

Analysis of Conditional Genetic Effects and Variance Components in Developmental Genetics

Jun Zhu

Agronomy Department, Zhejiang Agricultural University, Hangzhou 310029, China

Manuscript received March 13, 1995

Accepted for publication September 5, 1995

ABSTRACT

A genetic model with additive-dominance effects and genotype \times environment interactions is presented for quantitative traits with time-dependent measures. The genetic model for phenotypic means at time t conditional on phenotypic means measured at previous time $(t - 1)$ is defined. Statistical methods are proposed for analyzing conditional genetic effects and conditional genetic variance components. Conditional variances can be estimated by minimum norm quadratic unbiased estimation (MINQUE) method. An adjusted unbiased prediction (AUP) procedure is suggested for predicting conditional genetic effects. A worked example from cotton fruiting data is given for comparison of unconditional and conditional genetic variances and additive effects.

UNDERSTANDING gene expression is one of the major goals in developmental genetics. The influences of a single gene or several major genes on qualitative traits have been extensively studied in developmental genetics. The phenotypic variation for many important traits is essentially continuous and is influenced by both quantitative genes and environment effects. Therefore quantitative variation for time-dependent measures is difficult to interpret in developmental genetics.

For quantitative traits of developmental behavior, the genetic effect ($G_{(t)}$) at time t is the sum of the genetic effect ($G_{(t-1)}$) at time $(t - 1)$ and the extra genetic effect ($G_{(d)}$), which is most likely to be correlated with ($G_{(t-1)}$). Quantitative variation of a population is usually analyzed by genetic variance components for traits measured at different times. Genetic covariances or correlations between traits at time $(t - 1)$ and time t are also estimated sometimes. These analyses can provide inferences for the cumulative gene effects from initial time to t but not for the independent effects of gene expression in the period $(t - 1)$ to t .

By univariate or multivariate analysis, conditional random variables with conditional variances have been used for detecting effects conditional on specific situations. Estimation methods for conditional variances were proposed for analyzing autoregressive conditional heteroscedastic (ARCH) model in economics (ENGLE 1982) and also for estimating genetic variances conditional on selected base populations in animal breeding (HENDERSON 1985, 1988). ENGLE's method, which explicitly parameterizes the time dependency of the conditional variance for the error process of time series models, cannot estimate conditional variance components due to random genetic effects other than noise. Although HENDERSON's method can estimate condi-

tional genetic variance given selected base populations, it is not applicable to estimating conditional genetic variance components for time-dependent measures in developmental quantitative genetics.

In this paper statistical methods are proposed for analyzing conditional genetic effects and conditional genetic variance components based on genetic models for quantitative traits with time-dependent measures. Cotton data for fruiting behavior within the whole blooming period in 2 yr is analyzed as an example to illustrate the use of these new methods.

GENETIC MODEL FOR QUANTITATIVE TRAITS WITH TIME-DEPENDENT MEASURES

Among various genetic mating designs, the diallel design (HALLAUER and MIRANDA 1981) is the one widely used by plant and animal breeders for analyzing quantitative traits and also for developmental behavior studies (HENDERSON 1990; ZHU *et al.* 1993a, 1993b).

Based on the diallel model with additive-dominance effects and genotype \times environment interactions, the phenotypic mean of a trait measured at time t for the cross between inbred lines i and j in the k th block within environment h can be partitioned as

$$y_{hijk(t)} = E_{h(t)} + A_{i(t)} + A_{j(t)} + D_{ij(t)} + AE_{hi(t)} + AE_{hj(t)} + DE_{hij(t)} + B_{hk(t)} + \epsilon_{hijk(t)}, \quad (1)$$

where $y_{hijk(t)}$ is the phenotypic mean at time t for genetic entry $i \times j$ in block k within environment h ; $E_{h(t)}$ is a fixed effect at time t for environment h ; $A_{i(t)}$ is the additive effect of genes from parental line i at time t , $A_{i(t)} \sim (0, \sigma_{A(i)}^2)$; $A_{j(t)}$ is the additive effect of genes from parental line j at time t , $A_{j(t)} \sim (0, \sigma_{A(j)}^2)$; $D_{ij(t)}$ is the dominance effect of genes from the cross of $i \times j$ at time t , $D_{ij(t)} \sim (0, \sigma_{D(ij)}^2)$; $AE_{hi(t)}$ is the interaction effect

of $A_{i(t)} \times E_{h(t)}$ at time t , $AE_{hi(t)} \sim (0, \sigma_{AE(t)}^2)$; $AE_{hj(t)}$ is the interaction effect of $A_{j(t)} \times E_{h(t)}$ at time t , $AE_{hj(t)} \sim (0, \sigma_{AE(t)}^2)$; $DE_{hij(t)}$ is the interaction effect of $D_{ij(t)} \times E_{h(t)}$ at time t , $DE_{hij(t)} \sim (0, \sigma_{DE(t)}^2)$; $B_{hk(t)}$ is random effect of block k within environment h at time t , $B_{hk(t)} \sim (0, \sigma_{B(t)}^2)$; and $\epsilon_{ijk(t)}$ is a residual effect at time t , $\epsilon_{ijk(t)} \sim (0, \sigma_{\epsilon(t)}^2)$.

The additive-dominance model can be written in a matrix form of the mixed linear model,

$$\begin{aligned} \mathbf{y}_{(t)} &= \mathbf{X}\mathbf{b}_{(t)} + \mathbf{U}_A\mathbf{e}_{A(t)} + \mathbf{U}_D\mathbf{e}_{D(t)} + \mathbf{U}_{AE}\mathbf{e}_{AE(t)} \\ &\quad + \mathbf{U}_{DE}\mathbf{e}_{DE(t)} + \mathbf{U}_B\mathbf{e}_{B(t)} + \mathbf{e}_{\epsilon(t)} \\ &= \mathbf{X}\mathbf{b}_{(t)} + \sum_{u=1}^6 \mathbf{U}_u\mathbf{e}_{u(t)}, \end{aligned} \tag{2}$$

where $\mathbf{y}_{(t)}$ is the vector of phenotypic mean for all entries of the mating design; $\mathbf{b}_{(t)}$ is the vector of fixed environment effects at time t ; \mathbf{X} is the known incidence matrix with coefficients 1 or 0 relating to the fixed environment effects; $\mathbf{e}_{u(t)}$ is the vector of random effects at time t , $\mathbf{e}_{u(t)} \sim (0, \sigma_{u(t)}^2\mathbf{I})$; \mathbf{U}_u is the known coefficient matrix relating to the random vector $\mathbf{e}_{u(t)}$; \mathbf{U}'_u is the transpose of \mathbf{U}_u ; $\mathbf{U}_6 = \mathbf{I}$ is an identity matrix.

The vector of random variables $\mathbf{y}_{(t)}$ has variance-covariance matrix

$$\begin{aligned} \text{Var}(\mathbf{y}_{(t)}) &= \sigma_{A(t)}^2 \mathbf{U}_A \mathbf{U}'_A + \sigma_{D(t)}^2 \mathbf{U}_D \mathbf{U}'_D + \sigma_{AE(t)}^2 \mathbf{U}_{AE} \mathbf{U}'_{AE} \\ &\quad + \sigma_{DE(t)}^2 \mathbf{U}_{DE} \mathbf{U}'_{DE} + \sigma_{B(t)}^2 \mathbf{U}_B \mathbf{U}'_B + \sigma_{\epsilon(t)}^2 \mathbf{I} \\ &= \sum_{u=1}^6 \sigma_{u(t)}^2 \mathbf{U}_u \mathbf{U}'_u \\ &= \mathbf{V}_{(t)}. \end{aligned}$$

For developmental behavior traits, quantitative genetic effects at time t are not independent of those at time $(t - 1)$, $\text{Cov}(\mathbf{e}_{u(t-1)}, \mathbf{e}'_{u(t)}) = \sigma_{u(t-1,t)} \mathbf{U}_u \mathbf{U}'_u$. Therefore, phenotypic means of quantitative traits measured at time t will be correlated with the means measured at time $(t - 1)$,

$$\begin{aligned} \text{Cov}(\mathbf{y}_{(t-1)}, \mathbf{y}'_{(t)}) &= \sigma_{A(t-1,t)} \mathbf{U}_A \mathbf{U}'_A + \sigma_{D(t-1,t)} \mathbf{U}_D \mathbf{U}'_D \\ &\quad + \sigma_{AE(t-1,t)} \mathbf{U}_{AE} \mathbf{U}'_{AE} + \sigma_{DE(t-1,t)} \mathbf{U}_{DE} \mathbf{U}'_{DE} \\ &\quad + \sigma_{B(t-1,t)} \mathbf{U}_B \mathbf{U}'_B + \sigma_{\epsilon(t-1,t)} \mathbf{I} \\ &= \sum_{u=1}^6 \sigma_{u(t-1,t)} \mathbf{U}_u \mathbf{U}'_u \\ &= \mathbf{C}_{(t-1,t)}, \end{aligned}$$

where $\mathbf{C}_{(t-1,t)}$ is symmetric ($\mathbf{C}_{(t-1,t)} = \mathbf{C}'_{(t-1,t)}$) and identical to $\mathbf{C}_{(t,t-1)}$.

Differences between observed means at sequential times ($\mathbf{y}_{(d)} = \mathbf{y}_{(t)} - \mathbf{y}_{(t-1)}$) are sometimes used for studying quantitative genetic effects on developmental behavior during time $(t - 1)$ to t . Analyses based on $\mathbf{y}_{(d)}$ will not give net genetic variation within the period $(t - 1)$ to t , since

$$\begin{aligned} \text{Var}(\mathbf{y}_{(d)}) &= \sum_{u=1}^6 \sigma_{u(d)}^2 \mathbf{U}_u \mathbf{U}'_u \\ &= \sum_{u=1}^6 (\sigma_{u(t)}^2 - 2\sigma_{u(t-1,t)} + \sigma_{u(t-1)}^2) \mathbf{U}_u \mathbf{U}'_u \\ &= \mathbf{V}_{(t)} - 2\mathbf{C}_{(t-1,t)} + \mathbf{V}_{(t-1)}, \end{aligned}$$

and $\mathbf{y}_{(d)}$ is not independent to $\mathbf{y}_{(t-1)}$, $\text{Cov}(\mathbf{y}_{(d)}, \mathbf{y}'_{(t-1)}) = \mathbf{C}_{(t-1,t)} - \mathbf{V}_{(t-1)}$.

Instead of using differences between two random variables, conditional random variables can be used for detecting conditional genetic effects as well as conditional variance components. The genetic model for phenotypic means at time t conditional on phenotypic means measured at time $(t - 1)$, $(\mathbf{y}_{(t)} | \mathbf{y}_{(t-1)})$, is defined as

$$\begin{aligned} \mathbf{y}_{(t)} | \mathbf{y}_{(t-1)} &= \mathbf{X}\mathbf{b}_{(t|t-1)} + \mathbf{U}_A\mathbf{e}_{A(t|t-1)} + \mathbf{U}_D\mathbf{e}_{D(t|t-1)} \\ &\quad + \mathbf{U}_{AE}\mathbf{e}_{AE(t|t-1)} + \mathbf{U}_{DE}\mathbf{e}_{DE(t|t-1)} \\ &\quad + \mathbf{U}_B\mathbf{e}_{B(t|t-1)} + \mathbf{e}_{\epsilon(t|t-1)} \\ &= \mathbf{X}\mathbf{b}_{(t|t-1)} + \sum_{u=1}^6 \mathbf{U}_u\mathbf{e}_{u(t|t-1)} \end{aligned}$$

with the conditional variance-covariance matrix

$$\begin{aligned} \text{Var}(\mathbf{y}_{(t|t-1)}) &= \sigma_{A(t|t-1)}^2 \mathbf{U}_A \mathbf{U}'_A + \sigma_{D(t|t-1)}^2 \mathbf{U}_D \mathbf{U}'_D \\ &\quad + \sigma_{AE(t|t-1)}^2 \mathbf{U}_{AE} \mathbf{U}'_{AE} + \sigma_{DE(t|t-1)}^2 \mathbf{U}_{DE} \mathbf{U}'_{DE} \\ &\quad + \sigma_{B(t|t-1)}^2 \mathbf{U}_B \mathbf{U}'_B + \sigma_{\epsilon(t|t-1)}^2 \mathbf{I} \\ &= \sum_{u=1}^6 \sigma_{u(t|t-1)}^2 \mathbf{U}_u \mathbf{U}'_u \\ &= \mathbf{V}_{(t|t-1)}, \end{aligned}$$

where $\mathbf{b}_{(t|t-1)}$ is the vector of conditional environment effects at time t , and $\mathbf{e}_{u(t|t-1)}$ is the vector of conditional random effects at time t , $\mathbf{e}_{u(t|t-1)} \sim (0, \sigma_{u(t|t-1)}^2 \mathbf{I})$.

Conditional genetic effects at time t are independent of the genetic effects at time $(t - 1)$, $\text{Cov}(\mathbf{e}_{u(t|t-1)}, \mathbf{e}'_{u(t-1)}) = 0$. Conditional genetic effects $\mathbf{e}_{u(t|t-1)}$ and the conditional variance $\sigma_{u(t|t-1)}^2$ can be used in developmental genetic analyses for illuminating the independent effects of gene expression within the period $(t - 1)$ to t .

METHODS FOR ANALYZING CONDITIONAL GENETIC EFFECTS AND VARIANCE COMPONENTS

The conditional variance-covariance matrix can be directly estimated by the following equation (KRZANOWSKI 1988; BOX *et al.* 1994):

$$\hat{\mathbf{V}}_{(t|t-1)} = \hat{\mathbf{V}}_{(t)} - \hat{\mathbf{C}}_{(t-1,t)} \hat{\mathbf{V}}_{(t-1)}^{-1} \hat{\mathbf{C}}_{(t-1,t)}. \tag{3}$$

Estimated variance components $\hat{\sigma}_u^2$ in $\hat{\mathbf{V}}_{(t)}$ and $\hat{\mathbf{V}}_{(t-1)}$ can be obtained by method of minimum norm quadratic unbiased estimation (MINQUE) (RAO 1970, 1971) or restricted maximum likelihood (REML) (PATTERSON

and THOMPSON 1971; CORBEIL and SEARLE 1976). Estimated covariance components $\hat{\sigma}_{u(t-1,0)}$ in $\hat{\mathbf{C}}_{(t-1,0)}$ can be calculated by the method suggested by ZHU and WEIR (1994a, 1995).

Estimation of conditional genetic variance components ($\sigma_{u(t-1)}^2$) or prediction of conditional genetic effects ($\mathbf{e}_{u(t-1)}$) cannot be derived directly from the estimate of conditional variance-covariance matrix ($\hat{\mathbf{V}}_{(t-1)}$). The indirect approaches are suggested for analyzing conditional genetic effects and their variance components. For experiments on developmental quantitative behavior with phenotypic means measured at time series ($t - 1$) and t , a new random vector \mathbf{y}_* is defined as

$$\mathbf{y}_* = \mathbf{y}_{(t)} - \mathbf{C}_{(t-1,0)} \mathbf{V}_{(t-1)}^{-1} (\mathbf{y}_{(t-1)} - \mathbf{X} \mathbf{b}_{(t-1)}). \quad (4)$$

Random vector \mathbf{y}_* has variance $\text{Var}(\mathbf{y}_*) = \mathbf{V}_{(t)} - \mathbf{C}_{(t-1,0)} \times \mathbf{V}_{(t-1)}^{-1} \mathbf{C}_{(t-1,0)}$, which is identical to the conditional variance-covariance matrix $\mathbf{V}_{(t|t-1)}$. In practice, unknown parameters in Equation 4 can be replaced by their unbiased estimates.

If random vector \mathbf{y}_* is fitted to the additive-dominance model as in Equation 2,

$$\begin{aligned} \mathbf{y}_* &= \mathbf{X} \mathbf{b}_* + \mathbf{U}_A \mathbf{e}_{A^*} + \mathbf{U}_D \mathbf{e}_{D^*} + \mathbf{U}_{AE} \mathbf{e}_{AE^*} \\ &\quad + \mathbf{U}_{DE} \mathbf{e}_{DE^*} + \mathbf{U}_B \mathbf{e}_{B^*} + \mathbf{e}_{c^*} \\ &= \mathbf{X} \mathbf{b}_* + \sum_{u=1}^6 \mathbf{U}_u \mathbf{e}_{u^*} \end{aligned} \quad (5)$$

with variance-covariance matrix

$$\begin{aligned} \text{Var}(\mathbf{y}_*) &= \sigma_{A^*}^2 \mathbf{U}_A \mathbf{U}_A' + \sigma_{D^*}^2 \mathbf{U}_D \mathbf{U}_D' + \sigma_{AE^*}^2 \mathbf{U}_{AE} \mathbf{U}_{AE}' \\ &\quad + \sigma_{DE^*}^2 \mathbf{U}_{DE} \mathbf{U}_{DE}' + \sigma_{B^*}^2 \mathbf{U}_B \mathbf{U}_B' + \sigma_{c^*}^2 \mathbf{I} \\ &= \sum_{u=1}^6 \sigma_{u^*}^2 \mathbf{U}_u \mathbf{U}_u' \\ &= \mathbf{V}_*, \end{aligned}$$

which turns out to be the conditional variance-covariance matrix $\mathbf{V}_{(t|t-1)}$.

Unbiased estimations for variance components $\sigma_{u^*}^2$ are obtainable by using the mixed model approaches (ZHU *et al.* 1993c; ZHU and WEIR 1994a, 1995). Therefore, $\hat{\sigma}_{u^*}^2$ can be regarded as unbiased estimates for the conditional variance components $\sigma_{u(t-1)}^2$. Since $\sigma_{u^*}^2$ are equivalent to $\sigma_{u(t-1)}^2$, genetic effects \mathbf{e}_{u^*} also have an equivalency to the conditional genetic effects $\mathbf{e}_{u(t-1)}$. Conditional genetic effects $\mathbf{e}_{u(t-1)}$ (or \mathbf{e}_{u^*}) can be predicted by analyzing \mathbf{y}_* with several methods. The best linear unbiased prediction (BLUP) (HENDERSON 1963) and linear unbiased prediction (LUP) (ZHU *et al.* 1993c; ZHU and WEIR 1994a) tend to give unbiased estimates for means but underestimated variances for predicted genetic effects (ZHU 1993; ZHU and WEIR 1995). An adjusted unbiased prediction (AUP) is therefore suggested for predicting the conditional genetic effects (ZHU 1993; ZHU and WEIR 1995).

WORKED EXAMPLE OF DEVELOPMENTAL ANALYSES FOR COTTON FRUITING BEHAVIOR

Number of bolls per plant is one of the major yield components of Upland cotton (*Gossypium hirsutum* L.). There is a high correlation between number of bolls and lint yield for cultivars and selected lines of Upland cotton (WOODWARD and MALM 1976). It was indicated, by comparing obsolete and modern cultivars, that modern cultivars produced more bolls during the early growth season (WELLS and MEREDITH 1984). The final number of bolls depends on the number of bolls per plant setting in different stages of the blooming period (from early July to early September).

A genetic analysis was conducted for fruiting behavior of Upland cotton within the whole blooming period. Four parental lines were used for diallel mating in 1981 and 1985. Four parents and their six F₁ hybrids and six reciprocal crosses were included in 1981. But only four parents and their six F₁ hybrids were included in 1985. Randomized complete block design with three replications was used for both years. From July 1 to September 3 in each year, bolls per plant were counted for every 5 days in 10 plants from each plot. Cell means of these 2-yr data were analyzed by the additive-dominance model including genotype \times environment interactions. Unconditional analyses were conducted for genetic effects and their variances with observed mean \mathbf{y}_t and difference between observed means at sequential times $\mathbf{y}_{(t)}$ = $\mathbf{y}_{(t)} - \mathbf{y}_{(t-1)}$. The gap from time ($t - 1$) to time t is 5 days. Two kinds of conditional data, $\mathbf{y}_{(t|t-1)}$ for phenotypic means at time t conditional on means at time ($t - 1$) and $\mathbf{y}_{(t)}$ for final observation means conditional on means at different previous time t , were analyzed.

Variances were estimated by MINQUE(1) method, and genetic effects were predicted by AUP method (ZHU 1993; ZHU and WEIR 1995). Jackknifing over blocks within year was used for estimates and standard errors of variances and also for predictors and standard errors of genetic effects (MILLER 1974; ZHU 1989). There were six blocks for 2 years, and the degrees of freedom were 5 for the jackknifing. A t -test was employed for testing significance of genetic parameters studied. When estimated unconditional variances $\hat{\sigma}_u^2$ were obtained, unconditional genetic variance components were estimated by $\hat{V}_A = 2\hat{\sigma}_A^2$, $\hat{V}_D = \hat{\sigma}_D^2$, $\hat{V}_{AE} = 2\hat{\sigma}_{AE}^2$, and $\hat{V}_{DE} = \hat{\sigma}_{DE}^2$. Estimates of genetic variance components for difference and conditional variables were also calculated in the same way. All the data analyses were conducted by a PC computer with programs written in C.

Estimates of unconditional genetic variance components (in Figure 1, A and B) and conditional genetic variance components (in Figure 2, A and B) were presented for bolls per plant in 13 developmental stages of the blooming period. It was indicated by the unconditional genetic variance components (Figure 1A) that genes had different genetic effects on number of bolls

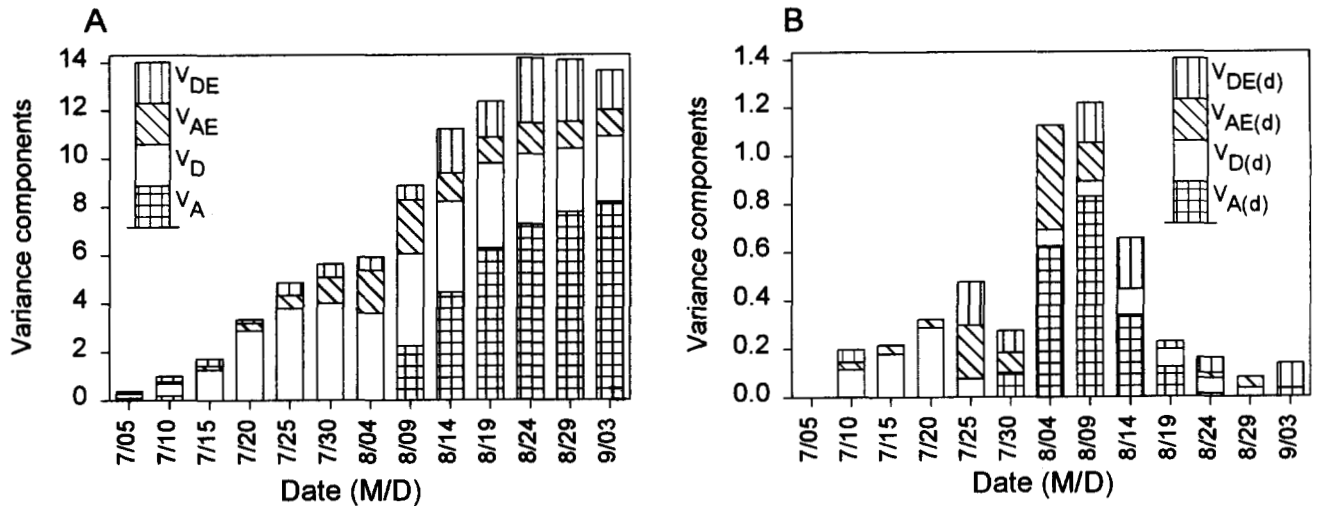


FIGURE 1.—Variance components for bolls per plant of Upland cotton. (A) Estimates of variance components for cumulative data y_t : V_A = additive variance, V_D = dominance variance, V_{AE} = additive \times environment interaction variance, V_{DE} = dominance \times environment interaction variance. (B) Estimates of variance components for difference between observed means at sequential times $y_{(d)} = y_{(t)} - y_{(t-1)}$: $V_{A(d)}$ = additive variance, $V_{D(d)}$ = dominance variance, $V_{AE(d)}$ = additive \times environment interaction variance, $V_{DE(d)}$ = dominance \times environment interaction variance.

per plant in different developmental stages. Compared to genotypic variance ($V_G = V_A + V_D$) genotype \times environment interaction variance ($V_{GE} = V_{AE} + V_{DE}$) was relatively small.

At the initial stage 7/05, V_D (0.17 ± 0.04) and V_A (0.11 ± 0.03) were two major components to the total variation. V_A was then decreased rapidly to zero on 7/15. Additive variation was not detected during the period from 7/15 to 8/04. Additive variation increased steadily later on. V_D increased up to 7/30 (4.00 ± 1.04) and then decreased slowly. The major contribution of genetic variation for number of bolls per plant was due

to dominance variance component in early period of blooming but due to additive variance component in late period of blooming.

Figure 1B showed genetic variance components for bolls per plant setting in 5 days between times t and $t - 1$. Additive variation $V_{A(d)}$ was not detected from 7/10 to 7/25 but was found during 7/30 (0.10 ± 0.03) to 8/19 (0.12 ± 0.02) with peak at 8/09 (0.83 ± 0.00). Dominance variation $V_{D(d)}$ was the major component and increased until 7/20 (0.29 ± 0.01). Large interaction variances $V_{AE(d)}$ and/or $V_{DE(d)}$ were observed from 7/25 to 8/14. Since variance of differences between

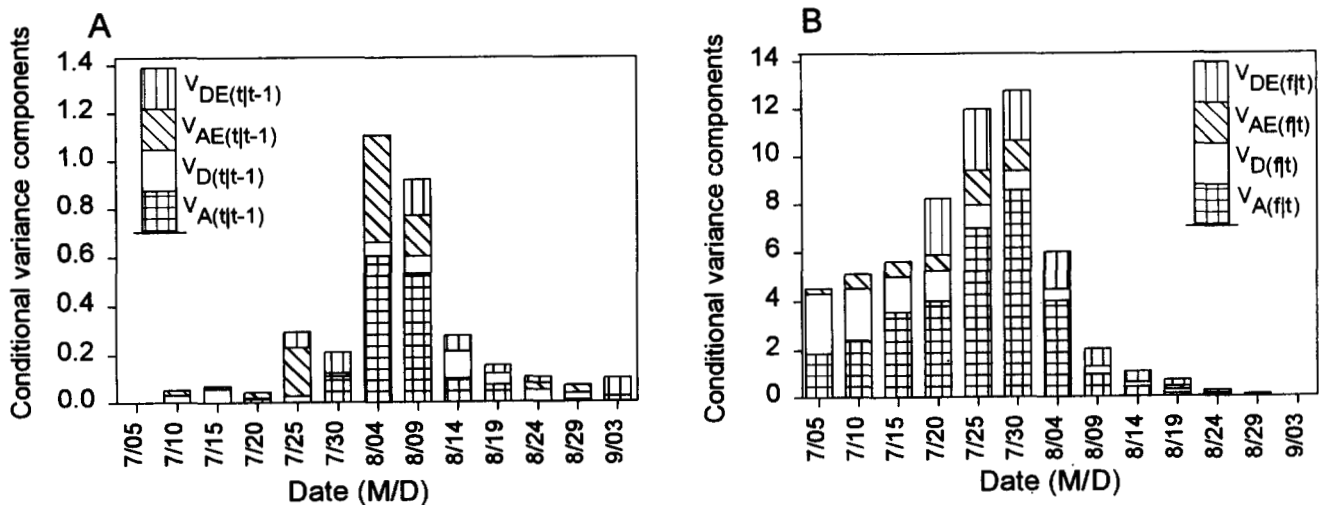


FIGURE 2.—Conditional variance components for bolls per plant of Upland cotton. (A) Estimates of conditional variance components for phenotypic means at time t conditional on means at time $(t - 1)$: $V_{A(t|t-1)}$ = conditional additive variance, $V_{D(t|t-1)}$ = conditional dominance variance, $V_{AE(t|t-1)}$ = conditional additive \times environment interaction variance, $V_{DE(t|t-1)}$ = conditional dominance \times environment interaction variance. (B) Estimates of conditional variance components for final observation means conditional on means at different previous time t : $V_{A(f|t)}$ = conditional additive variance, $V_{D(f|t)}$ = conditional dominance variance, $V_{AE(f|t)}$ = conditional additive \times environment interaction variance, $V_{DE(f|t)}$ = conditional dominance \times environment interaction variance.

sequential random effects $V_{(t|t-1)}$ consists of components for both variances and covariance, it does not show the real variation due to the net effects within time t and time $t - 1$. Therefore, the changes in genetic variance components cannot be appropriately explained by the results of unconditional analysis for developmental behavior of cotton fruiting.

The analysis results for conditional variance components $V_{(t|t-1)}$ was presented in Figure 2A. After the initial stage (7/05), conditional variation was not detected for conditional additive effects $A_{(t|t-1)}$ but was significant for conditional dominance effects $D_{(t|t-1)}$ from 7/10 ($(V_{D(t|t-1)} = 0.03 \pm 0.01)$) to 7/15 ($(V_{D(t|t-1)} = 0.06 \pm 0.02)$). These induced the increase of dominance variation but the decrease of additive variation of cotton bolls in the early developmental stages. No significant $V_{D(t|t-1)}$ was detected since 7/20, which indicated the turning off of the gene expression for dominance effects. But unconditional dominance variances for both cumulative effects V_D and difference effects $V_{D(d)}$ were still large and significant at 7/20. Conditional additive variances $V_{A(t|t-1)}$ were found significantly for number of bolls per plant from 7/30 (0.11 ± 0.06) to 8/19 (0.07 ± 0.03), indicating the turning on of the gene expression for additive effects for this period. The most dramatic increase of $V_{A(t|t-1)}$ was found at 8/04 (0.60 ± 0.12), when unconditional V_A was not even detectable and $V_{A(d)}$ was not the largest one. The dynamic changes of gene expression resulted in the decrease of dominance variation and the increase of additive variation for boll setting during later developmental stages.

Conditional additive variation was observed 10 days before unconditional additive variances V_A were detectable. This implies that expression of quantitative genes might start several days before the accumulated genetic effects for developmental behavior can be detected by unconditional genetic analysis.

Conditional variance $V_{A(t|t-1)}$ and $V_{AE(t|t-1)}$ accounted for $\sim 70\%$ of total variation for number of bolls per plant on 8/04. But no significance was observed for these unconditional variance at that time. $V_{AE(t|t-1)}$ was found to be significantly large on 7/25 (0.20 ± 0.08), 8/04 (0.44 ± 0.17) and 8/09 (0.17 ± 0.07). The difference of environment effects in 2 yr could influence the gene expression for additive effects of parents and F_1 hybrids in these two stages.

Analysis results were presented in Figure 2B for conditional genetic variances ($V_{(f)}$) of the final number of bolls per plant ($f = 9/03$) conditional on observed number of bolls per plant at different previous time t . As compared with the cumulative variance components in Figure 1A on 9/03, conditional variance components enlarged from 7/05 to 7/30 and then diminished rapidly. After 8/14 conditional variances $V_{(9/03|8/14\sim 29)}$ were very small. It was implied that gene expression contributed to the final number of bolls was negligible after 8/14. It was also suggested by the largest conditional

variances $V_{(9/03|7/30)}$ that gene expression in later July (7/25 \sim 7/30) had little influence on genetic variation of final number of bolls per plant. It was found by the conditional analysis that the period from 7/30 to 8/09 was a critical time for bolls setting with lots of new gene expression (large $V_{(t|t-1)}$) and enormous influence on final number of bolls (small $V_{(f)}$).

Predicted unconditional and conditional additive effects for four parents were presented in Figures 3 and 4, respectively. As compared to the other three parents, the early-season parent (P_3) had more bolls per plant in early July but less later on. Both P_2 and P_4 belonged to the later-season parents with similar phenotypic means during whole developmental stages.

Unconditional additive effects were positive for A_3 but negative for A_2 until 7/10. Additive effects were not found for all the parents from 7/15 to 8/04, but later on decreased A_3 and increased A_2 and A_4 were detectable while A_1 still remained near zero. Unconditional additive effects of 5-day bolls setting ($A_{(d)}$) and conditional additive effects ($A_{(t|t-1)}$) for three parents were already detected on 7/30 and 8/04. Therefore, quantitative genes expressed positive additive effects for P_2 and P_4 but negative for P_3 in 10 days before they can be detected by unconditional genetic analysis for cumulative data. The dissimilarity of boll-setting patterns for P_2 and P_4 in mid- and later stages of development can be detected by conditional additive effects $A_{(t|t-1)}$. Both $A_{2(t|t-1)}$ and $A_{4(t|t-1)}$ were positive on 7/30 and 8/04. After that time, $A_{2(t|t-1)}$ continued to have positive effects, but $A_{4(t|t-1)}$ had negative effects.

The predicted conditional additive effects $A_{(f)}$ could further reveal that gene expression of parents in the blooming period could have different effects on the final number of bolls per plant. Since conditional additive effects $A_{(9/03|7/30)}$ had similar magnitude as unconditional additive effects of final number of bolls ($A_{(9/03)}$) for all four parents, gene expression during 7/25 to 7/30 would not affect the additive variation of parents for final number of bolls. For the later-season parents, large influence of gene expression was detected for P_4 in 8/04 \sim 8/19 but for P_2 before 7/25 and in 8/09 \sim 8/19. For the early-season parent (P_3), gene expression had effects on additive variation of final number of bolls in two periods (before 7/20 and in 8/04 \sim 8/14).

This study shows that conditional analyses of genetic effects and genetic variance components are more powerful than unconditional analyses for finding gene expression of quantitative traits of developmental behaviors.

DISCUSSION

Genetic behavior measured at time t is the confounded result of genes expressed before time $(t - 1)$ and within the period $(t - 1)$ to t . Genetic variation

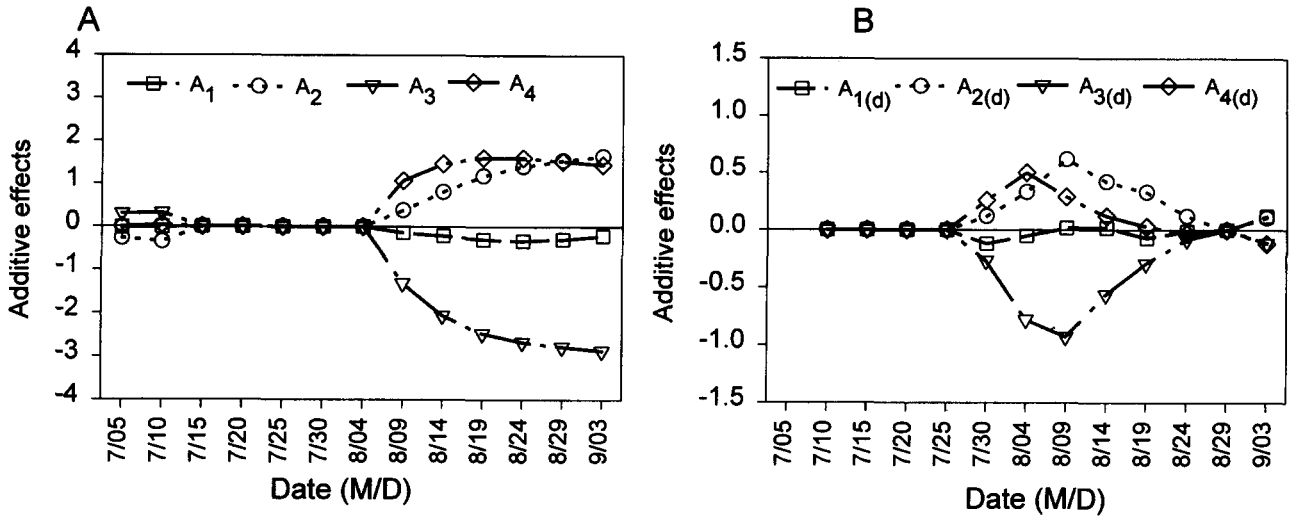


FIGURE 3.—Additive effects of four parents for bolls per plant of Upland cotton. (A) Predicted additive effects for cumulative data y_i : A_1 = additive effect of parent one, A_2 = additive effect of parent two, A_3 = additive effect of parent three, A_4 = additive effect of parent four. (B) Predicted additive effects for difference between observed means at sequential times $y_{(d)} = y_{(t)} - y_{(t-1)}$: $A_{1(d)}$ = additive effect of parent one, $A_{2(d)}$ = additive effect of parent two, $A_{3(d)}$ = additive effect of parent three, $A_{4(d)}$ = additive effect of parent four.

revealed by analyzing developmental behavior at specific stages could provide inference only for accumulated genetic effects at that time. Although differences between observed means at sequential times are sometimes used for studying quantitative genetic effects on developmental behavior, these analyses are still affected by the variables at previous times. Linear time series models have been proposed for development and age-related changes (EAVES *et al.* 1990). In these models phenotypic means measured at time t are expressed as linear functions of previous phenotypic means at time $(t - 1)$ and currently expressed genetic and environmental effects. This methodology still encounters the

problem of dependency of the current effects on the previous behavior.

Analysis of conditional genetic effects and their variance components provides a way for exploring gene expression on quantitative traits in different developmental stages. Statistical methods proposed in this paper are not only suitable for an additive-dominance model presented in this paper, but are also applicable for other genetic models expressed in mixed linear models. For some genetic models with correlated genetic factors, such as seed or endosperm models (ZHU and WEIR 1994a,b), MINQUE(0/1) can be applied for estimating variance and covariance components (ZHU

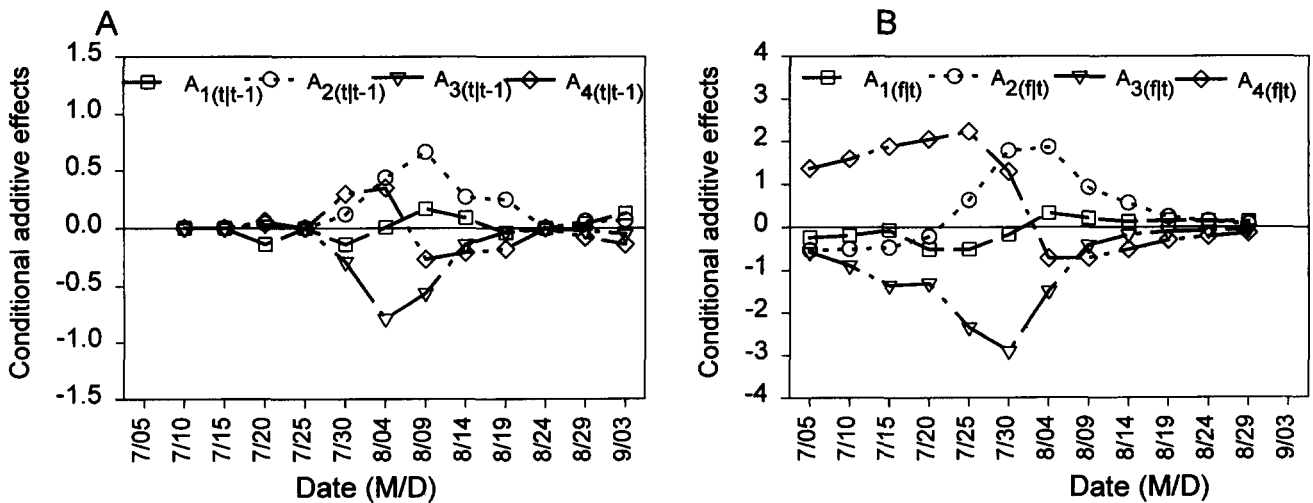


FIGURE 4.—Predicted additive effects of four parents for bolls per plant of Upland cotton. (A) Predicted additive effects for phenotypic means at time t conditional on means at time $(t - 1)$: $A_{1(t|t-1)}$ = additive effect of parent one, $A_{2(t|t-1)}$ = additive effect of parent two, $A_{3(t|t-1)}$ = additive effect of parent three, $A_{4(t|t-1)}$ = additive effect of parent four. (B) Predicted additive effects for final observation means conditional on means at different previous time t : $A_{1(f|t)}$ = additive effect of parent one, $A_{2(f|t)}$ = additive effect of parent two, $A_{3(f|t)}$ = additive effect of parent three, $A_{4(f|t)}$ = additive effect of parent four.

and WEIR 1994a). MINQUE(0/1) is a MINQUE method setting 0 for all the prior covariances and 1 for all the prior variances.

Analysis of conditional genetic effects and conditional variance components can be used not only for developmental quantitative traits but also for some agronomic traits. For example, the number of bolls per plant, boll size and lint% are three components for lint yield of cotton. Instead of analyzing genetic effects and variance components for lint yield of cotton, we can evaluate extra gene effects and genetic variation of lint yield conditional on any one of its components. For example, estimated genetic variance components for lint yield (LY) were $V_{A(LY)} = 79.7 \pm 13.6$, $V_{D(LY)} = 0.0 \pm 0.0$, $V_{AE(LY)} = 25.9 \pm 13.0$, and $V_{DE(LY)} = 81.0 \pm 23.9$. The conditional variance components of LY given the number of bolls (NB) observed were $V_{A(LY|NB)} = 37.5 \pm 7.5$, $V_{D(LY|NB)} = 0.0 \pm 0.0$, $V_{AE(LY|NB)} = 11.1 \pm 4.1$, and $V_{DE(LY|NB)} = 79.3 \pm 22.6$. The proportion of $V_{(LY|NB)}/V_{(LY)}$ might uncover gene effects without the influence of the given component (NB) on the final lint yield (LY), and $1.0 - V_{(LY|NB)}/V_{(LY)}$ could be used for measuring the proportion of gene effects of NB on lint yield LY. It was suggested that genetic effects of number of bolls per plant contributed ~45% variation of additive effects and additive by environment effects but only 2% variation of dominance by environment effects for lint yield of cotton.

I am grateful to BRUCE S. WEIR and SHAOBANG ZENG for helpful comments on the manuscript. I also thank two anonymous reviewers for useful comments and suggestions on the earlier version of the manuscript. This research was supported in part by National Natural Science Foundation of China (39170417) and National Institutes of Health grant GM-32518.

LITERATURE CITED

- BOX, G. E. P., G. M. JENKINS and G. C. REINSEL, 1994 *Time Series Analysis: Forecasting and Control*, Ed. 3. Prentice Hall: Englewood Cliffs, NJ.
- CORBEIL, R. R., and S. R. SEARLE, 1976 Restricted maximum likelihood (REML) estimation of variance components in the mixed model. *Technometrics* **18**: 31–38.
- EAVES, L., J. K. HEWITT, J. MEYER and M. NEALE, 1990 Approaches to the quantitative genetic modeling of development and age-related changes, pp. 266–280 in *Developmental Behavior Genetics*, edited by M. E. HAHN, J. K. HEWITT, N. D. HENDERSON and R. H. BENNO. Oxford University Press, New York.
- ENGLER, R. F., 1982 Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**: 987–1007.
- HALLAUER, A. R., and J. B. MIRANDA, 1981 *Quantitative Genetics in Maize Breeding*. Iowa State University Press, Ames, Iowa.
- HENDERSON, C. R., 1963 Selection index and expected genetic advance, Publication 982, pp. 141–163 in *Statistical Genetics and Plant Breeding*, edited by W. D. HANSON and H. F. ROBINSON. National Academy of Science, National Research Council, Washington, DC.
- HENDERSON, C. R., 1985 Best linear unbiased prediction using numerator relationship matrices derived for selected base populations. *J. Dairy Sci.* **68**: 443–448.
- HENDERSON, C. R., 1988 A simple method to account for selected base populations. *J. Dairy Sci.* **71**: 3399–3404.
- HENDERSON, N. D., 1990 Genetic analysis as a route to understanding the evolution of animal behavior: examples using the diallel cross, pp. 190–216 in *Developmental Behavior Genetics*, edited by M. E. HAHN, J. K. HEWITT, N. D. HENDERSON and R. H. BENNO. Oxford University Press, New York.
- KRZANOWSKI, W. J., 1988 *Principles of Multivariate Analysis: A User's Perspective*. Clarendon Press, Oxford.
- MILLER, R. G., 1974 The jackknife: a review. *Biometrika* **61**: 1–15.
- PATTERSON, H. D., and R. THOMPSON, 1971 Recovery of inter-block information when block sizes are unequal. *Biometrika* **58**: 545–554.
- RAO, C. R., 1970 Estimation of heteroscedastic variances in linear models. *J. Am. Stat. Assoc.* **65**: 161–172.
- RAO, C. R., 1971 Estimation of variance and covariance components MINQUE theory. *J. Multivariate Analysis* **1**: 257–275.
- WELLS, R., and W. R. MEREDITH, JR., 1984 Comparative growth of obsolete and modern cotton cultivars. II. Reproductive dry matter partitioning. *Crop Sci.* **24**: 863–868.
- WOODWARD, T. W., and N. R. MALM, 1976 Influence of lint percentage on yield, boll, and fiber characteristics in Acala strains of Upland cotton. *Crop Sci.* **16**: 594–596.
- ZHU, J., 1989 Estimation of genetic variance components in the general mixed model. Ph.D. Dissertation, NC State University, Raleigh, NC.
- ZHU, J., 1993 Methods of predicting genotype value and heterosis for offspring of hybrids. (Chinese). *J. Biomathematics (Chinese)*, **8**: 32–44.
- ZHU, J., D. JI and F. XU, 1993a Genetic analysis for flowering and fruiting of different sites on cotton plant. (Chinese). *Acta Gossypii Sinica*, **5**: 25–32.
- ZHU, J., D. JI, and F. XU, 1993b Genetic analysis for flowering and fruiting behaviors of Upland cotton (*Gossypium hirsutum* L.), pp. 302–312 in *Proceedings of Beijing International Symposium for Cotton Genetics and Breeding*, edited by Chinese Society of Cotton. Chinese Agricultural Science Press, Beijing.
- ZHU, J., D. JI and F. XU, 1993c A genetic approach for analyzing Inter-cultivar heterosis in crops. *Chinese J. Genet.* **20**: 183–191.
- ZHU, J., and B. S. WEIR, 1994a Analysis of cytoplasmic and maternal effects: I. A genetic model for diploid plant seeds and animals. *Theor. Appl. Genet.* **89**: 153–159.
- ZHU, J., and B. S. WEIR, 1994b Analysis of cytoplasmic and maternal effects: II. Genetic models for triploid endosperm. *Theor. Appl. Genet.* **89**: 160–166.
- ZHU, J., and B. S. WEIR, 1995 Diallel analysis for sex-linked and maternal effects. *Theor. Appl. Genet.* (in press).

Communicating editor: P. D. KEIGHTLEY