Linkage disequilibrium mapping of molecular polymorphisms at the *scabrous* locus associated with naturally occurring variation in bristle number in *Drosophila melanogaster*

RICHARD F. LYMAN, CHAOQIANG LAI† AND TRUDY F. C. MACKAY*

Department of Genetics, Box 7614, *North Carolina State University*, *Raleigh*, *NC* 27695, *USA*

Summary

We evaluated the hypothesis that the *Drosophila melanogaster* second chromosome gene *scabrous* (*sca*), a candidate sensory bristle number quantitative trait locus (QTL), contributes to naturally occurring variation in bristle number. Variation in abdominal and sternopleural bristle number was quantified for wild-derived *sca* alleles in seven genetic backgrounds: as homozygous second chromosomes (C2) in an isogenic background, homozygous lines in which approximately 20 cM including the *sca* locus had been introgressed into the isogenic background (*sca* BC), as C2 and *sca* BC heterozygotes and hemizygotes against a *P* element insertional *sca* allele and a *P*-induced *sca* deficiency in the same isogenic background, and as *sca* BC heterozygotes against the wild-type *sca* allele of isogenic strain. Molecular restriction map variation was determined for a 45 kb region including the *sca* locus, and single-stranded conformational polymorphism (SSCP) was examined for the third intron and parts of the third and fourth exons. Associations between each of the 27 molecular polymorphisms and bristle number were evaluated within each genotype and on the first principal component score determined from all seven genotypes, separately for each sex and bristle trait. Permutation tests were used to assess the empirical significance thresholds, accounting for multiple, correlated tests, and correlated markers. Three sites in regulatory regions were associated with female-specific variation in abdominal bristle number, one of which was an SSCP site in the region of the gene associated with regulation of *sca* in embryonic abdominal segments.

1. Introduction

In recent years, the discovery of abundant, polymorphic molecular markers and advances in statistical methods for mapping quantitative trait loci (QTL) by linkage to these markers has facilitated the dissection of genetically complex traits in many species at the level of individual QTL (Lynch & Walsh, 1997). A major challenge for the future is to resolve QTL into genetic loci; to determine the homozygous, heterozygous, epistatic and pleiotropic effects and environmental sensitivities of alleles at these loci; and to determine the causal molecular variants associated with variation in phenotype. An understanding of the genetic architecture of quantitative traits at the level of allelic effects and frequencies at loci causing genetic variation in the trait is essential for accurate risk assessment for complex human diseases, for understanding the evolutionary mechanisms maintaining standing variation in nature as well as adaptation and divergence between species, and for introgression of favourable alleles between agriculturally important strains and species.

The first step towards this goal is to obtain genetic evidence that the QTL maps to the same location as a candidate gene or genes affecting the trait, by fine-scale recombination mapping, and, in species for which experimental crosses are possible, by quantitative complementation of QTL alleles with alleles of the candidate gene (Long *et al.*, 1996; Mackay & Fry, 1996; Lyman & Mackay, 1998). The second step is to use linkage disequilibrium mapping in random mating populations to determine whether molecular variation at the candidate gene is associated with phenotypic variation in the quantitative trait and to map the

* Corresponding author. Tel: +1 (919) 515 5810. Fax: +1 (919) 515 3355. e-mail: Trudy_Mackay@ncsu.edu
† Current address: Cereon Genomics, 270 Albany Street, Cambridge, MA 02139, USA.

molecular variant(s) causally associated with variation in the trait (Lander & Schork, 1994; Risch & Merikangas, 1996). Currently, these methods are restricted to genetically tractable and well-characterized organisms, and to quantitative traits for which the genetic pathway leading to the ultimate phenotype is known at the molecular level. *Drosophila melanogaster* sensory bristle number is one such model system. Bristle numbers have a long history of use as model quantitative traits (reviewed by Mackay, 1996). *Drosophila* bristles are external sense organs of the peripheral nervous system (PNS) and many candidate genes affecting bristle and PNS development (Jan & Jan, 1993; Kania *et al.*, 1995; Salzberg *et al.*, 1997) and/or that have major mutant effects on bristle number (Lindsley & Zimm, 1992; Lyman *et al.*, 1996) have been identified and characterized at the molecular level.

Several lines of evidence implicate the *Drosophila scabrous* (*sca*; 2- 66·7 cM) locus, which encodes a secreted protein involved in lateral inhibition during neurogenesis (Baker *et al.*, 1990; Mlodzik *et al.*, 1990), as a genetic locus corresponding to a bristle number QTL. (1) *sca* alleles that (presumably) segregated at low frequency in nature have contributed to response to selection for bristle number (McBride & Robertson, 1963). (2) Spontaneous *sca* mutations have contributed to response of initially inbred lines to selection for bristle number (Hollingdale, 1971; Frankham, 1980). (3) Spontaneous mutations affecting bristle number fail to complement *sca* mutant alleles (Mackay & Fry, 1996). (4) *P* element insertions at *sca* have quantitative effects on bristle number (Lyman *et al.*, 1996). (5) QTL affecting variation in bristle number map to the gene region containing *sca* (Gurganus *et al.*, 1998). (6) A larger proportion of molecular polymorphisms at *sca* than expected by chance were associated with phenotypic variation in bristle number among second chromosome substitution lines sampled from nature (Lai *et al.*, 1994). In this report we provide further evidence that *sca* is a bristle number QTL by introgressing *sca* alleles from nature into a common inbred background, testing for quantitative failure of *sca* mutations to complement naturally occurring *sca* alleles, and testing for linkage disequilibrium between molecular polymorphisms at *sca* alleles and bristle number phenotypes in multiple genetic backgrounds.

2. Materials and methods

(i) Drosophila *stocks*

A sample of 50 wild *sca* alleles were derived from isofemale lines established from a collection from the Raleigh, NC Farmer's Market in 1988. The effects of *sca* alleles were evaluated in seven genetic back-

grounds: (1) Homozygous whole second chromosome substitution lines (C2, *n* = 46) in the background of the highly inbred *Sam*; *ry*[506] strain (Lyman *et al.*, 1996); these are the lines used by Lai *et al.* (1994). (Four of the original homozygous second chromosome lines had accumulated lethal mutations at the time the data described here were collected.) (2) Whole second chromosome substitution lines as heterozygotes against a hypomorphic *P* element insertion *sca* allele (C2/*sca*[P], *n* = 50), which was derived in the *Sam*; *ry*[506] background (Lyman *et al.*, 1996). (3) Whole second chromosome substitution lines as hemizygotes against a *P*-element-induced *sca* deficiency (C2/*Df*(*sca*), *n* = 50), also co-isogenic with the *Sam*; *ry*[506] background (Lyman *et al.*, 1996). (4) Lines homozygous for the *sca* region, derived by 10 generations of backcrossing females heterozygous for the wild-derived second chromosomes and the *sca*[P] allele to *sca*[P] homozygous males (*sca* BC, *n* = 46). Two independent backcross lines were generated for each wild *sca* allele. After 10 backcross generations, each backcross line is expected to contain the wild *sca* allele plus a linked segment of wild-derived chromosome approximately 10 cM to either side of *sca* (Crow & Kimura, 1970, pp. 94–95), with the remainder of the genetic background homozygous *Sam*; *ry*[506]. (5) *sca* BC lines as heterozygotes against the *sca*[P] allele (*sca* BC/*sca*[P], *n* = 50). (6) *sca* BC lines as hemizygotes against the *Df*(*sca*) allele (*sca* BC/*Df*(*sca*), *n* = 50). (7) *sca* BC lines as heterozygotes against the *Sam sca* allele (*sca* BC/*sca*[Sam], *n* = 50). All stocks were reared in shell vials with 10 ml cornmeal–agar–molasses medium, at 25 °C.

(ii) Bristle number phenoytpes

Abdominal bristle number (the number of hairs on the most posterior sternite; segment six of females and segment five of males) and sternopleural bristle number (the total number of bristles on the right and left sternopleural plates) was recorded on 10 males and 10 females from each of two replicate vials per line (or per replicate backcross line). A total of 21 520 individuals were scored for both bristle traits.

(iii) Quantitative genetic analyses

Variation in male and female bristle number among the homozygous, heterozygous and hemizygous C2 substitution lines was partitioned into sources attributable to line (L), vial replicate (R) within line and error (E) by random effects analysis of variance (ANOVA), according to the model: $y = \mu + L + R(L) + E$. μ is the overall mean and nested effects are in parentheses. For the analyses within sexes of the homozygous, heterozygous and hemizygous backcross lines, an additional nested random term attributable

to backcross replicate (BCR) within line was added to the model: $y = \mu + L + BCR(L) + R(BCR) + E$. Variation in bristle numbers for the full data sets included the fixed cross-classified effect of sex (S) and its interactions. The mixed model ANOVAs were $y = \mu + L + S + L \times S + R(L) + S \times R(L) + E$ for the C2 lines and $y = \mu + L + S + L \times S + BCR(L) + S \times BCR(L) + R(BCR) + S \times R(BCR) + E$ for the BC lines. Quantitative complementation (interaction) of wild-type *sca* alleles to mutant and wild-type *sca* alleles was evaluated for the heterozygous and hemizygous genotypes by including an additional cross-classified, fixed effect of cross (C) in the above analyses of C2 and *sca* region BC lines. The full three-way factorial model for the BC lines was $y = \mu + L + S + C + L \times S + L \times C + S \times C + L \times S \times C + BCR(L) + S \times BCR(L) + C \times BCR(L) + S \times C \times BCR(L) + R(BCR \times C) + S \times R(BCR \times C) + E$. The full model for the C2 lines was similar, but excluded terms involving BCR, and the separate sex analyses for the *sca* introgression lines and C2 dropped all terms involving S. Tests of significance of *F* ratios and estimates of variance components of the random effects were obtained using SAS procedures GLM and VARCOMP, respectively (SAS Institute, 1988).

(iv) *Molecular variation at* sca

Molecular restriction map variation among the homozygous viable alleles in the 45 kb region including the *sca* locus was assessed using three six-base-cutter restriction enzymes (Lai *et al.*, 1994). Finer-scale single-stranded conformational polymorphism (SSCP) variation was determined for three SSCP 'loci' spanning a 776 bp region including the 3′ end of exon 3 (SSCP 1683), the entire third intron (SSCP 1447) and the 5′ end of exon 4 (SSCP 1826). There were 24 polymorphic (i.e. the frequency of the least common allele was greater than 5%) restriction site and insertion/deletion variants. The 1683, 1447 and 1826 SSCP loci had 8, 9 and 14 mobility alleles. The number of alleles at each SSCP locus was reduced to 6, 6 and 8 for SSCP 1683, 1447 and 1826, respectively, after pooling all rare alleles for each locus into a single category (Lai *et al.*, 1994). Although alleles at some adjacent polymorphic sites were in strong linkage disequilibrium, the null hypothesis of linkage equilibrium was not rejected for most pairs of sites in this sample, indicating considerable historical recombination in the *sca* gene region.

(v) *Molecular marker/bristle number phenotype associations*

Associations between each of the 27 polymorphic marker loci and bristle number phenotypes were

evaluated, separately for each bristle character, sex and genetic background, by testing for significance of the marker allele (M) term in analyses of variance of bristle number performed on line means: $y = \mu + M + L(M) + E$. These tests were also performed on the difference in line means between wild *sca* alleles as heterozygotes or hemizygotes against standard mutant or wild-type *sca* alleles. There were four such complementation effects: C2/*Df*(*sca*)–C2/*sca*^{P}, *sca* BC/*sca*^{P}–*sca* BC/*sca*^{Sam}, *sca* BC/*Df*(*sca*)–*sca* BC/*sca*^{Sam} and *sca* BC/*Df*(*sca*)–*sca* BC/*sca*^{P}. Thus, for each of the 11 *sca* genotypes, we performed multiple tests using the same phenotypic data, and the molecular markers were correlated to varying extents. To determine an appropriate *P* value for the experiment-wise Type I error rate (*a*), we permuted trait phenotypes among the marker haplotypes 1000 times. The lowest *P* value for the effect of marker was recorded for each permuted data set, and the distribution of *P* values under the null hypothesis of no marker–phenotype association was obtained. The 50th lowest *P* value corresponds to $\alpha = 0.05$ under the null hypothesis. (The permutation tests were done with *P* values, rather than *F* statistics, because the number of alleles per marker genotype was not constant – two alleles for the restriction site and insertion/deletion markers and 6–8 alleles for the SSCP loci.) The seven homozygous, heterozygous and hemizygous genotypes were also correlated. To reduce the dimensionality of the data set, the seven line means of each wild *sca* allele for each sex and trait were converted to principal component scores, after transformation of each set of line means to zero mean and unit variance. The associations between the 27 marker genotypes and the first principal components were evaluated as described above for the untransformed bristle number phenotypes. The significance of associations between marker genotypes and the principal component scores were also determined by permutation tests.

3. Results

(i) *Quantitative genetic analyses*

Distributions of line means for the two bristle traits in each of the seven genetic backgrounds are given in Fig. 1. Variation among lines was often significantly skewed and leptokurtic, particularly for abdominal bristle number. Abdominal bristle number did not differ significantly from a Normal distribution for *sca* BC/*sca*^{P} and *sca* BC/*Df*(*sca*) males, and for both sexes of the *sca* BC/*sca*^{Sam} genotype. Line means for sternopleural bristle number were not significantly different from Normal in all of the BC genotypes, but were significantly skewed and/or leptokurtic in the whole chromosome genotypes.
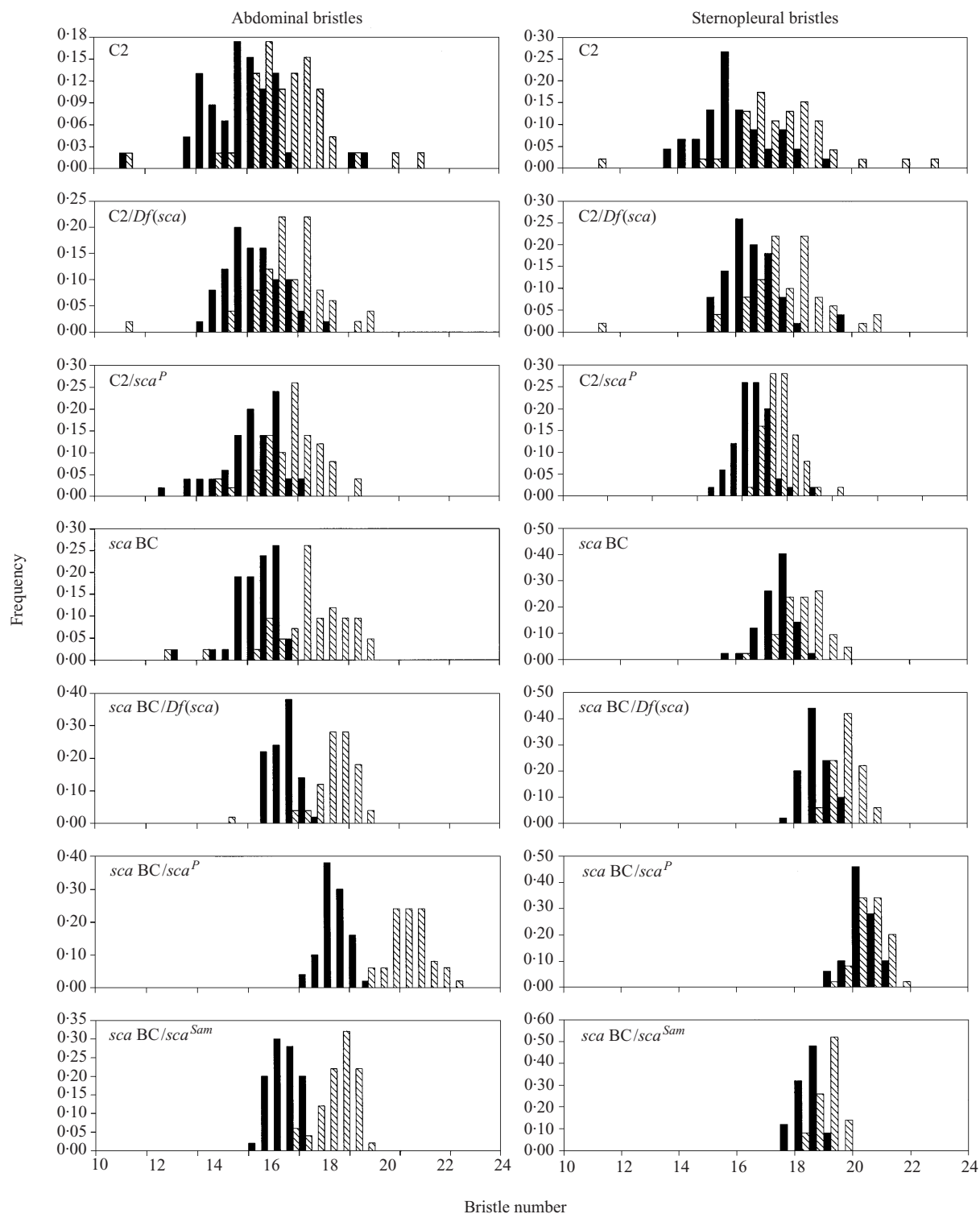
Fig. 1. Distributions of line means of bristle number in seven genetic backgrounds, for *scabrous* alleles sampled from the Raleigh population. Male and female data are indicated by filled and cross-hatched bars, respectively.

Among-line variance components, estimated from ANOVA of bristle number for sexes pooled and considered separately, are given for each of the seven genotypes in Table 1. Most of the estimates of among-line variance were highly significant; the exceptions were male abdominal bristle number in *sca* BC/*Df*(*sca*)

hemizygotes and abdominal bristle number of both sexes in *sca* BC/*sca*^Sam^ heterozygotes. Both bristle characters exhibit highly significant sex dimorphism, whereby females have on average one or two more bristles than males in all genotypes. This is apparent from the distributions of line means in Figure 1. More

Table 1. *Among-line and sex × line variance components and significance of* F *ratio tests of significance from ANOVA of bristle number*

Genotype	Abdominal bristles				Sternopleural bristles			
	Females	Males	Sexes Combined	Sex × Line	Females	Males	Sexes Combined	Sex × Line
C2	2·696***	2·403***	2·545***	−0·005 ns	1·295***	1·326***	1·282***	0·028 ns
C2/sca^P	1·153***	1·035***	0·831***	0·264*	0·466***	0·541***	0·479***	0·025 ns
C2/Df(sca)	1·892***	0·817***	1·205***	0·150*	0·761***	0·796***	0·730***	0·049 ns
BC sca	2·038***	0·512***	1·006***	0·271***	0·347***	0·186***	0·234***	0·032*
BC sca/sca^P	0·409***	0·148**	0·181**	0·097**	0·102*	0·120**	0·097*	0·014 ns
BC sca/Df(sca)	0·586***	0·078 ns	0·184*	0·148**	0·137***	0·100**	0·116***	0·002 ns
BC sca/sca^{Sam}	0·025 ns	0·067 ns	0·048 ns	−0·003 ns	0·067*	0·057*	0·082**	−0·020 ns

***$P < 0.001$; **$0.001 < P < 0.01$; *$0.01 < P < 0.05$; ns $P > 0.05$.

formally, the effect of sex in the full mixed model ANOVA for each genotype was significant at $P < 0.0001$ (data not shown). However, abdominal and sternopleural bristle number differ markedly in the extent to which there is genetic variation in sex dimorphism, as indicated by a significant sex × line (S × L) interaction term in the ANOVAs including sex as a cross-classified effect. The only genotype exhibiting significant S × L interaction variance for sternopleural bristle number was the BC homozygotes, and the significance was marginal ($P = 0.04$). In contrast, sex × genotype interaction variance for abdominal bristle number was highly significant for all genotypes except the C2 homozygotes and sca BC/sca^{Sam} heterozygotes. Thus loci on the second chromosome and/or the sca introgression region that interact with mutant sca alleles, and loci in the homozygous sca introgression region, harbour naturally occurring variation for sex dimorphism of abdominal bristle number. The observation that sex × genotype interaction variance for abdominal bristle number is not significant for the whole C2 homozygotes but is highly significant for the sca BC homozygous lines introgression genotypes indicates that C2 loci affecting this trait have variable pleiotropic effects on sex dimorphism The nature of the sex × genotype interaction in all cases is that expressed genetic variance is greater in females than males. This effect is particularly apparent for the BC region introgression lines, where female among-line variance is greater than male among-line variance of abdominal bristle number by a factor of 4 for the homozygous lines, 2·8 for the heterozygotes against the *P* element insertional sca allele, and 7·5 as hemizygotes against the sca deficiency allele.

If there are a large number of loci affecting natural variation in bristle number, and the effects of these loci are equal and individually small (the infinitesimal model), one expects the contribution of an introgressed gene region to overall variation in bristle number to be proportional to its genetic length. The *Drosophila* second chromosome is 110 map units long (Lindsley & Zimm, 1992). The average length of the sca region introgressions is expected to be 20 cM; therefore under the infinitesimal expectation the sca region introgressions would account for roughly 18 % of the total chromosome 2 variation in bristle number. The observed proportions of total variation accounted for by the sca region introgressions were, respectively, 75·6 % and 21·3 % for female and male abdominal bristle number, and 26·8 % and 14·0 % for female and male sternopleural bristle number. The standard errors of the among-line variance components are not small, and the true lengths of the introgressions are not known. Nevertheless, it appears that a gene or genes in the sca introgressions explains a larger than expected fraction of the variation in female abdominal bristle number, but not male abdominal bristle number nor sternopleural bristle number in either sex.

Further genetic evidence supporting the hypothesis that naturally occurring sca alleles have female-specific effects on abdominal bristle number comes from the results of quantitative complementation tests. In these analyses, QTL alleles are crossed to two different candidate locus alleles that have been obtained in, or introgressed into, the same homozygous genetic background. If there is significant variation in the difference in heterozygous effects of the QTL alleles in the background of the two candidate locus alleles (i.e. a significant cross × line effect in ANOVA), there is quantitative failure of the QTL alleles to complement mutations at the candidate locus (Long *et al.*, 1996; Mackay & Fry, 1996; Lyman & Mackay, 1998). As in any complementation test, failure to complement can arise from interactions between QTL alleles at the candidate locus and candidate locus alleles, or between QTL alleles at a different locus and candidate locus alleles (epistasis). The results of quantitative comple-

Table 2. *Cross × line and cross × sex × line variance components and significance of* F *ratio tests of significance from ANOVA of bristle number*

	Abdominal bristles				Sternopleural bristles			
Genotypes	Females	Males	Sexes Combined	Sex × Line	Females	Males	Sexes Combined	Sex × Line
C2/sca^P, C2/Df(sca)	0·593***	0·158*	0·337***	0·039 ns	0·295***	0·355***	0·329***	−0·004 ns
BC sca/sca^P, BC sca/sca^{sam}	−0·022 ns	0·024 ns	0·021 ns	−0·020 ns	0·023 ns	0·011 ns	0·018 ns	−0·002 ns
BC sca/Df(sca), BC sca/sca^{sam}	0·051 ns	−0·024 ns	−0·033 ns	0·047 ns	0·013 ns	0·013 ns	0·014 ns	−0·002 ns
BC sca/sca^P, BC sca/Df(sca)	0·219**	0·030 ns	0·025 ns	0·099**	−0·010 ns	0·042 ns	0·007 ns	0·009 ns

***$P < 0.001$; **$0.001 < P < 0.01$; *$0.01 < P < 0.05$; ns $P > 0.05$.

mentation tests to *sca* alleles are given in Table 2. There was significant failure of QTL alleles on C2 to complement abdominal and sternopleural bristle number effects of mutant *sca* alleles. However, significant failure of QTL alleles in the *sca* BC region to complement mutant *sca* alleles was observed only for female abdominal bristle number. The most parsimonious interpretation of these results is that there is naturally occurring allelic variation at C2 QTL outside the *sca* introgression region affecting abdominal and sternopleural bristle number in both sexes that interact with *sca*; and that there is additionally naturally occurring allelic variation at QTL with female-specific effects on abdominal bristle number inside the *sca* introgression region, perhaps at *sca* itself, that interact with *sca*.

(ii) *Molecular and phenotypic associations at* sca

We tested whether there were differences in bristle number between alleles at polymorphic markers in the *sca* gene region. These tests were done for each of the 27 polymorphic sites for each of the seven homozygous and heterozygous genotypes, the first principal component derived from these estimates, and the four measures of complementation effects, separately for the two bristle traits and both sexes. In total, 1296 association tests were performed. To address the serious problem that multiple, non-independent tests pose for declaring an appropriate experiment-wise significance threshold, empirical significance thresholds were determined by permutation within each of the 48 'genotypes' (2 bristle traits × 2 sexes × 12 estimates of bristle effect).

Given the complexity of the analysis, the results of these tests were remarkably straightforward. None of the polymorphic sites at *sca* were significantly associated with variation in sternopleural bristle number in either sex, or with abdominal bristle
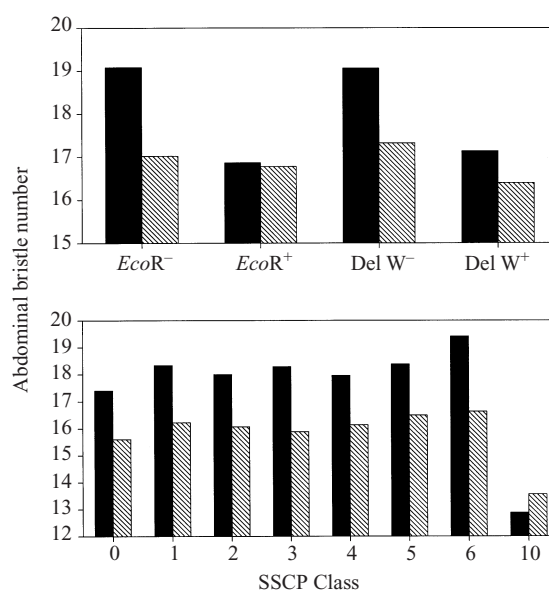

Fig. 2. Mean abdominal bristle number of marker alleles associated with significant variation of this trait. Male and female data are indicated by filled and cross-hatched bars, respectively. Upper panel: *Eco*RI(20·2) and DelW variants, in *sca* BC/Df(sca) hemizygotes. Lower panel: SSCP 1836 alleles, in *sca* BC homozygotes.

number in males. However, three sites – *Eco*RI(20.2) and DelW(27.6) in the 3′ flanking region (Lai *et al.*, 1994) and SSCP 1836 – were significantly associated with abdominal bristle number in females (Fig. 2). Marker allele frequencies at these sites did not depart significantly from linkage equilibrium in this sample (Lai *et al.*, 1994).

The *Eco*RI(20.2) site was significant in the *sca* BC/Df(sca) genetic background ($P = 0.0002$), and for two of the BC complementation effect measures: *sca* BC/Df(sca)–sca BC/sca^{Sam} ($P = 0.0003$) and *sca* BC/Df(sca)–sca BC/sca^P ($P = 0.0003$). The presence of this site in the *sca* BC/Df(sca) genetic background was associated with a reduction of 2·22 abdominal

bristles in females. DelW was also significant in the *sca* BC/*Df*(*sca*) genetic background (*P* = 0·0002), and for the *sca* BC/*Df*(*sca*)–*sca* BC/*sca*^Sam (*P* = 0·0001) complementation effect. The presence of DelW in the *sca* BC/*Df*(*sca*) genetic background was associated with a reduction of 1·93 abdominal bristles in females. The SSCP 1836 'locus' was significantly associated with female abdominal bristle number in *sca* BC homozygotes (*P* = 0·0001), *sca* BC/*sca*^Sam heterozygotes (*P* = 0·0009), and the first principal component score (*P* = 0·0004). It should be noted that the nominal *P* value for the first principal component analysis of male abdominal bristle number at SSCP 1836 was 0·0005, which was not quite formally significant on the permutation test, but suggestive. Variation at SSCP 1836 was partitioned into 8 'allelic' classes. The variation in abdominal bristle number was associated with 'allele' number 10 (Lai *et al.*, 1994), which had on average 5·37 abdominal bristles fewer than the other seven SSCP 1836 alleles in female *sca* BC homozygotes. The marker alleles associated with variation in bristle number were not at intermediate frequency in this sample. The *Eco*RI(20.2) restriction site and DelW each had a frequency of 6 % (3/47) in the Raleigh C2 homozygous lines, and there were only two SSCP 1836 '10' alleles (Lai *et al.*, 1994).

4. Discussion

Previously, we evaluated the hypothesis that *sca* contributes to naturally occurring variation in bristle number among the Raleigh C2 whole chromosome substitution lines by testing for significant associations between molecular variation at *sca* and phenotypic variation in bristle number (Lai *et al.*, 1994). In the previous analysis, permutation was used to deal with the issue of multiple tests by determining whether more nominally significant associations were observed than expected by chance. As this was the case, *sca* was imputed as a bristle number QTL, but the significance of each of the nominally significant sites was not tested by permutation analysis. Here, we have extended the analysis of phenotypic and molecular variation at *sca* to individual sites and to multiple genetic backgrounds. In this analysis, none of the sites in the C2 substitution line data set met the more rigorous criterion for individual statistical significance. However, the site with the lowest *P* value in the analysis of Lai *et al.* (1994) was SSCP 1447, which is in strong linkage disequilibrium with SSCP 1683 and SSCP 1836. In the analysis reported here, SSCP 1836 was significantly associated with variation in female abdominal bristle number in two genetic backgrounds and in the analysis of principal component scores. Thus, the results of the two analyses do not differ qualitatively in that *sca* is associated with variation in bristle number, and the same 776 base pair region of

the gene gave the highest statistical signal in each study. The quantitative differences between the two studies are partially attributable to the increased statistical power in the BC genotype – loosely linked C2 bristle number QTL do not contribute to the estimate of line means, and twice as many individuals were scored per line.

The most striking feature of these results is that molecular variation at *sca* is associated with female-specific effects on abdominal bristle number. Similar results were obtained in the analyses of associations between molecular variation and phenotypic variation in bristle number at the candidate genes *Delta* (*Dl*, Long *et al.*, 1998) and the *achaete-scute* complex (*ASC*, Long *et al.*, 1999). In both these studies, one polymorphic site in the candidate gene region was associated with variation in sternopleural bristle number in both sexes, and one site was associated with abdominal bristle number in females only. Other independent lines of evidence indicate that sex-specific effects of bristle number QTL are common. Significant sex × genotype interactions are observed for spontaneous (Mackay *et al.*, 1995) and *P*-element induced mutations (Lyman *et al.*, 1996) affecting bristle number, for individual bristle number QTL effects (Long *et al.*, 1995; Gurganus *et al.*, 1999; Nuzhdin *et al.*, 1999) and even for QTL × environment interaction effects (Gurganus *et al.*, 1998). To the extent that *Drosophila* bristle number is a model quantitative trait, these results suggest that sex-specific effects may be a general feature of the genetic architecture of quantitative traits, in *Drosophila* and other organisms. Such effects should be explicitly incorporated into models of the evolution of quantitative traits.

Two of the sites significantly associated with abdominal bristle number in females in this study, DelW and *Eco*RI(20.2), had the unusual property that the bristle number effects were apparent as BC hemizygotes against a *sca* deficiency, but not as BC homozygotes. It is impossible to deduce from these data whether the gene dosage effect is caused by genetic interactions between *sca* alleles from nature and the deficiency, or whether the *sca* deficiency interacts with naturally occurring variation at a linked locus.

The frequency of sites affecting variation in phenotype (quantitative trait nucleotides, or QTN) is an important parameter in models of the maintenance of quantitative genetic variation in nature. Rare alleles of large effect are consistent with mutation–selection balance models (reviewed by Barton & Turelli, 1989; Lynch & Walsh, 1997), while intermediate-frequency QTN are consistent with either selective neutrality or maintenance of variation by a balancing selection mechanism. We do not know the frequency of the QTN, but that of the markers associated with the QTN. Further, the sample sizes of experiments to date

have been on the order of 50 chromosomes, which are too small for precise estimates of gene frequency. With these caveats, intermediate-frequency molecular polymorphisms were associated with variation in bristle number at *Dl* and *ASC* (Long *et al*, 1998, 1999), and individually rare large insertions were associated as a class with variation in bristle number at *ASC* (Mackay & Langley, 1990; Long *et al*., 1999). Such heterogeneity across loci and within loci in the frequency of molecular markers associated with effects on bristle number suggests that there may not be one predominant mechanism responsible for equilibrium levels of quantitative genetic variation in nature, but that the question needs to be addressed on a locus-by-locus and site-by-site basis. In this study we observed that large bristle number effects were associated with low-frequency molecular variants at *sca*. However, it would be premature to infer that some fraction of the standing variation at *sca* may therefore be attributable to rare QTN of large effect, maintained by mutation–selection balance. This is because the upper confidence limit for the frequency of the bi-allelic polymorphisms at *sca* that are associated with variation in bristle number is 13%, which is not rare. Further, the SSCP loci have multiple alleles for which the actual nucleotide differences are not known, and pooling unique alleles per locus into a single category could bias the analysis. It will be necessary to obtain DNA sequence polymorphism information for the SSCP loci and larger sample sizes before making inferences about the frequency of *sca* bristle number QTN.

Two of the three molecular variants associated with variation in bristle number, DelW and *Eco*RI(20.2), were in the 3′ regulatory region of the gene. The third, SSCP 1836, includes part of the third intron and fourth exon. This latter region corresponds to binding sites for the homeotic proteins, *Ultrabithorax*, *abdominal-A* and *Abdominal-B*, which are associated with down-regulation of *sca* in abdominal segments of wild-type embryos (Graba *et al*., 1992). For this 'locus', one can imagine pursuing a mechanistic explanation linking changes in tissue-specific gene expression with quantitative variation in phenotype through further molecular and developmental analyses of alleles from nature.

The inferences drawn from these results are limited by the power of the experimental design. Through genetic manipulation (chromosome substitution, introgression) we were able to reduce the genetic variation in bristle number to that contributed by a single chromosome or chromosome region. Construction of homozygous stocks further increased the power to detect allelic differences, and replicated measurements of phenotypes enabled accurate estimates of genotypic effects. Nevertheless, only large effects on bristle number could be detected as significant in a sample of 50 alleles. Thus, the

conclusion that *sca* alleles with large effects on bristle number occur in nature must be tempered by our inability to detect alleles with moderate, let alone small, effects. Likewise, the observation that molecular variants associated with bristle number phenotype occurred in regulatory regions does not mean that polymorphisms in coding regions are not important: the molecular restriction map survey was capable of detecting only a small fraction of the existing variation, and very few restriction site polymorphisms were found in coding regions. The significance of an association between a molecular marker and a QTN will depend on the effect of the QTN and its physical distance from the marker, as linkage disequilibrium falls off rapidly with physical distance. At *sca*, for example, there is no detectable linkage disequilibrium between markers greater than two kilobases apart (Lai *et al*., 1994). It follows that efforts to localize QTN using this method will require a sufficiently high density of molecular markers that allele frequencies at adjacent markers are in linkage disequilibrium. For this reason, it is not surprising that the SSCP 1836 site, which was in a region with the highest marker density, was in fact the site associated with bristle number across all genotypes. Finally, one cannot neglect the important caveat that associations of molecular polymorphisms in candidate genes with phenotypic variation can occur for historical reasons, such as population admixture. Independent replication in another population is always necessary to infer a causal relationship between a polymorphic site and a putative QTN. Therefore, future work to map and estimate effects of QTN affecting bristle number must utilize much larger samples of alleles from at least two populations, and a much higher density of molecular markers in the candidate gene region.

Recently, linkage disequilibrium mapping has been suggested as a powerful method to map QTN affecting complex traits in humans (Lander & Schork, 1994; Risch & Merikangas, 1996). Some guidelines for the design of these experiments can be taken from the work with *Drosophila* bristles. Clearly, very large samples of individuals will be necessary. Without the ability to manipulate and replicate genotypes, the effects of any one candidate gene will be averaged over all other genomic loci affecting variation in the phenotype, and there is no control over the environment. The possible existence of sex × genotype and genotype × genotype interaction effects translates operationally to an even greater sample size, to enable evaluation of associations within males and females separately, for example. Finally, it is important to search for associations between polymorphic sites in introns and other regulatory regions as well as coding regions, and not to confine these studies to exons, as has been proposed; and to have a high density of molecular markers across the gene region of interest.

References

Baker, N. E., Mlodzik, M. & Rubin, G. M. (1990). Spacing differentiation in the developing *Drosophila* eye: a fibrinogen-related lateral inhibitor encoded by *scabrous*. *Science* **250**, 1370–1377.

Barton, N. & Turelli, M. (1989). Evolutionary quantitative genetics: How little do we know? *Annual Reviews of Genetics* **23**, 337–370.

Crow, J. F. & Kimura, M. (1970). *An Introduction to Population Genetics Theory*. New York: Harper and Row.

Frankham, R. (1980). Origin of genetic variation in selection lines. In *Selection Experiments in Laboratory and Domestic Animals*, (ed. A. Robertson), pp. 56–68. Slough, UK: Commonwealth Agricultural Bureau.

Graba, Y., Aragnol, D., Laurenti, P., Garzino, V., Charmot, D., Berenger, H. & Pradel, J. (1992). Homeotic control in *Drosophila*: the *scabrous* gene is an *in vivo* target of *Ultrabithorax* proteins. *European Molecular Biology Organization Journal* **11**, 3375–3384.

Gurganus, M. C., Fry, J. D., Nuzhdin, S. V., Pasyukova, E. G., Lyman, R. F. & Mackay, T. F. C. (1998). Genotype–environment interaction for quantitative trait loci affecting sensory bristle number in *Drosophila melanogaster*. *Genetics* **149**, 1883–1898.

Gurganus, M. C., Nuzhdin, S. V. & Mackay, T. F. C. (1999). High resolution mapping of quantitative trait loci for sternopleural bristle number in *Drosophila melanogaster*. *Genetics* (in press).

Hollingdale, B. (1971). Analyses of some genes from abdominal bristle selection lines in *Drosophila* nucleotide data. *Genetics* **116**, 153–159.

Jan, Y. N. & Jan, L. Y. (1993). The peripheral nervous system. In *The Development of* Drosophila melanogaster, vol. 2 (ed. M. Bate & A. Martinez Arias), pp. 1207–1244. Plainview, NY: Cold Spring Harbor Laboratory Press.

Kania, A., Salzberg, A., Bhat, M., D'Evelyn, D., He, Y., *et al.* (1995). *P*-element mutations affecting embryonic peripheral nervous system development in *Drosophila melanogaster*. *Genetics* **139**, 1663–1678.

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

Lander, E. S. & Schork, N. J. (1994). Genetic dissection of complex traits. *Science* **265**, 2037–2048.

Lindsley, D. L. & Zimm, G. G. (1992). *The Genome of* Drosophila melanogaster. San Diego: Academic Press.

Long, A. D., Mullaney, S. L., Reid, L. A., Fry, J. D., Langley, C. H. & Mackay, T. F. C. (1995). High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. *Genetics* **139**, 1273–1291.

Long, A. D., Mullaney, S. L., Mackay, T. F. C. & Langley, C. H. (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**, 1497–1518.

Long, A. D., Lyman, R. F., Langley, C. H. & Mackay, T. F. C. (1998). Two sites in the *Delta* gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genetics* **149**, 999–1017.

Long, A. D., Lyman, R. F., Morgan, A. H., Langley, C. H. & Mackay, T. F. C. (1999). Both naturally occurring insertions of transposable elements and intermediate frequency polymorphisms at the *achaete-scute* complex are associated with variation in bristle number in *Drosophila melanogaster*. *Genetics*, submitted.

Lyman, R. F. & Mackay, T. F. C. (1998). Candidate quantitative trait loci and naturally occurring variation for bristle number in *Drosophila melanogaster*: the *Delta-Hairless* gene region. *Genetics* **149**, 983–998.

Lyman, R. F., Lawrence, F., Nuzhdin, S. V. & Mackay, T. F. C. (1996). Effects of single *P* element insertions on bristle number and viability in *Drosophila melanogaster*. *Genetics* **143**, 277–292.

Lynch, M. & Walsh, B. (1997). *Genetics and Analysis of Quantitative Traits*. Sunderland MA: Sinauer Associates.

Mackay, T. F. C. (1996). The nature of quantitative genetic variation revisited: lessons from *Drosophila* bristles. *BioEssays* **18**, 113–121.

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

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

Mackay, T. F. C., Lyman, R. F. & Hill, W. G. (1995). Polygenic mutation in *Drosophila melanogaster*: non-linear divergence among unselected strains. *Genetics* **139**, 849–859.

McBride, G. & Robertson, A. (1963). Selection using assortative mating in *Drosophila melanogaster*. *Genetical Research* **4**, 356–369.

Mlodzik, M., Baker, N. E. & Rubin, G. M. (1990). Isolation and expression of *scabrous*, a gene regulating neurogenesis in *Drosophila*. *Genes and Development* **4**, 1848–1861.

Nuzhdin, S. V., Dilda, C. L & Mackay, T. F. C. (1999). The genetic architecture of selection response: inferences from fine-scale mapping of bristle number quantitative trait loci in *Drosophila melanogaster*. *Genetics*, in press.

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

Salzberg, A., Prokopenko, S. N., He, Y., Tsai, P., *et al.* (1997). *P*-element insertion alleles of essential genes on the third chromosome of *Drosophila melanogaster*: mutations affecting embryonic PNS development. *Genetics* **147**, 1723–1741.

SAS Institute (1988). *SAS/STAT User's Guide*, release 6.03 edition. Cary, NC: SAS Institute.