Journal of Heredity 2003:94(3):236–242
DOI: 10.1093/jhered/esg021
© 2003 The American Genetic Association

The Mode of Evolution of Molecular Markers in Populations of House Mice Under Artificial Selection for Locomotor Behavior

T. J. Morgan, T. Garland, Jr., B. L. Irwin, J. G. Swallow, and P. A. Carter

From the School of Biological Sciences, Washington State University, Pullman, WA 99164-4236 (Morgan, Irwin, and Carter); Department of Biology, University of California, Riverside, CA 92521 (Garland); and Department of Biology, University of South Dakota, Vermillion, SD 57069 (Swallow). T. J. Morgan is currently at the Department of Genetics, College of Agriculture and Life Sciences, North Carolina State University, Raleigh, NC 27695.

Address correspondence to Theodore J. Morgan at the address above, or e-mail: tjmorgan@unity.ncsu.edu.

Abstract

A complete understanding of the mode of evolution of molecular markers is important for making inferences about different population genetic parameters, especially because a number of studies have reported patterns of allelic variation at molecular markers that are not in agreement with neutral evolutionary expectations. In the present study, house mice (*Mus domesticus*) from the fourteenth generation of a selection experiment for increased voluntary wheel-running activity were used to test how selection on a complex behavior affects the distribution of allelic variation by examining patterns of variation at six microsatellite and four allozyme loci. This population had a hierarchical structure that allowed for simultaneous testing of the effects of selection and genetic drift on the distribution of allelic variation by comparing observed patterns of allele frequencies and estimates of genetic divergence at multiple hierarchical levels to expectations under models of neutral evolution. The levels of genetic divergence among replicate lines and between selection groups, estimated from microsatellite data or pooled microsatellite and allozyme data, were not significantly different from expectations under neutral evolution. Furthermore, the pattern of change of allele frequencies between the base population and generation 14 was largely in agreement with expectations under neutral evolution (although the *PGM* locus exhibited a pattern of change within populations that was difficult to explain under neutral evolution). Overall the results generally provide support for the neutral evolution of molecular markers.

The patterns of genetic variation present within and among a subdivided system of populations are the result of selection, genetic drift, founder effects, migration, and mutation (Wright 1951). In populations that have been recently isolated and are finite in size, selection intensity at different loci will vary depending on the function of the gene products, but random genetic drift will affect all loci equally (Crow and Kimura 1970). Since the discovery of allozyme variation (Harris 1966; Lewontin and Hubby 1966), molecular markers have been used in many different types of studies which have revealed, among other things, insights into processes such as migration, genetic differentiation, founder effects, and hybridization in natural and experimental populations (Doherty et al. 1995; Lehmann et al. 1996; Wolf and Soltis 1992; Wolfe et al. 1998).

Although the utility of molecular markers in population biology has been high, how they evolve has been debated in the literature for decades (see reviews in Lewontin 1991; Mitton 1997; Watt 1994). Work that has attempted to address the mode of evolution of molecular markers usually takes one of two forms: studies either determine the correlation between single- or multilocus molecular marker heterozygosity and one or more components of Darwinian fitness (e.g., Pogson and Zouros 1994; Savolainen and Hedrick 1995; Watt et al. 1985), or they compare population genetic statistics (e.g., F_{ST}) calculated from multiple classes of molecular markers (e.g., Allendorf and Seeb 2000; Karl and Avise 1992; Latta and Mitton 1997; Pogson et al. 1995).

Several studies have applied both of these methods. For example, Pogson and Zouros (1994) found significant correlations between fitness and multilocus heterozygosity

at allozyme but not nuclear restriction fragment length polymorphism (RFLP) loci in the scallop. They argued that this discrepancy between allozyme and RFLP markers supported an overdominant model of selection acting on allozyme loci but not on RFLP loci. In another study, Karl and Avise (1992) found significant differences in allozyme and single-copy nuclear RFLP allele frequencies among populations of the American oyster, and concluded that the data were best explained by balancing selection on allozyme loci. Pogson et al. (1995) found similar patterns in six populations of Atlantic cod, in which variation at nuclear RFLP loci was enormous in contrast to the pattern observed at allozyme loci; again, the discrepancy was interpreted as evidence for selection acting on the allozyme variation across the populations sampled. Finally, Allendorf and Seeb (2000) genotyped sockeye salmon from four Alaskan populations at allozyme, microsatellite, randomly amplified polymorphic DNA (RAPD), and mitochondrial markers. Unlike the studies summarized above, they found similar patterns of allelic variation and genetic differentiation among the four classes of markers and concluded that all classes of markers were evolving neutrally.

The general result that some molecular markers may evolve in a nonneutral manner in some systems is important, because the conclusions of most studies of population structure generally rely on the assumption that molecular markers are evolving solely as a result of genetic drift, mutation, and migration. However, if molecular markers are evolving in a nonneutral manner, then estimation of parameters, such as the time since populations became isolated or migration rates, is difficult with simple population genetic models (Hartl and Clark 1989).

A complimentary method to those used in natural populations is to test the mode of evolution of molecular markers in a hierarchical population with a known evolutionary history at each level in the hierarchy. Other studies, primarily with *Drosophila melanogaster*, have investigated the evolution of molecular markers in laboratory populations with known evolutionary histories. For example, Cavener and Clegg (1981) found that selection for ethanol tolerance in *Drosophila* caused a change in the allele frequencies at the alcohol dehydrogenase, α-glycerophosphate dehydrogenase, malate dehydrogenase, and 6-phosphogluconate dehydrogenase loci. In addition, Oakeshott (1979) showed that the alcohol dehydrogenase locus responded to temperature selection, while Luckinbill et al. (1989) found that only glucose-6-phosphate dehydrogenase responded to selection on life span, and finally, Deckert-Cruz et al. (1997) found that selection for longevity resulted in a change in allele frequencies at phosphoglucomutase and CuZn-superoxide dismutase.

Multilevel hierarchical populations that have been produced and maintained in a laboratory are quite desirable for evaluating the mode of evolution of molecular markers because the allelic variation measured at the different hierarchical levels is attributable to a particular evolutionary process. For example, replicated artificial selection experiments have two or more selection groups, which differ in selection regimen, and each selection group contains two or more replicate lines, which should differ only as a result of stochastic processes (e.g., founder effects, genetic drift). Such hierarchical populations allow the total allelic variation to be partitioned among selection groups, among lines within selection groups, and among individuals within lines. Thus partitioning the effect of selection (i.e., variation among selection groups) and genetic drift (i.e., variation among lines within selection groups) on the patterns of variation observed at molecular markers is possible.

In the present study we investigated the distribution of allelic variation at microsatellite and allozyme loci in hierarchical populations of house mice from an artificial selection experiment for high levels of voluntary wheel running (Garland 2003; Swallow et al. 1998). Although the relationship between fitness and movement in natural populations varies, in this system wheel running is directly related to fitness in the selected lines because only individuals who run the most are allowed to breed. In addition, wheel-running activity may reflect components of many different behaviors, is affected by morphological, physiological, and biochemical traits, and may be of ecological relevance if it reflects general propensities or abilities for locomotion during dispersal, patrolling the territory, and traversing the home range (e.g., Boggs and Frappell 2000; Garland and Carter 1994; Kelt and Van Vuren 2001; Kramer and McLaughlin 2001; Miles et al. 2000). The populations used in this study have been isolated and maintained at known population sizes for 14 generations, while experiencing one of two selection regimens (i.e., selection for high wheel running versus unselected control). Because the experimental design has four replicate lines within each of the two selection groups, the effects of selection (i.e., differences between selection groups) and genetic drift (i.e., differences among lines) on the genetic divergence can be partitioned into different levels in the population hierarchy. If the assumption of neutral evolution of the microsatellite and allozyme loci is correct, then divergence among lines within selection groups and divergence between selection groups is not expected to differ significantly from theoretical expectations of neutrality. On the other hand, if the molecular markers or loci in linkage disequilibrium with these markers were influenced by artificial selection, then the neutrality of molecular markers would not be supported.

Materials and Methods

Establishment of Lines and Selection Regimen

The establishment of lines and selection regimen has already been described (Swallow et al. 1998), so only a brief description will be provided here. Male and female (112 of each) laboratory house mice (*Mus domesticus*) of the outbred Hsd:ICR strain were purchased from Harlan Sprague-Dawley (Indianapolis, IN). These individuals were paired randomly to produce generation −1. From generation −1, one male and one female were chosen randomly from each litter, and these individuals were paired randomly with the provision of no sib mating. Thirteen of these pairs were

assigned randomly to each of eight lines, and four lines were randomly assigned to each selection group (selection or control). Offspring from these pairings were designated as generation 0 and selection was begun at generation 1. Lines were maintained with 13 pairs per generation, although only 10 pairs were used per generation through generation 14.

In each generation, mice were weaned from the dams at 21 days of age, weighed, toe-clipped for individual identification, and housed in groups of four by sex. At 6–8 weeks of age the mice were placed in cages with activity wheels for 6 consecutive days. The selection trait was the average number of wheel revolutions on days 5 and 6. Within-family selection was used to reduce the effects of inbreeding; in selected lines, the highest-running male and the highest-running female from each family were chosen to breed, while in control lines, one male and one female from each family were randomly chosen to breed. Breeders were paired randomly within lines for 14 generations, with the condition of no sib mating.

Sampling of Animals for Genotyping

A subset of the animals (10 males and 10 females) from generation 14 of this artificial selection experiment were chosen for genotyping by randomly choosing one male and one female from each family. After production of generation 15 was complete, these individuals were sacrificed and their livers were removed and stored at −80°C until genotyping.

Microsatellite Genotyping

DNA was extracted by proteinase K/SDS digestion and phenol:chloroform:isoamyl alcohol extraction methods (Sambrook et al. 1989). DNA was resuspended in TE and stored at −20°C until used for polymerase chain reaction (PCR). Four microsatellite loci were chosen from Dietrich et al. (1992) (*D11Mit16*, *D7Mit18*, *D13Mit14*, *D15Mit16*) and two from Hearne et al. (1991) (*144* on chromosome 12 and *150* on chromosome 11). Microsatellite loci were chosen that were highly polymorphic across inbred strains of mice and were not in close linkage with each other. Microsatellite genotypes were determined following Dietrich et al. (1992), except 10 µl PCR reactions with 10–25 ng of DNA were used.

Samples were run on 8% denaturing acrylamide gels. After the run was complete, gels were visualized by silver staining. Gels were first fixed by gentle rocking in 10% ethanol for 5 min, then transferred to 1% nitric acid for 3 min, and then rinsed twice for 2 min in ddH_2O. The gel was then stained for 15–20 min in 5 mM $AgNO_3$ and 0.03% formaldehyde. The gel was then rinsed once in ddH_2O for 15–20 s, and then placed in developer (0.28 M sodium carbonate, 0.111% formaldehyde, and 0.016 mM sodium thiosulfate) until bands began to appear. Developing was terminated by adding an equal volume of 10% acetic acid and allowing the gel to rock for an additional 5 min. The gel was then rinsed with ddH_2O and a scanned into a computer using Adobe Photoshop 5.0.

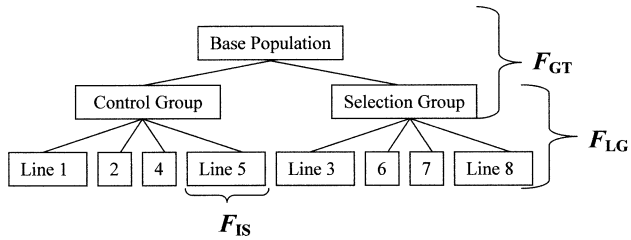

Figure 1. Population hierarchy. F_{LG} measures divergence between lines within selection groups, F_{GT} measures divergence between selection groups within the total population, and F_{IS} measures within-line inbreeding levels.

Allozyme Genotyping

Allozyme genotypes were assayed using the same liver tissue as for the microsatellites. Allozyme genotypes were determined using cellulose acetate electrophoresis (Hebert and Beaton 1989). Slight modifications to Hebert and Beaton's (1989) protocols were used following Carter et al. (1999). *PGI*, *PGM*, *MDH*, and *6PGD* were the only allozyme loci genotyped in generation 14 because an exhaustive survey of numerous allozyme loci found only these four loci to be polymorphic in the base population (Carter et al. 1999). *PGI*, *PGM*, *MDH*, and *6PGD* were not in close linkage with one another, with *PGI* residing on chromosome 7, *PGM* residing on chromosome 4, *MDH* residing on chromosome 5, and *6PGD* residing on chromosome 4.

Statistical Methods

Allele frequencies and levels of genetic divergence were estimated using Wright's (1951) F statistics in a three-level hierarchical analysis of variance (ANOVA) with sources of variation between selection groups, among lines within selection groups, among individuals within lines, and between alleles within individuals (Weir 1996, chap. 5; Weir and Cockerham 1984). Because of the additional hierarchical level (i.e., among lines within selection groups; Figure 1), F_{ST} was subdivided into F_{LG} and F_{GT} (Smouse and Long 1987). F_{LG} corresponds to the probability that two randomly chosen alleles within the same line are identical by descent relative to the particular selection group where they are nested, and F_{GT} corresponds to the probability that two randomly chosen alleles within the same selection group are identical by descent relative to the base population. Calculations of Wright's F statistics were computed using GDA version 1.0 (Lewis and Zaykin 2001). The significance of the F statistics was evaluated using 95% confidence intervals (CIs) that were calculated by 1,000 bootstrap replicates of the loci. This technique was used because it allows evaluation of the significance of each statistic without making assumptions about the distribution of the statistic.

The expected values of F_{LG} and F_{GT} under a neutral model were obtained by simulating the expected level of divergence among lines within selection groups and between

Table 1. Allele frequencies for control lines (C) and selected lines (S)

Locus	Line 1 (C)	Line 2 (C)	Line 4 (C)	Line 5 (C)	Line 3 (S)	Line 6 (S)	Line 7 (S)	Line 8 (S)
PGI (fast)	0.917	0.735	0.000	0.737	0.895	0.722	0.647	0.844
PGI (slow)	0.083	0.265	1.000	0.263	0.105	0.278	0.353	0.156
PGM (fast)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PGM (slow)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MDH (fast)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MDH (slow)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6PGD (fast)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
6PGD (slow)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
D11Mit16 (126)	0.528	0.176	0.618	0.316	0.132	0.237	0.139	0.438
D11Mit16 (135)	0.472	0.824	0.382	0.684	0.868	0.763	0.861	0.563
D7Mit18 (120)	0.361	0.588	0.444	0.921	0.763	0.684	0.806	0.531
D7Mit18 (131)	0.639	0.412	0.556	0.079	0.237	0.316	0.194	0.469
D13Mit14 (145)	0.765	0.735	1.000	0.605	0.342	0.684	0.472	0.531
D13Mit14 (150)	0.118	0.147	0.000	0.395	0.658	0.316	0.111	0.156
D13Mit14 (153)	0.118	0.118	0.000	0.000	0.000	0.000	0.417	0.313
144 (198)	0.444	0.676	0.250	0.289	0.316	0.132	0.528	0.531
144 (210)	0.556	0.324	0.694	0.711	0.421	0.842	0.361	0.469
144 (230)	0.000	0.000	0.056	0.000	0.263	0.026	0.111	0.000
D15Mit16 (127)	0.139	0.382	0.176	0.395	0.105	0.316	0.028	0.500
D15Mit16 (130)	0.222	0.235	0.059	0.079	0.184	0.158	0.111	0.125
D15Mit16 (133)	0.611	0.294	0.765	0.526	0.632	0.421	0.444	0.219
D15Mit16 (137)	0.028	0.088	0.000	0.000	0.079	0.105	0.417	0.156
150 (100)	0.765	0.088	0.000	0.368	0.211	0.395	0.375	0.000
150 (102)	0.147	0.147	0.188	0.079	0.000	0.000	0.000	0.219
150 (108)	0.059	0.000	0.000	0.000	0.342	0.316	0.063	0.000
150 (110)	0.029	0.765	0.813	0.553	0.447	0.289	0.563	0.781

Locus names are in italics with allele names in parentheses.

selection groups, based on the estimated effective population size of each line (~35) for the first 14 generations. The effective population size in each line of each generation was calculated as

$$N_{e_i} = \frac{4N_i - 2}{V_{k_i} + 2},$$

where N_i is the census population size and V_{ki} is the among-individual variance in progeny number within each line. The simulation took a number of random samples equal to the effective population size of a particular line from a binomial distribution, with the probability of sampling a particular allele equal to the frequency of that allele in the previous generation. The simulation was run 1,000 times based on the effective population size of each line and an initial allele frequency of 0.5 (although this does not influence the null distribution of F_{LG} and F_{GT}). F_{LG} and F_{GT} were estimated from each simulated allele frequency data set following Weir and Cockerham (1984). The observed values of F_{LG} and F_{GT} were then evaluated relative to the null distribution of F_{LG} and F_{GT} by the percentile method to determine if the observed levels of divergence were significantly different than the level expected under neutrality (Manly 1997).

Results

All of the allozyme loci assayed in this study were polymorphic in the base population (Carter et al. 1999). During the 14 generations of the artificial selection experiment, *PGM*, *MDH*, *and 6PGD* went to fixation for the same allele at each locus in all four selected lines and in all four control lines, while *PGI* remained polymorphic in seven of eight lines (Table 1). In the base population, the allele frequencies were highly skewed at *MDH* and *6PGD* (0.928 and 0.981, respectively) for the fast allele, which became fixed by generation 14. *PGM* and *PGI* had less skewed allele frequencies for the most common allele (0.639 and 0.577, respectively) (Carter et al. 1999), and microsatellite loci were not scored in the base population.

All of the lines in both the selected and control groups had values of F_{IS} that were not significantly different from zero (Table 2), suggesting that within-family selection and disallowing sib mating has been effective in limiting the levels of inbreeding within lines. For the microsatellite loci, estimates of F_{GT} ranged from –0.032 to 0.052, with a mean estimate of 0.001, which was not significantly different from zero (Table 3). The mean estimate of F_{GT} with all of the loci (i.e., including the data for *PGM*) was –0.003, which was also not significantly different from zero. These results suggest that no significant genetic divergence had occurred between the selection groups. In contrast, the estimates of F_{LG} from microsatellite loci ranged from 0.095 to 0.218, and the mean estimate over all microsatellite loci was 0.149, which was significantly different from zero (Table 3). The mean estimate of F_{LG} with both microsatellite loci and the one allozyme locus included was 0.173, which was also significantly different from zero (Table 3). These results indicate that a significant amount of genetic divergence has

Table 2. Estimates of the reduction in heterozygosity (F_{IS}) caused by nonrandom mating within lines with 95% CIs generated by bootstrapping the loci

Line	F_{IS}	95% CI
Control lines		
Line 1	−0.066	−0.158, 0.033
Line 2	−0.046	−0.200, 0.158
Line 4	0.036	−0.347, 0.330
Line 5	0.016	−0.161, 0.154
Selected lines		
Line 3	0.135	−0.138, 0.431
Line 6	−0.064	−0.232, 0.130
Line 7	0.095	−0.120, 0.291
Line 8	0.173	−0.043, 0.342

occurred among lines within each selection group. In addition, the observed values of F_{LG} and F_{GT} based on microsatellite and pooled data were not significantly different from the expected level of divergence under neutrality for F_{LG} ($P_{microsat} = .065$ and $P_{pooled} = .084$) and F_{GT} ($P_{microsat} = .666$ and $P_{pooled} = .654$; Table 3 and Figure 2).

Table 3. Wright's F_{ST}

	Subdivided F_{ST}	
Locus	F_{LG}	F_{GT}
PGI	0.348	−0.032
D11Mit16	0.146	0.039
D7Mit18	0.129	−0.011
D13Mit14	0.208	0.052
144	0.095	−0.032
D15Mit16	0.096	−0.001
150	0.218	−0.036
$Mean_{microsat}$	0.149 (0.108–0.192)	0.001 (−0.023–0.029)
$Mean_{pooled\ data}$	0.173 (0.120–0.236)	−0.003 (−0.025–0.022)
$P_{microsat}$	.065	.666
$P_{pooled\ data}$	.084	.654

F_{ST} is subdivided into among-line within selection group divergence (F_{LG}) and between selection group divergence (F_{GT}). $Mean_{microsat}$ is the mean estimate based only on the microsatellite loci with 95% CIs in parentheses, while the $Mean_{pooled\ data}$ is the estimate based on both microsatellite and allozyme data with 95% CIs in parentheses. *P* values test whether the observed values of F_{LG} and F_{GT} are significantly different from the distribution under neutrality (significance is evaluated based on a two-tailed test; Figure 2).

Discussion

Within each line, all of the microsatellite loci and one of four allozyme loci were polymorphic in generation 14 (Table 1). Three of the four allozyme loci that were polymorphic in the base population (Carter et al. 1999) went to fixation in all lines, while *PGI* remained polymorphic in all lines with the exception of line 4 (Table 1). The fixation of *MDH* and *6PGD* was expected because of the highly skewed base population allele frequencies, and the probability of fixation by drift alone was .55 for *MDH* and .858 for *6PGD*. In addition, the finding that *PGI* remained polymorphic in seven of eight populations was also expected, as the *PGI* alleles were at frequencies of 0.557 and 0.422 in the base population. However, the *PGM* locus, which was polymorphic in the base population with allele frequencies of 0.639 and 0.361, went to fixation for the same allele in all of the lines in both the selected and control groups. This is unexpected because by genetic drift alone, the probability of fixation is quite small ($P = .028$).

Multiple explanations are possible for the changes in allele frequencies at *PGM* during the 14 generations of selection. First, this pattern may suggest the influence of a deterministic process, like selection, acting on *PGM* or on loci closely linked with *PGM*. However, the selection seems not to be related to the experimenter-imposed artificial selection, because *PGM* is fixed for the same allele in both selected and control groups. This nonneutral pattern of evolution measured at *PGM* is not uncommon in studies of allozyme variation (e.g., Carter 1997; Carter et al. 1999; Carter and Watt 1988; Mitton 1997; Pogson 1991; Pogson and Zouros 1994; Watt et al. 1985). However, most studies find that the patterns of allozyme variation generally fit a model of overdominance acting to maintain polymorphism. The pattern of variation measured here suggests that the region of the genome containing *PGM* appears to be evolving by directional selection. This argument is perplexing, because the effective population size within the lines is smaller than that of the base population; hence if uncontrolled laboratory selection was acting on one allele at *PGM* we would expect it to be fixed in the base population, as selection in the base population would be less affected by random genetic drift. In addition, unintentional directional selection for a single allele at *PGM* might have occurred at

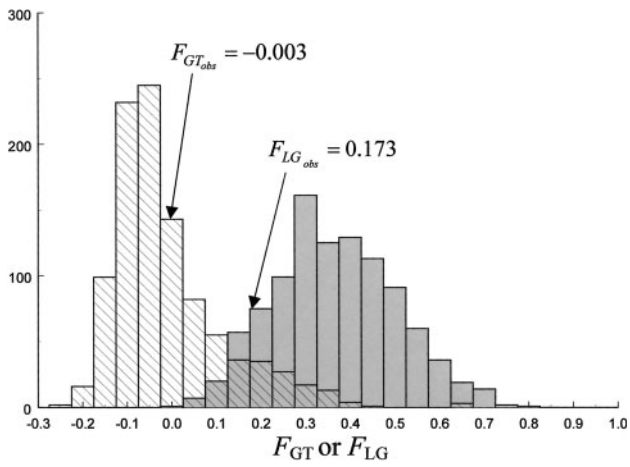

Figure 2. The null neutral distribution of F_{GT} (cross-hatched bars) and F_{LG} (gray bars) produced from 1,000 simulated data sets based on the effective population sizes of each line and the time since the populations became isolated. $F_{GT(obs)}$ and $F_{LG(obs)}$ are the observed levels of divergence estimated from the pooled molecular marker data.

the University of Wisconsin during the initiation and duration of the selection experiment.

A second explanation centers on the effects of sampling the base population to produce the founding population of the lines. *PGM* may be fixed in all lines as a result of a founder effect that was the result of the sampling of mice from the base population when the lines were founded. Although the probability of drawing an entire sample of 160 mice all containing the same allele at *PGM* is highly unlikely, the possibility exists that the sampling resulted in the first generation of mice having an increased frequency of the more common allele than in the base population. If this were the case, then random genetic drift might be strong enough to cause the fixation of this more common allele in 14 generations with highly reduced effective population sizes within each line. Unfortunately, determining which mode of evolution is influencing *PGM* is not possible given the restrictions of our data, but the second neutral explanation is in agreement with the dynamics of the other microsatellite and allozyme loci assayed in this study.

F_{IS} was also used to determine the level of genetic diversity within lines. F_{IS} is a measure of the observed level of heterozygosity within a line relative to the expected levels of heterozygosity under the assumption of random mating within a subpopulation (Table 2). If individuals within a population are nonrandomly mating, then F_{IS} will be significantly greater than zero. F_{IS} estimates were not significantly different from zero in any of the lines within either selection group (Table 2). This indicates that using within-family selection with no sib mating in the selected lines and random mating and no sib mating in the control lines results in minimal amounts of inbreeding within each line.

The level of genetic differentiation among populations was evaluated between selection groups using F_{GT} and among lines within a selection group using F_{LG}. At the hierarchical level of the line within the selection group (Figure 1), each line should be experiencing the same selection as the other lines within that particular selection group, and divergence among lines within a selection group should be the result of founder effects, genetic drift, and mutation [see also Garland et al. (2002) concerning an allele with major effects on muscle mass]. Significant levels of genetic divergence occurred among lines within selection groups (i.e., $F_{LG} \neq 0$). F_{LG} was significantly different from zero when estimated using microsatellites or the pooled data. In addition, the expected value of F_{LG} under a neutral model was not significantly different from F_{LG} estimated from microsatellite or pooled microsatellite and allozyme data (Table 3), which indicated that the microsatellite and allozyme markers are evolving neutrally among lines within selection groups. At the level of selection group, essentially no genetic divergence was expected between the selection groups under a neutral model (Table 3 and Figure 2), and the observed genetic divergence (F_{GT}) between selection groups does not differ from this expectation (Table 3). In addition, the estimated value of F_{GT} for the microsatellite and pooled data were not significantly different from zero, and the estimated 95% CI included the mean estimate of F_{GT} under the neutral model (Table 3). If the molecular markers were associated with the imposed directional selection on wheel running, then higher than expected divergence should have been measured between selection groups. Given that the observed divergence is in agreement with the expected value at both the hierarchical levels within the population, we can conclude that molecular markers are evolving essentially neutrally among lines within selection groups as well as between selection groups.

In summary, genetic divergence between selection groups, genetic divergence among replicate lines within a selection group, and potentially the patterns of change in the allelic variation between the base population and generation 14 at all of the allozyme loci were not different from expectations under a model of neutral evolution. These results on the distribution of allelic variation within and among multiple levels of a hierarchical laboratory population with a known evolutionary history largely support the neutral evolution of molecular markers in closed populations undergoing selection for an integrated locomotor behavior. However, as demonstrated in many studies of natural populations, allozymes (e.g., *PGM*) may not consistently agree with expectations of neutral evolutionary change.

Acknowledgments

We would like to thank M. Morgan, M. Evans, N. Copeland, and two anonymous reviewers for helpful comments and discussions on previous versions of this manuscript. This research was supported in part by National Science Foundation grants DEB-0083638 (to P.A.C.), DEB-0105079 (to P.A.C. and T.J.M.), and IBN-9728434 and IBN-0212567 (to T.G.).

References

Allendorf FW and Seeb LW, 2000. Concordance of genetic divergence among sockeye salmon population at allozyme, nuclear DNA, and mitochondrial DNA markers. Evolution 54:640–651.

Boggs DF and Frappell PB, 2000. Unifying principles of locomotion: foreword. Physiol Biochem Zool 73:647–650.

Carter PA, 1997. Maintenance of the Adh polymorphism in *Ambystoma tigrinum* (tiger salamanders). I. Genotypic differences in time to metamorphosis in extreme oxygen environments. Heredity 78:101–109.

Carter PA, Garland TJ, Dohm MR, and Hayes JP, 1999. Genetic variation and correlations between genotype and locomotor physiology in outbred laboratory house mice (*Mus domesticus*). Comp Biochem Physiol A 123:155–162.

Carter PA and Watt WB, 1988. Adaptation at specific loci. V. Metabolically adjacent enzyme loci may have very distinct experiences of selective pressures. Genetics 119:913–924.

Cavener DR and Clegg MT, 1981. Multigenic response to ethanol in *Drosophila melanogaster*. Evolution 35:1–10.

Crow JF and Kimura M, 1970. An introduction to population genetics theory. New York: Harper and Row.

Deckert-Cruz DJ, Tyler RH, Landmesser JE, and Rose MR, 1997. Allozymic differentiation in response to laboratory demographic selection of *Drosophila melanogaster*. Evolution 51:865–872.

Dietrich W, Katz H, Lincoln SE, Shin HS, Friedman J, Dracopoli NC, and Lander ES, 1992. A genetic map of the mouse suitable for typing intraspecific crosses. Genetics 131:423–447.

Doherty PJ, Planes S, and Mather P, 1995. Gene flow and larval duration in seven species of fish from the Great Barrier Reef. Ecology 76:2373–2391.

Garland T Jr, 2003. Selection experiments: an underutilized tool in biomechanics and organismal biology In: Vertebrate biomechanics and evolution (Bels VL, Gasc J-P, and Casinos A, eds). Oxford: BIOS Scientific Publishers; 23–56.

Garland T Jr and Carter PA, 1994. Evolutionary physiology. Annu Rev Physiol 56:579–621.

Garland T Jr, Morgan MT, Swallow JG, Rhodes JS, Girard I, Belter JG, and Carter PA, 2002. Evolution of a small-muscle polymorphism in lines of house mice selected for high activity levels. Evolution 56:1267–1275.

Hartl DL and Clark AG, 1989. Principles of population genetics. Sunderland, MA: Sinauer Associates.

Harris H, 1966. Enzyme polymorphisms in man. Proc R Soc Lond B 164:298–310.

Hearne CM, McAleer MA, Love JM, Aitman TJ, Cornall RJ, Ghosh S, Knight AM, Prins JB, and Todd JA, 1991. Additional microsatellite markers for mouse genome mapping. Mamm Genome 1:273–282.

Hebert PD and Beaton MJ, 1989. Methodologies for allozyme analysis using cellulose acetate electrophoresis: a practical handbook. Beaumont, TX: Helena Laboratories.

Karl SA and Avise JC, 1992. Balancing selection at allozyme loci in oysters: implications from nuclear RFLPs. Science 256:100–102.

Kelt DA and Van Vuren DH, 2001. The ecology and macroecology of mammalian home range area. Am Nat 157:637–645.

Kramer DL and McLaughlin RL, 2001. The behavioral ecology of intermittent locomotion. Am Zool 41:137–153.

Latta R and Mitton J, 1997. A comparison of population differentiation across four classes of gene marker in limber pine (*Pinus flexilis* James). Genetics 146:1153–1163.

Lehmann T, Hawley WA, Kamau L, Fontenille D, Simard F, and Collins FH, 1996. Genetic differentiation of *Anopheles gambiae* populations from east and west Africa: comparison of microsatellite and allozyme loci. Heredity 77:192–200.

Lewis PO and Zaykin D, 2001. Genetic data analysis: computer program for the analysis of allelic data, version 1.0 (d16c). Available at http://lewis.eeb.uconn.edu/lewishome/software.html.

Lewontin RC, 1991. Twenty-five years ago in genetics: electrophoresis in the development of evolutionary genetics: milestone or millstone? Genetics 128:657–662.

Lewontin RC and Hubby JL, 1966. A molecular approach to the study of genetic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. Genetics 54:595–609.

Luckinbill LS, Grudzien TA, Rhine S, and Weisman G, 1989. The genetic basis of adaptation to selection for longevity in *Drosophila melanogaster*. Evol Ecol 2:85–94.

Manly BFJ, 1997. Randomization, bootstrap and Monte Carlo methods in biology. London: Chapman and Hall.

Miles DB, Sinervo B, and Frankino WA, 2000. Reproductive burden, locomotor performance, and the cost of reproduction in free ranging lizards. Evolution 54:1386–1395.

Mitton JB, 1997. Selection in natural populations. Oxford: Oxford University Press.

Oakeshott JG, 1979. Selection affecting enzyme polymorphisms in laboratory populations of *Drosophila melanogaster*. Oecologia 143:341–354.

Pogson GH, 1991. Expression of overdominance for specific activity at the phosphoglucomutase-2 locus in the Pacific oyster, *Crassostrea gigas*. Genetics 128:133–141.

Pogson GH, Mesa KA, and Boutilier RG, 1995. Genetic population structure and gene flow in the Atlantic cod, *Gadus morhua*: a comparison of allozyme and nuclear RFLP loci. Genetics 139:375–385.

Pogson GH and Zouros E, 1994. Allozyme and RFLP heterozygosities as correlates of growth rate in the scallop *Placopecten magellanicus*: a test of the associative overdominance hypothesis. Genetics 137:221–231.

Sambrook J, Fritsch EF, and Maniatis T, 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Savolainen O and Hedrick P, 1995. Heterozygosity and fitness: no association in Scots pine. Genetics 140:755–817.

Smouse PE, and Long JC, 1987. A comparative F-statistics analysis of the genetic structure of human populations from lowland South America and highland New Guinea. In: Proceedings of the Second International Conference on Quantitative Genetics (Weir BS, Goodman MM, Eisen EJ, and Namkoong G, eds). Sunderland, MA: Sinauer Associates; 32–46.

Swallow JG, Carter PA, and Garland T Jr, 1998. Artificial selection for increased wheel-running behavior in house mice. Behav Genet 28:227–237.

Watt W, Carter PA, and Blower S, 1985. Adaptation at specific loci. IV. Differential mating success among glycolytic allozyme genotypes of *Colias* butterflies. Genetics 109:157–817.

Watt WB, 1994. Allozymes in evolutionary genetics: self-imposed burden or extraordinary tool? Genetics 136:11–16.

Weir B, 1996. Genetic data analysis II. Sunderland, MA: Sinauer Associates.

Weir B and Cockerham C, 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370.

Wolf P and Soltis P, 1992. Estimates of gene flow among populations, geographic races, and species in the *Ipomopsis aggregata* complex. Genetics 130:639–647.

Wolfe AD, Xiang QY, and Kephart SR, 1998. Assessing hybridization in natural populations of Penstemon (Scrophulariaceae) using hypervariable intersimple sequence repeat (ISSR) bands. Mol Ecol 7:1107–1125.

Wright S, 1951. The genetical structure of populations. Ann Eugen 15: 323–354.

Received January 28, 2002
Accepted December 16, 2002

Corresponding Editor: Neal Copeland