Theor Appl Genet (2006) 112: 391–399
DOI 10.1007/s00122-005-0042-z

ORIGINAL PAPER

Jixiang Wu · Johnie. N. Jenkins · Jack C. McCarty
Sukumar Saha · David M. Stelly

An additive-dominance model to determine chromosomal effects in chromosome substitution lines and other gemplasms

Received: 18 April 2005 / Accepted: 2 July 2005 / Published online: 9 December 2005

Abstract When using chromosome substitution (CS) lines in a crop breeding improvement program, one needs to separate the effects of the substituted chromosome from the remaining chromosomes. This cannot be done with the traditional additive-dominance (AD) model where CS lines, recurrent parent, and their hybrids are used. In this study, we develop a new genetic model and software, called a modified AD model with genotype × environment interactions, which can predict additive and dominance genetic effects attributed to a substituted alien chromosome in a CS line as well as the overall genetic effects of the non-substituted chromosomes. In addition, this model will predict the additive and dominance effects of the same chromosome of interest (i.e. chromosome 25 of cotton in this study) in an inbred line, as well as the effects of the remaining chromosomes in the inbred line. The model requires a CS line, its recurrent parent and their F_1 and/or F_2 hybrids between the substitution lines and several inbred lines. Monte Carlo simulation results showed that genetic variance components were estimated with no or slight bias when we considered this modified AD model as random. The correlation coefficient between predicted effects and true effects due to the chromosomes of interest varied from zero to greater than 0.90 and it was positively relative to the difference between the CS line and the recurrent line. To illustrate the use of this new genetic model, an upland cotton, G*ossypium hirsusum* L, CS line (CS-B25), TM-1 (the recurrent parent), five elite cultivars, and the F_2 hybrids from test-crossing these two lines with the five elite cultivars were grown in two environments in Mississippi. Agronomic and fiber data were collected and analyzed. The results showed that the CS line, CS-B25, which has chromosome 25 from line 3 to 79, *Gossypium barbadense* substituted into TM-1, had positive genetic associations with several fiber traits. We also determined that Chromosome 25 from FiberMax 966 had significantly positive associations with fiber length and strength; whereas, chromosome 25 from TM-1 and SureGrow 747 had detectable negative genetic effects on fiber strength. The new model will be useful to determine effects of the chromosomes of interest in various inbred lines in any diploid or amphidiploid crop for which CS lines are available.

Communicated by B. Friebe

J. Wu
Department of Plant & Soil Sciences,
Mississippi State University, Mississippi State, MS, 39762 USA

Johnie. N. Jenkins (✉) · J. C. McCarty · S. Saha
USDA-ARS, Crop Science Research Lab, P.O. Box 5367,
Mississippi State, MS, 39762 USA
E-mail: JJenkins@msa-msstate.ars.usda.gov
Tel.: +1-662-3207386
Fax: +1-662-3207528

D. M. Stelly
Department of Soil & Crop Sciences, Texas A&M University,
College Station, TX, 77843-2474 USA

Introduction

Near-isogenic (NI) lines derived by backcrossing are useful for quantitative trait analysis. Chromosome substitution (CS) lines, which are divergent for one chromosome pair while NI to the recurrent parent for the remaining chromosome pairs, also offer a great potential for dissecting quantitative traits of interest, without or in conjunction with DNA markers available. CS lines have been used to detect genes for quantitative traits like yield and grain quality associated with specific chromosomes in wheat (Law 1966, 1967; Law et al. 1976; Al-Quadhy et al. 1988; Zemetra et al. 1986, 1988; Mansur et al. 1990; Berke et al. 1992a, 1992b). Chromosomes 3A and 6A from wheat cultivar Wichita were determined to have major quantitative trait loci (QTLs) that increased grain yield and kernel weight when present in cultivar Cheyenne, while Cheyenne had major QTLs on 3A and 6A that decreased grain yield and kernel weight in cultivar Wichita (Berke et al. 1992a). Kohel et al. (1977) and Ma

and Kohel (1983) evaluated six CS lines and indicated several quantitative genes to these chromosomes in upland cotton. QTLs for boll size, lint percentage, fiber length, and fiber elongation were mapped to chromosome 16 of Pima 3–79 using 178 families from the cross with TM-1 (Ren et al. 2002). Chromosome 25 of Pima 3–79 in a TM-1 background (CS-B25) has been identified to have desirable genes affecting fiber micronaire, length, and strength, while chromosomes 16 and 18 of 3–79 have genes negatively associated with yield (Saha et al. 2003, 2004).

A significant deviation of a CS line from its recurrent parent for a specific trait is usually considered to show association of that chromosome with the specific quantitative trait. When a specific chromosome in a recurrent parent is replaced by the respective chromosome from a donor parent, genes on the alien chromosome are expressed in the genetic background of the recurrent parent. Thus, the total genetic effect of a CS line may be due to the gene(s) on the substituted chromosome, the genes on the remaining chromosomes of its recurrent parent, and an interaction between the gene(s) on the substituted chromosome and the remaining chromosomes of its recurrent parent. With only data from one CS line, a donor parent, and a recurrent parent, the chromosome effects and interaction effects between the substituted chromosome and the remaining chromosomes of the recurrent parent cannot be separated by comparative analysis. When data from the reciprocal CS lines and their two recurrent parents are evaluated, the substituted chromosome and interaction effects with remaining chromosomes can be determined by the two-way analysis of variance (ANOVA) method (Berke et al. 1992a, 1992b). To detect dominant effects for quantitative traits due to a specific chromosome, F_1 hybrids between CS lines and their recurrent parents should be analyzed (Yen et al. 1997).

The development of a set of CS lines is usually very time-consuming. For example, it took more than 25 years to complete CS line development in wheat (Zemetra et al. 1986; Berke et al. 1992a), and more than 20 years for the same in cotton (unpublished data). Up to now, most studies focused on detecting chromosome effects using CS lines, recurrent parents, and/or their hybrids or progenies (Law 1966, 1967; Law et al. 1976; Berke et al. 1992a, 1992b; Saha et al. 2003, 2004). The above-mentioned studies have provided genetic information relative to the substituted chromosome association with traits of importance; however, they were not able to determine the genetic effects due to the chromosomes of interest in other inbred lines. Therefore, it should be more helpful to develop a new genetic model and analytical method through which the CS lines can be extensively used to determine desirable genetic effects of specific chromosomes when crossed with other cultivars or inbred lines. Such a study should greatly improve the use of CS lines in both genetic mapping and breeding programs.

In this research, we propose a new genetic model and develop the corresponding software package that can be used to dissect the effects of specific chromosomes when a CS line and its recurrent parent are crossed with various cultivars or inbred lines. Monte Carlo simulation technique was used to detect the estimated variance components and predicted genetic effects. To illustrate the use of the new genetic model, we utilized our two-location data set from CS-B25 and its recurrent parent TM-1 test-crossed with five upland cotton cultivars which are usually maintained by open pollination due to its predominant selfing nature.

Genetic models and methodology

Genetic model

The phenotypic value for any genotype with replications under multiple environments can be expressed as follows,

$$y = \mu + E + G + GE + B + e \tag{1}$$

where, μ is the population mean; E is the environmental effect, G is the genotypic effect; GE is the genotype × environment interaction effect; B is the block effect if applicable, e is the random error.

The total genotypic value (G) for a quantitative trait in Eq. 1 for a genotype relative to a CS line includes the effects of the specific substituted chromosome ($G_{(1)}$), effects of the remaining non-substituted chromosomes of the recurrent parent ($G_{(2)}$), and the possible interaction effects between the specific substituted chromosome and the remaining chromosomes of recurrent parent ($G_{(12)}$). Therefore, G and GE can be rewritten as,

$$G = G_{(1)} + G_{(2)} + G_{(12)} \tag{2}$$

$$GE = GE_{(1)} + GE_{(2)} + GE_{(12)} \tag{3}$$

Now we consider a more complicated case. Assume that there are one CS line, its recurrent parent, and n inbred lines, and top-crosses (e.g. F_1 or F_2) of the CS line and its recurrent parent crossed with these n inbred lines, and that both $G_{(1)}$ and $G_{(2)}$ include only additive and dominance effects, then, we call this genetic model a modified AD model. We define P_{ij} as a parent used for a cross, where i represents the index for the specific chromosome ($i = 1, \ldots, n+2$), and j the index for the remaining non-specific chromosomes ($j = 1,\ldots, n+1$). $P_{i'j}$ is called the CS line of P_{ij} when a specific chromosome i in P_{ij} is replaced by its respective chromosome from a donor parent i'. Then genotypic value G for a parental line can be expressed as follows,

$$G(P_{ij}) = 2A_{i(1)} + D_{ii(1)} + 2A_{j(2)} + D_{jj(2)} \tag{4}$$

Genotypic value for F_1 between P_{ij} and P_{kl}

$$G(P_{ij} * P_{kl})(F_1) = A_{i(1)} + A_{k(1)} + D_{ik(1)} + A_{j(2)} + A_{l(2)} + D_{jl(2)} \quad (5)$$

Genotypic value for F_2 between P_{ij} and P_{kl}

$$G(P_{ii} * P_{kl})(F_2) = A_{i(1)} + A_{k(1)} + 0.25D_{ii(1)} + 0.25D_{kk(1)} + 0.5D_{ik(1)} + A_{j(2)} + A_{l(2)} + 0.25D_{jj(2)} + 0.25D_{ll(2)} + 0.5D_{jl(2)} \quad (6)$$

If $i = k$, then Eqs. 5 and 6 can be expressed as Eqs. in 7 and 8, respectively,

$$G(P_{ij} * P_{kl})(F_1) = 2A_{i(1)} + D_{ii(1)} + A_{j(2)} + A_{l(2)} + D_{jl(2)} \quad (7)$$

$$G(P_{ii} * P_{kl})(F_2) = 2A_{i(1)} + D_{ii(1)} + A_{j(2)} + A_{l(2)} + 0.25D_{jj(2)} + 0.25D_{ll(2)} + 0.5D_{jl(2)} \quad (8)$$

If $j = l$, then Eqs. 5 and 6 can be expressed as Eqs. in 9 and 10, respectively,

$$G(P_{ij} * P_{kj})(F_1) = A_{i(1)} + A_{k(1)} + D_{ik(1)} + 2A_{j(2)} + D_{jj(2)} \quad (9)$$

Genotypic value for F_2 between P_{ij} and P_{kl}

$$G(P_{ii} * P_{kj})(F_2) = A_{i(1)} + A_{k(1)} + 0.25D_{ii(1)} + 0.25D_{kk(1)} + 0.5D_{ik(1)} + 2A_{j(2)} + D_{jj(2)} \quad (10)$$

Where $A_{(1)}$ and $D_{(1)}$ are additive effects and dominance effects due to the chromosomes of interest, and $A_{(2)}$ and $D_{(2)}$ are additive effects and dominance effects due to the remaining chromosomes,

The GE can be also partitioned into its respective $G*E$ components following above Eqs. 4, 5, and 6.

Analytical approaches

For a specific data set, these genetic effects can be considered as fixed and may be analyzed by a general linear model (GLM) approach; however, some of the coefficients in these genetic models may not be 0 or 1 and the data set may be unbalanced. The GLM approach cannot analyze the data with the complicated genetic model. In this study, we consider all genetic effects and block effects as random. Environmental effects can be considered as fixed or random which we are not interested in. In this study the environmental effects were considered as random. The advantage of this consideration is that it allows both the estimation of variance components and the prediction of genetic effects possible (Searle et al. 1992).

The Eqn. 1 can be extended based on Eqs. 4, 5, and 6 accordingly and it can be also expressed in the form of matrices and vectors in Eq. 11,

$$\begin{aligned}\mathbf{y} &= \mathbf{1}\mu + \mathbf{U}_E\mathbf{e}_E + \mathbf{U}_{A(1)}\mathbf{e}_{A(1)} + \mathbf{U}_{D(1)}\mathbf{e}_{D(1)} + \mathbf{U}_{A(2)}\mathbf{e}_{A(2)} \\ &\quad + \mathbf{U}_{D(2)}\mathbf{e}_{D(2)} + \mathbf{U}_{AE(1)}\mathbf{e}_{AE(1)} + \mathbf{U}_{DE(1)}\mathbf{e}_{DE(1)} \\ &\quad + \mathbf{U}_{AE(2)}\mathbf{e}_{AE(2)} + \mathbf{U}_{DE(2)}\mathbf{e}_{DE(2)} + \mathbf{U}_B\mathbf{e}_B + \mathbf{e}_e \\ &= \mathbf{1}\mu + \sum_{r=1}^{11}\mathbf{U}_u\mathbf{e}_u \end{aligned} \quad (11)$$

where the constant μ is the population mean, and $\mathbf{U}_u$ is the known incidence matrix relative to the vector $\mathbf{e}_u \sim N(\mathbf{0}, \sigma_u^2\mathbf{I}_u)$.

The variance components for genetic effects can be obtained using the minimum norm quadratic unbiased estimation (MINQUE) (Rao 1971; Searle et al. 1992). Genetic effects can be predicted by the adjusted unbiased prediction (AUP) or the linear unbiased prediction (LUP) approach (Zhu 1993; Zhu and Weir 1994).

Monte Carlo simulations were conducted in this study to evaluate the estimations of variance components by the MINQUE approach. The efficiency of prediction with the LUP method was also compared by simulation. Pseudo-random normal deviates with mean zero and preset variance were generated by the Polar algorithm (Devroye 1986). For each case, 500 simulations were run to obtain sample means of estimates, bias and mean square error (MSE) which is defined as $\text{MSE} = \text{var}(\hat{\theta}) + \text{bias}^2$. For simplicity, randomized complete block designs, with four replication, within each of the 2 years, was used in this study. Assume that there are one CS line, its recurrent parent, and five inbred lines, and top-crosses of the CS line and its recurrent parent crossed with these five inbred lines. We set three cases. Case 1 includes parents and F_1 hybrids, case 2 parents and F_2 hybrids, and case 3 parents, F_1 and F_2 hybrids. So there were 17, 17, and 27 genotypes for the cases one, two, and three, respectively. The real data set in section Example has the same genetic design and replicate number as the case two in our simulation study.

Monte Carlo Simulation Results

Estimation of variance components

The simulation results for bias and MSE are summarized in Table 1 for variance component. Variance components were estimated almost with no bias by this modified AD model. The data set with both F_1 and F_2 hybrids did not give noticeable improvement for variance estimation. On an average, the variance components due to the chromosome of interest were not affected by the genetic effects related to the remaining chromosomes. The MSE obtained by parents and both F_1 and F_2 hybrids were slightly lower than that by parents and F_2 hybrids. The additive variance component due to the chromosome of interest was slightly underestimated while the additive variance component due to the remaining chromosomes was slightly overestimated. A MSE is related to the bias and variation

Table 1 Results for estimating variance components based on 500 simulations using the modified AD model

	True value	P + F1		P + F2		P + F1 + F2	
		Bias	MSE	Bias	MSE	Bias	MSE
σ^2_{A1}	20	3.83	19.64	−1.57	6.65	0.41	5.19
σ^2_{D1}	20	0.73	1.55	−0.69	2.03	0.93	1.75
σ^2_{A2}	0	0.00	5.13	4.82	27.47	0.33	5.04
σ^2_{D2}	0	0.00	0.69	0.00	1.12	0.18	0.65
σ^2_{AE1}	20	0.05	2.09	0.85	2.84	1.80	5.02
σ^2_{DE1}	20	−0.01	0.45	−1.21	2.67	−0.66	0.86
σ^2_{AE2}	0	0.00	1.86	0.00	1.79	0.00	1.75
σ^2_{DE2}	0	0.00	0.30	2.12	5.36	0.04	0.32
σ^2_{ε}	20	0.13	0.04	0.01	0.02	0.01	0.01
σ^2_{A1}	20	0.01	4.11	−2.48	10.69	−0.98	5.27
σ^2_{D1}	20	2.77	8.63	1.77	4.85	0.02	1.05
σ^2_{A2}	20	−0.50	7.38	3.22	17.82	3.42	19.84
σ^2_{D2}	20	−1.10	3.91	0.97	4.89	0.14	2.42
σ^2_{AE1}	20	−0.52	2.00	0.20	2.27	−1.32	3.24
σ^2_{DE1}	20	0.10	0.54	0.82	2.10	0.19	0.48
σ^2_{AE2}	20	0.51	3.05	−2.12	7.49	1.18	4.16
σ^2_{DE2}	20	0.71	1.60	−0.72	2.36	0.06	1.07
σ^2_{ε}	20	0.09	0.03	0.04	0.02	1.26	1.59

among different simulation. In this study, the high MSEs were mainly due to the large biases.

Prediction of genetic effects

Researchers are interested not only in the overall genetic information and genetic variance components, but also in the genetic effects related to the chromosome of interest and due to the remaining chromosomes. In this simulation study, we also compared the predicted genetic effects with the true genetic effects using correlation analysis. We ran the simulations based on random and fixed effects models. When the genetic effects due to the chromosome of interest were fixed, the additive correlation coefficient between the predicted values and true values due to the chromosome of interest could be greater than 0.90 on average, and it was positively related to the difference between the CS line and the recurrent line (results not shown). Dominant correlation coefficient due to the chromosome of interest was generally less than 0.80. Similar results were found when the genetic effects were set as random, the additive correlation coefficients ranged from 0 > 0.90 depending on the difference between the CS line and the recurrent line (data not shown).

Example

Materials and experiments

In this study, TM-1 and one upland cotton CS line, CS-B25, are near isogenic, with the exception of the replacement of a specific homologous pair of chromosome 25 from 3–79 (*Gossypium barbadense*) into the recurrent parent TM-1, upland cotton (*Gossypium hirsutum*) (Stelly et al. 2004). TM-1 is an inbred line derived from the commercial variety Deltapine 14 by Kohel et al. (1970). CS-B25 and TM-1 were used as male and top-crossed with five elite cultivars in 2002 at Mississippi State University. The five elite cultivars are: 'Deltapine 90' (DP90); 'FiberMax 966'(FM 966); 'Stoneville 474' (ST 474); 'Phytogen 355'(PSC 355); and 'SureGrow 747' (SG 747). CS-B25, TM-1, the 5 cultivars, and 10 F_2 hybrids were planted at two locations with replicated plots in 2003 at the Plant Science Research Center at Mississippi State, MS. Soil type for location 1 was a Marietta loam (Fine-loamy, siliceous, active, fluvaquentic Eutrudepts) and for location 2 it was a Leeper silty clay loam (Fine, smectitic, nonacid, thermic Vertic Epiaquept). Plots were planted in a plant two skip one row pattern on 28 May and harvested on 3 November, 2003 at location 1 and 31 October at location 2. Standard cultural practices were followed and the environmental conditions during the growing season were above average for each location. A 25-boll sample per plot was hand harvested from the first fruited positions from the middle nodes of plants to determine boll weight and fiber properties. Samples were ginned on a ten-saw laboratory gin to determine lint percentage and the lint samples were sent to StarLab (StarLab, Inc., Knoxville, TN, USA) for single instrument fiber measurements. Micronaire (MIC), elongation (EL), 2.5% span length (SL), and fiber strength (T1) were measured. After the boll samples were harvested, all plots were harvested with a commercial cotton picker modified to bag seed cotton from each plot and lint weight (kg) ha^{-1} was calculated.

The data were analyzed by using the above modified AD model by the MINQUE(1) approach (Zhu 1989). The proportion of each variance component to the phenotypic variance was also calculated. The variance component used for calculation of the corresponding proportion was based on F_1 generation: $V_{A1} = 2\sigma^2_{A1}$,

$V_{D1} = \sigma^2_{D1}$, $V_{A2} = 2\sigma^2_{A2}$, $V_{D2} = \sigma^2_{D2}$; $V_{AE1} = 2\sigma^2_{AE1}$, $V_{DE1} = \sigma^2_{DE1}$, $V_{AE2} = 2\sigma^2_{AE2}$, $V_{DE2} = \sigma^2_{DE2}$, $V_e = \sigma^2_e$, and $V_P = V_{A1} + V_{D1} + V_{A2} + V_{D2} + V_{AE1} + V_{DE1} + V_{AE2} + V_{DE2} + V_e$. Genetic effects were predicted by the AUP approach (Zhu 1993). The resampling (jackknife) method was applied to calculate the standard error (SE) for each parameter by removal of each block within each of two locations. The *t*-test was used to detect the significance of each parameter (Miller 1974) and the degrees of freedom were 7. One and two-tail tests were used to test the significance of the variance and the genetic effects, respectively. The data analyses were conducted by using self-written programs in C++ for the modified AD genetic models.

Variance components

Estimated proportions of variance components to the phenotypic variance based on the modified AD genetic model for all traits are summarized in Table 2. No additive (*A1*) or dominant (*D1*) effects were detected due to chromosome 25 for lint percentage, seed cotton yield, or lint yield. Both additive (*A1*) and dominant effects (*D1*) were significant for micronaire and 2.5% span length. *D1* was significant for boll weight and A1 for elongation and strength.

Significant additive effects for lint percentage, boll weight, seed cotton yield, lint yield, and elongation due to the remaining 25 chromosomes (*A2*) and dominant effects for strength due to the remaining chromosomes (*D2*) were detected (Table 2). Strong additive effects for lint percentage (77%) were due to *A2*. Lint percentage, seed cotton yield, lint yield, elongation, and fiber strength were affected by dominance × environment interaction effects due to chromosome 25 (D1E) (17, 29, 25, 22, and 33%, respectively). Dominance by environment interaction effects due to chromosomes, other than chromosome 25, were significant for micronaire, elongation, and 2.5% span length. Residual variance ranged from 10 to 21% for all traits except lint percentage (5%). These results indicated that substituted chromosome 25 has genetic association with boll weight, micronaire, elongation, 2.5% span length, and fiber strength, while it does not have genetic association with lint percentage or cotton yield.

Predicted genetic effects

The additive effects of chromosome 25 (*A1* effects) for each of the parental lines varied for different traits, (Table 3). Chromosome 25 from TM-1 and FM966 had a positive additive effect on boll weight, while chromosome 25 from the other 5 lines had a nil or negative additive effect. Chromosome 25 from CS-B25 reduced micronaire, but chromosome 25 from TM-1, ST474, and PS 355 increased micronaire. Chromosome 25 from CS-B25, DP 90, and FM966 reduced fiber elongation, while chromosome 25 from TM-1, PSC355, and SG747 increased elongation. Chromosome 25 from CS-B25 and FM966 increased 2.5% span length, whereas chromosome 25 from TM-1, and ST474 reduced 2.5% span length. Chromosome 25 from CS-B25 and from FM966 increased strength, while chromosome 25 from TM-1, ST474, and SG747 reduced strength.

Table 2 Estimated proportions of variance components for two genetic models using F_2 and Parents

	Modified AD model							
	LP	BW	YLD	LY	MIC	EL	SL	T1
V_{A1}/V_P	0.00	0.17	0.00	0.00	0.20*	0.13*	0.18*	0.31**
V_{D1}/V_P	0.00	0.37**	0.00	0.00	0.34**	0.07	0.25*	0.00
V_{A2}/V_P	0.77**	0.06*	0.15*	0.36**	0.00	0.22*	0.00	0.00
V_{D2}/V_P	0.01	0.17	0.26*	0.14*	0.00	0.00	0.00	0.21*
V_{A1E}/V_P	0.00	0.01	0.00	0.00	0.02	0.00	0.04	0.00
V_{D1E}/V_P	0.17*	0.00	0.29	0.25	0.00	0.22*	0.00	0.33*
V_{A2E}/V_P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05*
V_{D2E}/V_P	0.00	0.12	0.12	0.11	0.29**	0.17*	0.32**	0.00
V_e/V_P	0.05**	0.10**	0.19**	0.14**	0.15**	0.19**	0.21*	0.10*
	Traditional AD model							
	LP	BW	YLD	LY	MIC	EL	SL	T1
V_A/V_P	0.78**	0.27**	0.17*	0.43**	0.18**	0.40**	0.23**	0.43**
V_D/V_P	0.00	0.49*	0.22*	0.00	0.35*	0.00	0.21	0.00
V_{AE}/V_P	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
V_{DE}/V_P	0.15*	0.10*	0.39*	0.39*	0.23*	0.32*	0.23*	0.38**
V_e/V_P	0.07**	0.14**	0.22**	0.19**	0.24**	0.28**	0.33*	0.19**

LP lint percentage, *BW* boll weight, *YLD* seed cotton yield, *LY* lint yield, *MIC* micronaire, *EL* elongation, *SL* 2.5% span length, and *T1* fiber strength

A1 additive for chromosome 25, *D1* dominance for chromosome 25, *A2* additive for chromosomes other than 25, *D2* dominance for chromosomes other than 25, *A1E* additive × environment for chromosome 25, *D1E* dominance × environment for chromosome 25, *A2E* additive × environment for chromosomes other than 25, *D2E* dominance × environment for chromosomes other than 25, *A* additive for all chromosomes, *D* dominance for all chromosome, *AE* additive× environment for all chromosomes, and *DE* dominance × environment for all chromosomes

* and ** are probability levels of 0.05 and 0.01, respectively

Table 3 Additive effects ± standard errors due to chromosome 25 (*A1*) for boll weight and four fiber traits based on the modified AD model

A1	BW (g)	MIC	EL (%)	SL (mm)	T1(kNm/kg)
CS-B25	−0.13 ± 0.09	−0.34 ± 0.11	−0.28 ± 0.15	0.66 ± 0.29	14.08 ± 1.52
TM1	0.30 ± 0.15	0.14 ± 0.07	0.39 ± 0.17	−0.50 ± 0.28	−15.02 ± 1.52
DP90	−0.19 ± 0.08	−0.09 ± 0.03	−0.25 ± 0.08	−0.02 ± 0.18	2.47 ± 1.33
FM966	0.30 ± 0.11	−0.02 ± 0.02	−0.68 ± 0.12	0.55 ± 0.14	13.28 ± 1.89
ST474	−0.16 ± 0.05	0.11 ± 0.03	0.04 ± 0.06	−0.44 ± 0.09	−3.99 ± 1.09
PS355	−0.13 ± 0.05	0.17 ± 0.04	0.39 ± 0.09	0.03 ± 0.07	1.24 ± 1.45
SG747	0.00 ± 0.02	0.04 ± 0.03	0.38 ± 0.09	−0.30 ± 0.04	−12.10 ± 0.92
Grand $\bar{X}$	5.68	4.56	8.42	29.34	204.9

BW Boll weight, *MIC* micronaire, *EL* elongation, *SL* 2.5% span length, and *T1* fiber strength

Homozygous dominance effects, $D_{ii(1)}$ and $D_{jj(2)}$, measure the degree of inbreeding depression and should be included (Zhu, 1993) when more than two parents are used in crosses. Dominant effects due to chromosome 25 (*D1*) for four traits are summarized in Table 4. Chromosome 25 from CS-B25 and all five commercial cultivars had negative homozygous dominant effects for boll weight. On average, the heterozygous effects for boll weight between chromosome 25 from CS-B25 and the chromosome 25 from five commercial cultivars was greater than that between chromosome 25 from TM-1 and chromosome 25 from five commercial cultivars, possibly indicating that more diversity between chromosome 25 from CS-B25 and commercial cultivars causes the larger heterozygous dominant effects for boll weight. Chromosome 25 of CS-B25 showed positive dominant homozygous effects for micronaire. Negative dominant heterozygous effects for micronaire between CS-B25 and ST474, CS-B25 and SG747, and TM-1 and DP90 were found due to chromosome 25. Positive dominant heterozygous effect for micronaire between TM-1 and ST474 was found due to chromosome 25. Negative dominant heterozygous effects for span length between CS-B25 and FM966, and TM-1 and ST474, and positive dominant effects for this trait between CS-B25 and ST474, and TM-1 and FM966 were found due to chromosome 25.

The additive effects of the remaining 25 chromosomes (A2) from TM-1 reduced lint percentage, but the remaining 25 chromosomes from the other cultivars, except DP 90, increased lint percentage. These 25 chromosomes from TM-1 and DP 90 reduced seed cotton yield and lint yield, but those from the other cultivars increased yield (Table 5).

Significant *D2* effects for seed cotton yield, lint yield, and fiber strength was detected (Table 2). Thus, there were no significant difference among all *D2* effects for lint percentage and boll weight (Table 6). In general, heterozygous dominant effects were greater than homozygous dominant effects for seed cotton yield and lint yield due to the remaining 25 chromosomes, indicating that large heterosis for cotton yield at early

Table 4 Dominant effects ± standard errors due to chromosome 25 (*D1*) for boll weight and three fiber traits based on the modified AD model

$D^{a}_{ij}(1)$	BW (g)	MIC	EL (%)	SL (mm)
1×1[b]	−0.85 ± 0.33	0.33 ± 0.08	0.08 ± 0.37	0.40 ± 0.44
2×2	0.20 ± 0.23	−0.02 ± 0.12	−0.23 ± 0.50	−1.04 ± 0.73
3×3	−0.17 ± 0.04	0.15 ± 0.08	−0.11 ± 0.20	0.01 ± 0.35
4×4	−0.17 ± 0.06	0.12 ± 0.07	−0.22 ± 0.25	−0.03 ± 0.08
5×5	−0.20 ± 0.08	0.06 ± 0.09	−0.29 ± 0.19	−0.21 ± 0.11
6×6	−0.17 ± 0.08	0.04 ± 0.04	0.28 ± 0.27	−0.37 ± 0.15
7×7	−0.16 ± 0.06	0.14 ± 0.06	0.07 ± 0.12	−0.21 ± 0.20
1×3	0.18 ± 0.15	0.18 ± 0.21	−0.05 ± 0.11	−0.45 ± 0.47
1×4	0.86 ± 0.30	−0.11 ± 0.22	−1.26 ± 0.91	−1.49 ± 0.35
1×5	0.69 ± 0.43	−0.73 ± 0.20	0.82 ± 0.51	1.54 ± 0.83
1×6	−0.22 ± 0.18	0.16 ± 0.12	0.23 ± 0.29	0.05 ± 0.45
1×7	0.00 ± 0.10	−0.50 ± 0.24	−0.12 ± 0.50	0.28 ± 0.30
2×3	−0.08 ± 0.14	−0.57 ± 0.23	0.00 ± 0.25	0.39 ± 0.69
2×4	−0.16 ± 0.21	−0.14 ± 0.17	0.97 ± 0.81	2.11 ± 0.57
2×5	−0.49 ± 0.34	0.73 ± 0.20	−0.18 ± 0.27	−1.58 ± 0.83
2×6	0.39 ± 0.25	−0.08 ± 0.10	−0.42 ± 0.28	0.75 ± 0.32
2×7	0.32 ± 0.10	0.26 ± 0.19	0.43 ± 0.63	−0.14 ± 0.38

BW Boll weight, *MIC* micronaire, *EL* elongation, *SL* 2.5% span length, and *T1* fiber strength

[a] *Dij* Dominant effects due to the chromosome 25 in cotton, if $i=j$, then *Dij* is the homozygous dominant effect, if $i \neq j$, then *Dij* is the heterozygous dominant effect

[b] *1* CS-B25, *2* TM1, *3* DP90, *4* FM966, *5* ST474, *6* PS355, and *7* SG747

Table 5 Additive effects ± standard error from remaining 25 chromosomes (*A2*) based on the modified AD model

	LP (%)	BW (g)	YLD (kg/ha)	LY (kg/ha)	EL (%)
TM1	−4.57 ± 0.06	0.14 ± 0.12	−526 ± 216	−431 ± 59	0.16 ± 0.08
DP90	−0.05 ± 0.12	−0.18 ± 0.13	−275 ± 85	−123 ± 19	−0.35 ± 0.10
FM966	1.67 ± 0.12	0.27 ± 0.21	301 ± 131	210 ± 45	−0.91 ± 0.19
ST474	1.38 ± 0.14	−0.13 ± 0.11	96 ± 63	102 ± 27	0.06 ± 0.09
PS355	0.77 ± 0.11	−0.12 ± 0.08	285 ± 140	158 ± 48	0.53 ± 0.12
SG747	0.80 ± 0.11	0.01 ± 0.01	118 ± 85	84 ± 39	0.51 ± 0.15
Grand $\bar{X}$	38.96	5.67	4687	1816	8.42

LP Lint percentage, *YLD* seed cotton yield, and *LY* lint yield

Table 6 Dominant effects ± standard error from remaining 25 chromosomes (D2) based on the modified AD model

	LP	BW	YLD	LY	T1
2×2[a]	0.40 ± 0.79	−0.61 ± 0.45	−1456 ± 431	−478 ± 273	−4.01 ± 3.69
3×3	0.04 ± 0.13	−0.16 ± 0.13	−2 ± 109	9 ± 33	8.84 ± 4.50
4×4	−0.04 ± 0.18	−0.11 ± 0.12	−138 ± 103	−3 ± 40	13.22 ± 5.61
5×5	−0.76 ± 0.87	−0.22 ± 0.14	−131 ± 143	45 ± 36	−8.41 ± 5.04
6×6	−0.29 ± 0.52	−0.17 ± 0.14	−156 ± 210	9 ± 66	−11.73 ± 3.77
7×7	−0.56 ± 0.70	−0.16 ± 0.11	−452 ± 220	−119 ± 81	−4.68 ± 2.11
2×3	−0.09 ± 0.29	0.09 ± 0.10	−286 ± 188	−133 ± 63	−14.11 ± 7.92
2×4	−0.65 ± 0.74	0.58 ± 0.51	615 ± 246	213 ± 95	−8.23 ± 9.09
2×5	0.96 ± 0.91	0.25 ± 0.16	368 ± 293	11 ± 87	11.38 ± 8.85
2×6	0.21 ± 0.55	0.19 ± 0.19	619 ± 477	139 ± 166	25.01 ± 8.80
2×7	0.77 ± 0.91	0.31 ± 0.24	1018 ± 499	308 ± 201	−7.30 ± 5.09

LP Lint percentage, *BW* boll weight, *YLD* seed cotton yield, *LY* lint yield, and *T1* fiber strength
[a] *2* TM1, *3* DP90, *4* FM966, *5* ST474, *6* PS355, and *7* SG747

generations could be expected. Homozygous dominant effects on seed cotton yield for TM-1 and SG747 were significantly negative due to the remaining 25 chromosomes, while heterozygous dominant effects on seed cotton yield between TM-1 and FM966, and TM-1 and SG747 were significantly positive due to the remaining chromosomes. High homozygous dominant effect on fiber strength due to the remaining 25 chromosomes for FM966 and low homozygous dominant effect due to the remaining 25 chromosomes for PS355 were detected. Low heterozygous dominant effect on fiber strength due to the remaining 25 chromosomes between TM-1 and DP90 was detected while high heterozygous dominant effect between TM-1 and PS355 was detected.

Discussion

Most researchers have focused on dissection of genetic effects that were whole genome-based. The genetic effects may include additive effects, dominance effects (Cockerham, 1980) and their G×E interaction effects (Zhu 1994). CS lines can bring desirable genetic resources into a recurrent parent with possibly unwanted DNA fragments from other chromosomes being minimized. On the other hand, each CS line is divergent for only one pair of chromosomes from the recurrent parent, thus, CS lines can be considered as near iso-genetic to the recurrent parent except for the substituted chromosome, and thus provide important potential for dissecting genetic effects on quantitative traits of importance. When a CS line (or several CS lines) is (or are) crossed with its (or their) recurrent parent are crossed, additive and dominant effects due to the substituted chromosome(s) can be obtained (Yen et al. 1997). Such crosses may detect the genetic differences of one specific chromosome between the donor parent and the recurrent parent; however, such genetic information is only limited to the two genotypes.

Gilbert (1985a, 1985b), Wu et al. (2000), and Lou and Zhu (2002a, 2002b) proposed several genetic models to separate single gene effects based on the AD models and other genetic models. Among these studies, the single gene effects were considered as fixed effects; however, the mixed linear model approaches such as REML and MINQUE do not estimate the variation contribution of fixed effects to the total phenotypic variance. When a number of genotypes are used as parents to top-cross with a CS line and its recurrent parent, the above methods cannot be directly used to analyze this type of data when chromosome effects (A1 vs. A2, D1 vs. D2) need to be separated. It is thus reasonable to consider the chromosome effects from different genetic resources as random. On the other hand, a mixed linear model approach can estimate the variance components and predict the genetic effects simultaneously (Searle et al. 1992). The modified AD genetic model proposed in this study, which includes the effects due to a specific chromosome (*A1* and *D1*) and the remaining chromosomes (*A2* and *D2*), was extended

from our previously proposed genetic model (Wu et al. 2000). However, this modified AD model differs from our previous one in three aspects: (1) a specific chromosome was measured in a number of genotypes (rather than two); (2) chromosome effects are random; and (3) genotype × environment interaction effects are included. Monte Carlo simulation results showed that genetic variance components were estimated with no or slight bias when we considered this modified AD model as random. The correlation coefficient between predicted effects and true effects due to the same number chromosomes varied from zero to greater than 0.90 and it was positively related to the difference between the CS line and the recurrent line. The results indicated that when a CS line is greatly different from its recurrent parent, this CS line can be efficiently used as a probe to detect the genetic effects due to the chromosome of interest in other inbred lines. Thus, this modified AD model combined with mixed linear model approach has several additional advantages: (1) it can be used to determine the chromosome association of traits of importance; (2) each CS line can be used to probe the desirable genes located on specific chromosomes when crossed with different inbred lines; and (3) once a strong genetic association with a specific chromosome is detected, researchers may focus on that specific chromosome to sequence and clone specific genes. Thus, the model proposed in this study greatly extends the use of CS lines both in genetic mapping and breeding studies. We also developed a software package, which is available upon request.

If there are no *A2* or *D2* effects for a trait (i.e. micronaire and span length), which are due to remaining 25 chromosomes, it does not mean that any other chromosomes does not have association with this trait because *A2* or *D2* are cumulative effects from the remaining 25 chromosomes, and individual chromosome effects could be positive or negative.

The mating design used in our example can be also considered as North Carolina II design. We also analyzed the data set using the traditional AD model. Comparing variance components and their proportions to the phenotypic variance based on the modified AD model and the traditional AD model, we found that the additive variance (V_A), dominance variance (V_D), and their GE interaction variance components (V_{AE} and V_{DE}) obtained by the AD model were close to the sum of the additive variances ($V_{A1} + V_{A2}$), dominance variances ($V_{D1} + V_{D2}$), and their GE interaction variances ($V_{A1E} + V_{A2E}$ and $V_{D1E} + V_{D2E}$) obtained by the modified AD genetic model (data not presented). The proportions of variance components for the traditional AD model are reported in last five rows of Table 2. The predicted additive effects obtained by the AD model are cumulative effects of chromosome 25 (*A1*) and the remaining 25 chromosomes (*A2*). Generally, the additive effects obtained by the AD model (*A*) had the similar trends with the additive effects due to *A1* + *A2* (data not shown). Above results indicated that the additive effects, dominance effects, and GE interaction effects in the AD model can be partitioned into the specific chromosome effects ($G_{(1)}$) and remaining chromosome effects ($G_{(2)}$) using the modified AD model. This showed the utility of the modified AD model and its superiority over the AD model.

Chromosome 25 of 3–79 (CS-B25) in TM-1 had an important association with several fiber traits, including decreasing micronaire, increasing fiber length and strength compared to TM-1, which agreed with our previous study (Saha et al. 2004; Jenkins et al. 2004). We also found that chromosome 25 of FM966 had similar and positive genetic associations with fiber length and strength. This indicates that the CS-B25 line with chromosome 25 from 3–79 as well as chromosome 25 from other gemplasms from *G hirsutum* can provide genes for fiber quality improvement.

The genetic model proposed in this study is extendable. For this modified AD model, F_1 or F_2 populations including a CS line and its recurrent parent test-crossed with other gemplasms and all parent lines are required. If more generations (such as at least F_1 and F_2 or F_2 and F_3) are included, a modified ADAA model may also be used to detect the additive × additive espistasis effects between a specific chromosome and the remaining chromosomes.

Chromosome substitution lines allow the net effects of a whole chromosome to be studied. Recombinant substituted (RS) inbred lines can be used to identify and map gene(s) controlling agronomic traits and fiber traits by linkage with molecular markers (Kaeppler 1997; Lander and Botstein 1989; Zeng 1993, 1994; Shah et al. 1999). RS inbred lines are superior to recombinant inbred (RI) lines for identifying genes or QTLs of quantitative traits more precisely because RS inbred lines have a more uniform genetic background with the recurrent parent with only one divergent chromosome segment rather than whole chromosome or chromosome arm. Some of these studies have been done in wheat (Chen et al. 1994; Joppa et al. 1997; Shah et al. 1999; Campell et al. 2003, 2004). Efforts are underway to develop several RS inbred populations in cotton, which can be more effectively used to locate the QTLs of cotton yield and fiber quality.

Software for analysis can be obtained from authors on request.

Acknowledgements The authors want to express their special appreciations to Drs. Daryl Bowman and Jinfa Zhang for careful peer review and helpful comments. The authors also like to thank one editor and two anonymous reviewers for their supportive comments for improving the quality of this paper.

References

Al-Quadhy RS, Morris WR, Mumm RF (1988) Chromosomal locations of genes for traits associated with lodging in winter wheat. Crop Sci 28:631–635

Berke TG, Baenziger PS, Morris WR (1992a) Chromosomal location of wheat quantitative trait loci affecting agronomic performance, using reciprocal chromosome substitutions. Crop Sci 32:621–627

Berke TG, Baenziger PS, Morris WR (1992b) Chromosomal location of wheat quantitative trait loci affecting stability of six traits, using reciprocal chromosome substations. Crop Sci 32:628–633

Campbell BT, Baenziger PS, Gill KS, Eskridge KM, Budak H, Erayman M, Dweikat I, Yen Y (2003) Identification of QTLs and environmental interactions associated with agronomic traits on chromosome 3A of wheat. Crop Sci 43:1493–1505

Campbell BT, Baenziger PS, Eskridge KM, Budak H, Streck NA, Weiss A, Gill KS, Erayman M (2004) Using environmental covariates to explain genotype × environment and QTL × environment interactions for agronomic traits on chromosome 3A of wheat. Crop Sci 44:620–627

Chen Z, Devey M, Tuleen NA, Hart GE (1994) Use of recombinant substitution lines in the construction of RFLP-based genetic maps of chromosomes 6A and 6B of tetraploid wheat (*Triticum turgidum* L.). Theor Appl Genet 89:703–712

Cockerham CC (1980). Random and fixed effect in plant genetics. Theor Appl Genet 56:119–131

Devroye L (1986) Non-uniform random variate generation. Berlin Heidelberg NewYork

Gilbert DG (1985b) Estimating single gene effects on quantitative traits. II. Statistical properties of five experimental methods. Theor Appl Genet 69:631–636

Gilbert DG (1985a) Estimating single gene effects on quantitative traits. I. A diallel method applied to Est 6 in *D. melanogaster*. Theor Appl Genet 69:625–629

Jenkins JN, McCarty MC, Wu J, Saha S, Stelly D (2004). Chromosome substitution lines from *Gossypium barbadense* L. as sources for *G. hirsutum* L. improvement. In: 4th International crop science congress. Brisbane Convention & Exhibition Centre, Queensland

Joppa LR, Changheng D, Hart GE, Hareland GA (1997) Mapping gene(s) for grain protein in tetraploid wheat (*Triticum turgidum* L.) using a population of recombinant inbred chromosome lines. Crop Sci 37:1586–1589

Kaeppler SM (1997) Quantitative trait locus mapping using set of near-isogenic lines: relative power comparisons and technical considerations. Theor Appl Genet 95:384–392

Kohel RJ, Richmond TR, Lewis CF (1970) Texas marker-1. Description of a genetic standard for *Gossypium hirsutum* L. Crop Sci 10:670–671

Kohel RJ, Endrizzi JE, White TG (1977) An evaluation of Gossyppium barbadense L. chromosome 6 and 17 in the *G. hirsutum* L. genome. Crop Sci 17:404–406

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–191

Law CN (1966) The location of genetic factors affecting a quantitative character in wheat. Genetics 53:487–498

Law CN (1967) The location of genetic factors controlling a number of quantitative characters in wheat. Genetics 56:445–461

Law CN, Worland AJ, Giorgi B (1976) The genetic control of ear-emergence time by chromosome 5A and 5D in wheat. Heredity 36:49–58

Lou XY, Zhu J (2002a) Analysis of genetic effects of major genes and polygenes on quantitative traits: I. Genetic model for diploid plants and animals. Theor Appl Genet 104:414–421

Lou XY, Zhu J (2002b) Analysis of genetic effects of major genes and polygenes on quantitative traits: II. Genetic model for seed traits of crops. Theor Appl Genet 105:964–971

Ma JZ, Kohel RJ (1983) Evaluation of 6 substitution lines in cotton. Acta Agron Sinica 9:145–150

Mansur LM, Qualset CO, Kasarda DD, Morris R (1990) Effects of 'Cheyenne' chromosomes on milling and baking quality in 'Chinese Spring' wheat in relation to glutenin and gliadin storage protein. Crop Sci 30:593–602

Miller RG (1974) The jackknife: a review. Biometrika 61:1–15

Patterson HD, Thompson R (1971) Recovery of inter-block information when block sizes are unequal. Biometrika 58:545–554

Rao CR (1971) Estimation of variance and covariance components MINQUE theory. J Multivariate Anal 1:257–275

Ren L, Guo W, Zhang T (2002) Identification of quantitative trait loci (QTLs) affecting yield and fiber properties in chromosome 16 in cotton using substitution line. Acta Botanica Sinica 44:815–820

Saha S, Jenkins JN, McCarty JC, Wu J, Gutiérrez OA, Cantrell RG, Percy RG, Raska DA, Stelly DM (2003) Chromosomal effect on fiber and agronomic traits utilizing *Gossypium barbadense* chromosome-specific F_2hybrid population of *G. hirsutum*. Would Cotton Conference III, South Africa

Saha S, Wu J, Jenkins JN, McCarty JC, Gutiérrez OA, Percy RG, Raska DA, Stelly DM (2004) Association of agronomic and fiber traits with specific Pima 3–79 chromosomes in a TM-1 Background. J Cotton Sci 8:162–169

Searle SR, Casella G, McCulloch CE (1992) Variance Components. Wiley, NewYork

Shah MM, Gill KS, Baenziger PS, Yen Y, Kaeooler SM, Ariyarathne HM (1999) Molecular mapping of loci for agronomic traits on chromosome 3A of bread wheat. Crop Sci 39:1728–1932

Wu J, Jenkins JN, McCarty JC, Cheatham C (2000) Separation of single genetic effects from additive-dominance genetic models. Proc Conf Appl Stat in Agri In Kansas State University, pp 239–249

Yen Y, Baenziger PS, Bruns R, Reeder J, Moreno-Sevilla B, Budak N (1997) Agronomic performance of hybrids between cultivars and chromosome substitution lines. Crop Sci 37:396–399

Zemetra RS, Morris WR (1988) Effects of an intercultivaral chromosome substitution on winetrhardiness and vernalization in wheat. Genetics 119:453–456

Zemetra RS, Morris WR, Schmidt JW (1986) Gene locations for heading date using reciprocal chromosome substitutions in winter wheat. Crop Sci 26:531–533

Zeng ZB (1993) Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proc Natl Acad Sci USA 90:10972–10976

Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1468

Zhu J (1989) Estimation of genetic variance components in the general mixed model. PhD Dissertation, North Carolina State University, Raleigh, U.S.A.

Zhu J (1993) Methods of predicting genotype value and heterosis for offspring of hybrids (Chinese). J Biomath 8(1):32–44

Zhu J (1994) General genetic models and new analysis methods for quantitative traits. J Zhejianiang Agri Univ 20(6):551–559

Zhu J, Weir BS (1994) Analysis of cytoplasmic and maternal effects I. A genetic model for diploid plant seeds and animals. Theor Appl Genet 89:153–159