Summary.** Reciprocal sets of homozygous inbred backcross lines were developed by crossing two pure line varieties (Baart 46 and Ramona) of *Triticum aestivum* L., followed by two backcrosses to each of the two parent varieties, and six to eight generations of selfing. Data on each inbred backcross line was obtained from twelve plots (from replications in three years). Five genes were responsible for over 95% of the genetic variation for heading date. These genes had pleiotropic effects on plant height that were proportional to their effects on heading date. Two additional genes had detectable effects on plant height. The genes with a measurable effect on height accounted for 90% of the genetic variation in the Baart 46 genetic background. One gene affected seed weight. In the Ramona background, this gene accounted for 80% of the genetic variation in seed weight and 16% of the genetic variation in seed yield. Two genes, responsible for the earliest and latest heading date classes, had large pleiotropic effects on seed yield. They accounted for 60% of the genetic variation in yield. One gene, with no effect on heading date, caused a detectable reduction in yield of 23% in the Baart 46 inbred backcross lines. This gene had no apparent effect in the Ramona genetic background. Quantitative trait genes are sparsely distributed in the genome: fewer than one in four chromosome arms carries a gene with a detectable effect. Gene effects on quantitative traits are not small and similar. The distribution of 22 gene effects for heading date and height is slightly skewed to the right: as the magnitude of effect increases, the frequency of genes having the effect decreases.

**Key words:** Polygenes – Quantitative trait loci – Inbred backcross lines

## Introduction

Traits that vary continuously are generally thought to be influenced by numerous genes with small and similar effects (Mather and Jinks 1982; Falconer 1981). The state of our knowledge regarding gene number is summarized by Hallauer and Miranda (1981) in their book on the use of quantitative genetics in maize (*Zea mays* L.) breeding where they state: "Gene number for traits such as vigor, health, and productivity are unknown but they must be extremely great in all instances". This may be true in large populations if very nearly monomorphic genes are counted.

Nevertheless, a few genes are known to have relatively large effects on quantitative characters (cf. Robertson 1967; Soller and Brody 1976; Allendorf et al. 1983). Genes with larger effects can be detected in certain kinds of experiments (e.g., Allard 1956; Thoday 1961; Wehrhahn and Allard 1965; Mulitzke and Baker 1985; Kahler and Wehrhahn 1986), and it is possible to determine the proportion of variation due to these genes. The extent to which genes with large effect occur has especially important implications for selection theory where a common assumption is that selection is too weak to change allele frequencies appreciably (Lande 1980).

In this paper we report the results of experiments involving large numbers of inbred backcross lines derived from two homozygous varieties, Baart 46 and Ramona, of *Triticum aestivum* L. The objective of the study was to determine the effects of individual loci on several quantitative characters.

## Materials and methods

Reciprocal sets of inbred backcross lines (IBLs) were developed by crossing the two homozygous varieties Baart 46 and Ramona

to obtain an $F_1$ hybrid. This was followed by two successive generations of backcrossing to each parent. Each line was propagated from a single randomly chosen plant in each of the first four generations of selfing. The pedigree of lines backcrossed twice to Baart 46 (parent 1) and Ramona (parent 2) are denoted by B11(n)0 and B22(n)0, respectively: the 1's represent backcrosses to parent 1, 2's represent backcrosses to parent 2 and (n)0 represents the number, n, of selfed generations.

To produce an IBL with the Baart 46 genetic background, seed from a single B11(4)0 plant was used to produce a row of B11(5)0 plants which was harvested in bulk to produce seed of the B11(6)0 generation. Each B11(8)0 IBL was obtained from one plot of the B11(7)0 generation which, in turn, came from a plot of the B11(6)0 generation. Seed of each Ramona background IBL was obtained in the same way. After seed increase 76 B11(n)0 and 78 B22(n)0 IBLs were available for use in replicated field experiments.

The 76 B11(6)0 Baart 46 genetic background lines and five Baart 46 parent families were included in a $10 \times 10$ repeated simple lattice design with four replicates and seeded on dryland summerfallow at the University of Saskatchewan, Saskatoon, Canada. Additional Ramona and Baart 46 families and hybrids were included to fulfill the requirement of 100 plots per replicate. Each plot consisted of two 5 m rows spaced 0.3 m apart, with a spacing of 0.6 m between plots. The seeding rate was 400 seeds per plot. The 78 B22(6)0 Ramona genetic background lines and five Ramona parent families were included in a similar experiment. The experiments were repeated using B11(7)0, B22(7)0, B11(8)0, and B22(8)0 lines in two subsequent years.

Data were obtained on each plot for the following characters: (1) days to heading – recorded as the number of days after July 1 required for 80% of the first spikes of plants in a plot to emerge completely from the boot; (2) height – average height in cm from the ground to the top of the spike at six random places in a plot; (3) seed weight – weight in g of 200 seeds; (4) yield – weight of seeds in g/plot.

*Methods of data analysis*

The data from each field experiment was analysed by the standard method for the repeated simple lattice design with four replicates (Cochran and Cox 1957). Mean values used in subsequent analyses were adjusted for block effects. The effective error mean square per plot (EMS) for an experiment was calculated by using a formula given by Cochran and Cox (1957, p. 409). The error variance of the mean for an inbred backcross line in any year is, therefore, EMS/4.

*Theory and strategy for detecting genes*

After two backcrosses to the recurrent parent, the probability that a plant is heterozygous is 1/4. Subsequently, after four generations of selfing, the probability that a plant is heterozygous at a locus is $1/(4 \times 16) = 1/64$ and the probability that it is homozygous for the allele of the nonrecurrent (donor) parent is 15/128. The probability that an inbred backcross line is either homogeneous or segregating for an allele from the nonrecurrent parent is, therefore, $q = 17/128 = 0.1328$. The number of IBLs carrying a particular allele from a nonrecurrent parent has a binomial distribution with a mean of nq and a standard deviation of $[nq(1-q)]^{1/2}$. The expected number of lines which are carriers of an allele of a gene from the nonrecurrent parent is $10.1 \pm 3.0$ for the $n = 76$ Baart 46 background IBLs. The expected number of carrier lines for each nonrecurrent allele is $10.4 \pm 3.0$ for the 78 Ramona IBLs.

Because the proportion of lines that carry an allele has a binomial distribution, the arcsin square root transformation of q is nearly normal with a variance of 1/4n (Sokal and Rohlf 1981). The transformation was used to find approximate 95% confidence limits for the number of lines that carry a particular allele from the nonrecurrent parent. The 95% confidence intervals are 5–17 lines. If there are two genes with similar effects, the probability that a line is a carrier of a donor allele of either one, but not both, of the genes is $Q = 2q(1-q) = 0.230$. In this case the 95% confidence limits for the number of lines that carry one of two donor alleles, are 13–28 and 14–29 for Baart 46 and Ramona IBLs, respectively.

**Table 1.** Effective error mean square (per plot), from a repeated simple lattice experiment with four replicates per generation, of each B11(n)0 or B22(n)0 line

| Character | Generation | Error mean square (EMS) |
|---|---|---|
| Days to head (after July 1) | B11(6)0 | 0.644 |
| | B11(7)0 | 0.264 |
| | B11(8)0 | 0.636 |
| | B22(6)0 | 0.360 |
| | B22(7)0 | 0.372 |
| | B22(8)0 | 0.856 |
| Height (cm) | B11(6)0 | 19.69 |
| | B11(7)0 | 16.88 |
| | B11(8)0 | 14.76 |
| | B22(6)0 | 8.52 |
| | B22(7)0 | 8.90 |
| | B22(8)0 | 12.10 |
| Seed weight (g per 200) | B11(6)0 | 0.257 |
| | B11(7)0 | 0.084 |
| | B11(8)0 | 0.274 |
| | B22(6)0 | 0.225 |
| | B22(7)0 | 0.067 |
| | B22(8)0 | 0.200 |
| Seed yield (g per plot) | B11(6)0 | 10,968 |
| | B11(7)0 | 7,964 |
| | B11(8)0 | 9,524 |
| | B22(6)0 | 7,016 |
| | B22(7)0 | 7,272 |
| | B22(8)0 | 7,940 |

The strategy employed to detect genes was the following: first, scatter diagrams of the means of one or two generations versus another generation were plotted; then a 99% probability ellipse centered at the mean for the recurrent parent was constructed; next, starting with the most extreme lines, each cluster of lines was enclosed in a 99% ellipse; finally the number of lines in each ellipse was compared to the confidence intervals for one, two and three genes. When a very few lines were found in a region far from the recurrent parent's mean, these were regarded as carrying alleles of more than one gene from the nonrecurrent parent. Of the possible two dimensional scatter diagrams, the one that most clearly showed discontinuities between clusters of lines is presented.

*Intervals and ellipses enclosing a specified percentage of lines*

If they are identical, the difference, $\bar{d} = \bar{X}_L - \bar{X}_P$, in the mean, $\bar{X}_L$, of an IBL, with four replicates per generation, and the mean, $\bar{X}_P$, of its recurrent parent with five families or 20 replicates per generation, has a variance of $\sigma_{\bar{d}}^2 = \text{EMS}(1/4 + 1/20)$. Because EMS (Table 1) has over 300 df, 99% of lines identical in mean to

their recurrent parent are expected to lie within the interval $\bar{X}_P - 2.57\,\sigma_{\bar{d}} < \bar{X}_L < \bar{X}_P + 2.57\,\sigma_{\bar{d}}$ (Sokal and Rohlf 1981).

For lines identical to their recurrent parent in two generations, the standardized mean differences $y_1 = (\bar{X}_{L1} - \bar{X}_{P1})/\sigma_{\bar{d}1}$ and $y_2 = (\bar{X}_{L2} - \bar{X}_{P2})/\sigma_{\bar{d}2}$ should have independent standard normal distributions. The joint distribution of $y_1$ and $y_2$ will then be

$$dF = \frac{1}{\sqrt{2\pi}} e^{-y_1^2/2}\,dy_1 \times \frac{1}{\sqrt{2\pi}} e^{-y_2^2/2}\,dy_2$$

$$= \frac{1}{2\pi} e^{-(y_1^2+y_2^2)/2}\,dy_1\,dy_2, \quad (1)$$

which is the bivariate normal distribution with zero covariance (Kendall and Stuart 1969).

Transform the equation, $r^2 = y_1^2 + y_2^2$, for a circle of radius r about zero to polar coordinates by letting $y_1 = r\cos\theta$ and $y_2 = r\sin\theta$ (Walters and Wehrhahn 1987). Equation (1) becomes $dF = (1/2\pi)\exp(-r^2/2)\,r\,dr\,d\theta$. The probability that a standardized mean difference has a magnitude greater than $R = \sqrt{y_1^2 + y_2^2}$ without regard to its direction, $\theta$, can be obtained by integrating this equation with respect to $\theta$, from 0 to $2\pi$, and with respect to r, from R to $\infty$. The probability is

$$P(r > R) = \frac{1}{2\pi}\int_0^{2\pi} d\theta \int_R^{\infty} r\,e^{-r^2/2}\,dr$$

$$= \frac{1}{2\pi} \times 2\pi \times (-1) e^{-r^2/2}\Big|_R^{\infty} = e^{-R^2/2}. \quad (2)$$

Conversely, the value of $R^2$ for any specified probability is

$$R^2 = -2\ln P(r > R) = -2\ln[1 - P(r \le R)]. \quad (3)$$

In particular, for ellipses containing 67 and 99 percent of lines identical in mean to their recurrent parent, $R^2 = -2\ln(1-0.67) = 2.22$ and $R^2 = -2\ln(1-0.99) = 9.21$, respectively.

To obtain coordinates of the ellipse, note that $y_2^2 = R^2 - y_1^2$ and $y_2 = \pm[R^2 - y_1^2]^{1/2}$ or

$$\bar{X}_{L2} - \bar{X}_{P2} = \pm\sigma_{\bar{d}2}[R^2 - (\bar{X}_{L1} - \bar{X}_{P1})^2/\sigma_{\bar{d}1}^2]^{1/2}, \quad (4)$$

where $\sigma_{\bar{d}1}^2 = \mathrm{EMS1}(1/4 + 1/20)$ and $\sigma_{\bar{d}2}^2 = \mathrm{EMS2}(1/4 + 1/20)$.

If means from two generations are averaged, the variance of the mean difference between an IBL and its recurrent parent is $\sigma_{\bar{d}}^2 = \overline{\mathrm{EMS}}(1/8 + 1/40)$, where $\overline{\mathrm{EMS}}$ is the mean EMS for the two generations from Table 1.

## Results

Table 1 includes the effective error mean squares (EMS's) per plot for four quantitative characters studied in three generations. These were used to construct intervals and ellipses within which either 67% or 99% of inbred backcross lines should lie if they are genetically identical to their recurrent parent.

### *Inheritance of days to heading in Baart 46 genetic background lines*

For inbred backcross lines with the Baart 46 genetic background, a scatter diagram of the average heading date of B11(6)0 and B11(67)0 versus the heading date for the B11(8)0 generation is shown in Fig. 1. Ellipses about

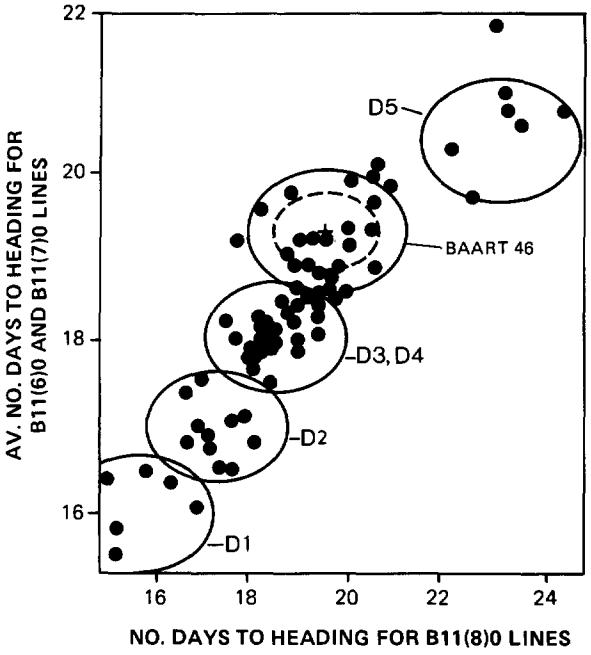


**Fig. 1.** Mean number of days to heading (after July 1) for the average of Baart 46 genetic background B11(6)0 and B11(7)0 versus descendent B11(8)0 inbred backcross lines; solid, and dashed, ellipses are expected to include 99%, and 67%, of lines with a specific genotype, respectively; the mean for Baart 46 families is indicated by a cross; D1–D5 are loci responsible for clusters of lines that differ in heading date from Baart 46

the mean number of days to heading (after July 1) for Baart 46 were constructed using Eqs. (3) and (4) and the error mean squares given in Table 1. The dashed ellipse should include 67% and the solid Ellipse should enclose 99% of all lines whose true mean (centroid) is the same as the mean for the recurrent (Baart 46) parent.

Four 99% ellipses in addition to the parental one are needed to enclose most of the lines. The number of lines expected to carry the Ramona allele of a single gene is $10.1 \pm 3.0$ with a 95% confidence interval of 5–17 lines. The two lowest and the highest ellipses of Fig. 1 enclose numbers of lines consistent with the single locus hypothesis.

The lowest (D1) ellipse encloses six lines which are an average of 3.4 days earlier than the Baart 46 parental mean. The locus responsible for the lines in the ellipse is designated D1 in Fig. 2. The D2 locus is responsible for the second lowest ellipse, containing a distinct cluster of 11 lines with a mean effect on heading date of −2.4 days (Table 2).

The third lowest ellipse (adjacent to the one for Baart 46) contains 26 lines. This is more than the upper 95% confidence limit of 17 lines expected if the ellipse was due to one locus. A group of nine to eleven of these lines are earlier than Baart 46 in the B11(6)0 and B11(7)0 genera-

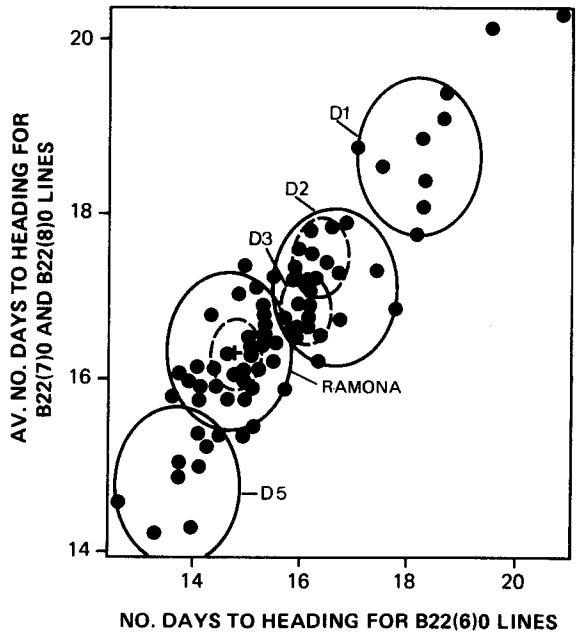


**Fig. 2.** Mean number of days to heading (after July 1) for the average of Ramona genetic background B22(7)0 and B22(8)0 lines versus ancestral B22(6)0 inbred backcross lines; solid, and dashed, ellipses are expected to include 99%, and 67%, of lines with a specific genotype, respectively; the mean for Ramona families is indicated by a cross; clusters of lines that differ from Ramona for the D1 and D5 heading date loci are indicated; the D2 and D3 ellipses enclose most of the lines attributable to D2 and D3

**Table 2.** Mean effects of loci on quantitative traits; designations are "D" for days to head, "H" for plant height, "R" for leaf rusting and "S" for shattering

| Locus | Days to head | Plant height (cm) | Seed weight (g/200) | Seed yield (g/plot) | Qualitative association |
|---|---|---|---|---|---|
| Baart 46 Background | | | | | |
| Locus D1 | −3.4 | −8.1 | | −122 | |
| Locus D2 | −2.4 | −6.1 | | | |
| Locus D3 | −1.5 | −3.6 | | | |
| Locus D4 | −0.8 | −2.0 | | | |
| Locus D5 | 2.5 | 10.4 | | 186 | Awn inhibition |
| Locus H1 | | 6.2 | | | |
| Locus H2 | | −5.9 | | | |
| Locus R1[a] | | | −0.57 | −76 | Fast leaf rusting |
| Locus S1 | | | | −219 | Shattering |
| Ramona Background | | | | | |
| Locus D1 | 3.0 | 14.2 | | 74 | |
| Locus D2 | 1.5 | 4.6 | | | |
| Locus D3 | 1.1 | 2.8 | | | |
| Locus D4 | | | | | |
| Locus D5 | −1.4 | −4.1 | | −114 | Awnlet development |
| Locus H1 | | −5.7 | | | |
| Locus H2 | | 4.1 | | | |
| Locus R1[a] | | | 0.75 | 94 | Slow leaf rusting |
| Locus S1 | | | | 0 | |

[a] B11(8)0 and B22(8)0 only

tions but are nearly identical to the Baart 46 parental lines in the B11(8)0 generation. The locus (designated D4) responsible for this group of lines has an effect of less than one day (Table 2). The remaining lines in the ellipse, which are somewhat later and differ from the Baart 46 parent to roughly the same extent in all generations, are accounted for by locus D3.

A group of seven very late lines, with a mean effect of 2.5 days on heading date (Table 2) is accounted for by locus D5. Plants in these lines have inhibited awn expression, instead of being fully awned like Baart 46.

### *Inheritance of days to heading in Ramona genetic background lines*

For inbred backcross lines with the Ramona genetic background, the average number of days to heading after July 1 is plotted in Fig. 2 for the average of the B22(7)0 and B22(8)0 generations versus the B22(6)0 generation.

The uppermost ellipse with a discrete cluster of 8 lines is attributable to locus D1 with an effect of 3.0 days (Table 2). The IBRs in this ellipse descended primarily from Wehrhahn's and Allard's (1965) locus 1 lines which had a mean effect under California growing conditions of 16.5 days, which is over five times the observed effect of 3.0 days in Canada.

Twenty five lines are in the 99% probability ellipse adjacent and to the upper right of the one for the Ramona parental group. This number of lines is intermediate between the numbers expected if two genes ($21 \pm 3.9$ lines) or three genes ($31 \pm 4.3$ lines) accounted for the ellipse. Two 67% probability ellipses are shown within the 99% ellipse. The topmost (D2) of these ellipses encloses lines that are later than Ramona in all generations. The other ellipse (D3) encloses lines that are slightly (1.1 days in Table 2) later than Ramona, particularly in the B22(6)0 generation.

The nine lines that are significantly earlier than Ramona are due to the early allele of locus D5. All of these lines were awnletted instead of awnless like Ramona.

There are two very late lines with an average heading date of about July 20, but so few lines would be expected with a probability of less than 0.01 if these lines were due to a single allele from Baart 46. The two lines are more likely to carry the late allele of the D1 gene as well as a late allele of either D2 or D3.

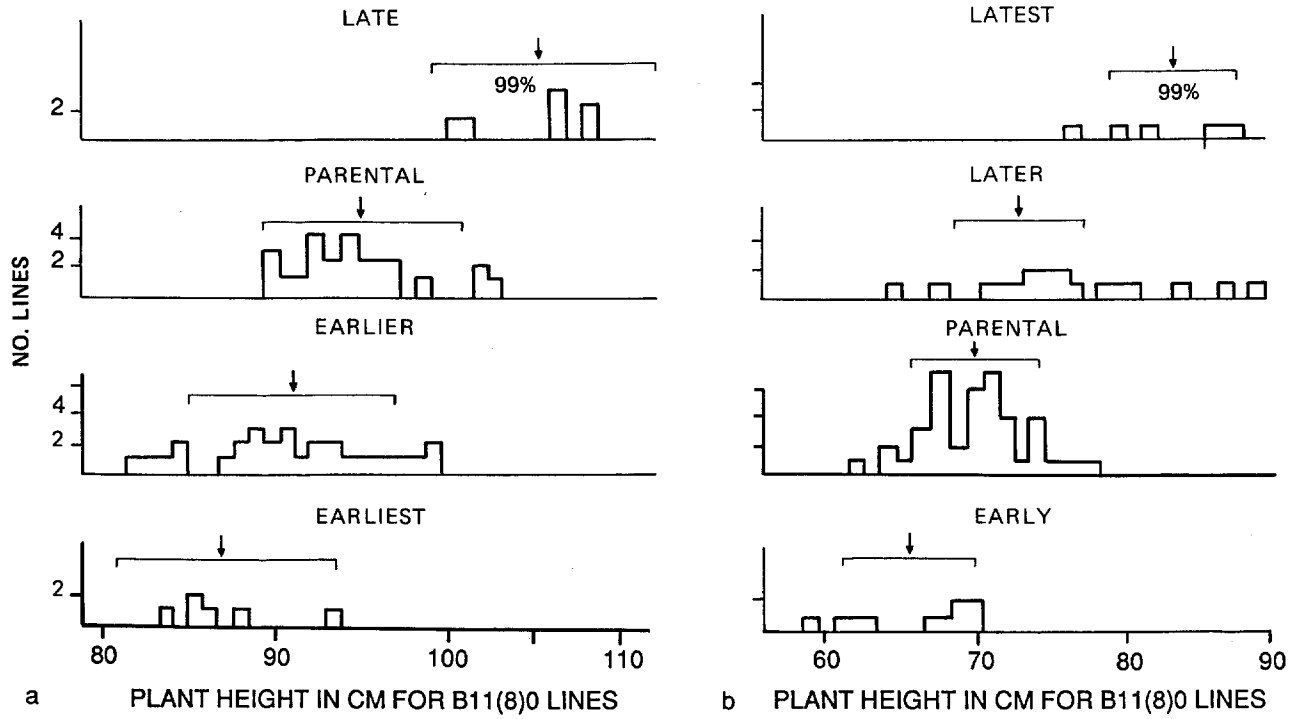


**Fig. 3a and b.** Plant height distribution of **a** Baart 46 background B11(8)0 lines and **b** Ramona background B22(8)0 lines in each of four heading date classes; the mean height of each distribution is indicated by an arrow; the horizontal lines indicate intervals in which 99% of genetically identical lines should lie; in **a**, the "earlier" lines include those carrying the early alleles of loci D2, D3 and D4; in **b**, the "later" lines include those carrying the late alleles of loci D2 and D3

*Inheritance of height*

Plant height was correlated with heading date in all experiments. Consequently, genes that have an effect on heading date also influence height. To study the effects of heading date genes on height, lines were separated according to their heading date genotypes, and the distribution for height within each heading date genotype was studied.

In the Baart 46 genetic background the distributions for plant height of lines in the earliest (locus D1), early (locus D2, D3 and D4), parental, and late (locus D5) heading date groups are presented in Fig. 3a for B11(8)0 lines. Clearly, the pleiotropic effects of heading date loci on height are proportional to their effects on the number of days to heading. The difference from the mean height of lines in the Baart 46 parental class is −8.1 cm for locus D1 and 10.4 cm for the late allele of locus D5 (Table 2). The "earlier" group of lines is an average of 3.86 cm shorter than the parental group of lines. Hence, loci D2, D3 and D4 have effects of −6.1, −3.6 and −2.0 cm respectively, if the effect on height of each gene is proportional to its effect on heading date.

Height also increases with heading date in plants with a Ramona genetic background (Fig. 3b). The difference of genotypic group mean heights from the mean of lines in the Ramona parental class is 14.2 cm for the latest genotype (locus D1), an average of 3.35 cm for "later" ones (loci D2, and D3), and −4.1 cm for the early allele of locus D5 (Table 2).

A 99% probability interval, centered at the mean height, is shown for each height distribution in Fig. 3. Some inbred backcross lines lie outside of these intervals which means that additional genes affecting height are present. If additional variation was solely due to many genes with small effects, the Central Limit Theorem would ensure that the distributions for height would be nearly Normal. Instead of being Normal, the distributions are highly platykurtic which implies that the residual variation for height is partly due to a gene or genes with relatively large effects.

Figure 4 is a scatter diagram of B11(7)0 and B11(8)0 heights for lines in and nearest the Baart 46 heading date ellipse in Fig. 1. The expected number of lines carrying a Ramona allele of a specific gene is $3.3 \pm 1.6$. Two lines due to the tall allele of locus H1 are significantly taller than Baart 46 and three lines due to the short allele of locus H2 are significantly shorter. The tall allele of H1 increased height 6.2 cm, and the short allele of H2 decreased height by 5.9 cm (Table 2).

Figure 5 is a scatter diagram of (B22(6)0 + B22(7)0)/2 versus B22(8)0 heights for the 41 lines in and near the

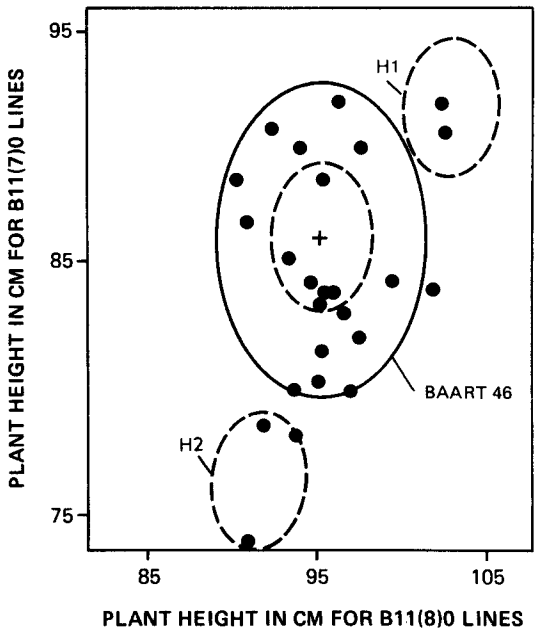


**Fig. 4.** Mean plant heights of Baart 46 genetic background lines that belong to the Baart parental heading date class; the mean height for Baart 46 families is indicated by a cross; dashed 67% probability ellipses enclose lines with the tall allele of H1 and the short allele of H2

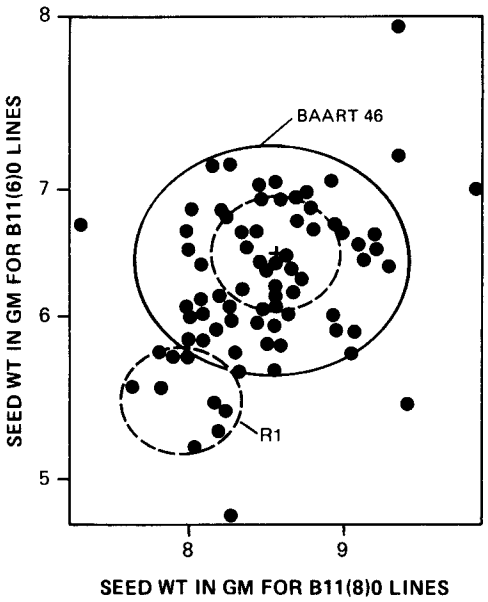


**Fig. 6.** Mean weights per 200 seeds of Baart 46 genetic background inbred backcross lines in the two generations when leaf rust was present; the mean of Baart 46 families is indicated by a cross; the lines enclosed by the R1 67% probability ellipse were very susceptible to leaf rust

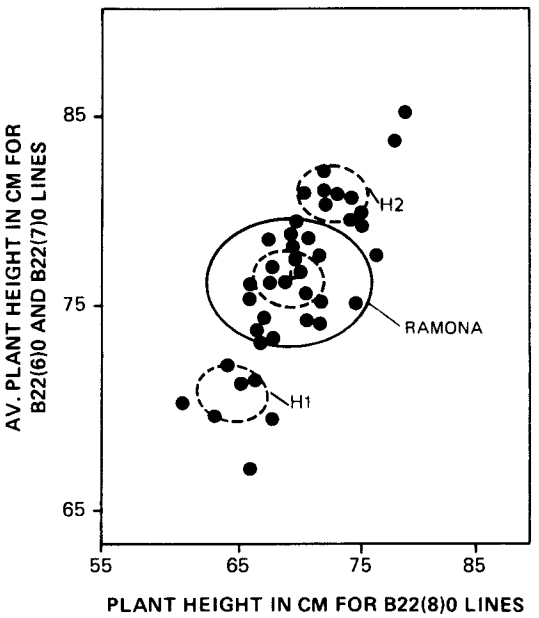


**Fig. 5.** Mean plant heights of Ramona genetic background lines that belong to the Ramona parental heading date class; the mean of Ramona families is indicated by a cross; dashed 67% probability ellipses enclose lines with the short allele of H1 and the tall allele of H2

Ramona heading date ellipse in Fig. 2. Nine lines are significantly taller than Ramona and are clustered together. Because the expected number of lines homozygous for a specific allele from Baart 46 is $5.4 \pm 2.0$ with a 95% confidence interval of 2–10 lines, this cluster is probably due to one gene (H2) with an effect of 4.1 cm (Table 2) on height. Seven lines are significantly shorter than Ramona in Fig. 4. The effect on height of the short allele of this locus (H1) is $-5.7$ cm (Table 2).

*Inheritance of seed weight (in g per 200 seeds)*

None of the genes with detectable effects on heading date and height had obvious effects on seed weight.

Figure 6 is a scatter diagram for seed weight (g per 200 seeds) of B11(6)0 versus B11(8)0 lines. A cluster of nine lines, enclosed by a 67% probability ellipse (R1), is outside of the 99% probability ellipse for the Baart 46 parental group, suggesting that the nonrecurrent Ramona parent has the low seed weight allele of a gene.

In the seed weight distribution of B22(6)0 versus B22(8)0 lines (Fig. 7) there is a group of ten lines with high weights, suggesting that Baart 46 has the high seed weight allele of the same gene.

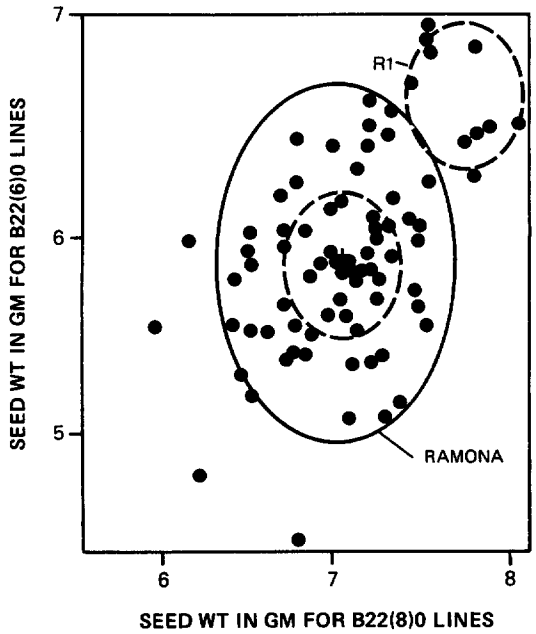


**Fig. 7.** Mean weights per 200 seeds of Ramona genetic background inbred backcross lines in the two generations when leaf rust was present; the mean of Ramona families is indicated by a cross; the lines enclosed by, and near, the R1 67% probability ellipse were scored as "slow leaf rusters"

Data on the severity of rust diseases was recorded in all experiments. None of the inbred backcross lines were resistant to leaf rust (*Puccinia recondita*). However, leaf rust developed more quickly on the lines with the Ramona genetic background than on those with the Baart 46 background. One gene was found to account for much of this difference (Tai unpublished). The lines in the low seed weight cluster of Fig. 6 carried the Ramona allele for fast rust development. The lines in the high seed weight cluster of Fig. 7 carried the Baart 46 allele for slow rust development.

Because the primary effect of the seed weight locus was to modify the virulence of the leaf rust pathogen, the locus is denoted as R1 in Figs. 6 and 7 and Table 2. The locus had no apparent effect on seed weight in the B11(7)0 and B22(7)0 experiments which were essentially free of rust. The effect (Fig. 6) of the rust susceptibility allele in the B11(8)0 lines was to reduce seed weight by 0.57 g per 200 seeds, which is 6.7% of the mean seed weight (8.5 g per 200 seeds) of the recurrent parental class. The consequent effect on seed yield (with a parent mean of 1130 g in the B11(8)0 generation) was −76 g (Table 2). The effect of the rust tolerance allele in the B22(8)0 lines was to increase the weight per 200 seeds by 0.75 g, or 11%. The resulting effect on seed yield per plot was 94 g (Table 2).

*Inheritance of seed yield*

Yield is highly correlated with the number of days to heading. Yield distributions of Baart 46 background inbred backcross lines, for the four major heading date categories, are presented in Fig. 8a. Some of the lines were susceptible to seed shattering. These were shaded in Fig. 8a and were removed before estimates of heading date gene effects on yield were calculated.

The early allele of locus D1, with an effect on heading date of −3.4 days, reduced yield by 122 g (Table 2); i.e., by 13% of the parental Baart 46 mean of 960 g. The yield of the "earlier" group of lines, carrying early alleles from loci D2, D3 and D4, do not differ greatly from the Baart 46 parental group. The late heading date allele of locus D5 increased yield by 186 g or 19%. This is a large pleiotropic effect on yield compared to the heading date effect of only 2.5 days (Table 2) which is less than 5% of the time from seeding to heading.

Yield distributions of Ramona background inbred backcross lines, for the four major heading date categories, are presented in Fig. 8b. Lines carrying the late allele of locus D1 had a 74 g (Table 2), or 9%, higher yield per plot than the mean of 832 g for Ramona families. This is only slightly lower than the effect of the locus in the Baart 46 genetic background. Again lines carrying "later" alleles of loci D2, D3 and D4 have about the same mean yield (834 g) as Ramona parent lines, suggesting that these loci do not influence yield directly or indirectly. The "early" class of lines, carrying the early allele of locus D5, had a mean yield of 718 g which is 14% less than the average for Ramona families. The effect on yield is almost as great as the 19% effect in the Baart 46 background. Because the direct effect, of the early allele of locus D5, on heading date was only −1.4 days (Table 2), it seems unlikely that the effect of the locus on yield is solely an indirect effect through the correlation between days to heading and yield.

The scatter diagram (Fig. 9) for yield of B11(7)0 versus B11(8)0 lines in the Baart 46 heading date category, has a group of seven lines with far lower (22%) yields than would be expected if the lines were genetically identical to Baart 46. These lines were found to be the ones susceptible to shattering in Fig. 8a. Lines with the shattering susceptibility allele of locus S1 had a 219 g lower mean yield (Table 2) than the 975 g for lines in the parental ellipse of Fig. 9.

The scatter diagram (Fig. 10) of B22(7)0 versus B22(8)0 seed yields of lines from the Ramona heading date class showed no additional loci with sufficiently large effects to be detectable.

No Ramona background inbred lines showed any evidence of shattering. Apparently the background in which locus S1 is expressed determines its effect on shattering and consequently on yield. Ramona is awnless

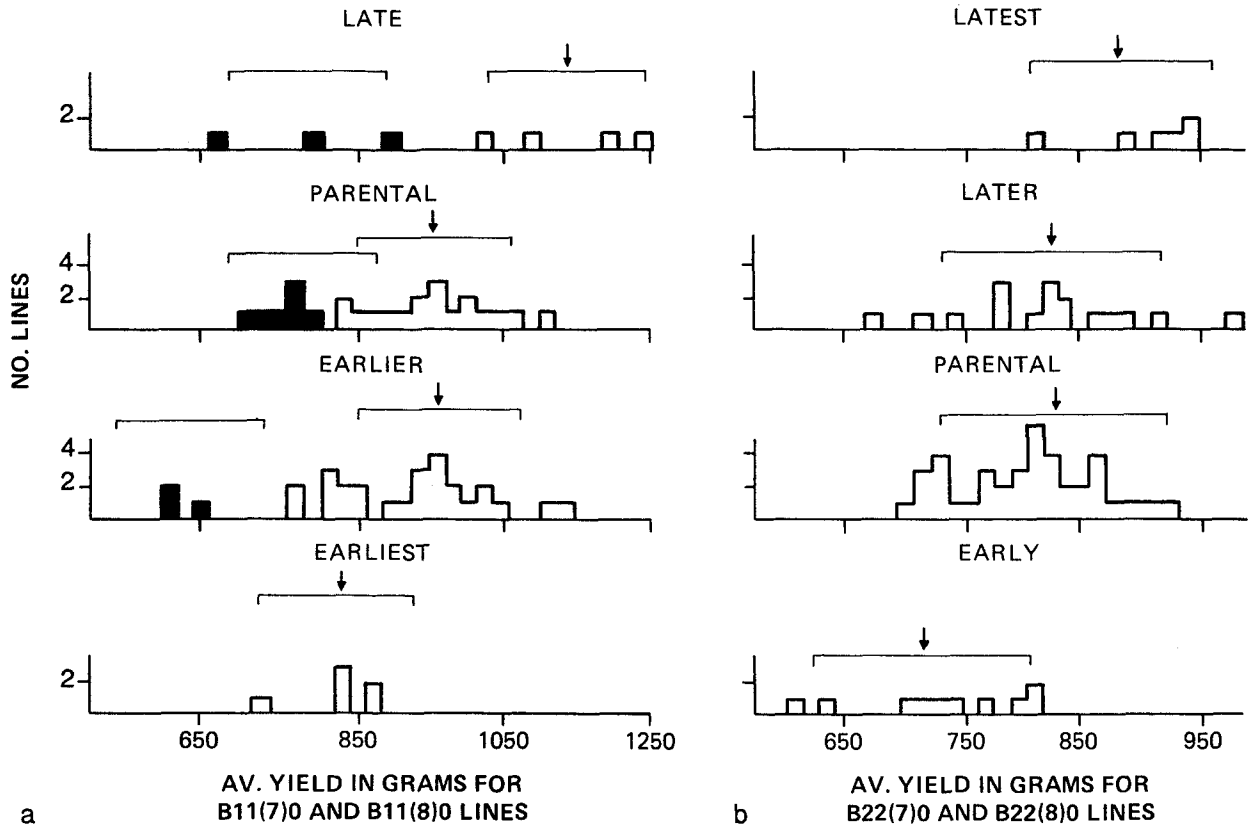


**Fig. 8a and b.** Seed yield distributions for different heading date classes of **a** Baart 46 genetic background and **b** Ramona genetic background inbred backcross lines; lines susceptible to shattering are shown in black; the mean yield of each distribution is indicated by an arrow; the horizontal lines indicate intervals in which 99% of genetically identical lines should lie; in **a**, the "earlier" lines include those carrying the early alleles of loci D2, D3 and D4; in **b**, the "later" lines include those carrying the late alleles of loci D2 and D3

**Table 3.** Total ($V_p$) and residual ($V_{rp}$) phenotypic variances, error variance ($V_e$), ratio ($V_{rg}/V_g$) of the residual genetic variance ($V_{rg} = V_{rp} - V_e$) to the total genetic variance ($V_g = V_p - V_e$), and F test for the hypothesis that the residual phenotypic variance is zero. The values are based on data from B11(7)0, B11(8)0, B22(7)0 and B22(8)0 inbred backcross lines

| | Variance | | | Residual | |
|---|---|---|---|---|---|
| | $V_p$ | $V_{rp}$ | $V_e$ | $V_{rg}/V_g$ | $F = V_{rp}/V_e$ |
| Baart 46 Background | | | | | |
| Days to head | 2.0349 | 0.1051 | 0.0562 | 2.5% | 1.87[b] |
| Plant height (cm) | 26.492 | 3.925 | 1.979 | 9.8% | 2.22[b] |
| Seed weight (g/200)[a] | 0.2328 | 0.1329 | 0.0664 | 39.2% | 1.94[b] |
| Yield (g/plot) | 16,781 | 6,050 | 1.093 | 31.6% | 5.53[b] |
| Yield with R1 gene | 16,781 | 5,452 | 1,093 | 27.8% | 4.99[b] |
| Ramona Background | | | | | |
| Days to head | 1.6730 | 0.1681 | 0.0766 | 5.7% | 2.19[b] |
| Plant height (cm) | 26.479 | 7.549 | 1.313 | 24.8% | 5.75[b] |
| Seed weight (g/200)[a] | 0.2255 | 0.0862 | 0.0500 | 20.6% | 1.72[b] |
| Yield (g/plot) | 6,696 | 3,177 | 950 | 38.8% | 3.34[b] |
| Yield with R1 gene | 6,696 | 2,263 | 950 | 22.9% | 2.38[b] |

[a] Based on B11(8)0 and B22(8)0 data
[b] Significant at the 5% level

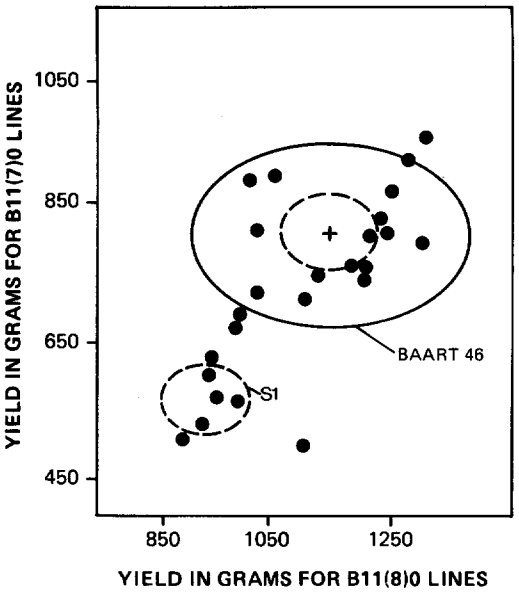


**Fig. 9.** Mean seed yields per plot of Baart 46 genetic background lines that belong to the Baart 46 parent's heading date class; the mean yield of Baart 46 families is indicated by a cross; the seven low yielding lines enclosed by or near the S1 ellipse are the ones found to be susceptible to shattering in Fig. 8

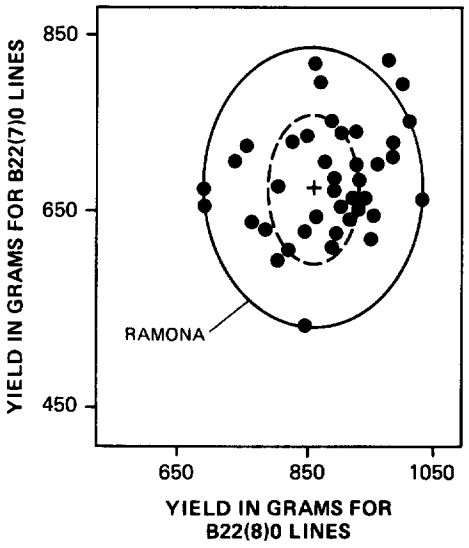


**Fig. 10.** Mean seed yields per plot of Ramona genetic background lines that belong to the Ramona parent's heading date class; the mean of Ramona families is indicated by a cross

whereas Baart 46 has long awns that help make spikes susceptible to shattering.

*Residual genetic variation*

A quantitative character may be affected by genes whose effects are too minuscule to be detected in our experiments. The residual phenotypic variance, $V_{rp}$ (Table 3), was calculated from the means of those lines which were in the parental class and, presumably, did not differ genetically from the recurrent parent with respect to the loci, listed in Table 2, that had detectable effects on the character. The residual genetic variance, $V_{rg}$, not accounted for by loci with detectable effects was calculated as $V_{rp}-V_e$, where $V_e$ is the average effective error mean square per replicate (Table 1) for generations (7)0 and (8)0 divided by the number (eight) of replicates per line in the two generations.

The variance of inbred backcross line means was used as an estimate of the total phenotypic variance, $V_p$ (Table 3), and the total genetic variance was calculated as $V_g=V_p-V_e$. These values can be used to estimate the ratio, $V_{rg}/V_g$, of the residual genetic variance to the total genetic variance (Table 3).

More than 94% of the total genetic variance in number of days to heading is attributable to the five heading date loci listed in Table 2. More than 90% of the genetic variance in plant height for lines with a Baart 46 genetic background is due to the seven genes with detectable effects on height.

When locus R1 is included, more than 72% of the genetic variation in yield is due to genes with detectable effects.

The ratio $V_{rp}/V_e$ is an F statistic (last column, Table 3) for a test of the hypothesis that the parental group lines are identical. The F statistics for all characters are significant at the 5% level, suggesting that all characters are affected by genes with undetectably small effects.

## Discussion

Table 2 contains a summary of the effects of loci on four quantitative characters.

Five loci have detectable effects (Fig. 1) on the number of days to heading (after July 1) in lines with the Baart 46 genetic background. These loci accounted for 97% of the total genetic variation in heading date (Table 3). The experiments involving lines with the Ramona genetic background are in accord with the hypotheses that either four or five loci have substantial effects on heading date (Fig. 2). The loci accounted for 94% of the total genetic variation in heading date (Table 3). Four of the five loci were previously detected in partially inbred lines by Wehrhahn and Allard (1965) and Tai (1968).

Locus D1 has the largest effect (Table 2; Wehrhahn and Allard 1965). Allard (1956) carried out extensive studies of $F_1$, $F_2$ and backcross progenies of the cross between Baart 46 and strain 3859. The effect of the late allele of locus D1 in Baart 46 was so large that skewed and bimodal distributions occurred. Using the "factorial method" Allard found that, if he postulated the existence of a second locus with one-fifth the effect of D1, he could

obtain a satisfactory fit to all of the distributions in a 1953 study.

When the partially inbred lines were grown in the short day length and cool (sufficient to vernalize winter wheats) winter conditions of California by Wehrhahn and Allard (1965) they had a heading date range of 28 days. The range in Canada under warm, long day length conditions was less than 10 days. Nevertheless, the same four loci (D1, D2, D3 and D5) were responsible for most of the heading date differences in the two locations and the heading date order was the same. Although D1 remained the locus with the largest effect (Figs. 1 and 2), its effect, relative to other loci, was much smaller (3.0–3.4 days) in Canada (Table 2) than in California (11.8–16.5 days). None of the other three loci are particularly sensitive to vernalization or to day length because temperature and day length do not alter their relative effects appreciably (Tai 1968).

As Wehrhahn and Allard (1965) noted, their heading date locus 4 (i.e., D5) was associated with awn expression. This was also the case in our study (Table 2). All of the lines in the D5 cluster of Fig. 2 were either homozygous or segregating for awnletted plants. Because two backcrosses and four generations of selfing did not break the linkage between the heading date effect of D5 and awn expression, locus D5 must be very closely linked to one of the genes controlling awn expression. Possibly D5 is located on the long arm of chromosome 5A which has a major gene *bl* for awn expression, as well as *Vrn1* for winter versus spring habit and other genes affecting heading date (Snape et al. 1985).

Height is influenced by the time available for vegetative growth prior to heading and, therefore, heading date loci have pleiotropic effects on height that are roughly proportional to their effects on heading date (Fig. 3). Height is also affected by two additional genes (Figs. 4 and 5) for a total of seven detectable genes (Table 2). These genes account for more than 90% of the total genetic variation of height in the Baart 46 genetic background (Table 3).

Seed weight (g per 200 seeds) was influenced by one detectable gene (Figs. 6 and 7). The primary effect of the gene was to modify the virulence of leaf rust. In the Baart 46 genetic background, the effect (Table 2) was to reduce seed weight by 6.7%: in the Ramona genetic background, the effect was to increase seed weight by 11%. Because almost all rust damage occurred after anthesis, the effect of the R1 gene on yield was proportional to its effect on seed weight (Table 2).

Yield per plot was influenced by heading date locus D1 and locus D5 (Fig. 8). Other heading date loci did not appear to affect seed yield appreciably.

In the Baart 46 genetic background, the Ramona allele of locus S1 reduced yield by 219 g per plot (Fig. 9, Table 2) or 22%. However, locus S1 had no measurable effect on lines with the Ramona genetic background (Fig. 10). The Baart 46 background lines with low yields were found to be susceptible to seed shattering. The Ramona allele of locus S1 apparently makes long awned plants like Baart 46 susceptible to shattering, but does not affect awnless plants such as Ramona.

In the Baart 46 genetic background, four genes (S1, R1, D1 and D5) account for over 72% of the genetic variation in yield (Table 3). In the Ramona genetic background, two genes (D1 and D5) with primary effects on heading date account for 61% of the genetic variation for yield (Table 3). These two genes, together with the R1 gene for rust tolerance, account for a remarkable 77% of the total genetic variation in seed yield of B22(8)0 lines.

### *Quantitative trait genes are sparsely distributed among chromosomes*

*Triticum aestivum* L. has 21 pairs of chromosomes. Only five genes were found to affect heading date. If the genes are randomly distributed throughout the genome it is very improbable that two genes are close enough to be regarded as "effective factors", in the terminology of Mather and Jinks (1982), rather than genes. In any case, genes with relatively large effects are sparsely distributed throughout the genome: less than one in four chromosome arms has any of the nine genes detected in our experiments. These results do not support the belief (Mather and Jinks 1982, p. 369) that "Polygenes, or polygenic complexes, may be built up by selection to have effects large enough for them to be located in the chromosomes by special techniques (Thoday 1961); but generally polygenes have effects sufficiently small by comparison with the residual variation to prevent them being followed as individuals in genetic analysis. Their effects are also similar to one another".

### *Distribution of the magnitude of gene effects*

The distribution of the magnitude of gene effects is of particular interest. Of the nine heading date effects in Table 2, one is less than 1 day in magnitude, four are betwen 1 and 2 days, three are between 2 and 3 days, and only one is greater than 3 days. Of the thirteen height effects in Table 2, six are less than 5 cm, five are between 5 cm and 10 cm, and only two are greater than 10 cm.

In general, absolute gene effects have a distribution which is intermediate between the uniform distribution and a reverse J shaped distribution. The distribution of gene effects is slightly skewed to the right: as the magnitude of effect increases the frequency of genes having the effect decreases.

Knowledge that a few genes have relatively large effects is especially important. The genes with large effects are the ones that are most subject to selection and most useful for plant improvement. In the case of heading date,

locus D1 accounts for over 50% of the variation attributable to known loci. In a selection experiment of the sort carried out by Allard and Harding (1963), locus D1 should undergo rapid allele frequency change and become fixed for one allele after a very few generations. At this point the heritability and the response to selection would fall to less than half of the initial value. By the time the first three or four loci are close to fixation, the remaining heritability would drop to less than ten percent of its original value, resulting in an apparent selection plateau.

*Gene-genetic background interactions*

The use of reciprocal sets of inbred backcross lines turned out to be informative. Although the effects of loci in the two genetic backgrounds differed somewhat, their effects were in the same direction and sometimes similar in both backgrounds. This was certainly the case for the D1 locus where the difference in growing conditions between Canada and California resulted in more than a three fold change in heading date effect in both the Ramona (3.0 vs 16.5 days) and Baart 46 (3.4 vs 11.8 days) genetic backgrounds.

Although the effect on heading date of locus D1 is similar in both Baart 46 and Ramona backgrounds, the pleiotropic effect of D1 on height in the Baart 46 background was 8.1 cm or 57% of its effect of 14.1 cm in the Ramona background (Table 2). The corresponding pleiotropic difference, due to genetic background, on seed yield was 165%. The D5 locus, which is associated with awn development, has pleiotropic effects for seed yield, heading date, and height that are 180 to 250 percent as large in the Baart 46 as in the Ramona genetic background (Table 2). Clearly, genetic background influences the effects of the two loci. Furthermore, the spectrum of pleiotropic effects of the two loci differ markedly.

The R1 gene had similar effects on seed weight and seed yield in both backgrounds when leaf rust was present and no effect when it was absent. The S1 locus was exceptional because its effect on yield was solely due to shattering in Baart 46 background plants.

*Alternative methods of analysis*

Mulitze and Baker (1985) developed a considerable body of simulation results and theory pertaining to the gene counting aspects of the inbred backcross line method. In their Table 3, they show that the gene number estimate is dependent on the type I error level. For heading date in the partially inbred Ramona genetic background lines of Wehrhahn and Allard (1965), their estimates were 2.5 genes for $\alpha = 0.01$, 3.5 for $\alpha = 0.05$, and 4.0 genes for $\alpha = 0.10$. For the inbred backcross lines with Baart 46 as the recurrent parent, the estimates were 3.4 genes for $\alpha = 0.01$, 4.0 for $\alpha = 0.05$ and 4.5 genes for $\alpha = 0.10$. These estimates are surprisingly close to the four genes with appreciable effects tentatively identified, through subjective classification of lines, by Wehrhahn and Allard. The estimates are also in reasonable accord with the observation, reported in this paper, of five genes with appreciable effects on heading date in inbred backcross lines with Baart 46 as the recurrent parent and either four or five genes in lines with Ramona as the recurrent parent.

*Acknowledgements.* We are grateful to K. K. Wu for technical assistance and helpful suggestions. We are also grateful to the University of Saskatchewan for use of its facilities.

## References

Allard RW (1956) Biometrical approach to plant breeding. In: Genetics in plant breeding. Brookhaven Symp Biol 9:69–88

Allard RW, Harding J (1963) Early generation analysis and prediction of gain under selection in derivatives of a wheat hybrid. Crop Sci 3:454–456

Allendorf FW, Knudsen KL, Leary RF (1983) Adaptive significance of differences in the tissue-specific expression of a phosphoglucomutase gene in rainbow trout. Proc Natl Acad Sci USA 80:1397–1400

Cochran WG, Cox GM (1957) Experimental designs, 2nd edn. Wiley, New York

Falconer DS (1981) Introduction to quantitative genetics, 2nd edn. Longman, London

Hallauer AR, Miranda JB (1981) Quantitative genetics in maize breeding. Iowa State Press, Ames/IA

Kahler AL, Wehrhahn CF (1986) Associations between quantitative traits and enzyme loci in the $F_2$ population of a maize hybrid. Theor Appl Genet 72:15–26

Kendall MG, Stuart A (1969) The advanced theory of statistics, vol 1: Distribution theory. Hafner, New York

Lande R (1980) The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215

Mather K, Jinks JL (1982) Biometrical genetics, 3rd edn. Cornell University Press, Ithaca/NY

Mulitze DK, Baker RJ (1985) Evaluation of biometrical methods for estimating the number of genes 2. Effect of type I and type II statistical errors. Theor Appl Genet 69:559–566

Robertson A (1967) The nature of quantitative genetic variation. In: Brink RA (ed) Heritage from Mendel. University of Wisconsin Press, Madison/WI, pp 265–280

Snape JW, Law CN, Parker BB, Worland AJ (1985) Genetical analysis of chromosome 5A of wheat and its influence on important agronomic characters. Theor Appl Genet 71:518–526

Sokal RR, Rohlf FJ (1981) Biometry, 2nd edn. Freeman, San Francisco

Soller M, Brody T (1976) On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theor Appl Genet 47:35–39

Tai GCC (1968) A study of the genic basis for differences in quantitative characters between two varieties of *Triticum aestivum* L. PhD Thesis, University of Saskatchewan, Saskatoon, Canada

Thoday JM (1961) Location of polygenes. Nature 199:368–370

Walters RFC, Wehrhahn K (1987) Calculus 1. Carslaw, Syndey

Wehrhahn C, Allard RW (1965) The detection and measurement of the effects of individual genes involved in the inheritance of a quantitative character in wheat. Genetics 51:109–119