Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 3562–3566, May 1986
Population Biology

Genetic uniformity in two populations of *Drosophila melanogaster* as revealed by filter hybridization of four-nucleotide-recognizing restriction enzyme digests

(polymorphism/restriction polymorphism/insertion/deletion variation)

MARTIN KREITMAN* AND MONTSERRAT AGUADÉ†

Museum of Comparative Zoology Laboratories, Harvard University, Cambridge, MA 02138

Communicated by C. Clark Cockerham, December 19, 1985

ABSTRACT **A filter hybridization method is described for identifying restriction-site and insertion/deletion variation by using restriction enzymes that recognize four-nucleotide sequences and denaturing polyacrylamide gels for separating fragments. Eighty-seven lines of *Drosophila melanogaster* representing two natural populations were surveyed over a 2.7-kilobase region encompassing the alcohol dehydrogenase locus. Fifty distinct haplotypes were identified from 17 restriction-site and 11 insertion/deletion polymorphisms and from one allozyme polymorphism. There was no evidence for genetic differentiation between an East-Coast and a West-Coast (North American) sample. This technique has widespread applications in screening for DNA polymorphism.**

Two techniques dominate attempts to identify allelic variation at the DNA level—Southern blot analysis of restriction endonuclease digestions (1, 2) and direct DNA sequencing (3). However, neither technique fully satisfies two requirements for the study of variation in natural populations: (*i*) to allow sampling of a large number of alleles and (*ii*) to resolve variable sites over relatively short stretches of DNA. The problem of finding nucleotide polymorphisms is not that they are too rare. For example, in his sequence survey of 11 alleles coding for alcohol dehydrogenase (Adh) in *Drosophila melanogaster*, Kreitman (3) identified 43 polymorphic sites, including insertions/deletions in a 2.7-kilobase (kb) region. However, without technical innovation, the extension of this work to population surveys has remained impractical.

One such innovation that allows DNA sequences to be obtained directly from genomic DNA (gDNA) has recently been reported by Church and Gilbert (4). In this paper, a modification of their technique is described, which uses enzymes that recognize four-nucleotide sequences to reveal restriction-site and insertion/deletion polymorphisms within small regions of DNA. We have applied this technique to estimate the extent of polymorphism and the degree of genetic divergence at the *Adh* structural locus in two populations of *D. melanogaster*.

MATERIAL AND METHODS

Fly Samples. Isofemale lines were established from wild-caught flies collected on banana bait. One sample was collected at Farmers Market, Raleigh, NC (October 1983) and one at Putah Creek, Davis, CA (October 1983). Isochromosomal lines were established by the Curly extraction procedure using SM5/B1L[2] as extractor stock (5). Adh allozyme phenotype was determined as described by Kreitman (6).

DNA Preparation and Electrophoresis. Total nucleic acid was extracted from approximately 0.5 g of frozen (−70°C) adult flies by cell lysis with potassium acetate/NaDodSO$_4$ and extraction with phenol (7). Two to three micrograms of gDNA (estimated from ethidium bromide-stained agarose gels containing total nucleic acid) were digested first for 4–6 hr with the following enzymes: *Alu* I, *Ban* I, *Dde* I/*Bam*HI, *Hae* III, *Hha* I, *Msp* I, *Sau*3A, *Sau*96 I, or *Taq* I; then the samples were digested for 0.5–1 hr with RNase I (100 μg/ml). Samples were precipitated with ethanol, washed, dried under reduced pressure, and resuspended in 3 μl of formamide loading buffer (94% formamide/0.05% xylene cyanol/0.05% bromophenol blue/10 mM Na$_2$EDTA, pH 7.2). After incubation in a boiling water bath or 90°C dry bath for 5–10 min, 1.5–2.0 μl of each sample was loaded with a Hamilton syringe onto a standard 30 cm × 40 cm × 0.4 mm 5% polyacrylamide/7 M urea buffer gradient DNA sequencing gel (8). The buffer gradient was 50–500 mM in Tris borate/EDTA, pH 8.3. Fifty to sixty samples were loaded on one gel in 0.35-mm sharks' teeth-comb-formed slots. Gels were run at 1200–1300 V until the bromophenol blue reached 35–40 cm from the origin. Electrophoretic transfer of DNA from the gel to New England Nuclear/DuPont GeneScreen (NEF 972) and subsequent UV crosslinking was performed as described in (4).

Probe Preparation. A 2.7-kb gel-purified *Sal* I–*Cla* I DNA fragment containing the *Adh* structural locus (3) was nick-translated by using 3000 or 5000 Ci (1 Ci = 37 GBq) per mmol of [α-^{32}P]dATP (DuPont/New England Nuclear) to a specific activity of 8–10 × 10^8 cpm/μg of DNA (9). The probe was boiled for 5–10 min before being added to prewarmed (65°C) hybridization buffer (1% crystalline grade bovine serum albumin /1 mM EDTA/0.5 M sodium phosphate, pH 7.2/7% NaDodSO$_4$).

Hybridization/Wash. Prehybridization/hybridization was at 55–65°C either using polyethylene/polyester-laminated bags as described (4) (5–10 × 10^7 cpm in 12 ml of hybridization buffer) or more recently as follows: one to five 30 × 40 cm wet (50 mM Tris borate/10 mM EDTA, pH 8.3) filters placed on top of one another were rolled tightly around a 10-ml sterile disposable pipette and placed in a 35 × 2.54 cm (i.d.) polycarbonate (Lexan) tube sealed at one end. Approximately 20 μl of hybridization buffer (65°C) per cm^2 was added, and the open end of the tube was sealed with a rubber cork. The tube was placed on a modified Wheaton tissue culture roller set at 5 rpm in an incubator at 55–60°C for a minimum of 5 min. This solution was replaced with 12–18 ml of hybridization buffer containing 5–50 × 10^7 cpm of probe prewarmed to 65°C. It is important to orient the filters so they "unwind" when rolling against the inner surface of the tube

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: bp, base pair(s); gDNA, genomic DNA; Adh, alcohol dehydrogenase; Adh^S and Adh^F, isozymes whose relative migration is slow and fast in electrophoresis.

*Current address: National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709.
†Permanent address: Departament de Genètica, Universitat de Barcelona, Av. Diagonal, 645, 08028 Barcelona, Spain.

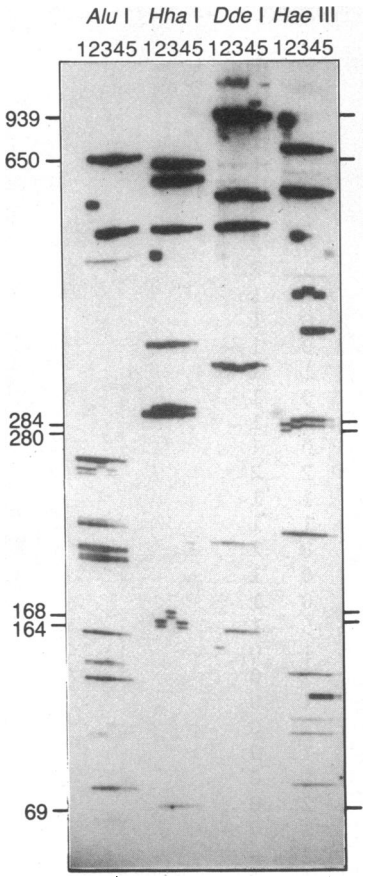

FIG. 1. Autoradiograph of five *D. melanogaster* DNA samples digested with four restriction enzymes and probed with a 2.7-kb homologous probe encompassing the *Adh* locus. The range of fragment sizes in bp is shown at the left. For some fragments (e.g., 164/168 bp), the two DNA strands have slightly different electrophoretic mobilities. A 4-bp insertion/deletion difference in sample 3 can be easily seen in the *Hha* I (164–168 bp) fragments. A 1-bp insertion/deletion difference between samples 2 and 3 can be seen in the *Hae* III (283–284 bp) fragments.

and remain stationary relative to the tube. Hybridization was for 12–18 hr.

The filters were rinsed several times with prewarmed (57°C) wash solution (1 mM Na_2EDTA/40 mM sodium phosphate/1% $NaDodSO_4$) after first removing the probe solution from the tube. Filters were then transferred to a large tub and washed on a rotary shaker either eight times for 5 min (each wash at room temperature) or three times for 30 min (each wash at 57°C; 1500–2000 ml of prewarmed washsolution was used for each wash for four to eight filters). After the last wash, filters were blotted on Whatman 3 MM paper to remove excess liquid, placed in SaranWrap or Seal-a-Meal bags, and autoradiographed for 2–4 days on Kodak XAR film with one intensifying screen.

RESULTS

Fig. 1 shows an autoradiograph of five DNA samples, each digested with *Alu* I, *Dde* I, *Hae* III, and *Hha* I; a 2.7-kb *Adh* probe was used. We were able to reliably score fragments as small as 60–70 base pairs (bp) after 3-day exposures and insertion/deletion differences of only one base. More than 90% of the four-nucleotide-recognizing restriction enzyme sites and essentially every insertion/deletion difference could be detected.

In a random sequence, restriction enzymes that recognize four-nucleotide sequences are expected to cut, on average, every 256 bp, a 16-fold improvement over those that recognize six-nucleotide sequences. Knowledge of the complete sequence of the 2.7-kb *Adh* probe (3) allows direct calculation of the fraction of all possible base changes that would be detected in this sequence given any constellation of enzymes. Assuming that all changes are equally likely, we estimate that ≈19% of all possible changes would be detected (discounting sites producing fragments that would be too small to score—e.g., < 70 bp) with the 10 enzymes used in this study. Thus, all changes at 526 "site equivalents" (0.19 × 2723) are detectable in addition to all insertion/deletion differences within the probed region.

Fig. 2 shows the distribution of polymorphic sites in a sample of 87 isochromosomal lines: 27 from Putah Creek and 60 from Raleigh. Not including the allozyme polymorphism, a total of 28 polymorphisms were scored representing 17 restriction sites and 11 insertions/deletions. Two restriction sites (*Alu* I:1068 and *Dde* I:1551) and two insertions/deletions (655–685 and ≈3320) appear only once in the 87 lines. The remaining polymorphisms are multiply represented.

Fig. 3 shows the restriction, insertion/deletion, and allozyme haplotypes for the 87 lines. Three insertion/deletion sites are assigned more than two lengths. Site 29 contains insertions/deletions located 3′ to the probed region and were detected only in *Dde* I digestions. It is likely that these "alleles" occur at more than one position within the approximately 1000-bp fragment, and some may be combinations of two or more other insertions/deletions. The other two multiallelic sites, sites 23 and 26, are likely to represent variation in the lengths of homonucleotide sequences (3).

The sample of 87 alleles contains a total of 50 haplotypes, 37 in the larger Raleigh sample (n = 60) and 22 in the Putah

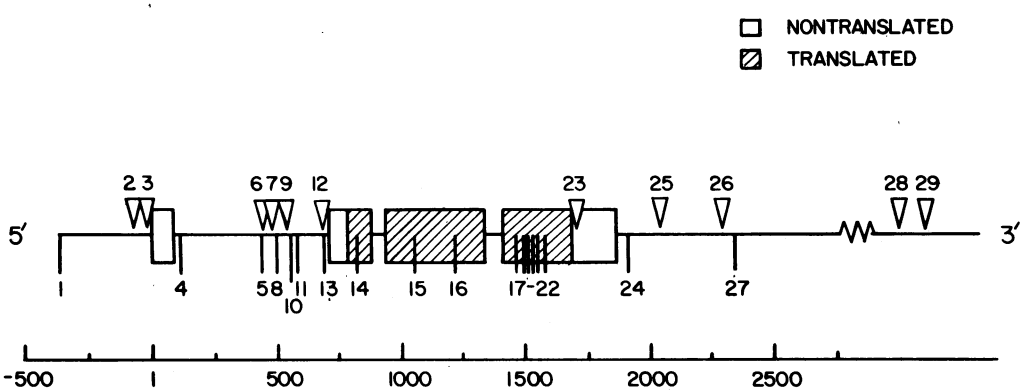

FIG. 2. Distribution of 29 polymorphic sites. Restriction sites are shown below the line and insertions/deletions are shown above the line. Insertions are numbered 3, 6, 7, 9, 12, 23, 25, 26; deletions are numbered 2, 26, 28, 29. Precise locations are given in Table 2.

HAP. #	SITE NUMBER 1 10 20 29	FREQ. Ra	Pu
1:	+ − − + − − + − − + + − + + + + S + − − + + 4 + − − + + − −	0	1
2:	+ − − − − + + + − + + − + + + + F + − − + + 4 + − − + + − 5	0	1
3:	+ − − − − − + + − + + − + + + S − + − + + 4 + − − + + − −	0	1
4:	+ − − − − − + − − + + − + − + + F + − − + + 4 + − − + + − 4	0	1
5:	+ − − + − − + + − + + − + + + + S + − − + + 4 + − − + + − −	8	2
6:	+ − − − + + + + − + + − + − + + F + − − + + 4 + − − + + − −	1	2
7:	+ − − − − − + + − + + − + + + + S − + − + + 3 + − − + + − −	3	2
8:	+ − − − − + + + + + + − + + + + F + − − + + 4 + − − + + − 4	1	2
9:	+ − − − − − + + − − + − − − + + S − + − + − 4 + − − + + − −	0	1
10:	+ − − − − + + + − + + − + − + + F + − − + + 4 + − − + + − 5	0	1
11:	+ − − − − + + − − + + − + + + + F + − − + + 4 + − − + + − −	0	1
12:	+ − − − − + + + + + + − + + + + F + − − + + 4 + − − + + − −	1	1
13:	+ − − + − − + + − + + − + + + + S + − − + + 4 + − − + + − 4	0	1
14:	+ − − + − − + + − + + − + − + + F + − − + + 4 + − − + + − −	0	1
15:	+ − − − − + + + + + + − + + + + F + − − + + 5 + − − + + − 2	0	1
16:	− − − + − − + + − + + − + − + + S − + − + − −1 + − − + + − −	2	2
17:	+ − − − − − + + − + + − + + + + F + − − + + 4 + − − + + − −	1	1
18:	+ − − − − + + + − + − − + + + − F + − − + + 4 + − −1 + + − −	1	1
19:	+ − − − − + + + + + + − + + + + F + − − + + 4 + − − + + − 3	0	1
20:	+ − − − − − + + − − + − − − + + S − + − − − − + − +1 + + − −	0	1
21:	+ − − − + + + + − + + − + + + + F + − + + + 4 + − −1 + + − −	0	1
22:	+ − − − − + + + − + + − + − + + F + − − + + 4 + − − + + − −	5	1
23:	+ − − − − + + − − + + − + − + + F + − − + + 5 + − −1 + + − −	1	0
24:	+ − − + − − − + − + + − + + + + S + − − + + 4 − − − − − − −	2	0
25:	+ − − − − − + + − − + − − − − + S − + − + − − + − +1 + + − −	1	0
26:	+ − − + − − + + − + + − + − + + S + − − + + 4 + − − + + − −	2	0
27:	+ + − + − − − + − + + − + + + + S + − − + + 5 − − − − − − −	1	0
28:	+ − − + − − + + − + + − + − + − S − + − + − − + + − + + − −	1	0
29:	+ − − + − − + + − + + − + + + + F + − − + + 4 + − − + + − −	2	0
30:	+ − − − − − + + − − + − − − + + S + − − + − − + − +1 + + − −	1	0
31:	+ − − + − − − + − + + − + + + + S + − − + + 5 − − − − − − −	1	0
32:	+ − − − − + + − − + + − + − + + F + − − + + 4 + − − + + − −	1	0
33:	+ − − + − − + + − + + − + − + + S − + − + − − + + − + + − −	1	0
34:	+ − − − − − + + − − + − − − + + S − + − + − − + − +1 + + − −	1	0
35:	− − + − − − + + − + + − + − + + S + − − + + 3 + − − + + − −	6	0
36:	+ − − + − − + + − + + − + − + + S + − − + + 4 + − − + + + 6	1	0
37:	+ − − + − − + + − + + − + − + + S + − − + + 4 + − − + + − 1	1	0
38:	+ − − − − + + + − + + − + + + + F + − − + + 5 + − − + + − −	1	0
39:	+ − − + − − + + − + + − + − + + S + − − − + 5 + − − + + − −	1	0
40:	+ − − + − − + + − + + − + − + + F + − − − + 3 + − −1 + + − −	1	0
41:	− − + − − − + + − + + − + − + + S + − − + + 4 + − − + + − −	1	0
42:	+ − − − − − + + − + + − + − + + S − + − + − − + − − + + − −	1	0
43:	+ − − + − − + + − + + − + − + + S − + − + + 3 − − +2 + + − −	1	0
44:	+ + − + − − − + − + + + + + + + S + − − + + 4 − − − − − − −	2	0
45:	+ − + − − + + + − + + − + − + + F + − − + + 4 + − − + + − −	1	0
46:	+ − − + + − + + − + + − + − + + S + − − + + 4 + − − + + − −	1	0
47:	+ − − − − − + + − + + − + + + + S + − − + + 4 + − − + + − −	1	0
48:	+ − + − − − + + − − + − − − + + S − + − + − − + − − + + − −	1	0
49:	+ − − + − − + + − − + + + + + + S + − − + + 4 + − − + + − −	1	0
50:	+ − − − − + + + + + + − + − + + F + − − + + 4 + − − + + − 4	1	0

FIG. 3. Fifty haplotypes (Hap.) identified from 29 polymorphisms in 87 lines. Frequencies (Freq.) in the two samples (Ra, Raleigh; Pu, Putah Creek) are shown to the right of the haplotypes.

Creek sample ($n = 27$). There is an average of 5.75 differences between alleles for the combined sample of 87, including restriction-site and insertion/deletion polymorphisms (Table 1). The two populations are strikingly similar in this respect, with the Raleigh sample having 5.70 and the Putah Creek sample having 5.45 average differences between alleles.

Table 1 also provides a summary for the two Adh allozymes. Consistent with previous evidence suggesting a more recent ancestry for the Adh^F allele (3, 10), the Adh^S allele (average no. of differences = 3.74) has more than twice the average number of restriction-site differences compared to Adh^F (average no. of differences = 1.50). Surprisingly though, the average number of insertions/deletions are essentially identical (1.71 and 1.69 for Adh^S and Adh^F, respectively).

Table 2 gives the individual frequencies for each of the 29 polymorphic sites in the two populations. By using Fisher's exact probability test to compare frequencies, only two sites, 6 and 17, are significant at the 5% level, and one site, 29, is significant at the 1% level. As described above, this latter site may conflate insertions/deletions at several different sites. Site 17 is the Adh^F/Adh^S allozyme polymorphism. The frequency of Adh^S is higher in the Raleigh sample (70%) than in the Putah Creek sample (40%). Other than these two sites, only one additional site shows a significantly different frequency in the two populations. Since, at the 5% significance level, 1 in 20 sites are expected to be significant under the null hypothesis of no difference, the similarity of frequencies at 26 of 29 sites offers clear evidence for genetic homogeneity of the two populations.

Comparison of haplotype frequencies in the two population samples is a potentially more powerful method for identifying genetic differences. The distribution of 87 alleles among the 50 distinct haplotypes in the two populations is shown in Fig.

Table 1. *Adh* locus restriction-site and insertion/deletion (Ins/Del) haplotype variation in 87 lines

Sample	Type of polymorphism		
	Restriction + ins/del	Restriction only	Ins/del only
Ra + Pu (n = 87)			
No. of haplotypes	50	30	30
Mean differences	5.75	3.65	2.1
Haplotype diversity	0.96	0.93	0.89
Ra (n = 60)			
No. of haplotypes	37	24	26
Mean differences	5.70	3.60	2.10
Haplotype diversity	0.95	0.92	0.90
Pu (n = 27)			
No. of haplotypes	22	15	12
Mean differences	5.54	3.61	1.93
Haplotype diversity	0.95	0.90	0.87
Adh^S (n = 53)			
No. of haplotypes	29	20	19
Mean differences	5.45	3.74	1.71
Haplotype diversity	0.92	0.90	0.86
Adh^F (n = 34)			
No. of haplotypes	21	10	13
Mean differences	3.19	1.50	1.69
Haplotype diversity	0.93	0.80	0.84

The average number of differences (Mean differences) is the number of sites segregating between two alleles. Haplotype diversity is calculated as $1 - \Sigma(p_i)^2$ where p_i is the frequency of the ith haplotype. Ra, Raleigh; Pu, Putah Creek; Adh^F, allele for alcohol dehydrogenase-fast, isozymes that migrate faster in electrophoresis than isozymes specified by Adh^S, the allele for alcohol dehydrogenase-slow.

3. The distribution is strongly skewed—35 of the 50 haplotypes are represented only once; 14 haplotypes are multiply represented, 9 of which are present in both population samples. There is no significant difference between the distribution of haplotypes in the two samples ($P > 0.1$). The third most abundant haplotype, 35, is found only in the Raleigh sample. However, because this haplotype is an Adh^S allele, and the number of slow alleles is so much greater in the Raleigh sample (42 vs. 11), it is not unexpected that some Adh^S haplotypes would have no Putah Creek representatives. Therefore, taking the Adh allozyme frequency difference into account, there is a striking similarity in the pattern of haplotype representation in the two samples. This offers additional support for genetic homogeneity of the two populations.

DISCUSSION

Technical Considerations. In preparing the gels, we used a steep buffer gradient to allow greater migration of large fragments without losing small ones. Five percent polyacrylamide gels provide adequate resolution of small fragments and reasonable separation of fragments up to 1000 bp in length. We noticed a tendency for enzyme digestions producing an excess of large fragments (e.g., digestions with enzymes that recognize a six-nucleotide sequence) to resolve poorly, possibly a result of the poor entry of large fragments into the gel matrix. We detected no significant increase in ^{32}P-labeled filter background when probing from one to five filters in as high as 5×10^8 cpm/18 ml of hybridization solution.

The use of Lexan tubes for hybridization makes the technique simpler, safer, and less expensive (requires less probe per filter). We noticed no difference in signal for single filters hybridized in bags or from one to five filters simultaneously hybridized in one tube. This allows hybridization of up to 20 filters at a time in four tubes.

Applicability. The technique with restriction enzymes that recognize a four-nucleotide sequence should be useful not only in evolutionary studies but also in identifying specific haplotypes associated with genetic disease. In our study, most of the variable restriction sites and all of the insertion/deletion sites were detected by using only three or four enzymes. The ability to identify small insertions/deletions using only a few enzymes may be of particular value for genetic screening, since this type of mutation is abundant and in some cases is "multiallelic" (e.g., site 23).

Population Survey. Comparison of the 87 lines reported in this study with 11 sequenced *Adh* alleles from a world-wide collection (3) provides some evidence for genetic differentiation of populations. None of the 10 restriction-site and insertion/deletion haplotypes predicted from the sequences of the 11 alleles are exactly represented in the sample of 87 lines, although 8 are closely related to 1 or more of the 50 haplotypes. However, one distinct slow haplotype, which was represented twice in 6 slow alleles [Fl^{1s} (southern Florida) and Wa^s (Seattle, Washington)], is completely absent in our sample of 53 slow alleles. It is of interest, then, whether this haplotype is completely absent in both populations. Since only three slow alleles were sequenced in the Florida and Washington populations and the haplotype in question appeared twice, it is presumably not rare in these populations.

A distinction must be made between the evidence presented here suggesting a lack of genetic divergence between the two populations and other morphological (11–15) and biochemical (16) evidence suggesting the contrary (although see ref. 17). In this study, except for the allozyme polymorphism, individual sites as well as haplotypes are at similar frequencies in the two populations. This implies either a sufficient migration rate between populations or a sufficiently recent common ancestry to preclude differentiation by genetic drift.

Such factors would not prevent genetic differentiation by natural selection. It is possible, for example, that natural selection is sufficient to maintain genetic differences between populations in spite of interpopulational migration. The difference in allozyme frequency in the two populations may exemplify a polymorphism being subject to natural selection.

The analysis of haplotype frequencies contains information about population structure that is not contained in the analysis of individual nucleotide site frequencies. In the present case, by comparing haplotypes within allozymes, we can essentially ignore the possibility of natural selection acting on the Adh allozymes. Thus, the similarity in haplotype frequencies within allozymes suggests that the allozyme frequency difference between the two populations is a consequence of natural selection rather than a founder effect or genetic drift. This leads to the specific prediction that, in populations where Adh allozyme frequencies differ, such as at the geographic extremes of Adh allozyme clines (16), the same haplotypes will be represented within each allozyme class. Because of linkage disequilibrium between the allozyme polymorphism and other sites, the same is not necessarily true on a site-by-site basis.

For questions such as ones we address here, direct sequence information would not be expected to reveal quantitatively or qualitatively different results. The 29 polymorphic sites in this study can be considered a random sample of a larger number of polymorphisms within the region. Although some of the unidentified polymorphisms might have different frequencies in the two populations, only 3 of 29 sites in this study are statistically different. Therefore, there is no reason to expect a complete sequence comparison to reveal a pattern of genetic differentiation different from the restriction analysis with enzymes that recognize four-nucleotide sequences.

However, restriction polymorphism studies will not replace the need for direct DNA sequencing. In fact, our prior

Table 2. Allele frequencies at 29 sites in Raleigh (Ra) and Putah Creek (Pu) samples

Site	Position	Allele	Pu	Ra	Site	Position	Allele	Pu	Ra
1	−349 to −346	+	25	51	18	1518	+	20	48
		−	2	9			−	7	12
2	−62 to 69	+	0	3	19	1527	+	7	12
		−	27	57			−	20	48
3	−62 to 69	+	0	9	20	1551	+	1	0
		−	27	51			−	26	60
4	105	+	7	29	21	1563–1566	+	26	58
		−	20	31			−	1	2
5	423	+	3	2	22	1596	+	23	51
		−	24	58			−	4	9
6*	448	+	13	14	23	1698	−1	2	2
		−	14	46			0	1	7
7	497	+	0	6			+3	2	11
		−	27	54			+4	21	35
8	502–505	+	24	58			+5	1	5
		−	3	2	24	1925	+	27	53
9	551	+	5	3			−	0	5
		−	22	57	25	2081	+	0	2
10	571–574	+	25	55			−	27	58
		−	2	5	26	2303	−1	2	3
11	586	+	26	59			−	24	53
		−	1	1			+1	1	3
12	655–685	+	0	1			+2	0	1
		−	27	59	27	2348–2351	+	27	54
13	687–690	+	25	56			−	0	6
		−	2	4	28	≈3320	+	0	1
14	816	+	15	25			−	27	59
		−	12	35	29†	3′	−	19	56
15	1068	+	27	59			+1	0	1
		−	0	1			+2	1	0
16	1235	+	26	58			+3	1	0
		−	1	2			+4	4	2
17*	1490	S	11	42			+5	2	0
		F	16	18			+6	0	1

Frequency of polymorphism at 29 sites in Putah Creek (Pu) and Raleigh (Ra) samples; + and − show the presence or absence of restriction-site and insertion or deletion variation. Multiple insertions or deletions assigned to the same site are numbered corresponding to the relative size of the insertion or deletion. Positions of polymorphisms within the *Adh* region sequence are the same as in ref. 3. S, slow; F, fast.
*Significant at $P < 0.05$.
†Significant at $P < 0.01$.

knowledge about sequence polymorphism of *Adh* aided us in the analysis of the autoradiographs. For many evolutionary questions, sequencing remains the method of choice. In such cases, we advocate a stratified approach in which large samples are first subdivided into smaller groups of related haplotypes based on analysis with restriction enzymes that recognize four-nucleotide sequences, and these groups then are sampled for subsequent sequence analysis. In this regard, Hudson and Kaplan (18) have made progress in describing the expected neutral frequency distribution of mutation under such a sampling scheme.

We thank George Church and Harry Nick for technical help, Walter Gilbert for providing laboratory space and supplies during preliminary work, Fred Cohan and Chip Aquadro for flies, and also D. Lewis, L. Skins, C. Langley, R. Hudson, and S. Sprawls. This work was supported by National Institutes of Health Grant GM29301 to R. C. Lewontin.

1. Langley, C. H., Montgomery, E. & Quattlebaum, W. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5631–5635.
2. Cooper, D. N. & Schmidtke, J. (1984) *Hum. Genet.* **66**, 1–16.
3. Kreitman, M. (1983) *Nature (London)* **304**, 412–417.
4. Church, G. M. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1991–1995.
5. Lindsley, D. L. & Grell, E. H. (1968) *Genetic Variations of Drosophila melanogaster* (Carnegie Institution of Washington, Washington, DC), Publ. 627.
6. Kreitman, M. (1981) *Genetics* **95**, 467–475.
7. Bender, W., Spierer, P. & Hogness, D. S. (1983) *J. Mol. Biol.* **168**, 17–33.
8. Biggins, M. D., Gibson, T. J. & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3963–3965.
9. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
10. Bodmer, M. & Ashburner, M. (1984) *Nature (London)* **309**, 425–430.
11. Teissier, G. (1958) *Ann. Genet.* **1**, 2–10.
12. Petit, C. (1969) *Bull. Soc. Zool. Fr.* **94**, 641–648.
13. David, J. R. & Bocquet, C. (1975) *Experientia* **31**, 164–166.
14. David, J. R. & Bocquet, C. (1975) *Nature (London)* **257**, 588–590.
15. Takanashi, E. & Kitagawa, O. (1977) *Drosoph. Inf. Serv.* **52**, 28–29.
16. Oakeshott, J. G., Gibson, J. B., Anderson, P. R., Knibb, W. R., Anderson, D. G. & Chambers, G. (1983) *Evolution* **36**, 86–96.
17. Henderson, N. R. & Lambert, D. M. (1982) *Nature (London)* **300**, 437–440.
18. Hudson, R. & Kaplan, N. (1986) *Genetics*, in press.