DOI: 10.1534/genetics.104.028712

Effects of Population Structure and Sex on Association Between Serotonin Receptors and Drosophila Heart Rate

Naruo Nikoh,* April Duty[†] and Greg Gibson[†,1]

[†]*Department of Genetics, North Carolina State University, Raleigh, North Carolina 27695-7614 and *University of the Air, Chiba City 261-8586, Japan*

Manuscript received March 12, 2004
Accepted for publication July 19, 2004

ABSTRACT

As a first step toward population and quantitative genetic analysis of neurotransmitter receptors in *Drosophila melanogaster*, we describe the parameters of nucleotide variation in three serotonin receptors and their association with pupal heart rate. Thirteen kilobases of DNA including the complete coding regions of *5-HT1A*, *5-HT1B*, and *5-HT2* were sequenced in 216 highly inbred lines extracted from two North American populations in California and North Carolina. Nucleotide and amino acid polymorphism is in the normal range for Drosophila genes and proteins, and linkage disequilibrium decays rapidly such that haplotype blocks are typically only a few SNPs long. However, intron 1 of *5-HT1A* consists of two haplotypes that are at significantly different frequencies in the two populations. Neither this region of the gene nor any of the common amino acid polymorphisms in the three loci associate with either heart rate or heart rate variability. A cluster of SNPs in intron 2 of *5-HT1A*, including a triallelic site, do show a highly significant interaction between genotype, sex, and population. While it is likely that a combination of weak, complex selection pressures and population structure has helped shape variation in the serotonin receptors of Drosophila, much larger sampling strategies than are currently adopted in evolutionary genetics will be required to disentangle these effects.

MOLECULAR dissection of the genetic basis of complex multifactorial traits, such as cardiovascular physiology and behavior, presents numerous challenges (Glazier *et al.* 2002; Berry *et al.* 2003; Merikangas and Risch 2003). These include finding the genes that are responsible for small quantitative effects (Mackay 2001), establishing how particular alleles are distributed within and among populations (*e.g.*, Rockman *et al.* 2003), and working out the architecture of genotypic and environmental interactions. Linkage disequilibrium (LD) mapping has emerged as the most powerful approach to localizing genetic factors of small effect (Risch and Merikangas 1996; Jorde 2000), and in humans genome-wide association scans using thousands of SNP markers have been effective in localizing genetic effects to <1 cM (Cardon and Bell 2001). Such analysis depends on high levels of haplotype structure, while more fine-structure dissection requires comprehensive sampling of nucleotide variation in the region of interest (Johnson *et al.* 2001; Zondervan and Cardon 2004). The ability to identify specific single nucleotide polymorphisms (SNPs) within genes that contribute to phenotypic variation is also an essential component of efforts to understand the evolutionary forces that shape genomes.

In an organism like Drosophila that exhibits high levels of polymorphism with very little linkage disequilibrium, association studies must focus on candidate genes. To date, most progress in associating genotypic with phenotypic variation in Drosophila has been in relation to bristle number (Mackay 1995, 2001). Chromosome substitution into a common genetic background has been used to increase the relative contribution of single chromosomes, facilitating demonstration that several candidate loci with roles in neurogenesis appear to contribute to variation for abdominal and sternopleural bristle number (Lai *et al.* 1994; Long *et al.* 1998, 2000; Lyman *et al.* 1999; Robin *et al.* 2002). Most of these studies involved genotyping of just a subset of the most common SNPs and so approximated linkage disequilibrium mapping, but since haplotype blocks are only a few nucleotides long, every third or fourth SNP provides an independent test of association. Consequently, resolution may be to the level of individual quantitative trait nucleotides (QTN), but for a locus 10 kb long, several hundred tests of association need to be made for full coverage, and the gain in resolution is offset somewhat by reduced power due to multiple comparison issues (Long and Langley 1999). Most recently, complete sequencing has been used to identify QTN in *Dopa decarboxylase* for longevity (De Luca *et al.* 2003) and in *Egfr* for wing shape (Palsson and Gibson 2004) and cryptic variation for photoreceptor determination (Dworkin *et al.* 2003). Each of these studies discuss subtle nuances in experimental design and interpretation, and while they sug-

[1]*Corresponding author:* Department of Genetics, North Carolina State University, Raleigh, NC 27695-7614. E-mail: ggibson@unity.ncsu.edu

gest that sample sizes of a few hundred near isogenic lines are sufficient to detect QTN that account for several percent of the variation for a trait, much remains to be learned concerning the most appropriate experimental designs and sampling strategies.

Two central issues relate to the ability to replicate the associations in outbred populations (Dworkin *et al.* 2003, and in preparation; Genissel *et al.* 2004), and the effect of population stratification on test statistics (Pritchard *et al.* 2000; Thornsberry *et al.* 2001; Palsson and Gibson 2004). Although *Drosophila melanogaster* is often regarded as panmictic, there is increasing evidence for morphological and genetic differentiation between populations (Begun and Aquadro 1993; Andolfatto and Wall 2003; Caracristi and Schlötterer 2003; Glinka *et al.* 2003; Palsson and Gibson 2004). One of the objectives of the current study was to assess the effect of differences in SNP allele frequency between two populations from North Carolina and California on association mapping in Drosophila. By sampling of the order of 100 alleles from each population, we also have the opportunity to evaluate simultaneously population and quantitative genetic approaches to inference of potential functional differentiation of alleles.

Considerable attention has been given to the relationship between variation in the serotonin and dopamine axes and psychological variation in humans (Berry *et al.* 2003). Serotonin is a biogenic amine, which serves a neurotransmitter function in synaptic transmission in animal phyla (Hay-Schmidt 2000). It also plays an important role in the regulation of locomotion (Segalat *et al.* 1995; Lucki 1998) and affects physiological functions such as heart rate (Johnson *et al.* 1997; Zornik *et al.* 1999) and other targets of the sympathetic nervous system in mammals as well as in Drosophila (Monastirioti 1999; Blenau and Baumann 2001). Both serotonin receptors and transporters have been subject to extensive association studies in humans. Particular attention has focused on a handful of marker SNPs, notably T102C in *5-HT2A* and G861C in *5-HT1B* (Arranz *et al.* 1995; Williams *et al.* 1996; Lappalainen *et al.* 1998), although more extensive sequence surveys are beginning to reveal that these sites are in strong linkage disequilibrium with other sites, including promoter polymorphisms that affect transcription (Sanders *et al.* 2001; Duan *et al.* 2003). With this in mind, and recognizing the conserved function of serotonin in neurotransmission in vertebrates and invertebrates, we have undertaken a reverse quantitative genetic study in flies exploring the function of polymorphism in three *D. melanogaster* serotonin receptors (*5-HT1A*, *5-HT1B*, and *5-HT2*) and the candidate traits of pupal heart rate and physiological response to drug ingestion (N. Nikoh, R. Carrillo and G. Gibson, unpublished results).

The dorsal vessel of Drosophila is homologous to the vertebrate heart and pumps hemolymph rhythmically (Rizki 1978; Bodmer and Venkatesh 1998; Curtis *et al.* 1999). Aspects of both developmental specification and physiological function are highly conserved between mammals and invertebrates (Cripps and Olson 2002). Transfusion of neurotransmitters and neural peptides into pupae, larvae, or embryos has shown that regulation of the heart-beat length varies according to developmental stage and that serotonin is a major cardiac modulator (Johnson *et al.* 1997; Zornik *et al.* 1999). Genetic approaches have also been applied to dissect Drosophila heart rate and have revealed considerable genetic variation for heart rate in prepupae (Robbins *et al.* 1999). A genome-wide deficiency screen implicated the *5-HT1A* and *5-HT1B* loci as modulators of prepupal heart rate, since hemizygosity for this region of the genome reduced beat length (Ashton *et al.* 2001). Here we report the sequence of a total of 13 kb of the three serotonin receptors in 200 nearly isogenic lines from two natural populations and test for association with measures of prepupal heart rate and heart rate variability. Rather than focusing on a single major-effect SNP, the data reported here suggest that multiple SNPs throughout the locus may have phenotypic consequences.

MATERIALS AND METHODS

Fly stocks: *D. melanogaster* isofemales were collected in July 2000 from West End, North Carolina, and used to establish 130 nearly isogenic lines by between 15 and 20 generations of sib pair-mating (Palsson and Gibson 2004). A total of 85 nearly isogenic lines from a California population were kindly provided by S. Nuzhdin. These lines were collected in 1998 from Wolfskill, California and inbred by nearly 50 generations of sib pair mating (Yang and Nuzhdin 2003). While in the lab, all lines were maintained in vials containing standard cornmeal media with a few granules of live yeast and were kept in an incubator at 25° on a 12-hr light:dark cycle.

Sequencing: Fly DNA was extracted from a single male of each near-isogenic line. The fly was homogenized in 50 μl squishing buffer (10 mm Tris pH 8.2; 1 mm EDTA, 25 mm NaCl, 200 mg/ml proteinase K), incubated at 37° for 30 min and then at 95° for 2 min, and chilled for storage at −20°. Amplification of specific fragments was performed by 35 cycles of PCR, each 30 sec at 94°, 30 sec at 50°, and 3 min at 72°, in 1.5 mm $MgCl_2$. Amplified fragments were extracted from 0.8% agarose gels using QIAquick columns according to manufacturer's instructions (QIAGEN, Chatsworth, CA), and directly sequenced using Big Dye II dye-terminator sequencing (PE Biosystems, Boston) on an ABI 3700 with average read lengths in the range of 700–750 nucleotides. All oligonucleotide primers for PCR and sequencing are listed in supplementary Table 1 at http://www.genetics.org/supplemental/. Sequence data from North Carolina were checked and aligned with Contig Express in Vector NTI (InforMax, Frederick, MD). Chromatograms of sequence data were checked for polymorphic sites and manually edited where necessary. Sequence data from California were aligned using Phred (Ewing *et al.* 1998), Phrap (http://www.phrap.org), and BioLign (http://www.maizegenetics.net) programs. All polymorphic sites lower than phred score 20 were confirmed by eye. Singleton substitutions were confirmed by resequencing, and regions of heterozygosity were omitted from the analysis. GenBank accession numbers for all of the sequences are AY564239–AY564734 and AY564785–AY566107.

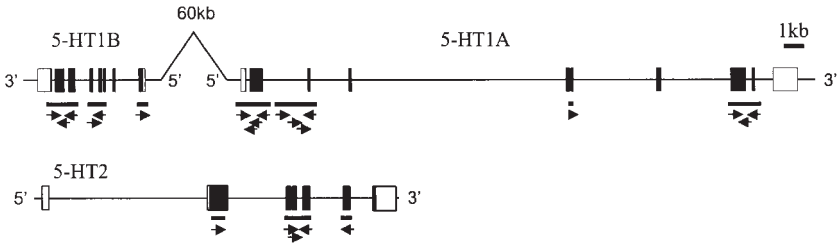

FIGURE 1.—Genomic structure of *5-HT1A*, *5-HT1B*, and *5-HT2*. The extent of the sequenced region in each of the three genes is shown as a solid bar underneath the gene structures. Exons are depicted by squares (solid coding regions) and flanking regions and introns as thin lines. Arrows indicate sequence reads in each contig. In this figure, the centromere is to the left of *5-HT1B*, which is transcribed from right to left, divergently from *5-HT1A*. The two genes are separated by 60 kb and several genes. An alternative promoter for *5-HT1A* 20 kb upstream of the indicated promoter (not shown) appeared in the Version 3 release of the Drosophila genome annotation.

Sequence analysis: Sequence alignment was performed manually using GeneDoc (NICHOLAS *et al.* 1997). For association tests, F_{ST}, LD, and sliding-window analyses, only biallelic sites with common polymorphisms (defined as the more rare allele frequency >5%) were used. Sites that contain indel substitution were also omitted from our analysis because of complex mutational patterns. Parameters of population genetic structure were computed using an analysis of variance framework (WEIR and COCKERHAM 1984; EXCOFFIER *et al.* 1992; WEIR 1996). Fixation indices for each polymorphic site were separately calculated using Arlequin (SCHNEIDER *et al.* 2000), and significance was assessed using either 1000 or 100,000 permutations. Linkage disequilibrium between pairs of biallelic sites was calculated with TASSEL (http://www.maizegenetics.net) or our own code, and significance was estimated by Fisher's exact test with the standard Bonferroni correction for multiple tests. Sliding-window analysis of the number of haplotypes was conducted on the common polymorphic sites, with a window length of 300 bp and a step size of 20 bp. Energy minimization was used to predict the secondary structures of the two *5-HT1A* intron 1 RNAs, using the Vienna RNA secondary structure server (HOFACKER 2003).

Phenotyping heart rate and heart rate variability: Heart rates of 129 isogenic lines from North Carolina and 85 isogenic lines from California were assayed as described (ROBBINS *et al.* 1999; ASHTON *et al.* 2001) by digital electronic capture of the fluctuation in transmitted light intensity due to movement of the dorsal vessel of a white prepupa placed on a microscope stage. The fluctuation in light intensity was recorded for 1–3 min. The flies were maintained in an incubator at 25° and were assayed at this temperature (±2°) between 2:00 and 4:00 PM. Pupal heart rates were calculated as the inverse of the mean length of 20 consecutive beats, averaged over three sets of nonoverlapping beats. The index of heart rate variability was calculated as the coefficient of variation for 10 sets of the inverse of the mean length of 5 consecutive beats. To normalize the distribution of the coefficient of variation, we added 1 to the coefficient of variation for each individual and took the base 10 logarithm. The sex of 90% of the prepupae was determined after eclosion of the adult in a microcentrifuge tube. In general, 20 prepupae per North Carolina line and 10 per California line were scored, although a few lines with very low fecundity were represented by fewer individuals. These did not show a trend toward extreme heart rates.

Statistical analysis: The significance of differences in trait values between sexes and between populations was performed using PROC TTEST or PROC NPAR1WAY in SAS Version 8.2 (SAS Institute, Cary, NC). Quantitative genetic parameters were assessed using JMP Version 5 (SAS Institute) or PROC MIXED in SAS. Lines were assumed to be nearly isogenic for the estimation of heritability (FALCONER and MACKAY 1996). Line means for the two heart-beat traits were estimated by analysis of variance using PROC MIXED in SAS based on measurements of each individual fly. This ensured that among-fly variability was the source of residual error. Two-way factorial analyses of variance of heart rate and heart rate variability were conducted with sex as a fixed effect and line as a random effect. Genetic correlations between the traits were computed following ROBERTSON (1959) as the difference between the line and trait mean squares divided by the sum of the line and trait mean square minus twice the error mean square. Similarly, the genetic correlations between the sexes were computed with mean squares from analysis of variance with sex substituting for trait.

Associations between common polymorphic sites in serotonin receptors and heart rate phenotypes were assessed using PROC MIXED with mixed models of the form:

$$\text{Heart beat trait} = \mu + G + S + P + G \times S + G \times P + S \times P + G \times S \times P + L + \varepsilon,$$

where SNP genotype (G), sex (S), and population (P) denote fixed effects, and line (L) denotes the random effect of line. In this case, heart-beat traits were the least square mean for each sex and line within a population. Fitting line as a random effect controls for pseudoreplication between the sexes, which are highly correlated. Omission of this term greatly inflates *P*-values, as confirmed by permutation tests keeping sexes within lines and lines within sexes constant (see supplementary Table 3 at http://www.genetics.org/supplemental/). Only common polymorphic sites were tested, since rare alleles do not provide a large enough sample size to detect effects.

RESULTS

Sequence variation in the serotonin receptors of *D. melanogaster*: We sequenced 13 kb of three of the four known serotonin receptors from 216 highly inbred lines. Two of the genes, *5-HT1A* and *5-HT1B*, are located within 60 kb of one another at 56B1-56B5 on chromosome 2R, while the third, *5-HT2*, is located near the centromere of chromosome 3 at 82C5. On the basis of a preliminary survey of variation in cDNA clones from a dozen lines (W. JIN and G. GIBSON, unpublished data), we focused on the regions diagrammed in Figure 1, excluding monomorphic exons and the 3′ UTR (exon 10) of *5-HT1A*. No attempt was made to include all of the regulatory regions associated with each gene, which

TABLE 1

Summary of nucleotide variation in the serotonin receptors

Gene	Population[a]	No. of lines	Length (bp)	Success rate[b]	No. of SNPs	No. of indels	No. of 3 or 4 variants[c]	No. of haplotypes	Tajima's *D*
5-HT1A	NC	120	6414	96.8	275	28	8	98	−0.157
	CA	82	90.3	238	26	7	72	0.026	
	Total	202	94.2	321	35	9	167	−0.344	
5-HT1B	NC	121	3454	96.6	135	13	1	84	−0.176
	CA	80		91.0	105	12	0	57	0.471
	Total	201	94.4	142	16	1	130	−0.137	
5-HT2	NC	95	3011	84.3	33	5	0	28	−0.617
	CA	83	77.0	26	2	0	22	−0.681	
	Total	178	80.8	44	5	0	42	−1.150	

[a] North Carolina (NC), California (CA), or combined data set (Total).
[b] The percentage of sequence coveraged, reduced for 5-HT2 due to higher heterozygosity.
[c] SNPs segregating three or four different nucleotides.

in any case have not yet been defined experimentally. In total, 6.5 of the 30 kb of *5-HT1A* was covered in four contiguous fragments, 3.5 of the 6 kb of 5-HT1B in three contigs, and 3.0 of the 20 kb of *5-HT2* in three contigs.

Heterozygosity was three times greater in the *5-HT2* than in the *5-HT1* region (~31 *vs.* 11%) and 50% greater in the North Carolina than in the California lines. Both results probably relate to the effect of linked deleterious sites on the rate of fixation due to inbreeding (Caballero and Keightley 1998): the California lines were sib mated for three times as many generations as were the North Carolina ones, and beacause the *5-HT2* gene is proximal to the centromere, recombination is reduced relative to *5-HT1*. Final sequence coverage was 90 and 97% for 82 California and 121 North Carolina *5-HT1* chromosomes, respectively, and between 85 and 90% for 83 California and 95 North Carolina *5-HT2* chromosomes. Due to the expense of obtaining in excess of 2 Mb of sequence, no attempt was made to sequence each allele in both directions, so estimates of sequence diversity must be considered with this caveat in mind. However, as for our similar study of *Egfr* (Palsson and Gibson 2004), >10% sequence overlap was obtained from different reads, from which the error rate in detection of polymorphic sites can be estimated as <0.1%.

Parameters of nucleotide variation are summarized in Table 1. In total, 463 SNPs and 51 indels were detected in <10 kb of the *5-HT1* region, corresponding to a polymorphism every 20 bp. Correspondingly, several sites segregate three or even four nucleotides, and most alleles represent novel haplotypes across the full extent of the locus. Sliding-window analysis of nucleotide polymorphism and haplotype number in Figure 2 demonstrates a tendency toward reduced nucleotide diversity in exons. A full breakdown by exon is available online in supplemental Table 2 at http://www.genetics.org/supplemental/. The ratio of synonymous to replacement substitutions was 78 to 48, and while only 12 of the replacements are in *5-HT1B*, comparison with the *D. simulans* allele by the HKA test (Hudson *et al.* 1987) did not indicate evidence for positive selection in the region. However, Figure 3 clearly shows the constraining effect of purifying selection on the serotonin

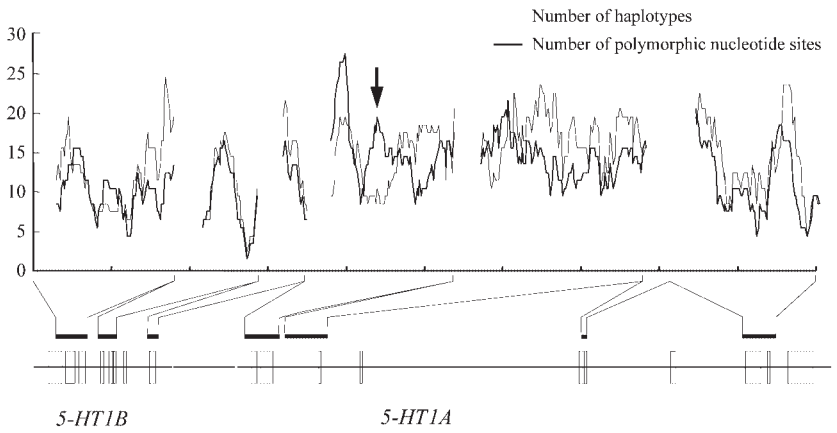

Figure 2.—Sliding-window analysis of sequence and haplotype diversity. The number of polymorphic nucleotide sites and the number of haplotypes in a sliding window of 300 bp, at 20-bp intervals, are plotted along the sequenced region. Structural features of each gene are indicated along the bottom of each panel for *5-HT1B* and *5-HT1A*. The arrow highlights the location of a haplotype block with significant population structure.

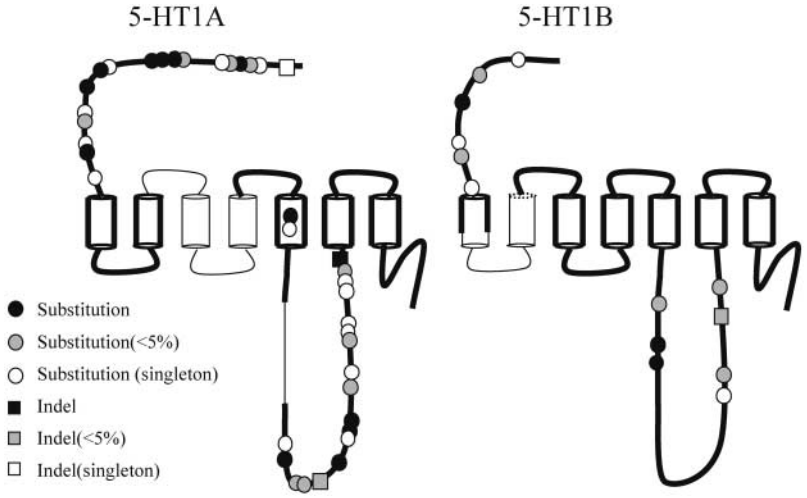

FIGURE 3.—Constraint on the location of amino acid polymorphisms. The location of singleton (open), rare (shaded), and common (solid) amino acid variants is shown relative to their location on the seven-transmembrane structure of the serotonin receptors for *5-HT1A* and *5-HT1B*. Nucleotide and indel polymorphisms are shown as circles and squares, respectively. Note that polymorphism is concentrated in the extracellular N-terminal domain and intracellular loop, but that a few polymorphisms are detected in the membrane-spanning regions. Sequences that were not sequenced are shown as dotted lines.

receptors, as amino acid polymorphism in *D. melanogaster* is largely excluded from the transmembrane domains and the carboxy-terminal tail. As expected for a centromere-proximal locus, sequence diversity was much reduced in *5-HT2*, and all subsequent analyses focus on the *5-HT1* region.

Population structure and linkage disequilibrium in the *5-HT1* region: Two types of data point to significant population structure between the two North American samples. Both populations harbor a large number of rare private alleles. The difference in sign of Tajima's *D* for each locus separately suggests a slight excess of singletons and doubletons in the North Carolina population, although this is not significant. The more compelling observation is that there is a stretch of 500 bp in intron 1 of *5-HT1A* that shows highly significant divergence in allele frequency between North Carolina and California. This region stands out in Figure 4, a sliding-window plot of F_{ST} computed by AMOVA (EXCOFFIER *et al.* 1992), with a peak value of 0.2, which is significant at $P < 0.0003$. Interestingly, haplotype diversity is relatively low in this intron (Figure 2, arrow), consistent with recent selection driving differentiation between the two populations. Throughout the remainder of the loci, allele frequencies of common SNPs are very similar, and F_{ST} is not significantly greater than zero.

Plots of the significance of linkage disequilibrium in Figure 5 demonstrate that haplotype structure is very similar between the two populations (NC on the left, CA on the right). Typically only sites within 200 bp of one another show strong LD, as observed for other loci in regions of normal recombination (SCHAEFFER and MILLER 1993; LANGLEY *et al.* 2000; AQUADRO *et al.* 2001; PALSSON and GIBSON 2004). However, a large 1.5-kb block of LD centered on intron 1 and including most of exon 2 stands out in both populations and actually extends to two noncontiguous regions in intron 2 and exon 3. Decay of LD is thus not monotonic as generally expected, and fairly sharp breaks between haplotype blocks may suggest unequal distribution of recombination and/or gene conversion breakpoints within the locus. The higher significance of LD in North Carolina can be attributed to the larger sample size. There is some suggestion of long-range LD between regions of *5-HT1A* and *5-HT1B*, particularly involving a second LD block including exons 8 and 9 of *5-HT1A*, but it is unclear whether this is significant after correction for multiple comparisons.

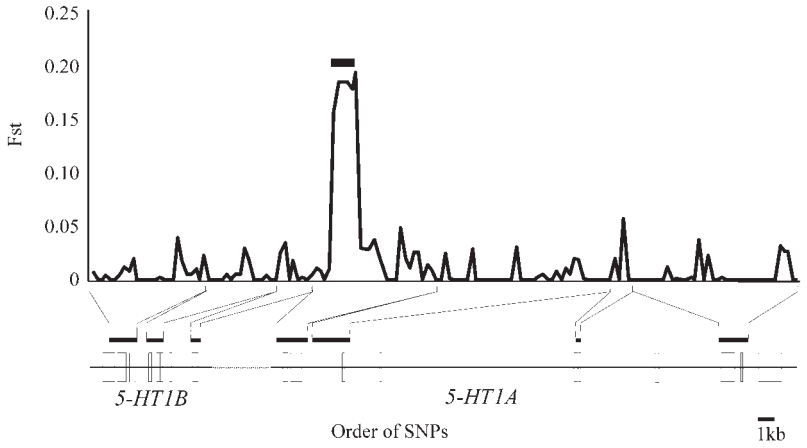

FIGURE 4.—Population differentiation in intron 1 of 5-HT1A. Fixation indices (F_{ST}) estimated by AMOVA between North Carolina and California populations are plotted according to position in the *5-HT1B*–*5-HT1A* gene region. Open boxes on the line at the bottom of the figure represent exons. Note that the position along the gene has been made artificially contiguous in the plot by telescoping each block of sequence (represented as solid bars above the gene structure diagram) as indicated. A solid bar above the plot highlights the region, corresponding to intron 1 of *5HT1A*, in which F_{ST} departs significantly from the empirical null model (testwise $P < 0.0003$; experimentwise $P < 0.05$ after Bonferroni correction).

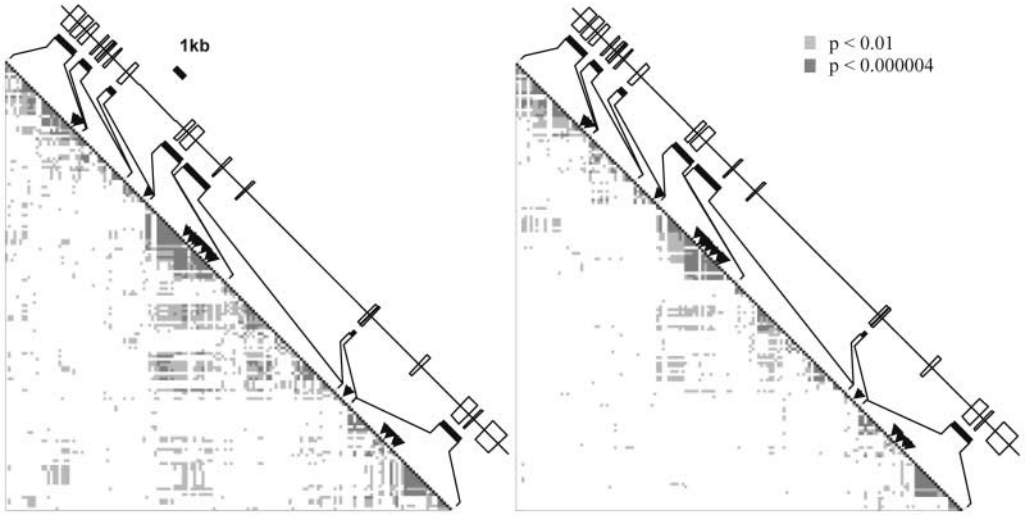

FIGURE 5.—Linkage disequilibrium in the 5-HT1B–5-HT1A gene region. Significance of linkage disequilibrium between pairs of common polymorphic sites (less common allele frequency >0.05) is plotted separately for North Carolina (left) and California (right) populations. Significance was assessed using Fisher's exact test, with $P < 0.01$ denoted by shading, and $P < 0.000004$ (the experimentwise threshold after Bonferroni correction) as solid. Arrowheads indicate amino acid replacement sites.

Since intron 1 of *5-HT1A* stands out as unusual in both the population structure and LD analyses, we examined the structure of the region by phylogenetic network analysis (Figure 6). Intron 1 ranges in length from 202 to 230 bases and includes six common SNPs and one indel that define two major haplotypes observed in 97% of the alleles. Only six recombinant alleles defining three novel haplotypes were observed in the sample of 192 alleles. The high F_{ST} is due to the low frequency of one of the haplotypes (0.21) in California, whereas it is equally prevalent in North Carolina (0.49). Since a *D. simulans* allele shows elements of both haplotypes, it is not possible to polarize the direction of selection in the region in the two populations. Prediction of the free energy of duplex formation suggests that the two haplotypes are likely to fold into remarkably different secondary structures (Figure 6B), which could influence splicing efficiency.

Association between heart rate and nucleotide variation: Cardiograms were recorded from 3427 white prepupae by digital capture of the fluctuation in light intensity with each beat. Heart rate was estimated in beats per second from 60 beats by averaging the time required for three sets of 20 contiguous beats, as described in MATERIALS AND METHODS. Since the Drosophila dorsal vessel periodically skips beats or even stops for several seconds, it is rarely possible to measure 60 consecutive beats. Although there is variation in the number and severity of these stoppages, it is difficult to quantify, so we chose a more robust measure of arrhythmia, namely heart rate variability, the logarithm of the coefficient of variation of the heart rate in five sets of 10 beats. Both

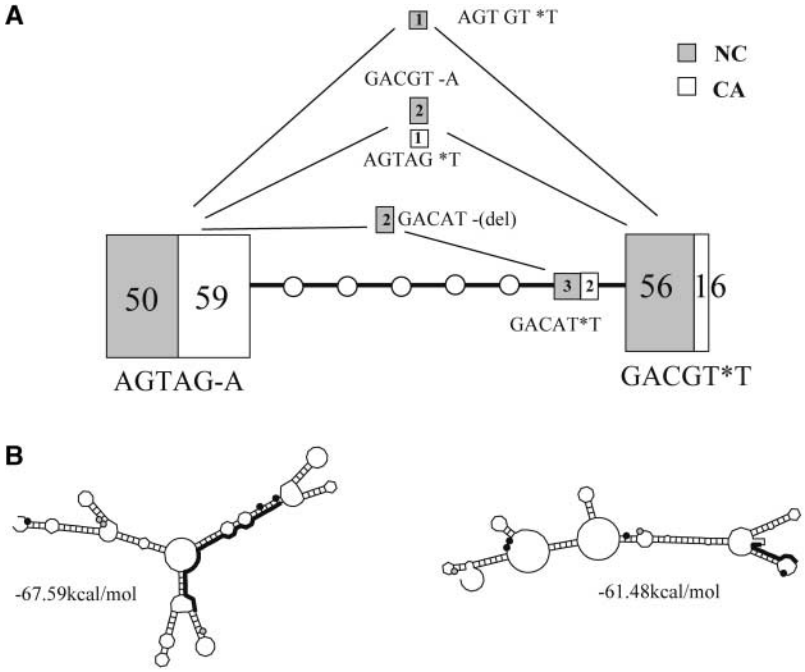

FIGURE 6.—Structure and distribution of the *5-HT1A* intron 1 haplotype. (A) The haplotype structure involving six SNPs and one indel in intron 1 of *5-HT1A* is indicated with numbers inside boxes denoting the number of lines with each haplotype. Circles represent expected ancestral haplotypes that were not observed, and segmented lines represent inferred recombination events. Shaded boxes indicate frequency in the North Carolina population, and open boxes the frequency in the California population. (B) Secondary structure of the two most common haplotypes, predicted on the basis of the free energy of association (HOFACKER 2003). Nucleotide and indel polymorphisms are shown as circles and bars, respectively, with derived alleles in solid and sites shared in *D. simulans* in shading.

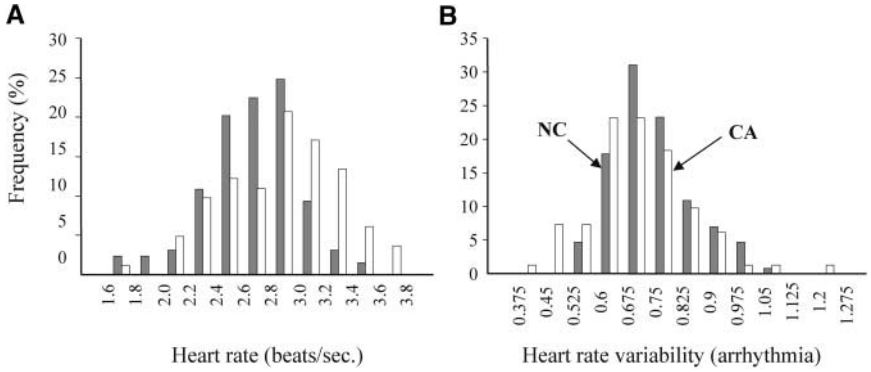

FIGURE 7.—Distribution of heart rate and heart rate variability phenotypes. Histograms show the distribution of heart rate (A) and heart rate variability (B) between the North Carolina (solid bars) and California (open bars) populations.

heart rate, which is slightly higher in females than in males, and heart rate variability, are approximately normally distributed, and are significantly different in the North Carolina and California populations (Figure 7). Significant negative phenotypic and genetic correlation between the two traits documented in Table 2 suggests that variability is greater for more rapidly beating hearts.

The simplest test of association is a *t*-test of the significance of the difference in phenotype between each SNP, ignoring effects of sex and population. Figure 8, A and

TABLE 2

Estimates of quantitative genetic parameters for heart beat phenotypes

	Total		NC		CA	
Parameter	HR	HRV	HR	HRV	HR	HRV
Pooled						
M	2.7225	0.7455	2.6733	0.7549	2.8791	0.7157
V_G	0.0667	0.0108	0.0575	0.0047	0.0862	0.0093
CV_G	9.4870	13.9461	8.9700	9.0667	10.1992	13.5078
V_E	0.1205	0.0464	0.1240	0.0470	0.1076	0.0438
CV_E	12.7481	28.8871	13.1725	28.7103	11.3931	29.2393
V_P	0.1872	0.0572	0.1815	0.0517	0.1938	0.0531
h^2	0.3564	0.1890	0.3168	0.0907	0.4449	0.1759
r_{GT}	−0.5540		−0.5422		−0.5427	
r_{PT}	−0.4683		-0.4846		−0.4039	
r_{GS}	0.9571	0.8846	0.9791	1.0477	0.8942	0.5504
r_{PS}	0.8383	0.5657	0.8636	0.6688	0.7811	0.4515
Female						
M	2.7478	0.7467	2.6912	0.7600	2.9305	0.7048
V_G	0.0638	0.0054	0.0557	0.0038	0.0795	0.0102
CV_G	9.1931	9.8519	8.7664	8.1185	9.6193	14.3017
V_E	0.1167	0.0452	0.1191	0.0463	0.1080	0.0404
CV_E	12.4330	28.4775	12.8221	28.3187	11.2132	28.5261
V_P	0.1805	0.0506	0.1747	0.0501	0.1874	0.0506
h^2	0.3535	0.1069	0.3185	0.0759	0.4239	0.2009
Male						
M	2.6986	0.7367	2.6504	0.7426	2.8334	0.7204
V_G	0.0628	0.0052	0.0572	0.0051	0.0738	0.0050
CV_G	9.2869	9.7658	9.0207	9.6578	9.5864	9.7912
V_E	0.1248	0.0480	0.1303	0.0481	0.1076	0.0478
CV_E	13.0911	29.7352	13.6206	29.5224	11.5757	30.3416
V_P	0.1876	0.0532	0.1875	0.0532	0.1814	0.0528
h_2	0.3348	0.0974	0.3049	0.0967	0.4068	0.0943

The parameter estimates are for the mean (M); genotypic, environmental, and phenotypic variance (V_G, V_E, V_P, respectively); coefficients of genotypic (CV_G) and environmental (CV_E) variance; heritability (h^2); genetic and phenotypic correlation between the traits (r_{GT}, r_{PT}, respectively); and genetic and phenotypic correlation between the sexes (r_{GS}, r_{PS}, respectively). HR and HRV refer to heart rate and heart rate variability traits as defined in MATERIALS AND METHODS.

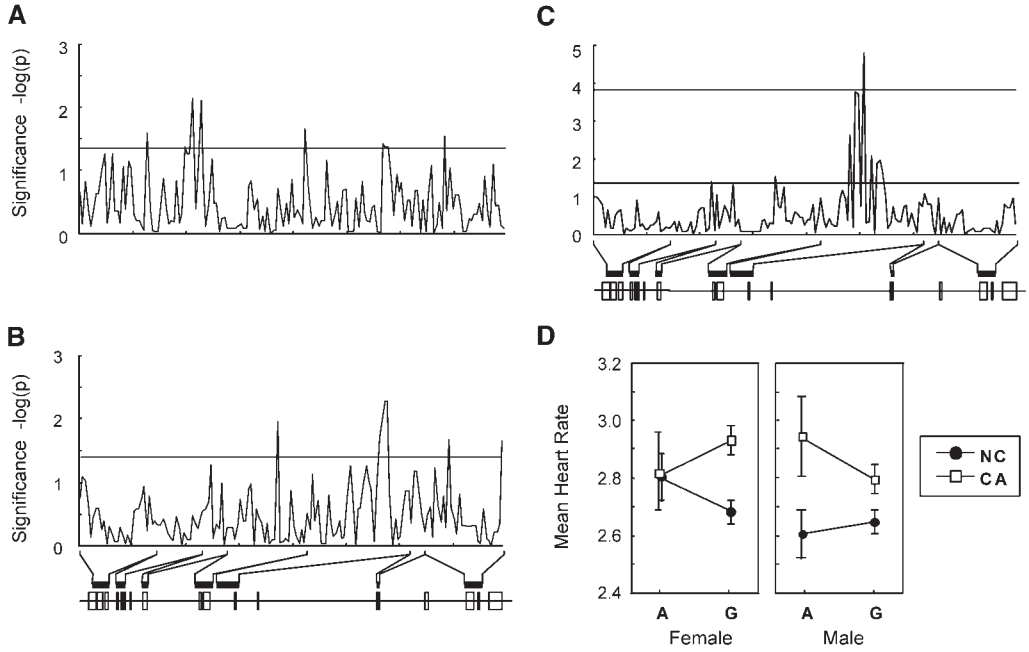

FIGURE 8.—Association between nucleotide polymorphisms and heart rate phenotypes. Significance as the negative logarithm of the *P*-value is plotted against location in each gene for (A) heart rate and (B) heart rate variability according to a simple one-way ANOVA and (C) the genotype × sex × population interaction for heart rate according to a mixed-model analysis of variance. The nature of the significant three-way interaction for site G221856A is plotted in D. See supplementary Figure 1 at http://www.genetics.org/supplemental/ for a complete set of plots for each of the terms in the mixed-model analysis.

B, plot the negative logarithm of the *P*-value for heart rate (HR) and heart rate variability (HRV), respectively, for each of the 159 common variants along the locus from centromere to telomere (namely from the 3′ end of *5-HT1B* to the 3′ end of *5-HT1A*, since the genes are transcribed divergently). Allowing for strong linkage disequilibrium between adjacent sites, five SNPs show nominal association at $P < 0.05$ for HR and seven for HRV. These numbers are expected by chance. Bonferroni correction for the multiple comparisons and both traits requires a significance threshold of 0.000157 (0.05/318), but none of the sites approach this value. The cluster of SNPs adjacent to exon 3 that have the strongest evidence for association with HRV also associate weakly with HR, a result that is not expected by chance, but the permutation test significance for the summed likelihood of two-trait associations, given the genetic correlation between the traits, is only suggestively significant ($P = 0.006$), given the number of comparisons.

To account for the fixed effects of sex and population differences, we fit a mixed model with line as a random effect including each of the interaction terms among genotype, sex, and population (DWORKIN *et al.* 2003; PALSSON and GIBSON 2004). Neither of the two-way interactions involving genotype resulted in any *P*-values <0.01, but Figure 8C shows a cluster of sites in intron 2 of *5-HT1A* that show highly significant association as a function of both sex and population. One of these sites (G221856A in AE003797) plotted in Figure 8D exceeds the Bonferroni threshold, formally suggesting that the G allele has opposite effects on heart rate in females in North Carolina and California, whereas it is the A allele that differentiates the populations in males. Such a complex interaction implies sex-dependent epistatic interaction between the SNP and an unknown site or sites that are at a different frequency in the two populations, but might also simply be a false-positive result. Intriguingly, several other linked sites in intron 2 and exon 3 of *5-HT1A* have three common segregating variants, one of which also show signs of a three-way interaction and has allele frequencies of 0.66, 0.20, and 0.14 (site 221719). This is a very uncommon situation and is at least consistent with the idea that weak but complex selection operates on this region of the gene.

It is also notable, however, that no evidence for association with either heart rate trait was observed either with the sites in intron 1 that show population stratification or with any of the common amino acid polymorphisms, as documented in Table 3.

DISCUSSION

This investigation of association between nucleotide polymorphisms and cardiophysiological variation is part of a larger effort to evaluate the potential for mapping quantitative genetic variation to the nucleotide level in Drosophila. Although successful linkage disequilibrium mapping has been reported for bristle number (LONG *et al.* 1998 and 2000), insecticide resistance (DABORN *et al.* 2002), and immunity (LAZZARO *et al.* 2004), replication of association in natural populations is not straightforward (GENISSEL *et al.* 2004), and it seems that complete sequencing may be necessary to identify the actual site(s) that cause phenotypic differences (DWORKIN *et al.* 2003). An emerging feature of these studies is that extra insight may be gained through the combination of population and quantitative genetic inferences from a single experiment (PALSSON *et al.* 2004; PALSSON and GIBSON 2004), so it is interersting to note that these two types of signature in this study point to different gene

TABLE 3

Association statistics for common amino acid polymorphic sites

			Heart rate				Heart rate variability				
Site	Amino acid	*t*-test	*G*	*G* × *S*	*G* × *P*	*G* × *S* × *P*	*t*-test	*G*	*G* × *S*	*G* × *P*	*G* × *S* × *P*
5-HT1B											
C153318T	G S	0.20	0.21	0.12	0.69	0.27	0.58	0.85	0.19	0.67	0.27
C153320T	R H	0.11	0.04	0.00	0.33	0.34	0.75	0.71	0.23	0.59	0.75
C157121T	V I	0.47	0.31	0.05	0.13	0.51	0.71	0.73	0.46	*1.40*	0.46
5-HT1A											
C219249A	D E	0.07	0.21	0.21	0.30	0.08	0.39	0.58	0.28	0.03	0.12
T219352A	S T	0.83	1.22	0.96	0.04	0.41	0.97	1.31	0.33	0.36	0.91
C219368A	P Q	0.36	0.07	0.08	0.81	0.18	0.01	0.20	0.84	0.60	0.97
T219383C	V A	0.55	0.51	0.58	0.46	0.31	0.57	0.79	0.03	0.38	0.22
T219532A	S T	0.28	0.39	0.85	0.24	*1.53*	0.37	0.57	0.07	0.73	*1.58*
C219608T	A V	0.41	0.71	1.27	0.19	0.73	0.35	0.59	0.19	0.64	1.07
C219690A	Q H	0.00	0.11	0.42	0.11	1.28	0.05	0.30	0.29	0.97	0.27
C237412A	L M	0.67	0.40	0.15	0.18	0.56	0.10	0.32	0.00	0.03	0.02
A246559G	T A	0.00	0.08	0.05	0.29	0.05	0.14	0.10	0.07	0.11	1.05
T246665A	V E	0.50	0.18	0.64	0.62	0.06	0.64	0.43	0.12	0.72	0.19
C246716T	T M	*1.54*	1.15	0.68	0.41	0.55	0.78	0.58	0.57	0.13	0.13
C246719G	T R	0.34	0.55	0.59	0.63	0.56	0.29	0.20	0.11	0.34	0.33

Each value represents the negative logarithm of the *P*-value (NLP) associated with either a simple *t*-test (disregarding sex and population) or the *F*-ratios for genotype, and the genotype interaction terms with sex and population. Only four tests (indicated by italics) exceed the testwise threshold of $P < 0.05$ (NLP > 1.3), but none of these are significant experimentwise. Site numbers correspond to GenBank scaffold file AE003797.

regions of the *5-HT1A* gene. In particular, two pieces of evidence relate to the recent operation of natural selection on the loci, and two different aspects of the data are suggestive of genotype-phenotype association.

Selection on the serotonin receptors of *D. melanogaster*: The most conspicuous feature of the molecular evolution of the 5-HT receptor loci is the nonrandom distribution of amino acid polymorphisms relative to structural features of the predicted proteins, which implies the operation of purifying selection. The C-terminal intracellular domain and the membrane-spanning helices are relatively devoid of replacement substitutions, whereas the N-terminal extracellular domain and cytoplasmic loop are quite polymorphic. Compared with another membrane receptor protein that we have also sequenced in the same lines, the tyrosine kinase EGFR, which harbors just a single common amino acid variant, 5-HT1A and 5-HT1B are more variable than might be expected of pleiotropic proteins with essential physiological roles. However, the ratio of replacement to synonymous substitutions is in the normal range for Drosophila proteins, and there is no statistical evidence for departure from assumptions of neutral theory for the loci as a whole.

On the other hand, there is a clear footprint of divergence between the North Carolina and California populations in intron 1 of *5-HT1A* for a haplotype that consists of seven polymorphisms. Circumstantial evidence suggests that the alignment of these sites into two haplotypes could affect splicing efficiency (Kirby *et al.* 1995; Chen and Stephan 2003). Strong positive selection such as that observed at the cytochrome P450 *Cyp6g1* locus, which is attributed to the emergence of insecticide resistance, leaves a footprint covering tens of kilobases (Daborn *et al.* 2002) that is unlike the highly restricted nature of the differentiated haplotype detected here. The drop-off in both linkage disequilibrium and population stratification occurs sharply within a hundred bases while the haplotype itself is <500 bases. This pattern is also inconsistent with admixture, and most likely reflects weak selection, with the sharp boundaries of the haplotype maintained either as a result of unequal recombination probabilities along the gene (Li and Stephens 2003; Wall and Pritchard 2003) or as selection in favor of the integrity of the entire haplotype (Kelly and Wade 2000; but see discussion in Baines *et al.* 2002; Dworkin *et al.* 2003). Differential selection pressures in the two populations could most simply reflect some environmental difference or an epistatic interaction between the haplotype and some other site that is at a different frequency in North Carolina and California. Molecular and physiological surveys along geographic clines could conceivably resolve this issue, although it is quite possible that the heart rate phenotype is simply a correlated response with another, unknown, phenotype that is the actual target of selection.

Association between the serotonin receptors and heart rate phenotypes: The most direct evidence for association is between a cluster of sites adjacent to exon 3

and both heart rate and heart rate variability. Neither association alone exceeds the standard multiple comparison threshold, but weak association with both traits was detected. However, permutation testing suggests this might be due to the genetic correlation between the two traits. The stronger association was actually observed with HRV, which can be regarded as a measure of arrhythmia. Physiologically, it is not unreasonable to suppose that a serotonin receptor would actually regulate the consistency of heart rate over several beats, rather than the basal rate itself, but we have no idea over what number of beats such a homeostasis mechanism would operate. More fundamentally, it is not known what normal physiological role serotonin plays in regulating the Drosophila dorsal vessel, which may have an intrinsic rate of contraction set by molecules such as the potassium channels, but is also innervated by cardioactive neuropeptide-secreting cells (Dulcis and Levine 2003). In the absence of any knowledge of the molecular effect of the SNPs (222824, 222844, and 222853 in AE003797), the association should be regarded as suggestive at best and may simply represent a false positive.

The other evidence for association was specifically with heart rate, but involves a highly complex interaction among genotype, sex, and population for one SNP that exceeds the conservative experimentwise threshold and several linked triallelic SNPs. These sites do not form a haplotype, and several other SNPs in the region show similar levels of linkage disequilibrium but do not associate with heart rate, so it is likely that variation in the region is shaped by effects of two or more SNPs. No evidence for synergistic or antagonistic interactions was observed, but with a sample of just 200 chromosomes power to detect epistasis is low. Again, the heart rate association might be a correlated response with a more biologically relevant behavioral or physiological effect. It will be difficult to tease apart proximate causes, but our analysis does suggest that subtle and complex interactions contribute to the genetics of synaptic transmission in flies.

By contrast, we have not been able to find evidence that amino acid polymorphisms in the serotonin receptors have a quantitative effect on heart rate, a trait that is known to be affected by serotonin. Similarly, preliminary analysis of two potassium channel genes failed to provide evidence for association with heart rate (N. Nikoh, A. Duty and G. Gibson, unpublished results), so this is a situation in which common polymorphisms in five of the most likely candidate proteins do not account for more than a few percent of the trait variance. No attempt has been made here to comprehensively survey regulatory regions of the serotonin receptors, so it remains possible that a handful of major-effect regulatory SNPs contribute to the variation. Alternatively, summation of small effects of numerous sites throughout the surveyed loci and/or large effects of rare alleles that would go unnoticed in association studies could be contributing to trait variation. Taken together, the evidence reported here is consistent with the hypothesis that very weak selection that is sex specific and structured by population helps shape neurophysiological variation in Drosophila. This is hardly a surprising conclusion, but is worth bearing in mind as attention increasingly focuses on those handful of cases in which sites of major effect seem to account for highly complex behavioral phenotypes.

We thank Arnar Palsson, Ian Dworkin, Zhao-Bang Zeng, and Ed Buckler for discussions; Sergey Nuzhdin for the near-isogenic Californian population sample; Bryon Sosinski and Regina Brierley at the North Carolina State University Genome Research Laboratory for sequencing; and Rebecca Berger and Roland Carrillo for help with establishment and inbreeding of the North Carolina lines. The project was funded by National Institutes of Health grant P01-GM45344 to G.G.

LITERATURE CITED

Andolfatto, P., and J. Wall, 2003 Linkage disequilibrium patterns across a recombination gradient in African *Drosophila melanogaster*. Genetics **165:** 1289–1305.

Aquadro, C. F., D. Bauer, V. DuMont and F. A. Reed, 2001 Genome-wide variation in the human and fruitfly: a comparison. Curr. Opin. Genet. Dev. **11:** 627–634.

Arranz, M., D. Collier, M. Sodhi, D. Ball, G. Roberts *et al.*, 1995 Association between clozapine response and allelic variation in 5–HT2A receptor gene. Lancet **346:** 281–282.

Ashton, K., A. P. Wagoner, R. Carrillo and G. Gibson, 2001 Quantitative trait loci for the monoamine-related traits heart rate and headless behavior in *Drosophila melanogaster*. Genetics **157:** 283–294.

Baines, J. F., Y. Chen and W. Stephan, 2002 DNA sequence variation at a duplicated gene: excess of replacement polymorphism and extensive haplotype structure in the *Drosophila melanogaster bicoid* region. Mol. Biol. Evol. **19:** 989–998.

Begun, D. J., and C. F. Aquadro, 1993 African and North American populations of *Drosophila melanogaster* are very different at the DNA level. Nature **365:** 548–550.

Berry, N., V. Jobanputra and H. Pal, 2003 Molecular genetics of schizophrenia: a critical review. J. Psychiatry Neurosci. **28:** 415–429.

Blenau, W., and A. Baumann, 2001 Molecular and pharmacological properties of insect biogenic amine receptors: lessons from *Drosophila melanogaster* and *Apis mellifera*. Arch. Insect Biochem. Physiol. **48:** 13–38.

Bodmer, R., and T. V. Venkatesh, 1998 Heart development in Drosophila and vertebrates: conservation of molecular mechanisms. Dev. Genet. **22:** 181–186.

Caballero, A., and P. D. Keightley, 1998 Inferences on genome-wide deleterious mutation rates in inbred populations of Drosophila and mice. Genetica **102/103:** 229–239.

Caracristi, G., and C. Schlötterer, 2003 Genetic differentiation between American and European *Drosophila melanogaster* populations could be attributed to admixture of African alleles. Mol. Biol. Evol. **20:** 792–799.

Cardon, L. R., and J. I. Bell, 2001 Association study designs for complex diseases. Nat. Rev. Genet. **2:** 91–99.

Chen, Y., and W. Stephan, 2003 Compensatory evolution of a precursor messenger RNA secondary structure in the *Drosophila melanogaster Adh* gene. Proc. Natl. Acad. Sci. USA **100:** 11499–11504.

Cripps, R. M., and E. N. Olson, 2002 Control of cardiac development by an evolutionarily conserved transcriptional network. Dev. Biol. **246:** 14–28.

Curtis, N. J., J. M. Ringo and H. B. Dowse, 1999 Morphology of the pupal heart, adult heart, and associated tissues in the fruit fly, *Drosophila melanogaster*. J. Morphol. **240:** 225–235.

Daborn, P. J., J. Yen, M. Bogwitz, G. Le Goff, E. Feil *et al.*, 2002 A single P450 allele associated with insecticide resistance in Drosophila. Science **297:** 2253–2256.

De Luca, M., N. V. Roshina, G. L. Geiger-Thornsberry, R. F. Lyman, E. G. Pasyukova *et al.*, 2003 Dopa decarboxylase (Ddc) affects variation in *Drosophila* longevity. Nat. Genet. **34:** 429–433.

Duan, J., A. R. Sanders, J. E. Molen, L. Martinolich, B. J. Mowry *et al.*, 2003 Polymorphisms in the 5′-untranslated region of the human serotonin receptor 1B (*HTR1B*) gene affect gene expression. Mol. Psychiatry **8:** 901–910.

Dulcis, D., and R. B. Levine, 2003 Innervation of the heart of the adult fruit fly, Drosophila melanogaster. J. Comp. Neurol. **465:** 560–578.

Dworkin, I., A. Palsson, K. Birdsall and G. Gibson, 2003 Evidence that *Egfr* contributes to cryptic genetic variation for photoreceptor determination in natural populations of *Drosophila melanogaster*. Curr. Biol. **13:** 1888–1893.

Ewing, B., L. Hillier, M. C. Wendl and P. Green, 1998 Basecalling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. **8:** 175–185.

Excoffier, L., P. E. Smouse and J. M. Quattro, 1992 Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics **131:** 479–491.

Falconer, D. S., and T. F. C. Mackay, 1996 *Introduction to Quantitative Genetics*, Ed. 4. Longman, London.

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.

Glazier, A. M., J. H. Nadeau and T. J. Aitman, 2002 Finding genes that underlie complex traits. Science **298:** 2345–2349.

Glinka, S., L. Ometto, S. Mousset, W. Stephan and D. De Lorenzo, 2003 Demography and natural selection have shaped genetic variation in *Drosophila melanogaster*: a multi-locus approach. Genetics **165:** 1269–1278.

Hay-Schmidt, A., 2000 The evolution of the serotonergic nervous system. Proc. R. Soc. Lond. B Biol. Sci. **267:** 1071–1079.

Hofacker, I. L., 2003 Vienna RNA secondary structure server. Nucleic Acids Res. **31:** 3429–3431.

Hudson, R. R., M. Kreitman and M. Aguadé, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

Johnson, E., J. Ringo and H. Dowse, 1997 Modulation of Drosophila heartbeat by neurotransmitters. J. Comp. Physiol. **167:** 89–97.

Johnson, G. C., L. Esposito, B. J. Barratt, A. N. Smith, J. Heward *et al.*, 2001 Haplotype tagging for the identification of common disease genes. Nat. Genet. **29:** 233–237.

Jorde, L. B., 2000 Linkage disequilibrium and the search for complex disease genes. Genome Res. **10:** 1435–1444.

Kelly, J. K., and M. J. Wade, 2000 Molecular evolution near a two-locus balanced polymorphism. J. Theor. Biol. **204:** 83–101.

Kirby, D. A., S. V. Muse and W. Stephan, 1995 Maintenance of pre-mRNA secondary structure by epistatic selection. Proc. Natl. Acad. Sci. USA **92:** 9047–9051.

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 polymorphisms at the scabrous locus of *Drosophila melanogaster*. Science **266:** 1697–1702.

Langley, C. H., B. P. Lazzaro, W. Phillips, E. Heikkinen and J. M. Braverman, 2000 Linkage disequilibria and the site frequency spectra in the *su(s)* and *su(w(a))* regions of the *Drosophila melanogaster* X chromosome. Genetics **156**: 1837–1852.

Lappalainen, J., J. C. Long, M. Eggert, N. Ozaki, R. W. Robin *et al.*, 1998 Linkage of antisocial alcoholism to the serotonin *5–HT1B* receptor gene in 2 populations. Arch. Gen. Psychiatry **55:** 989–994.

Lazzaro, B. P., B. K. Sceurman and A. G. Clark, 2004 Genetic basis of natural variation in D. melanogaster antibacterial immunity. Science **303:** 1873–1876.

Li, N., and M. Stephens, 2003 Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. Genetics **165:** 2213–2233.

Long, A. D., and C. H. Langley, 1999 The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. Genome Res. **9:** 720–731.

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.

Lucki, I., 1998 The spectrum of behaviors influenced by serotonin. Biol. Psychiatry **44:** 151–162.

Lyman, R. F., C. Lai and T. F. C. Mackay, 1999 Linkage disequilibrium mapping of molecular polymorphisms at the scabrous locus associated with naturally occurring variation in bristle number in *Drosophila melanogaster*. Genet. Res. **74:** 303–311.

Mackay, T. F. C., 1995 The genetic basis of quantitative variation: numbers of sensory bristles of *Drosophila melanogaster* as a model system. Trends Genet. **11:** 464–470.

Mackay, T. F. C., 2001 The genetic architecture of quantitative traits. Annu. Rev. Genet. **35:** 303–339.

Merikangas, K. R., and N. Risch, 2003 Will the genomics revolution revolutionize psychiatry? Am. J. Psychiatry **160:** 625–635.

Monastirioti, M., 1999 Biogenic amine systems in the fruit fly *Drosophila melanogaster*. Microsc. Res. Tech. **45:** 106–121.

Nicholas, K. B., H. B. Nicholas, Jr. and D. W. Deerfield, II, 1997 GeneDoc: analysis and visualization of genetic variation. EMBNEWS **4:** 14.

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.

Pritchard, J. K., M. Stephens, N. A. Rosenberg and P. Donnelly, 2000 Association mapping in structured populations. Am. J. Hum. Genet. **67:** 170–181.

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

Rizki, T. M., 1978 *The Genetics and Biology of Drosophila*, Vol. 2b, Chap. 35. Academic Press, London.

Robbins, J., R. Aggarwal, R. Nichols and G. Gibson, 1999 Genetic variation affecting heart rate in *Drosophila melanogaster*. Genet. Res. **74:** 121–128.

Robertson, A., 1959 The sampling variance of the genetic correlation coefficient. Biometrics **15:** 469–485.

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.

Rockman, M. V., M. W. Hahn, M. Soranzo, D. B. Goldstein and G. A. Wray, 2003 Positive selection on a human-specific transcription factor binding site regulating IL4 expression. Curr. Biol. **13:** 2118–2123.

Sanders, A. R., Q. Cao, J. Taylor, T. E. Levin, J. A. Badner *et al.*, 2001 Genetic diversity of the human serotonin receptor 1B (*HTR1B*) gene. Genomics **72:** 1–14.

Schaeffer, S. W., and E. L. Miller, 1993 Estimates of linkage disequilibrium and the recombination parameter determined from segregating nucleotide sites in the alcohol dehydrogenase region of *Drosophila pseudoobscura*. Genetics **135:** 541–552.

Schneider, S., D. Roessli and L. Excoffier, 2000 Arlequin: a software for population genetics data analysis, Version 2.00. Genetics and Biometry Lab, Department of Anthropology, University of Geneva (http://lgb.unige.ch/arlequin/).

Segalat, L., D. A. Elkes and J. M. Kaplan, 1995 Modulation of serotonin-controlled behaviors by Go in *Caenorhabditis elegans*. Science **267:** 1648–1651.

Thornsberry, J. M., M. M. Goodman, J. Doebley, S. Kresovich, D. Nielsen *et al.*, 2001 *Dwarf8* polymorphisms associate with variation in flowering time. Nat. Genet. **28:** 286–289.

Wall, J. D., and J. K. Pritchard, 2003 Assessing the performance of the haplotype block model of linkage disequilibrium. Am. J. Hum. Genet. **73:** 502–515.

Weir, B. S., 1996 *Genetic Data Analysis II: Methods for Discrete Population Genetics Data*. Sinauer Associates, Sunderland, MA.

Weir, B. S., and C. C. Cockerham, 1984 Estimating F-statistics for the analysis of population structure. Evolution **38:** 1358–1370.

Williams, J., G. Spurlock, P. McGuffin, J. Mallet, M. M. Nothen *et al.*, 1996 Association between schizophrenia and T102C poly-

morphism of the 5-hydroxytryptamine type 2a-receptor gene. European Multicentre Association Study of Schizophrenia (EMASS) Group. Lancet **347:** 1294–1296.

Yang, H. P., and S. V. Nuzhdin, 2003 Fitness costs of Doc expression are insufficient to stabilize its copy number in *Drosophila melanogaster*. Mol. Biol. Evol. **20:** 800–804.

Zondervan, K. T., and L. R. Cardon, 2004 The complex interplay among factors that influence allelic association. Nat. Rev. Genet. **5:** 89–101.

Zornik, E., K. Paisley and R. Nichols, 1999 Neural transmitters and a peptide modulate Drosophila heart rate. Peptides **20:** 45–51.

Communicating editor: J. Bergelson