Theor Appl Genet (1985) 69:559–566

Evaluation of biometrical methods for estimating the number of genes

2. Effect of type I and type II statistical errors*

D. K. Mulitze** and R. J. Baker

Crop Development Centre, University of Saskatchewan, Saskatoon, Saskatchewan S7N 0WO, Canada

Received July 15, 1984; Accepted August 4, 1984
Communicated by A. L. Kahler

Summary. Computer simulation of several genetic models was used to assess the effect of type I and type II statistical errors on estimating the number of genes by the inbred-backcross and genotype assay procedures. Depending upon the actual number of genes, heritability, and the probability of type I errors (α), substantial upward and downward biases were observed in estimates of the number of genes from both methods. The estimated number of genes increased as α was increased from 0.01 to 0.30 and as heritability increased. With high α and/or high heritability, the estimated number of genes often exceeded the actual number. Downward biases occurred with low α and low heritability, and tended to become greater as the number of genes in the model was increased. Large type II errors were associated with downward biases. The choice of α had a greater impact on biases in estimates from the genotype assay procedure than from the inbred-backcross procedure. Increasing the number of backcrosses in the inbred-backcross procedure or delaying the assay generation in genotype assay increased the probability of upward biases in the estimated number of genes. Unbiased estimates can be obtained only by choice of an optimum α. There is no known way to choose the optimum α in practice. This fact reduces the value of estimates of the number of genes by genotype assay or by the inbred-backcross methods.

Key words: Inbred-backcross method – Genotype assay – Number of genes – Quantitative genetics

* Part of this study is based on the Ph.D. thesis of the senior author

** Present address: Cereal Improvement Program, International Centre for Agricultural Research in the Dry Areas (ICARDA), P.O. Box 5466, Aleppo, Syria

Introduction

The inbred-backcross and genotype assay procedures were developed for estimating the number of genes governing quantitative traits in autogamous diploids. Both procedures rely on binary classification of genotypes based on tests of statistical hypotheses.

The inbred-backcross procedure (Wehrhahn and Allard 1965) requires production of inbred lines following several backcrosses to the recurrent parent and their subsequent classification as different from or not different from the recurrent parent. Inbred lines with means falling outside the confidence interval or confidence ellipse of the recurrent parent are classified as being non-parental. The estimated proportion of non-parental lines, $\hat{d}$, is then used to estimate the number of genes by the formula

$$\hat{k} = \ln(1-\hat{d})/\ln(1-\tfrac{1}{2}^{b+1}),$$

where b is the number of backcrosses (Mulitze and Baker 1984).

The genotype assay procedure (Jinks and Towey 1976; Towey and Jinks 1977) requires an estimate of the proportion (P_h) of randomly chosen F_n plants that are heterozygous for at least one locus. P_h is estimated by comparing means of two or more F_{n+1}-derived F_{n+2} lines developed by selfing each F_n assay plant. F_n plants which give rise to grandprogeny lines whose means are declared unequal by conventional statistical tests are classified as being heterozygous. The estimated proportion of heterozygous plants, $\hat{P}_h$, is then compared to theoretical expectations in order to estimate number of genes.

Researchers generally have chosen type I error (α) levels of 0.01, 0.05, or 0.10 when estimating the number of genes by the inbred-backcross or genotype assay procedures. Wehrhahn and Allard (1965) studied the genetic differences in days to heading between two spring wheat (*Triticum aestivum* L.) cultivars, 'Ramona' and 'Baart 46'. They constructed 90 and 50% probability ellipses for the 'Ramona' set of inbred-backcross lines and 95 and 67% probability ellipses for the 'Baart 46' set. They used α-levels of 0.10 for the 'Ramona' set and 0.05 for the 'Baart' set to support the hypothesis that four genes controlled days to heading in both sets of inbred-backcross lines.

Tai (1968) also used the 'Ramona'/'Baart 46' sets of inbred-backcross lines to study the genetics of days to heading in a much shorter growing season. Tai used $\alpha = 0.01$ and concluded that four genes accounted for most of the variation in days to heading. Wu et al. (1975) studied the inheritance of sedimentation value in the same material. They used $\alpha = 0.01$ to conclude that three genes accounted for most of the difference in sedimentation value between 'Ramona' and 'Baart 46'.

Talukdar (1972) analyzed inbred-backcross lines from a cross of 'Thatcher' and 'Selkirk' spring wheats. Using $\alpha = 0.05$, Talukdar concluded that three genes accounted for most of the differences in seed weight. Baker (1978) analyzed inbred-backcross lines from a cross of 'Neepawa' and 'Pitic 62' spring wheats. Based on tests at $\alpha = 0.05$, Baker concluded that at least five genes, and quite possible many more, controlled the difference between the parents in kernels/spike.

Towey and Jinks (1977) used the genotype assay procedure to study the inheritance of quantitative traits in a *Nicotiana rustica* cross. Estimates of the number of genes controlling flowering time were 1, 4, 5, 7, and 19 when tests were made with $\alpha = 0.05$ and plants from the F_2, F_3, F_4, F_5, and F_6 were assayed. When tests based on $\alpha = 0.01$ were used, comparable estimates were much lower at 1, 2, 2, 5, and 10. Similar estimates for final height were 2, 3, 4, 6, and 19 when $\alpha = 0.05$ and 1, 2, 3, 5, and 15 when $\alpha = 0.01$. Choice of α-level had considerable impact on estimated number of genes.

Both the inbred-backcross procedure and the genotype assay procedure depend upon adequate binary classification. Ideally, genotypic means should be estimated without error and then correctly classified. Any error in the classification procedure can result in biased estimates of d or P_h and thus in biased estimates of the number of genes. In order to minimize the probability of classifying a parental line as non-parental or a homozygous F_n plant as heterozygous, researchers using these two methods have chosen small type I error levels. The potential impact of type II errors, i.e. classifying non-parental lines as parental or heterozygous F_n plants as homozygous, has not been considered. The purpose of this study was to determine if type I and type II statistical errors have a significant impact on estimates of the numbers of genes obtained by the inbred-backcross and genotype assay procedures.

Methods

Computer simulation of genetic models was carried out by using the FORTRAN language on the University of Saskatchewan DEC 2060 computer. Subroutines from the International Mathematical and Statistical Library (IMSL) were utilized. For the simulation of genotypes, random deviates from the U(0, 1) distribution were generated by IMSL subroutine GGUW, a multiplicative congruential random number generator which shuffles random deviates from GGUBS (Anon. 1980, Vol. 2). Environmental deviates were generated by use of IMSL subroutine GGNPM to generate N(0, 1) deviates by the polar method (Knuth 1969). Environmental variance was computed as $\sigma_e^2 = \sigma_g^2(1-h^2)/h^2$, where h^2 was the heritability specified as an input parameter to the simulation program and σ_g^2 was the expected genetic variance for the model under consideration. Phenotypic values were computed as $P = G + \sigma_e z$, where G was the genetic value as determined by the genetic model and the individual's genotype, z was a N (0, 1) deviate from GGNPM, and σ_e was the environmental standard deviation.

For the equal effects model, genotypic values of 0.0, 1.0, and 2.0 were assigned to genotypes aa, Aa, and AA, respectively, at each locus. For the unequal additive effects model, genotypic values of 0.0, i, and 2i were assigned to genotypes aa, Aa, and AA, respectively, at locus i = 1, 2, ... k. Genotypes were stored in two-dimensional arrays with the integers 0 and 1 representing the a and A alleles, respectively. Only additive models without linkage were considered. The simulation programs used random walk and other genetic simulation techniques as described by Fraser and Burnell (1970) and by Crosby (1973).

Genotypes for homozygous inbred-backcross lines were generated in the following way. If the backcross plant under consideration was homozygous aa or AA at a particular locus, then the inbred line was considered to be homozygous for the same allele. If the backcross plant was heterozygous, then a U (0, 1) random deviate was generated and the inbred line was considered to be homozygous aa if the deviate was less than 0.5 and homozygous AA otherwise. Input parameters for the inbred-backcross simulation program, IB.FOR, included an initial seed for the random number generator, number of runs, number of inbred-backcross lines (m = 500), number of backcrosses (b = 1, 2 or 3), number of loci (k = 1, 2, ... 10), genetic model, heritability, and the t-value for the desired type I error level. In calculating σ_e, IB.FOR was programmed to compute the theoretical genetic variance among inbred-backcross lines.

Each inbred-backcross line was classified as parental or non-parental in the following way. For type I error (α) equal to 0.10, for example, those lines which deviated from the expected genotypic mean of the recurrent parent by more than 1.645 σ_e were classified as being non-parental. The number of non-parental lines were tallied to estimate the proportion of non-parental lines (d). The estimated number of genes controlling the trait was then given by $\hat{k} = \ln(1-\hat{d})/\ln(1-\frac{1}{2}^{b+1})$. In order to estimate the probability of a type I (α) or type II (β) error, the "true" situation was recorded for each inbred-backcross line by comparing its genotype with that of the recurrent parent. Simulation results were recorded as the averages of 20 runs.

Genotypes for the F_n individuals used for genotype assay were simulated as follows: (0, 0) representing the genotype aa was assigned to a locus if the U (0, 1) random deviate, r, was less than $(2^{n-1}-1)/2^n$; (1, 1) representing the genotype AA was assigned if r exceeded $1-(2^{n-1}-1)/2^n$; (0, 1) representing the genotype Aa was assigned if r was between those two limits. Genotypes for progeny and grandprogeny of assay plants were simulated by a random walk procedure. Input parameters for the genotype assay simulation program, GA.FOR, included an initial seed for random number generation, number of runs, assay generation (n = 3, 5), number of F_n plants (m = 200), number of loci (k = 1, 2, ... 8), heritability, and genetic model.

In calculating the environmental variance, GA.FOR was programmed to compute the theoretical genetic variance among F_2 plants given the input parameters specified above. For each run, a completely randomized design with s plants of each of p lines nested within each of m families was simulated (Table 1). The subroutine FTEST performed m F-tests per run, each testing the mean square for variation among lines within a family (DF = p−1) against the pooled within-line error variance (DF = mp(s−1)). FTEST used the IMSL subroutine MDFD to calculate the probability of each observed F-ratio. Families with probabilities less than 0.10, 0.05, and

Table 1. Expected mean squares for the completely randomized design used in the analysis of genotype assay simulations

Source of variation	Degrees of freedom	Expected mean square
Families	m-1	$V_e + s\,V_{lines} + ps\,V_{families}$
Lines within families	m(p-1)	$V_e + s\,V_{lines}$
Error	mp(s-1)	V_e

0.01 were tallied to provide estimates of numbers of genes at each of those type I error levels. In order to estimate α and β, the "true" situation was recorded for each family by comparing the mean genotypic values of the p F_{n+2} grandprogeny lines within each family. Simulation results were recorded as the average of 20 runs.

Validation of the simulation programs included tests of the random number generators as well as comparisons between theoretical and observed genotypic distributions. The random number generators were tested by several nonparametric goodness-of-fit and association tests (Kolmogorov-Smirnov one-sample test, Chi-square, Good's serial test, and the d^2 test) using IMSL subroutines. GGUW and GGNPM generated acceptable ($P < 0.05$) samples from the U(0, 1) and N(0, 1) distributions, respectively. Lewis et al. (1969) reached the same conclusions with a more stringent series of tests.

Program logic for IB.FOR and GA.FOR was considered correct when average simulated proportions of non-parental lines or segregating F_n plants failed to depart significantly from theoretical expectations. As a further check, the subroutine THEORY was written to calculate the expected numbers of inbred-backcross lines deviating from the recurrent parent at r=0, 1, ... k loci and the expected numbers of F_n plants segregating at r=0, 1, ... k loci. These expectations also were compared to observed data in the simulation programs. Since the inbred-backcross and genotype assay methods of analysis assume normal diploid meiosis and absence of selection, any simulation runs with populations that deviated from theoretical expectations at a significance level of 0.10 were considered to be products of poor random number sequences and were therefore bypassed. Furthermore, simulated type I error levels, genetic and environmental variances did not depart significantly from expected values.

Results

Inbred-backcross procedure

Estimates of numbers of loci carrying genes which affect a trait ($\hat{k}$) were computed as the means of 20 simulations of 500 inbred-backcross lines. Coefficients of variation for $\hat{k}$ and for the probability of a type II error (β) ranged from three to nine percent for all simulations. Standard errors of $\hat{k}$ ranged from 0.03 to 0.18, generally increasing with actual number of loci simulated. In order to more clearly indicate the direction and magnitude of the biases in $\hat{k}$, some results were plotted as bias ratios, the ratios of estimated numbers of loci to simulated numbers ($\hat{k}/k$).

Simulation of the equal and unequal additive effects models with k=1, 2, ... 10 and with b=2 backcrosses resulted in both upward and downward biases in the estimates of numbers of loci (Figs. 1 and 2). With equal effects, estimates generally exceeded those from the unequal effects model with the same number of loci. Type I error (α) levels had a significant impact on $\hat{k}$ with $\hat{k}$ being reduced as α was decreased. Although the bias ratio curves are not uniformly smooth because of the stochastic nature of genetic simulation (Figs. 1 and 2), average bias ratios for the three α-levels at any one heritability and any number of loci were always significantly different as determined by Tukey's w-procedure ($P=0.01$, 57 error degrees of freedom, three treatments). Differences between $\hat{k}$ at $\alpha=0.01$ and 0.10 increased from approximately 0.80 (k=1) to 2.00 (k=10), irrespective of heritability or genetic model. Decreased heritability of line means resulted in wider

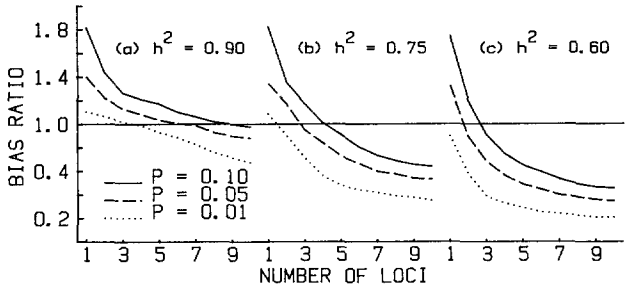

Fig. 1. Bias ratios ($\hat{k}/k$) for simulated inbred-backcross data for an equal additive effects genetic model with two backcrosses. Ratios are given for type I error levels (P) of 0.10, 0.05, and 0.01 and for heritabilities of (*a*) 0.90, (*b*) 0.75 and (*c*) 0.60

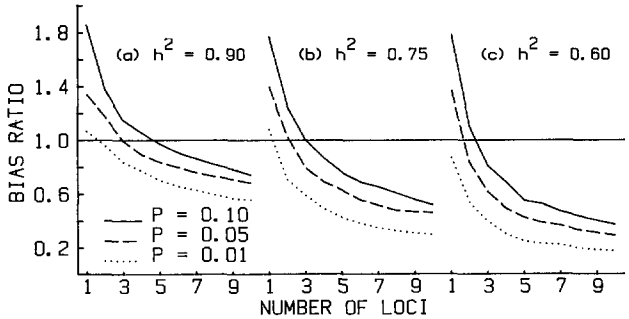

Fig. 2. Bias ratios ($\hat{k}/k$) for simulated inbred-backcross data for an unequal additive effects genetic model with two backcrosses. Ratios are given for type I error levels (P) of 0.10, 0.05, and 0.01 and for heritabilities of (*a*) 0.90, (*b*) 0.75 and (*c*) 0.60

confidence intervals about the recurrent parent. Therefore, more non-parental lines were classified as parental lines at low heritabilities. Large underestimates were observed at $h^2=0.60$ (Figs. 1c and 2c). Except for $h^2=0.90$ and $\alpha > 0.01$ (Fig. 1a), all $\hat{k}$ were biased downward for $k > 5$.

Type II error probabilities ($\hat{\beta}$) were estimated for simulations with α-levels of 0.01, 0.10 and 0.30 (Table 2). $\hat{\beta}$ generally increased as α decreased as would be expected from statistical theory. Each $\hat{k}$ was a function of α and β, which in turn was a function of heritability and the genetic model. Simulation with $k=10$, $\alpha=0.01$, $h^2=0.75$, and an equal additive effects model, for example, resulted in an average $\hat{\beta}$ of 0.499 (Table 2). The expected proportion of non-parental lines is $d=1-(1-\frac{1}{2}^{b+1})^k=0.737$. With 500 inbred-backcross lines and $\hat{\beta}=0.499$, one would expect $500\times0.737=368$ to be non-parental and, of these, $(1-0.499)\times368=184$ would be correctly classified as non-parental. Including a proportion of parental lines incorrectly classified as non-parental would add $(1-0.737)\times0.01\times500=2$, approximately, to the estimated number of non-parental lines. Thus, $\hat{d}=(184+2)/500=0.372$ is considerably below the expected proportion of 0.737 and the resulting estimated number of loci, i.e. 3.5, is biased downward. An α of 0.30, on the other hand, resulted in $\hat{\beta}=0.135$ and $\hat{k}=9.5$, close to the simulated number of loci (Table 2).

$\hat{\beta}$ increased as heritability decreased, resulting in greatly downward biased estimates of the number of loci at a heritability of 0.60 (Table 2). As heritability decreased, the difference between $\hat{k}$ at the three α-levels increased notably. As α increased from 0.01 to 0.10, $\hat{k}$ increased by roughly 35, 80, and 120% at heritabilities of 0.90, 0.75 and 0.60, respectively. Increasing α to 0.30 from 0.01 increased $\hat{k}$ by roughly 100, 200, and 300%for the three heritabilities. The α-level had a progressively greater impact on the estimated number of loci as the heritability decreased.

$\hat{\beta}$ also increased as k increased. With each addition of a locus into the genetic model, both the genetic and environmental variances were increased to maintain a constant heritability. This resulted in a decreased ratio of the additive effect at a locus to the environmental standard deviation and an increase in the type II error rates.

Table 2. Average[a] estimates of the number of loci ($\hat{k}$) and type II error ($\hat{\beta}$) for simulated inbred-backcross data with two backcrosses and with type I errors (α) of 0.01 to 0.30

α	Additive effects	Actual no. loci simulated					
		2		6		10	
		$\hat{k}$	$\hat{\beta}$	$\hat{k}$	$\hat{\beta}$	$\hat{k}$	$\hat{\beta}$
				($h^2=0.90$)			
0.01	Equal	2.1±0.04	0.000	5.3±0.08	0.090	6.7±0.11	0.199
	Unequal	2.0±0.04	0.039	3.9±0.06	0.265	5.6±0.08	0.290
0.10	Equal	2.8±0.05	0.000	6.6±0.09	0.012	9.7±0.11	0.056
	Unequal	2.8±0.04	0.003	5.5±0.07	0.152	7.4±0.08	0.169
0.30	Equal	4.5±0.06	0.000	8.5±0.10	0.002	12.0±0.12	0.017
	Unequal	4.7±0.06	0.001	7.9±0.14	0.088	10.2±0.13	0.104
				($h^2=0.75$)			
0.01	Equal	1.8±0.04	0.125	2.6±0.04	0.481	3.5±0.04	0.499
	Unequal	1.4±0.03	0.289	2.3±0.03	0.527	3.0±0.04	0.560
0.10	Equal	2.7±0.04	0.021	4.8±0.05	0.222	6.3±0.09	0.262
	Unequal	2.5±0.04	0.118	4.1±0.04	0.309	5.3±0.07	0.348
0.30	Equal	4.7±0.08	0.003	7.5±0.08	0.093	9.5±0.08	0.135
	Unequal	4.5±0.06	0.047	6.5±0.10	0.189	8.2±0.09	0.203
				($h^2=0.60$)			
0.01	Equal	1.1±0.03	0.440	1.5±0.03	0.678	2.0±0.03	0.680
	Unequal	1.1±0.04	0.468	1.4±0.04	0.705	1.8±0.04	0.717
0.10	Equal	2.4±0.06	0.156	3.5±0.06	0.406	4.6±0.09	0.416
	Unequal	2.2±0.04	0.249	3.2±0.05	0.456	3.9±0.05	0.486
0.30	Equal	4.5±0.06	0.062	6.2±0.07	0.227	7.4±0.09	0.252
	Unequal	4.2±0.05	0.136	5.5±0.05	0.291	6.7±0.08	0.304

[a] Average of 20 runs each with 500 inbred-backcross lines

Genetic models hypothesized by Wehrhahn and Allard (1965); Tai (1968), and Wu et al. (1975) for inheritance of traits in 'Ramona'/'Baart 46' inbred-backcross lines also were simulated (Table 3). The estimated number of loci ($\hat{k}$) for each of the models of Wehrhahn and Allard (1965), and Tai (1968) increased from about three to about six as α was increased from 0.01 to 0.30. Similar results were obtained for each of the three-locus models of Wu et al. (1975). Conclusions regarding the impact of α, β, and h^2 on $\hat{k}$ were identical for simulations involving the unequal effects model and models postulated in the literature.

Simulation trials with one or three backcrosses, but using environmental variances associated with simulation of two backcrosses and $h^2 = 0.90$, were conducted to assess the impact of changing the number of backcrosses on estimation of number of loci (Table 4). Repeated backcrossing reduced the genetic variance

Table 3. Average[a] estimates of the number of loci for simulated inbred-backcross data using models cited for days to heading and sedimentation value in a cross of 'Ramona' and 'Baart 46' spring wheats

Genetic model					Estimated no. of loci with type I error level (α) of:			
Additive effect of locus								
1	2	3	4	h^2	0.01	0.05	0.10	0.30
(Days to heading; Wehrhahn and Allard 1965)								
16.5	6.0	2.7	−3.5	0.955	2.5±0.03	3.5±0.04	4.0±0.06	6.3±0.07
−11.8	−3.0	−3.0	4.7	0.960	3.4±0.03	4.0±0.05	4.5±0.06	6.7±0.08
(Days to heading; Tai 1968)								
3.06	1.04	1.02	−0.73	0.951	3.4±0.04	4.0±0.06	4.4±0.07	6.4±0.08
−2.40	−0.90	−1.01	1.02	0.950	3.6±0.04	4.1±0.05	4.5±0.07	6.2±0.07
(Sedimentation value; Wu et al. 1975)								
11.9	6.5	−9.3	–	0.909	2.7±0.02	3.1±0.03	3.6±0.05	5.6±0.06
19.3	5.8	−8.9	–	0.880	2.1±0.02	2.7±0.03	3.6±0.04	5.4±0.06
−7.2	−7.2	7.7	–	0.877	2.8±0.03	3.0±0.04	3.4±0.05	5.3±0.06
−10.6	−10.6	12.9	–	0.835	2.5±0.03	3.1±0.03	3.4±0.05	5.4±0.07

[a] Average of 20 runs each with 500 inbred-backcross lines
[b] Heritability of inbred-backcross line means

Table 4. Average[a] estimates of the number of loci ($\hat{k}$) and type II error ($\hat{\beta}$) for simulated inbred-backcross data with one or three backcrosses and with type I errors (α) of 0.01 to 0.30

α	Additive effects	Actual no. of loci simulated					
		2		6		10	
		$\hat{k}$	$\hat{\beta}$	$\hat{k}$	$\hat{\beta}$	$\hat{k}$	$\hat{\beta}$
(One backcross; $h^2 = 0.94$)[b]							
0.01	Equal	2.0±0.02	0.000	5.3±0.07	0.056	7.2±0.09	0.074
	Unequal	2.0±0.02	0.028	3.9±0.04	0.177	5.8±0.06	0.142
0.10	Equal	2.4±0.03	0.000	6.3±0.07	0.009	9.1±0.11	0.023
	Unequal	2.3±0.03	0.003	5.1±0.07	0.089	7.4±0.14	0.075
0.30	Equal	3.3±0.04	0.000	7.2±0.10	0.007	10.6±0.11	0.007
	Unequal	3.2±0.05	0.000	6.2±0.09	0.055	8.8±0.09	0.046
(Three backcrosses; $h^2 = 0.83$)[b]							
0.01	Equal	2.2±0.04	0.000	5.4±0.10	0.100	6.6±0.10	0.275
	Unequal	2.1±0.04	0.026	4.0±0.08	0.321	5.4±0.10	0.393
0.10	Equal	3.6±0.07	0.000	7.5±0.12	0.020	10.0±0.14	0.086
	Unequal	3.7±0.06	0.002	6.3±0.09	0.173	8.3±0.15	0.245
0.30	Equal	7.5±0.11	0.000	11.6±0.20	0.002	15.2±0.17	0.022
	Unequal	7.5±0.00	0.000	10.6±0.17	0.109	12.7±0.18	0.150

[a] Average of 20 runs each with 500 inbred-backcross lines
[b] Heritability adjusted to maintain environmental variance at same level as used with two backcrosses and a heritability of 0.90

and, since the environmental variance was held constant, the heritability. Heritabilities were 0.94, 0.90 and 0.83 under both genetic models with b = 1, 2 and 3, respectively. Although $\hat{\beta}$ increased as heritability decreased with each additional backcross, $\hat{k}$ increased for $\alpha > 0.01$. With increased backcrossing, type I errors took on a relatively more important role than type II errors and led to upward biases. With k = 10, for example, d = 0.944 at b = 1 and d = 0.474 at b = 3. With $\alpha = 0.30$, $\hat{d}$ would be biased upward by (1–0.944) $\times 0.30 = 0.017$ for b = 1 and by $(1-0.474) \times 0.30 = 0.158$ for b = 3 because of parental lines being incorrectly classified as non-parental lines (type I error). For the unequal effects model, $\hat{\beta}$ increased from 0.046 at b = 1 to 0.150 at b = 3. With these levels of type II errors, $\hat{d}$ would be biased downward by $0.944 \times 0.046 = 0.043$ at b = 1 and by $0.474 \times 0.150 = 0.071$ at b = 3 because of non-parental lines being incorrectly classified as parental. The net results were downward biased estimates of d and k at b = 1 ($\hat{d} = 0.944 + 0.017 - 0.043 = 0.918$; $\hat{k} = 8.76$) and upward biased estimates at b = 3 ($\hat{d} = 0.474 + 0.158 - 0.071 = 0.561$; $\hat{k} = 12.70$) in this example (Table 4). Comparisons of one or three backcrosses to two backcrosses from simulation trials at other heritabilities led to the same conclusions. In general, estimates of gene number from the inbred-backcross procedure will increase as the number of backcrosses increases.

Genotype assay procedure

Simulation trials for genotype assay involved primarily the equal additive effects model with no dominance ($P_{int.A}$ of Towey and Jinks 1977) for k = 1, 2, ... 8. GA.FOR calculated environmental variances relative to F_2 single plant heritabilities of 0.80, 0.60, and 0.40. For assay of F_3 plants, simulated heritabilities of means of F_3-derived F_5 families were approximately 0.92, 0.90, and 0.86 while those for differences among means of F_4-derived F_5 lines within families were approximately 0.86, 0.75, and 0.62 for the corresponding F_2 heritabilities of 0.80, 0.60 and 0.40. For assay of F_5 plants, the corresponding heritabilities for family means were approximately 0.98, 0.97 and 0.94 while those for line means (mean of ten plants) within families were 0.76, 0.47 and 0.40. Average heritabilities of differences among line means within families decreased because inbreeding decreased the genetic variance among lines within families. Heritability of differences among line means within families is critical for detecting heterozygous plants in genotype assay.

In all simulations, coefficients of variation for $\hat{k}$ ranged from three to 13%. Simulations of F_3 assay with α equal to 0.01 to 0.10 revealed a consistent and significant downward bias in $\hat{k}$ for $k > 3$ (Fig. 3).

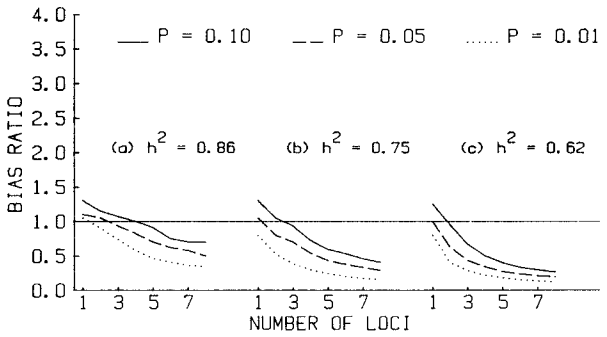

Fig. 3. Bias ratios ($\hat{k}/k$) for simulated F_3 genotype assay data for an equal additive effects genetic model. Ratios are given for type I error levels (P) of 0.10, 0.05 and 0.01 and for heritabilities of (*a*) 0.86, (*b*) 0.75 and (*c*) 0.62

Except for $k > 7$ and $h^2 = 0.62$ (Fig. 3c), all three bias ratios for the same k and the same h^2 were significantly different ($P = 0.05$). As heritability decreased, fewer differences were detected among grandprogeny line means within families. As α was increased from 0.01 to 0.10, more families with lines having equal means were incorrectly classified as having lines with unequal means. This resulted in the classification of more families as being derived from heterozygous assay plants and an upward bias in the estimated number of loci. At $\alpha = 0.10$, an upward bias was observed for 2–3 loci (Fig. 3). As k increased, more small differences went undetected with the result that the downward bias in $\hat{k}$ was increased.

Simulation trials with the number of plants per line (s) other than ten also were conducted. Simulations for k = 8 loci, $\alpha = 0.01$, $h^2 = 0.86$ for F_3-derived F_5 family means, and s = 4, 6, 10, 12 and 16 resulted in estimates of the number of loci $\hat{k} = 1.0 \pm 0.02$, 1.8 ± 0.04, 2.4 ± 0.04, 2.5 ± 0.04, and 3.0 ± 0.05, respectively. Increased sample sizes gave increased power to the F-tests and resulted in higher $\hat{k}$. A sample size greater than s = 10 would increase the bias ratios plotted in Fig. 3. With many loci and low heritability, s must be impractically large to counter the downward bias in estimated numbers of loci.

A greater potential for upward bias was observed in F_5 assay compared to F_3 assay (Fig. 4). Differences between $\hat{k}$ at the three α-levels were greater than in the F_3 and were significant for all k and h^2. Substantial upward biases in $\hat{k}$ were observed for $k < 5$ at all heritabilities simulated. With further inbreeding, more F_n plants were homozygous and there were therefore more cases where F_{n+1}-derived lines were equal. Because of this, type I errors had a greater impact in F_5 than in F_3. The expected proportion of detectable heterozygous plants ($P_{int.A}$) in the F_5 was small for few loci. Thus, a type I error level of 0.10, for example, resulted in a greatly increased $\hat{k}$. $P_{int.A}$ for one locus, for

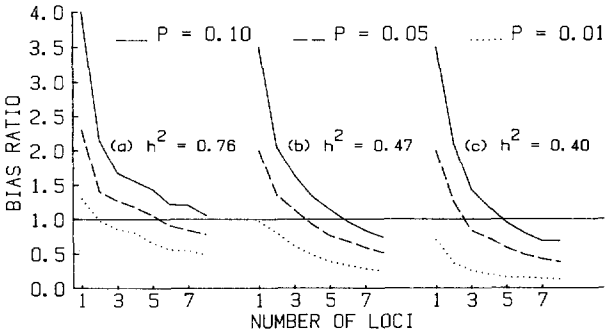

Fig. 4. Bias ratios ($\hat{k}/k$) for simulated F_5 genotype assay data for an equal additive effects genetic model. Ratios are given for type I error levels (P) of 0.10, 0.05 and 0.01 and for heritabilities of (*a*) 0.76, (*b*) 0.47 and (*c*) 0.40

example, is 0.039 in the F_5. With $\alpha=0.10$, the estimated proportion would be approximately 0.139; almost equal to $P_{int.A}=0.144$ expected with four loci. The bias ratio under high heritability was approximately 4.0 as expected from the above consideration (Fig. 4a). As k increased, the frequency of F_n homozygotes decreased and fewer families with equal line means were available for misclassification due to type I errors. Furthermore, the pooled error term for the F_5 was less than the pooled error for F_3 because of reduced genetic variance among plants within each line. Consequently, the power of the F-test was increased and smaller differences were detected. Except in the case of low h^2 and small α, $\hat{k}$ was higher in F_5 than in F_3.

As in the F_3, unbiased estimates of the number of loci required use of the proper level of type I error. An α of 0.05, for example, was required for an unbiased estimate when $k=5$ and $h^2=0.76$ (Fig. 4a) while $\alpha=0.10$ was required for an unbiased estimate when $h^2=0.40$ (Fig. 4c). Estimates of the number of loci using conventional statistical criteria in F_5 genotype assay can be biased upward or downward depending upon heritability and the actual number of loci.

Other simulation trials for later generations showed the same phenomena as observed in F_5 but with even greater bias ratios. Simulations of genotype assay with other genetic models (unequal effects, presence or absence of dominance, etc.,) led to the same conclusions as reported for the equal additive effects model. The biases observed seem to be independent of the underlying genetic model.

Genotype assay appears to be sensitive to the choice of α-levels. Type I error levels beyond the conventional range of 0.01 to 0.10 may be required for unbiased estimates of the number of loci, particularly for assay in later generations and in early generations when the actual number of loci exceed 5 or 6. It is apparent that, once α has been chosen, β will depend upon the actual number of loci and heritability. The imbalance of type I and type II errors may lead to estimates of the number of loci which are biased upward or downward. Estimates in practice will be unreliable because they are complex functions of h^2, k, α and β.

Discussion

In this study, analysis of simulated data for inbred-backcross lines was based on the classification of each line as parental or non-parental. This approach differs from that used by Wehrhahn and Allard (1965) and others who classified lines into somewhat subjective groups. While the approach used in this study does not afford estimates of the effects of individual genes, it was chosen to avoid the subjectivity that is apparent in papers such as those by Wehrhahn and Allard (1965), and Tai (1968). The approach of Wehrhahn and Allard (1965) often excludes lines from the analysis because their means fall outside all group limits. Moreover, their methods do not allow for multi-gene deviates and fail to recognize that there may be errors in assigning lines to various groups. Re-analysis of inbred-backcross data reported by Wehrhahn and Allard (1965); Tai (1968) and others using the type I error level reported invariably resulted in greater estimates of numbers of genes than reported.

In estimating the number of loci carrying genes which control a quantitative trait, unbiased estimates will result only if one can obtain unbiased estimates of the proportion of non-parental lines in the inbred-backcross method of Wehrhahn and Allard (1965) or of the proportion of heterozygous plants in the genotype assay method of Jinks and Towey (1976). The simulation results reported herein show that there is a need for the proper balance of type I and type II errors if unbiased estimates are to be obtained. The nature of this balance can be formalized in the following way.

Let d_i be the difference between means of two lines within the i^{th} family in the genotype assay method or the difference between the mean of the i^{th} inbred-backcross line and the recurrent parent in the inbred-backcross method. Then, the required proportion, P, of segregating plants or non-parental lines is estimated by carrying out a series of tests of the null hypothesis $H_0: d_i=0$. A type I error, α, is the probability of rejecting H_0 when it is true while a type II error, β, is the probability of accepting H_0 when it is false. $H_0: d_i=0$ is true with expected frequency (1–P) and false with expected frequency P. The estimate of P will be equal to the probability that H_0 is true (1–P) times the probability that it will be incorrectly rejected (α) plus the probability that H_0 is false (P) times the probability that it will correctly be rejected (1–β). That is, $\hat{P}=(1-P)\alpha+P(1-\beta)=P+(1-P)\alpha-P\beta$, from whence

it is clear that $\hat{P}$ will be an unbiased estimate of P only if $(1-P)\alpha = P\beta$.

In practice, it is impossible to choose the correct level of α because the correct value depends on P which in turn depends on the number of loci affecting the trait. The simulation results show that type II errors will be important when the actual number of loci is large or when heritabilities are low. In those cases, estimated numbers of loci will be biased downward. Upward biases may develop if there are few genes and high heritability. This will be particularly true when using inbred-backcross procedures with three or more backcrosses or the genotype assay procedure in later generations.

Snape et al. (1984), in discussing genotype assay of doubled haploid lines, indicated that the method will detect only those effective factors with effects which exceed the level of sensitivity of the experiment. A similar conclusion with respect to the inbred-backcross method was implied by Wehrhahn and Allard (1965) who concluded that "among the genes which differentiate the wheat varieties 'Ramona' and 'Baart 46', there are four which have large enough effects on heading date to be detected by an 'inbred backcross line' experiment of modest size". Such interpretation of the results of genotype assay or the inbred-backcross method may be an oversimplification. For example, after two backcrosses 41.4% of inbred-backcross lines are expected to differ from the recurrent parent at one or more loci if the parents differ at four loci. If, on the other hand, the parents differ at ten loci but the experiment was sufficient precision to detect only those lines which deviate from the recurrent parent by two or more loci, then 36.1% would be classified as non-parental. In the latter case, it would be incorrect to interpret the results as evidence for the presence of four genes of sufficiently large effect to be detected. Without a priori knowledge that there are few major genes, it is doubtful that one can interpret results in the way suggested by Snape et al. (1984) and Wehrhahn and Allard (1965). As pointed out by Wehrhahn and Allard, traits controlled by few genes should show discontinuity in the distribution of inbred-backcross lines. Unfortunately, there does not appear to be an objective test for discontinuity and some researchers do not see this as a requirement for application of the method.

Mulitze and Baker (1984) have shown that both methods require large sample sizes in order to give reasonably precise estimates of the number of loci.

In this paper, we show that those estimates, however precise, may be biased. Because the bias may be upward or downward, we are not able to say that estimates of the number of genes are either upper or lower bounds to the actual number. Our evaluation serves to point out that there are major difficulties in attempts to infer precise genetic models from the study of continuous variation.

Acknowledgements. This research was carried out while the senior author was in receipt of a scholarship from the Canadian Wheat Board. The authors are grateful for statistical advice from Professor W. H. Laverty, Mathematics Department, University of Saskatchewan.

References

Anon (1980) International mathematical and statistical library, 9th edn. IMSL Inc, Houston, Texas

Baker RJ (1978) Evaluation of the inbred-backcross method for studying the genetics of continuous variation. Can J Plant Sci 58:7–12

Crosby JL (1973) Computer simulation in genetics. John Wiley and Sons, New York

Fraser AS, Burnell D (1970) Computer models in genetics. McGraw-Hill, New York

Knuth DE (1969) The art of computer programming. Vol II. Seminumerical algorithms. Addison-Wesley, Reading Mass

Jinks JL, Towey P (1976) Estimating the number of genes in a polygenic system by genotype assay. Heredity 37:69–84

Lewis PAW, Goodman AS, Miller JM (1969) A pseudo-random number generator for the System/360. IBM Systems J 8:136–146

Mulitze DK, Baker RJ (1984) Evaluation of biometrical methods for estimating the number of genes. 1. Effect of sample size. Theor Appl Genet 69:553–558

Snape JW, Wright AJ, Simpson E (1984) Methods for estimating gene numbers for quantitative characters using doubled haploid lines. Theor Appl Genet 67:143–148

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

Talukdar BS (1972) The genetic basis for the difference in seed weight between Thatcher and Selkirk wheats. Can J Genet Cytol 14:667–673

Towey P, Jinks JL (1977) Alternative ways of estimating the number of genes in a polygenic system by genotype assay. Heredity 39:399–410

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

Wu KK, Sosulski FW, Wehrhahn CF (1975) The genic basis for differences in sedimentation value and protein content between two cultivars of *Triticum aestivum*. Can J Genet Cytol 17:433–439