Copyright © 1998 by the Genetics Society of America

Two Sites in the *Delta* Gene Region Contribute to Naturally Occurring Variation in Bristle Number in *Drosophila melanogaster*

Anthony D. Long,* Richard F. Lyman,† Charles H. Langley* and Trudy F. C. Mackay†

**Center for Population Biology, University of California, Davis, California 95616 and*
†Department of Genetics, North Carolina State University, Raleigh, North Carolina 27695

Manuscript received August 29, 1997
Accepted for publication February 27, 1998

ABSTRACT

A restriction enzyme survey of a 57-kb region including the gene *Delta* uncovered 53 polymorphic molecular markers in a sample of 55 naturally occurring chromosomes. A permutation test, which assesses the significance of the molecular marker with the largest effect on bristle variation in four genetic backgrounds relative to permuted data-sets, found two sites that were independently associated with variation in bristle number. A common site in the second intron of *Delta* affected only sternopleural bristle number, and another common site in the fifth intron affected only abdominal bristle number in females. Under an additive genetic model, the polymorphism in the second intron may account for 12% of the total genetic variation in sternopleural bristle number due to third chromosomes, and the site in the fifth intron may account for 6% of the total variation in female abdominal bristle number due to the third chromosomes. These results suggest the following: (1) models that incorporate balancing selection are more consistent with observations than deleterious mutation-selection equilibrium models, (2) mapped quantitative trait loci of large effect may not represent a single variable site at a genetic locus, and (3) linkage disequilibrium can be used as a tool for understanding the molecular basis of quantitative variation.

MANY characters of evolutionary, medical, and agricultural importance are quantitative in nature. Variation in these traits can be partitioned into environmental and genetic components, the genetic component presumed to be due to the segregation of alleles at a number of loci that affect the trait (Falconer and Mackay 1996). A major goal of both medical genetics and agricultural genetics is the identification of the genes underlying quantitative traits. An understanding of the number of loci contributing to quantitative variation, the number and effects of alleles at these loci, how sensitive allelic effects are to variants segregating at other loci (*i.e.*, epistasis) and changes in the environment (*i.e.*, genotype by environment and genotype by sex interactions), and the molecular nature of the variants that give rise to quantitative variation, is also crucial to models that attempt to explain the nature of standing variation in quantitative traits (Barton and Turelli 1989), and in the determination of how the adaptations that distinguish closely related species become established (Orr and Coyne 1992). Despite the importance of understanding the nature of the genes contributing to quantitative variation, our understanding of quantitative traits is still largely limited to statistical descriptions of observed phenotypic variation under assumed underlying genetic models (Barton and Turelli 1989; Falconer and Mackay 1996).

We have been dissecting the genetic basis of standing variation in abdominal and sternopleural bristle number in *Drosophila melanogaster* to understand the genetic basis of quantitative variation. Bristle number in Drosophila is a good system for addressing this question, as it has been well characterized using classic quantitative genetic approaches (Falconer and Mackay 1996), and a number of studies have documented stabilizing selection on bristle number (Linney *et al.* 1971; Nuzhdin *et al.* 1995; Garcia-Dorado and Gonzalez 1996). Abdominal and sternopleural bristle are external sensory organs of the peripheral nervous system. A large number of candidate genes have been identified through mutants that affect bristle pattern and spacing, and the signaling pathways these genes act in are partially characterized (reviewed in Jan and Jan 1994). In addition, previous work on the mapping of quantitative trait loci (QTLs) important in short-term response to artificial selection on bristle number (Long *et al.* 1995), quantitative complementation testing of chromosomes containing high and low QTLs to mutants at candidate loci (Long *et al.* 1996), and associations between molecular variants in a candidate gene region and phenotypic variation among a set of wild-derived chromosomes (Mackay and Langley 1990; Lai *et al.* 1994) have suggested that a small number of loci make large contributions to standing variation in bristle number and that these loci appear to correspond to candidate genes that

Corresponding author: Tony Long, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697.
E-mail: tdlong@ucdavis.edu

are known to be important in bristle development. The generality of these initial results will rely on the results of a thorough survey of candidate loci. This work extends these previous results and demonstrates that naturally occurring variants at *Delta* contribute to standing variation in bristle number.

The QTL mapping study of Long *et al.* (1995) mapped a factor to the cytogenetic interval 89D1 to 92E1 that had an effect of approximately three abdominal bristles and one sternopleural bristle; this cytogenetic interval contains the *Delta* genetic locus. Additional support that *Delta* may be contributing to naturally occurring variation in bristle number came from quantitative complementation testing studies, in which a mutant of *Dl* genetically interacted with chromosomes containing high and low bristle QTLs fixed by artificial selection and likely segregating in nature (Long *et al.* 1996), as well as chromosomes which had accumulated mutations with quantitative effects on bristle number (Mackay and Fry 1996). *Delta* is an important and obvious locus at which naturally occurring genetic variation might be expected to contribute to variation in bristle number. *Delta* is known to be a component of the lateral inhibition process in bristle forming regions of the pupae (Parks and Muskavitch 1993; Parks *et al.* 1997). Lateral inhibition involves cell-cell signaling between cells expressing the *Delta* ligand and the *Notch* receptor in patches of cells that have the potential to form neurogenic structures (Haenlin *et al.* 1994; Kunisch *et al.* 1994; Artavanis-Tsakonas *et al.* 1995; Parks *et al.* 1997). Cells initially expressing slightly lower levels of the *Delta* ligand become committed to a non-neurogenic fate through a down regulation, via the Notch receptor pathway [whose downstream members include *H*, *Su(H)*, *E(spl)*, and *ASC*], and subsequent failure to up regulate additional *Dl* expression in the down regulated cell (Artavanis-Tsakonas *et al.* 1995). Experiments in which overall *Dl* expression levels are artificially raised (or lowered) tend to decrease (or increase) bristle number, consistent with the molecular understanding of *Delta* function (Heitzler and Simpson 1991). Based on the molecular biology of *Delta* it seems plausible that molecular variants in the *Delta* region of more subtle, quantitative effect may be responsible for naturally occurring variation in bristle number and patterning.

We report here the results of an experiment to associate DNA polymorphisms in the *Delta* gene region with variation in abdominal and sternopleural bristle number among a representative sample of 55 chromosomes extracted from a natural population. This association study approach has been used in previous studies of bristle number variation (Mackay and Langley 1990; Lai *et al.* 1994), and is an approach that will be useful in the future study of human diseases (Jorde 1995; Risch and Merikangas 1996; Long *et al.* 1997). If a neutral molecular marker (referred to as the marker site) and a site contributing to variation in a quantitative trait (the quantitative trait nucleotide or QTN) are physically close to one another, they are likely to be in linkage disequilibrium. This linkage disequilibrium will result from the two sites only rarely recombining from one another and hence sharing a common evolutionary history. Disequilibrium between the marker site and QTN can be detected as a significant regression of marker site allelic state on the phenotypic measure. The work presented here differs from previous studies in two main respects. First, we have used a permutation-based statistical approach for assessing marker/phenotype associations that is robust with respect to the number of correlated molecular markers used (Churchill and Doerge 1994; Doerge and Churchill 1996); and second, we test for marker/phenotype associations in a number of different genetic backgrounds to increase the power of detecting significant associations. Using this approach, we were able to identify two molecular markers in the *Delta* region that are at intermediate frequency and are associated with bristle variation.

MATERIALS AND METHODS

Isogenic lines: The derivation of the lines employed in this study, Drosophila culturing conditions, the crosses and genetic backgrounds employed, and the experimental design used to determine average abdominal and sternopleural bristle number for each line are described in Lyman and Mackay (1998). Briefly, all lines are derived from a wild collection of Drosophila from the Raleigh, NC Farmer's Market in 1988. Average bristle number was assessed in four different genetic backgrounds for each line: (1) homozygous third chromosomes in an otherwise isogenic *Samarkand* (*SAM*) background (*W*), (2) homozygous introgressed fragments of the third chromosome, including *Dl*, made by repeated rounds of backcrossing through females mutant for *Delta* in an otherwise isogenic *SAM* background (*B*), (3) the same introgressed fragment heterozygous against the *Dl* mutant (*T*), and (4) the same introgressed fragment heterozygous against the *SAM* chromosome (*C*). In the case of the *B*, *T*, and *C* bristle measurements were made on two independent introgression lines per whole third chromosome (Lyman and Mackay 1997). The lines used in this study were examined cytologically for the presence of large inversions on the right arm of chromosome three.

Restriction map analysis: The *Delta* region was examined at two levels of resolution to find sites detectable with restriction endonucleases which could be associated with variation in bristle number. Genomic DNA was prepared from approximately 200 flies either homozygous for the entire third chromosome, or a backcross segment including the *Delta* region, using an SDS lysis, organic extraction, ethanol precipitation protocol (Jowett 1986). Initially, each of the lines was examined with four restriction enzymes (*Eco*RI, *Hin*dIII, *Bam*HI, and *Pst*I) with a six base pair recognition sequence and Southern blotting technique (Southern 1975). Filters were made using ~1–2 μg of genomic DNA per restriction enzyme per line blotted onto S&S Nytran neutral charge membranes (Scheicher & Schuell, Keene, OH). These filters were sequentially probed with four phage clones spanning ~52 kb including the *Delta* transcriptional unit (see Figure 1; kindly provided by M. Muscavitz). Probes were labeled using the Random Prime kit from Boehringer Mannheim (Indianapolis, IN) and either 200 μCi of [α-P^{32}] dCTP or 200 μCi of both [α-P^{32}] dCTP

TABLE 1

Primers used in the survey of four-cutter variation in the *Delta* region

Name	Primer sequence	Anneal	Size (kb)
Dl-5′-F	5′-GGACAAGCGGTTGGAATTCCAGGAATCGCC	64	5.6
Dl-5′-R	5′-CTGCACGATGACTGTGAAGCAAATGAATGCTG		
Dl-In1-F	5′-TAACAGAATTCATTTGCTTCACAGTCATCGTGC	64	5.8
Dl-Inl-R	5′-GTGAAGCCCTTGTTCTGGAAGCGCTGG		
Dl-In2-F	5′-CTACGGGGACGTGATCACGC	59[a]	3.9
Dl-In2-R	5′-CGCATTGCCGCTATTGTTCG		
Dl-In3-F	5′-GTCGAGGCCTGGCATGATACG	59	4.2
Dl-In3-R	5′-CCATCCGGTCAAACAGATAATTTCG		
Dl-In4-F	5′-GATCTCAACTACTACGGATCCGGCTGTGCC	61	3.3
Dl-In4-R	5′-TTGGGTTTGTCGCAATGTCCATGTTCACAGCC		
Dl-In5-F	5′-CATTGCGACAAACCCAATCAATGC	58[b]	2.7
Dl-In5-R	5′-GTTTATGCATGAATTCGGACTGC		
Dl-3′-F	5′-TGGGTTGGGTTCAAAATGTGAGAGAGACGCCC	61	2.1
Dl-3′-R	5′-CAACACCAACAATAACAGTTAAAAGACAGCGG		

All primers with "In" in their name are located in *Dl* exons and amplify the intron they flank. Dl-5′-R and Dl-3′-F are located in *Dl* exons, and the positions of Dl-5′-F and Dl-3′-R can be inferred from the size of the amplicon. Adjacent amplicons do not necessarily overlap, thus the entire coding region was not surveyed. PCR conditions are described in text, with exceptions noted below.

[a] 0.8 μl of 25 mM $Mg(0Ac)_2$ added per 50 μl reaction.

[b] 1.5 μl of 25 mM $Mg(0Ac)_2$ added per 50 μl reaction.

and [α-P^{32}] dGTP, with unincorporated nucleotides removed over a Sephadex column (Sambrook *et al.* 1989). Hybridization, washing, and autoradiography conditions were carried out as in Krietman and Aguade (1986; modified from Church and Gilbert 1984), with similar hybridization (10 mg/ml BSA (Sigma, St. Louis, MO), 1 mm EDTA pH 7.2, 525 mm NaPi pH 7.2, 7% SDS) and wash (40 mm NaPi pH 7.2, 1mm EDTA pH 7.2, 1% SDS) solutions.

The above six-cutter survey resulted in only a small number of polymorphic sites throughout the *Delta* region, so a second higher resolution survey was carried out using "Long" PCR (Barnes 1994; Cheng *et al.* 1994) coupled with digestion of the PCR products with four base pair recognition sequence endonucleases. This approach used a series of primers designed using published *Delta* cDNA sequence to amplify introns of unknown sequence. In the case of Dl-5′ and Dl-3′ λ clone-associated restriction fragments were subcloned into pBluescript and sequenced into the subclone far enough to generate a sequence tagged site 5′ or 3′ to the *Delta* transcription unit large enough to design PCR primers. The following PCR conditions were used to amplify genomic DNA: 50 μl reaction volume, 1× Tricine Buffer [5× Tricine Buffer: 0.1 m Tricine pH 8.9, 0.42 m KOAc, 0.01 m $Mg(OAc)_2$], 0.5 mm of both forward and reverse primer (Table 1), 0.2 mm of each dNTP, ~20 ng of gDNA, 2 μl Taq polymerase, and 0.5 μl Taq Extender (Stratagene, La Jolla, CA). Cycling conditions were: an initial 2-min denaturation step at 92° followed by 35 cycles with 45 sec at 92°, 45 sec at the correct annealing temperature for a given primer (Table 1), and 2–4 min at 70° depending on the expected size of the PCR product. From each PCR reaction, 10 μl was transferred to each of 4 microtiter plates and 0.5 μl of a four-cutter restriction enzyme (*Alu*1, *Cfo*1, *Hae*III, and *Hpa*II) was added to each sample in each of the plates. In the case of Dl-I2 and Dl-I5 the entire experiment was repeated with four additional enzymes (*Dde*I, *Rsa*I, *Scr*FI, and *Taq*I). Plates were sealed with adhesive microplate film and incubated at 37° (65° in the case of *Taq*I) for ~3 hr. Under the described conditions, these eight four-cutter enzymes appear to completely digest the PCR products (although partial digestions were occasionally observed with *Hpa*II). After digestion, 3 μl of 6X Ficoll loading buffer (Sambrook *et al.* 1989, with the addition of EDTA pH 8.0 to 0.01 m) was added to the restricted Long PCR product and the products were run overnight at low power on a submarine "Synergel" [1.5% Synergel (Research Products International, Mt. Prospect, IL), 0.5% agarose, 1× TAE], and visualized by staining with ethidium bromide.

As restriction maps were not available for the unsequenced regions of *Delta*, polymorphisms were scored as the presence or absence of bands. Thus, unlike the data from the six-cutter survey of the *Delta* region, which could be analyzed using standard methods for studying nucleotide variation, the data from the four-cutter survey was analyzed using a band sharing approach to estimate nucleotide variation (Nei and Li 1979). A 95% confidence interval was constructed for the sample estimate of nucleotide variation by estimating the same parameter from 1000 bootstrap estimates of the data with each estimated correct for bias by a factor of $n/(n - 1)$, where n is the number of bootstrap samples drawn from the data which is also equal to the number of lines surveyed (Weir 1996). Introns 2 and 5 were sequenced, allowing sites located in the PCR fragments corresponding to these regions to be ordered. These sequences are deposited in GeneBank (accession numbers AF035257 and AF035258).

Statistical analysis of molecular marker/phenotypic associations: For each line in the study, bristle phenotypes were assessed in each of four genetic backgrounds (*W*, *B*, *T*, and *C*). For analyses of molecular marker/bristle number associations, arithmetic mean bristle phenotypes were used for each background/line/sex (see Lyman and Mackay 1998 for a complete description of the lines). In the case of the *W* genetic background this mean is based on 20 individuals of each sex (ten individuals per vial times two vials), and for the *B*, *T*, and *C* genetic backgrounds it is based on 40 individuals of each sex (ten individuals per vial times two vials times two independent introgression replicates). The molecular marker data used to test for marker/phenotype associations is listed in appendix a. The marker data is a combination of the six-cutter sites

and four-cutter bands, with four-cutter bands having identical patterns of presence and absence among the lines collapsed to a single site for the purpose of association testing. In previous work we have carried out quantitative complementation testing on mapped QTLs (Long *et al.* 1996), and mutation accumulation lines (Mackay and Fry 1996). In this study we apply an analogous contrast to that employed in previous work, by testing for associations between molecular markers and the difference in bristle number between flies of genotype B_i/Df^3 *(T)* and those of genotype B_i/Sam *(C)* over the $i = 1$ to n lines of the study. The $T - C$ contrast can be thought of as the effect of failure to complement with a control for the additive effect of wild-type alleles. All subsequent significance testing was carried out on a transformed data-set consisting of three response variables, the first two are homozygous entire third and backcross chromosomes and the third the effect of failure complement (W, B, and $T - C$). Since the variance among line means is different for each of the three response variables, the significance of molecular marker/phenotypic associations were only carried out on line means transformed by sex and response variable to have zero mean and unit variance (effectively weighting each sex and response variable equally). The original average bristle counts in the four genetic backgrounds and the three scaled response variables will be referred to as the "four genetic background" and "three genetic background" data-sets, respectively.

To assess the significance of marker/phenotype associations, a test statistic was constructed that combines phenotypic information from the three genetic backgrounds. Since the measures of bristle number are correlated over the genetic backgrounds (Lyman and Mackay 1998), an ideal test statistic will take these correlations into account. This requirement was met by performing a principal component analysis on the three response variable data-set to reduce the dimensionality of the observations to one for each character and sex. Principal components were separately estimated for each character and sex from the covariance matrix associated with the three genetic backgrounds, and principal component scores were subsequently created by multiplying the three genetic background z-scores by the eigenvector associated with the largest eigenvalue (*i.e.*, the first principal component score; Chatfield and Collins 1980). This resulted in a set of four principal component scores (male abdominal, female abdominal, male sternopleural, and female sternopleural bristle number) which were separately analyzed for each character and molecular marker with the linear model $Y_{ijk} = \mu + s_i + m_j + (s{*}m)_{ij} + c(m)_{k(j)} + s{*}c(m)_{ik(j)}$, where Y_{ijk} is the principal component score of the ith sex with the jth allele for the lth wild chromosome at the molecular marker under consideration. The F statistics, ϕ_m and ϕ_{s*m}, associated with the model terms of marker and sex X marker, respectively (tested with the chromosome in marker and sex X chromosome in marker mean squares, respectively), can now be assessed for significance. Principal component analysis has advantages and disadvantages relative to other approaches we examined, which are discussed later.

At this point there are two statistics, ϕ_m, and ϕ_{s*m}, for each bristle character/molecular marker combination, and we wish to assess if variation in bristle number is associated with any of the m molecular markers of the study. Since the m markers under consideration are correlated with one another, it is difficult to derive a theoretical threshold that the ϕ's must exceed to denote statistical significance. Therefore, we assessed the statistical significance of our results using a permutation testing approach for mapping QTLs (Churchill and Doerge 1994) combined with sequential removal of significant markers (Doerge and Churchill 1996) applied in the disequilibrium mapping context of this work. This method has the advantage of assessing significance using a distribution free statistical approach, while preserving the linkage disequilibria structure of the markers (*i.e.*, the correlational structure of the markers) over permutations. Significance was assessed by calculating the largest value of ϕ, ϕ^{max}, over the m markers of our study for a given character/statistic combination, and then comparing this value to the distribution of ϕ^{max} derived by randomly permuting entire haplotypes for each line with respect to phenotypic data 1000 times and calculating ϕ^{max} each replication. If the ϕ^{max} associated with the nonpermutated haplotypes is in the upper 5% of the permutation tests (*i.e.*, only 50 or fewer values of ϕ^{max} from the permutated data-sets exceed it), the effect of that molecular marker is statistically removed from each of the principal component scores. After statistical removal of any significant molecular markers, the entire permutation testing procedure is repeated on the residuals generated in this manner until no further markers remain significant.

The effect of a given marker in a given sex and genetic background is calculated as the difference between the mean of all lines having that marker versus the mean of all lines lacking the marker. One standard error on the effect of a given marker can be calculated as

$$\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_0^2}{N_0}},$$

where σ_x^2 is the variance among lines that have allele x at the molecular marker and N_x is the number of lines that have allele x at the molecular marker. The variance due to a given marker is calculated using the VARCOMP procedure of SAS. Within VARCOMP the following model was fitted for molecular markers that showed sex limited effects: $Y_{ijklm} = \mu + M_i + L(M)_{j(i)} + B(L\ M)_{k(ji)} + R(B\ L\ M)_{l(kji)} + \varepsilon_{mlkji}$, where Y_{ijklm} is the bristle number of the mth individual nested in the lth replicated vial, nested in the kth backcross replicate (this effect was excluded in the case of homozygous third chromosomes that were not replicated), nested in the jth line, nested in the ith marker class. All effects were considered random for the purpose of estimation, except the effect of marker which is fixed. For markers that showed phenotypic effects in both sexes the above model was fitted with the inclusion of an additional fully crossed fixed effect of sex. Under a strictly additive model of quantitative variation $2FV_G = \sigma_2^2 + 2\sigma_{S*L}^2 + \sigma_M^2 + 2\sigma_{S*M}^2$, and similarly $2FV_M = \sigma_M^2 + 2\sigma_{S*M}^2$, where F is the inbreeding coefficient (Falconer and Mackay 1996). For completely homozygous lines $F = 1$; thus, obvious estimators of the total additive genetic variation (V_G) and the total additive genetic variation due to marker (V_M) are suggested by the above equations.

RESULTS

Patterns of bristle variation among Raleigh third chromosomes: appendix a gives the mean bristle count by sex and genetic background for each of the lines presented in this study. A complete analysis of the patterns of variation among the introgression lines of this study is given in Lyman and Mackay (1998): here we briefly review some of these patterns. Small standard errors were generally associated with estimates of line means (*i.e.*, 0.26, 0.27, 0.38, 0.35 pooled over sexes for abdominal bristle number in the W, B, T, and C genetic backgrounds, and 0.23, 0.20, 0.25, 0.18 for sternopleural bristle number), indicating that variation between lines

is primarily genetical in origin. Within a given sex and bristle character, bristle numbers were comparable over genetic backgrounds, the exception being the complementation testing cross background, which generally had a higher average bristle number. The observation of higher bristle number on average in a genetic background containing a *Dl* loss-of-function allele is consistent with previous phenotypic results involving *Delta* mutants (Long *et al.* 1996; Mackay and Fry 1996), and the proposed function of *Delta* in which loss-of-function mutants cause an over-commitment to a neurogenic fate. Also in accord with previous work, within a given bristle trait and genetic background, females generally had a greater average number of bristles than males. This effect of sexual dimorphism was more pronounced for abdominal bristle number (range 2.09 to 3.88) than sternopleural number (range 0.19 to 2.36). For both bristle characters, the ordering of the four genetic backgrounds with respect to sexual dimorphism was the same with: $T > B > C > W$ chromosome. Over all experimental treatments, the variance among line means in abdominal bristle number ranged from 0.33 (males in the *C* genetic background) to 5.36 (females in the *W* genetic background), and from 0.16 (females in the *C* genetic background) to 6.30 (females in the *W* genetic background) for sternopleural bristle number. For both bristle characters, the ordering of the four genetic backgrounds with respect to variance in bristle number was the same, with $W > B > T > C$. For both bristle characters, the variance among lines was greater for females than for males (with the exception of the *C* genetic background for sternopleural bristle number). This effect was more pronounced for abdominal bristle number, where the variance among lines in females was approximately twice that of males (the ratio of female to male variance in abdominal bristle number ranged from 1.59 to 3.33), than for sternopleural bristle number, where the variance among lines was, at most, 26% greater in females than in males (the ratio of female to male variance in sternopleural bristle number ranged from 0.58 to 1.26).

Molecular population genetics of the *Delta* region: The region surveyed with four six-cutter enzymes was ~57 kb, of which 22 kb includes the entire transcriptional unit of the *Dl* locus (Figure 1). appendix a gives the molecular marker genotypes for each of the lines presented in this study. Of the 59 restriction sites surveyed with six-cutters, 18 were polymorphic. Of these 18 polymorphic sites, 15 were 5′ to the *Dl* transcriptional start site and are thus unlikely to be useful for detecting molecular marker/phenotype associations in the coding region. From the six-cutter data we estimate θ, whose expectation is $4N_e\mu$ for selectively neutral substitutions, to be equal to 6.53×10^{-3}, with sample variances of 2.39×10^{-6} and 1.18×10^{-5} under infinite and zero recombination, respectively (Hudson 1982). The average per-site heterozygosity, π, which was estimated using three different methods, gave very similar values of 7.13×10^{-3} (with variance 1.54×10^{-5}; Nei and Li 1979), 7.14×10^{-3} (with variance 1.54×10^{-5}; Nei and Tajima 1981), and 6.96×10^{-3} (Engels 1981). Under a neutral model of evolution, the expectations of θ and π are the same, and departures from neutrality, due to forces such as genetic hitch-hiking, can be detected as the normalized difference between θ and π, a statistic referred to a Tajima's D (Kaplan *et al.* 1989; Tajima 1989; Braverman *et al.* 1996). Tajima's D for the six-cutter of data of this sample is 0.272, providing little evidence for a departure from neutrality. The six-cutter survey revealed an atypically low amount of insertion/deletion variation, with no large insertions, and only three small deletions. Two of the deletions were observed once in the set of 55 lines (a 50-bp deletion between -0.63 and 0.0 seen in line 34, and a 150-bp deletion between 5.9 and 6.2 seen in line 8), the other deletion was seen in four of the 55 lines (a 50-bp deletion between 23.7 and 24.6 in lines 17, 36, 56, and 81).

Using a survey approach that combines Long-PCR, and subsequent restriction digests with a series of four-cutter enzymes, followed by electrophoretic separation, we were able to score 112 polymorphic and 156 monomorphic bands throughout ~27 kb of the *Dl* region that includes the entire *Dl* transcription unit. This survey was carried out by dividing *Dl* into eight PCR fragments corresponding to the 5′ region, Introns 1 through 5, Exon 6, and the 3′ untranslated region (the most 5′ part of the survey was −3.15 and the most 3′ was +23.7). This four-cutter survey resulted in data that was difficult to interpret with respect to the presence or absence of restriction sites. Four-cutter surveys are generally carried out on segments with a known consensus sequence that can be used to interpret banding patterns relative to the gain or loss of sites (Kreitman and Aguade 1986), but *Dl* sequence is not available for much of the region surveyed. Although many sites could be easily interpreted as the gain or loss of sites, some sites could not be interpreted in this manner, so we estimated per-site heterozygosity using a band sharing approach (Nei and Li 1979). Using this approach on the four-cutter data, we estimate π to be 9.84×10^{-3} for the *Dl* region, with a bootstrap 95% confidence interval on the sample estimate of 9.13×10^{-3} to 1.06×10^{-2}. Estimates of nucleotide variation in the Delta region are typical relative to other loci that have been surveyed in *D. melanogaster* despite the observation of zero variable sites in Intron 4 and Exon 6. (These two fragments show 20 and 14 monomorphic bands, respectively). Estimates of π associated with the other fragments were typical, ranging from 6.11×10^{-3} to 1.69×10^{-2}. As a number of the surveyed bands could not easily be interpreted as site variants, it is possible they represent small insertion/deletion variants not detectable by the six-cutter survey. Thus, it is not possible to partition the estimate of π into variation due to insertions and deletions versus restriction

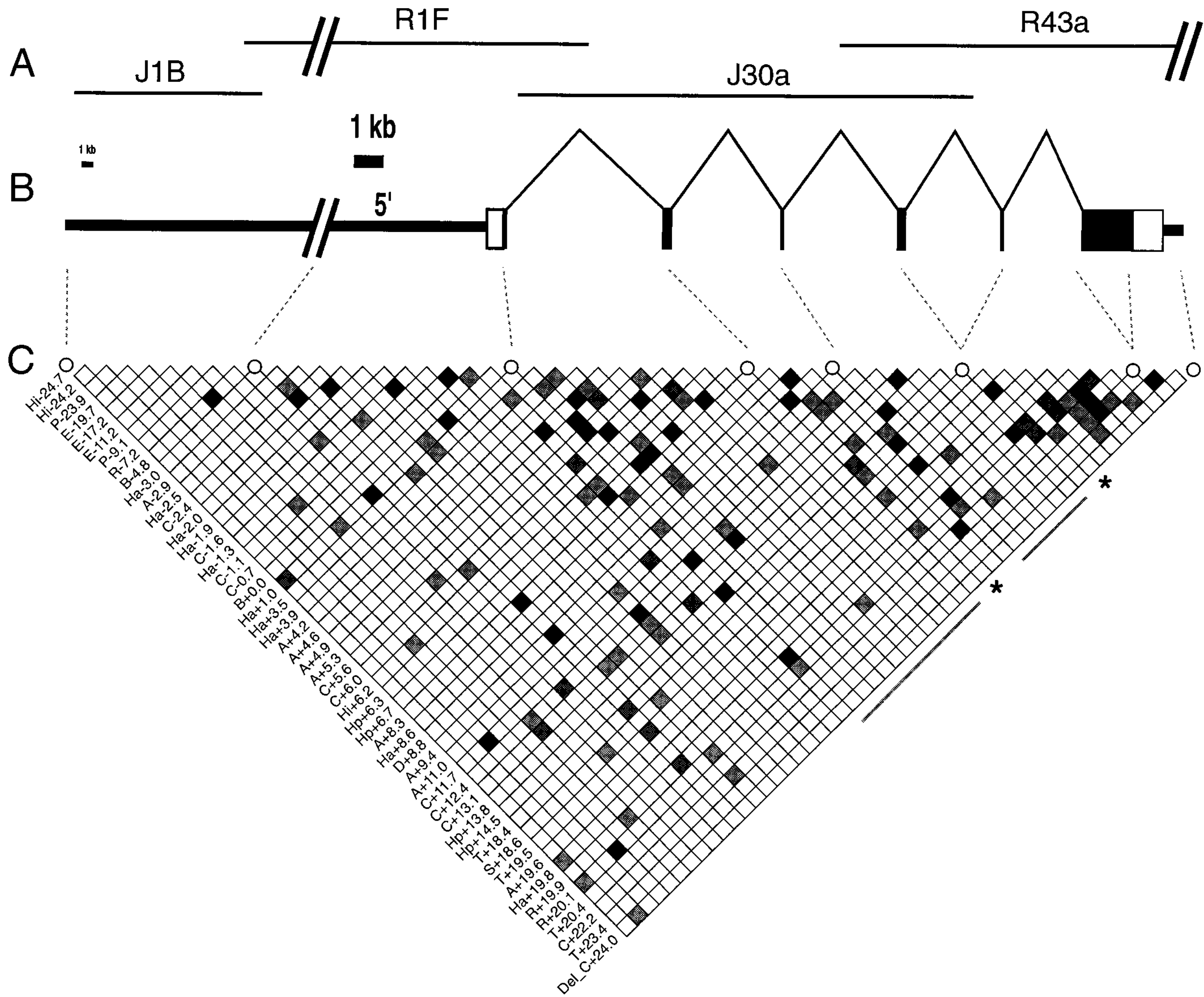


Figure 1.—Summary of restriction map variation in the *Delta* region. (A) Four phage clones covering the *Delta* region used in the six-cutter survey. (B) The exon/intron structure of the *Delta* gene region and flanking DNA (Kopczynski *et al.* 1988; Haenlin *et al.* 1990; M. A. T. Muscavitz, personal communication). The five prime- and three prime-most markers are at −24.7 and +24.0, respectively, and the point at which the scale of the figure changes is −3.5. A *Bam*HI site that is 1637 bp 5′ to the AUG translation initiation codon (and 964 bp 5′ to a strong transcription initiation site) is designated as position 0.0. Thick lines are exons, unfilled are untranslated, and filled are translated; thin lines are flanking DNA (the 5′ flanking region is drawn at two different scales as indicated by 1-kb labels). (C) A summary of linkage disequilibria in the *Delta* region. Hollow circles connecting the disequilibrium table with exon/intron structure represent boundaries of the different fragments surveyed by long PCR and four-cutter enzymes. Labels on the left of the disequilibrium table correspond to those described in appendix a (those prefixed with Hi, E, P, or B are six-cutter enzymes, and the remaining are four-cutter enzymes). Those sites with a gray bar to the right of the disequilibrium table are unordered but within the interval specified by the bar. Those sites with an asterisk on the right side of the table are the two sites that are significantly associated with variation in bristle number. Squares within the disequilibrium grid are shaded according to the significance of the disequilibrium between pairs of sites (not corrected for multiple tests) assessed using a Fisher's exact test: black squares $P < 0.005$, gray squares $P < 0.01$, light gray squares $P < 0.05$, and unshaded squares $P > 0.05$.

site polymorphisms. As the objective of this study was to find molecular markers throughout the *Dl* region that may be in linkage disequilibrium with a putative site causing bristle differences in natural populations, it is not important at this stage to distinguish between these two types of molecular events. The Long-PCR/four-cutter approach of this work does seem an efficient means to generate molecular markers in a large candidate gene region of interest.

All polymorphisms that had a frequency of greater than three in the sample of 55 lines were examined for pairwise linkage disequilibrium using Fisher's exact test. Figure 1 shows the patterns of disequilibria for the *Delta* region. Of the 1378 pairs of sites tested, 40, 57, and

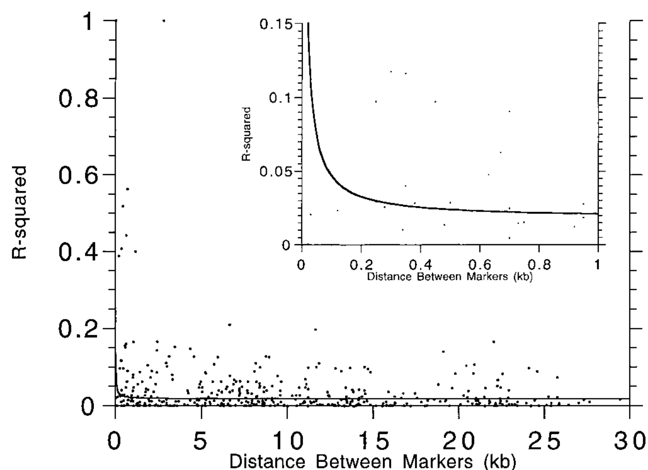


Figure 2.—A plot of R^2, the correlation coefficient associated with pairs of molecular markers, versus the estimated physical distance between markers for all markers with a frequency of 0.20 or greater. The fitted line is for the value of k that minimizes the sum of the squared deviations between the observed value of R^2 and its expectation of $1/(1 + k * \text{distance}) + 1/55$ (Weir and Hill 1986); k is a constant relating physical distance to units of $4Nc$ (Γ). The inset figure is the same plot for the region containing the inflection point of the fitted line; the inflection point is at ~0.16 kb, which is equal to ~1 Γ. R^2 falls off very quickly with distance, indicative of very low average levels of linkage disequilibria in the *Delta* gene region. All sites with $R^2 > 0.3$ are within 5 kb of one another.

118 were significant at 0.5, 1.0, and 5.0 percent levels, respectively (not corrected for multiple tests). The number of pairs of sites for which significant values of the Fisher's exact test were observed is slightly more than what one would expect by chance: 2.9, 4.1, and 8.6 percent were observed significant at 0.5, 1.0, and 5.0 percent levels, respectively. From Figure 2 it appears this small excess of linkage disequilibria, over what would be expected by chance alone, is primarily due to sites physically close to one another. When R^2 (the correlation coefficient is a measure of disequilibrium) for all pairs of markers with frequencies of 20% or greater is plotted against distance between the markers, all pairs of markers with values of R^2 greater than 0.3 are within 5 kb of one another. Hudson's estimator of 4Nc (an estimate of four times the effective population size times the probability of a recombinational event per gamete per generation for the region under consideration, subsequently referred to as Γ) was applied to the data to derive an estimate of the number of recombination events that have occurred in the history of the sample (Hudson 1987). Based on the six-cutter data, Γ was 725, or 14.9 Γ/kb, whereas the estimate of Γ from the four-cutter data was 165, or 6.2 Γ/kb. In the four-cutter survey, we observed 41 polymorphic sites in 26.8 kb; it follows that the average distance between polymorphic sites was 0.65 kb or 4 Γ. It follows from Hudson's estimator of Γ for the four-cutter data that one Γ is approximately 175 bp, which coincides with the inflexion point in the theoretical curve fit to the relationship between R^2 and distance between markers of Figure 2.

Associations between molecular variants and bristle number: We tested whether the 53 polymorphic molecular markers of the survey (those with a frequency of 3 or greater in the survey) were associated with variation in bristle number using a principal component analysis which combined data from the three genetic backgrounds (*W*, *B*, and *T-C*) coupled with a permutation testing approach. Table 2 lists the results of the analyses carried out to detect associations between molecular variants at *Delta* and variation in bristle number (marker names are defined in appendix a). Marker associations carried out on the first principal component (which accounts for 48, 62, 56, and 62% of the variation in male abdominal, female abdominal, male sternopleural, and female sternopleural bristle number, respectively) resulted in Ha + 8.6 being significantly associated with variation in sternopleural bristle number ($P < 0.009$) and S + 18.6 significantly associated with an abdominal bristle number by sex interaction ($P < 0.037$). The effect of Ha + 8.6 was subsequently statistically removed from the component scores in both sexes and for both bristle characters and the permutation testing procedure repeated. Marker S + 18.6 remained significantly ($P < 0.028$) associated with an abdominal bristle number by sex interaction. This demonstrates that Ha + 8.6 and S + 18.6 are independently associated with variation in sternopleural and abdominal bristle number. When the effects of both Ha + 8.6 and S + 18.6 are statistically removed from the component scores, no sites remain significantly associated with variation in bristle number. Although not significant, once the effects of Ha + 8.6 and S + 18.6 are statistically removed from the model, Ha − 2.5 approaches statistical significance for abdominal bristle number ($P < 0.075$).

The second principal component scores accounted for 33, 21, 26, and 23% of the variation in male abdominal, female abdominal, male sternopleural, and female sternopleural bristle number, respectively, in the three genetic background data-set. Molecular marker/phenotype association analyses were carried out on the second principal component scores and no sites were significantly associated with variation in bristle number. That variation in the second principal component scores was not associated with any molecular markers may be expected, given that the second principal component generally accounted for much less of the total variation in bristle number than the first principal component.

Table 2 gives P-values associated with four rounds of testing for statistically significant associations between DNA sequence polymorphisms and variation in bristle number for each of the four characters in the study, the four characters studied being the two bristle characters (abbreviated AB and SB) and each character for a marker

TABLE 2

Results of permutation testing procedure used to evaluate the significance of marker/bristle trait associations

		AB		SB		AB*X		SB*X	
Statistic[a]	Data[b]	Mark	$P <$	Mark	$P <$	Mark	$P <$	Mark	$P <$
PC1	*W, B, T-C*	Ha − 2.5	0.424	Ha + 8.6	0.009	S + 18.6	0.037	A + 19.6	0.403
	Remove Ha + 8.6	Ha − 2.5	0.243	Ha − 2.0	0.483	S + 18.6	0.028	A + 19.6	0.302
	Remove S + 18.6	Ha − 2.5	0.075	Ha − 2.0	0.660	Hp + 6.3	0.817	A + 19.6	0.304
PC2	*W, B, T-C*	A + 11.0	0.711	C + 5.6	0.721	A + 11.0	0.603	A + 11.0	0.973

[a] Statistic refers to the statistic tested for significance using the permutation testing procedure. PC1 and PC2 refer to F statistics calculated from a univariate ANOVA of the first and second principal component scores, respectively, separately derived from the data for each sex and bristle trait (see text).

[b] Data refers to the data used in the analysis (*W* = whole chromosome, *B* = backcross chromosome, *T* = backcross chromosome over Df^3, and *C* = backcross chromosome over *Samarkand*). All analyses were carried out on measures transformed (by sex and background) to zero mean and unit variance, and a number preceeded by "remove" refers to the set of phenotypic measures immediately above with the average effect of that marker (by sex) statistically removed from each of the data points.

by sex interaction (abbreviated AB*X and SB*X). Although the second principal component scores are independent of the first principal component scores (by definition), a Bonferroni correction to the first principal component scores would be overly conservative. Since the second principal component score accounts for much less of the total variation in the data-set than the first principal component score (on average 21% *vs.* 62%), tests for molecular marker/phenotype associations with the second principal component score are *a priori* less likely to be significant compared to those for the first principal component. Abdominal and sternopleural bristle counts are correlated phenotypically, but only weakly genetically correlated, therefore it is unclear how to correct for multiple tests over these two characters. As with the second principal component scores, marker by sex interactions for a given bristle trait have lower statistical power *a priori* than a marker site main effect test. Thus, it is also unclear how statistical significance of marker by sex interactions should be assessed relative to marker main effects. Ultimately, replication is desirable to firmly establish the significance of any marker/phenotype associations, nonetheless, the significance levels observed in the study indicate these associations are indeed statistically significant.

Figures 3 and 4 are plots of the value of the F statistics (*i.e.*, the ϕ's) associated with the first principal component derived from the three genetic background data-set for each of the molecular markers in this study as a function of marker position. These F statistics are presented separately for sternopleural and abdominal bristle number in Figures 3 and 4, respectively, with the two panels within each figure showing the F statistics for a model that tested the main effect of a given marker over both sexes (labeled SB and AB in the two figures), or for a model that tested the effect of a marker by sex interaction (labeled SB*sex and AB*sex in the two figures). These figures provide a visual means of assessing the significance of different markers relative to their position. In Figure 3 (bottom panel) the marker with a F statistic of 14.6 in Intron 2 is the marker found significant using the permutation test for SB (Ha + 8.6), and in Figure 4 (top panel) the marker with an F statistic of 13.4 in Intron 5 is the marker found significant using the permutation test for AB by sex (S + 18.6). As the significance of associations were assessed using a permutation testing approach, there is no meaningful threshold to place on these figures to indicate statistical significance. A helpful guide may be to recall that an $F_{1,54}$-statistic of 8.6 is normally considered significant at $P < 0.005$ for independent tests, and an $F_{1,54}$-statistic of 12.2 corresponds to an experiment-wise significant level of $P < 0.05$ (corrected for the 53 markers considered). Considering these thresholds, no markers exceed the theoretical critical F of 8.6 but fail to exceed 12.3, and only two markers exceed 12.3 (Ha + 8.6 for SB and S + 18.6 for AB*X). These are the same two markers that are significant by the permutation testing procedure. In Figures 3 and 4 the large F statistics appear to be clustered. This is to be expected, as markers close to one another are often in linkage disequilibrium and, as a result, are likely detecting the same marker/phenotype association. A lemma of this is that when the effect of a significant marker is statistically removed, "satellite" F statistics disappear.

Figure 5 gives the effect of Ha + 8.6 on sternopleural bristle number in males (top panel) and females (bottom panel) in each of the four genetic backgrounds and for the first principal component score associated with the three genetic background data-set. It can be seen in Figure 5 that the effect of Ha + 8.6 on sternopleural bristles is very consistent over the two sexes. The estimated effect of this polymorphism is 0.6 and 0.7 bristles in the male and female backcross (*B*) lines, respectively. The whole chromosome (*W*) lines gave the largest estimated effect, while the introgressed chromo-

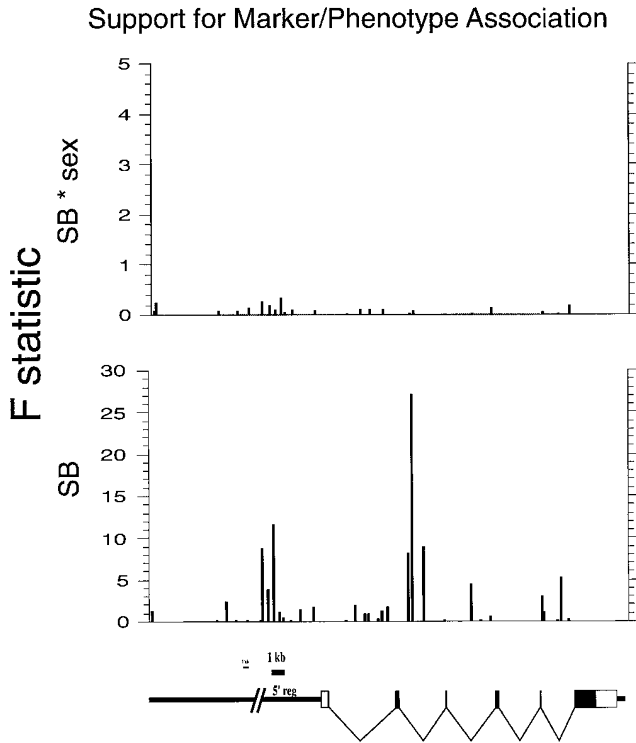


Figure 3.—Plots of the values of the F statistics of models testing for molecular markers that are associated with variation in the first principal component score derived from the three genetic background data-set (*W*, *B*, and *T-C*) for sternopleural bristle number. F statistics are plotted as a function of the estimated physical position of the marker site in the *Delta* gene region. The locations of some markers 5′ to the *Delta* gene are drawn at a different scale as indicated in the small figure of the *Dl* transcription unit. The top panel shows F statistics associated with a model that tests for a marker by sex association with the phenotype, and the bottom panel a model testing only for the main effect of the molecular marker. The F statistics are proportional to the variance attributable to a given marker.

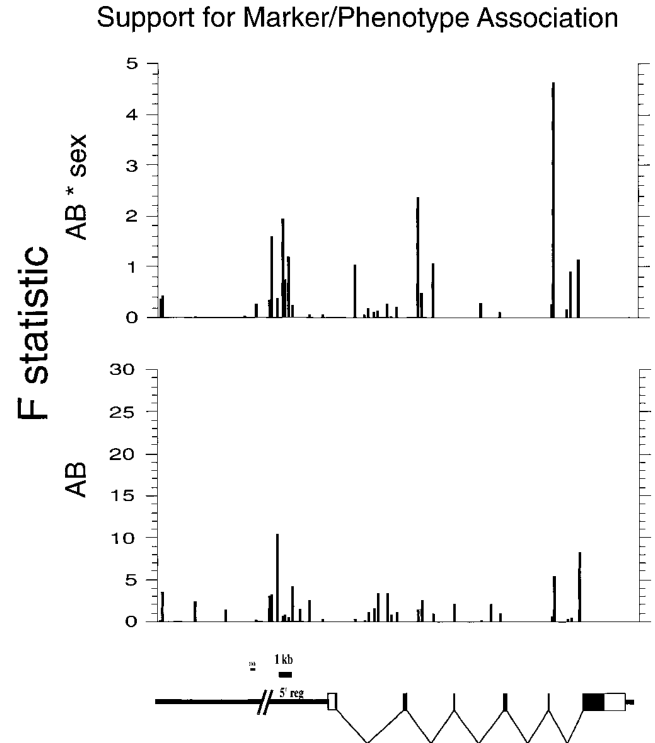
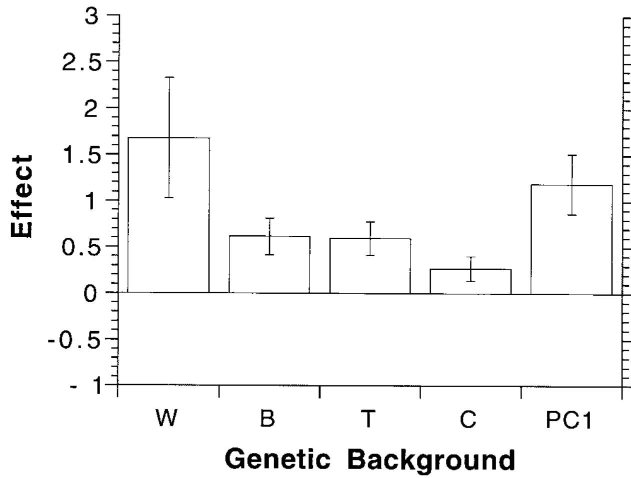


Figure 4.—Plots of the values of the F statistics of models testing for molecular markers that are associated with variation in the first principal component score derived from the three genetic background data-set (*W*, *B*, and *T-C*) for abdominal bristle number.

some over a *Delta* loss-of-function mutant (*T*) lines gave an estimate very comparable to the backcross lines, and the introgressed chromosome over a wild-type chromosome (*C*) effect was about half as large as the backcross lines. In all four genetic backgrounds, effects were in the same direction, with the presence of the Ha + 8.6 marker associated with increased sternopleural bristle number.

Figure 6 gives the effect of S + 18.6 on abdominal bristle number in males (top panel) and females (bottom panel) in each of the four genetic backgrounds and for the first principal component scores associated with the three genetic background data-set. Unlike the effect of Ha + 8.6 on sternopleural bristle number, the effect of S + 18.6 on abdominal bristle number is very different in the two sexes. From the figure, it appears that this polymorphism has no measurable effect in males. The observation of sex-specific allelic effects associated with molecular markers in the *Delta* region for abdominal bristle number is consistent with a phenotypic analysis of the backcross chromosomes where a large sex by line component of variation was observed (Lyman and Mackay 1998). The estimated homozygous effect (from the backcross lines) of S + 18.6 in females is approximately 1.1 abdominal bristles, which is in almost complete agreement with the estimates from the whole chromosome lines and the introgression over *Delta* loss-of-function lines. The estimated effect associated with the introgressed chromosome over wild-type chromosome lines (*C*) is 37% of the estimate from the backcross lines, consistent with additivity. Like the effect of Ha + 8.6 on sternopleural bristle number, the effect of S + 18.6 on female abdominal bristle number was in the same direction in all genetic backgrounds (presence of the site associated with decreased bristle number).

Table 3 gives the variance attributable to sites Ha + 8.6 and S + 18.6, and the total genetic variance, in each of the four genetic backgrounds of this study. The effects associated with the two sites significantly associated with bristle variation appear large relative to total additive genetic variation. For example, considering only effects and variance in females for abdominal bristle number, the effect associated with S + 18.6 is 0.78 whole third chromosome background genetic standard deviations (*i.e.*, $effect_W/\sqrt{V_{G,W}}$), and 1.06 homozygous backcross



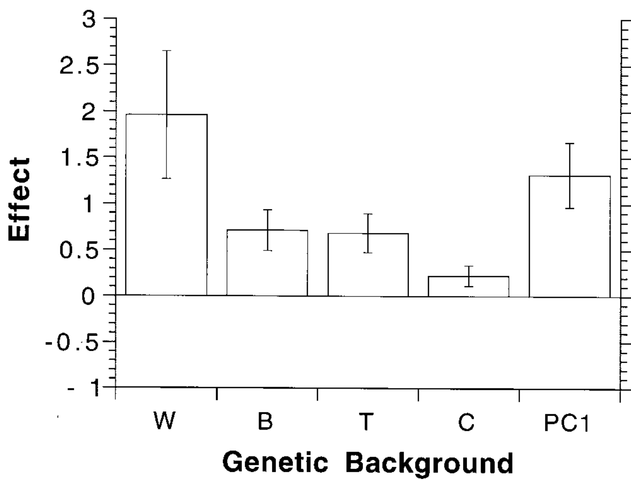
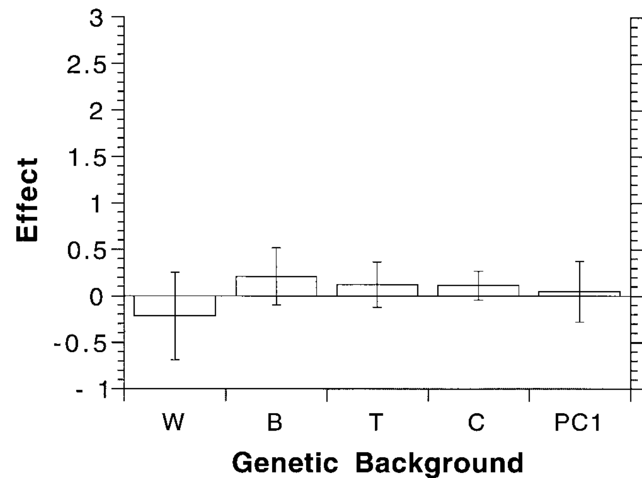


Figure 5.—Sternopleural bristle number effects associated with molecular marker Ha + 8.6 (a *Hae*III site in Intron 2) in each of the genetical backgrounds in which it was measured (*W*, *B*, *T*, and *C*), and the effect associated with the first principal component score (*PC1*). The top panel gives estimates of effects in males and the bottom in females. Effects are the mean of all the lines having a value of 1 for that marker (presence of a band or restriction site) minus the mean of all the lines having a 0 for that marker (absence of a band or restriction site), with error bars representing one standard error on the estimate of the effect (see text). All effects are expressed in sternopleural bristles except those associated with the principal component scores. Effects associated with marker Ha + 8.6 are very consistent over sexes.



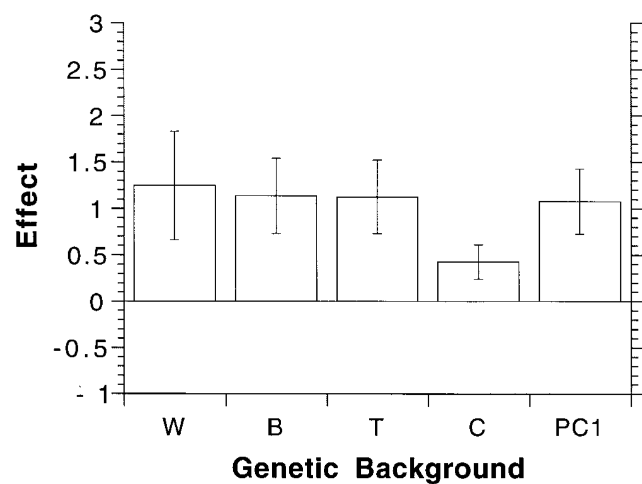


Figure 6.—Abdominal bristle number effects associated with molecular marker S + 18.6 (a *Scr*FI site in Intron 5) in each of the genetical backgrounds in which it was measured (*W*, *B*, *T*, and *C*) and the effect associated with the first principal component score (*PC1*). The top panel gives estimates of effects in males and the bottom in females. Effects are the mean of all the lines having a value of 0 for that marker minus the mean of all the lines having a 1 for that marker, with error bars representing one standard error on the estimate of the effect (see text). All effects are expressed in abdominal bristles except those associated with the principal component scores. Effects associated with S + 18.6 are only significantly different from zero in females, that is, the effect of S + 18.6 appears sex-limited.

background genetic standard deviations. Similarly, for sternopleural bristle number averaged over sexes, the effect of Ha + 8.6 is 1.06 whole third chromosome background genetic standard deviations, and 1.26 homozygous backcross background genetic standard deviations. The two molecular markers associated with bristle number variation may also contribute a significant amount to standing genetic variation in bristle number. The additive genetic variation in abdominal bristle number attributable to S + 18.6 in females is 5.7% and 8.2% of the total genetic variation due to homozygous third chromosomes and homozygous backcross segments, respectively. Similarly, the additive genetic variation in sternopleural bristle number attributable to Ha + 8.6 averaged over sexes is 12.4% and 16.6% of the total genetic variation due to homozygous third chromosomes and homozygous backcross segments, respectively. Estimates of variance attributable to these sites is expected to be greater in the homozygous backcross background than the homozygous entire third chromosome background since the total genetic variation due

TABLE 3

Variance components associated with significant marker alleles over all genetic backgrounds

Variance compounds[a]	Genetic background			
	W	*B*	*T*	*C*
Abdominal bristle number				
S + 18.6 (females only)				
Marker	0.147	0.094	0.139	0.018
Total genetic	2.568	1.142	0.692	0.142
Percentage	5.7	8.2	20.1	12.5
Sternopleural bristle number				
Ha + 8.6 (averaged over sexes)				
Marker	0.362	0.046	0.045	0.006
Total genetic[b]	2.917	0.277	0.134	0.017
Percentage	12.4	16.6	33.9	34.9

[a] Variance components were estimated using the VARCOMP procedure of SAS and then set equal to their expectations under a strictly additive model of quantitative variation.

[b] For all genetic backgrounds, the effect associated with sex*-marker was very small or zero, so it was set equal to zero for the calculations here.

to backcross chromosomes is less than that due to homozygous entire third chromosomes, possibly reflecting the presence of other factors on the third chromosome located outside the introgressed region.

Polymorphic inversions and variation in bristle number: All the lines included in the survey were Standard cytotype, with the exception of lines 38 [In(3R)86D-90D], 94 [In(3R)Payne], and 97 [In(3R)92E-100E]. In (3R)Payne is a common cosmopolitan inversion often observed in collections of *D. melagaster* from the wild, and the other two inversions are unique endemics Lemeunier and Aulard 1992). Data were reanalysed to determine if the associations observed between DNA polymorphisms in the *Delta* gene region and variation in bristle number could be due to the presence of inversions. The data were re-analysed with the combined effect of the three inversion containing lines statistically removed from the data, as well as with the three inversion containing lines dropped from the analysis. In both cases, the *P*-values associated with marker Ha + 8.6 changed by less than one percent. Those associated with S + 18.6 changed less than 1% when the effect of inversions was statistically removed and by 3.3 percent when the three inversion containing lines were dropped from the analysis. In addition, the *P*-values associated with an analysis of variance of the effect of inversion (and an inversion by sex interaction) on variation in abdominal and sternopleural bristle number did approach statistical significance.

DISCUSSION

Two polymorphic markers in the *Delta* gene region are significantly associated with bristle variation: Ha + 8.6 in Intron 2 is associated with variation in sternopleural bristle number in both males and females and S + 18.6 in Intron 5 is associated in a sex-limited manner with abdominal bristle number in females. These two sites are not in linkage disequilibrium with one another, and the effect associated with each site is apparently unaffected by statistical removal of the other. Thus, these two sites appear to be independently acting on variation in bristle number, despite their close physical proximity (10 kb), and likely mode of action through the same gene product (*Dl*). The observation of at least two sites at the *Delta* locus contributing to standing variation may be inconsistent with discussions in the literature regarding how adaptation may occur. In these discussions *genes* (or mapped factors) of large effect are often equated to *mutations* of large effect (Orr and Coyne 1992). Although adaptation may appear to be due to a single large effect genetic locus, the effect of this locus may result from multiple substitutions over evolutionary time. As no genes contributing to adaptive differences between species have been dissected at the resolution necessary to distinguish between one or multiple substitution events at the same locus, we cannot conclude at this time how evolution proceeds. The thesis that multiple substitutions at the same locus may contribute to short-term evolutionary change appears supported by these observations at *Delta*, where at least two sites contribute to standing variation in bristle number. Additional support comes from previous studies on *Adh*, where at least three sites in the *Adh* gene region contribute to standing variation in *Adh* protein levels (Stam and Laurie 1996).

The estimated effects associated with both Ha + 8.6 and S + 18.6 are large relative to standing genetic variation in bristle number; likely too large to be consistent with Gaussian allelic effects models for the maintenance of additive genetic variation by the mutation selection balance theory (Barton and Turelli 1989). It is also unlikely that the QTNs detected in this study are rare. For example, if it is assumed the QTN affecting sternopleural bristle number detected by Ha + 8.6 in the *B* background is at a frequency of less than 5% in the sample, then for the effect associated with this QTN to be consistent with a 95% lower bound on the estimated effect of Ha + 8.6 the QTN's effect would have to be on the order of four backcross genotypic standard deviations {see Lai *et al.* 1994, footnote #24, frequency of Ha + 8.6 is 0.40, 95% lower bound is [effect of Ha + 8.6 − (1.96 × standard error from Figure 5)], genotypic standard deviation from Table 3}. QTNs of this magnitude would have been detected by inspection of the phenotypic data, thus the QTNs detected in this study are likely too common to be consistent with House of Cards models for the maintenance of additive genetic variation by the mutation selection balance theory (Barton and Turelli 1989). Recent experiments suggest that there are detectable levels of stabilizing selection on bristle number, so it is difficult to conclude that the variation of this study is neutral with respect to fitness

(Linney *et al.* 1971; Nuzhdin *et al.* 1995; Garcia-Dorado and Gonzalez 1996). It is possible to explain the presence of variants of large effect at intermediate frequency if the net effect of such variants with respect to fitness is zero, but then it is difficult to explain the apparent stabilizing selection on the trait (Keightley and Hill 1990). Selective models, with the variants being maintained as balanced polymorphisms (possibly aided by spatial or temporal variation in the environment), or through antagonistic pleiotropy, seem more amenable to further empirical testing at this time. Antagonistic pleiotropy is a particularly attractive hypothesis to test in the case of S + 18.6, where allelic effects are different in the two sexes, as this may represent an example of sex-specific trade-offs.

The estimated bristle number effects associated with Ha + 8.6 and S + 18.6 are smaller than those estimated via a QTL mapping approach of earlier work (1.4 sternopleural bristles averaged over sexes, and 3.6 and 2.4 abdominal bristles in males and females, respectively; Long *et al.* 1995). The discrepancy in the estimates of effects from the two experiments does not necessarily imply these two studies are inconsistent with one another. QTL mapping tends to over-estimate effects if there are linked QTLs, or multiple sites at the same locus, affecting a character. In contrast, it can be shown that association studies tend to under-estimate the variance attributable to a QTN by a factor of $1/R^2$, where R^2 is the correlation between marker and QTN (Lai *et al.* 1994, footnote #24). Even considering this discrepancy, the large differences in estimated effects makes it appear that the factors detected in this association study do not completely correspond to those detected in earlier QTL mapping work. For example, the QTL affecting abdominal bristle number that mapped to the *Delta* region (Long *et al.* 1995) had a larger effect in males than in females, whereas the factor detected in this association study has no effect in males and a large effect in females. It seems likely that the discrepancy in estimates of effects using the two different approaches to characterizing quantitative variation is due to one, or a combination of, the following reasons: (1) QTLs linked to (but not at) the *Delta* locus led to inflated estimates of effects in the earlier QTL mapping experiment (another candidate gene known to affect bristle number, *Hairless*, is only 3.5 cM from *Delta*), (2) the selection experiment used to generate the high and low parental strains used in the QTL mapping experiment may have fixed variants at the *Delta* locus that are rare or not represented in the sample of 55 random chromosomes used in this study, or (3) differences in estimates of effects from the two experiments result from epistatic interactions between factors at *Delta* and loosely linked factors that exist in the set of third chromosome recombinant inbred lines of the earlier QTL mapping experiment, but do not exist in the backcross lines of the current experiment. The possibility that the difference between estimates of allelic effects in the two studies is due to epistatic interactions among factors is plausible, as significant epistatic effects were observed in the earlier study of Long *et al.* (1995).

Two previous studies have detected associations between DNA polymorphisms at candidate loci important in peripheral nervous system development and variation in bristle number. Mackay and Langley (1990) found that insertions of transposable elements in the *achaete-scute* region were associated with a reduction in bristle number, consistent with *achaete-scute* weak loss-of-function mutations. As the frequency of any given transposable element insertion is rare in samples from natural populations, these results are consistent with deleterious mutation-selection equilibrium models for the maintenance of quantitative variation. A second study found that DNA polymorphisms in the *scabrous* gene region were associated with variation in bristle number more often than one would expect by chance alone, and that insertions of transposable elements in the *scabrous* region were not associated with bristle number variation (Lai *et al.* 1994). These results appeared more consistent with models incorporating balancing selection than deleterious mutation-selection models, as the sites that were associated with bristle number variation were generally common. The statistical significance of the associations observed in the *scabrous* study was assessed by permutation testing the *number of markers significant* at a marginal significance level of $P < 0.05$, yielding a set of marginally significant markers. In the present study, significance was assessed by permutation testing the F-statistic associated with the *most significant marker*, yielding a prioritized set of associated markers. The performance of these two statistical tests under different models of quantitative variation is an interesting topic for further investigation.

The protein product of *Delta* is conserved in both sequence and function, is utilized in a number of different tissues and times during development, yet contributes to standing variation in bristle number: The product of the *Delta* gene is the ligand, and that of the *Notch* gene is the receptor, in the well-characterized *Notch* signal transduction pathway (Artavanis-Tsakonas *et al.* 1995). Bristle patterning and spacing is believed to be the result of a process known as lateral inhibition, whereby cell-cell interactions cause feed back loops between cells expressing *Notch* and *Delta,* resulting in some cells assuming a neuronal fate (*i.e.*, a bristle) and others a non-neuronal fate (Heitzler and Simpson 1991). *Delta* homologues have also been identified in human, mouse, chick, Xenopus, and *Caenorhabditis elegans*, where the protein sequence is conserved, as is its apparent role in the *Notch* signaling pathway. In addition, the function of *Delta*, in defining a commitment to a neuronal versus non-neuronal fate, appears to be conserved in Xenopus (Chitnis *et al.* 1995), chick (Henrique *et al.* 1995), and mice (Hrabe de Angelis *et al.*

1997; in mice *Delta* appears to have assumed an additional role in the maintenance of somite borders). *Delta-Notch* signaling is involved in many important developmental processes in addition to bristle patterning and spacing, as evidenced by the pleiotropy observed for different *Delta* mutants (Vassin *et al.* 1987; Alton *et al.* 1988) and the many different times and tissues in which *Delta* is expressed during development (Haenlin *et al.* 1990; Parks and Muskavitch 1993). The observation that the *Delta* region has very few insertions of transposable elements (only three small insertions) relative to other loci that have been surveyed in Drosophila is consistent with the idea that the *Delta* region has a number of cis-acting regulatory sequences for which a change in physical distance from each other and/or the *Delta* transcription unit is strongly deleterious (*i.e.*, insertions in the *Delta* region are quickly eliminated in natural populations). The large number of *Delta* spontaneous mutants that have been identified suggests that the *Delta* locus is as much of a target for transposable element insertions as other loci (Lindsley and Zimm 1992). The typical levels of DNA heterozygosity observed for the *Delta* region provides no support for the hypothesis that *Delta* has experienced recent selective sweeps (Kaplan *et al.* 1989).

The two sites, Ha + 8.6 and S + 18.6, that are significantly associated with bristle variation are located in the large second (4.0 kb) and fifth (2.9 kb) introns of *Delta*, respectively. Although the permutation testing approach we describe does not localize the causative site, physical locations close to Ha + 8.6 and S + 18.6 are more likely to harbor the causative site than locations further away. It is likely that the QTNs we have localized to somewhere in the *Delta* region are in one of the introns or the translated portion of *Delta*, rather than either the 5′ or 3′ untranslated flanking regions. It is possible that the *Delta* variants that are responsible for variation in bristle number are site polymorphisms in binding domains for proteins that regulate *Delta* expression levels. The first intron of *Delta* is known to contain an enhancer(s) of expression in the embryo (Haenlin *et al.* 1994), so it is possible that the other introns also have, as yet unidentified, enhancer elements important later in development (0–21 hr and 7–30 hr after puparium formation for macrochaetae and microchaetae, respectively) when normal *Delta* expression is important for bristle spacing (Parks and Muskavitch 1993). Our data do not necessarily implicate such regulatory variants as contributing to standing variation though, as we can not rule out polymorphisms resulting in amino acid changes, or variants that affect splicing efficiency of the *Delta* pre-mRNA. With regard to the last possibility, a peculiar feature of *Delta* transcription is that the introns of *Delta* are not immediately degraded after being spliced out of the pre-mRNA (as with other genes), and also remain localized to the sites of transcription (Kopczynski and Muskavitch 1992).

The permutation testing approach provides a robust and general method for testing for associations between molecular markers and phenotypic variation: In addition to applying the permutation testing procedure to F-statistics derived from a univariate analysis of pricipal component scores, we also applied the permutation testing procedure to likelihood ratio test statistics estimated from applying a multivariate general linear model to the data (Kendall and Stuart 1979; results not shown). The two approaches gave comparable results, although the approach based on the likelihood ratio seemed less powerful than that based on the principal component scores. The difference in power is likely due to a biological constraint that was difficult to incorporate into the multivariate likelihood ratio test statistic. Although we may expect the phenotypic effect of a causative site in linkage disequilibrium with one of our marker sites to have different effects in the different genetic backgrounds (and hence response variables) studied, we expect the direction of the effects to be the same in all genetic backgrounds. Since the multivariate likelihood ratio test statistic is completely a function of the covariance structure of the error (or residual) matrix that results from fitting a specific model, it follows that a marker whose effects are in different directions in the different genetic backgrounds, but always large, would not weaken the model (*i.e.*, the column vector of *B*'s associated with the marker effect of the design matrix could have both large positive and negative values without weakening the overall likelihood; Kendall and Stuart 1979). Under almost any genetic model that purports to explain the nature of quantitative variation, the assumption that alleles should act in the same direction in a number of genetic backgrounds will be met. Thus we consider the permutation testing approach based on the principal component scores to be a more powerful statistical approach for detect marker/phenotype associations than the likelihood based test statistic. If it is discovered in the future that the alleles that affect quantitative traits are massively epistatic to one another (enough that genetic background can result in reversals of allelic effects), then statistical analyses based the multivariate likelihood ratio test statistic may be more appropriate. If this were the case, principal components other than the first are expected to account for a significant proportion of the total variation over genetic backgrounds, which was not true for *Delta.*

A molecular marker can be significantly associated with bristle number variation using the permutation test for one of three reasons: (1) there is no QTN in the population and the association between marker and phenotype is due to chance, (2) there is a QTN in the population that is in linkage equilibrium with markers in the *Dl* gene region, and by chance it is in linkage disequilibrium with a marker in the *Dl* region in the sample of this study, or (3) there is a QTN in linkage disequilibrium with sites in the *Dl* region in the popula-

tion and this association is reflected in the sample. The *P*-value associated with the permutation testing procedure is the probability that the association is due to outcomes (1) or (2). Given the small *P*-values of the permutation tests associated with Ha + 8.6 and S + 18.6, the implication is that sites in linkage disequilibrium with polymorphic sites in the *Dl* region in the population are contributing to bristle number variation. Past studies have observed linkage disequilibrium between inversions and allozyme loci. Disequilibrium between allozymes and inversions on the same arm can be large and replicable over different studies, whereas disequilibrium between inversions and allozymes on opposite arms is small or nonexistent (reviewed in Lemeunier and Aulard 1992). As only three of the lines surveyed contained inversions on the right arm of the third chromosome, and these inversions have little impact on observed associations between the two significant markers and variation in bristle number, it is unlikely that the presence of inversions can account for observed marker/phenotype associations. In Drosophila, linkage disequilibrium is rarely observed over large distances in sets of inversion-free chromosomes (*e.g.*, Aquadro *et al.* 1986; Miyashita and Langley 1988), therefore it is unlikely that polymorphic sites in the *Delta* region are in linkage disequilibrium with QTNs outside the *Dl* region. The permutation test indicates that the two molecular markers (Ha + 8.6 and S + 18.6) are likely to be in linkage disequilibrium with a QTN in the surveyed region or nearby.

Can disequilibrium mapping be used to localize the sites that cause phenotypic variation to specific intervals? The permutation testing approach employed in this work, although useful with respect to statistical inference, does not estimate the position of the QTN. The observed level of linkage disequilibrium between a marker site of known position and a site of unknown position that causes a simple Mendelian disease (*e.g.*, cystic fibrosis) can be used to estimate the position of the disease-causing site in equilibrium populations. With only a single marker site the likelihood surface is fairly flat (*i.e.*, a site causing the simple Mendelian trait cannot be localized with much accuracy; Hill and Weir 1994). For QTLs the situation is worse, as the disequilibrium between the molecular marker and the QTN is unknown. The phenotypic variation attributable to a marker site estimates only the product of linkage disequilibrium between the QTN and the marker site, and the variance attributable to the QTN. Likelihood-based approaches that simultaneously consider all markers may eventually allow efficient localization of the sites contributing to quantitative variation, although these approaches do not seem currently applicable (Kuhner *et al.* 1995; Griffiths and Marjoram 1996). Population genetics theory shows that

$$E[\sigma^2_{\text{marker}}] = E[R^2]\sigma^2_{\text{QTN}} = \frac{\sigma^2_{\text{QTN}}}{1 + 4Nc}$$

(Hill and Robertson 1968; Lai *et al.* 1994, extension of footnote #24). In plots of the F statistics for each marker site as a function of its position (*e.g.*, Figures 3 and 4), local peaks in the F statistic with an exponential decay on each side suggest positions of QTNs (since F statistics are proportional to the variance due to marker and physical distance is assumed to be proportional to genetic distance). Unfortunately, equating observed measures of σ^2_{marker} to their expectations and minimizing squared differences does not allow effective localization of the QTN (simulation results not shown). This may be expected, as population genetics theory predicts a large stochastic variance associated with independent evolutionary realizations of R^2. Furthermore, pairs of marker sites are not independent of one another (Figure 1).

The power to detect associations between sites affecting a quantitative trait and molecular variants segregating in a candidate gene region ultimately depends on having a sufficient density of molecular markers to ensure that some markers will be in strong linkage disequilibrium with a QTN. In the four-cutter portion of the survey presented here, we assessed molecular marker/phenotype associations for 41 markers over a region of 26.8 kb, resulting in an average of one polymorphic marker site every 4 Γ. Typical of previous work in Drosophila, linkage disequilibrium falls off quickly with increasing physical distance between markers (*e.g.*, Aquadro *et al.* 1986; Miyashita and Langley 1988; Figure 2). A marker density of at least one marker every Γ is desirable, as such a density implies that most markers would be in linkage disequilibrium with at least one other marker, and hence one or more markers are likely to be in linkage disequilibrium with any QTNs in the region (see Figure 2 and note that Γ is estimated to be 160 bp). In any given study, the likelihood of detecting significant associations will depend on the actual marker sites surveyed and their history relative to the QTN (*e.g.*, the distribution of mutations and recombination events in the history of the sampled genomes). It follows that our experiment may have failed to detect additional sites in the *Delta* region that contribute to bristle number variation because the marker density was too low. The significant markers of this experiment account for 5–12% of the total genetic variation due to homozygous third chromosomes and 8–17% of the total genetic variation associated with the *Delta* region (*i.e.*, the homozygous backcross chromosomes). Since $\sigma^2_{\text{marker}} = R^2\sigma^2_{\text{QTN}}$, if R^2 is small, then the QTN would have an effect large enough to be detected as a segregating variant. Thus it seems likely that Ha + 8.6 and S + 18.6 are both in strong disequilibrium, and physically close to QTNs which are contributing to standing variation in bristle number.

We thank B. Dixon and B. Marsh for technical assistance, Sasha Langley for assistance with cytological work, M. Muscavitz and A. Parks for providing phage clones and other unpublished materials, and Z-B. Zeng, J. Fry, and W. G. Hill for suggestions on how to

analyze the data. This work was supported by a Natural Sciences and Engineering Research Council of Canada Fellowship to A.D.L., and grants GM45344 and GM45146 from the National Institutes of Health to T.F.C.M.

LITERATURE CITED

Alton, A. K., K. Fechtel, A. L. Terry, S. B. Meikle and M. A. Muskavitch, 1988 Cytogenetic definition and morphogenetic analysis of *Delta*, a gene affecting neurogenesis in *Drosophila melanogaster*. Genetics **118:** 235–245.

Aquadro, C. F., S. F. Deese, M. M. Bland, C. H. Langley and C. C. Laurie-Ahlberg, 1986 Molecular population genetics of alcohol dehydrogenase gene region of *Drosophila melanogaster*. Genetics **114:** 1165–1190.

Artavanis-Tsakonas, S., K. Matsuno and M. E. Fortini, 1995 *Notch* signaling. Science **268:** 225–232.

Barnes, W. M., 1994 PCR amplification of up to 35-kb DNA with high fidelity and high yeild from λ bacteriophage templetes. Proc. Natl. Acad. Sci. USA **91:** 2216–2220.

Barton, N. H., and M. A. Turelli, 1989 Evolutionary quantitative genetics: How little do we know? Annu. Rev. Genet. **23:** 337–370.

Braverman, J. M., R. R. Hudson, N. L. Kaplan, C. H. Langley and W. Stephan, 1996 The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics **140:** 783–796.

Chatfield, C., and A. J. Collins, 1980 *Introduction to Multivariate Analysis*. Chapman and Hall, New York.

Cheng, S., C. Fockler, W. M. Barnes and R. Higuchi, 1994 Effective amplification of long targets from cloned inserts and human genomic DNA. Proc. Natl. Acad. Sci. USA **91:** 5695–5699.

Chitnis, A., D. Henrique, J. Lewis, D. Ish-Horowicz and C. Kintner, 1995 Primary neurogenesis in *Xenopus* embryos regulated by a homologue of the *Drosophila* neurogenic gene *Delta*. Nature **375:** 761–766.

Church, G. M., and W. Gilbert, 1984 Genomic sequencing. Proc. Natl. Acad. Sci. USA **81:** 1991–1995.

Churchill, G. A., and R. W. Doerge, 1994 Empirical threshold values for quantitative trait mapping. Genetics **138:** 963–971.

Doerge, R. W., and G. A. Churchill, 1996 Permutation tests for multiple loci affecting a quantitative character. Genetics **142:** 285–294.

Engels, W. R., 1981 Estimating genetic divergence and genetic variability with restriction endonucleases. Proc. Natl. Acad. Sci. USA **78:** 6329–6333.

Falconer, D. S., and T. F. C. Mackay, 1996 *Introduction to Quantitative Genetics*, 4th ed., Addison Wesley Longman, Harlow, Essex.

Garcia-Dorado, A., and J. A. Gonzalez, 1996 Stabilizing selection detected for bristle number in *Drosophila melanogaster*. Evolution **50:** 1573–1578.

Griffiths, R. C., and P. Marjoram, 1996 Ancestral inference from samples of DNA sequences with recombination. J. Comp. Bio. **3:** 479–502.

Haenlin, M., B. Kramatschek and J. A. Campos-Ortega, 1990 The pattern of transcription of the neurogenic gene *Delta* of *Drosophila melanogaster*. Development **110:** 905–914.

Haenlin, M., M. Kunisch, B. Kramatschek and J. A. Campos-Ortega, 1994 Genomic regions regulating early embryonic expression of the *Drosophila* neurogenic gene *Delta*. Mech. Dev. **47:** 99–110.

Heitzler, P., and P. Simpson, 1991 The choice of cell fate in the epidermis of *Drosophila*. Cell **64:** 1083–1092.

Henrique, D., J. Adam, A. Myat, A. Chitnis, J. Lewis *et al.*, 1995 Expression of a *Delta* homologue in prospective neurons in the chick. Nature **375:** 787–790.

Hill, W. G., and A. Robertson, 1968 Linkage disquilibrium in finite populations. Theor. Appl. Genet. **38:** 226–231.

Hill, W. G., and B. S. Weir, 1994 Maximum-likelihood estimation of gene location by linkage disequilibrium. Am. J. Hum. Genet. **54:** 705–714.

Hrabe de Angelis, M., J. McIntyre III and A. Gossler, 1997 Maintenance of somite borders in mice requires the *Delta* homologue *Dll1*. Nature **386:** 717–720.

Hudson, R. R., 1982 Estimating genetic variability with restriction endonucleases. Genetics **100:** 711–719.

Hudson, R. R., 1987 Estimating the recombination parameter of a finite population model without selection. Genet. Res. **50:** 245–250.

Jan, Y. N., and L. Y. Jan, 1994 Genetic control of cell fate specification in *Drosophila* peripheral nervous system. Ann. Rev. Genet. **28:** 373–93.

Jorde, L. B., 1995 Linkage disequilibria as a gene mapping tool. Am. J. Hum. Genet. **56:** 11–14.

Jowett, T., 1986 Preparation of nucleic acids, pp. 275–286 in *Drosophila: A Practical Approach*, edited by D. B. Roberts. IRL Press, Oxford.

Kaplan, N. L., R. R. Hudson and C. H. Langley, 1989 The "hitchhiking effect" revisited. Genetics **123:** 887–899.

Keightley, P. D., and W. G. Hill, 1990 Variation maintained in quantitative traits with mutation-selection balance: pleiotropic side-effects on fitness traits. Proc. R. Soc. Lond. Ser. B **242:** 95–100.

Kendall, M., and A. Stuart, 1979 *The Advanced Theory of Statistics*. MacMillan Publishing Co., New York.

Kopczynski, C. C., A. K. Alton, K. Fechtel, P. J. Kooh and M. A. Muskavitch, 1988 *Delta*, a *Drosophila* neurogenic gene, is transcriptionally complex and encodes a protein related to blood coagulation factors and epidermal growth factor of vertebrates. Genes Dev. **2:** 1723–1735.

Kopczynski, C. C., and M. A. Muskavitch, 1992 Introns excised from the *Delta* primary transcript are localized near sites of *Delta* transcription. J. Cell Biol. **119:** 503–512.

Kreitman, M., and M. Aguade, 1986 Genetic uniformity in two populations of *Drosophila melanogaster* as revealed by filter hybridization of four-nucleotide-recognizing restriction enzyme digests. Proc. Natl. Acad. Sci. USA **83:** 3562–3566.

Kuhner, M. K., J. Yamato and J. Felsenstein, 1995 Estimating effective population size and mutation rate from sequence data using Metropolis-Hastings sampling. Genetics **140:** 1421–1430.

Kunisch, M., M. Haenlin and J. A. Campos-Ortega, 1994 Lateral inhibition mediated by the *Drosophila* neurogenic gene *Delta* is enhanced by proneural proteins. Proc. Natl. Acad. Sci. USA **91:** 10139–10143.

Lai, C., R. F. Lyman, A. D. Long, C. H. Langley and T. F. Mackay, 1994 Naturally occurring variation in bristle number and DNA polymorphisms at the *scabrous* locus of *Drosophila melanogaster*. Science **266:** 1697–1702.

Lemeunier, F., and S. Aulard, 1992 Inversion polymorphisms in *Drosophila melanogaster*, pp. 339–405 in *Drosophila Inversion Polymorphism*, edited by C. B. Krimbas and J. R. Powell. CRC Press, Boca Raton, FL.

Lindsley, D. L., and G. G. Zimm, 1992 *The Genome of Drosophila melanogaster*. Academic Press, San Diego.

Linney, R., B. W. Barnes and M. J. Kearsey, 1971 Variation for metrical characters in *Drosophila melanogaster*. III. The nature of selection. Heredity **27:** 163–174.

Long, A. D., S. L. Mullaney, L. A. Reid, J. D. Fry, C. H. Langley *et al.*, 1995 High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. Genetics **139:** 1273–1291.

Long, A. D., S. L. Mullaney, T. F. C. Mackay and C. H. Langley, 1996 Genetic interactions between naturally occurring alleles at quantitative trait loci and mutant alleles at candidate loci affecting bristle number in *Drosophila melanogaster*. Genetics **144:** 1497–1510.

Long, A. D., M. N. Grote and C. H. Langley, 1997 Genetic analysis of complex diseases. Science **275:** 1328.

Lyman, R. F., and T. F. C. Mackay, 1998 Candidate quantitive trait loci and naturally occurring phenotypic variation for bristle number in *Drosophila melanogaster*: the *Delta-Hairless* gene region. Genetics **149:** 993–998.

Mackay, T. F. C., and J. D. Fry, 1996 Polygenic mutation in *Drosophila melanogaster*: genetic interactions between selected lines and candidate quantitative trait loci. Genetics **144:** 671–688.

Mackay, T. F., and C. H. Langley, 1990 Molecular and phenotypic variation in the *achaete-scute* region of *Drosophila melanogaster*. Nature **348:** 64–66.

Miyashita, N., and C. H. Langley, 1988 Molecular and phenotypic variation of the *white* locus region in *Drosophila melanogaster*. Genetics **120:** 199–212.

Nei, M., and W. Li, 1979 Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA **76:** 5269–5273.

Nei, M., and F. Tajima, 1981 DNA polymorphism detectable by restriction endonucleases. Genetics **97:** 145–163.

Nuzhdin, S.V., J.D. Fry and T. F. C. Mackay, 1995 Polygenic mutation in *Drosophila melanogaster*: the causal relationship of bristle number to fitness. Genetics **139:** 861–872.

Orr, H. A., and J. A. Coyne, 1992 The genetics of adaptation—a reassessment. Am. Nat. **140:** 725–742.

Parks, A. L., and M. A. Muskavitch, 1993 *Delta* function is required for bristle organ determination and morphogenesis in *Drosophila.* Dev. Bio. **157:** 484–496.

Parks, A. L., S. S. Huppert and M. A. Muskavitch, 1997 The dynamics of neurogenic signalling underlying bristle development in *Drosophila melanogaster.* Mech. Dev. **63:** 61–74.

Risch, N., and K. Merikangas, 1996 The future of genetic studies of complex human diseases. Science **273:** 1516–1517.

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989 *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Southern, E., 1975 Detection of specific sequences among DNA sequences separated by gel electrophoresis. J. Mol. Biol. **98:** 537–544.

Stam, L. F., and C. C. Laurie, 1996 Molecular dissection of a major gene effect on a quantitative trait: the level of *alcohol dehydrogenase* expression in *Drosophila melanogaster.* Genetics **144:** 1559–1564.

Tajima, F., 1989 Evolutionary relationship of DNA sequences in finite populations. Genetics **105:** 437–460.

Vassin, H., K. A. Bremer, E. Knust and J. Campos-Ortega, 1987 The neurogenic gene *Delta* of *Drosophila melanogaster* is expressed in neurogenic territories and encodes a putative transmembrane protein with EGF-like repeats. EMBO J. **6:** 3431–3440.

Weir, B. S., 1996 *Genetic Data Analysis II.* Sinauer Associates, Inc., Sunderland, Mass.

Weir, B. S., and W. G. Hill, 1986 Nonuniform recombination within the human β-globin gene cluster. Am. J. Hum. Genet. **38:** 776–778.

Communicating editor: P. D. Keightley

APPENDIX

Phenotype and polymorphic restriction site data for the *Delta* region

Line	Abdominal bristles								Sternopleural bristles							
	W		*B*		*T*		*C*		*W*		*B*		*T*		*C*	
	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
2	20.45	23.25	18.95	20.90	25.68	28.45	17.18	20.28	24.10	24.75	19.43	21.20	20.00	22.93	18.58	19.23
4	18.20	19.30	19.05	21.03	25.38	25.65	17.28	20.08	19.30	19.70	18.45	19.63	19.78	22.23	17.98	19.10
5	18.10	18.35	18.83	22.50	24.38	27.50	17.00	20.05	21.50	20.95	19.18	19.80	19.95	23.13	18.33	19.18
6	22.55	25.90	19.25	22.95	25.85	30.68	17.45	21.03	23.05	21.90	19.03	19.78	19.75	23.33	18.43	19.10
8	19.95	24.10	19.13	21.95	25.03	29.58	17.35	20.48	23.70	24.95	19.43	20.30	21.15	22.58	18.43	19.30
9	22.50	25.30	20.90	25.45	27.68	33.05	18.45	21.55	19.25	20.45	20.80	21.45	20.85	23.20	19.98	19.78
10	22.95	27.05	21.70	24.23	26.20	30.85	18.35	21.33	25.00	25.55	19.73	20.43	21.68	23.63	19.60	19.45
12	20.10	20.00	*	*	25.63	28.80	16.73	20.20	25.30	24.70	*	*	20.90	23.00	18.40	18.95
13	20.30	21.85	17.95	21.40	25.03	28.05	17.40	19.65	19.35	19.30	18.08	19.10	19.15	21.73	18.10	18.60
14	19.45	20.45	19.23	22.00	25.30	28.93	17.50	20.20	26.75	27.25	19.13	20.13	20.55	23.20	19.15	18.85
15	21.40	24.45	19.38	23.15	25.48	29.53	17.60	21.50	22.60	22.15	20.03	21.18	21.15	23.63	18.78	19.58
16	21.05	23.75	20.13	24.80	26.35	32.33	17.95	21.25	20.95	20.75	18.73	19.80	19.70	22.05	18.50	19.03
17	20.60	21.95	20.15	23.33	26.90	30.48	17.88	20.83	22.95	22.10	20.55	22.20	22.28	24.65	19.35	19.83
19	20.15	21.55	17.23	20.43	24.53	28.33	16.33	19.13	23.05	21.80	19.10	19.63	19.73	22.60	18.05	19.40
22	19.30	20.05	19.30	22.93	24.95	29.43	17.35	20.80	20.65	20.75	18.35	19.20	19.65	21.20	18.55	18.75
23	20.75	24.75	18.40	22.10	25.40	29.90	17.58	20.80	21.10	22.65	18.10	19.38	19.48	22.93	18.18	18.73
24	20.00	24.70	18.55	22.05	24.10	29.05	17.43	20.83	24.95	26.60	19.08	19.18	20.20	22.20	19.30	19.65
31	18.20	20.10	18.95	21.48	25.03	28.45	17.95	19.83	16.10	16.70	17.05	17.95	18.83	20.30	17.23	17.93
32	20.35	21.85	19.80	21.88	24.98	29.58	17.23	20.08	20.05	21.00	18.45	19.40	20.15	23.30	18.43	18.88
33	17.45	21.20	18.78	22.73	24.60	29.30	17.10	20.53	18.95	19.00	18.53	18.75	19.63	21.88	18.03	18.55
34	20.80	22.80	17.98	23.13	24.85	31.08	17.08	20.53	19.40	18.95	18.08	19.28	19.93	22.53	18.30	19.40
36	19.35	23.80	20.30	22.33	26.18	28.25	17.93	20.68	20.70	21.35	18.50	19.23	20.23	22.33	18.75	19.25
38	20.55	22.60	19.13	23.73	24.03	30.43	17.10	21.00	19.60	19.70	18.80	19.88	20.20	22.58	18.45	19.28
39	20.50	23.65	19.05	23.00	25.25	28.75	17.53	20.43	18.10	18.40	17.90	18.63	19.85	21.40	17.83	18.75
41	18.35	21.25	18.68	23.95	23.78	30.45	16.65	21.28	21.75	21.15	18.73	18.80	19.33	21.80	18.08	18.75
44	21.45	23.75	19.88	23.15	24.73	29.65	17.08	20.43	18.95	19.40	19.95	19.45	19.73	21.95	18.98	18.80
45	19.70	20.25	19.33	22.58	25.65	28.25	17.73	20.43	18.35	17.85	18.30	19.00	20.38	22.63	18.23	18.90
46	19.10	20.35	17.68	19.95	24.43	24.88	17.00	19.18	15.85	15.95	18.73	19.23	19.63	21.53	18.28	18.68
48	17.95	22.20	18.80	23.05	24.58	28.40	17.08	20.38	21.95	22.05	19.33	20.40	20.45	22.38	17.98	18.93
50	18.55	21.85	18.45	22.33	24.08	28.25	16.90	20.33	22.15	21.80	20.20	21.50	21.25	23.68	19.73	19.75
51	19.25	21.55	19.03	23.30	25.10	29.53	16.58	20.00	18.70	18.65	17.80	18.80	18.98	21.08	17.90	18.25
53	21.00	24.00	18.88	22.73	25.78	30.98	17.45	19.95	18.85	17.65	18.15	19.33	19.73	21.98	18.53	18.98
55	22.25	23.50	20.45	24.45	24.65	27.60	16.83	19.98	22.05	21.65	19.60	21.20	20.85	23.48	18.65	19.05
58	19.75	20.75	18.40	19.93	25.28	28.70	17.20	19.75	24.40	23.85	17.35	18.78	19.45	21.08	18.08	18.80
59	22.40	25.50	20.20	24.45	25.85	30.85	17.68	21.63	23.25	22.45	18.33	20.03	20.50	23.08	19.08	19.53
74	16.65	15.10	21.03	23.33	25.25	27.98	17.63	19.80	23.20	22.85	19.60	21.18	20.70	22.73	17.98	19.58
75	21.00	22.50	19.75	22.38	24.85	28.40	16.68	19.98	21.75	22.50	19.10	19.78	20.60	23.38	18.30	18.40
81	20.45	23.90	18.50	22.98	24.43	29.30	16.90	19.55	24.20	27.15	18.90	20.08	19.95	22.43	18.75	19.33
84	20.80	23.30	19.30	21.23	26.25	27.98	17.60	20.33	23.10	24.55	20.13	20.75	20.83	24.23	18.35	19.23
85	20.45	22.00	18.83	20.73	25.63	28.18	17.28	20.75	19.20	19.55	18.35	19.80	20.05	22.60	17.80	19.00
86	21.80	24.60	19.23	24.43	24.58	31.18	17.93	21.55	19.35	18.65	18.33	18.75	20.05	22.13	18.35	18.50
87	21.45	24.25	19.60	22.85	25.53	29.95	17.90	20.38	21.80	22.90	19.30	19.53	20.03	22.55	18.48	18.75
94	19.60	22.30	*	*	25.18	28.68	17.53	19.53	21.90	22.30	*	*	21.33	23.15	19.25	19.80
95	21.20	21.00	16.63	17.75	23.98	25.88	15.88	19.00	20.65	21.75	19.15	18.58	19.70	22.30	17.83	18.90
97	22.10	25.00	21.85	26.13	24.45	30.28	18.95	22.40	22.05	22.35	18.40	19.98	19.88	22.48	18.08	18.95
101	26.65	25.05	21.68	23.03	26.03	28.58	18.68	21.35	23.60	24.75	19.73	21.20	21.30	22.83	19.05	19.50
105	18.65	20.75	18.05	22.13	23.15	29.05	16.75	20.10	21.15	21.10	17.95	19.45	19.55	21.98	18.38	19.15
106	22.55	25.35	20.10	21.60	24.98	29.13	17.65	19.90	23.65	23.65	19.40	20.30	21.18	22.83	18.88	19.48
107	20.35	22.10	21.45	23.20	26.25	29.63	18.10	21.08	18.40	18.25	18.38	19.78	19.83	22.18	17.85	18.88
111	18.75	19.30	17.83	18.95	24.08	28.20	16.75	19.35	21.05	20.80	18.80	20.05	19.88	22.45	18.43	18.78
113	17.50	19.00	18.30	21.43	25.73	28.50	16.78	20.25	23.65	23.55	19.28	20.30	20.33	22.90	18.40	19.40
115	15.85	16.85	18.55	21.20	25.38	27.75	17.80	19.53	21.90	22.30	18.08	19.43	19.60	22.68	18.58	18.75
116	19.85	23.15	21.80	26.30	26.28	29.83	17.90	20.33	21.60	21.90	21.38	22.53	21.50	23.00	18.58	19.08
117	20.95	22.80	19.03	24.28	25.55	30.30	17.93	21.33	22.00	22.35	18.18	19.50	20.50	23.13	18.03	18.83
119	21.65	22.00	19.78	20.48	25.38	26.40	17.18	19.15	22.45	22.55	19.40	20.38	20.75	23.55	18.90	19.08

This section of the table gives mean abdominal and sternopleural bristle numbers in both males and females for each of the lines surveyed in each of the four genetic backgrounds studied. In the analyses of this paper the eight cells with missing data were replaced with the means for that sex/genetic background/bristle trait. The following sections give molecular marker genotype for markers in the *Delta* gene region;

APPENDIX *continued*

Line	Sites																			
	5′									5′ reg										
	Hi-24.7	Hi-24.2	P-23.9	E-19.7	E-17.2	E-11.2	P-9.1	R-7.2	B-4.8	Ha-3.0	A-2.9	Ha-2.5	C-2.4	Ha-2.0	Ha-1.9	C-1.6	Ha-1.3	C-1.1	C-0.7	B+0.0	Ha+1.0
2	0	0	1	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	1	1	0
4	0	0	0	1	1	0	0	1	1	0	0	1	0	1	1	0	1	0	0	1	0
5	1	0	0	1	1	0	0	1	1	1	0	1	0	1	1	0	1	1	1	0	0
6	1	0	1	0	1	0	0	1	1	1	0	0	1	1	0	0	0	1	1	1	0
8	1	0	1	1	1	0	0	1	1	0	0	0	0	1	0	0	0	0	1	0	0
9	0	0	1	0	0	0	0	1	1	*	0	0	0	0	0	0	0	0	1	1	0
10	1	0	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	1	0
12	1	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	1	1	0
13	0	0	1	0	1	0	0	1	1	1	0	0	1	1	0	0	0	1	1	0	0
14	0	0	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	1	0	0
15	0	0	1	1	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	1	1
16	1	0	1	1	1	0	1	1	1	0	1	1	0	0	0	0	0	1	1	1	0
17	1	1	1	0	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	1	0
19	0	0	1	0	1	0	0	1	1	1	0	0	1	1	0	0	0	1	1	0	0
22	1	0	1	1	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	1	0
23	1	0	1	1	1	0	1	0	1	1	0	0	1	1	0	0	0	1	1	0	0
24	0	0	1	0	1	0	0	1	1	1	0	0	1	1	0	0	0	1	1	1	0
31	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	1	0
32	0	0	1	1	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	1	1
33	0	0	1	1	1	0	1	1	1	0	0	1	0	0	0	0	0	1	1	0	0
34	1	0	1	0	0	0	0	1	1	1	0	1	0	0	1	0	1	0	1	1	0
36	0	0	1	0	1	1	1	1	1	0	0	1	0	0	0	0	0	0	1	0	0
38	1	0	1	0	1	0	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0
39	0	0	1	0	1	0	0	0	1	0	0	1	0	0	0	0	0	1	1	1	0
41	1	0	1	0	0	0	0	1	1	*	0	0	0	0	0	0	0	0	1	1	0
44	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	1	0
45	0	0	0	0	1	0	0	1	1	1	0	0	1	1	0	0	0	1	1	1	0
46	1	0	1	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	1	0
48	0	0	0	0	1	0	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0
50	0	0	1	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0
51	0	1	1	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1
53	0	0	1	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	1	0	1
55	1	0	1	1	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	1	0
58	1	0	1	0	0	1	1	0	1	0	0	1	0	0	0	0	0	1	1	1	0
59	1	0	1	1	0	0	0	1	1	0	0	1	0	0	0	0	0	1	1	0	0
74	1	0	1	0	0	1	1	1	1	0	0	1	0	0	0	0	0	1	1	0	0
75	0	0	1	1	1	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0
81	0	0	1	1	0	0	0	1	1	1	0	1	1	1	1	0	0	0	0	1	0
84	0	0	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	1	0	0
85	0	0	1	1	0	0	0	1	1	1	0	1	1	1	1	0	1	0	0	1	0
86	0	0	1	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	1	1
87	0	0	1	1	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	1	0
94	0	0	1	1	1	1	0	1	1	1	0	0	1	1	1	0	1	1	1	1	1
95	1	0	1	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0
97	0	0	1	0	0	0	1	1	1	1	0	0	0	1	0	0	0	0	0	1	0
101	0	0	1	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	1	1	1
105	0	0	1	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	1	0
106	0	0	1	0	0	0	0	0	1	0	0	1	0	1	1	0	1	1	1	0	0
107	0	0	1	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	1	1	1
111	0	0	1	0	1	0	0	1	1	0	0	1	0	0	1	1	1	0	0	1	1
113	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	1	1	0
115	0	0	1	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	1	1
116	1	0	1	1	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	1	0
117	0	0	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	1	0
119	1	1	1	1	1	0	1	1	1	1	0	0	0	1	0	0	0	1	1	1	0

1 indicates the presence of a site or band and 0 the absence of a site or band. The column heading "Location" gives the relative location in the *Delta* gene region, and the column heading "Site" gives the site name. The site name consists of a one- or two-letter designation indicating the restriction enzyme used to detect that site (A = *Alu*I, B = *Bam*HI, C = *Cfo*I, D = *Dde*I, E = *Eco*RI, Ha = *Hae*III, Hi = *Hin*dIII, Hp = *Hpa*II, P = *Pst*I, R = *Rsa*I, S = *Scr*FI, T = *Taq*I) followed by a number indicating the exact position of the site in the *Delta* region (with some exceptions

APPENDIX *continued*

	Sites: Intron 1											Intron 2				Intron 3						Intron 5								3′ UTR		
Line	Ha+3.5	Ha+3.9	A+4.2	A+4.6	A+4.9	A+5.3	C+5.6	C+6.0	Hi+6.2	Hp+6.3	Hp+6.7	A+8.3	Ha+8.6	D+8.8	A+9.4	A+11.0	C+11.7	C+12.4	C+13.1	Hp+13.8	Hp+14.5	T+18.4	S+18.6	T+19.5	A+19.6	Ha+19.8	R+19.9	R+20.1	T+20.4	C+22.2	T+23.4	ΔC+24.0
2	1	1	0	0	0	0	0	0	1	1	0	0	1	1	1	0	0	0	0	1	1	0	1	0	1	0	0	1	0	0	0	0
4	1	1	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0
5	1	0	1	0	1	0	1	0	1	1	1	0	1	1	1	1	1	0	1	0	0	0	1	0	1	1	0	1	1	0	0	0
6	1	1	1	0	0	0	0	0	0	1	0	0	1	1	1	0	1	0	0	1	1	0	0	0	1	0	0	1	0	1	1	0
8	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	0	0	0	0	1	0	0	0	1	1	0	1	1	0	1	0
9	1	1	0	0	0	0	1	0	1	1	0	*	*	*	*	0	0	0	0	1	1	0	0	0	1	1	0	1	1	0	0	0
10	1	1	0	1	0	0	0	0	1	0	0	0	1	1	1	1	1	0	0	0	1	0	0	0	1	0	0	1	1	0	0	0
12	1	1	0	1	0	1	0	0	0	0	0	0	1	1	1	0	1	0	0	1	0	0	1	0	1	0	0	1	0	0	0	0
13	1	1	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	1	0	1	0	0	1	0	0	0	0
14	0	1	0	0	0	1	1	0	1	1	0	0	0	1	1	0	0	0	0	1	1	0	1	0	1	0	0	1	0	1	1	0
15	1	1	1	0	0	0	0	0	1	1	0	0	1	1	1	1	0	1	1	0	1	0	1	0	0	0	0	0	0	1	1	0
16	0	1	0	0	0	1	1	0	1	1	0	*	*	*	*	0	0	0	0	1	1	0	0	0	1	1	0	1	1	0	0	0
17	1	1	0	0	0	0	0	0	0	1	0	0	1	1	1	0	0	0	0	1	1	0	1	0	1	0	0	1	0	0	0	1
19	1	1	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	1	0	1	0	0	1	0	0	0	0
22	1	1	0	0	0	1	0	0	1	1	0	0	0	1	1	1	1	0	0	0	1	0	1	0	1	1	0	1	1	0	0	0
23	1	1	1	0	1	0	0	1	1	1	1	0	1	1	1	1	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	0
24	1	1	1	0	0	0	0	0	0	1	0	1	1	1	1	0	1	0	0	1	1	0	1	0	1	0	0	1	0	1	1	0
31	1	1	0	0	0	0	1	0	1	1	0	0	0	0	1	0	0	0	0	1	1	0	1	0	1	0	0	1	0	1	1	0
32	1	1	1	0	0	0	1	0	1	1	0	0	1	1	1	1	0	1	1	0	1	0	1	0	1	0	0	0	0	0	0	0
33	0	1	0	0	0	1	1	0	1	1	0	0	0	1	1	0	0	0	0	1	1	0	0	0	1	1	0	1	1	0	0	0
34	0	1	0	0	0	1	1	0	1	1	0	0	0	1	1	0	0	0	0	1	1	0	1	0	0	0	0	1	0	1	1	0
36	1	1	1	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	1	1	1	0	0	0	1	0	0	1	0	0	0	1
38	0	1	0	0	0	1	1	0	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	1	1	0	1	1	1	1	0
39	1	1	1	0	0	0	0	0	0	1	0	0	1	1	1	1	1	0	0	0	1	0	1	0	1	1	0	1	1	0	0	0
41	1	1	0	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	1	1	0	0	0	1	1	0	1	1	0	0	0
44	1	1	1	0	0	0	0	0	0	1	0	0	1	1	1	0	1	0	0	1	1	0	1	0	1	1	0	1	1	0	1	0
45	0	1	0	0	0	1	1	0	1	1	0	0	0	0	1	0	0	0	0	1	1	0	1	0	0	1	0	1	0	0	1	0
46	1	0	0	0	1	0	0	1	1	0	1	0	0	0	1	0	0	0	0	1	1	0	1	0	1	1	0	1	0	0	0	0
48	0	1	0	0	0	1	0	0	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	1	1	0	1	1	0	0	0
50	0	1	0	0	0	1	1	0	1	1	0	0	1	1	1	1	1	0	0	0	1	0	0	0	1	1	0	1	1	0	0	0
51	1	1	0	1	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	1	0	1	0	1	1	0	1	0	1	1	1	1
53	1	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	1	0	1	0	1	1	0	1	0	1	1	1	0
55	1	1	0	0	0	0	1	0	1	1	0	0	0	0	1	0	0	0	0	1	1	0	1	0	1	0	0	1	0	1	0	0
58	1	1	0	0	0	1	0	0	0	1	0	0	1	0	1	1	0	0	0	0	1	0	1	0	1	0	0	1	0	0	0	0
59	1	0	0	0	1	0	0	1	1	1	1	0	1	1	1	0	0	0	0	1	1	0	1	0	1	1	0	1	1	0	1	0
74	1	0	1	0	1	0	0	1	1	1	1	0	1	1	1	1	0	0	1	0	0	0	1	0	1	1	0	1	1	0	0	0
75	1	0	1	0	0	0	0	1	1	1	1	0	1	1	1	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0
81	1	1	0	0	0	0	0	0	1	0	0	0	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	*	*	1
84	1	0	1	0	1	0	1	0	1	1	1	0	1	1	1	0	0	0	0	1	1	0	1	0	0	0	0	1	0	1	1	0
85	1	1	0	0	0	0	0	0	1	0	0	0	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	*	*	0
86	1	1	0	0	0	0	0	0	1	1	0	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0
87	1	1	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	1	0	1	0	1	0	0	1	0	0	0	0
94	1	1	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	1	0	0	1	1	0
95	0	1	0	0	0	1	0	1	1	1	0	0	0	0	1	0	0	0	0	1	1	0	1	0	1	0	0	1	0	1	0	0
97	1	1	1	0	0	0	0	0	0	1	0	0	0	1	1	1	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0
101	1	1	0	0	0	0	0	0	0	1	0	0	1	1	1	1	0	0	0	0	1	0	1	0	1	1	0	1	1	0	1	0
105	1	0	0	0	1	0	0	1	1	0	1	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0
106	1	0	1	0	1	0	0	1	1	1	1	0	1	1	1	1	0	0	1	0	1	0	1	0	1	1	0	1	1	0	1	0
107	1	1	0	0	0	1	1	0	1	1	0	0	0	0	1	0	1	0	0	1	1	0	0	0	1	0	0	1	0	0	0	0
111	1	1	0	0	0	0	0	0	0	1	0	0	1	1	1	1	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0
113	1	1	0	0	0	1	0	1	1	1	0	0	0	0	1	0	0	0	0	1	1	0	1	0	0	0	0	1	0	1	1	0
115	1	1	1	0	0	0	0	0	1	1	0	0	1	1	1	1	0	1	1	0	1	0	1	0	1	0	0	1	0	0	0	0
116	0	1	0	0	0	0	1	0	1	1	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	0	0	1	0	0	0	0
117	1	1	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1	0	0	0	1	0	0	0	1	0	0	1	1	0	0	0
119	1	1	0	0	0	0	0	0	1	0	0	0	1	1	1	0	1	0	0	1	1	0	1	0	1	1	0	1	1	0	0	0

noted in Figure 1). Some of the markers had missing molecular marker data (in the case of those located in Intron 2, this was likely due to a polymorphism in the PCR priming site). In the case of missing molecular marker data, analyses were carried out with the lines with missing markers dropped for that particular marker. It is important to note that missing markers were permuted with the other markers, so the actual lines for which phenotypic measures were dropped varied over permutations.