Adaptive evolution of a candidate gene for aging in *Drosophila*

Paul S. Schmidt*†, David D. Duvernell*‡, and Walter F. Eanes

Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, NY 11794-5245

Communicated by Robert R. Sokal, State University of New York at Stony Brook, Stony Brook, NY, July 20, 2000 (received for review May 11, 2000)

Examination of the phenotypic effects of specific mutations has been extensively used to identify candidate genes affecting traits of interest. However, such analyses do not reveal anything about the evolutionary forces acting at these loci, or whether standing allelic variation contributes to phenotypic variance in natural populations. The *Drosophila* gene *methuselah* (*mth*) has been proposed as having major effects on organismal stress response and longevity phenotype. Here, we examine patterns of polymorphism and divergence at *mth* in population level samples of *Drosophila melanogaster*, *D. simulans*, and *D. yakuba*. *Mth* has experienced an unusually high level of adaptive amino acid divergence concentrated in the intra- and extracellular loop domains of the receptor protein, suggesting the historical action of positive selection on those regions of the molecule that modulate signal transduction. Further analysis of single nucleotide polymorphisms (SNPs) in *D. melanogaster* provided evidence for contemporary and spatially variable selection at the *mth* locus. In ten surveyed populations, the most common *mth* haplotype exhibited a 40% cline in frequency that coincided with population level differences in multiple life-history traits including lifespan. This clinal pattern was not associated with any particular SNP in the coding region, indicating that selection is operating at a closely linked site that may be involved in gene expression. Together, these consistently nonneutral patterns of inter- and intraspecific variation suggest adaptive evolution of a signal transduction pathway that may modulate lifespan in nature.

The prevalence of genes of major effect for complex phenotypic traits such as lifespan remains unresolved. Although the number of genes underlying many quantitative traits may be quite high, there is mounting evidence that mutations at individual loci can substantially increase longevity in both *Caenorhabditis elegans* and *Drosophila melanogaster* (1, 2). The potential for such candidate genes to contribute to genetic variance for lifespan is predicated on the existence of functionally significant variation segregating at these loci. Presently, there is relatively little information on patterns of variation and evolutionary forces acting at these genes. Furthermore, it is unknown whether genetic variation in natural populations is actively maintained by selection (3, 4), or results from a balance between mutational input and selection against deleterious variants (5). These competing hypotheses generate distinct predictions regarding patterns of intraspecific polymorphism and interspecific divergence. Population level DNA sequence data may therefore be used to distinguish between the effects of neutral and nonneutral evolutionary forces on specific genes implicated in aging and senescence.

The *methuselah* (*mth*) gene has been proposed as such a candidate in *D. melanogaster*. Flies homozygous for a single P-element insertion in the third intron lived 35% longer and exhibited increased stress resistance when compared with the progenitor line (6). *mth* encodes a novel and yet uncharacterized GTP-binding protein-coupled seven-transmembrane domain receptor (GPCR), a diverse and functionally important group of receptor proteins involved in signal transduction. If age- and stress-related trait variation is modulated by signal transduction, it is of interest to determine which particular signaling elements of these pathways represent regulatory steps and may respond adaptively to selection on aging phenotype. Due to their position in the signal transduction pathway, regulatory control of the phenotypic response to a ligand-mediated signal may be primarily determined by the receptor (7). Consequently, genetic variation at receptor loci may have a larger effect on phenotypic expression than does variation at downstream loci. Directed mutations in other GPCRs have been shown to affect agonist-independent receptor activity (8), which can determine organismal phenotype (9).

In this paper, we examine evolutionary dynamics at the *mth* locus as well as geographic patterns of allele frequency variation in natural populations of *D. melanogaster*. The data reveal consistently nonneutral patterns of amino acid divergence in the intracellular (IC) and extracellular (EC) loop domains of the *mth* protein, whereas the null hypothesis of evolutionary neutrality cannot be rejected for transmembrane regions. The common *mth* haplotype exhibits a strong cline in frequency along the east coast of the United States, providing further indication of natural selection at the *mth* locus. These data do not address whether *mth* is evolving in response to selection on longevity phenotype in natural populations. Rather, the consistent patterns of selection and adaptation we observed at the *mth* locus are consistent with its proposed effects on organismal phenotype.

Materials and Methods

Population Samples. *D. melanogaster* and *D. simulans* (s) isofemale lines were collected in the fall of 1996 and 1997 from Homestead, FL (HFL97, HFL97s), Merritt Island, FL (MFL97), Jacksonville, FL (JFL97), Macon, GA (GA96s), Spartanburg, SC (SC96), Eutawville, SC (SC97), Smithfield, NC (NC97), Richmond, VA (VA97), Eastern Shore, VA (VA96s), Churchville, MD (MD97), Davis Peach Farm, Long Island, NY (DPF96, DPF96s), Middlefield, CT (CT97, CT96s), Concord, MA (MA97), and Whiting, VT (VT97). To generate homozygous third chromosomes, either the TM3 balancer chromosome was used in standard extractions of lethal-free third chromosomes (*D. melanogaster*) or isofemales lines were subjected to 14 generations of full-sib mating (*D. simulans*). East African *D. melanogaster* samples (ZIMH and ZIMS) were donated by the lab of C.-I. Wu (University of Chicago). One homozygous *D. yakuba* line was obtained from the Bowling Green Stock Center

Abbreviations: *mth*, methuselah; SNP, single nucleotide polymorphism; GPCR, GTP-binding protein-coupled seven-transmembrane domain receptor; EC, extracellular; IC, intracellular.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF280552–AF280589).

*P.S.S. and D.D.D. contributed equally to this work.

†To whom reprint requests should be addressed. E-mail: pschmidt@life.bio.sunysb.edu.

‡Present address: Department of Biological Sciences, Southern Illinois University, Edwardsville, IL 62026.

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

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.190338897. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.190338897

Table 1. McDonald–Kreitman tests of neutrality

	Fixed		Polymorphic		
	Syn.	Rep.	Syn.	Rep.	G
Complete coding region					
mel/sim	27	28	32	17	2.79
mel/sim/yak	107	123	53	31	6.83**
N-terminal region (219 codons)					
mel/sim	15	7	11	7	0.22
mel/sim/yak	57	69	24	17	2.20
Loop domains (157 codons)					
mel/sim	9	16	12	6	4.00*
mel/sim/yak	27	40	17	7	6.74**
Transmembrane domains (139 codons)					
mel/sim	3	5	9	4	2.05
mel/sim/yak	23	14	12	7	0.005

*, $P < 0.05$; **, $P < 0.01$.

(BG1013) and six isofemale lines were donated by the lab of T. F. C. Mackay (North Carolina State University).

Sequencing. Genomic DNA was extracted from whole flies (10). The *mth* gene was isolated by PCR using overlapping sets of primers; *D. melanogaster* primers were designed from genomic sequence obtained by a BLAST search of the Berkeley *Drosophila* Genome Database with the published *mth* cDNA sequence (6). Where necessary, species-specific primers were designed for *D. simulans* and *D. yakuba*, and both the 5′ and 3′ ends of the *mth* sequence in *D. yakuba* were obtained by inverse PCR (11). PCR products were purified and sequenced directly. Primary sequences for each species, ambiguous regions and replacement substitutions, and heterozygous sites in *D. yakuba* isofemale lines were verified on both strands. All sequences have been deposited in GenBank under accession numbers AF280552–AF280589.

Sequence Analysis. Sequences were aligned manually. DnaSPv3.14 (12) was used to generate estimates of the mutation parameter $4N_e\mu$ (θ; ref. 13) and average number of pairwise differences (π; ref. 14) and was also used to examine sequence divergence between species. The neutrality of the *mth* gene was tested by comparing ratios of synonymous to replacement changes for polymorphic and divergent sites (15).

SNPs Survey. A total of 533 third chromosomes were extracted by using the TM3 balancer from the following ten *D. melanogaster* populations: HFL97 ($n = 50$; 25.2°N), MFL97 ($n = 52$; 28.3°N), JFL97 ($n = 55$; 30.2°N), SC97 ($n = 55$; 33.2°N), NC97 ($n = 55$; 35.3°N), VA97 ($n = 55$; 37.3°N), MD97 ($n = 52$; 39.3°N), CT97 ($n = 55$; 41.2°N), MA97 ($n = 55$; 42.0°N), and VT97 ($n = 49$; 43.6°N). Population surveys of single nucleotide polymorphisms (SNPs) were conducted by restriction enzyme digestion of PCR products using *Bsp*HI (nucleotide position 531), *Bsm*I (pos. 1048), *Msp*I (pos. 1183), *Tsp*509I (pos. 1189), and *Pal*I (pos. 1244).

Results and Discussion

Divergence. Comparisons between the sibling species *D. melanogaster* and *D. simulans* and an outgroup, *D. yakuba*, reveal that *mth* is one of the fastest evolving genes in *Drosophila*. Twenty-eight replacement fixed differences were observed between *D. melanogaster* and *D. simulans* (Table 1), and approximately one in every six amino acids is different between these species and *D. yakuba*. This extreme rate of amino acid evolution is not the result of increased mutational pressure, because the level of silent divergence at *mth* between *D. melanogaster* and *D. simulans* ($K_s = 0.090$) is similar to that reported for other *Drosophila*

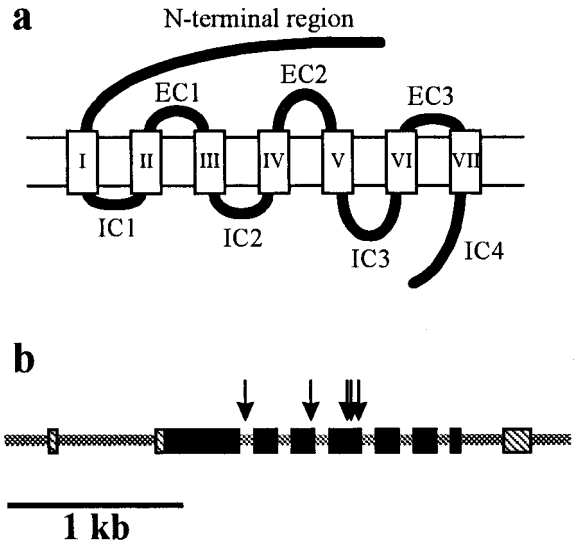

Fig. 1. (*a*) General two-dimensional model of a GPCR molecule. The *mth* N-terminal region is 219 aa long. The transmembrane domains (I–VII) include 139 aa, and the IC and EC loops, combined, include 157 aa. (*b*) Locations of the five SNP sites used in the analysis of geographic variation. Filled boxes indicate exons, and hatched boxes untranslated regions.

genes (16). The *mth* gene is clearly under selective constraints, because the level of interspecific divergence for replacement sites remains much lower than that for synonymous sites.

If the elevated level of amino acid replacements at *mth* relative to other *Drosophila* genes was solely a consequence of low functional constraints, neutral theory predicts that the ratio of silent to replacement substitutions should be the same for polymorphisms within species and fixed differences between species (15). This prediction forms the basis of the McDonald–Kreitman tests in Table 1. For the complete *mth* coding region, the total number of mutational events observed across all three species indicated a significant excess of replacement fixed differences. This pattern is in sharp contrast to that exhibited by other rapidly evolving *Drosophila* genes, for which the high level of amino acid divergence appears to result from a combination of reduced constraints and fixation of slightly deleterious variants (17). Given the level of silent divergence at *mth* between *D. yakuba* and *D. melanogaster*/*D. simulans* ($K_s = 0.26$), less than 5% of silent sites would be expected to have mutated more than once. Adjusting the dataset to account for multiple hits at silent sites does not substantially alter the significance of the McDonald–Kreitman test.

To examine which portions of the *mth* molecule had experienced adaptive evolution, we divided the coding region into functional domains based on homology of *mth* to other GPCRs and Kyte–Doolittle hydropathy profiles (Fig. 1*a*) (6). The large N-terminal segment of the *mth* protein (codons 1–219) projects outside the cell membrane and may be involved in ligand binding. McDonald–Kreitman tests for this region fail to reject the null hypothesis of neutrality (Table 1). The C-terminal region of the molecule (codons 220–514) comprises the more highly conserved seven transmembrane domains that characterize GPCRs and the intervening EC and IC loops. The evolutionary history of these two functionally disparate regions has been quite distinct. Whereas a neutral pattern of evolution was observed in the transmembrane domains, a significant excess of replacement fixed differences was evident in the loops for both the *D. melanogaster*/*D. simulans* comparison as well as the total number of mutational events across all three species (Table 1). The EC and IC loops are responsible for the correct insertion (18) and

Table 2. Nucleotide variation in *mth*

	n	Polymorphic sites			π		θ	
		Rep.	Syn.	Intron	Total	Syn.	Total	Syn.
D. melanogaster	31	9	10	5				
North America	22	4	6	1	0.0017	0.0058	0.0017	0.0047
Zimbabwe	9	8	8	5	0.0041	0.0089	0.0044	0.0084
D. simulans	13	8	22	5	0.0083	0.0251	0.0067	0.0202
D. yakuba	11*	14	21	12	NA	NA	0.0085	0.0193

*Three homozygous, four heterozygous individuals; NA, not available.

geometrical arrangement of the transmembrane domains (19) in other GPCRs, and also determine the activation state of both the receptor (20) and the coupled G-protein (21). Thus, the adaptive amino acid divergence appears localized to those regions of the protein that modulate signal transduction (7).

When the *D. yakuba* sequence was used to assign fixed differences to either *D. melanogaster* or *D. simulans*, no lineage-specific effects were evident. Excluding sites characterized by multiple substitutions, 11 of the 28 replacement fixed differences occurred on the *D. melanogaster* lineage and 11 on the *D. simulans* lineage. Of the 27 fixed differences at silent sites, 15 were unambiguously assigned to *D. melanogaster* and 7 to *D. simulans*. This discrepancy was not significant by the relative rate test of Tajima (22) ($\chi^2 = 2.91$, $P < 0.09$) but may reflect a faster rate of synonymous site evolution in *D. melanogaster*.

Patterns of codon usage at *mth* are highly unusual. The codon bias value we calculated for *mth* (codon adaptation index (CAI) = 0.264) is in the lowest ten percent for *Drosophila* genes (R. Kliman, unpublished data), and weak selection at silent sites would not be predicted (23). Fixed silent differences were again assigned to the *D. melanogaster* and *D. simulans* lineages by polarization with the outgroup, *D. yakuba*. In *D. simulans*, an equal number of fixed silent mutations were from an unpreferred to a preferred state ($n = 3$) and from a preferred to an unpreferred ($n = 2$) codon. However, a significant deviation from expectation was observed in *D. melanogaster*; only one fixed silent difference was from an unpreferred to a preferred codon, but ten fixed silent differences were from a preferred to an unpreferred state ($G = 8.55$, $P < 0.01$). This pattern may be typical of high codon bias genes (23), thought to reflect the reduced efficacy of weak selection in *D. melanogaster* because of lower effective population size (N_e), but is anomalous in a gene with such an apparent absence of codon bias.

Polymorphism. Similar to the data we obtained for divergence among species, the level of silent polymorphism at *mth* across species was similar to that of other genes, but amino acid polymorphism was unusually high (Table 2) (16). The proportion of polymorphisms that result in an amino acid change was higher in *D. melanogaster* (47%) than in *D. simulans* (27%), a commonly documented pattern that may reflect both the slightly deleterious nature of many amino acid polymorphisms and differences in N_e between species (16). Also consistent with this hypothesis is our observation that eight of nine amino acid polymorphisms in *D. melanogaster* were segregating at relatively low frequency (Fig. 2). It should be noted, however, that the proportion of replacement polymorphism for both *D. melanogaster* and *D. simulans* was approximately two times higher than the average for other genes in each species (16). A very high level of amino acid polymorphism was also observed in *D. yakuba* (Table 1).

The historical effects of positive or differentiating selection at *mth*, as indicated by the McDonald–Kreitman tests, did not coincide with a noticeably reduced level of intraspecific polymorphism in any species. Sufficient time has apparently elapsed since the most recent selective sweep(s) to allow the accumulation of polymorphism at *mth*. This inference was supported by examinations of site frequency distributions, which did not reveal any departures from neutrality by the Tajima test (24) (*D. melanogaster*: $D = 0.018$, $P > 0.10$; *D. simulans*: $D = 1.00$, $P > 0.10$). Similarly, the null hypothesis of evolutionary neutrality was not rejected by the Fu and Li test (25) (*D. melanogaster*: $D = 0.260$, $P > 0.10$; *D. simulans*: $D = 1.41$, $P > 0.05$).

Geographic Variation. We then investigated whether selection may be operating on specific polymorphic sites in the *mth* coding region in *D. melanogaster*. If variation at this locus contributes to genetic variance for lifespan, *mth* allele frequencies would be predicted to vary among populations that differ in longevity and other age-related traits. Isofemale lines collected from ten populations along a latitudinal gradient in North America demonstrated substantial clines in metabolic pools and life-history traits, and the northern- and southernmost populations

```
                      1111111111111
                 11333455780001112445788
                 26699533981142884179922
                360946115205686394110507
HFL97 15        GACTTCGgaCCTTCTCAACCCtta
DPF96  3        .......a.T..............
JFL97  1        .......a...C.T..........
JFL97  9        .......a...C.T..........
HFL97  8        .......a...C.T...C......
HFL97 12        .......a...C.T...C......
JFL97  5        .......a...C.T...C......
SC96  19        .......a...C.T...C......
CT97   4        .......a...C.T...C......
DPF96 74        .......a...C.T.ATC......
HFL97 16        .......a.....T.ATC......
MFL97  6        .......a.....T.ATC......
CT97   1        .......a.....T.ATC......
MA97   1        .......a.....T.ATC......
MA97   4        .......a.....T.ATC......
VT97   1        .......a.....T.ATC......
VT97  39        .......a.....T.ATC......
VT97  41        .......a.....T.ATC......
MA97   6        ..A....a.....T.ATC......
MFL97  1        ......A........ATC......
MFL97  3        ..............CATC......
CT97   3        .G....A........ATC......

ZIMH  26        .......a.....T.ATC......
ZIMH  44        ....G...t...G..ATC......
ZIMS  49        ....G...t...G..ATC......
ZIMH  23        .G.C.T...TT....ATC.TT...
ZIMS  24        .G.C.T...TT....ATC..T...
ZIMS  35        .G.C.T.........ATCA..c..
ZIMS  15        .G.C.T.........ATC...c.c
ZIMH  39        .G.C.T.........ATC...ccc
ZIMS  37        T.....A.t...............
```

Fig. 2. Variable sites among *D. melanogaster* nucleotide sequences of the *mth* coding region and six intervening introns. Variable intron sites are presented in lowercase and replacement sites are bold.

EVOLUTION

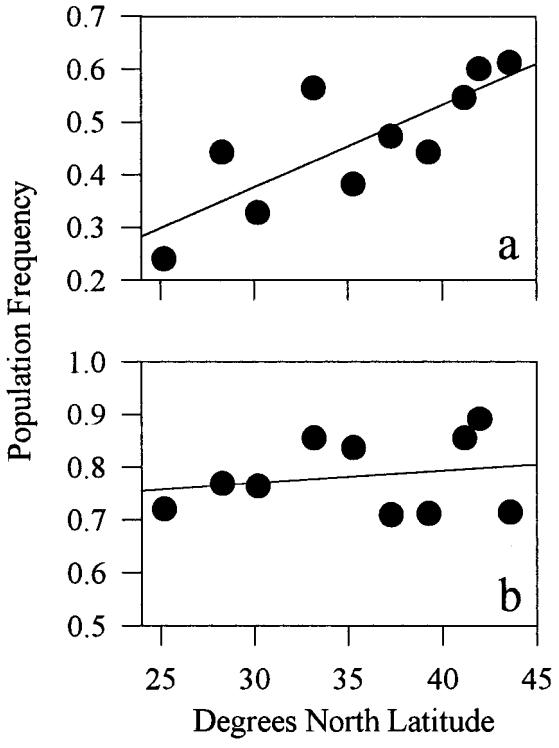

Fig. 3. Geographic variation in *mth* allele frequencies. (*a*) Frequency of the common ''ATATC'' haplotype vs. latitude (regression for arcsine square-root transformed data: $R^2 = 0.645$; $F = 14.29$, $P < 0.006$). (*b*) Frequency of SNP site 531 vs. latitude (analysis on transformed data: $R^2 = 0.044$; $F = 0.365$, $P > 0.56$).

significantly differed in average lifespan; Vermont males and females lived an average of 18% longer than Florida lines (L. M. Matzkin & W. F. Eanes, unpublished data). The pattern of geographic variation for longevity supports the life-history proposition that northern populations are characterized by increased lifespan. We scored 533 extracted third chromosomes from these ten populations for five of the six SNPs that reached appreciable frequencies in the 22 North American *mth* sequences listed in Fig. 2. The distribution of these SNP sites along the *mth* gene is illustrated in Fig. 1*b*. This analysis included a single intron site, three silent sites, and a charge-changing histidine/asparagine amino acid polymorphism in the receptor's third IC loop. In other GPCRs, this loop is responsible for G-protein coupling, and mutations of specific codons can have dramatic effects on receptor activity (8, 26). Although the frequency of the His/Asn replacement polymorphism was significantly variable among populations ($G = 22.24$, $P < 0.01$), none of the five individual SNP sites demonstrated a pattern deviating from the null model of no clinal variation (e.g., Fig. 3*b*). Given our relatively large sample size, which includes *mth* sequences from the highly divergent Zimbabwe population (27), it is unlikely that we failed to observe any moderate to high frequency amino acid polymorphisms that may be segregating in *D. melanogaster* populations.

However, *mth* haplotypes defined by the combination of nucleotide state at all of the five SNP sites demonstrated a strikingly clinal pattern; the most common *mth* haplotype, which constituted 36% of the haplotypes in the North American samples (Fig. 2), exhibited a 40% cline in frequency across populations (Fig. 3*a*). The clinally favored haplotype is present only in a single Zimbabwe allele, suggesting it may have risen to high frequency in northern North American populations subsequent to the species' introduction. The contrast between individual polymorphic sites and *mth* haplotypes suggests that selection is not operating at a particular site within the *mth* coding region in *D. melanogaster* populations (28), but rather on an allelic haplotype that is defined by a combination of substitutions related by genealogy. This inference of spatially variable selection is further supported by evidence of high rates of gene flow in this species (29) and the absence of clinal patterns for synonymous sites at other loci (28).

We interpret the *mth* cline as generated by disequilibrium between the overall coding region and a site that is under strong selection that varies along the sampled latitudinal gradient. *Mth* is located at band position 61C on the left arm of the third chromosome, a region predicted to be characterized by a slightly reduced rate of recombination (30). However, DNA sequence variation at band 61 is still predicted to be relatively high ($\theta = 0.0055–0.0085$ depending on strength of background selection and degree of dominance) (30), and our observed nucleotide diversity estimates for *mth* (Table 1) are only slightly below average (16). In our sample of 533 third chromosomes, the *In(3L)Payne* inversion was at very low frequency (<3%) and not associated with any one particular *mth* haplotype, therefore precluding any potential role of the inversion in generating the steep cline at *mth*. There is no evidence for larger regional effects of low recombination and genetic hitch-hiking seen in more extreme cases such as the tip of the X chromosome (31), and multiple recombination events are also evident in the *mth* dataset. Based on the SNPs data, a minimum of three recombination events were inferred between the following sites: 531–1048, 1048–1183, and 1189–1244. Furthermore, sequence data (Fig. 2) revealed two additional recombination events between positions 126–451 and 451–535.

In principle, the cline in haplotype frequencies at *mth* could be generated by selection on an unrelated but linked site or on a site within the functional *mth* gene itself. Either a recently established disequilibrium with a distant but functionally unrelated site, or a close physical association between the *mth* coding region and the site under selection, would be required to explain our observations. The SNP/haplotype analysis is currently being extended to include additional SNP sites encompassing the entire 5′ end of the *mth* transcript as well as sites in the two loci flanking *mth* in both the 5′ and 3′ directions. Further examination of patterns of linkage disequilibrium in this chromosomal region and geographic variation for individual polymorphic sites and multilocus haplotypes may localize which particular sites generate the observed cline.

Conclusions

Signal transduction through GPCRs may be modulated in two primary ways: either by variation in the effective number of receptors available to the ligand or by structural modifications of various interactions between the receptor and other molecules that determine signal amplification and termination (7). These interactions are determined by the GPCR loop domains, which are the portions of the molecule accessible for posttranslational modification. Our data suggest adaptive differentiation of the *mth* signal transduction pathway among *Drosophila* species. It remains to be seen whether other GPCR loci, in particular the various potential *mth* paralogs, evolve in a similar manner. As previously indicated (6), variation in the expression of *mth* may be a crucial determinant of the organismal response to stress and may greatly affect lifespan. In natural populations, selection on *mth* may involve a pleiotropic tradeoff between longevity and other fitness-associated traits that are negatively affected. These trade-offs may vary in a latitudinal fashion. Our observations are consistent with this hypothesis and suggest that variation in cis-acting regulatory regions may be driving the *mth* cline in *D. melanogaster*. However, the link between existing variation at *mth* and genetic variance for longevity in natural populations has

not been addressed. Functional studies of *mth* alleles identified in this study and further examination of the association between variation in *mth* allele frequency and longevity phenotype among natural populations should provide additional insights regarding the role of this gene in aging and senescence.

We thank Luciano Matzkin and Daniel MacDonald for generously providing access to unpublished data on clinal patterns of longevity and metabolic pools, and Brian Verrelli for supplying frequency data for *In(3L)Payne* in the *D. melanogaster* populations. This research was supported by U.S. Public Health Service Grant GM-45247 to W.F.E.

1. Kimura, K. D., Tissenbaum, H. A., Liu, Y. X. & Ruvkun, G. (1997) *Science* **277,** 942–944.
2. Orr, W. C. & Sohal, R. S. (1994) *Science* **263,** 1128–1130.
3. Gillespie, J. H. (1984) *Genetics* **107,** 321–330.
4. Gillespie, J. H. & Turelli, M. (1989) *Genetics* **121,** 129–138.
5. Barton, N. H. (1990) *Genetics* **124,** 773–782.
6. Lin, Y., Seroude, S. & Benzer, S. (1998) *Science* **282,** 943–946.
7. LeVine, H. (1999) *Mol. Neurobiol.* **19,** 111–149.
8. Lefkowitz, R. J., Cotecchia, S., Samama, P. & Costa, T. (1993) *Trends Pharmacol. Sci.* **14,** 303–307.
9. Orgiazzi, J. & Pradec, A. M. (1994) *Rev. Prat.* **44,** 1184–1191.
10. Winnepenninckx, B., Backeljau, T. & De Wachter, R. (1993) *Trends Genet.* **9,** 407.
11. Triglia, T., Peterson, M. G. & Kemp, D. J. (1988) *Nucleic Acids Res.* **16,** 8186.
12. Rozas, J. & Rozas, R. (1999) *Bioinformatics* **15,** 174–175.
13. Watterson, G. A. (1975) *Theor. Popul. Biol.* **7,** 256–276.
14. Nei, M. (1987) *Molecular Evolutionary Genetics* (Columbia Univ. Press, New York).
15. McDonald, J. H. & Kreitman, M. (1991) *Nature (London)* **351,** 652–653.
16. Moriyama, E. N. & Powell, J. R. *Mol. Biol. Evol.* **13,** 261–277.
17. Schmid, K. J., Nigro, L., Aquadro, C. F. & Tautz, D. (1999) *Genetics* **153,** 1717–1729.
18. Cramer, W. A., Engelman, D. M., Von Heijne, G. & Rees, D. C. (1992) *FASEB J.* **6,** 3397–3402.
19. Noda, K., Saad, Y., Graham, R. M. & Karnik, S. S. (1994) *J. Biol. Chem.* **269,** 6743–6752.
20. Mijares, A., Lebesque, D., Arjibay, J. & Hoebeke, J. (1996) *FEBS Lett.* **399,** 188–191.
21. Liu, J., Conklin, B. R., Blin, N., Yun, J. & Wess, J. (1995) *Proc. Nat. Acad. Sci. USA* **92,** 11642–11646.
22. Tajima, F. (1993) *Genetics* **135,** 599–607.
23. Akashi, H. (1995) *Genetics* **139,** 1067–1076.
24. Tajima, F. (1989) *Genetics* **123,** 585–595.
25. Fu, Y.-Y. & Li, W.-H. (1993) *Genetics* **133,** 693–709.
26. Hill-Eubanks, D., Burstein, E. S., Spalding, T. A., Brauner-Osborne, H. & Brann, M. R. (1996) *J. Biol. Chem.* **271,** 3058–3065.
27. Begun, D. J. & Aquadro, C. F. (1993) *Nature (London)* **365,** 548–550.
28. Berry, A. J. & Kreitman, M. (1993) *Genetics* **134,** 869–893.
29. Hale, L. R. & Singh, R. S. (1991) *Genetics* **129,** 103–117.
30. Hudson, R. R. & Kaplan, N. L. (1995) *Genetics* **141,** 1605–1617.
31. Begun, D. J. & Aquadro, C. F. (1992) *Nature (London)* **356,** 519–520.