DARWINIAN *VERSUS* NON-DARWINIAN EVOLUTION IN NATURAL POPULATIONS OF *DROSOPHILA*

FRANCISCO J. AYALA
UNIVERSITY OF CALIFORNIA, DAVIS

1. Scientific hypotheses: natural selection

The goal of science is to discover patterns of relations among recorded phenomena, so that a few principles can explain a large number of propositions concerning these phenomena [2]. The scientific value of a theory depends on its explanatory power, that is, on its ability to encompass many subsidiary hypotheses into a single comprehensive set of mutually consistent principles. But in order to be accepted in science the applicability of a theory needs to be demonstrated.

Demonstration, or proof, of a hypothesis or theory concerning the empirical world is a gradual process which is never irrevocably completed. The process of demonstration requires, first, to show that the hypothesis or theory is consistent with the relevant facts. Moreover, the hypothesis or theory needs to be confirmed by empirical tests. Empirical tests are experiments or observations which may conceivably have diverse outcomes only some of which are compatible with the hypothesis tested while most of them would lead to its rejection. If the tests are of such a nature that any conceivable outcome or state of affairs be compatible with the hypothesis tested they contribute nothing to the scientific verification of the hypothesis. The value of an empirical test is measured by the *a priori* likelihood of its outcome being incompatible with the hypothesis.

The synthetic theory of evolution, or the theory of evolution by natural selection has a considerable explanatory power. The central concept of the theory is the principle of natural selection—the differential reproduction of genetic variants. Natural selection is the main process directing the evolution of organisms by promoting their adaptation to their environments. The principle of natural selection, together with some subsidiary and generally well authenticated hypotheses, can explain a large number of facts concerning the living world; like the diversity of organisms, their gradual change through historical time, and their remarkable adaptations to the environments where they live. The synthetic theory of evolution by natural selection is, indeed, the single most encompassing biological theory.

The theory of evolution by natural selection appears to be consistent with all or most facts known in every field of biology. The principle of natural selection can account for different patterns, rates and outcomes of evolutionary processes. Phylogenetic radiations as well as lack of phyletic diversification, rapid and slow rates of evolutionary change, abundant and limited genetic variation within a population—these and many other alternative occurrences can all be explained by postulating the existence of appropriate environmental challenges. This is, of course, as it should be. If evolutionary processes are to be explained by a single major directional process, this process must be able to explain the multifarious ways in which evolution in fact occurs. However, it is often difficult or impossible to obtain direct evidence demonstrating the role of natural selection in a given evolutionary situation. This is in part due to the historical character of biological evolution; it is difficult to get information concerning events which occurred in a more or less remote past. But there is also a methodological handicap. Obtaining direct evidence is hampered precisely by the broad explanatory power of the theory of evolution by natural selection. It is difficult to plan observations or to design experiments most of whose conceivable outcomes are incompatible with natural selection. The scientific validity of the theory of evolution by natural selection rests largely on its consistency with the entire range of available facts. Empirical evidence for the theory is, more often than not, indirect—it supports the theory of natural selection because it is incompatible with alternative hypotheses.

The hypothesis has recently been advanced that much evolutionary change may be the result of random processes [11], [15]. This has been labeled "non-Darwinian evolution" [15], or more appropriately, "evolution by random walk" [8]. According to this hypothesis most genetic variants, particularly variants observed at the molecular level, are adaptively neutral, and therefore not subject to natural selection. The proponents of the theory agree that a large fraction of all new mutations is harmful. These mutants are eliminated or kept at very low frequencies by natural selection. It is, however, argued that a substantial fraction of all occurring mutations is adaptively neutral. Carriers of alternative genotypes do not differ in their adaptedness to the environment. The frequencies in populations of these adaptively neutral variants is, accordingly, not subject to natural selection. Since no natural population consists of an infinite number of individuals, random sampling affects the frequencies of adaptively neutral variants and leads ultimately to the elimination of some variants and the fixation of others. It would then be possible that many genetic differences between species—especially differences observed at the molecular level—have no adaptive or evolutionary significance. They are evolutionary "noise" rather than "signal."

The suggestion that random processes play a major role in evolutionary change is an interesting addition to modern development of the theory of evolution. Evolution by random walk may provide a mechanistic explanation, alternative to natural selection, of evolutionary change. The theory of evolution by

random walk has the advantage of leading to some rather precise predictions concerning the pattern of genetic variation. It, therefore, lends itself to observational verification or refutation. Experiments and observations to test the predictions of random walk evolution test also, at least indirectly, the validity of the theory of natural selection.

Evolution by natural selection and evolution by random walk are not incompatible. As noted above, proponents of random walk evolution admit that some genetic variants have deleterious effects and are, therefore, subject to natural selection. Similarly, the theory of natural selection admits that gene frequencies are affected by other processes besides natural selection, like mutation, migration, and random sampling. The issue between the two theories is whether most of the genetic variation occurring in natural populations, and most of the genetic differences between species, are the result of random processes or are the result of natural selection. The rest of this paper gives a partial summary of a series of observations and experiments made to decide that question.

2. Genetic variation in *Drosophila*

These studies deal with a group of neotropical species related to *Drosophila willistoni*. The group includes at least twelve species, whose overall geographic distribution extends from Mexico and southern Florida, through Central America and the West Indies, to tropical South America, as far down as southern Brazil and northern Argentina. Six species are siblings, nearly indistinguishable from each other by external morphology. Four of the siblings, *D. willistoni*, *D. tropicalis*, *D. equinoxialis*, and *D. paulistorum* have wide, largely overlapping geographic distributions [29].

Genetic variation in the *D. willistoni* group has been studied with standard techniques of starch gel electrophoresis [3], [4], [5], [24]. A summary of the amount of genetic variation in *D. willistoni* is given in Table I. For each of 27 gene loci the table gives the total number of wild genes sampled. This is simply the number of wild individuals studied multiplied by two, except for sex linked genes which are carried by males in single dose. The number of populations sampled is the number of different local populations studied. The amount of genetic variation at each locus is measured in two ways. First, the proportion of populations in which the gene is polymorphic. A gene is taken as polymorphic when the most common allele in the population has a frequency no higher than 0.95. The second, more precise, measure of variation is the mean frequency of individuals which are heterozygous at a given locus. The estimates of frequency of heterozygotes are made assuming that the locus is at Hardy-Weinberg equilibrium.

The information in Table I is summarized, first, by the proportion of polymorphic loci per population. This is simply the average over all loci of the proportion of polymorphic loci per population. This statistic is 56.3 ± 6.2 per cent for the genes sampled in the table. The data are summarized also by the

TABLE I

GENETIC VARIATION AT 27 GENE LOCI OF *Drosophila willistoni*

Gene	Sample size	Number of populations sampled	Frequency of polymorphic populations	Frequency of heterozygous individuals
Lap-5	10348	80	1.00	.629 ± .007
Est-2	7012	61	.668	.117 ± .008
Est-3	3914	42	.727	.108 ± .009
Est-4	9692	78	.885	.270 ± .012
Est-5	10432	81	.333	.089 ± .006
Est-6	2418	48	1.00	.285 ± .030
Est-7	6819	67	1.00	.601 ± .009
Aph-1	2041	10	.800	.136 ± .020
Acph-1	3066	24	.588	.102 ± .018
Acph-2	874	10	.900	.151 ± .023
Adh	5916	56	.515	.103 ± .015
Mdh-2	6680	58	.175	.040 ± .008
α-Gpdh	7032	57	.044	.021 ± .005
Idh	3168	63	.370	.086 ± .023
G3pdh	190	27	.500	.148 ± .053
Odh-1	1088	30	.765	.180 ± .019
Odh-2	40	4	.333	.060 ± .024
Me-1	2882	61	.275	.079 ± .013
Me-2	1368	15	.733	.235 ± .044
To	4949	22	.333	.131 ± .029
Tpi-2	2478	59	.105	.041 ± .010
Pgm-1	2636	28	.815	.186 ± .025
Adk-1	2150	40	1.00	.527 ± .019
Adk-2	2580	60	.722	.171 ± .023
Hk-1	1620	44	.333	.099 ± .016
Hk-2	2228	49	.556	.138 ± .032
Hk-3	2060	47	.071	.039 ± .009

Mean frequency of polymorphic loci per population 0.563 ± .062
Mean frequency of heterozygous loci per individual 0.177 ± .031

proportion of genes at which an individual is heterozygous. This is obtained by averaging over all loci the frequency of heterozygotes at each locus. On the average, an individual of *D. willistoni* is heterozygous at 17.7 ± 3.1 per cent of its genes. A similar degree of genetic variation has been found in the sibling species, *D. tropicalis* (18.6 ± 3.3) and *D. equinoxialis* (21.3 ± 3.0).

Genetic variation of the same order of magnitude has also been observed in other species of *Drosophila*. The following are estimates of the per cent of the genome heterozygous per individual: *D. simulans*, 8.0 [16], [20]; *D. athabasca*, 8.7 [16]; *D. persimilis*, 10.5 [21]; *D. ananassae*, 10.8 [9]; *D. pseudoobscura*, 12.3 [23]; *D. melanogaster*, 17.8 [16], [20]; *D. affinis*, 25.3 [16]. Similarly, the proportion of heterozygous loci per individual is estimated to be about five or six per cent in organisms as different as man [10], horseshoe crab [28], and old-field mouse [26]. The proportion of heterozygous loci per individual is 8 to 11 per cent in the house mouse [25], [27].

The genes examined in *D. willistoni*, as well as in all the studies just quoted, were chosen because simple assay techniques were available. They can be considered a random sample with respect to variation of loci coding for soluble proteins, since they were selected independently of how variable they were. However, they all belong to a single class of genes, namely, those coding for soluble proteins. Two other major classes of genes exist, regulatory genes and those coding for structural proteins. We do not know how variable regulatory genes are. Our ignorance about variation in genes coding for structural proteins is also nearly complete (see [19], however, for variation in fibrinopeptides among species of artiodactyls). It must also be pointed out that even for the genes coding for soluble proteins the estimates of heterozygosity are only gross approximations. The biochemical techniques used, as well as the sampling procedures are potentially subject to various biases which may affect the estimates in an undetermined direction and by an undetermined amount. It is, however, likely that they underestimate the amount of variation [3], [17]. The remarkable fact is that different investigators using different techniques to study different genes in different organisms have obtained estimates of genetic variation falling within a narrow range of values. It may be concluded that in organisms as different as *Drosophila* flies, horseshoe crabs, mice, and men, about 50 per cent of the genes coding for soluble proteins are polymorphic in a given population, and an individual is heterozygous at about ten per cent of its genes. These estimates are likely to be correct as to the order of magnitude.

To estimate the absolute number of polymorphic loci in a given species we need to know the total number of functional gene loci in the species. Estimates obtained by dividing the number of nucleotide pairs per haploid DNA (between 10^9 and 10^{10} for insects and mammals) by the average number of nucleotide pairs (about 500) per gene (cistron) run in the millions. These estimates give the upper limit of the number of functional loci per individual. Considerations based on classical genetic analyses give a lower limit of 10,000 loci per individual in organisms like man and *Drosophila*. It is likely that the number of loci in these organisms is of the order of 10^5. The total number of genes which are polymorphic in a given population must, then, be at least in the thousands. This is, indeed, a considerable amount of genetic variation, and considerably more than what many geneticists were willing to accept as recently as five years ago.

The question which concerns us now is: what is the evolutionary significance of all this genetic variation? Are most genetic polymorphisms adaptively neutral or, on the contrary, are they maintained by natural selection? To answer this question we must look at the *pattern* of the genetic variation.

3. Genetic variation between populations

The amount of genetic variation varies considerably from locus to locus of *Drosophila willistoni*. The proportion of heterozygous individuals is greater than 50 per cent at the *Lap-5*, *Est-7*, and *Adk-1* loci, but four per cent or less at

TABLE II

GENETIC VARIATION AT THE *Lap-5* LOCUS OF *Drosophila willistoni*

Locality	Sample size	Alleles .96	.98	1.00	1.03	1.05	Other	Frequency of heterozygous individuals
Jaque	194	.03	.18	.34	.37	.06	.02	.709
Teresita	260	.03	.14	.32	.36	.14	.02	.728
P. Lopez	402	.002	.08	.30	.57	.04	.002	.578
Betoyes	180	.01	.13	.26	.54	.06	.01	.618
Mitu	160	.03	.09	.23	.41	.24	.01	.710
Caracas	316	.01	.11	.23	.52	.12	—	.646
Macapa	56	—	.11	.23	.63	.04	—	.543
Belem	74	—	.10	.22	.47	.22	—	.674
Santarem	492	.02	.14	.39	.43	.02	.01	.649
Tefe	172	—	.11	.34	.42	.13	—	.678
Mirassol	1806	.01	.07	.25	.57	.09	.002	.618

Mdh-2, *α-Gpdh* and *Hk-3* (Table I). In sharp contrast, the amount and the pattern of the variation remain remarkably constant from locality to locality [4], [5]. To illustrate this difference I have selected six loci, two very polymorphic (*Lap-5* and *Est-7*), two moderately polymorphic (*Est-5* and *Pgm-1*), and two with little polymorphism (*α-Gpdh* and *Mdh-2*). Tables II through VII give for each of the six loci the number of genes sampled, the allelic frequencies, and the proportion of heterozygous individuals expected on the assumption of Hardy-Weinberg equilibrium. At each locus one allele, usually the most common, has been arbitrarily designated 1.00. Others are named with reference to that standard. For instance an allele .98 codes for a protein that in the gels migrates towards the anode 2 mm less than the standard. Only a few representative populations are included in the tables. Their geographic position is

TABLE III

GENETIC VARIATION AT THE *Est-7* LOCUS OF *Drosophila willistoni*

Locality	Sample size	Alleles .96	.98	1.00	1.02	1.05	Other	Frequency of heterozygous individuals
Jaque	174	.03	.27	.46	.20	.03	—	.673
Teresita	216	.02	.20	.45	.30	.03	—	.669
P. Lopez	294	.02	.15	.58	.20	.05	.003	.560
Betoyes	111	.06	.14	.53	.22	.05	—	.646
Mitu	224	.06	.25	.50	.14	.05	.004	.658
Macapa	41	—	.12	.39	.29	.20	—	.709
Belem	61	—	.10	.64	.16	.10	—	.545
Santarem	406	.01	.15	.55	.21	.07	—	.620
Tefe	147	.02	.12	.45	.31	.11	—	.679
Mirassol	920	.01	.13	.57	.26	.01	.02	.586

TABLE IV

GENETIC VARIATION AT THE *Est-5* LOCUS OF *Drosophila willistoni*

Locality	Sample size	Alleles .95	1.00	1.05	Other	Frequency of heterozygous individuals
Jaque	188	.01	.98	.01	—	.032
Teresita	238	.02	.97	.01	—	.058
P. Lopez	400	.04	.95	.01	.002	.101
Betoyes	178	.02	.96	.02	—	.076
Mitu	224	.04	.94	.02	—	.115
Caracas	212	.02	.96	.02	—	.078
Macapa	48	.02	.96	.02	—	.081
Belem	58	.02	.98	—	—	.034
Santarem	444	.03	.96	.01	—	.079
Tefe	174	.03	.96	.01	.01	.078
Mirassol	1976	.03	.96	.01	—	.086

TABLE V

GENETIC VARIATION AT THE *Pgm-1* LOCUS OF *Drosophila willistoni*

Locality	Sample size	Alleles .96	1.00	1.04	Other	Frequency of heterozygous individuals
Jaque	42	—	.95	.05	—	.091
Teresita	36	—	.92	.08	—	.153
P. Lopez	188	.03	.88	.08	.01	.222
Betoyes	134	.02	.98	.01	—	.044
Macapa	12	—	.83	.17	—	.320
Santarem	20	.05	.85	.10	—	.265
Tefe	20	.05	.85	.10	—	.265
Mirassol	40	.05	.88	.08	—	.226

TABLE VI

GENETIC VARIATION AT THE *α-Gpdh* LOCUS OF *Drosophila willistoni*

Locality	Sample size	Alleles .94	1.00	1.06	Other	Frequency of heterozygous individuals
Jaque	48	—	1.00	—	—	.000
Teresita	48	—	1.00	—	—	.000
P. Lopez	476	.002	.99	.002	.002	.013
Betoyes	256	—	.99	.004	.004	.016
Caracas	222	.01	.95	.05	—	.105
Macapa	54	—	1.00	—	—	.000
Belem	74	.01	.99	—	—	.027
Santarem	498	.002	.996	—	.002	.008
Tefe	174	—	1.00	—	—	.000
Mirassol	1102	.004	.98	.02	.004	.046

TABLE VII

GENETIC VARIATION AT THE *Mdh-2* LOCUS OF *Drosophila willistoni*

Locality	Sample size	Alleles .86	.94	1.00	1.06	Other	Frequency of heterozygous individuals
Jaque	50	—	—	.98	.02	—	.039
Teresita	48	—	—	1.00	—	—	.000
P. Lopez	476	.004	.004	.99	—	—	.017
Betoyes	256	—	.004	.996	—	—	.008
Caracas	220	—	.02	.93	.05	.005	.168
Macapa	54	—	—	1.00	—	—	.000
Belem	74	.03	.03	.93	.01	—	.129
Santarem	486	.004	.002	.99	—	.002	.016
Tefe	178	—	.03	.93	.02	.02	.095
Mirassol	1024	.001	.03	.95	.01	.004	.090

indicated in Figure 1. The localities in Tables II through VII are all in continental South America. They embrace an enormous territory extending from Panama (Jaque) to the Amazon delta (Belem), and from Caracas to south Brazil (Mirassol). For additional details concerning these populations see [5], [29].

The observed pattern of genetic variation in *D. willistoni* suggests the following generalizations: (1) there is considerable variation from locus to locus in the amount of polymorphism; (2) at a given locus there is great similarity from population to population as to the amount and pattern of the genetic variation; (3) nevertheless, differences between localities occur, with some allelic frequencies being characteristic of certain regions or localities. For instance, as illustrated in Table II, allele 1.03 is the most common in continental populations of *D. willistoni*, allele 1.00 is the second most common, and alleles .98, 1.05 and .96 occur at lower frequencies. These five alleles have been found in every adequately sampled population. Yet allele 1.05 reaches frequencies of 0.24 and 0.22 in Mitu and Belem, respectively, but only 0.02 in Santarem, 0.04 in P. Lopez and Macapa. Any explanation of the genetic variation must account for the observed pattern of the variation.

The ultimate source of gene variability is mutation. In a local population gene variants may also be introduced or removed by migration between neighboring populations. The theory of random walk evolution argues that most of the observed genetic variation is adaptively neutral and, therefore, not subject to natural selection. Gene frequencies are the result of random sampling which occurs every generation. Then, the effective number of neutral alleles which can be maintained in a population is

$$n = 4Nu + 1, \tag{1}$$

where N is the effective size of the population and u is the rate of mutation to neutral alleles [12]. The use of (1) is handicapped by lack of accurate estimates

FIGURE 1

Geographic origin of the samples used in this study: 6, Jaque, Panama; 14, Teresita, Colombia; 18, Mitu, Colombia; 22, southeast of Caracas, Venezuela; 32, Macapa, Brazil; 33, Belem, Brazil; 34, Santarem, Brazil; 38, Tefe, Brazil; 45, Mirassol, Brazil; 50, P. Lopez, Colombia; 52, Betoyes, Colombia; 53, Martinique; 54, St. Lucia; 55, St. Vincent; 56, Bequia; 57, Carriacou; 58, Grenada.

concerning N and u. The rate of mutation to neutral alleles is unlikely to be much higher than 10^{-6}, and is likely to be of the order of 10^{-7} per gene per generation. We do not know the effective size of a breeding population of *D. willistoni*. If it is approximately one tenth of the reciprocal of the mutation rate, the average effective number of alleles would be 1.4 per locus; and the average heterozygosity per locus would be about 29 per cent. This is approximately what is found in populations of *D. willistoni*.

There are, however, some difficulties. First, assuming, as is reasonable, that the mutation rate remains constant from population to population, differences in average population size should result, unless they are all effectively infinite, in differences in the effective number alleles. Consider two isolated patches of tropical forest, one with an effective population size of *D. willistoni* four times as large as the other. Four times as many alleles per locus should be found in one as in the other population. We find that, on the contrary, the number of alleles at a given locus remains fairly constant from population to population. A conceivable escape from this difficulty consists in arguing that independently of the total number of flies in a given locality, the population is subdivided in breeding colonies of approximately equal size. These breeding colonies should all be of a size approximately one tenth of the reciprocal of the mutation rate, or about 10^6 individuals. It seems difficult to understand why such remarkable constancy of effective population size would in fact occur.

A more serious difficulty is that, according to the *theory of random evolution*, different sets of alleles should occur in different populations. Whenever the same alleles are found in different populations, their frequencies should be uncorrelated. This prediction clearly follows from the claim that allelic frequencies are the result of random sampling. This prediction is completely negated by the observations. At every locus the same alleles are found again and again in different populations, at frequencies which are highly correlated.

4. The role of migration

The great degree of similarity between populations in the pattern of genetic polymorphisms is incompatible with the hypothesis that these polymorphisms are the result of unrelated, random processes. It has been recently suggested, however, that if a certain amount of migration occurs between neighboring populations, the species may effectively approximate a single panmictic population [13]. This may seem a possible hypothesis because it could explain, even with selective neutrality, the similarity of allelic frequencies throughout the distribution of the species. Different localities would effectively represent samples of one single interbreeding population. The similarity of allelic frequencies between populations can be accounted for if

$$Nm > 4, \tag{2}$$

where N is the effective size of the *local* population, and m is the rate of migration per generation between neighboring populations. The observed levels of heterozygosity per individual and of polymorphic loci per population (10 and 50 per cent, respectively) would obtain [13] if

$$4N_e u \approx 0.1, \tag{3}$$

where N_e is the effective size of the *species*, and u is the mutation rate as in (1).

It is difficult to estimate even approximately the effective population size of *D. willistoni*. The geographic distribution of this species extends over several million square kilometers. Throughout this enormous territory *D. willistoni* is often the most abundant drosophilid. An experienced collector typically obtains several hundred to several thousand individuals in two or three hours within a few hundred square meters. There is no indication that the removal of these individuals affects substantially the size of the population. Collections made in consecutive days in the same site yield approximately equal numbers of flies. The total number of *D. willistoni* flies living at a given time is doubtless much larger than 10^9. Taking 10^9 as a lower estimate of the effective size of the species and assuming that the mutation rate is of the order of 10^{-7} [13], we obtain $4N_e u + 1 = 401$. If the species *D. willistoni* forms effectively a single panmictic population, we should observe at each locus a large number of alleles each at very low frequencies. We find instead a small number of alleles at high and intermediate frequencies.

The theory of evolution by random walk does not claim that selective differences between genotypes are in fact zero in all places and at all times. This is, on theoretical grounds, extremely unlikely. The theory claims rather that allelic frequencies will be regulated by random processes rather than by selection whenever the differences in selective value between genotypes is sufficiently small. If N_e is the effective size of the population and s measures the selective advantage or disadvantage of the alleles, random sampling will have an important effect in loci satisfying the condition ([7], [14])

$$|N_e s| \leqq 1. \tag{4}$$

If the whole species approximates a single panmictic population, N_e will be greater than 10^9 in *D. willistoni*. Even extremely small differences in selective value, say of the order of 10^{-7} or 10^{-8}, would be the major factor determining gene frequencies.

Within the framework of population genetics theory the most serious difficulty with the maintenance of large amounts of variation by natural selection is that it imposes a "genetic load" on the population. This matter cannot be fully discussed here. In outline the difficulty is as follows. If genetic variation is maintained at a given locus by natural selection, some genotype or genotypes will be, by definition, adaptively inferior to others. The average fitness of the population will, then, be lower than a population monomorphic for the most fit

genotype. If there are many polymorphic loci maintained by selection, the average fitness of the population will be substantially lower than that of the optimal genotype. For instance, assume that the average fitness of the population is reduced by one per cent at each polymorphic locus. If genes affect fitness independently of each other the mean fitness of a population segregating at 1000 loci is $(0.99)^{1000} = 4 \times 10^{-5}$. It seems impossible that a population would survive with a genetic load that reduced its fitness to 0.00004 of its potential reproductive capacity.

This difficulty has been answered in many ways. The proponents of random walk evolution believe that the objection is nevertheless valid and that, therefore, most genetic polymorphisms must be selectively neutral. As shown above if the species constitutes effectively a single breeding population, very small selective differences would be sufficient to maintain polymorphisms. The difficulty from genetic load theory consequently vanishes. For instance, if the mean fitness of the population is on the average, decreased at each polymorphic locus by 10^{-5}, the average fitness of a population polymorphic at 1000 independent loci would be $(0.99999)^{1000} = 0.99$. With 10,000 polymorphic loci the mean fitness of the population would be 0.90. A reduction of ten per cent from its optimal reproductive potential can doubtless be sustained by *D. willistoni*. A *Drosophila* female produces several hundred mature eggs throughout its lifetime. Even a human female possesses several hundred oocytes capable to develop into mature eggs. To maintain a constant population size only two need, on the average, to develop into reproducing adults.

5. Genetic variation in geographically isolated populations

The hypothesis that, with selective neutrality, migration explains the constancy of pattern of the genetic polymorphisms not only encounters serious theoretical difficulties, but is incompatible with empirical evidence. First, this hypothesis fails to account for the occurrence of local and regional differentiations. Why should two populations, such as P. Lopez and Betoyes, have very similar allelic frequencies at some loci, for example, *Est-5*, *α-Gpdh* and *Mdh-2*, but quite different at *Pgm-1*? Examples of local and regional differences in allelic frequencies have been found at essentially every locus of *D. willistoni* studied [4], [5]. Some examples are given below.

To ascertain whether migration may account for the similarity of gene frequencies we conducted a study of six oceanic islands: Martinique, St. Lucia, St. Vincent, Bequia, Carriacou, and Grenada. They belong to the Windward Group of the Lesser Antilles (see Figure 1). Although relatively small islands, they range in size from about ten square miles (Bequia and Carriacou) to four hundred (Martinique). These islands were not connected with each other or with continental South America in geological history. Their *Drosophila* populations must derive from small numbers of founders which reached them by

accidental transport. Substantial differences in the chromosomal polymorphisms between the islands and the continent, and between different islands, attest to their geographic isolation [4]. Allelic frequencies at the six loci listed for the continental populations are given in Tables VIII through XIII. If the poly-

TABLE VIII

GENETIC VARIATION AT THE *Lap-5* LOCUS IN SIX ISLAND POPULATIONS OF *Drosophila willistoni*

Locality	Sample size	Alleles .96	.98	1.00	1.03	1.05	Frequency of heterozygous individuals
Martinique	264	—	.06	.53	.40	.01	.555
St. Lucia	280	—	.03	.56	.39	.01	.530
St. Vincent	258	.004	.03	.59	.37	.004	.509
Bequia	354	—	.05	.59	.35	.02	.532
Carriacou	306	.003	.05	.49	.42	.03	.573
Grenada	266	.01	.04	.49	.42	.04	.584

TABLE IX

GENETIC VARIATION AT THE *Est-7* LOCUS IN SIX ISLAND POPULATIONS OF *Drosophila willistoni*

Locality	Sample size	Alleles .96	.98	1.00	1.02	1.05	Other	Frequency of heterozygous individuals
Martinique	153	.05	.14	.63	.16	.01	.01	.552
St. Lucia	156	.01	.06	.74	.16	.03	.01	.417
St. Vincent	230	.01	.10	.70	.17	.03	—	.478
Bequia	258	.01	.09	.67	.21	.02	—	.493
Carriacou	262	.01	.10	.65	.21	.03	—	.521
Grenada	233	.02	.12	.65	.17	.03	.004	.528

TABLE X

GENETIC VARIATION AT THE *Est-5* LOCUS IN SIX ISLAND POPULATIONS OF *Drosophila willistoni*

Locality	Sample size	Alleles .95	1.00	1.05	Other	Frequency of heterozygous individuals
Martinique	264	.02	.97	.02	—	.059
St. Lucia	278	.01	.91	.08	.004	.166
St. Vincent	258	.02	.97	.01	.004	.061
Bequia	352	.01	.96	.02	—	.072
Carriacou	306	.04	.90	.06	.003	.193
Grenada	262	.07	.93	.004	—	.135

TABLE XI

GENETIC VARIATION AT THE *Pgm-1* LOCUS IN SIX ISLAND POPULATIONS OF *Drosophila willistoni*

Locality	Sample size	Alleles .96	1.00	1.04	Frequency of heterozygous individuals
Martinique	286	—	.77	.23	.359
St. Lucia	234	—	.70	.30	.423
St. Vincent	174	.02	.79	.20	.342
Bequia	384	—	.91	.09	.166
Carriacou	218	.01	.98	.01	.036
Grenada	198	.02	.97	.02	.059

TABLE XII

GENETIC VARIATION AT THE *α-Gpdh* LOCUS IN SIX ISLAND POPULATIONS OF *Drosophila willistoni*

Locality	Sample size	Alleles .94	1.00	1.06	Frequency of heterozygous individuals
Martinique	292	—	1.00	—	.000
St. Lucia	280	.004	.99	.01	.021
St. Vincent	236	—	.99	.01	.025
Bequia	294	—	1.00	—	.000
Carriacou	220	—	1.00	—	.000
Grenada	194	.01	.99	—	.010

TABLE XIII

GENETIC VARIATION AT THE *Mdh-2* LOCUS IN SIX ISLAND POPULATIONS OF *Drosophila willistoni*

Locality	Sample size	Alleles .86	.94	1.00	1.06	Other	Frequency of heterozygous individuals
Martinique	292	.003	.01	.99	—	—	.027
St. Lucia	280	—	—	1.00	—	—	.000
St. Vincent	236	—	—	1.00	—	—	.000
Bequia	294	—	—	.997	—	.003	.007
Carriacou	220	—	.005	.995	—	—	.009
Grenada	194	—	—	.99	.01	—	.010

morphisms were adaptively neutral, allelic frequencies should be uncorrelated between the islands and the continental populations, and between different islands. In fact they are very similar.

The overall similarity of allelic frequencies throughout the distribution of *D. willistoni*, including the oceanic islands, could conceivably be explained by

postulating that allelic variants are selectively neutral but different alleles mutate at different rates. The frequencies of alleles would simply reflect the rates at which they arise by mutation. A most serious difficulty with this hypothesis is its *ad hoc* character. It is advanced *post facto* to account for an observed state of affairs. There is no evidence whatsoever to support the claim that alleles found in natural populations at high frequencies arise by mutation at higher rates than alleles occurring at low frequencies. Moreover, this hypothesis fails to account for the existence of local and regional differences in gene frequencies. Enough has been said above about local differentiation. Here I will point out two major instances of regional differentiation. Allele 1.03 of the *Lap-5* locus is the most common in the eleven continental populations of Table II, and in fact through most of the distribution area of *D. willistoni* [5]. In the six islands of Table VIII, allele 1.00 becomes the most common. At the *Est-7* locus, the frequency of allele 1.00 increases from about 0.52 in continental populations (see Table III and [5]) to about 0.67 in the six islands of the lesser Antilles (Table IX). Local differentiation and regional clines in gene frequencies have been observed in allozyme polymorphisms studied in other organisms [1], [23], [24].

6. Genetic variation between species

My colleagues and I have studied for several years allozyme variation in several species of the *D. willistoni* group, especially *D. willistoni* and its siblings, *D. tropicalis*, *D. equinoxialis*, and *D. paulistorum*. These species have largely overlapping geographic distributions. Often they can all be found in the same collection sites. The comparative study of genetic variation in these species provides important clues as to the evolutionary significance of the polymorphisms. Only a brief summary of the relevant facts can be presented here.

The first generalization concerns the amount of genetic variation. Similar degrees of overall variation are found in the sibling species. As stated above, the average proportion of polymorphic loci per individual is 17.7 per cent in *D. willistoni*, 18.6 per cent in *D. tropicalis*, and 21.3 per cent in *D. equinoxialis*. Moreover, at a given locus the amount of genetic variation is positively correlated between the species. If a locus is very polymorphic in one species, it is in most cases also highly polymorphic in the other sibling species. If a given locus is nearly monomorphic in one species, it is generally so in the other species (Table XIV). The correlation coefficients in the amount of heterozygosity per locus are: *D. willistoni-D. tropicalis*, 0.83; *D. willistoni-D. equinoxialis*, 0.60; *D. tropicalis-D. equinoxialis*, 0.70.

Let us now turn to the *pattern* of the variation. First, I shall consider loci with high levels of polymorphism. Tables XV and XVI give the allelic frequencies of *Lap-5* and *Est-7* in *D. tropicalis*. Comparison of these two tables with Tables II and III reveal that the constellation of allelic frequencies is remarkably similar in the two species. At *Lap-5*, 1.03 is the most common allele in both species,

TABLE XIV

GENETIC VARIATION AT 27 LOCI IN THREE SPECIES OF *Drosophila*

Gene	Frequency of heterozygous individuals		
	D. willistoni	*D. tropicalis*	*D. equinoxialis*
Lap-5	.629	.545	.439
Est-2	.117	.137	.126
Est-3	.108	.321	.566
Est-4	.270	.321	.373
Est-5	.089	—	.090
Est-6	.285	.247	.256
Est-7	.601	.528	—
Aph-1	.136	.183	.133
Acph-1	.102	.124	.308
Acph-2	.151	—	.296
Adh	.103	.335	.232
Mdh-2	.040	.009	.010
α-Gpdh	.021	.016	.027
Idh	.086	.078	.088
G3pdh	.148	.099	.196
Odh-1	.180	.125	.293
Odh-2	.060	—	.165
Me-1	.079	.120	.025
Me-2	.235	.197	.308
To	.131	.101	.131
Tpi-2	.041	.041	.037
Pgm-1	.186	.038	.444
Adk-1	.527	.525	.474
Adk-2	.171	.137	.096
Hk-1	.099	.094	.151
Hk-2	.138	.112	.146
Hk-3	.039	.049	.122

alleles 1.00 and 1.05 occur at intermediate, and allele .98 at somewhat lower, frequencies. At *Est-7*, allele 1.00 is the most frequent in both species, .98 and 1.02 occur at intermediate frequencies, and alleles .96 and 1.05 at still lower frequencies. *D. willistoni* and *D. tropicalis* are reproductively completely isolated from each other. No gene exchange whatsoever occurs between them (see [29] and references therein). The gene pools of *D. willistoni* and *D. tropicalis* have evolved independently for probably many millions of generations. It is incredible that independent random sampling of these two gene pools would have produced after such a long process so similar distributions of allelic frequencies. We must conclude that at these two loci some directional process like natural selection is maintaining a similar pattern of genetic variation in the two species. We can generalize this finding. At every highly polymorphic locus that we have studied the pattern of allelic frequencies found in one of these sibling species is also found in some other sibling species.

We turn now to those loci which have intermediate or low levels of poly-

TABLE XV

GENETIC VARIATION AT THE *Lap-5* LOCUS OF *Drosophila tropicalis*

Locality	Sample size	Alleles .96	.98	1.00	1.03	1.05	Other	Frequency of heterozygous individuals
Jaque	560	—	.04	.23	.51	.20	.01	.640
Teresita	490	—	.02	.20	.51	.25	.02	.635
P. Lopez	128	—	.02	.34	.64	.01	—	.475
Betoyes	348	—	.01	.14	.84	.02	—	.282
Mitu	50	—	.04	.20	.56	.16	.04	.599
Caracas	96	—	—	.28	.68	.04	—	.454
Macapa	180	—	—	.09	.74	.12	.04	.402
Belem	226	—	—	.06	.74	.19	.01	.414
Santarem	70	.01	.07	.23	.43	.16	.11	.724
Tefe	24	—	.04	.25	.58	.13	—	.580
Mirassol	12	—	—	.25	.75	—	—	.375

morphism. I will include in this category those loci in which the most frequent allele has an average frequency of 0.90 or higher. About half of the 30 loci that we have studied fall in this category. These loci can be divided into two classes, which I shall call rigid and flexible loci.

Rigid loci are those with a predominant allele which is identical (as judged by the electrophoretic mobility of the allozyme for which it codes) in all the sibling species (and in all other species of the group which have been studied so far). An example of a rigid locus is the *α-Gpdh* gene. Allele 1.00 has an average frequency of about 0.99 in each of seven species of the group studied in our laboratory. The allelic frequencies of *α-Gpdh* in several localities of *D. tropicalis*

TABLE XVI

GENETIC VARIATION AT THE *Est-7* LOCUS OF *Drosophila tropicalis*

Locality	Sample size	Alleles .96	.98	1.00	1.02	1.05	Other	Frequency of heterozygous individuals
Jaque	484	.03	.13	.52	.29	.03	.004	.625
Teresita	510	.03	.13	.58	.23	.03	.002	.593
P. Lopez	120	.02	.14	.73	.08	.03	—	.436
Betoyes	236	.03	.18	.64	.15	.01	—	.535
Mitu	62	.02	.11	.68	.18	.02	—	.459
Macapa	160	.03	.09	.53	.33	.02	—	.594
Belem	162	.02	.09	.44	.43	.02	—	.609
Santarem	66	—	.02	.52	.41	.06	—	.476
Tefe	17	—	—	.76	.18	.06	—	.381

TABLE XVII

GENETIC VARIATION AT THE *α-Gpdh* LOCUS OF *Drosophila tropicalis*

Locality	Sample size	Alleles .94	1.00	1.06	Other	Frequency of heterozygous individuals
Jaque	56	—	1.00	—	—	.000
Teresita	48	—	1.00	—	—	.000
P. Lopez	142	—	1.00	—	—	.000
Betoyes	1222	.001	.998	.001	.001	.005
Caracas	66	.02	.97	.02	—	.059
Macapa	54	—	1.00	—	—	.000
Belem	222	.005	.99	.005	—	.018
Santarem	68	.01	.99	—	—	.029
Tefe	24	—	1.00	—	—	.000

and *D. equinoxialis* are given in Tables XVII and XVIII. About half of the loci with intermediate or low levels of polymorphism are rigid. Of the genes listed in Table I, *α-Gpdh, Idh, Me-2, Hk-2,* and *Hk-3* are rigid loci.

The most likely causal explanation of these rigid polymorphisms is that one and the same allele is favored by natural selection in all the species of the group. Within the physiological background of these species one of the allozymes is more effective than the others. It might be that the physicochemical properties of the enzyme are such that only one of its multiple configurations can perform all its essential functions.

Flexible loci are those in which the predominant allele is not the same in all species. An example of a flexible locus is *Mdh-2*. Tables XIX and XX give the allelic frequencies in *D. tropicalis* and *D. equinoxialis*. In *D. willistoni* (Tables VII and XIII), allele 1.00 is the most frequent in every locality. The most common allele in *D. tropicalis* is .86, and .94 in *D. equinoxialis*. Allele 1.00 is absent or has frequencies no higher than 0.01 in every population of *D. equi-*

TABLE XVIII

GENETIC VARIATION AT THE *α-Gpdh* LOCUS OF *Drosophila equinoxialis*

Locality	Sample size	Alleles .94	1.00	1.06	Other	Frequency of heterozygous individuals
Jaque	42	.02	.96	.02	—	.092
Teresita	56	—	1.00	—	—	.000
P. Lopez	362	.01	.99	.003	—	.023
Betoyes	598	.01	.98	.005	.002	.033
Caracas	44	.02	.95	—	.02	.088
Macapa	22	—	1.00	—	—	.000
Belem	46	.02	.98	—	—	.043
Tefe	502	.01	.99	.002	.002	.028

TABLE XIX

GENETIC VARIATION AT THE *Mdh-2* LOCUS OF *Drosophila tropicalis*

Locality	Sample size	Alleles .82	.86	.94	Other	Frequency of heterozygous individuals
Jaque	44	—	1.00	—	—	.000
Teresita	48	—	1.00	—	—	.000
P. Lopez	142	—	1.00	—	—	.000
Betoyes	1206	.001	.997	.003	—	.007
Caracas	66	—	1.00	—	—	.000
Macapa	54	—	1.00	—	—	.000
Belem	224	—	.996	—	.004	.009
Santarem	68	—	.99	—	.01	.029
Tefe	24	—	1.00	—	—	.000

noxialis and *D. tropicalis*. Allele .86 is absent or very rare in every population of *D. willistoni* and *D. equinoxialis*. Allele .94 is rarely found in *D. willistoni* or *D. tropicalis*. Other flexible loci in the *D. willistoni* group are *Est-5*, *Mdh-2*, *G3pdh*, *Odh-2*, *Me-1*, *To*, and *Tpi-2*.

What is the causal explanation of the situation encountered in these flexible loci? Clearly, it is not that only one of the allozymes has the physiological properties necessary to be fully functional. Different allozymes occur with high frequencies in different species. A conceivable explanation is that allelic variation at these loci is selectively neutral. It would then be a historical accident—the result of random sampling—that one allele has high frequency in one species but low frequency in some other species. However, if the alleles were adaptively identical in a given species their frequencies should vary from one local population to another. Yet we find that within a given species the allelic frequencies are highly correlated among the populations, including those of geographically isolated islands. (For *Mdh-2* see Tables VII, XIII, XIX, and XX; in particular,

TABLE XX

GENETIC VARIATION AT THE *Mdh-2* LOCUS OF *Drosophila equinoxialis*

Locality	Sample size	Alleles .86	.94	1.00	Frequency of heterozygous individuals
Jaque	40	—	1.00	—	.000
Teresita	24	—	1.00	—	.000
P. Lopez	362	.006	.99	.003	.019
Betoyes	598	—	.997	.003	.007
Caracas	44	.02	.98	—	.044
Macapa	24	—	1.00	—	.000
Belem	46	—	1.00	—	.000
Tefe	494	.002	.99	.01	.016

compare Table VII with Table XIII.) Considering the constancy of allelic frequencies throughout a given species, a more likely explanation is that different alleles are selectively favored in different species. This can be due to the well known phenomenon of genetic coadaptation [8]. Organisms are well integrated systems. The selective value of a given genetic variant is unlikely to be the same regardless of the remainder of the genotype. Rather, it depends on the genetic background in which it exists. Genetic coadaptation affecting an allozyme locus has been demonstrated in *D. pseudoobscura* [22]. Experiments in our laboratory (see below) have shown that allelic variants at the *Mdh-2* are not selectively neutral.

7. Evidence of natural selection

The study of allelic variation in natural populations of the *D. willistoni* group provides conclusive evidence against the theory of random walk, or non-Darwinian, evolution. The pattern of the variation within and between species is inconsistent with the hypothesis that most of the variation is adaptively neutral. Laboratory experiments described below provide direct evidence of the operation of natural selection maintaining several allozyme polymorphisms.

The rationale of these experiments is simple. For the study of each locus, two or more experimental populations are started in laboratory cages. The environmental treatment (amount of food and space, temperature, and so forth) as well as the genetic background is made identical. The only variable is the initial allelic frequency. The populations are allowed to run for a number of generations. Predictions as to the course of events are derived from the hypotheses of natural selection and of random walk. If natural selection is operative the allelic frequencies will change in a determinate fashion as to direction and rate of change. In a typical situation allelic frequencies in the various cages will converge towards equilibrium values, which are independent of the initial frequencies. The rate of change is a function of the genotypic selective values and permits the estimation of these values.

If the genetic variation is selectively neutral, changes in gene frequency will be mostly the result of accidents of sampling through the generations. These changes will be erratic as to direction. Their magnitude will be an inverse function of population size. If population size is in the thousands, changes in gene frequencies from one to another generation will be very small. If the alleles are adaptively identical but mutate at different rates, a directional process is superimposed over the random process of change. Changes in gene frequency due to differential mutation rate are, however, very slow. The rate of change in the frequency of a given allele cannot be greater than the mutation rate to that allele.

Using experimental populations, we have studied five gene loci in the *D. willistoni* group. Two loci, *Lap-5* and *Est-7*, are highly polymorphic in natural populations; one locus, *Est-5*, is moderately polymorphic; and two loci, *α-Gpdh* and *Mdh-2*, are nearly monomorphic. The *α-Gpdh* is a *rigid* locus; *Mdh-2* is a

flexible locus. These studies are still in progress. Only preliminary and brief accounts can be given here. Yet the significance of the results is already clear.

The studies at the *Lap-5*, *Est-7*, and *Est-5* are being conducted by Mr. Jeffrey Powell, a graduate student in our laboratory. Several hundred inseminated females of *D. willistoni* collected in March 1970 in Mirassol, São Paulo, Brazil, were brought to the laboratory. Each female was placed in a separate culture giving raise to a "strain." The genetic constitution of the strain was ascertained. Flies from about 50 different strains were used to start each experimental population. By advisedly combining the number of flies of each strain introduced in a given cage, the initial allelic frequencies can be manipulated as desired. Each allele introduced in a cage was present in at least 30 of the strains used to establish that cage. The allele was thus introduced in many different genetic combinations, which must be a fair sample of the genetic combinations in which it occurs in the natural population. The results are given in Figures 2, 3, and 4, where the frequency of a given allele (ordinate) has been plotted through time (abscissa). Each line gives the allelic frequency in a separate cage. Average population size is about 5,000 individuals per cage; one full generation takes approximately 30 days. The cages are kept at 25°C. Similar experiments are being conducted at 19°C with results qualitatively similar to those given in the figures.

The results at the *Lap-5* and *Est-5* loci (Figures 2 and 3) are unambiguous. In each case the populations converge towards an equilibrium frequency. This

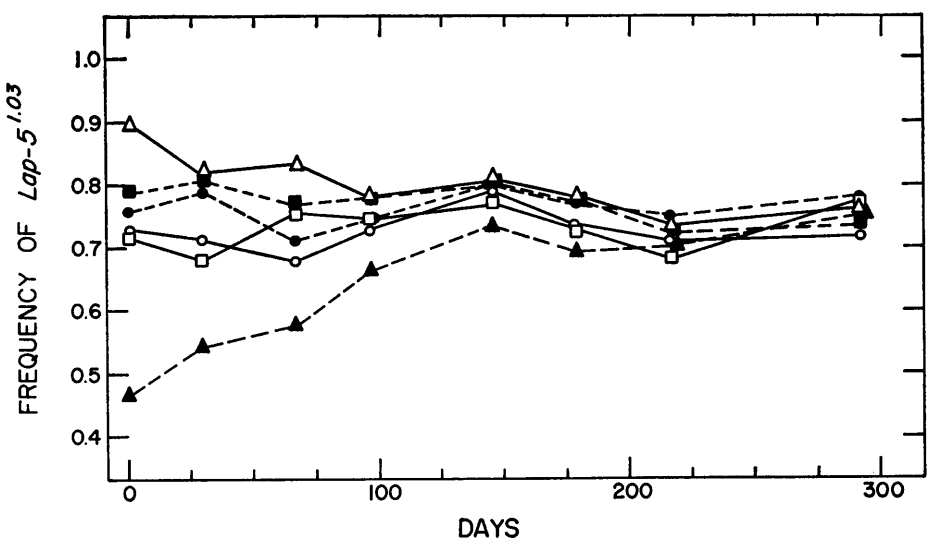

FIGURE 2

Changes in the frequency of allele 1.03 at the *Lap-5* locus in six experimental populations of *Drosophila willistoni*. The populations are kept at 25°C.

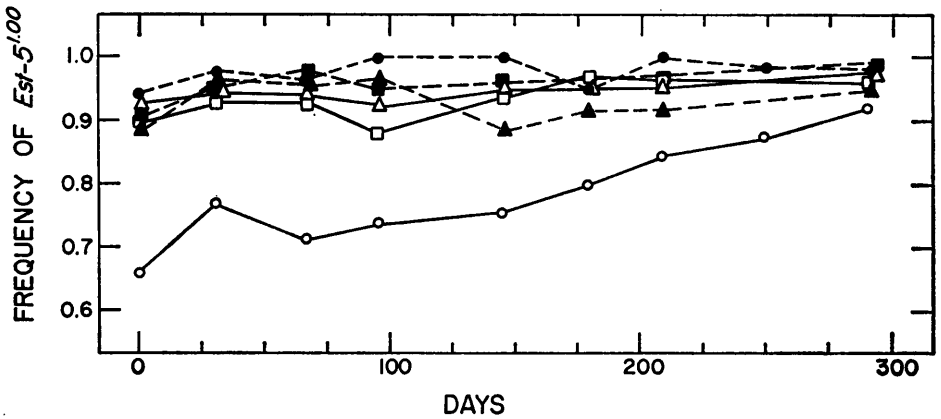

FIGURE 3

Changes in the frequency of allele 1.00 at the *Est-5* locus in six experimental populations of *Drosophila willistoni*. The populations are kept at 25°C.

is the outcome predicted by the hypothesis of natural selection. The results are incompatible with the hypothesis of adaptive neutrality. This latter hypothesis predicts that only erratic changes will occur in gene frequencies. There have not been enough generations for directional mutation to cause any appreciable change in allelic frequencies.

The results at the *Est-7* locus (Figure 4) are different. They show no large selective differences among the genotypes. Whether weak natural selection is occurring, or no selection at all is not clear. More generations are required to decide this question. The results are compatible with the hypothesis that the allelic variants in the experimental populations are selectively neutral. Experimental studies at the *Est-5* locus of *D. pseudoobscura* (which, like *Est-7* in *D. willistoni* is sex linked) have shown no evidence of natural selection either [30]. Ambiguous results have been obtained with a sex linked esterase locus in *D. melanogaster* [18]. We should recall, however, that in natural populations variation at the *Est-7* locus is not adaptively neutral.

The studies of the *α-Gpdh* and *Mdh-2* loci were made with the descendants of flies collected in the Llanos of Colombia. Two population cages with different initial frequencies were started for each species. Each population was started with about 30 different strains established as above. The results show that natural selection is operating at the *α-Gpdh* as well as at the *Mdh-2* locus. The initial frequencies of allele 1.00 at *α-Gpdh* were, in *D. equinoxialis*, 0.702 and 0.807. After three generations allele 1.00 has increased in both cages to frequencies 0.771 and 0.826, respectively. (The frequency of allele 1.00 in the natural population is 0.99.) The results in the *D. willistoni* and *D. tropicalis* cages are similar.

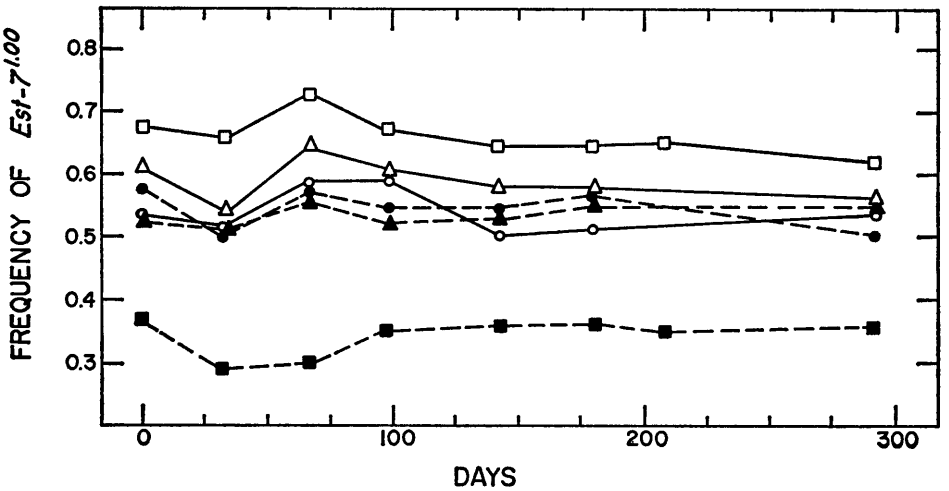

FIGURE 4

Changes in the frequency of allele 1.00 at the *Est-7* locus in six experimental populations of *Drosophila willistoni*. The populations are kept at 25°C.

Mdh-2 is a flexible locus in the *D. willistoni* group. In natural populations of *D. equinoxialis*, allele .94 has a frequency of about 0.99, while allele .86 occurs at frequencies usually lower than 0.01. In *D. tropicalis*, allele .94 occurs at frequencies lower than 0.01, while allele .86 reaches a frequency of 0.99 or higher. The results of this experiment are summarized in Table XXI. Although the environmental treatment is the same for all cages, the allelic frequencies follow opposite trends in the two species. Allele .86 decreases, while allele .94 increases in *D. equinoxialis*. In *D. tropicalis* allele .86 increases in frequency, while allele .94 decreases. There is evidence of natural selection at the *Mdh-2* in both species. Which allele is selectively favored depends on the genetic background in which it occurs. In the experimental populations of *D. equinoxialis*, as in natural

TABLE XXI

FREQUENCY OF ALLELE .86 AT THE *Mdh-2* LOCUS IN TWO EXPERIMENTAL POPULATIONS OF *Drosophila equinoxialis* AND TWO OF *D. tropicalis*

The frequency of allele .94 is one minus the frequency of allele .86.

Species	Cage	Initial frequency	Frequency after three generations
D. equinoxialis	I	0.161 ± 0.018	0.049 ± 0.010
D. equinoxialis	II	0.025 ± 0.002	0.009 ± 0.007
D. tropicalis	I	0.927 ± 0.013	0.977 ± 0.006
D. tropicalis	II	0.918 ± 0.014	0.970 ± 0.007

populations, allele .94 is at a selective advantage. In *D. tropicalis* allele .94 is selected against, while allele .86 is selectively favored.

8. Summary and conclusions

The theory of "non-Darwinian evolution" or "evolution by random walk" proposes that most genetic variation, particularly that observed at the molecular level, is adaptively neutral and therefore not subject to natural selection. The consequence follows that most of the genetic differences between species, espepecially differences responsible for amino acid substitutions in proteins, might be evolutionary "noise," with no adaptive significance. This paper summarizes a study of genetic polymorphisms in several species of the *Drosophila willistoni* group. The evidence indicates that most of the genic variation observed in these species is *not* adaptively neutral.

Drosophila willistoni, *D. equinoxialis*, and *D. tropicalis* are three sibling species whose geographic distributions overlap through much of the American tropics. These species possess large amounts of genetic variation. In *D. willistoni*, 56.3 ± 6.2 per cent loci are polymorphic in a given population. A locus is considered polymorphic when the frequency of the most common allele is no greater than 0.95. An individual *D. willistoni* is heterozygous at 17.7 ± 3.1 per cent of its loci. These figures are based on the study of 27 randomly selected genes coding for enzymes. Similar amounts of genetic variation exist in *D. equinoxialis* and *D. tropicalis*.

The amount of genetic variation varies considerably from locus to locus. However, at a given locus there is great similarity among different populations in the amount and in the pattern of genetic variation. Nevertheless, differences between populations occur, with some allelic frequencies being characteristic of certain regions or localities. The localities surveyed embrace an enormous territory extending from Panama to the Amazon delta and to the state of São Paulo in southern Brazil. They include six small oceanic islands. The similarity of the pattern of variation among continental populations covering such enormous territory, and between the continental and the isolated island populations is incompatible with the hypothesis that the polymorphisms are adaptively neutral. The similarity cannot be accounted for by migration between populations, since populations geographically isolated from each other have also similar genetic frequencies.

The amount of variation at a given locus is positively correlated among the sibling species. Comparisons between species as to the pattern of the variation are as follows. At very polymorphic loci similar distributions of allelic frequencies are found in different species. This result is incompatible with the hypothesis that allelic variation is selectively neutral at these highly variable loci.

Loci with intermediate and low levels of polymorphism can be classified as *rigid*—if one and the same allele is the most common one in all the species of the *willistoni* group, or *flexible*—when the most common allele varies from species

to species. The evidence supports the hypothesis that allelic frequencies at loci with low levels of variation are for the most part regulated by natural selection. Which allele is selectively favored depends in the case of flexible loci on the genetic background of the species.

The predominant role of natural selection in determining the amount and kind of genetic variation is confirmed by the study of change in gene frequencies in laboratory populations.

REFERENCES

[1] R. W. Allard and A. L. Kahler, "Patterns of molecular variation in plant populations,', *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability'* Berkeley and Los Angeles, University of California Press, 1972, Vol. 5, pp. 237–254.

[2] F. J. Ayala, "Biology as an autonomous science," *Amer. Sci.*, Vol. 56 (1968), pp. 207–221.

[3] F. J. Ayala, C. A. Mourão, S. Pérez-Salas, R. Richmond, and T. Dobzhansky, "Enzyme variability in the *Drosophila willistoni* group. I. Genetic differentiation among sibling species," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 67 (1970), pp. 225–232.

[4] F. J. Ayala, J. R. Powell, and T. Dobzhansky, "Enzyme variability in the *Drosophila willistoni* group. II. Polymorphisms in continental and island populations of *Drosophila willistoni*," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 68 (1971), pp. 2480–2483.

[5] F. J. Ayala, J. R. Powell, M. L. Tracey, C. A. Mourão, and S. Pérez-Salas, "Enzyme variability in the *Drosophila willistoni* group. IV. Genic variation in natural populations of *Drosophila willistoni*," *Genetics*, Vol. 70 (1972), pp. 113–139.

[6] J. M. Burns and F. M. Johnson, "Esterase polymorphism in the butterfly *Hemiargus isola*. Stability in a variable environment," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 68 (1971), pp. 34–37.

[7] J. F. Crow, "Darwinian and non-Darwinian evolution," *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1972, Vol. 5, pp. 1–22.

[8] T. Dobzhansky, *Genetics of the Evolutionary Process*, New York, Columbia University Press, 1970.

[9] J. H. Gillespie and K. Kojima, "The degree of polymorphism in enzymes involved in energy production compared to that in nonspecific enzymes in two *Drosophila ananassae* populations," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 61 (1968), pp. 582–585.

[10] H. Harris, *The Principles of Human Biochemical Genetics*, New York, Elsevier, 1970.

[11] M. Kimura, "Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles," *Genet. Res.*, Vol. 11 (1968), pp. 247–269.

[12] M. Kimura and J. F. Crow, "The number of alleles that can be maintained in a finite population," *Genetics*, Vol. 49 (1964), pp. 725–738.

[13] M. Kimura and T. Ohta, "Protein polymorphism as a phase of molecular evolution," *Nature*, Vol. 229 (1971), pp. 467–469.

[14] ———, "Population genetics, molecular biometry and evolution," *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1972, Vol. 5, pp. 43–68.

[15] J. L. King and T. H. Jukes, "Non-Darwinian evolution: Random fixation of selectively neutral mutations," *Science*, Vol. 164 (1969), pp. 788–798.

[16] K. Kojima, J. Gillespie, and Y. N. Tobari, "A profile of *Drosophila* species' enzymes assayed by electrophoresis. I. Number of alleles, heterozygosities, and linkage disequilib-

rium in glucose-metabolizing systems and some other enzymes," *Biochem. Genet.*, Vol. 4 (1970), pp. 627–637.

[17] R. C. LEWONTIN and J. L. HUBBY, "A molecular approach to the study of genic heterozygosity in natural populations. II. Amounts of variation and degree of heterozygosity in natural populations of *Drosophila psuedoobscura*," *Genetics*, Vol. 54 (1966), pp. 595–609.

[18] R. J. MACINTYRE and T. R. F. WRIGHT, "Response of esterase-6 alleles of *Drosophila melanogaster* and *Drosophila simulans* to selection in experimental populations," *Genetics*, Vol. 53 (1966), pp. 371–387.

[19] G. A. MROSS and R. F. DOOLITTLE, "Amino acid sequence studies on artiodactyl fibrinopeptides. II. Vicuna, elk, pronghorn, antelope and water buffalo," *Arch. Biochem. Biophys.*, Vol. 122 (1967), pp. 674–684.

[20] S. J. O'BRIEN and R. J. MACINTYRE, "An analysis of gene-enzyme variability in natural populations of *Drosophila melanogaster* and *D. simulans*," *Amer. Natur.*, Vol. 103 (1969), pp. 97–113.

[21] S. PRAKASH, "Genic variation in a natural population of *Drosophila persimilis*," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 62 (1969), pp. 778–784.

[22] S. PRAKASH and R. C. LEWONTIN, "A molecular approach to the study of genic heterozygosity in natural populations. III. Direct evidence of coadaptation in gene arrangements of *Drosophila*," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 59 (1967), pp. 398–405.

[23] S. PRAKASH, R. C. LEWONTIN, and J. L. HUBBY, "A molecular approach to the study of genic heterozygosity in natural populations. IV. Patterns of genic variation in central, marginal, and isolated populations of *Drosophila pseudoobscura*," *Genetics*, Vol. 61 (1969), pp. 841–858.

[24] R. C. RICHMOND, "Enzyme variability in the *Drosophila willistoni* group. III. Amounts of variability in the superspecies *D. paulistorum*," *Genetics*, Vol. 70 (1972), pp. 87–112.

[25] R. K. SELANDER, W. G. HUNT, and S. Y. YANG, "Protein polymorphism and genic heterozygosity in two European subspecies of the house mouse," *Evolution*, Vol. 23 (1969), pp. 379–390.

[26] R. K. SELANDER, M. H. SMITH, S. Y. YANG, W. E. JOHNSON, and J. B. GENTRY, "Biochemical polymorphism and systematics in the genus Peromyscus. I. Variation in the old-field mouse (*Peromyscus polyonotus*)," *Univ. Texas Publ.*, Vol. 7103 (1971), pp. 49–90.

[27] R. K. SELANDER and S. Y. YANG, "Protein polymorphism and genic heterozygosity in a wild population of the house mouse (*Mus musculus*)," *Genetics*, Vol. 63 (1969), pp. 653–667.

[28] R. K. SELANDER, S. Y. YANG, R. C. LEWONTIN, and W. E. JOHNSON, "Genetic variation in the horseshoe crab (*Limulus polyphemus*), a phylogenetic relic," *Evolution*, Vol. 24 (1970), pp. 402–414.

[29] B. SPASSKY, R. C. RICHMOND, S. PÉREZ-SALAS, O. PAVLOVSKY, C. A. MOURÃO, A. S. HUNTER, H. HOENIGSBERG, T. DOBZHANSKY, and F. J. AYALA, "Geography of the sibling species related to *Drosophila willistoni*, and of the semispecies of the *Drosophila paulistorum* complex," *Evolution*, Vol. 25 (1971), pp. 129–143.

[30] T. YAMAZAKI, "Measurement of fitness at the esterase-5 locus in *Drosophila pseudoobscura*," *Genetics*, Vol. 67 (1971), pp. 579–603.