Restriction-Map Variation in the *Notch* Region of *Drosophila melanogaster*[1]

Stephen W. Schaeffer,[2] *Charles F. Aquadro,*[3] *and Charles H. Langley*
Laboratory of Genetics, National Institute of Environment Health Sciences

A worldwide sample of 37 X chromosomes of *Drosophila melanogaster* was analyzed with four restriction endonucleases for a 60-kb region of the *Notch* locus. Any two randomly chosen homologous chromosomes were heterozygous at one in 143 nucleotides ($\hat{\theta} = 0.007$). The chromosomes that were sampled contained no more than one insertion/deletion. The four insertions and one deletion observed in the 37 chromosomes sampled were located 3′ to the *Notch* transcript; one insertion was represented twice in the sample. The amount of linkage disequilibrium in the *Notch* region appears to be lower than that of the alcohol dehydrogenase locus in *D. melanogaster.* The few instances of linkage disequilibrium observed could be due to geographic differentiation of African populations. The genetic variation estimates in the *Notch* region were comparable with those of the alcohol dehydrogenase region in *D. melanogaster,* suggesting that molecular genetic variation on the X chromosome is not dramatically reduced by selection against slightly deleterious alleles.

Introduction

The alcohol dehydrogenase gene (*Adh*) locus in *Drosophila melanogaster* provided one of the earliest glimpses of the level and organization of nucleotide diversity in natural populations (Langley et al. 1982; Kreitman 1983; Aquadro et al. 1986; Cross and Birley 1986). Any two randomly chosen chromosomes were estimated to be heterozygous at an average of 0.6% of their nucleotide positions for a 13-kb region around the *Adh* locus of *D. melanogaster* when restriction-endonuclease analysis was used. Many of the polymorphic restriction sites were nonrandomly associated with the naturally occurring ADH^F and ADH^S allozymes. Eighty percent of all chromosomes sampled from nature contained one or more sequence-length variant within 8 kb of the *Adh* locus. Most of the large (>500 bp) insertions were shown to be transposable elements. A very similar picture emerged from the early survey of restriction-map variability at the 87A *heat-shock* locus in *D. melanogaster* (Leigh Brown 1983).

The level and organization of nucleotide diversity for a 30-kb region of the *Adh* locus in *D. pseudoobscura* was shown to be substantially different from that in *D. melanogaster* (Schaeffer et al. 1987). Any two randomly chosen chromosomes differed at an estimated 2% of their nucleotide sites for this region of *D. pseudoobscura,* even though *Adh* is monomorphic for allozyme variation in this species (Prakash 1977). With one exception, all pairs of restriction sites were randomly associated with each

1. Key words: population genetics, molecular evolution, X linked, complex locus, insertion/deletion.

2. Present address: Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138.

3. Present address: Section of Genetics and Development, Cornell University, Ithaca, New York 14853.

Address for correspondence and reprints: Dr. Stephen W. Schaeffer, Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138.

Mol. Biol. Evol. 5(1):30–40. 1988.

0737-4038/88/0501-0003$02.00

other. Two of the 19 *D. pseudoobscura* lines surveyed contained a unique insertion located relatively far upstream from the *Adh* locus.

The above differences must be due to species-specific or locus-specific differences in the action of evolutionary forces on *Adh*. If *D. melanogaster* has a smaller population size or a lower mutation rate than *D. pseudoobscura,* we should see equivalent reductions in nucleotide heterogeneity across all loci within *D. melanogaster.* Alternatively, locus-specific selection could result in higher levels of sequence diversity in some loci and not in others. The goal of the present paper is to determine whether the restriction-map variation in the *Notch* locus of *D. melanogaster* resembles the variation at the *Adh* locus in *D. melanogaster* or the variation at the *Adh* locus in *D. pseudoobscura.*

We present here a restriction analysis of a 60-kb region encompassing the X-linked *Notch* locus in *D. melanogaster* (position 1–3.0. on the cytogenetic map). Mutations in the *Notch* region are dominant lethal, but in heterozygous females disruptions of the morphology of the wing and eye are observed. Recessive visible mutants that alter the wing and eye morphology also have been documented. During neurogenesis excess neuroblast formation at the expense of dermoblast formation is also caused by *Notch* mutations (Welshons 1958; Wright 1970; Welshons and Keppy 1981). The *Notch* locus codes for a 10.5-kb poly (A^+) RNA that is transcribed from across a 40-kb genomic region (Artavanis-Tsakonas et al. 1983; Kidd et al. 1983). The protein produced by this 10.5-kb transcript contains 36 tandem repeats of a 40–amino acid sequence similar to human epidermal growth factor (Wharton et al. 1985). This structure of the *Notch* protein may be important for cell-cell communication at the time of neuroblast formation in neurogenesis.

The *Notch* locus represents a new type of locus for the study of molecular population genetics. First, it is a large, complex locus with important developmental functions. Second, it is on the X chromosome and thus provides an initial opportunity to compare variation on the X with that on the autosomes. If much of molecular genetic variation were actually slightly deleterious, as Ohta (1973) and Kimura (1983) have argued, then we would expect the level of variation to be significantly reduced on the X, because mildly deleterious mutants are known to be partially recessive (Mukai and Yamaguchi 1974; Simmons and Crow 1977) and thus more quickly eliminated from the X chromosomes than from the autosomes (Haldane 1927).

Material and Methods

Strains

A list of strains with their collection localities, encompassing six continents, is given in table 1. These lines were begun as isofemale lines and maintained in the laboratory. As indicated in table 2, some of the lines were clearly segregating for two distinct alleles. When this was the case, each was scored and they were designated in table 2 by the suffixes "a" or "b."

Restriction Analysis

Total genomic DNA from each line was isolated according to the method of Bingham et al. (1981). DNAs were digested singly with restriction endonucleases (*Bam*HI, *Eco*RI, *Hin*dIII, and *Sal*I). The restriction fragments were size fractionated on 1% agarose gels in E buffer (McDonell et al. 1977; E buffer = 40 mM Tris[hydroxymethyl]aminomethane, 20 mM acetic acid, 2 mM ethylenedinitrilotetraacetic acid [EDTA]; pH 8.0), then transferred to aminophenylthioether cellulose

Table 1
Collection Locations of *Drosophila melanogaster* Lines

Line	Location	Obtained by[a]
1. Arg 3	La Plata, Argentina	A
2. Arg 4	La Plata, Argentina	A
3. Arg 6	La Plata, Argentina	A
4. Bel 2	Belize (British Honduras)	B
5. Bel 3	Belize (British Honduras)	B
6. FM 122	Raleigh, North Carolina	C
7. FM 167	Raleigh, North Carolina	C
8. FM 135	Raleigh, North Carolina	C
9. FM 149	Raleigh, North Carolina	C
10. FM 152	Raleigh, North Carolina	C
11. 7.30.11	Madison, Wisconsin	D
12. 7.29.11	Madison, Wisconsin	D
13. 7.29.12	Madison, Wisconsin	D
14. 7.29.8	Madison, Wisconsin	D
15. AV 15-25	Ottawa	B
16. FrV2-1	Villeurbanne, France	E
17. FrV3-1	Villeurbanne, France	E
18. FrV16-1	Villeurbanne, France	E
19. V13-1	Villeurbanne, France	E
20. V4	Villeurbanne, France	B
21. V14	Villeurbanne, France	B
22. V22-24	Villeurbanne, France	B
23. Ben16-C	Benin, West Africa	E
24. Ben11-C	Benin, West Africa	E
25. CAF-7	Central Africa	A
26. Viet12-1	Ho Chi Minh City, Vietnam	E
27. Viet13-1	Ho Chi Minh City, Vietnam	E
28. Viet15-1	Ho Chi Minh City, Vietnam	E
29. Taiw 2	Taiwan	E
30. Taiw 20	Taiwan	E
31. Aust 2	Fairfield, Australia	E
32. Aust 14	Fairfield, Australia	E
33. Haw 7	Hawaii	B

[a] A = Rama Singh (Singh and Coulthart 1982); B = George R. Carmody; C = Charles H. Langley; D = William J. Engels (Engels and Preston 1980); and E = Rama Singh (Singh et al. 1982).

(APT) filters by means of the transfer technique of Southern (1975) (APT filters are commercially available through Schleicher and Schull). Probes labeled with [α-^{32}P]dCTP (−3,000 Ci/mmol; Rigby et al. 1977) were hybridized to digested genomic DNA on DPT filters in 50% formamide at 42 C overnight. Nonspecifically hybridized probe was removed with three 5-min washes in a solution of 2 × SSC (1 × SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.5) and 0.1% sodium dodecyl sulfate (SDS) at room temperature, then in two 15-min washes in a solution of 0.1 × SSC and 0.1% SDS at 50 C. The filters were autoradiographed, and the probes stripped from the APT filters with 0.5 M NaOH for 30 min at room temperature. The filters were neutralized in a solution of 2 × SSPE (1 × SSPE = 0.18 M NaCl, 10 mM $NaPO_4$, 1 mM EDTA, pH 7.7) and 0.1% SDS for 30 min at room temperature and were ready to be hybridized with another probe.

Table 2
Eighteen *Notch*-Region Haplotypes from 19 Polymorphic Restriction-Map Variants for 37 Chromosomes of *Drosophila melanogaster*

		Restriction-Map Variants[a]																		
Line	Hap[b]	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s
Arg 3	B	+	+	−	+	−	−	−	+	+	+	−	−	+	−	−	+	−	−	−
Arg 4	B	+	+	−	+	−	−	−	+	+	+	−	−	+	−	−	+	−	−	−
Arg 6	B	+	+	−	+	−	−	−	+	+	+	−	−	+	−	−	+	−	−	−
Bel 2	A	+	−	−	+	+	−	−	+	+	+	−	−	+	−	−	+	−	−	−
Bel 3	C	+	−	−	+	+	−	+	+	+	+	−	−	+	−	−	+	−	−	−
Fm122	L	+	−	−	−	+	−	−	+	+	+	−	−	+	−	−	+	−	−	−
Fm167	A	+	−	−	+	+	−	−	+	+	+	−	−	+	−	−	+	−	−	−
Fm135	B	+	+	−	+	−	−	−	+	+	+	−	−	+	−	−	+	−	−	−
Fm149a	M	+	+	−	+	−	−	+	+	+	+	−	−	+	−	−	+	−	−	−
Fm149b	N	+	+	−	+	−	−	+	+	+	+	−	−	+	−	−	−	−	−	−
Fm152a	O	+	+	−	−	−	−	−	+	+	+	−	−	+	−	−	+	−	−	−
Fm152b	P	+	+	−	−	−	−	−	+	+	+	−	−	+	−	−	−	−	−	−
7.30.11	C	+	−	−	+	+	−	+	+	+	+	−	−	+	−	−	+	−	−	−
7.29.11	A	+	−	−	+	+	−	−	+	+	+	−	−	+	−	−	+	−	−	−
7.29.12	A	+	−	−	+	+	−	−	+	+	+	−	−	+	−	−	+	−	−	−
7.29.8	K	−	+	+	+	−	−	+	+	−	+	−	−	−	−	−	+	−	−	−
AV15-25	B	+	+	−	+	−	−	−	+	+	+	−	−	+	−	−	+	−	−	−
FrV2-1	G	−	+	−	+	−	−	−	+	+	+	−	−	+	−	−	+	−	+	−
FrV3-1	H	+	+	−	+	−	−	−	+	+	+	−	−	+	+	−	+	−	−	−
FrV16-1	I	+	+	−	+	−	−	+	+	+	+	−	+	+	−	−	+	−	−	−
V13-1	B	+	+	−	+	−	−	−	+	+	+	−	−	+	−	−	+	−	−	−
V4	C	+	−	−	+	+	−	+	+	+	+	−	−	+	−	−	+	−	−	−
V14	R	+	+	−	+	−	−	−	+	−	−	−	−	+	−	−	+	−	−	−
V22-24	M	+	+	−	+	−	−	+	+	+	+	−	−	+	−	−	+	−	−	−
Ben16-Ca	D	−	−	−	+	+	−	+	+	−	+	−	+	+	−	−	−	−	−	−
Ben16-Cb	E	−	−	−	+	+	+	+	+	−	+	−	+	+	−	−	−	−	−	−
Ben11-Ca	A	+	−	−	+	+	−	−	+	+	+	−	−	+	−	−	+	−	−	−
Ben11-Cb	F	+	−	−	+	+	+	−	+	+	−	+	−	+	−	−	+	−	−	−
CaF 7	A	+	−	−	+	+	−	−	+	+	+	−	−	+	−	−	+	+	−	−
Viet12-1	A	+	−	−	+	+	−	−	+	+	+	−	−	+	−	+	+	−	−	−
Viet13-1	J	+	+	−	+	−	−	−	+	+	+	−	−	+	−	−	−	−	−	−
Viet15-1	A	+	−	−	+	+	−	−	+	+	+	−	−	+	−	+	+	−	−	−
Taiw 2	B	+	+	−	+	−	−	−	+	+	+	−	−	+	−	−	+	−	−	−
Tai 20	Q	+	+	−	+	−	−	−	−	+	+	−	−	+	−	−	+	−	−	−
Aust 2	A	+	−	−	+	+	−	−	+	+	+	−	−	+	−	−	+	−	−	−
Aust 14	A	+	−	−	+	+	−	−	+	+	+	−	−	+	−	−	+	−	−	+
Haw 7	A	+	−	−	+	+	−	−	+	+	+	−	−	+	−	−	+	−	−	−

[a] Restriction-site coordinates (see fig. 1) are as follows: a = *Eco*RI −36.6; b = *Bam*HI −36.4; c = *Eco*RI −36.0; d = *Sal*I −35.2; e = *Bam*HI −24.4; f = *Bam*HI −18.6; g = *Eco*RI −9.7; h = *Hin*dIII −5.9; i = *Sal*I 2.4; j = *Sal*I 7.7; k = *Eco*RI 8.3; l = *Eco*RI 8.4; m = *Eco*RI 11.2; n = *Eco*RI 15.6; o = insertion 2.2 kb; p = *Hin*dIII 23.6; q = deletion 0.14 kb; r = insertion 1.2 kb; and s = insertion 1.8 kb.

[b] Haplotype designation. The haplotype designation does not include insertions or deletions.

We obtained four overlapping lambda-phage clones—NL35, NL22, N2, and NR311—from M. W. Young (Kidd et al. 1983). These clones probe a 60-kb region that includes the major transcriptional unit of the *Notch* locus. Phage DNA was isolated according to the method of Maniatis et al. (1982). The two independent groups that cloned the *Notch* locus orient their restriction maps differently (Kidd et al. 1983;

Grimwade et al. 1985). The zero point on the restriction map in figure 1 is that used by Kidd et al. (1983). The map from Kidd et al. (1983) can be converted to the scale in Grimwade et al. (1985) by adding 1.1 kb to the site location.

Estimates of Genetic Variability

We used two estimators of population genetic variation at the nucleotide level, viz., $\hat{\theta}$ and $\hat{\pi}$. Hudson (1982) presents the estimator of $\hat{\theta}$ as $\hat{\theta} = \hat{p}/\ln(n)$, where $\hat{p}$ is the proportion of polymorphic sites and n is the sample size. The estimator of $\hat{p}$ is given by $\hat{p} = k/(2m - k)j$, where k is the number of polymorphic sites, m is the total number of restriction sites, and j is the number of base pairs recognized by the restriction site. The variance of the point estimate $\hat{\theta}$ is given by $\mathrm{Var}(\hat{\theta}) = \theta^2/k$. Nei and Tajima (1981) estimate $\hat{\pi}$ as

$$\hat{\pi} = [n/(n-1)] \sum_{i \neq j}^{n} x_i x_j \pi_{ij},$$

where x_i is the frequency of the ith haplotype, n is the sample size, and π_{ij} is the number of nucleotide differences per site between the ith and the jth haplotypes. The parameter π_{ij} is estimated by $\pi_{ij} = -\ln S_{ij}/r$, where S_{ij} is the fraction of shared sites between the haplotypes i and j and r is the number of base pairs in the recognition sequence of the restriction site. S_{ij} is estimated by $S_{ij} = 2n_{ij}/(n_i + n_j)$, where n_{ij} is the number of shared restriction sites between haplotypes i and j and n_i is the number of restriction sites in haplotype i. The analysis of linkage disequilibrium consisted of calculation of linkage disequilibrium parameters D and D'. D is a measure of the

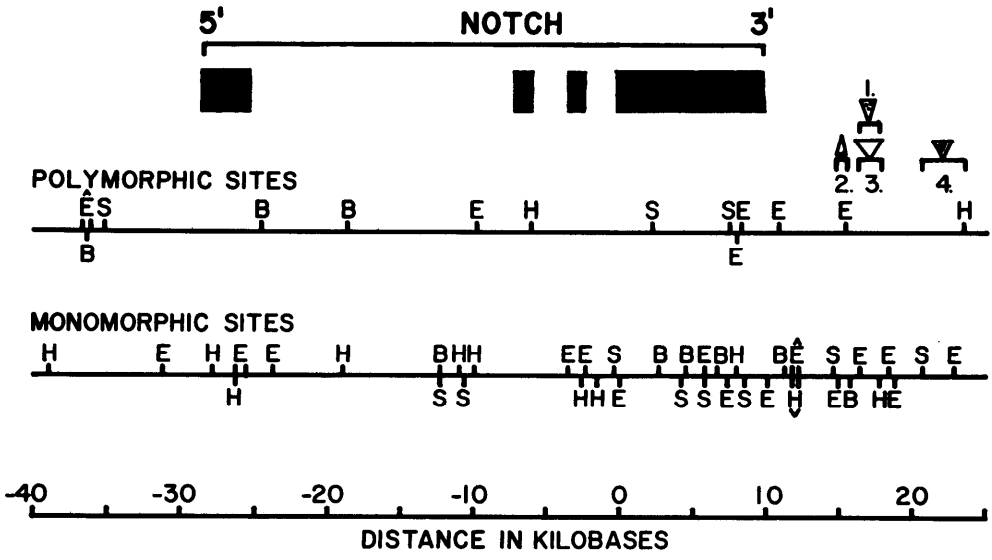

FIG. 1.—Restriction map of the *Notch*-locus region in *Drosophila melanogaster*. The shaded boxes above the restriction map show a simplified version of the fine structure of the coding region (after Kidd et al. 1983). The restriction sites are as follows: B = *Bam*HI; E = *Eco*RI; H = *Hin*dIII; and S = *Sal*I. Restriction sites with a circumflex (^) denote two sites. The insertions/deletions are as follows: 1., 1.20-kb insertion in line FrV2-1; 2., 0.14-kb deletion in line CaF 7; 3., 2.20-kb insertion in lines Viet12-1 and Viet15-1; and 4., 1.8-kb insertion in line Aust 14.

disparity of the observed gametic frequency from the theoretically expected gametic frequency determined by $D = X1 - p1q1$, where $X1$ is the frequency of the observed (++) gametic type, $p1$ is the frequency of the (+) allele at the first restriction site, and $q1$ is the frequency of the (+) allele at the second restriction site. D' is the ratio of the linkage disequilibrium parameter, D, to its theoretical maximum, D_{max} or $D' = D/D_{max}$ (Lewontin 1964). D_{max} is defined as $\min(p1q2, p2q1)$ if $D > 0$ or as $\min(p1q1, p2q2)$ if $D < 0$. (Note: the definition of D_{max} in the footnote to table 3 of Aquadro et al. (1986) is a misprint. The values in their table 3 are correct.) The statistical significance of deviations of D from zero was judged by means of Fisher's exact test of independence.

Results

The variation in the region of the *Notch* locus is presented in figure 1 and table 2. Figure 1 is a restriction map of the *Notch* region, depicting the transcription unit, exons and introns, and 10 kb of 5'- and 3'-flanking sequence. We have analyzed this variation in the following four ways: (1) comparison of the amounts of variation in different structural/functional subdivisions of the DNA of the *Notch*-locus region, (2) estimation of both nucleotide substitution and insertion/deletion variation in the natural population rather than in the sample, (3) analysis of the amount of linkage disequilibrium in the *Notch*-locus region, and (4) search for evidence of geographic variation in the distribution of variants.

Variation in Different Parts of the *Notch*-Locus Region

Figure 1 shows no obvious clustering of polymorphic restriction sites. Four of six sites are polymorphic in the 10 kb of flanking sequence 5' to the *Notch* gene. This 5'-flanking region appears to be more variable than the rest of the region. Although we see four polymorphic sites in the large exon, it also has more monomorphic sites. The three large insertions and one small deletion do seem to be clustered in the 15 kb of 3'-flanking sequence of the *Notch*-locus region.

Estimates of Population Variation

Fifteen of 58 restriction sites in the *Notch*-locus region (table 2) were polymorphic in the sample of 37 chromosomes. *Expected heterozygosity* is a good estimate of genetic variation in natural populations because it is unbiased by sample size. Expected heterozygosity is most easily interpreted as the expected proportion of nucleotide sites different between any two randomly chosen alleles of the *Notch*-locus region. The two estimates of expected heterozygosity most commonly used are $\hat{\theta}$ (Hudson 1982) and $\hat{\pi}$ (Nei and Tajima 1981). For the low per-nucleotide-site variation in this and other *Drosophila* studies and under the usual assumptions of the neutral theory of molecular evolution, estimates of θ and π and of $\hat{\theta}$ and $\hat{\pi}$, respectively, are approximately equal to one another and to $4Ne\mu$ where Ne is the effective population size and μ is the mutation rate per nucleotide per generation to selectively neutral alternate states. On the basis of the data in table 2, $\hat{\theta} = 0.007 \pm 0.002$ (Hudson 1982) and $\hat{\pi} = 0.005$ (Nei and Tajima 1981). The frequency of chromosomes bearing insertions or deletions >100 bp in length is 0.002/chromosome/kb of sequence (5 variants/37 chromosomes/60 kb). This is 10-fold less than the 0.024 (= 15 variants/49 chromosomes/13 kb) observed at the *Adh* locus in *D. melanogaster* (Aquadro et al. 1986). The haplotypes with insertions or deletions are in low frequency, as has been seen in previous studies (e.g., Aquadro et al. 1986).

Linkage Disequilibrium

The 10 polymorphic restriction sites and one polymorphic insertion were used to make 55 pairwise tests for linkage disequilibrium. A site was considered polymorphic if it had two copies of the rare variant. In six of 55 pairs of polymorphic sites there is significant departure from random association (table 3). For pairs of polymorphic sites for which the frequencies of the rare variants are small, D might be statistically significant, but, owing to a small sample size, there may not be enough power in the statistical test to show differences between observed D values and zero. This level of linkage disequilibrium appears to be lower than that observed at *Adh* in *D. melanogaster*. However, the density per kilobase for polymorphic sites scored is much higher for *Adh* than for *Notch*. Nonetheless, the frequency of pairs of sites in which all four gametic types were found was high. This includes four of six significant pairwise comparisons in table 3. There is a clear pattern of frequent occurrence of chromosomes bearing the rare alleles at both sites. This may have a simple explanation (see below).

Geographic Variation

Because the sampled chromosomes are not from a single population, the variation and linkage disequilibrium discussed above reflect both within- and between-population variation. Several of the sites show distinctly different frequencies in the African samples compared with those from other areas. All of the African populations have the *Bam*HI site at position −24.4 on the restriction map and are missing the *Bam*HI site at position −36.6 (table 2). This geographic differentiation led to the apparent gametic associations of several pairs of rare alleles, especially in the African-derived chromosomes. Thus the curious linkage disequilibrium discussed above should probably be viewed as evidence of geographic associations in allelic frequencies at several of the polymorphic sites.

Table 3
D′ and D among *Notch*-Region Restriction Sites

	Site										
Site	a	b	d	e	f	g	i	j	l	o	p
a ...		0.026	−1.000	−0.026	−0.439	−0.657	0.720**	−1.000	−0.626*	1.000	0.422
b ...	0.002		−0.315	−1.000***	−1.000	−0.026	0.026	0.026	−0.351	−1.000	−0.178
d ...	−0.009	−0.012		0.315	1.000	1.000	−1.000	−1.000	1.000	1.000	0.229
e ...	−0.002	−0.250	0.012		1.000	0.026	−0.026	−0.026	0.351	1.000	0.178
f	−0.021	−0.028	0.004	0.028		0.315	−0.439	−0.471	0.456	−1.000	−0.422
g ...	−0.052	−0.004	0.022	0.004	0.012		−0.657	1.000	1.000*	−1.000	−0.452
i	0.069	0.002	−0.009	−0.002	−0.021	−0.052		0.439	−0.626*	1.000	0.422
j	−0.006	0.001	−0.004	−0.001	−0.024	0.015	0.021		1.000	1.000	−1.000
l	−0.045	−0.015	0.007	0.015	0.023	0.059	−0.045	0.004		−1.000	−0.615*
o ...	0.006	−0.028	0.004	0.028	−0.003	−0.015	0.006	0.003	−0.004		1.000
p ...	0.039	−0.012	0.016	0.012	−0.020	−0.045	0.039	−0.007	−0.043	0.007	

NOTE.—D′ values are above the diagonal; D values are below the diagonal. A value of 0 indicates completely random association; a value of −1 or 1 represents complete association. Statistical significance of association between two sites was determined by means of Fisher's exact test of independence. The letters of the site designation are those used in table 2.

* $P < 0.05$.

** $P < 0.001$.

*** $P < 0.0001$.

Discussion

The *Notch* locus differs from *Adh* in several important ways. *Notch* is a well-studied, X-linked locus with complex genetics and developmental roles. The *Notch* transcript is derived from the central 40 kb of the 60-kb region that we have studied. This transcript is processed into a 10.5-kb mRNA (Artavanis-Tsakonas et al. 1983; Kidd et al. 1983). We surveyed the 60 kb of DNA in the *Notch*-locus region of 37 chromosomes collected from many different geographic locations throughout the world. The 0.007 ± 0.002 estimate of expected heterozygosity for the *Notch*-locus region indicates that the level of nucleotide variation is comparable with that found at *Adh* in *Drosophila melanogaster* (Aquadro et al. 1986), whereas the few occurrences of linkage disequilibrium and sequence-length variation in the *Notch* region appear to be more like those of *Adh* in *D. pseudoobscura* (Schaeffer et al. 1987).

Because the *Notch* locus is so large and complex, there was an opportunity to discern any large differences in the amounts or kinds of variation in the several functionally distinct regions. The few large insertions that are present among the 37 chromosomes are clustered in the 10 kb of 3′-flanking sequence. This is similar to the situation found at *Adh* in *D. melanogaster* in that most of the insertions are clustered in a small region 3′ to *Adh*. However, the frequency of insertions and deletions >100 bp in length is particularly low at *Notch* (0.002/chromosome/kb vs. 0.024 for *Adh* or 0.008 for *white;* Aquadro et al. 1986; Langley and Aquadro 1987). *Notch* is more similar to the *Adh* region of *D. pseudoobscura,* where the frequency of 0.002/chromosome/kb >100 bp in length (1 insertion/20 chromosomes/32 kb) was observed (Schaeffer et al. 1987). Apart from the suggestion of a cluster of polymorphic restriction sites in the 5′-flanking region, the variable restriction sites appear to be uniformly distributed across the region. Because relatively few nucleotides were surveyed, it is possible that there are significant differences in the levels of variation in the different domains of the *Notch* locus. Certainly the surveys of the *Adh* region in *D. melanogaster* with many 4-base-recognition restriction enzymes (Kreitman and Aguade 1986) and by means of direct DNA sequencing (Kreitman 1983) revealed significant differences among segments of the *Adh* region.

The expected heterozygosity at the nucleotide-site level estimated on the basis of the survey of the *Notch*-locus region is either 0.007 or 0.005, depending on the method used to calculate the value. This means that, on average, two independent *Notch* isoalleles will differ by 420 nucleotides over the 60-kb region. This amount of nucleotide variation is comparable with, although slightly higher than, that measured at the 87A *heat-shock* locus (0.002) and at *Adh* (0.006) in *D. melanogaster* (Leigh Brown 1983; Aquadro et al. 1986). Companion studies on the *white*-locus region and the *Adh* region of the same worldwide lines estimates the heterozygosity at the nucleotide-site level to be 0.015 (white) (Langley and Aquadro 1987) and 0.004 (*Adh*) (C. F. Aquadro and C. H. Langley, unpublished results). This observation lends no support to the expectation that the level of molecular genetic variation might be less because natural selection would be more efficient on X-linked genes.

In the *Adh* and 87A *heat-shock* regions, the amount of linkage disequilibrium or nonrandom association among polymorphic sites is high (Leigh Brown 1983; Aquadro et al. 1986). In contrast, only six of the 55 pairs of polymorphic sites in the *Notch*-locus region are clearly in linkage disequilibrium. Two points are relevant in interpreting this level of linkage disequilibrium. First, geographic differentiation in the frequencies

of polymorphic sites at *Notch* appears to have contributed to the linkage disequilibrium. In contrast, the data for *Adh* and the 87A *heat-shock* gene are from smaller geographic regions or populations.

The chromosomes sampled for this survey are from many different locales around the world (table 1). Several of the rarer variants are found more than once in the samples from Africa. And the − *Bam*HI (−36.6) and + *Bam*HI (−24.4) gametic types are clearly more common in Africa. These geographic differences lead to some apparent linkage disequilibrium when the chromosomes from these various locations are pooled. Thus, there is some evidence for subdivision and gene-frequency differentiation.

Second, the density of polymorphic sites in the *Adh* region is higher than that in the *Notch* region. The distance between polymorphic restriction sites tends to be greater for the *Notch* data; thus there is a greater chance for recombination to break up any associations between sites. The polymorphic restriction sites at *Adh* are tightly linked and tend to remain associated.

Thus, the *Notch*-locus region resembles the situation at *Adh* in *D. pseudoobscura,* in which few pairs of polymorphic sites were in linkage disequilibrium (Schaeffer et al. 1987). The white-locus region in *D. melanogaster* also exhibits very little linkage disequilibrium (Langley and Aquadro 1987). On a finer scale of restriction maps using 4-base-recognition restriction enzymes and DNA sequencing—i.e., within segments of DNA <1 kb in length—there may still be considerable linkage disequilibrium. But the preliminary impression is that the apparently higher level of linkage disequilibrium found at *Adh* in *D. melanogaster* may be the exception, reflecting a particular, recent evolutionary history involving two alleles with large differences in activity owing to a single amino acid replacement.

In summary, the variation in the restriction map of the *Notch*-locus region of *D. melanogaster* among 37 chromosomes sampled from around the world indicates that the per-nucleotide-site heterozygosity is comparable with that observed at *Adh* in *D. melanogaster.* However, the amount of length variation and the amount of linkage disequilibrium appear to be reduced compared with those of *Adh* in *D. melanogaster* and to be similar to those observed at *white* in *D. melanogaster* or at *Adh* in *D. pseudoobscura.* Finally, there is evidence for a small amount of geographic differentiation in gene frequencies. There is little evidence in these data or other recent surveys that either X-linked or large, complex loci are particularly distinct with respect to molecular population-genetic variation.

Acknowledgments

We would like to thank George R. Carmody, William J. Engels, and Rama Singh for providing us with *Drosophila* lines; Susan F. Deese and Robert Jennings for technical assistance; and Rosemary Redfield, Margaret Riley, and two anonymous reviewers for helpful comments on the manuscript.

LITERATURE CITED

AQUADRO, C. F., S. F. DEESE, M. M. BLAND, C. H. LANGLEY, and C. C. LAURIE-AHLBERG. 1986. Molecular population genetics of the alcohol dehydrogenase gene region of *Drosophila melanogaster.* Genetics **114**:1165–1190.

ARTAVANIS-TSAKONAS, S., M. A. T. MUSKAVITCH, and B. YEDVOBNICK. 1983. Molecular

cloning of *Notch,* a locus affecting neurogenesis in *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA **80**:1977–1981.

BINGHAM, P. M., R. LEVIS, and G. M. RUBIN. 1981. Cloning of DNA sequences from the white locus of *Drosophila melanogaster* by a novel and general method. Cell **25**:693–704.

CROSS, S. R. H., and A. J. BIRLEY. 1986. Restriction endonuclease map variation in the *Adh* region in populations of *Drosophila melanogaster.* Biochem. Genet. **24**:415–433.

ENGELS, W. R., and C. R. PRESTON. 1980. Components of hybrid dysgenesis in a wild population of *Drosophila melanogaster.* Genetics **95**:111–128.

GRIMWADE, B. G., M. A. T. MUSKAVITCH, W. J. WELSHONS, B. YEDVOBNICK, and S. ARTAVANIS-TSAKONAS. 1985. The molecular genetics of the *Notch* locus in *Drosophila melanogaster.* Dev. Biol. **107**:503–519.

HALDANE, J. B. S. 1927. A mathematical theory of natural and artificial selection. Part V. Selection and mutation. Proc. Camb. Phil. Soc. **23**:838–844.

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

KIDD, S., T. J. LOCKETT, and M. W. YOUNG. 1983. The *Notch* locus of *Drosophila melanogaster.* Cell **34**:421–433.

KIMURA, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, New York.

KREITMAN, M. 1983. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster.* Nature **304**:412–417.

KREITMAN, M. E., and M. AGUADE. 1986. Excess polymorphism at the *Adh* locus of *Drosophila melanogaster.* Genetics **114**:93–110.

LANGLEY, C. H., and C. F. AQUADRO, 1987. Restriction-map variation in natural populations of *Drosophila melanogaster: white*-locus region. Mol. Biol. Evol. **4**:651–663

LANGLEY, C. H., E. MONTGOMERY, and W. F. QUATTLEBAUM. 1982. Restriction map variation in the *Adh* region of *Drosophila.* Proc. Natl. Acad. Sci. USA **79**:5631–5635.

LEIGH BROWN, A. J. 1983. Variation at the 87A heat-shock locus in *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA **80**:5350–5354.

LEWONTIN, R. C. 1964. The interaction of selection and linkage. I. General considerations; heterotic models. Genetics **49**:49–67.

MCDONELL, M. W., M. N. SIMON, and F. W. STUDIER. 1977. Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J. Mol. Biol. **110**:119–146.

MANIATIS, T., E. F. FRITSCH, and J. SAMBROOK. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

MUKAI, T., and O. YAMAGUCHI. 1974. The genetic structure of natural populations of *Drosophila melanogaster.* XI. Genetic variability in a local population. Genetics **76**:339–366.

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

OHTA, T. 1973. Slightly deleterious mutant substitutions in evolution. Nature **246**:96–98.

PRAKASH, S. 1977. Further studies on gene polymorphism in the mainbody and geographically isolated populations of *Drosophila pseudoobscura.* Genetics **85**:713–719.

RIGBY, P. W. J., M. DIECKMAN, C. RHODES, and P. BERG. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.

SCHAEFFER, S. W., C. F. AQUADRO, and W. W. ANDERSON. 1987. Restriction-map variation in the alcohol dehydrogenase region of *Drosophila pseudoobscura.* Mol. Biol. Evol. **4**:254–265.

SIMMONS, M. J., and J. F. CROW. 1977. Mutations affecting fitness in *Drosophila* populations. Annu. Rev. Genet. **11**:49–78.

SINGH, R. S., and M. B. COULTHART. 1982. Genic variation in abundant soluble proteins of *Drosophila melanogaster* and *Drosophila pseudoobscura.* Genetics **102**:437–453.

SINGH, R. S., D. A. HICKEY, and J. DAVID. 1982. Genetic differentiation between geographically distant populations of *Drosophila melanogaster.* Genetics **101**:235–256.

SOUTHERN, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.

WELSHONS, W. J. 1958. The analysis of a pseudoallelic recessive lethal system at the *Notch* locus of *Drosophila melanogaster.* Cold Spring Harbor Symp. Quant. Biol. **23**:171–176.

WELSHONS, W. J., and D. O. KEPPY. 1981. The recombinational analysis of aberrations and the position of the *Notch* locus on the polytene chromosome of *Drosophila.* Mol. Gen. Genet. **181**:319–324.

WHARTON, K. A., K. M. JOHANSEN, T. XU, and S. ARTAVANIS-TSAKONAS. 1985. Nucleotide sequence from the neurogenic locus *Notch* implies a gene product that shares homology with proteins containing EGF-like repeats. Cell **43**:567–581.

WRIGHT, T. R. F. 1970. The genetics of embryogenesis in *Drosophila.* Adv. Genet. **15**: 261–395.

WALTER M. FITCH, reviewing editor

Received June 14, 1987; revision received August 28, 1987