Molecular Variation of *Adh* and *P6* Genes in an African Population of *Drosophila melanogaster* and Its Relation to Chromosomal Inversions

Véronique Bénassi, Sylvie Aulard, Suzanne Mazeau and Michel Veuille

Laboratoire de Biologie et Génétique Evolutives, Centre National de la Recherche Scientifique, 91198 Gif-sur-Yvette cedex, France

Manuscript received November 12, 1992
Accepted for publication March 17, 1993

ABSTRACT

Four-cutter molecular polymorphism of *Adh* and *P6*, and chromosome inversion polymorphism of chromosome *II* were investigated in 95 isogenic lines of an Ivory Coast population of *Drosophila melanogaster*, a species assumed to have recently spread throughout the world from a West African origin. The *P6* gene showed little linkage disequilibrium with the *In(2L)t* inversion, although it is located within this inversion. This suggests that the inversion and the *P6* locus have extensively exchanged genetic information through either double crossover or gene conversion. Allozymic variation in ADH was in linkage disequilibrium with *In(2L)t* and *In(2R)NS* inversions. Evidence suggests either that inversion linkage with the *Fast* allele is selectively maintained, or that this allele only recently appeared. Molecular polymorphism at the *Adh* locus in the Ivory Coast is not higher than in North American populations. New haplotypes specific to the African population were found, some of them connect the "Wa^S-like" haplotypes found at high frequencies in the United States to the other *slow* haplotypes. Their relation with *In(2L)t* supports the hypothesis that Wa^S recently recombined away from an *In(2L)t* chromosome which may be the cause of its divergence from the other haplotypes.

THE evolutionary genetics and variability of the *alcohol dehydrogenase* gene (*Adh*) in *Drosophila melanogaster* have been the subject of extensive investigation [reviewed by CHAMBERS (1988)]. The 2-kb *Adh* locus is organized into three exons separated by two small introns, linked by a longer intron to the leader. It codes for a 256-amino acid protein. Two *Adh* transcripts are alternatively produced from either a distal (adult) or a proximal (larval) promoter. Two main allozymes, ADH^F and ADH^S are found around the world. The frequency of these alleles show latitudinal clines in both hemispheres: Adh^F increases in frequency with increasing distance from the equator and has a higher enzyme activity [reviewed by CHAMBERS (1988)]. This distribution is generally considered to result from selection. VIGUE and JOHNSON (1973) argue that ADH^S is favored in hotter climates due to its greater thermostability. ADH^F, being more active, is thought to be responsible for the observed tolerance of northern populations to ethanol.

LANGLEY, MONTGOMERY and QUATTLEBAUM (1982) initiated molecular polymorphism surveys of the *Adh* gene, studying a large region around *Adh* using six-cutter restriction enzymes. Subsequent studies sequenced 11 *Adh* alleles from different origins (KREITMAN 1983) and carried out extensive four-cutter restriction enzyme polymorphism studies of American (KREITMAN and AGUADÉ 1986a; SIMMONS *et al.* 1989) and Spanish populations (AGUADÉ 1988). These studies revealed an excess polymorphism of the coding region in the vicinity of the F/S polymorphic site, as predicted under a selectively balanced allozyme polymorphism hypothesis (KREITMAN and AGUADÉ 1986b). KREITMAN and HUDSON (1991) showed that the excess silent polymorphism is within the Adh^S class, not between the allozyme classes. They pointed to the existence of two Adh^S lineages. One of them, "Washington-slow" (Wa^S) is found at high frequencies in some American populations and has no intermediate haplotypes to connect it to the other "*Adh*-slow" (Adh^S) alleles. They put forward two explanations. First, according to a genetic isolation hypothesis, Wa^S would have recombined away from a chromosome carrying a recombination-suppressing system such as the *In(2L)t* inversion. Second, according to a geographic isolation hypothesis, *D. melanogaster* would have been subdivided into two groups of populations in the past, one of them being in West Africa.

The study of African *D. melanogaster* allows us to address some of the questions raised by previous studies. Little genetic data are available from African populations, despite the presumed Tropical African origin of this species (TSACAS and LACHAISE 1974). According to LACHAISE *et al.* (1988) an initial northward spread from Africa to Europe was followed by a worldwide dispersion. *D. melanogaster* was presumably introduced to Tropical America from Africa some centuries ago. From biometrical and allozymic data, DAVID and CAPY (1988) distinguish three kinds of *D. melanogaster* populations: "ancestral" (African), "ancient" (European and Asian), and "new" populations (Australian and North American). Mitochondrial

DNA variation studies by HALE and SINGH (1991) are in agreement with this scheme, although mtDNA data cannot distinguish between European and African samples.

An important feature of West African populations is the high frequency of the *In(2L)t* inversion in natural populations. Substantial linkage disequilibrium between Adh^S and *In(2L)t* has been reported in natural populations elsewhere in the world, when the frequency of *In(2L)t* is high enough to allow statistical testing [reviewed by LEMEUNIER and AULARD (1992)]. Significant gametic disequilibrium is also observed between the *In(2R)NS* inversion and *Adh*, which are on opposite arms of the second chromosome (MUKAI and VOELKER 1977). However, this correlation is not found in all studied localities and is not stable over time (KNIBB 1982).

A new approach to four-cutter restriction mapping based on polymerase chain reaction (PCR) enabled us to study DNA polymorphism in two genes, *Adh* and *P6* in a West African population from Lamto in the Ivory Coast. This population is composed of 17.6% Adh^F, 61.4% *In(2L)t* and 21.0% *In(2R)NS*.

The purpose of this study is to document genetic variation of *Adh* in Africa. This study will address several interrelated questions. Is *Adh* molecular polymorphism different between West Africa and other parts of the world? Can it help us understand the history of *D. melanogaster* and resolve unanswered questions about *Adh* polymorphism? Is the ADH allozyme polymorphism in linkage disequilibrium with *In(2L)t* in West African populations as it has been found in previous studies?

This study includes the survey of molecular polymorphism in *P6*, a gene characterized and sequenced in *D. melanogaster* (RAT, VEUILLE and LEPESANT 1991) in *Drosophila pseudoobscura* (MEGHLAOUI and M. VEUILLE, unpublished results) and in *Sarcophaga peregrina* (MATSUMOTO *et al.* 1985). It is considered to be related to *Adh* by an ancient duplication event (RAT, VEUILLE and LEPESANT 1991). It is located on the same chromosome arm (*2L*) as *Adh*, in the middle of the region covered by the *In(2L)t* inversion. *P6* is organized into two exons separated by an intron. Although the catalytic activity of P6 is unknown, both ADH and P6 have the structure of short chain dehydrogenases. Both of them are expressed at high level in the larval fat body. Both genes have a strong codon bias for G + C content, and therefore probably the same constraints at the level of silent polymorphism. Although located on the same chromosome, which allows simultaneous sampling, they are far enough apart to allow wide recombinational exchange. These properties enabled us to use *P6* as a control of molecular variation in *Adh*. The proteins encoded by these genes are likely to be subject to the same structural constraints, but to different selective pressures. In other words, conservative selection would be the same in both proteins, while selection by the environment would be different since only ADH is known to be involved in the degradation of ethanol.

MATERIALS AND METHODS

Drosophila stocks: The sample was collected in Lamto (Ivory Coast) by D. LACHAISE in June 1989. Isogenic lines for chromosome *II* were obtained through a classical crossing procedure using the Cy^o of the Cy^o/ap^{Xa} stock as a balancer. The sample is composed of two kinds of lines referred to hereafter as samples E1 and E2. (1) E1 is a random sample of 88 lines. Lines L1–L57 were each established from a wild-caught male; lines L58–L88 were established from isofemale lines collected at the same time and kept for 1 year in the laboratory. No significant difference was found between lines L1–L57 and L58–L88. (2) E2 is an additional nonrandom sample of ADH^F lines, not included in statistical testing. These lines (LF1–LF10) were established from isofemale lines to provide additional qualitative information on the ADH^F haplotypes, since this allele was at a low frequency in the random sample.

Lines L31–L57, L74–L88 and LF6–LF10 were homozygous lethal and were maintained as balanced heterozygotes with Cy^o lines. Due to amplification failure, lines L2, L65 and L82 were only studied for *P6* and L20, L24 and L72 only in *Adh*. Failure of amplification in some lines may result from mismatch at primer annealing. The E1 sample was therefore composed of 85 lines for both *P6* and *Adh*, with 82 shared chromosomes.

DNA extraction. Three hundred flies from each line were frozen in liquid nitrogen and homogenized in 100 mM Tris-Cl (pH 8.0), 50 mM NaCl, 50 mM EDTA, proteinase K (0.1 mg/ml), 0.15 mM spermine, 0.5 mM spermidine, 1% sodium dodecyl sulfate, and incubated for 2 hr at 37°. The homogenate was extracted twice with phenol, twice with phenol/chloroform/iso-amyl-alcohol and once with chloroform/iso-amyl-alcohol. It was precipitated with ethanol (2 volumes) and 5 M NaCl (1/25 volume). The pellet was resuspended in 10 mM Tris-HCl (pH 7.5), 5 mM EDTA, 300 mM NaCl and treated by RNase (100 μl/ml) for 1 hr at 37°. The DNA was again phenol and chloroform extracted, precipitated, and stored in 1 × TE.

PCR amplification: DNA amplifications were carried out using 20-mers designed from standard *Adh* (KREITMAN 1983) (GenBank/DROADHA) and *P6* (RAT, VEUILLE and LEPESANT 1991) sequences. They encompassed the whole transcription unit and the TATA-box (distal and proximal TATA-boxes for *Adh*). Amplifications were carried out from 100 ng genomic DNA, according to the supplier's instructions (Promega), except for the addition of 0.5 μl of [α-^{35}S]dATP (>1000 mCi/mMol). For *P6* the primers were: 5′-CGATAACAGCTGGCCTCCAA-3′; 5′-CGACGCCAGCTAATGCAT-3′. For *Adh* the primers were: 5′-TGTGGCCCTACACTGTAAGA-3′; 5′-ACTGATTGGCGGAGGAACTT-3′. The amplified product was phenol extracted, ethanol precipitated and resuspended in 1 × TE.

Restriction enzyme polymorphism: A sample of 100 ng of the amplified DNA was digested overnight using the following enzymes: *Alu*I, *Ban*I, *Dde*I/*Bam*HI (*Adh*), *Dde*I/*Pst*I (*P6*), *Hae*III, *Hha*I, *Msp*I, *Sau*3A, *Sau*96I and *Taq*I, according to the supplier's instructions (Boehringer). Each sample was mixed with 6 μl denaturating buffer (10 mM Na_2EDTA (pH 7.2), 94% formamide, 0.05% xylene cyanol, 0.05% bromphenol blue), heated for 3 mn at 95° and loaded

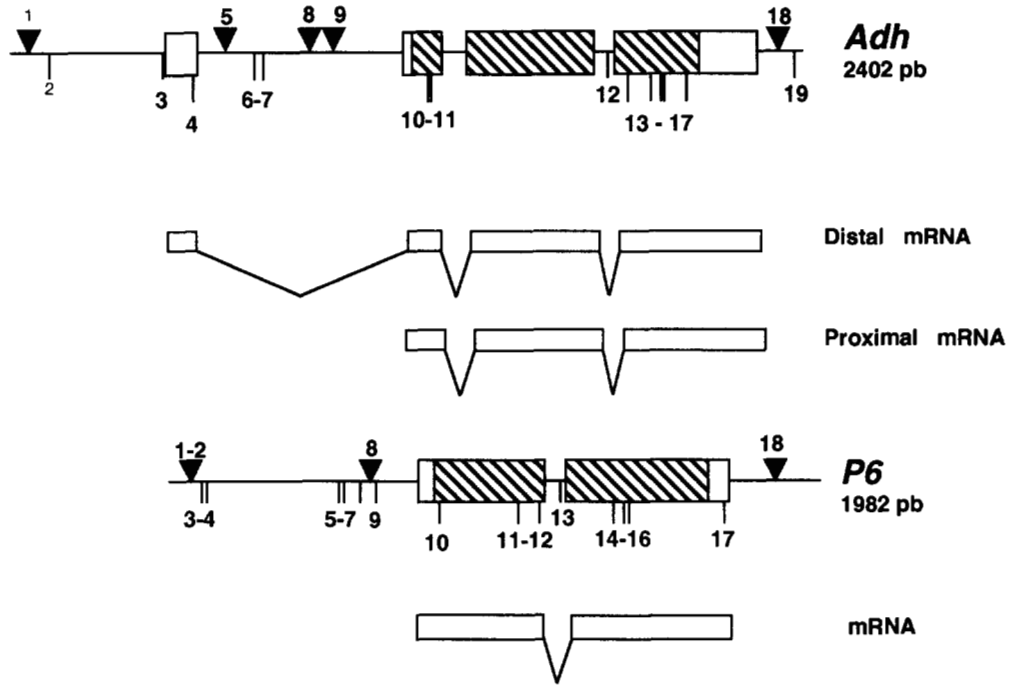

FIGURE 1.—Distribution of polymorphic sites in *Adh* and *P6* in Lamto. Insertion/deletions are represented by a triangle above the line. Polymorphic sites are shown below the line. Hatched regions represent the exons.

on a (30 cm × 40 cm × 0.4 mm) 5% polyacrylamide/7 M urea/TBE sequencing gel. Gels were run at 1600 V for 1 hr 30 min, fixed (10% methanol, 10% acetic acid), dried and exposed onto a Kodak X-O-Mat-AR film.

This technique allowed us to simultaneously score restriction fragments ranging from 50 to 1000 bp in a number of lines. The use of an already available reference sequence allowed us to locate polymorphic sites. Uncertain cases were resolved by restriction mapping.

For each line the fast or slow allele of ADH was determined by electrophoresis. A single fly was ground in 10 μl loading buffer (1 × TEB = 10.54 g Tris base, 0.54 g boric acid, 0.84 g tetrasodium EDTA per liter), 5% sucrose and bromphenol blue. Eight microliters were loaded onto a 7% acrylamide/1 × TEB gel and run at 150 V at 10° for 3 hr. Gels were stained with 50 mg NAD^+, 40 mg NBT, 9 ml isopropanol, and 5 mg PMS in 200 ml 50 mM Tris, pH 8.6.

Inversions: Males from each line were mated to virgin Canton-S females homozygous for standard chromosomes. Salivary glands from the F_1 progeny were dissected and chromosomes were stained using lacto-acetic orcein. Breakpoints of the inversions were established by comparing photographs with standard maps (LEFEVRE 1976).

RESULTS

Restriction polymorphism: A summary of restriction polymorphisms in the *Adh* (35B) and *P6* (30B) regions of the 95 second chromosomes from the Ivory Coast is shown in Figure 1. Tables 1 and 2 show the restriction and insertion/deletion haplotypes in the two samples. Nineteen polymorphisms were scored for *Adh*, corresponding to 12 restriction sites, 6 insertions/deletions and the allozyme polymorphism site. Two restriction sites and one insertion/deletion appeared only once in the 95 lines analyzed. An insertion (∇1) first described by KREITMAN (1983) and shown to act on adult ADH protein levels (LAURIE, BRIDGHAM and CHOUDHARY 1991), was found in all fast lines. Another large 37-bp insertion (∇2), previously found by KREITMAN (1983), appeared in five out of the 25 fast lines. This insertion is reportedly involved in larval ADH protein levels (MATTHEW *et al.* 1992). These two insertions were not found in slow lines from this sample. For *P6*, 18 polymorphisms were mapped, including 14 restriction sites (three being found only once) and four insertions/deletions. The exact positions of the *Taq*I 99 and *Taq*I 847 sites were characterized by double digestion. Insertions and deletions from 1 to 10 bp were scored.

Inversions: The random samples yielded a high frequency of paracentric inversions. Three cosmopolitan inversions, one recurrent endemic inversion and one unique endemic inversion were found:

- *In(2L)t*, 22DE;34A, a common cosmopolitan inversion, was found in 51 lines (61.4%) from sample E1.
- *In(2R)NS*, 52AB;56F, a common cosmopolitan inversion, was present in 17 lines (22%) from E1.
- *In(2R)*, 49B;56A, a recurrent endemic inversion recorded from the Ivory Coast, Benin, Gabon, Congo, Seychelles and Cameroon (AULARD 1990), was found in two E1 lines.
- *In(2L)NS*, 23E;35F, a rare cosmopolitan inversion, was observed in two E1 lines.
- *In(2R)*, 44E;53F is a new inversion found in one E1 line.

Inversion-carrying lines are listed in Tables 1 and 2.

TABLE 1

***Adh* haplotypes**

	I/D1	*Msp*I −346	*Dde*I 9	*Ban*I 102	I/D2	*Hae*III 287	*Dde*I 321	∇1	∇2	*Hae*III 816	*Msp*I 818	*Sau*3AI 1354	*Sau*96I 1424	F/S 1490	*Dde*I 1518	*Dde*I 1527	*Alu*I 1596	I/D5	*Hae*III 1925	*In(2L)t* In	*In(2L)t* St	*In(2R)NS* In	*In(2R)NS* St	Other inv.	E1	*N*
IC1	0	1	1	1	0	1	0	0	0	0	1	1	1	0	0	1	0	1	0	1			1	1d	1	1
IC2	0	1	1	0	0	1	0	0	0	0	1	1	1	0	1	0	0	1	1	1		1			1	1
IC3	0	1	1	0	0	1	0	0	0	0	0	1	1	0	1	0	0	1	1	3		3			3	3
IC4	0	1	0	0	0	1	1	0	0	0	0	1	1	0	1	0	0	1	1	1		1			1	1
IC5	0	1	1	1	0	1	0	0	0	0	1	0	0	0	0	1	1	0	1	1			1		1	1
IC6	0	0	1	0	1	1	0	0	0	0	1	1	1	0	0	1	0	0	0	3	1		4	1b	4	4
IC7	0	1	1	0	0	1	0	0	0	0	1	1	1	0	1	0	1	1	1	8			8		8	8
IC8	0	1	1	1	0	1	0	0	0	0	1	1	0	0	0	1	0	0	1	2		1	1		2	2
IC9	0	1	1	0	1	1	0	0	0	0	1	1	1	0	0	1	1	0	0	1			1		1	1
IC10	0	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1			1		1	1
IC11	0	1	1	1	0	1	0	0	0	0	1	1	1	0	1	0	0	0	1	1			1		1	1
IC12	0	0	1	1	0	1	0	0	0	0	1	1	1	0	1	0	1	1	1	3			3	1a	4	4
IC13	0	1	1	1	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1	10		2	8		10	10
IC14	0	0	1	0	0	1	0	0	0	0	1	1	1	0	1	0	1	1	1	2		1	1		2	2
IC15	0	0	1	0	1	1	0	0	0	0	1	1	1	0	0	1	1	0	0	1			1		1	1
IC16	0	1	1	1	0	1	0	0	0	0	1	1	1	0	0	1	0	0	1		6		6		6	6
IC17	0	1	1	1	0	1	0	0	0	0	1	1	1	0	0	1	0	1	1		1		1		1	1
IC18	0	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	1	0	1		1	1			1	1
IC19	1	1	1	0	0	1	0	0	0	1	1	1	1	0	0	1	0	0	0		1		1		1	1
IC20	0	1	1	1	0	1	0	0	0	1	1	1	1	0	0	1	0	1	1		1		1		1	1
IC21	0	1	1	1	0	1	0	0	0	0	1	0	0	0	0	1	0	1	1		1		1		1	1
IC22	0	1	1	1	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1		1		1		1	1
IC23	0	1	1	1	0	1	0	0	0	0	1	1	1	0	1	0	1	1	0		1		1		1	1
IC24	0	1	1	1	0	1	0	0	0	0	1	1	1	0	0	1	0	0	0	3	1		4		4	4
IC25	0	1	1	1	0	1	0	0	0	0	1	1	1	0	1	0	1	1	1	9	2	1	10	1a	12	12
IC26	0	1	1	0	0	1	0	1	0	0	1	1	1	1	1	0	1	1	1		10+8*	2+4*	8+4*	1c	10	18
IC27	0	1	1	0	0	1	0	1	0	1	1	1	1	1	0	1	1	1	1		1	1	3+2*		1	1
IC28	0	1	1	0	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1		3+2*				3	5
IC29	0	1	1	0	0	1	0	1	0	0	1	1	1	1	1	0	0	1	1		1		1	1b	1	1

1, Presence of a restriction site or insertion/deletion (I/D), or *Fast* allele. For consistency with other studies, polymorphic sites were numbered according to KREITMAN (1983). ∇1 and ∇2 correspond to KREITMAN (1983) insertions. E1, number of lines by sample for each haplotype; *N*, total number; *E2 sample. For *In(2L)t* and *In(2R)NS* inversions, the number of lines with inversion (In) or Standard (St) are indicated. Other inversions: (a) not determined; (b) *In(2L)NS*; (c) *In(2R)* 49B;56A; (d) *In(2R)* 44E;53F.

TABLE 2

***P6* haplotypes**

	I/D1	I/D2	*Sau*3A 91	*Taq*I 99	*Taq*I 527	*Alu*I 533	*Taq*I 588	I/D3	*Msp*I 641	*Taq*I 847	*Sau*96I 1090	*Hae*III 1155	*Dde*I 1222	*Sau*96I 1390	*Msp*I 1424	*Hha*I 1486	*Hae*III 1736	I/D4	*In(2L)t* In	*In(2L)t* St	*In(2R)NS* In	*In(2R)NS* St	Other inver.	E1	*N*
P1	0	0	1	0	1	1	0	0	0	0	0	0	0	0	1	1	1	0	7	5+1*	1	11+1*		12	13
P2	0	0	1	0	1	1	0	0	1	0	0	0	0	0	1	1	1	0	2	5	2	5	1c	7	7
P3	0	0	1	0	1	1	0	0	0	0	0	0	1	0	1	1	1	0	2	4	3	3		6	6
P4	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	1	0	1	2	1	2		3	3
P5	0	0	1	1	1	1	0	0	0	0	0	0	1	0	1	1	1	0	8	1		9	1c	9	9
P6	1	0	1	1	1	1	0	0	0	0	0	0	1	0	1	1	1	0	2	1*		2+1*		2	3
P7	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	1	1		2		2	2
P8	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	1	0	1			1		1	1
P9	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	0	3			3		3	3
P10	0	0	1	1	0	1	0	0	1	0	0	0	0	0	0	1	1	0	1	1		2	1b	2	2
P11	0	0	1	0	1	1	0	0	0	0	1	0	0	0	0	1	1	0	3			3		3	3
P12	1	0	1	0	1	1	0	0	0	0	0	0	1	0	1	1	1	0	1			1		1	1
P13	0	0	0	0	1	1	0	0	1	0	0	0	1	0	1	1	1	0	1			1	1a	2	2
P14	1	0	0	0	0	1	0	0	1	0	0	0	0	0	1	1	1	0	3			3	1d	3	3
P15	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	1			1		1	1
P16	0	0	1	1	1	1	0	0	0	0	0	0	1	0	1	1	0	0	1			1		1	1
P17	0	0	1	0	1	1	0	0	1	0	0	0	0	0	1	1	0	1	1			1		1	1
P18	0	0	1	0	1	1	0	0	0	0	0	0	0	1	0	1	1	0	1			1		1	1
P19	0	0	1	0	1	1	0	0	0	0	0	1	0	0	1	1	1	0	1			1		1	1
P20	0	0	1	0	1	1	0	0	0	0	0	0	0	0	1	1	0	1	1			1		1	1
P21	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0	1	1	0	3		2	1		3	3
P22	0	0	0	0	1	1	0	0	1	0	0	0	0	0	1	0	1	0	2		1	1		2	2
P23	0	0	1	1	1	1	0	0	1	0	0	0	0	0	1	1	1	0	1		1			1	1
P24	0	0	0	0	1	1	0	0	1	0	0	0	0	0	1	1	1	0	1		1		1a	2	2
P25	1	0	0	0	1	1	0	0	1	0	0	0	0	0	0	1	1	0	1		1			1	1
P26	1	0	1	0	1	1	0	0	1	0	0	0	0	0	0	1	1	0		1*		1*			1
P27	0	1	1	0	1	1	0	0	1	0	0	0	0	0	1	1	0	0		1*		1*			1
P28	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	1	1	0		2*		2*			2
P29	1	0	1	0	0	1	0	1	1	1	0	0	0	0	0	1	1	0		2*		2*			2
P30	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0		1*		1*			1
P31	1	0	1	0	0	1	0	1	1	0	0	0	0	0	0	1	1	0		1*		1*			1
P32	0	0	0	0	1	1	0	0	1	0	0	0	0	0	1	1	0	1		1		1		1	1
P33	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0		1		1		1	1
P34	0	1	1	0	0	1	0	1	1	1	0	0	0	0	0	1	0	0		1		1		1	1
P35	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0		1		1		1	1
P36	1	0	0	0	0	1	0	1	1	1	0	1	0	0	0	1	1	0		1		1		1	1
P37	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	1	1	0		1		1		1	1
P38	0	0	1	0	1	1	0	0	0	0	0	0	0	0	1	1	0	0		1		1		1	1
P39	0	0	1	0	1	1	0	0	0	0	0	0	1	0	1	1	0	0		1	1			1	1
P40	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0		1	1			1	1
P41	1	0	0	0	1	1	0	0	1	0	0	0	0	0	1	1	0	0		1	1			1	1
P42	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	1	1	0		4	1	3		4	4

1, Presence of a restriction site or insertion/deletion (I/D). E1, number of lines for sample E1 for each haplotype; *N*, total number; *E2 sample. For *In(2L)t* and *In(2R)NS* inversions, the number of lines with inversion (In) or Standard (St) is given. Other inversions: (1) not determined lines; (b) *In(2L)NS*; (c) *In(2R) 49B;56A; (d) In(2R)* 44E;53F.

TABLE 3

Haplotype diversity and heterozygosity

	n	*D*	*H*	*HL*	*HU*
Adh					
Total	85	0.936	0.005	0.0029	0.0115
Adh^F	15	0.542	0.002	0.0007	0.0980
Adh^S	70	0.926	0.005	0.0027	0.0116
t-Adh^S	51	0.903	0.005	0.0026	0.0122
St-Adh^S	17	0.870	0.006	0.0028	0.0180
P6					
Total	85	0.953	0.0074	0.0041	0.0160
t-*P6*	50	0.947	0.0068	0.0033	0.0156
St-*P6*	33	0.937	0.0072	0.0036	0.0185

Haplotype diversity was calculated as $D = (n/(n-1)(1-\sum(p_i)^2)$ where p_i is the frequency of the ith haplotype and n, the sample size. The range of H values was calculated according to KREITMAN and HUDSON (1991). This allows one to calculate the smallest (*HL*) and the largest (*HU*) values of H that are compatible with the observed number of polymorphic sites at the 0.005 probability level.

All Adh^F lines were Standard for the left arm of chromosome 2 both in the random and nonrandom samples. For this reason, the ADH enzyme polymorphism was in significant linkage disequilibrium with the *In(2L)t* inversion (Fisher exact test = 6.4×10^{-7}). It was also in significant linkage disequilibrium with the *In(2R)NS* inversion (Fisher exact test = 1.6×10^{-2}). In this case, the *Fast* allele was more frequently associated with the inversion.

Haplotype diversity and heterozygosity: Out of 85 chromosomes tested, 29 *Adh* restriction polymorphism haplotypes were found (17 being unique in the sample). The random sample (E1) contained 70 slow lines and 15 fast lines. Ten additional fast lines were described (E2) corresponding to previously found haplotypes; they were not considered in subsequent tests and calculations made hereafter. Thirty-six *P6* restriction polymorphism haplotypes were observed in sample E1: 20 haplotypes were unique. In the E2 sample six new haplotypes were described. Heterozygosity (H) and haplotype diversity (D) were estimated separately for *Adh*, Adh^F, Adh^S, Standard-Adh^S, *In(2L)t*-Adh^S, *P6*, Standard-*P6* and *In(2L)t*-*P6* chromosomes (Table 3).

Heterozygosity in *Adh* was consistent with levels found in previous surveys of *D. melanogaster* American populations ($H = 0.006$, LANGLEY, MONTGOMERY and QUATTLEBAUM 1982; KREITMAN 1983; AQUADRO *et al.* 1986; KREITMAN and AGUADÉ 1986b; $H = 0.004$, BIRLEY 1984; $H = 0.003$ for St-Adh^F and $H = 0.006$ for St-Adh^S, AGUADÉ 1988), Australian populations ($H = 0.008$, JIANG and GIBSON 1992), Chinese populations ($H = 0.005$, JIANG and GIBSON 1992), a Dutch population ($H = 0.007$, CROSS and BIRLEY 1986), an Australian population ($H = 0.005$, CROSS and BIRLEY 1986) and Japanese populations ($H = 0.005$, TAKANO, KUSAKABE and MUKAI 1991). These studies were carried out on regions of different size and they used different methods: four-cutter or six-cutter restriction mapping, and sequencing. This explains the variation in these estimates. The Lamto sample, although taken from the presumed area of origin of *D. melanogaster*, did not show a higher value. In agreement with previous studies, most of this variation is attributable to the *Slow* allele. In all studies, including this one, *Adh Fast* lines show a lower variability than the *Slow* lines.

Heterozygosity in *Adh Slow* reaches similar levels in standard ($H = 0.005$) and in *In(2L)t* ($H = 0.006$) chromosomes, as previously found in Spanish populations (AGUADÉ 1988).

Heterozygosity in *P6* is similar to that of *Adh*. Again, levels are comparable in standard ($H = 0.007$) and in *In(2L)t* ($H = 0.007$) chromosomes.

Tajima "*D* test": A general test, or "*D* test," has been proposed by TAJIMA (1989) to test such data. It assumes no recombination and no mutation. The D values were -0.003 for *P6* and 0.911 for *Adh*. KREITMAN and HUDSON (1991) found $D = 0.433$ for the *Adh* 5′-flanking region and $D = -0.008$ for the *Adh* gene. TAKANO, KUSAKABE and MUKAI (1991) found $D = 1.39$ and $D = 1.22$ in two Japanese populations. These values are consistently higher than expected, but in no case are they significant. It should be noted that the Tajima test is conservative since it uses the stochastic variance in the denominator.

Linkage disequilibrium: Linkage disequilibrium between pairs of polymorphic restriction sites where alleles occurred more than once was calculated using Fisher's exact test in sample E1 (Figure 2). The *Adh* restriction pattern showed a highly linked structure. A number of restriction sites were correlated, especially a group of five (*Sau*3AI 1354, *Sau*96I 1424, F/S 1490, *Dde*I 1518, *Dde*I 1527) centered on the F/S site, and encompassing 243 nucleotides. *P6* presented few associated sites, with no apparent grouping.

The *Adh* F/S site was in highly significant linkage disequilibrium with the *In(2L)t* inversion. Two sites, *Msp*I −346 and *Hae*III 816 showed significant disequilibrium with *In(2L)t*, but not with F/S. None of the restriction polymorphisms surrounding the F/S site were in significant linkage disequilibrium with either *In(2L)t* or *In(2R)NS*. This means that DNA variation around the *Adh* locus is randomized between standard and inversion bearing chromosomes. Insertion ∇1 from the leader intron showed significant linkage disequilibrium both with the inversion and F/S: all fast lines possessed the ∇1 insertion, which was absent in slow lines. The ∇2 insertion, which occurs only in Fast chromosomes, was in significant linkage disequilibrium with both inversions.

Only one P6 restriction site (*Taq*I 99) was significantly linked to *In(2L)t*. Two *Taq*I 99 bearing alleles out of 33 were found on standard chromosomes while

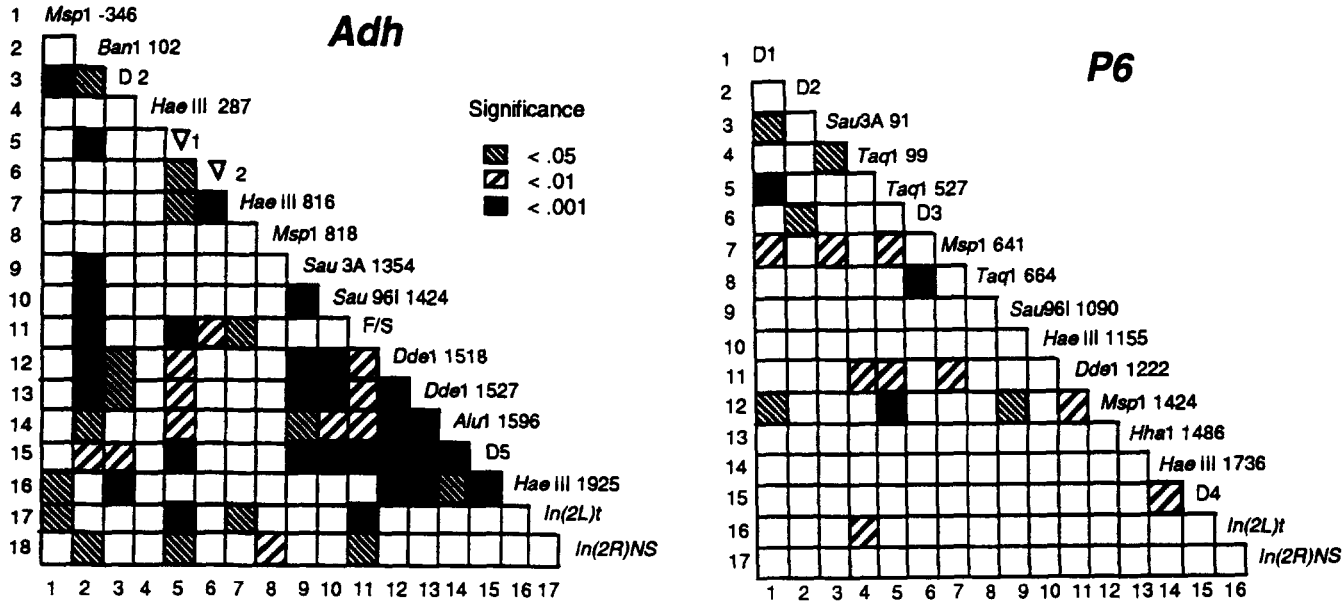

FIGURE 2.—Fisher's exact test of linkage disequilibrium among polymorphic sites and inversions in *Adh* and *P6* from the random (E1) sample. Only sites where both alleles occurred at least twice are represented.

16 out of 50 were present on *In(2L)t* chromosomes. Half of these *Taq*I 99/*In(2L)t* chromosomes make up only one haplotype, which is the second most frequent in Lamto. Both the high frequency of this allele, and its possible origin in a *In(2L)t* chromosome could explain this linkage disequilibrium.

DISCUSSION

Allozymes and inversions: Studies of linkage disequilibrium between *In(2L)t* and *In(2R)NS* inversions and ADH allozymic variation have been carried out on many populations from around the world (reviewed by LEMEUNIER and AULARD 1992). Only one such analysis was available from Africa, where ALFONSO *et al.* (1985) found significant linkage disequilibrium between ADH and *In(2L)t* in Morocco and in Senegal. The Lamto population, where both inversions are at a high frequency, revealed significant linkage disequilibrium with ADH. Molecular analysis allowed us to further investigate this result.

Molecular analysis of linkage disequilibrium between *Adh* and *In(2L)t*: At the molecular level, the "Fast" allele appears to consist of the allozymic fast variant plus the ∇1 insertion. These haplotypes sometime bear the ∇2 insertion which is always present in Fast haplotypes, although its low frequency does not allow linkage values to be significant. The *Fast* allele is also in highly significant linkage disequilibrium with the block of very close alleles overlapping the same region: sites *Sau*3A 1354, *Sau*96I 1424, *Dde*I 1518, *Dde*I 1527, and *Alu*I 1596. These polymorphic sites segregate together and have not been substantially separated by recombination. As a whole, they have been extensively exchanged between standard and *In(2L)t* chromosomes and show no linkage with them. The Fast allele most probably appeared in one particular combination of these sites and in one standard chromosome, and was not subjected to substantial recombination thereafter.

MUKAI and VOELKER (1977) analyzed linkage disequilibrium between *Adh* and *In(2L)t*, and found only one recombinant between *Adh* and *α-Gpdh* (located in the *In(2L)t* inversion) out of 4377 progeny gametes. They attributed this event to a double crossover within the inversion. In a later study, MALPICA *et al.* (1987) found no recombinant in the *In(2L)t* inversion out of 2761 progeny gametes. They reported a 0.00053 recombination rate (range: 6×10^{-5} to 1.9×10^{-3}) between *Adh* and the inversion, thus showing that recombination occurs between *Adh* and the inversion. Recombination within the inversion itself was not proven. Assuming no selection and 10 generations per year, this recombination rate would reduce linkage disequilibrium between *Adh* and *In(2L)t* to 50% in 130 years (range: 1155 to 36 years) and to 10% in 434 years (range: 3837 to 121 years). Even assuming a very recent age of one thousand years for the *Fast* allele, recombination with *In(2L)t* would be noticeable if no selection occurred.

Two mechanisms can be imagined to explain the linkage between the allozyme and the inversion. Either selection maintains Adh^F in a standard chromosome, or Adh^F appeared very recently in a standard chromosome and thereafter spread throughout the world in a few hundred years. In either case selection must have played an important role. The selective linkage hypothesis would imply the presence of some locus functionally linked to *Adh* in the area covered by *In(2L)t*. The recent finding of two *Adh* regulatory elements in this region, *Inr* (BIRCHLER, HIEBERT and

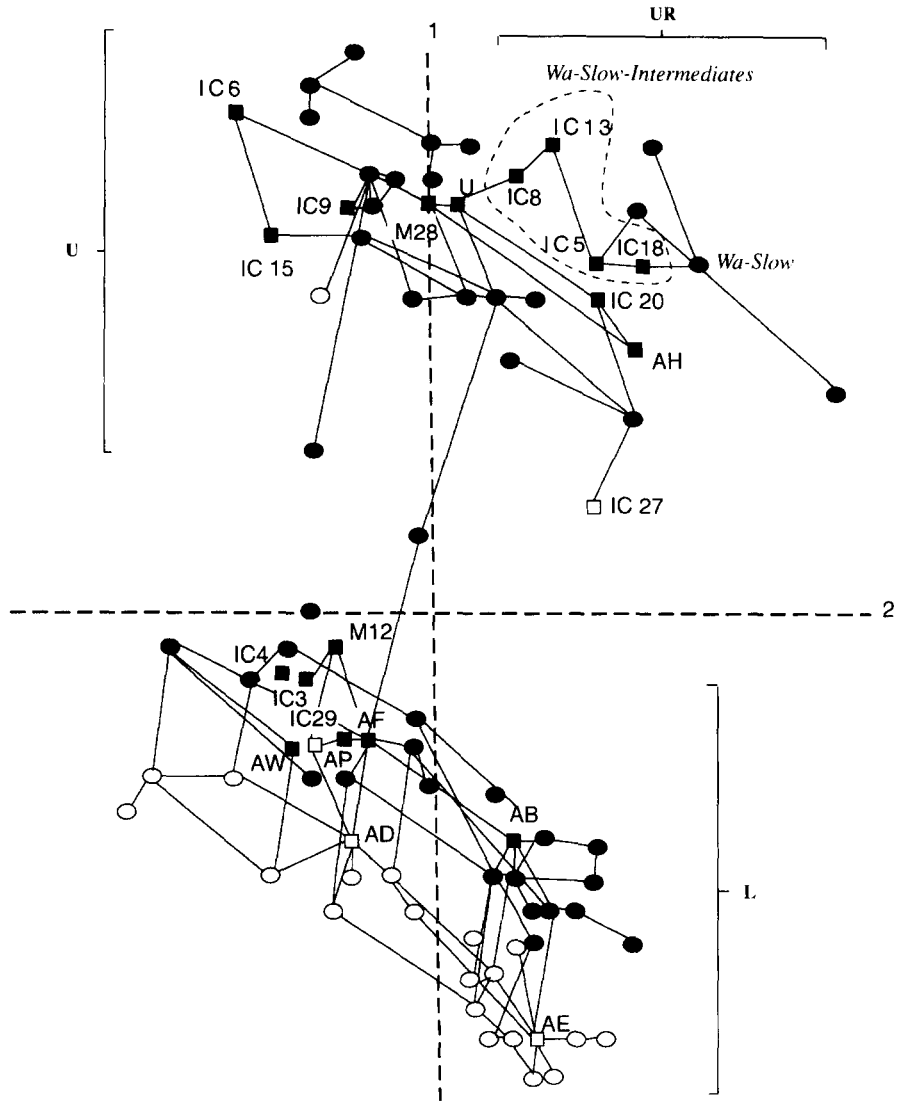

FIGURE 3.—Network of *Adh* haplotypes plotted on the first plane of a factorial correspondence analysis. Axes 1 and 2 account for 38% of the variance. Data are from Africa, America, and Spain. African haplotypes are named and represented by a square. IC: haplotypes found only in the Ivory Coast. For consistency with other studies, other haplotypes are named according to SIMMONS *et al.* (1989); AF corresponds to IC7 and IC25, AD to IC26, AE to IC28, AB to IC10, U to IC16 and IC17, AP to IC23, AW to IC12 and IC14, AH to IC19, M12 to IC2 and IC11, M28 to IC1 and IC24, IC13 to IC13 and IC21, IC18 to IC18 and IC22. AB and AD were found in all localities. Open symbols correspond to *fast* haplotypes and filled symbols to *slow* haplotypes. U corresponds to the first group of haplotypes (*Dde*I 1518 is absent and *Dde*I 1527 is present); L corresponds to the second group, including the fast haplotype (*Dde*I 1527 is absent and *Dde*I 1518 is present): UR corresponds to *Wa^S-like* and *Wa^S-intermediate*.

KENNETH 1990) and BBF-2 (ABEL, BHATT and MANIATIS 1992) suggests that such factors exist.

***P6* and the *In(2L)t* inversion:** Only one P6 restriction site (*Taq*I 99) was significantly linked to *In(2L)t*. No other significant linkage disequilibrium was found between P6 restriction site variation and the inversion. This raises the question of the unique origin of inversions and of their ability to restrict recombination.

In *D. pseudoobscura*, significant linkage disequilibria have been observed between restriction sites of the *amylase* region located within inverted regions of most chromosomal arrangements (Chiricahua, Santa Cruz, Arrowhead and Treeline). This shows that inversions sometimes maintain blocks of genes, thus counteracting the homogenizing effect of recombination (AQUADRO *et al.* 1991). According to KRIMBAS and POWELL (1992), only a few genes located within inversions are reported as not being in linkage disequilibrium. In the case of *In(2L)t*, which is a large inversion, the "middle gene hypothesis" (LOUKAS, KRIMBAS and VERGINI 1979), could explain the absence of disequilibrium with the *P6* molecular polymorphism and *α-Gpdh* allozyme polymorphism (VOELKER *et al.* 1978; MUKAI and VOELKER 1977).

Several hypotheses can explain the lack of linkage disequilibrium between restriction sites and *In(2L)t*. Both the inversion and the *P6* polymorphic sites could be old enough to have extensively exchanged genetic information through either double crossover or gene conversion. Another explanation could be that the inversion originated more than once. The latter hypothesis should be regarded as unlikely, since it would need another hypothesis to explain the strong linkage disequilibrium with the ADH^F allozyme.

***Adh* and *In(2R)NS*:** The linkage disequilibrium of *Adh^F* with the *In(2R)NS* inversion involves sites located on opposite chromosome arms. 46% of the Lamto *Adh^F* lines (±12%) were associated with this inversion. Only 16% of *Adh^S* (±4%) were *In(2R)NS*. MUKAI and VOELKER (1977) reported linkage disequilibrium be-

tween *Adh* and *In(2R)NS* in Raleigh (North Carolina) in 1968, 1969 and 1970, although the effect had disappeared by 1974. In this case, the Slow allele was associated with the inversion.

Although the map distance between *Adh* and *In(2R)NS* is theoretically 30 cM, their recombination rate is reduced to 1.75% when the inversion is present (MUKAI and VOELKER 1977). These authors concluded that a founder effect had taken place in Raleigh. If a founder event were at the root of the linkage between *In(2R)NS* and Adh^F in Lamto, it would have taken only four years for the degree of linkage to fall to its current value (assuming a recombination rate $r = 0.0175$ and 10 generations per year). This does not agree with the current high frequency of *In(2R)NS* in Lamto (21.7%) or with the haplotype diversity of *Adh*, which is similar for inverted and standard chromosomes, and is similar to values in other populations. Moreover, if the explanation was a founder effect, linkage to *In(2R)NS* would involve silent variation as well as the F/S polymorphism, since many silent polymorphic sites exist in *Adh*. The fact that only the F/S allozyme is linked to *In(2R)NS* cannot be accounted for by a founder effect. A selective explanation is, however, equally puzzling since *In(2R)NS* is negatively associated with Adh^F in America, and positively in Africa. Further studies are required before this enigma can be resolved.

Global haplotype variability: Molecular polymorphism of *Adh* in Lamto was compared with results from American populations surveys using similar restriction enzymes (Cherryfield 1985, Florida City 1985 and Raleigh 1984: data from SIMMONS *et al.* 1989; Raleigh 1983 and Putah Creek 1983: data from KREITMAN and AGUADE 1986a). The *Ban*I 102 restriction site, the small insertion/deletions I/D1, ∇1, I/D5 (<5 nucleotides) and the inversions were not taken into account as these data were not available for all surveys.

Populations were compared using the average number of restriction site differences between two haplotypes, d_A (NEI 1987). The average difference between African and American populations (0.286) is not significantly different from the average difference between American populations (0.280). No major geographic difference thus appears between the two continents.

Another way of comparing geographic samples is to show the relationship between haplotypes on a qualitative basis, not taking frequencies into account. For this analysis we also included data available from Spain (Spain 1986: data from AGUADÉ 1988). It is difficult to represent the wealth of information obtained from molecular analyses. Tree-fitting does not lend itself to the evolutionary relationships of haplotypes from the same species, since recombination occurs. SIMMONS *et al.* (1989) used a simple and convenient method whereby haplotypes are linked when they differ by one polymorphic site. Recombination, however, introduces many pathways between haplotypes. The relations between distant haplotypes are therefore subjective, as they depend on the choice of a particular haplotype arrangement. We used the same kind of connective diagram, but haplotypes were scattered on the diagram using a procedure which could be defined *a priori*. We used a multivariate projection, correspondence factor analysis (CFA), to position haplotypes (BENZECRI 1982) (Figure 3).

Polymorphic sites *Dde*I 1518 and *Dde*I 1527 are the two most heterozygous sites. They are very close to each other (11 bp). They generate two groups of haplotypes (hereafter called U for "upper" group and L for "lower" group). Within these groups (separated by a two step bridge), apparent recombination is extensive since several pathways link haplotypes within each group. Most Adh^F alleles cluster in a limited area of group L. The limited range of Adh^F alleles on this diagram is a confirmation that the haplotype separation is mostly due to sites other than F/S. Clearly, balanced selection for the amino acid variants is not an explanation of other components of *Adh* variation.

A secondary group of haplotypes (UR), corresponds to the absence of restriction sites *Sau*3I 1354 and *Hae*III 1425. These restriction sites are present in most haplotypes, but are absent in *Washington-slow* (Wa^S), one of the *Adh* genes sequenced by KREITMAN (1983). Wa^S-like North American restriction haplotypes were reported by SIMMONS *et al.* (1989) who found no intervening haplotypes to connect them to other *Adh* haplotypes [see also KREITMAN and HUDSON (1991)]. These intermediate haplotypes were found in the Ivory Coast population. They represent four steps: haplotypes IC8, IC13, IC5 and IC18, and make up 18.8% of the Lamto sample. KREITMAN and HUDSON (1991) suggested two explanations for the lack of intervening haplotypes between Wa^S and the other slow populations: a geographic isolation hypothesis and a chromosome isolation hypothesis (see Introduction). As emphasized by these authors, the geographic isolation hypothesis is very unlikely, since it would be equally true for other *D. melanogaster* genes for which there is no such evidence. Furthermore, West Africa would be the most likely geographic area from which an isolated population would have secondarily mixed with another population. But the Wa^S-intermediate and the other slow haplotypes are found in this population. According to the genetic isolation hypothesis Wa^S would have recombined away from an allele carrying a recombination suppressing factor. As KREITMAN and HUDSON did not find evidence of recombination suppression, they suggested that the *In(2L)t* inversion might be involved. Most (13 out of

16) of the intervening alleles between Wa^S and the other slow ("Wa^S-intermediate") found in Lamto are on an *In(2L)t* chromosome. The Wa^S-like haplotypes found in the USA are on standard chromosomes. This is in good agreement with KREITMAN and HUDSON'S genetic isolation hypothesis. The process could be viewed as follows: some slow haplotypes located on *In(2L)t* chromosomes would have lost the two restriction sites and accumulated mutations. Some of these would have recombined away to a standard chromosome and would have been involved in the global spread of *D. melanogaster*. The reason why Wa^S is at high frequencies in the United States while the intervening haplotypes are only found in Africa remains to be explained.

Most haplotypes found around the world make up a connected pattern of single change mutations. Five of them are present at high frequencies: AF, 13.5% of pooled data from all populations under consideration in this comparison; AD, 10.8%; AE, 10.3%; Wa^S-like, 8%; AB, 7.5% [haplotypes named according to SIMMONS *et al.* (1989)]. Two (AB and AD) are found in all studied localities. This agrees with the idea that all populations originate from a common stock. Most other haplotypes show a low frequency, and their absence or presence may be due to chance sampling. This does not hold true for the Wa^S-like haplotypes which are found at a high frequency (27.5%) in some North American samples (Cherryfield 1985; SIMMONS *et al.* 1989) while the haplotypes linking them to the remaining distribution are found only in Africa. A likely interpretation of this is that the African population is closer to the ancestral stock, and that only a subsection of it spread to North America through either random drift or selection.

We thank DANIEL LACHAISE for collecting the flies from the Lamto population (project supported by CNRS and Ecotrop), JEAN DAVID for helpful discussions, JEAN-PIERRE GAUTHIER for programming assistance and MATTHEW COBB for helpful criticism and careful reading of the manuscript.

LITERATURE CITED

ABEL, T., R. BHATT and T. MANIATIS, 1992 A *Drosophila* CREB/ATF transcriptional activator binds to both fat body and liver-specific regulatory elements. Genes Dev. **6:** 466–480.

AGUADÉ, M., 1988 Restriction map variation at the *Adh* locus of *D. melanogaster* in inverted and noninverted chromosomes. Genetics **119:** 135–140.

ALFONSO, J. M., M. HERNANDEZ, G. PADRON and A. M. GONZALEZ, 1985 Gametic non random association in north-west African populations of *Drosophila melanogaster*. Genetica **67:** 3–11.

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.

AQUADRO, C. F., A. L. WEAVER, S. W. SCHAEFFER and W. W. ANDERSON, 1991 Molecular evolution of inversions in *Drosophila* pseudoobscura: the *amylase* gene region. Proc. Natl. Acad. Sci. USA **88:** 305–309.

AULARD, S., 1990 Polymorphisme chromosomique de *Drosophila melanogaster*, en Afrique et dans les îles de l'Océan Indien. Doctoral Thesis, P. and M. Curie University.

BENZECRI, J. P., 1982 *L'Analyse des données-2 L'analyse des correspondances*. Edit. Dunod, Paris, 1982.

BIRCHLER, J. A., J. C. HIEBERT and K. PAIGEN, 1990 Analysis of autosomal dosage compensation involving the alcohol dehydrogenase locus in *Drosophila melanogaster*. Genetics **124:** 677–686.

BIRLEY, A. J., 1984 Restriction endonuclease map variation and gene activity in the *Adh* region in a population of *Drosophila melanogaster*. Heredity **52:** 103–112.

CHAMBERS, G. K., 1988 The *Drosophila alcohol dehydrogenase* gene-enzyme system. Advances in Genetics, vol **25:** 39–107.

CROSS, S. R., and A. J. BIRLEY, 1986 Restriction endonuclease map variation in the *Adh* region in population of *Drosophila melanogaster*. Biochem. Genet. **24:** 415–432.

DAVID, J. R., and P. CAPY, 1988 Genetic variation of *Drosophila melanogaster* natural populations. Trends Genet. **4:** 106–111.

HALE, L. R., and R. S. SINGH, 1991 A comprehensive study of genic variation in natural populations of *Drosophila melanogaster*. IV. Mitochondrial DNA variation and the role of history vs selection in the genetic structure of geographic populations. Genetics **129:** 103–117.

JIANG, J. C., and J. GIBSON, 1990 The Alcohol dehydrogenase polymorphism in natural population of *Drosophila melanogaster*: restriction map variation in the region of the *Adh* locus in populations from two hemispheres. Heredity **68:** 1–14.

KNIBB, W. R., 1982 Chromosome inversion polymorphisms in *D. melanogaster*. II. Geographic clines and climatic association in Australia, North America, and Asia. Genetica **58:** 213–221.

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

KREITMAN, M., and M. AGUADÉ, 1986a Genetic uniformity in two populations of *Drosophila melanogaster* as revealed by filter hybridation of four-nucleotide restriction enzyme digests. Proc. Natl. Acad. Sci. USA **83:** 3562–3566.

KREITMAN, M., and M. AGUADÉ, 1986b Excess polymorphism at the *Adh* locus in *Drosophila melanogaster*. Genetics **114:** 93–110.

KREITMAN, M., and R. R. HUDSON, 1991 Inferring the evolutionary histories of the *Adh* and *Adh-dup* loci in *Drosophila melanogaster* from patterns of polymorphism and divergence. Genetics **127:** 565–582.

KRIMBAS, C. B., and J. R. POWELL (EDITORS), 1992 *Drosophila Inversion Polymorphism*. CRC Press, Cleveland.

LACHAISE, D., M. L. CARIOU, J. R. DAVID, F. LEMEUNIER and L. TSACAS, 1988 Historical biogeography of the *Drosophila melanogaster* species subgroup. Evol. Biol. **22:** 159–225.

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.

LAURIE, C. C., J. T. BRIDGHAM and M. CHOUDHARY, 1991 Association between DNA sequence variation and variation in expression of the *Adh* gene in natural population of *Drosophila melanogaster*. Genetics **129:** 489–499.

LEFEVRE, G., JR., 1976 A photographic representation and interpretation of the polytene chromosomes of *Drosophila melanogaster* salivary glands, p. 32 in *The Genetics and Biology of Drosophila*, Vol. 1a, edited by M. ASHBURNER and E. NOVITSKI. Academic Press, New York.

LEMEUNIER, F., and S. AULARD, 1992 Inversion polymorphism in *Drosophila melanogaster*, pp. 339–405 in *Drosophila Inversion Polymorphism*, edited by C. B. KRIMBAS and J. R. POWELL. CRC Press, Cleveland.

LOUKAS, M., C. B. KRIMBAS and Y. VERGINI, 1979 The genetics of *Drosophila subobscura* populations. IX. Studies on linkage

disequilibrium in four natural populations. Genetics **93:** 497–523.

MALPICA, J. M., J. M. VASSALLO, A. FRIAS and F. FUENTE-BOL, 1987 On recombination among *In(2L)t*, *α-Gpdh* and *Adh* in *Drosophila melanogaster*. Genetics **115:** 141–142.

MATSUMOTO, N., K. SEKIMIZU, G. I. SOMA, Y. OHMURA, T. ANDOH, Y. NAKANISHI, M. OBINATA and S. NATORI, 1985 Structural analysis of a developmentally regulated 25-kDa protein gene of *Sarcophaga peregrina*. J. Biochem. **87:** 1501–1508.

MATTHEW, P., A. AGROTIS, A. C. TAYLOR and S. W. MCKECHNIE, 1992 An association between ADH protein levels and polymorphic nucleotide variation in the *Adh* gene of *Drosophila melanogaster*. Mol. Biol. Evol. **9:** 526–536.

MUKAI, T., and R. A. VOELKER, 1977 The genetic structure of natural populations of *Drosophila melanogaster*. XIII. Further studies on linkage disequilibrium. Genetics **86:** 175–185.

NEI, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.

RAT, L., M. VEUILLE and J. A. LEPESANT, 1991 *Drosophila* fat body protein P6 and alcohol dehydrogenase are derived from a common ancestral protein. J. Mol. Evol. **33:** 194–203.

SIMMONS, G. M., M. KREITMAN, W. F. QUATTLEBAUM and N. MIYASHITA, 1989 Molecular analysis of the alleles of *alcohol dehydrogenase* along a cline in *D. melanogaster*. Evolution **43:** 393–409.

TAJIMA, F., 1989 DNA polymorphism in a subdivided population: the expected number of segregating sites in the two-subpopulation model. Genetics **123:** 229–240.

TAKANO, T. S., S. KUSAKABE and T. MUKAI, 1991 The Genetic structure of natural populations of *Drosophila melanogaster*. XXII. Comparative study of DNA polymorphism in northern and southern natural population. Genetics **129:** 753–761.

TSACAS, L., and D. LACHAISE, 1974 Quatre nouvelles espèces de la Côte d'Ivoire du genre *Drosophila*, groupe *melanogaster* et discussion de l'origine du sous-groupe *melanogaster* (Diptera: Drosophilidae). Ann. Univ. Abidjan **E7:** 193–211.

VIGUE, C. L., and F. M. JOHNSON, 1973 Isozyme variability in species of the genus *Drosophila*. VI. Frequency-property-environment relationships of allelic *alcohol dehydrogenases* in *D. melanogaster*. Biochem. Genet. **9:** 213–227.

VOELKER, R. A., C. C. COCKERHAM, F. M. JOHNSON, H. E. SCHAFFER, T. MUKAI and I. E. METTLER, 1978 Inversions fail to account for allozyme clines. Genetics **88:** 515–527.

Communicating editor: M. SLATKIN