*A MOLECULAR APPROACH TO THE STUDY OF GENIC HETEROZYGOSITY IN NATURAL POPULATIONS, III. DIRECT EVIDENCE OF COADAPTATION IN GENE ARRANGEMENTS OF DROSOPHILA**

BY SATYA PRAKASH† AND R. C. LEWONTIN

DEPARTMENT OF ZOOLOGY, UNIVERSITY OF CHICAGO

Communicated by Theodosius Dobzhansky, November 15, 1967

Natural populations of many species of *Drosophila* maintain large amounts of genetic variation in the form of chromosomal inversions. In *D. pseudoobscura*, the polymorphism in natural populations is chiefly confined to the III chromosome and consists of 22 gene arrangements of which 15 are more or less common in natural populations. The closely related species *D. persimilis* has 11 arrangements[1] of the III chromosome, of which one arrangement, Standard (ST), is structurally identical with the Standard arrangement of *D. pseudoobscura*. The gene arrangements in these species have arisen one from another by a series of overlapping inversions so that it is possible to reconstruct the evolutionary sequence of events that gave rise to them. That is, the Arrowhead (AR) arrangement can be derived from the ST arrangement by a single inversion, and the Pikes Peak (PP) arrangement by a different inversion of ST, while the Cochise (CO) arrangement derives from a single inversion of AR. In this way the phylogeny of gene arrangements shown in Figure 1 has been constructed.[2] The arrangements we have studied in this paper are shown in large letters. The arrangement marked "hypothetical" in Figure 1 has never been found but must be postulated as an intermediate stage to complete the phylogeny. While the topology of the relationships shown is fixed, it is not possible to say which of the gene arrangements in each species were primitive and which were derived. Thus *D. persimilis* may have been the ancestral species with the Tuolumne gene arrangement, and all others may have evolved from it; but it is equally likely on the face of it that *D. pseudoobscura* was primitive with, say, Oaxaca as the aboriginal arrangement.

Because recombination is almost completely suppressed over the entire III chromosome in inversion heterozygotes,[3] even outside the inversion limits, there will be virtually no gene exchange between differently inverted chromosomes within a population, but free mixing of the gene contents among chromosomes of the same arrangement. We may then speak of separate "gene pools" in AR, CH, ST, etc., and these gene pools may diverge from each other by mutation, selection, and genetic drift. Evidence that the different gene arrangements do, in fact, differ in allelic frequencies at various loci comes from several lines of evidence. (1) Gene arrangements show cyclic seasonal changes in frequency in some populations but not in others.[4] (2) Some gene arrangements show altitudinal and geographic clines in frequency.[4] (3) Alternative gene arrangements in laboratory populations reach characteristic stable equilibrium frequencies when the arrangements are derived from the same population in nature,[5] and this equilibrium can be shown to result from heterosis of the hetero-

karyotypes.[6] (4) Different karyotypes in various *Drosophila* species have been found to differ in physiological components of fitness such as fecundity, viability, mating propensity, etc.[7-11] In addition to these evidences of general genetic differentiation, lethal alleles have been shown to be different in frequency in different gene arrangements, and the allelism of lethals is higher within karyotypes than between them.[12, 13]

None of these pieces of evidence, however, gives any quantitative estimate of the degree of genetic differentiation between karyotypes (except for the special case of lethal genes) nor do they throw any light on the more important and interesting problem of the relative roles of selection and chance in determining the differences. To what extent do genetic differences between inversions result from mutation and genetic drift and to what extent are they a product of natural selection? On the basis of experiments with gene arrangements from natural populations, Dobzhansky has hypothesized that the allelic contents of gene arrangements within a population have been *coadapted* by selection.[14]

The concept of coadaptation includes both selection of alleles at *different* loci *within* gene arrangements to produce a haploid genome that is physiologically balanced, and selection of alleles of the *same* loci *between* inversions to produce heterosis in heterokaryotypes. The chief evidence for this theory comes from the observation that heterokaryotypes are heterotic when the arrangements come from the same population but not when they come from different populations. Moreover, laboratory populations made by mixing genomes from different natural populations do not reach characteristic predictable equilibrium frequencies.[15] It is our purpose in this paper to use the method of electrophoresis of proteins[16, 17] to examine the genetic differentiation of the inversions of *D. pseudoobscura*. In particular we will ask, first, how much genetic divergence has occurred between different gene arrangements, and second, whether the genetic differentiation between gene arrangements is due to coadaptation. Does any evidence exist that selection is a major force holding internally and relationally balanced chromosome segments? We can give only a very preliminary answer to the first question because we have examined only two loci, the only two so far found by us to be segregating on the third chromosome of *D. pseudoobscura*. The results are such, however, that we can give an almost definitive answer to the second question. We will show that for the two loci in question, coadaptation by selection has certainly taken place to a very high degree.

Experimental Methods.—Electrophoresis in acrylamide gels was performed as described by Hubby and Lewontin.[16]

Enzyme assays: Pt-10 is a larval protein whose locus is on the III chromosome. The method used for electrophoresis and the assay and genetics of this locus is completely described by Hubby and Lewontin.[16] For α-amylase, gel assays were carried out by the method of Doane[18] modified for slab gels.

The strains used in the study were as follows:

D. pseudoobscura: Mather (California): Strains collected in 1959 by Professor Th. Dobzhansky. The population is highly polymorphic for inversions in the third chromosome. The strains were made homokaryotypic for different gene arrangements. These consisted of 10 Arrowhead (AR), 10 Standard (ST), 10 Treeline (TL), 7 Chiricahua (CH), and 16 Pikes Peak (PP). Five or more individuals were tested from each strain. *Straw-*

berry Canyon: Strains collected in 1966 by Dr. Christopher Wills and Mr. Alan Wick. The population is highly polymorphic for inversions in the third chromosome.[19] Five or more larvae from each of the 30 strains used in this study were tested for karyotype and corresponding Pt-10 or amylase genotype. The salivary chromosomes were examined in the same larva used for electrophoresis. In this manner, we obtained corresponding data for the karyotype and genotype of five or more larvae. *Mesa Verde (Colorado):* One hundred and twenty strains collected by Satya and Louise Prakash in August 1966. The population is virtually homozygous for AR. A single individual from each strain was examined to estimate the gene frequencies. *Austin (Texas):* Twenty-three strains collected by Dr. Michael Kambysellis in April 1967. The population is chromosomally polymorphic. Eight or more individuals were examined in each strain from nature in order to obtain information on all four chromosomes of the strain. The same larvae were used for karyotype and genotype determination as in the case of Strawberry Canyon. *Bogotá (Colombia):* Nineteen strains collected by Dr. A. S. Hunter in December 1966. The population is chromosomally polymorphic and represents an isolate separated from the main body of the species by a distance of 1500 miles. Five or more individuals in F_2 and F_3 were examined from each strain and again karyotype and genotype were determined.

D. persimilis: Mather (California): Strains collected in 1959 by Professor Th. Dobzhansky. The derived strains studied were all homokaryotypic for different gene arrangements. These consisted of 5 Klamath (KL), 11 Whitney (WT), 6 Standard (ST), and 3 Mendocino (MD). *Weott (California):* Strains collected in 1964 by Professor Eliot Spiess. The following derived homokaryotypic strains were studied: 21 KL, 18 MD, 1 ST, and 3 RW. For both populations of *D. persimilis*, five or more individuals were studied from each strain.

D. miranda: Five individuals each were studied from a strain from Big Basin, California, and a strain from Priest Lake, Idaho.

Results.—By convention we assume that each strain carries two genomes of independent origin, unless three alleles or gene arrangements are found in a strain, in which case we are forced to count all three. This convention causes *a slight overestimation of the frequencies of rare alleles.*

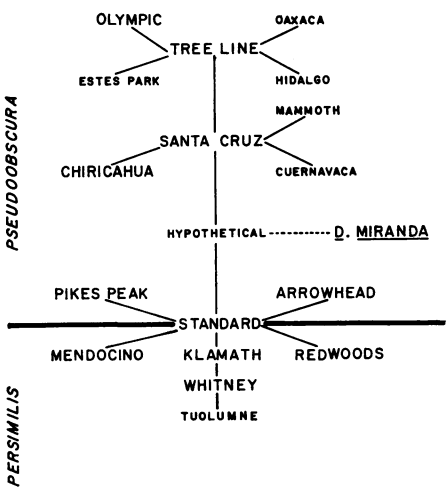

FIG. 1.—Phylogenetic relationships among gene arrangements on the third chromosome of *D. pseudoobscura* and *D. persimilis*. Modified from Dobzhansky.[3]

Pt-10 associations: It is clear from Table 1 that ST, AR, and PP gene arrangements have the allele *1.04* in frequencies close to or equal to 1.00 in all populations. In such cases as AR from Mesa Verde and PP from Austin where very large samples were available, one finds a rare allele, *1.02* or *1.06*, in an occasional chromosome. In contrast the gene arrangements SC, CH, TL, and OL are characterized by a high frequency of the allele *1.06*. Except for TL from Mather where one out of the three chromosomes was *1.04*, and CH from Mather where the alleles *1.04* and *1.06* are equally frequent, all chromosomes of this group have the allele *1.06*. Inspection of

TABLE 1. *Approximate gene frequencies of Pt-10 alleles in different gene arrangements of D. pseudoobscura from different localities.*

Gene arrangement	Allele	Population: Mather	Strawberry Canyon	Mesa Verde	Austin	Bogotá
		$n = 20$	$n = 33$	$n = 1$	$n = 5$	
Standard	*1.04*	1.00	1.00	X	1.00	—
	1.06	—	—	—	—	
		$n = 20$	$n = 6$	$n = 240$	$n = 7$	—
Arrowhead	*0.94*	0.10	—	—	—	—
	1.02	—	—	0.02	—	—
	1.04	0.90	1.00	0.97	1.00	—
	1.06	—	—	0.01	—	—
		$n = 26$			$n = 69$	
Pikes Peak	*1.02*	—	—	—	0.015	—
	1.04	1.00	—	1.00	0.985	—
			$n = 2$			$n = 38$
Santa Cruz	*1.06*	—	1.00	—	—	1.00
		$n = 14$	$n = 11$		$n = 1$	
Chiricahua	*1.04*	0.50	—	—	—	—
	1.06	0.50	1.00	—	X	—
		$n = 20$	$n = 20$		$n = 3$	$n = 38$
Treeline	*1.04*	—	—	—	0.33	—
	1.06	1.00	1.00	—	0.66	1.00
					$n = 1$	
Olympic	*1.06*				X	

X indicates presence of allele where frequency is not estimated; n = number of chromosomes.

Figure 1 shows that the arrangements ST, AR, and PP are phylogenetically related and form the so-called "Standard phylad," while SC, CH, TL, and OL forming the "Santa Cruz phylad" are removed from the Standard phylad by two steps, including the hypothetical gene arrangement. Thus Table 1 shows not only a nearly perfect association of alleles with inversion, irrespective of population from which they come, but an equally strong association within phylogenetically related inversion groups. The ST phylad as a whole is characterized by the allele *1.04*, while the SC phylad is essentially *1.06*. This phylad association leads to a prediction that *D. persimilis* ought to have the allele *1.04* since all its gene arrangements are derived from the ST phylad, as shown in Figure 1. All 25 strains of *D. persimilis* from Mather, including 6 strains of ST, 3 of MD, 5 of KL, and 11 of WT, turned out to be homozygous *1.04*. Finally the two strains of *D. miranda* were also homozygous *1.04*, but the sample is so small that nothing much can be made of this fact.

Associations of α-amylase: The balanced lethal marker stock of chromosome III, *Bl Sc pr or L* of *D. pseudoobscura*, was consistently heterozygous for α-amylase alleles *0.84/1.00*. Thus it was concluded that the locus for this enzyme (designated *Amy-1*) is on the third chromosome. Moreover, the amylase locus in *D. melanogaster* is on chromosome IIR,[20] which is homologous with the third chromosome of *D. pseudoobscura*. Simple Mendelian test crosses show that different alleles are specified by the same locus. Table 2 presents the frequencies of alleles of *Amy-1* in different gene arrangements of various

TABLE 2. *Approximate gene frequencies of α-amylase alleles in different gene arrangements of D. pseudoobscura from different localities.*

Gene arrangement	Allele	Mather	Strawberry Canyon	Mesa Verde	Austin	Bogotá
		$n = 20$	$n = 32$		$n = 5$	
Standard	*0.84*	0.05	0.12	—	0.80	—
	0.92	0.10	—	—	—	—
	1.00	0.85	0.88	—	0.20	—
		$n = 20$	$n = 10$	$n = 140$	$n = 7$	
Arrowhead	*0.84*	0.05	0.30	0.21	0.29	—
	1.00	0.95	0.70	0.79	0.71	—
		$n = 32$			$n = 69$	
Pikes Peak	*1.00*	1.00	—	—	1.00	—
			$n = 2$			$n = 38$
Santa Cruz	*0.84*	—	1.00	—	—	1.00
		$n = 14$	$n = 11$		$n = 1$	
Chiricahua	*0.84*	0.36	0.36	—	X	—
	1.00	0.64	0.64	—	—	—
		$n = 20$	$n = 14$		$n = 3$	$n = 38$
Treeline	*0.74*	—	0.14	—	—	—
	0.84	0.90	0.79	—	1.00	1.00
	1.00	0.10	0.07	—	—	—
					$n = 1$	
Olympic	*0.84*	—	—	—	X	—

TABLE 3. *Approximate gene frequencies of α-amylase alleles in different gene arrangements of D. persimilis from two localities in California.*

Gene arrangement	Allele	Mather	Weott
		$n = 12$	$n = 2$
Standard	*0.92*	0.33	1.00
	1.00	0.67	—
		$n = 6$	$n = 36$
Mendocino	*0.92*	—	0.05
	1.00	0.50	0.86
	1.05	0.50	—
	1.09	—	0.09
			$n = 6$
Redwood	*1.00*	—	1.00
		$n = 10$	$n = 42$
Klamath	*1.00*	1.00	0.83
	1.09	—	0.17
		$n = 22$	$n = 2$
Whitney	*0.84*	0.05	—
	1.00	0.27	—
	1.09	0.68	1.00

populations of *D. pseudoobscura* and Table 3 shows them in *D. persimilis*. The associations of *Amy-1* alleles with different gene arrangements are not as extreme as those of the Pt-10 locus but are still evident. In general the ST, AR, and PP gene arrangements of *D. pseudoobscura* have a high frequency (0.70–1.00) of the allele *1.00* with the single exception of ST from Austin where allele *0.84* is in high frequency. The SC and TL gene arrangements, in contrast, have a generally very high frequency of the allele *0.84*. CH chromosomes from Mather and Strawberry Canyon, however, seem to have a somewhat higher fre-

quency of the allele *1.00* than of *0.84*. *Amy-1* polymorphism in *D. persimilis* shows some interesting features (Table 3). As predicted from its derivation from the Standard phylad, *D. persimilis* is virtually free of allele *0.84*. However, further differentiation has taken place in this species so that in addition to a high frequency of *1.00* in most arrangements in most populations, there are three other alleles, *0.92*, *1.05*, and *1.09*. The WT arrangement, two steps removed from ST, has a very high frequency of allele *1.09*.

The amylase associations can be summarized by saying that the SC phylad is essentially *0.84*, and the ST phylad is essentially *1.00* with a further differentiation of other alleles in *D. persimilis*. *D. miranda* had yet another very dissimilar allele, *1.43*, in both strains.

Discussion.—Our results clearly show that very considerable genetic differentiation occurs between gene arrangements, so that at a typical locus different arrangements have not only different allelic frequencies, but may have quite different alleles altogether. If the two loci we have studied are at all typical, then for loci that are polymorphic there will be nearly complete genetic differentiation between a TL and a PP chromosome in the Mather population, for example. But what is most striking about the results is the apparent nonrandomness of the differentiation. If the genic differences between inversions were chiefly a result of their genetic isolation from each other coupled with random drift of gene frequencies, we would expect that TL in Bogotá should differ from TL in Mather about as much as TL and PP in Mather. The reverse is true. A given inversion is characterized by a particular allele over the entire range of the species, almost without exception. This strongly suggests that the genic differences date from the origin of the inversions and predate the present distribution of the species. Even more telling is the evidence from the association with phylads and the transgression of species boundaries of these phylad associations. The SC phylad is genically $Pt\text{-}10^{1.06}$ and $Amy^{0.84}$, while the arrangements of the ST phylad are generally $Pt\text{-}10^{1.04}$ and $Amy^{1.00}$. These two phylads are separated by the now extinct hypothetical arrangement which was probably polymorphic at both loci. Thus the genic difference between inversions traces back to the original split between the SC and ST phylads and also antedates the speciation event that separated *D. pseudoobscura* from *D. persimilis*. Epling[3] dates the gene arrangements to the Miocene, 13 million years ago, while the most conservative estimate is that the present distribution of *D. pseudoobscura* is Arcto-Tertiary (about 1 million years ago). Then the genic differences between arrangements, at the most conservative guess, are of 3–5 million generations' duration.

This long-maintained difference cannot be simply a relic of an original differentiation, because the "wrong" alleles are found in low frequencies in the various arrangements. $Pt\text{-}10^{0.94}$, $Pt\text{-}10^{1.02}$, and $Pt\text{-}10^{1.06}$ do occur in AR arrangements but at low frequencies and $Pt\text{-}10^{1.04}$ occurs in both CH and TL in appreciable frequency. Thus, we must assume that by mutation and occasional double-crossovers in heterozygotes, SC phylad genes are introduced into ST phylad chromosomes and vice versa. We are forced to conclude that selection is holding the allelic contents of the inversions and has held them over several million

TABLE 4. *Proportion of the genome estimated to be heterozygous in third-chromosome karyotypes for three of the populations studied.*

Mather	ST	AR	PP	TL	CH	
ST	0.132	0.145	0.075	0.935	0.469	
AR		0.138	0.075	0.977	0.462	
PP			0.000	0.950	0.430	
TL				0.090	0.556	
CH					0.480	
Strawberry Canyon	ST	AR		TL	CH	SS
ST	0.105	0.174		0.922	0.697	0.940
AR		0.210		0.857	0.722	0.850
TL				0.176	0.335	0.105
CH					0.230	0.320
SC						0.000
Austin	ST	AR	PP	TL		
ST	0.160	0.313	0.407	0.430		
AR		0.206	0.152	0.685		
PP			0.015	0.832		
TL				0.218		

generations. That is, the gene contents of the inversions are "coadapted" by selection and any introduction of foreign alleles is rejected. A concomitant of this observation is that the great genic variation observed by Lewontin and Hubby[17] is not simply isoallelic variation with no adaptive significance. At least for the genes on the third chromosome, genic differences are maintained by strong and long-term selective differences.

If the two loci we have studied are typical, all inversion polymorphisms do not have the same genic consequences. Inversion heterozygotes for gene arrangements from different phylads should be more heterozygous than heterokaryotypes between arrangements of the same phylad. Table 4 shows the genic heterozygosity of various karyotypes calculated from our gene frequency estimates in populations in which there is any appreciable genic and chromosomal polymorphism. In each case heterokaryotypes between different phylads, shown by the upper right-hand section of the table, are much the most heterozygous. In several cases homokaryotypes have more genic heterozygosity than heterokaryotypes within the same phylad, and the general impression from the table is that no very great increase in genic heterozygosity accompanies chromosomal heterozygosity if the arrangements are members of the same phylad. Applying the results of the analysis in Table 4 to the actual frequencies of in-

TABLE 5. *Proportion of the genome estimated to be heterozygous for the two third-chromosome loci in an average individual for each population of D. pseudoobscura studied in detail.*

Population	Frequency of inversion heterokaryotypes in the population	Relative proportion of gene arrangements of ST and SC phylads	Proportion of third-chromosome genome heterozygous per individual
Strawberry Canyon	0.710	0.59:0.41	0.460
Mesa Verde	0.030	1.00:0.00	0.195
Austin	0.370	0.92:0.08	0.171
Bogotá	0.469	0.00:1.00	0.000

versions in several populations of *D. pseudoobscura* has the result given in Table 5. There is no correlation between the *genic* heterozygosity on the third chromosome and the *chromosomal* heterozygosity, but there is a suggestion of a geographical pattern. Strawberry Canyon is in the region of high density of *D. pseudoobscura* and is representative of central populations of the species. Mesa Verde, Austin, and Bogotá are, on the other hand, near or at the eastern, southeastern, and southern boundaries of the species distribution. Bogotá in particular is a completely isolated southern outlier of the species range. It would appear that, at least for the third chromosome, geographical marginality means *genic* but not *chromosomal* homogeneity.

Summary.—Classification of electrophoretically separable alleles at two loci on the third chromosome of *D. pseudoobscura* and *D. persimilis* has been correlated with inversion polymorphism on the chromosome. It was shown that different inversions are genetically differentiated and the differentiation has been maintained over several million generations by natural selection.

We are most grateful to Mr. C. V. RamanMurthy for his valuable assistance during a number of phases of this work, to Th. Dobzhansky, Eliot Spiess, Christopher Wills, Michael Kambysellis, A. S. Hunter, Alan Wick, and Louise Prakash for providing us with the population samples, and to Winifred Doane for her instruction.

* The work reported here was performed under grant GB 3112 from the National Science Foundation.

† NIH postdoctoral trainee.

[1] Spiess, E. B., *Am. Naturalist*, **99**, 423–425 (1965).
[2] Dobzhansky, Th., and C. Epling, *Carnegie Inst. Wash. Publ. 554* (1944), pp. 1–183.
[3] Dobzhansky, Th., and C. Epling, these PROCEEDINGS, **34**, 137–141 (1948).
[4] Dobzhansky, Th., *Genetics*, **28**, 162–186 (1943).
[5] Wright, S., and Th. Dobzhansky, *Genetics*, **31**, 125–156 (1946).
[6] Dobzhansky, Th., *Genetics*, **35**, 288–302 (1950).
[7] Spiess, E. B., M. Ketchel, and B. Kinne, *Evolution*, **6**, 208–215 (1952).
[8] Levine, R. P., *Evolution*, **6**, 216–233 (1952).
[9] Spiess, E. B., and B. Langer, *Evolution*, **15**, 535–544 (1961).
[10] *Ibid.*, **18**, 430–444 (1964).
[11] Prakash, S., *Genetics*, **57**, 385–400 (1967).
[12] Epling, C., V. E. Tinderholt, and R. H. T. Mattoni, *Evolution*, **15**, 447–454 (1961).
[13] Mayhew, S. H., S. K. Kato, F. M. Ball, and C. Epling, *Evolution*, **20**, 646–662 (1966).
[14] Dobzhansky, Th., *Genetics and the Origin of Species* (New York: Columbia Univ. Press, 1951), 3rd ed.
[15] Dobzhansky, Th., and O. Pavlovsky, *Evolution*, **7**, 198–210 (1953).
[16] Hubby, J. L., and R. C. Lewontin, *Genetics*, **54**, 577–594 (1966).
[17] Lewontin, R. C., and J. L. Hubby, *Genetics*, **54**, 595–609 (1966).
[18] Doane, W. W., *Drosophila Information Service*, **41**, 192 (1966).
[19] Strickberger, M., and C. J. Wills, *Evolution*, **20**, 592–602 (1966).
[20] Kikkawa, H., and Z. Ogita, *Japan. J. Genet.*, **37**, 394–395 (1962).