

How Repeatable Are Associations Between Polymorphisms in *achaete-scute* and Bristle Number Variation in *Drosophila*?

Jonathan D. Gruber,^{*,1} Anne Genissel,[†] Stuart J. Macdonald[‡] and Anthony D. Long^{*}

^{*}Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697, [†]Section of Evolution and Ecology, University of California, Davis, California 95616 and [‡]Department of Ecology and Evolutionary Biology, Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66045

Manuscript received October 23, 2006
Accepted for publication January 20, 2007

ABSTRACT

Currently, the relevance of common genetic variants—particularly those significantly associated with phenotypic variation in laboratory studies—to standing phenotypic variation in the wild is poorly understood. To address this, we quantified the relationship between *achaete-scute* complex (*ASC*) polymorphisms and *Drosophila* bristle number phenotypes in several new population samples. MC22 is a biallelic, nonrepetitive-length polymorphism 97 bp downstream of the *scute* transcript. It has been previously shown to be associated with sternopleural bristle number variation in both sexes in a set of isogenic lines. We replicated this association in a large cohort of wild-caught *Drosophila melanogaster*. We also detected a significant association at MC22 in an outbred population maintained under laboratory conditions for ~25 years, but the phenotypic effects in this sample were opposite from the direction estimated in the initial study. Finally, no significant associations were detected in a second large wild-caught cohort or in a set of 134 nearly isogenic lines. Our ability to repeat the initial association in wild samples suggests that it was not spurious. Nevertheless, inconsistent results from the other three panels suggest that the relationship between polymorphic genetic markers and loci contributing to continuous variation is not a simple one.

MENDELIAN and biometric inheritances can be reconciled by the combined action of multiple genes giving rise to continuously distributed phenotypes in populations. However, no quantitative trait exists for which the major features of its genetic architecture have been amply described. Of the few options for expanding this knowledge, association studies among natural genotypes are the most straightforward means of defining the effects of segregating variants.

Unfortunately, associations between SNPs and variation in a complex trait have often proved difficult to detect and have often been even more difficult to independently repeat. Although “failure to replicate” was all too common in earlier studies of multifactorial diseases and still is a feature of the current literature (*e.g.*, DUAN *et al.* 2006; JEONG *et al.* 2006; STOGMANN *et al.* 2006; THIRUMALAI *et al.* 2006), the field has matured to the point where major discoveries are almost always documented in multiple data sets (*e.g.*, HAINES *et al.* 2005; KLEIN *et al.* 2005; AMUNDADOTTIR *et al.* 2006; GRAHAM *et al.* 2006; GRANT *et al.* 2006). Although failure to replicate can occasionally be reconciled through an examination of the confidence intervals on estimated phenotypic effects (ALTSCHULER *et al.* 2000), discordance

between association studies is also potentially attributable to a myriad of other factors; in humans, underpowered/overanalyzed experiments, publication bias, and admixture are the most highly cited culprits (CARDON and BELL 2001). Conversely, it has been noted that “replication” sometimes involves a different (but related) phenotype and/or a nearby (but not necessarily in LD) polymorphism (ROSAND *et al.* 2006). Perhaps such situations do not cast doubt on the influence of the implicated gene, but they do leave critical questions open about the identity and number of the causal site(s). In fact, because of the greater LD and impossibility of experimental manipulation in human genetics, causal sites may be more readily documented and verified in model organisms.

Thus, we sought to document evidence for “the” causal site(s) at a strong candidate locus in *Drosophila melanogaster*, using the model trait of bristle number. Theoretically, one could compare equivalent high-powered association studies in wild populations of any organism, but *Drosophila* have the additional advantage that we can also create panels of isogenic laboratory lines. Then it is possible to determine how often conclusions drawn from laboratory experiments are applicable to wild populations, where the natural complement of evolutionary forces has shaped the genome. Conversely, we can determine if allelic effects initially observed in natural populations are comparable in the laboratory,

¹Corresponding author: Ecology and Evolutionary Biology, 321 Steinhaus Hall, University of California, Irvine, CA 92697.
E-mail: gruberjd@uci.edu

potentially enabling future experiments aimed at dissecting the functional underpinning of continuous variation.

We focus on the classic model trait of bristle number in *D. melanogaster*. Many isogenic strain-based studies have been carried out previously, particularly on component genes of the *Notch* signaling pathway [*scabrous* (LAI *et al.* 1994); *Delta-Hairless* gene region (LYMAN and MACKAY 1998; LONG *et al.* 1998); *achaete-scute* complex (LONG *et al.* 2000); *hairy* (ROBIN *et al.* 2002)] as well as on genes from other pathways [*Dopa decarboxylase* (MACKAY and LYMAN 2005); *Catsup* (CARBONE *et al.* 2006)]. The associations between naturally occurring alleles at *achaete-scute* (*ASC*) (chromosome X, 1A6–1B3) and bristle phenotype were first noted by MACKAY and LANGLEY (1990), who found that as a class individually rare large insertion/deletion polymorphisms (indels) at *ASC* were significantly associated with reduction in both abdominal bristle number (ABN) and sternopleural bristle number (SBN). LONG *et al.* (2000) replicated this result for ABN and further observed that a polymorphic *transpac* transposable element insertion, located ~9 kb 5' of *asense*, was associated with ABN. They also described an association between a small indel near the *scute* transcript (MC22) and SBN. Finally, LONG *et al.* (2000) observed that a second small indel near *scute* (MC21) was associated with a genotype-by-sex interaction for ABN.

We genotyped these three polymorphisms in three collections of outbred individuals (~2000 individuals each), two from wineries in Northern California (GENISSEL *et al.* 2004; MACDONALD and LONG 2004; MACDONALD *et al.* 2005a) plus an additional laboratory/outbred population (ROSE 1984), to explore the replicability of phenotype-genotype associations. Very few association studies have been replicated at this level. We also performed a laboratory association study using a set of 134 nearly isogenic lines (NILs) derived from natural collections (YANG and NUZHIDIN 2003; PALSSON *et al.* 2004; S. NUZHIDIN, personal communication). We observed complex patterns of the significance and/or direction of allelic effects on bristle number, some consistent with prior studies, others finding no effect, and others quite novel. These results suggest that associations between genotypes and phenotypes may be sensitive to a range of factors that are not generally carefully controlled when carrying out association studies.

MATERIALS AND METHODS

Outbred populations: Each wild population consists of ~1000 *D. melanogaster* flies of each sex. Adults were captured from Kaz Vineyard and Winery (Kenwood, CA) October 15, 1996 ("SV96") (GENISSEL *et al.* 2004) and from Clos Pegase (Calistoga, CA) November 4, 2001 ("NV01") (MACDONALD and LONG 2004; MACDONALD *et al.* 2005a). These two wineries are ~10 miles from one another. A third collection, outbred

but adapted to the laboratory, was derived from the Ives population (provided by M. R. Rose) on May 26, 2004 ("IV04"). The Ives population was initiated from 200 males and 200 females in 1975 (ROSE and CHARLESWORTH 1981) and has subsequently been maintained at population sizes in the thousands (ROSE 1984). Parents of the experimental IV04 flies were allowed to lay eggs in a set of 16 half-pint bottles containing standard banana-molasses media. Parental flies were cleared after 3 days to ensure uncrowded conditions for the progeny. Culturing of the IV04 population was carried out at room temperature under 24-hr light conditions.

Nearly isogenic lines: We used NILs from the "Wolfskill Orchard" W1 (YANG and NUZHIDIN 2003; PALSSON *et al.* 2004) and W3 collections (S. NUZHIDIN, personal communication) and the "West End" WE collection (PALSSON *et al.* 2004). All were propagated through brother-sister mating for at least 15 generations. One or two DNA samples from individual females were used for line genotyping, and we observed no heterozygosity at the loci we genotyped. For bristle counting, each line was raised under 24-hr light at room temperature in four replicate vials; five flies of each sex were phenotyped per vial.

Phenotypes: ABN is defined as the bristles on the fifth (males) or sixth (females) abdominal sternite. SBN is defined as the total number of microchaete and macrochaete on the right and left sternopleural plates. Bristle count statistics from each population are given in Table 1. After phenotypes were recorded, flies were frozen until DNA extraction.

DNA extraction: Individuals from wild populations SV96 and NV01 were extracted in 1.5-ml microcentrifuge tubes using the Puregene DNA purification kit (Gentra Systems, Minneapolis), according to the manufacturer's instructions for *Drosophila*. Laboratory population IV04 and NIL flies were extracted with the Puregene kit, with the following modifications for high-throughput extraction. All steps were performed in 96-well plates, using multichannel pipettors. Homogenization with stainless-steel $\frac{5}{32}$ -in. grinding balls (BT&C/OPS Diagnostics, Bridgewater, NJ) in cell lysis solution was performed in the Geno/Grinder 2000 (BT&C/OPS Diagnostics) for 45 sec at 1500 rpm. Lysis incubation time at 65° was increased to 20 min. Stock RNase A solution (4 mg/ml) was diluted to 67 ng/ml in ddH₂O, and 30 μ l dilute enzyme were added to each sample. Samples were subjected to RNase A digestion (37°) for 1 hr. Protein precipitation time (on ice) was increased to 10 min, and protein precipitation centrifugation was performed for 25 min at 2300 \times g. DNA precipitation centrifugation was performed for 45 min at 2300 \times g, and the 70% ethanol wash was followed by 10 min centrifugation at 2300 \times g. Isopropanol/supernatant and 70% ethanol were removed by inverting the plate on a paper towel and centrifuging for 2 sec at minimum speed. All samples were resuspended in 200 μ l 0.1 \times TE, pH 8.0.

Genotyping: Most samples were genotyped using an oligo ligation assay (OLA) (GENISSEL *et al.* 2004; MACDONALD *et al.* 2005b). All oligonucleotide sequences for PCR and genotyping are given in supplemental Table 1 at <http://www.genetics.org/supplemental/>. Briefly, we amplified the region surrounding the site of interest by PCR (using one one-hundredth of the extracted DNA from each fly as template per 10- μ l reaction), and the amplified product was used as template for a ligation reaction. The ligation reaction introduced an allele-specific DNA "bar code," which was then amplified, printed onto filter arrays, and visualized using hybridization with radioactively labeled probes.

Amplification: MC22 and MC21 were amplified in a single amplicon (608, 616, or 621 bp, depending on genotype) that flanks both indels using the following cycling profile: 5 min at 95°; 35 cycles of 20 sec at 95°, 20 sec at 57°, 45 sec at 72°; 4 min at 72°; and hold at 4°. However, the *transpac* transposable

element insertion is 5249 bp long, making flanking PCR that amplifies both insertion and noninsertion alleles difficult. Therefore, forward and reverse primers that flank the insertion site were combined with one insert-specific reverse primer. This combination results in a 200-bp amplicon for the noninsert allele and a 491-bp amplicon for the insert allele. In a heterozygote, both amplify at approximately equal intensity (data not shown) under the following conditions: 2 min at 95°; 30 cycles of 20 sec at 95°, 20 sec at 58°, 2 min at 72°; 5 min at 72°; and hold at 4°. Although this assay can be typed on an agarose gel, for high-throughput purposes we used the PCR product as a template for an OLA (see supplemental Figure 1 at <http://www.genetics.org/supplemental/>).

Genotype calling: The remaining steps of the genotyping workflow (ligation, amplification, printing, and hybridization) were performed exactly as described previously for SNPs (MACDONALD *et al.* 2005b). Genotypes were determined from the consensus of two evaluations per OLA reaction for every population-polymorphism combination except IV04 × MC22 and IV04 × MC21 (assayed once). It is worth noting that although all the polymorphisms genotyped here are all indel variants (as opposed to SNPs), our OLA can easily be adapted to indels (see supplemental Figure 1 at <http://www.genetics.org/supplemental/>).

Genotyping NILs: The genotypes of most NILs were obtained by OLA; others were genotyped by sequencing (MC22, MC21) or by visualizing size differences on a 2% agarose gel (Transpac). For each locus, PCR was performed as described above. The amplicon containing MC22 and MC21 was presequenced enzymatically in the following reaction: 3 μl PCR product, 0.75 μl ddH₂O, 1 unit shrimp alkaline phosphatase (United States Biochemical, Cleveland), and 2.5 units Exonuclease I (United States Biochemical) were incubated at 37° for 1 hr and at 80° for 20 min. The equivalent of 1 μl unprocessed PCR product was used as template in a standard 5-μl sequencing reaction containing 1 μl BigDye, version 3.1 (Applied Biosystems, Foster City, CA), 0.5 μl “5× Sequencing Buffer” (Applied Biosystems), and 1.6 pmol sequencing primer (“scute-seqUTR,” supplemental Table 1 at <http://www.genetics.org/supplemental/>), cycled according to the manufacturer’s instructions.

Purging *D. simulans*: The wild populations SV96 and NV01 are potentially subject to contamination with *D. simulans* individuals. Probable *D. simulans* individuals were identified and removed from further consideration through methods previously described: 4 females and 2 males were identified in NV01 by clustering with the genotypes (at 200+ SNPs) of known *D. simulans* individuals (MACDONALD *et al.* 2005a); 168 females and 72 males from SV96 could not be positively identified as having the *D. melanogaster* allele in a PCR length difference assay (GENISSEL *et al.* 2004).

Haplotype estimation: For outbred males, haplotypes were determined directly from hemizygous genotype data. Among the 2596 outbred males successfully genotyped for the three markers, all but one belong to one of four common haplotypic classes (eight haplotypes expected under linkage equilibrium). Under the assumption that the male chromosomes represent all possible common haplotype configurations, nearly all outbred female diploid genotypes (2260 of 2286) could be deterministically decomposed into two constituent haplotypes.

Statistical analyses: All statistical tests were performed using the R statistical programming language (<http://www.r-project.org>). Each set of analyses was performed separately for each population, bristle character, and sex, except when testing population-genotype interactions.

Hardy-Weinberg equilibrium: We tested for Hardy-Weinberg genotype proportions using whole-population allele frequencies to calculate expected genotype frequencies in females. We used a χ^2 -test with 1 d.f.

Genotype-phenotype associations: The average effect of an allelic substitution (a) is estimated by the difference between allelic classes (hemizygous males) or the slope of a regression model (outbred females). In contrast, in the isogenic lines (including previously published results), the simple difference between (homozygous) classes estimates $2a$ for females. In all tables, figures, and text we present estimates of a , standardizing calculated effects and standard errors by halving where necessary.

For outbred females when the number of rare homozygotes was less than five, and for outbred males, we used Student’s t -test [$t.test()$ function in R, unpaired and two tailed, using separate variance estimations for each group and the Welch modification to degrees of freedom] to test for phenotypic differences between two genotypic classes. Otherwise (for outbred females) we tested for difference among the three genotypic classes by regressing the phenotype (y -axis) on the number of copies of the rarer allele (zero, one, or two copies, x -axis). An arbitrary dominance term was added to the additive regression by including an additional predictor that takes on a value of one if the female is a heterozygote and a zero otherwise. Confidence intervals (95%) on effect sizes were computed as the effect ± 1.96 times the standard error on the estimated effect.

An association with each phenotype was evaluated in the NIL data using the nested model $Y_{ijkl} = \mu + G_l + L_k(G_l) + V_j(L_k(G_l)) + \varepsilon_i$, where Y_{ijkl} is the number of bristles of the i th individual of the j th replicate Vial of the k th Line with l th Genotypic value. G_l is a fixed effect that corresponds to the number of copies of the rare allele ($l = \{0, 2\}$).

Haplotype-phenotype associations: For individuals who were successfully genotyped for all three polymorphisms, so that haplotypes are known for males or inferred for females, we tested for haplotype/bristle number associations using the model $Y_i = \mu + \sum_j h_{ij} + \varepsilon_i$, where Y_i is the number of bristles for the i th individual and h_{ij} is the number of copies of the j th haplotype harbored by the i th individual ($j = \{1, 2, 3, 4\}$). In males this model amounts to a one-way ANOVA on haplotype, whereas in females it tests the additive effect of haplotypes on phenotypic variation. We did not fit a haplotype model including dominance to the female data as it would involve estimating six dominance terms.

Model comparisons: We assessed the significance of some linear model terms (*e.g.*, population-genotype interaction, an additional polymorphism, etc.) by comparing a full to a reduced model. The test statistic $F = ((SS_{\text{resid, reduced}} - SS_{\text{resid, full}}) / (d.f._{\text{resid, reduced}} - d.f._{\text{resid, full}})) / ((SS_{\text{resid, full}}) / d.f._{\text{resid, full}})$ is distributed $F_{(d.f., \text{reduced} - d.f., \text{full}), (d.f., \text{total, full})}$ under the null hypothesis of no significant difference between models.

RESULTS

Phenotypic diversity: Means and standard deviations for bristle phenotypes from the three outbred populations are summarized in Table 1. Phenotypic measures for the two wild populations have been previously reported (GENISSEL *et al.* 2004; MACDONALD and LONG 2004, MACDONALD *et al.* 2005a). The additional populations presented were all raised in the laboratory. Contrasting the two “environments,” it appears that our laboratory conditions have no specific fixed effect on phenotype. Rather, either the environments seen by the flies are a random effect or the heterogeneity between

TABLE 1
Mean and standard deviations of population data

	SV96	NV01	IV04	W1	W3	WE
	Average bristle no.					
Male ABN	16.6 (2.44)	15.7 (2.34)	17.2 (2.08)	16.1 (1.61)	17.0 (2.39)	16.0 (1.73)
Male SBN	16.8 (2.09)	16.7 (2.16)	17.6 (1.64)	18.1 (2.34)	17.4 (2.27)	18.1 (2.16)
Female ABN	18.2 (2.68)	18.2 (2.76)	20.2 (2.61)	17.1 (2.65)	17.8 (3.53)	17.0 (3.06)
Female SBN	17.1 (2.09)	17.3 (2.19)	18.6 (1.63)	18.9 (2.35)	18.4 (2.19)	19.0 (2.32)
	Minor allele frequency					
MC22	0.266	0.256	0.0547	0.111	0.267	0.415
MC21	0.291	0.315	—	0.386	0.386	0.171
Transpac	0.110	0.0834	0.0553	0.0417	0.0667	0.097
<i>N</i> ^a	2580	2784	2836	48	45	41

^aTotal number of wild chromosomes in each genotyped panel (two per outbred female, one per outbred male or NIL).

sampled populations overwhelms any environmental signal.

Genotypic diversity: The three polymorphisms examined in this study are all length polymorphisms in which the shorter allele is more frequent. MC22 is 97 bp 3' of the *scute* transcript's polyadenylation signal. In Release 4.3 of the *D. melanogaster* genome, the common allele "MC22−" matches positions 254401–254408 (GCCTTGCA), while the MC22+ allele substitutes the sequence TATAACGCATTCCGCGT. An alignment of the region from *D. melanogaster*, *D. simulans*, and *D. sechellia* suggests that MC22+ (the rarer allele) is ancestral, although multiple mutations are needed to fully explain the alignment (data not shown). However, for "MC21" the rarer allele appears to be derived. About 175 bp from MC22, the MC21+ allele is a simple insertion of the sequence ATGTCAAATGTCA at genome position 254583. The third polymorphism we examined is a particular instance of a *transpac* element at genome position 308419 (we refer to this insertion polymorphism as "Transpac," equivalent to "InR" in LONG *et al.* 2000). Relative to the *D. simulans* and *D. sechellia* genomes, the insertion allele Transpac+ appears to be derived. In the sequenced *D. melanogaster* genome, the insertion would be ~42 kb downstream of *l'scute*, 9 kb upstream of *asense*, and ~15 kb upstream of *pepsinogen-like* (the only other gene between *l'sc* and *ase*, transcribed in the opposite orientation). Although vast tracts of regulatory DNA exist in *ASC* (CAMPUZANO *et al.* 1985), VISTA plots (<http://pipeline.lbl.gov/cgi-bin/gateway2?bg=dm1>) comparing sequence similarity between *D. melanogaster* and distantly related *Drosophila* do not show marked conservation in the regions containing MC22, MC21, or Transpac.

Allele frequencies are clearly very similar between the two wild populations, although they are evidently quite different in IV04 and the NILs (Table 1). Unlike the wild populations, in which we observe four common haplotypes, in IV04 we observe only two haplotypes total,

with a single haplotype (MC22−, MC21−, Transpac−) accounting for ~95% of all chromosomes observed. The NILs represent collections of wild genotypes, and pooling them with two previous studies (MARTÍN-CAMPOS *et al.* 1992; LONG *et al.* 2000) is anecdotally suggestive of geographic differences in MC22 and MC21 frequency. However, note that W1 and W3 were collected in the same orchard (different times of year), yet the frequency of MC22+ is 0.111 in W1 and 0.267 in W3. Therefore a detailed and comprehensive survey will be needed to legitimately assign frequency differences to season, geography, or sampling error.

The tip of the *D. melanogaster* X chromosome has a greatly reduced crossover rate relative to physical distance when compared to the remainder of the *Drosophila* genome, and thus linkage disequilibrium (LD) extends over large physical distances in this region (AGUADÉ *et al.* 1989). In this study we observe only three gametic types involving MC22 and either Transpac (Transpac+ is seen only on the MC22+ background) or MC21 (MC21+ is seen only on the MC22− background); hence there is no evidence for historical recombination events separating these markers. This contrasts with LONG *et al.*'s (2000) description of two MC22/Transpac recombinant chromosomes. We speculate that LONG *et al.*'s (2000) recombinants likely represent genotyping errors (Transpac's status was inferred using relatively crude Southern blots), and those errors perhaps colored the interpretation of genotype–phenotype associations in the region.

Departure from Hardy–Weinberg Equilibrium: In wild populations SV96 and NV01, we note several significant departures from Hardy–Weinberg proportions in females (SV96, $P_{MC21} = 0.017$, $P_{Transpac} = 0.016$; NV01, $P_{MC22} = 0.011$). In all cases (including non-significant tests) there appears to be a slight excess of heterozygotes. Although our tests for departure from Hardy–Weinberg equilibrium (HWE) are highly significant, it is unlikely they would have achieved statistical

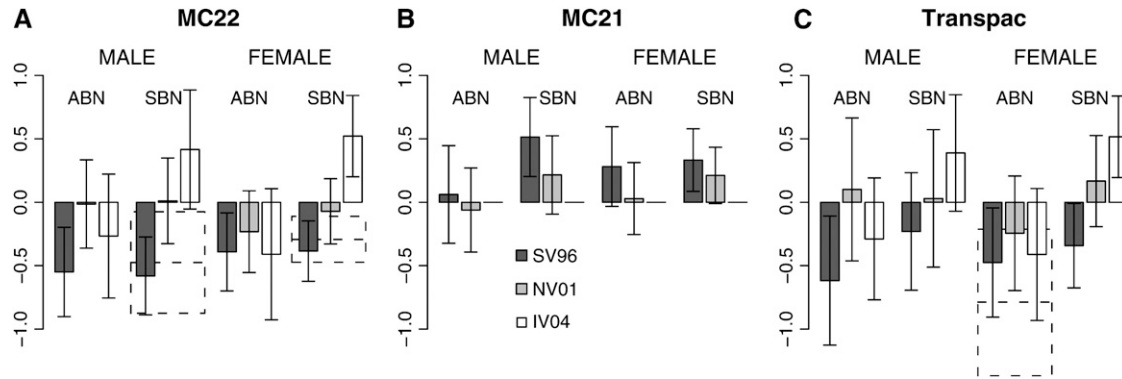


FIGURE 1.—Estimated effects of markers on bristle phenotypes in outbred populations. Allelic effects, a , are shown in terms of the change in bristle number if a rare allele were substituted for a common allele, grouped by polymorphism: (A) MC22, (B) MC21, (C) Transpac. Error bars indicate 95% confidence intervals on the mean; nonoverlap with 0 indicates nominal significance. Dashed rectangles show the effects of the alleles observed by LONG *et al.* (2000) in chromosome extraction lines (centerline, mean; top and bottom edge, $\pm 1.96 \times \text{SEM}$).

significance if the study had employed a more modest sample of wild individuals (MACDONALD *et al.* 2005b). An excess of heterozygotes is not a pattern commonly observed with this genotyping method (on balance, “no calls” generally cause a slight *deficit* of heterozygotes). There is also no evidence for any sort of systematic genotyping error. For these reasons we do not believe the observed departure from HWE is due to methodological error. We note that tests for phenotype–genotype associations are not affected by departures from HWE as long as genotypes are correct.

MC22 genotype–phenotype associations: Previous work (LONG *et al.* 2000) detected a significant association between MC22 and SBN in both males and females, an association between Transpac and ABN in females, and a sex-by-genotype interaction between MC21 and ABN. To ascertain if these associations replicate in large outbred populations and in a set of NILs, we regressed bristle phenotypes on genotypes at these three markers. Surprisingly, associations were highly dependent on the population examined.

In the SV96 population, MC22+ is associated with a reduction of bristle number for both ABN and SBN in both males and females (ABN, $P_{\text{male}} = 0.003$, $P_{\text{female}} = 0.011$; SBN, $P_{\text{male}} = 0.0002$, $P_{\text{female}} = 0.001$). Figure 1 demonstrates similarities between LONG *et al.* (2000) and the SV96 wild population by replicating the association between MC22 and SBN, with SV96 being additionally significant for ABN. However, results are different in the other populations. In the IV04 sample we observe MC22 to be significantly ($P = 0.001$) associated with female SBN, *but with effects in the opposite direction* from that predicted by previous studies. A similar nonsignificant ($P = 0.081$) trend is apparent for male SBN. However, in a second wild-caught population, NV01, there is no apparent association between MC22 and bristle number variation. (Regardless of population or trait, arbitrary dominance terms were never significant, $P > 0.09$.)

For the two wild populations, linear models testing for an interaction between MC22 and population of origin are nominally significant for both male ABN and SBN ($P = 0.038$ and $P = 0.010$, respectively). This suggests that observed differences in male bristle number effects at MC22 are biological in origin and not simply due to chance. Whereas, for female ABN and SBN there is no significant interaction between wild population of origin and MC22 genotype ($P = 0.474$ and $P = 0.074$). When all three outbred populations are considered, SBN shows significant MC22-by-population interactions ($P_{\text{male}} = 0.001$ and $P_{\text{female}} = 0.0003$), but ABN does not ($P_{\text{male}} = 0.097$ and $P_{\text{female}} = 0.726$).

We also sought to replicate associations in a moderately sized panel of 134 NILs. Similar to our NV01 collection of outbred flies, the entire panel of NILs also showed no convincing association between bristle number and genotype of any of the *ASC* loci examined, with the marginal exception of a nominal association between Transpac and female ABN ($P = 0.043$). These lines were collected in two localities (California and North Carolina) and at different times of year; additionally, they have different allele frequencies (Table 1). Therefore we analyzed the Californian collections (W1 + W3, ~ 90 lines), as well as each individual collection (> 40 lines each) separately. MC22 is not associated with any bristle phenotype in the subpopulations. There are four marginally significant associations between bristle phenotypes and MC21 or Transpac, similar in direction and magnitude to those observed in our SV96 or by LONG *et al.* (2000). However, they are all based on extremely small representation (2, 4, or 7 lines) of the minor allele class, and none were consistent over multiple NIL collections.

Testing associations between *ASC* haplotypes and bristle number phenotypes: With three polymorphisms in strong LD, each suggestive of a phenotypic effect in SV96, two scenarios deserved investigation: either the haplotype defined by the markers has a phenotypic

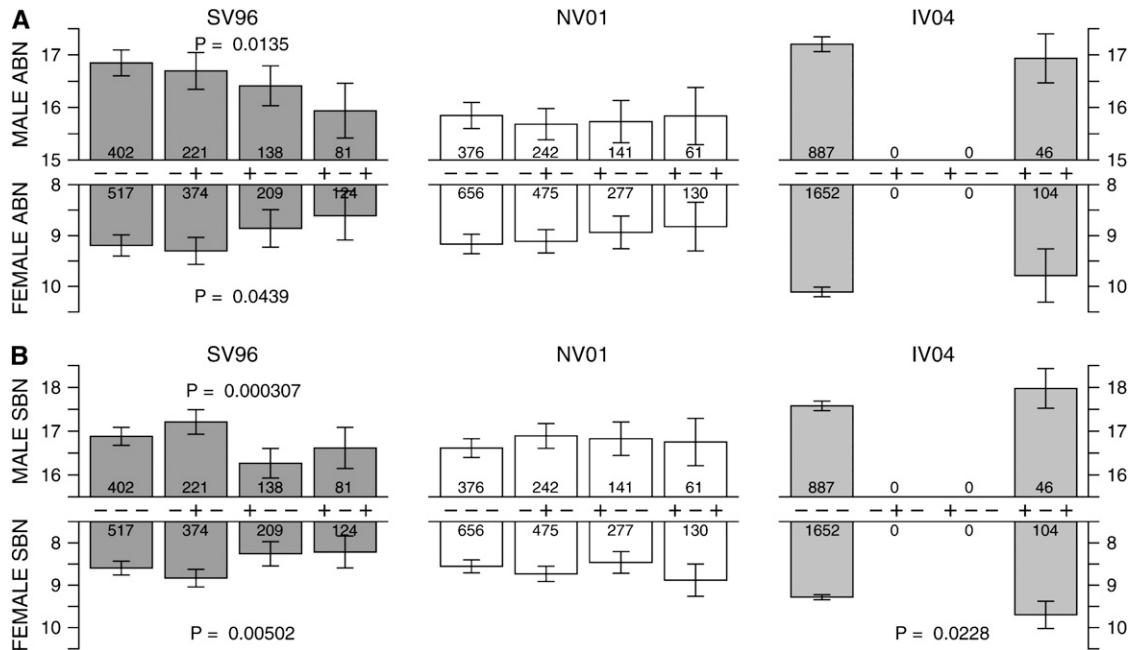


FIGURE 2.—Phenotypic estimates of haplotype classes in outbred populations. Haplotypes are ordered in terms of frequency in wild populations (allele counts given inside the bars). Each haplotype is defined by the absence (“−”) or presence (“+”) of the insertion allele, in the genomic order MC22, MC21, Transpac. Error bars indicate 95% confidence intervals ($\pm 1.96 \times \text{SEM}$). Significant differences among haplotypes have the associated P -value displayed (for purposes of analysis in IV04, the lone rare-homozygote female was excluded). The top axis plots male phenotypes and the bottom axis plots female phenotypes (scale is halved because each actual female is composed of two alleles). (A) Abdominal bristle number; (B) sternopleural bristle number.

consequence or all polymorphisms are providing evidence of an identical effect through LD with each other or with an untyped causal polymorphism. In our outbred populations, an inspection of Transpac’s “effects” (Figure 1C) reveals striking similarities to MC22 (Figure 1A), although with generally less-significant statistical support. Also, the measured effect of MC21+ in SV96 is an increase in bristle number (Figure 1B). Both of these observations are in accord with the coupling of Transpac+ with MC22+ and the repulsion between MC21+ and MC22+. Therefore, we wished to determine if MC21 and Transpac have effects on bristle number independent of the effects associated with MC22 in SV96 and NV01. (This is not possible in IV04, as MC21 is not polymorphic, and the pair of loci MC22 and Transpac are completely associated; *i.e.*, $R^2 = 1$.) To this end, we fit a linear model of MC22’s effect on each phenotype in each population, as well as a slightly more complex model that additionally included a term for an effect of MC21 or Transpac. At $P < 0.05$, in no case did the model with an additional locus result in a significantly better fit to the data in either wild population. We also performed association analyses focusing on three-locus haplotypes. Figure 2 shows the outbred populations’ phenotypic effects associated with the four common haplotypic classes at ASC. Even for phenotypes where haplotype is nominally significant (SV96, all traits; IV04, female SBN), inspection of pairs of haplotypes that differ only by Transpac (*i.e.*, $+ - -$ vs. $+ - +$)

or by MC21 (*i.e.*, $- - -$ vs. $- + -$) does not reveal significant differences. Thus analyses at the level of haplotypes confirm analyses at the level of single markers: the marker state at MC22 is sufficient to explain any and all associations between genotype and phenotype observed at ASC. Finally, in the panel of NILs, no significant effects of haplotype on bristle phenotype were observed in either the concatenated data set or the separate collections.

We tested for the presence of a population–haplotype interaction by concatenating the two outbred wild data sets and reanalyzing. No trait was significantly affected by the interaction term. It is worth noting that when SV96 and NV01 were analyzed together, the main effect of haplotype on bristle phenotype was always nonsignificant.

DISCUSSION

We have characterized the impact of three sites, observed to be associated with bristle number variation in prior studies (LONG *et al.* 2000), with bristle number variation in three large outbred panels of *Drosophila* and a moderately sized panel of NILs. Bristle number, in addition to being a classic quantitative model system, is one of just a handful of traits that has proved quick and reliable to measure in large numbers of wild organisms regardless of their age or size, thereby making

it amenable to this varied array of association study strategies.

In this study, our paradigm for authenticating an effect of variant on phenotype was that of repeatability in additional populations. This is the most widely applicable stratagem, as it can be applied to any organism and is the most straightforward means of observing genetic effects in natural environments. Another advantage will be increasingly obvious as whole-genome association studies become realistic: repeating such studies will address the association or nonassociation of many polymorphisms simultaneously. The repeatability paradigm's shortcomings are most apparent in instances of failure to replicate, as the comparison provides no experimental explanations but ample room for speculation. There are other options in model organisms. In *Drosophila*, crosses of putative variants into various genetic backgrounds have been used to verify associations from inbred lines, although at least one group specifically notes that resampling large outbred populations would have been more efficient (PALSSON *et al.* 2005). It has long been possible to insert genes at random positions in the *Drosophila* genome, and with sufficient replication this has been used to dissect quantitative traits, such as the classical effects of *Alcohol dehydrogenase* alleles (STAM and LAURIE 1996). Moreover, the field of *Drosophila* quantitative genetics is on the edge of having access to genetic manipulation of single endogenous nucleotides, such as that performed at the *desaturase2* gene (GREENBERG *et al.* 2003). However, even this method is potentially subject to controversy, as some have claimed that the effect of *desaturase2* on cold tolerance and sexual isolation cannot be repeated in the manipulated lines (COYNE and ELWYN 2006). This controversy reminds us that technical advances do not substitute for robust controls and replication (particularly over many genetic backgrounds) in all evolutionary genetics experiments. In any case, genetic manipulations as a whole are inherently conducted gene-by-gene and are likely to remain labor intensive for the foreseeable future.

In the present study, all three polymorphisms are located in the ASC complex, which contains the proneural genes *achaete* (*ac*), *scute* (*sc*), *lethal of scute* (*l'sc*), and *asense* (*ase*). In particular, the expression of *ac* and *sc* specifies bristle placement (SIMPSON 1990). Although no specific biochemical consequence has been assigned to these natural alleles, previous work on this complex has found that bristle number is associated with natural variation at small indels at intermediate frequency (LONG *et al.* 2000) and with transposable element insertions at both low and intermediate frequency (MACKAY and LANGLEY 1990; LONG *et al.* 2000).

Our modern genotyping technology and large sample sizes lead us to conclude that the natural variants we genotype at ASC are in complete LD ($D' = 1$), having never been separated by recombination. Therefore

associations between each polymorphism and bristle phenotypes are not independent. For this reason we consider only MC22 in subsequent discussion because it shows the strongest association with bristle number variation. We observed a general concordance between the inbred/laboratory study of LONG *et al.* (2000) and the outbred/wild population SV96, having observed association between MC22 and SBN in both studies. However, our additional (nearly isogenic) laboratory panel as well as another wild population does not show associations between bristle phenotypes and this locus. Finally, the outbred/laboratory population IV04 shows significant associations but in the opposite direction from other types of cohorts.

We think it unlikely that detecting a robust, significant effect of MC22 on SBN would occur spuriously using two different experimental strategies on genetic material collected thousands of miles apart (SV96 here and that in LONG *et al.* 2000), although false-positive associations at $P < 0.0013$ in one of three replicate follow-up populations remain a formal possibility. Left unexplained is why associations between genotype and phenotype are not always present (or if present, not always detected), even when levels of genotypic and phenotypic variation appear similar. As they pertain to these data, we explore the concepts of LD, sampling variation, genotype–environment interaction, and epistasis (see also a comprehensive review of bristle genetics in MACKAY and LYMAN 2005).

Linkage disequilibrium: In the process of conducting these studies, we collected genotype data from thousands of outbred flies, about half of which are hemizygous males. In all these data, no evidence for historical recombination between markers was detected, in spite of MC22 and Transpac being 50 kb apart. This is perhaps expected given the low levels of crossover in this region (AGUADÉ *et al.* 1989), but it also suggests that previously observed associations between bristle number variation and various polymorphisms in ASC may be explained by LD with an ungenotyped polymorphism. Unfortunately, we have not proved that MC22 is the causal site, and the prospects of this are particularly daunting given the region's LD. It is likely that there are additional sites strongly associated with MC22, any of which are equally good *a priori* candidate quantitative trait nucleotides (QTNs).

Might a breakdown in association between markers typed in this study and an untyped causative SNP be responsible for our variable associations with phenotype? For at least two sites we have identified, the answer is no. An attractive site in the 3'-UTR of the *sc* transcript is completely associated with MC22 in samplings of both the (significant) SV96 and (nonsignificant) NV01 population (data not shown). The disequilibrium between MC22 and a SNP ~150 bp from the Transpac insertion was also equally strong in both wild populations ($R^2 = 0.88$; data not shown). However, as long as any common

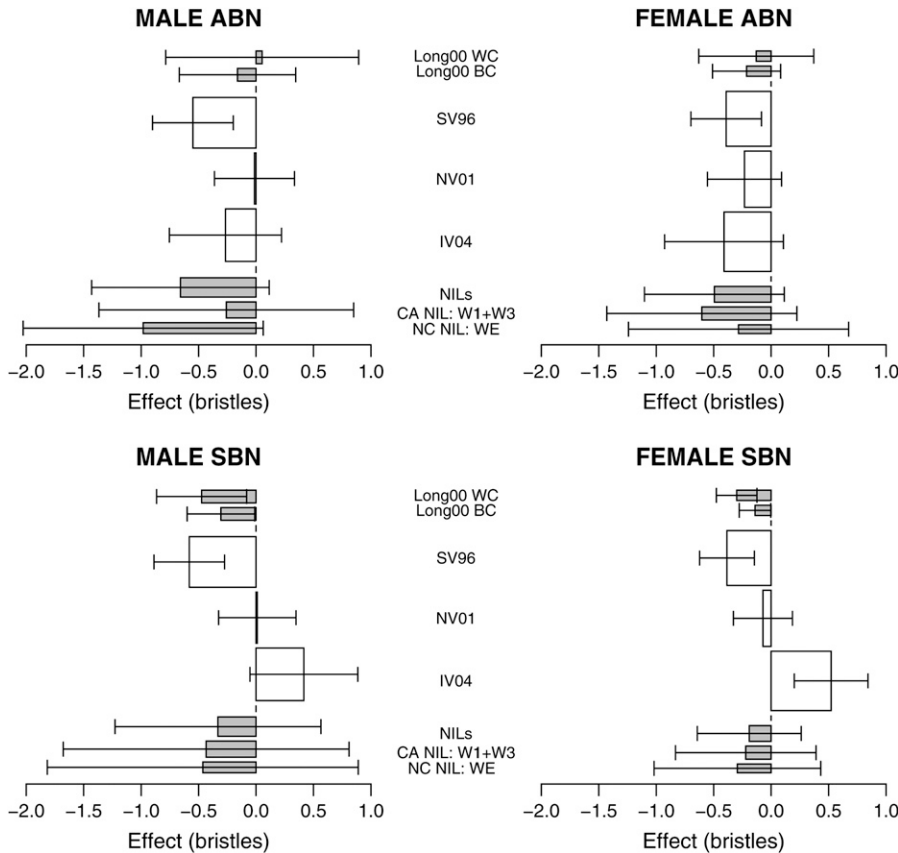


FIGURE 3.—Compiled estimates of MC22's effect on bristle number phenotypes. Plotted are estimates of a , the average effect of substituting a rare allele MC22+ in the place of the common allele MC22-, from the literature (LONG *et al.* 2000) and our association study panels. Error bars indicate 95% confidence intervals. Each section surveys a different number of wild chromosomes, and the height of each bar is proportional to square root of this number. Open bars are outbred populations, in which every female contributes two wild chromosomes to the overall panel. Shaded bars represent studies of isogenic strains; each individual phenotyped is homozygous, but many individuals are phenotyped per strain. The "Long00 WC" strains were isogenic except for the whole X chromosome, the "Long00 BC" strains were backcrossed to be isogenic except for the tip of the X chromosome (LONG *et al.* 2000). NILs were collected from the wild as isofemale lines and inbred by brother-sister mating for 15–40 generations (YANG and NUZHIDIN 2003; PALSSON *et al.* 2004).

polymorphisms in the region remain unknown and ungenotyped, they remain a possible explanation for any associations.

An additional consequence of the low recombination rate at the tip of the X chromosome is that large physical stretches of DNA sequence can "hitchhike" on the background of a selected allele. Our IV04 collection comes from the Ives laboratory population, which has undergone a consistent but unnatural selection regime for hundreds of generations and probably differs in its genetic history from typical wild populations in other, unknown ways. Our results demonstrate that "outbred/laboratory" populations are not necessarily a panacea to resolving the hazards of association studies of wild individuals. One possibility is that the unusual haplotype spectrum we observe, as well as the "backward" association between MC22 and bristle number, could both be explained by a sweep (the result of a laboratory selective pressure) that randomly associates MC22 with another bristle QTN of larger, opposite effect. Allele frequencies for autosomal SNPs in IV04 do not show such extreme frequency differences (data not shown), suggesting that demographic effects such as a bottleneck are implausible.

Sampling variation: SV96 and NV01, which are similar at the levels of geography, environment, ASC genotype spectrum/LD, and bristle number variation, differ in whether they replicate the original associations

between bristle phenotypes and MC22 (LONG *et al.* 2000). Figure 3 compares the MC22 effect estimates across all ASC association studies. It is striking that the large majority of studies are biased toward a negative estimate of a , even when nonsignificant. A potential nonbiological explanation of these results is that our significant associations overestimate (and nonsignificant associations underestimate) the true effect of MC22. If the IV04 population is set aside because of its unique ASC haplotype frequencies, confidence intervals of all other studies do overlap, reminiscent of the human PPAR γ studies in which a metaanalysis concludes that a true association exists for the polymorphism, albeit with a much smaller effect size than the initial study indicated (ALTSHULER *et al.* 2000). Power analysis on the SV96 and NV01 populations shows 90% power to detect an effect as small as 0.38 sternopleural bristles (MACDONALD and LONG 2004), but if the true effect is smaller, power is sharply reduced (see Figure 2 of MACDONALD and LONG 2004).

Although it can be argued that all the populations are consistent with a single small effect on bristle number, such a hypothesis is inconsistent with the observation that male phenotypes from the two outbred/wild populations display a significant population-genotype interaction. Although a significant but artifactual population-genotype interaction could be expected if one of the populations had served to initially ascertain

that MC22 is associated with bristle number variation (BEAVIS 1998), this does not apply here because both wild populations were genotyped to replicate an independent association initially obtained using isogenic lines in the laboratory (LONG *et al.* 2000). This suggests that the differences observed here are biological and not statistical in origin.

Two other studies in *Drosophila* have encountered similar transitory associations, albeit when comparing associations in inbred/laboratory panels outbred/wild populations. PALSSON and GIBSON (2004) found a significant association between *D. melanogaster* wing-shape metrics and a SNP near *Egfr*, using ~100 NILs in the laboratory. A follow-up study (DWORKIN *et al.* 2005) found that the same allele had a significant effect in a cohort of 871 wild-caught males. However, an extremely important distinction made in the follow-up study was that *the effect of the allele was 10-fold weaker in the natural cohort* and the wing-shape character examined was slightly different than it had been in the inbred lines. Furthermore, expanded genotyping to fill in missing data from the original study resulted in the California-derived part of the original experiment becoming decidedly nonsignificant, although the effect of the allele was detectable in a number of other collections of genetic material (PALSSON *et al.* 2005). The other relevant comparison is a follow-up association study on outbred/wild samples (MACDONALD and LONG 2004) of a putative SBN QTN at *hairy*. The original, highly significant association was identified in an inbred/laboratory study (ROBIN *et al.* 2002). Despite ample statistical power, the follow up failed to replicate in either of two natural populations (SV96 and NV01, as used in this study).

In each of these examples, it is impossible to say whether the differences in effect are a function of comparing inbred *vs.* outbred, nature *vs.* laboratory, or one association study *vs.* another (of any kind). Unfortunately, our set of studies illustrates discrepancies in potentially all of these categories. We feel that future experiments to rigorously test each possibility will be consequential.

Genotype–environment interactions: The nature *vs.* laboratory possibility suggests the importance of genotype–environment interactions (GEI) for quantitative traits. GEI has been demonstrated for genes and quantitative trait loci (QTL) of many traits. Specifically, GURGANUS *et al.* (1998) demonstrated GEI variance for bristle number on QTL throughout the genome. In that study, a marker linked to the ASC region showed no significant effect at 18° and 25°. Yet at 29°, the marker explained more genetic variation of female ABN than any other marker. GEIGER-THORNSBERRY and MACKAY (2002) describe GEI (for the association between bristle phenotypes and polymorphisms in *Delta*) when environments differ by media type or rearing temperature. One potential pitfall of the previous work was that only three representative strains per genotype were examined.

Today, no technological or economic barriers remain that would preclude experiments that seek to describe GEI at either *Delta* or the ASC in large, outbred populations reared under different conditions. The bristle number phenotype is well suited for further study of GEI with a phenotypic association to a molecular marker, as larval crowding (KEARSEY and BARNES 1970; IMASHEVA and BUBLIY 2003), temperature (IMASHEVA *et al.* 1998), and nutrition (IMASHEVA *et al.* 1999) have an effect on phenotypic variation for ABN and SBN and are easily manipulated in the laboratory.

Epistasis: Although initially believed to be a trait dominated by additive interactions, recent studies have documented significant epistatic variation for bristle phenotypes (MACKAY and LYMAN 2005). One of the first QTL studies of bristle phenotypes using neutral genetic markers noted a QTL in the vicinity of ASC and many epistatic interactions among autosomal QTL, but could not document an epistatic interaction between the “ASC” QTL and other sex-linked loci (LONG *et al.* 1995). Creating introgression panels on a grander scale, *i.e.*, breeding an identical panel of natural X chromosomes into multiple inbred backgrounds and looking for associations, would require large but not unprecedented quantities of bristle counting and only trivial amounts of genotyping. This may address the issue of epistasis between ASC and autosomal loci. Epistatic interactions themselves have been shown to be sensitive to environmental factors for *Drosophila* life span (LEIPS and MACKAY 2000), so integrating different environmental conditions as discussed above would also be fruitful.

Using bristle number variation and a polymorphism in the ASC of *D. melanogaster*, we show that associations between variants and phenotypes may be difficult to replicate in populations of different genetic makeups or obtained from samplings at different environments. Although there is no strong evidence that bristle number is an atypical polygenic trait (in terms of GEI or epistasis), the ASC is an atypical genomic region and it is possible that this would lead to aberrant patterns of genotype–phenotype associations.

The tip of the X chromosome of *D. melanogaster* is characterized by little recombination (AGUADÉ *et al.* 1989). All else being equal, a selective sweep (either due to the fixation of a beneficial allele or due to the elimination of a deleterious one) will affect more genes here than in a highly recombining region. The result is that a focal locus at the tip of the X chromosome will experience sweeps more commonly than other regions. In IV04, the reduced haplotype diversity and unexpected direction of the effects at MC22 are consistent with a sweep scenario.

In the two wild populations, we also observe a slight deficit of homozygotes relative to HWE. This pattern could result from closely linked deleterious recessive alleles. Reduced recombination implies that selection is

less efficient at removing specific deleterious variants (due to Muller's ratchet), and thus strong "deleterious" alleles contributing to bristle number variation may segregate at higher frequencies than elsewhere in the genome. It is possible that our results are influenced by transient associations with haplotypes harboring deleterious bristle alleles that have not yet been eliminated by natural selection. Two previous studies did suggest that, as a class, strains with *ASC* alleles harboring one or more transposable element insertions have fewer bristles than strains with insertion-free alleles (MACKAY and LANGLEY 1990; LONG *et al.* 2000).

We have endeavored to address whether MC22, a strong candidate as a polymorphism having an effect on a complex phenotype, behaves predictably in different populations and under diverse experimental strategies. This collection of association studies makes clear that much remains to be understood about the relationship between candidate loci, molecular markers, and phenotypic variation. As long as this is true, our understanding of the relationship between genetic variation and evolution is incomplete at best and naïve at worst. Systematically altering the environmental or genomic context of a population that shows association between *ASC* genotypes and bristle number will lend considerable insight into the relationship between genetic and phenotypic variation.

We thank B. P. Lazzaro for help with bristle counting. For providing fly strains, we thank M. R. Rose and C. Rauser for the Ives population; S. Nuzhdin for the W1 and W3 NIL panels; and G. Gibson, L. Goering, and M. Arbeitman for WE nearly isogenic lines. We also thank two anonymous reviewers for helpful comments on the manuscript. This work was supported by National Institutes of Health grant GM-58564 to A.D.L., by National Science Foundation grant DEB-0614429 to A.D.L., by National Institutes of Health service award 5 T15 LM007443 to J.D.G., and by a fellowship from the Achievement Rewards for College Scientists Foundation of Orange County to J.D.G.

LITERATURE CITED

- AGUADÉ, M., N. MIYASHITA and C. H. LANGLEY, 1989 Reduced variation in the *yellow-achaete-scute* region in natural populations of *Drosophila melanogaster*. *Genetics* **122**: 607–615.
- ALTSHULER, D., J. N. HIRSCHHORN, M. KLANNEMARK, C. M. LINDGREN, M.-C. VOHL *et al.*, 2000 The common PPAR γ Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat. Genet.* **26**: 76–80.
- AMUNDADOTTIR, L. T., P. SULEM, J. GUDMUNDSSON, A. HELGASON, A. BAKER *et al.*, 2006 A common variant associated with prostate cancer in European and African populations. *Nat. Genet.* **38**: 652–658.
- BEAVIS, W. D., 1998 QTL analyses: power, precision, and accuracy, pp. 145–162 in *Molecular Dissection of Complex Traits*, edited by A. H. PATERSON. CRC Press, New York.
- CAMPUZANO, S., L. CARRAMOLINO, C. V. CABRERA, M. RUIZ-GOMEZ, R. VILLARES *et al.*, 1985 Molecular genetics of the *achaete-scute* gene complex of *D. melanogaster*. *Cell* **40**: 327–338.
- CARBONE, M. A., K. W. JORDAN, R. F. LYMAN, S. T. HARBISON, J. LEIPS *et al.*, 2006 Phenotypic variation and natural selection at *Catsup*, a pleiotropic quantitative trait gene in *Drosophila*. *Curr. Biol.* **16**: 912–919.
- CARDON, L. R., and J. I. BELL, 2001 Association study designs for complex diseases. *Nat. Rev. Genet.* **2**: 91–99.
- COYNE, J. A., and S. ELWYN, 2006 Does the *desaturase2* locus in *Drosophila melanogaster* cause adaptation and sexual isolation? *Evolution* **60**: 279–291.
- DUAN, S., J. DU, Y. XU, Q. XING, H. WANG *et al.*, 2006 Failure to find association between TRAR4 and schizophrenia in the Chinese Han population. *J. Neural Transm.* **113**: 381–385.
- DWORKIN, I., A. PALSSON and G. GIBSON, 2005 Replication of an *Egfr*-wing shape association in a wild-caught cohort of *Drosophila melanogaster*. *Genetics* **169**: 2115–2125.
- GEIGER-THORNSBERRY, G. L., and T. F. C. MACKAY, 2002 Association of single-nucleotide polymorphisms at the *Delta* locus with genotype by environment interaction for sensory bristle number in *Drosophila melanogaster*. *Genet. Res.* **79**: 211–218.
- GENISSEL, A., T. PASTINEN, A. DOWELL, T. F. C. MACKAY and A. D. LONG, 2004 No evidence for an association between common nonsynonymous polymorphisms in *Delta* and bristle number variation in natural and laboratory populations of *Drosophila melanogaster*. *Genetics* **166**: 291–306.
- GRAHAM, R. R., S. V. KOZYREV, E. C. BAECHLER, M. V. P. L. REDDY, R. M. PLENGE *et al.*, 2006 A common haplotype of interferon regulatory factor 5 (*IRF5*) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. *Nat. Genet.* **38**: 550–555.
- GRANT, S. F. A., G. THORLEIFSSON, I. REYNISDOTTIR, R. BENEDIKTSSON, A. MANOLESCU *et al.*, 2006 Variant of transcription factor 7-like 2 (*TCF7L2*) gene confers risk of type 2 diabetes. *Nat. Genet.* **38**: 320–323.
- GREENBERG, A. J., J. R. MORAN, J. A. COYNE and C.-I. WU, 2003 Ecological adaptation during incipient speciation revealed by precise gene replacement. *Science* **302**: 1754–1757.
- GURGANUS, M. C., J. D. FRY, S. V. NUZHIDIN, E. G. PASYUKOVA, R. F. LYMAN *et al.*, 1998 Genotype-environment interaction at quantitative trait loci affecting sensory bristle number in *Drosophila melanogaster*. *Genetics* **149**: 1883–1898.
- HAINES, J. L., M. A. HAUSER, S. SCHMIDT, W. K. SCOTT, L. M. OLSON *et al.*, 2005 Complement factor H variant increases the risk of age-related macular degeneration. *Science* **308**: 419–421.
- IMASHEVA, A. G., and O. A. BUBLIY, 2003 Quantitative variation of four morphological traits in *Drosophila melanogaster* under larval crowding. *Hereditas* **138**: 193–199.
- IMASHEVA, A. G., V. LOESCHKE, L. A. ZHIVOTOVSKY and O. E. LAZEBNY, 1998 Stress temperatures and quantitative variation in *Drosophila melanogaster*. *Heredity* **81**: 246–253.
- IMASHEVA, A. G., D. V. BOSENKO and O. A. BUBLIY, 1999 Variation in morphological traits of *Drosophila melanogaster* (fruit fly) under nutritional stress. *Heredity* **82**: 187–192.
- JEONG, S. H., E. J. JOO, Y. M. AHN, K. Y. LEE, Y. S. KIM *et al.*, 2006 Investigation of genetic association between human Frizzled homolog 3 gene (*FZD3*) and schizophrenia: results in a Korean population and evidence from meta-analysis. *Psychiatry Res.* **143**: 1–11.
- KEARSEY, M. J., and B. W. BARNES, 1970 Variation for metrical characters in *Drosophila* populations. II. Natural Selection. *Heredity* **25**: 11–21.
- KLEIN, R. J., C. ZEISS, E. Y. CHEW, J.-Y. TSAI, R. S. SACKLER *et al.*, 2005 Complement factor H polymorphism in age-related macular degeneration. *Science* **308**: 385–389.
- LAI, C., R. F. LYMAN, A. D. LONG, C. H. LANGLEY and T. F. C. MACKAY, 1994 Naturally occurring variation in bristle number and DNA polymorphism at the *scabrous* locus of *Drosophila melanogaster*. *Science* **266**: 1697–1702.
- LEIPS, J., and T. F. C. MACKAY, 2000 Quantitative trait loci for life span in *Drosophila melanogaster*: interactions with genetic background and larval density. *Genetics* **155**: 1773–1788.
- 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., R. F. LYMAN, C. H. LANGLEY and T. F. C. MACKAY, 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., R. F. LYMAN, A. H. MORGAN, C. H. LANGLEY and T. F. C. MACKAY, 2000 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* **154**: 1255–1269.
- LYMAN, R. F., and T. F. C. MACKAY, 1998 Candidate quantitative trait loci and naturally occurring phenotypic variation for bristle number in *Drosophila melanogaster*: the *Delta-Hairless* gene region. *Genetics* **149**: 983–998.
- MACDONALD, S. J., and A. D. LONG, 2004 A potential regulatory polymorphism upstream of *hairy* is not associated with bristle number variation in wild-caught *Drosophila*. *Genetics* **167**: 2127–2131.
- MACDONALD, S. J., T. PASTINEN and A. D. LONG, 2005a The effect of polymorphisms in the *Enhancer of split* gene complex on bristle number variation in a large wild-caught cohort of *Drosophila melanogaster*. *Genetics* **171**: 1741–1756.
- MACDONALD, S. J., T. PASTINEN, A. GENISSEL, T. W. CORNFORTH and A. D. LONG, 2005b A low-cost open-source SNP genotyping platform for association mapping applications. *Genome Biol.* **6**: R105.
- MACKAY, T. F. C., and C. H. LANGLEY, 1990 Molecular and phenotypic variation in the *achaete-scute* region of *Drosophila melanogaster*. *Nature* **348**: 64–66.
- MACKAY, T. F. C., and R. F. LYMAN, 2005 *Drosophila* bristles and the nature of quantitative genetic variation. *Philos. Trans. R. Soc. B* **360**: 1513–1527.
- MARTÍN-CAMPOS, J. M., J. M. COMERÓN, N. MIYASHITA and M. AGUADÉ, 1992 Intraspecific and interspecific variation at the *y-ac-sc* region of *Drosophila simulans* and *Drosophila melanogaster*. *Genetics* **130**: 805–816.
- PALSSON, A., and G. GIBSON, 2004 Association between nucleotide variation in *Egfr* and wing shape in *Drosophila melanogaster*. *Genetics* **167**: 1187–1198.
- PALSSON, A., A. ROUSE, R. RILEY-BERGER, I. DWORKIN and G. GIBSON, 2004 Nucleotide variation in the *Egfr* locus of *Drosophila melanogaster*. *Genetics* **167**: 1199–1212.
- PALSSON, A., J. DODGSON, I. DWORKIN and G. GIBSON, 2005 Tests for the replication of an association between *Egfr* and natural variation in *Drosophila melanogaster* wing morphology. *BMC Genet.* **6**: 44.
- ROSAND, J., N. BAYLEY, N. ROST and P. I. W. DE BAKKER, 2006 Many hypotheses but no replication for the association between *PDE4D* and stroke. *Nat. Genet.* **38**: 1091–1092.
- ROSE, M. R., 1984 Laboratory evolution of postponed senescence in *Drosophila melanogaster*. *Evolution* **38**: 1004–1010.
- ROSE, M. R., and B. CHARLESWORTH, 1981 Genetics of life history in *Drosophila melanogaster*: I. Sib analysis of adult females. *Genetics* **97**: 173–186.
- ROBIN, C., R. F. LYMAN, A. D. LONG, C. H. LANGLEY and T. F. C. MACKAY, 2002 *hairy*: a quantitative trait locus for *Drosophila* sensory bristle number. *Genetics* **162**: 155–164.
- SIMPSON, P., 1990 Lateral inhibition and the development of the sensory bristles of the adult peripheral nervous system of *Drosophila*. *Development* **109**: 509–519.
- 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.
- STOGMANN, E., A. ZIMPRICH, C. BAUMGARTNER, A. GLEISS and F. ZIMPRICH, 2006 Lack of association between a GABA receptor 1 gene polymorphism and temporal lobe epilepsy. *Epilepsia* **47**: 437–439.
- THIRUMALAI, S., J. LAWRENCE, D. QUESTED, J. PIMM, N. BASS *et al.*, 2006 Failure to confirm genetic association between schizophrenia and markers on chromosome 1q23.3 in the region of the gene encoding the regulator of G-protein signaling 4 protein (RGS4). *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **141**: 296–300.
- YANG, H.-P., and S. V. NUZHIDIN, 2003 Fitness costs of *Doc* expression are insufficient to stabilize its copy number in *Drosophila melanogaster*. *Mol. Biol. Evol.* **20**: 800–804.

Communicating editor: G. GIBSON