Estimating the Influence of Selection on the Variable Amino Acid Sites of the Cytochrome *b* Protein Functional Domains

*David A. McClellan**† *and Kevin G. McCracken*‡

*Department of Zoology, Brigham Young University, Provo, Utah; †Institute of Statistical Mathematics, Tokyo, Japan, and Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama, Japan; and ‡Institute of Arctic Biology, and Department of Biology and Wildlife, University of Alaska, Fairbanks, Alaska

We evaluated the effects of selection on the molecular evolution of the functional domains of the mammalian cytochrome *b* gene as it relates to physicochemical properties shown to correlate with rates of amino acid replacement. Two groups of mammals were considered: pocket gophers of the rodent family Geomyidae, and cetaceans and ungulates of the monophyletic taxon Cetartiodactyla. Several characteristics of cytochrome *b* evolution were common to both mammal groups. The evolution of the matrix domain reflected the region's relative lack of function. Goodness of fit to neutral expectations indicated that external influences have had very little effect on the evolution of the matrix, although in some cases conservative and moderate changes have been favored. Although rates of synonymous nucleotide substitution have been relatively high, the transmembrane domain exhibited poor goodness of fit to neutral expectations. However, the evolution of the transmembrane domain has been constrained by negative selection, allowing a preponderance of conservative and moderate amino acid replacements. We hypothesize that a high rate of substitution is maintained in spite of negative selection because the codons of the transmembrane coding region are predisposed to conservative changes in all amino acid properties. The evolutionary patterns of the intermembrane domain in pocket gophers and cetartiodactyls, however, were very different. Changes inferred from the pocket gopher phylogenetic tree exhibited a significant fit to neutral expectations for each of the amino acid properties. Changes inferred from the cetartiodactyl tree exhibited significant fit to neutral expectations for polarity and isoelectric point, but not for composition, molecular volume, polar requirement, or hydropathy. In each case, lack of fit was due to selection that promoted conservative or moderate change, with the noteworthy exception of polar requirement. We detected an unexpectedly large change in polar requirement (from aspartic acid to threonine) in two separate lineages (*Camelus bactrianus* and all cetaceans) at amino acid position 159. This inferred change occurred in a region of the cyt-*b* protein that directly interacts with external surface proteins of the cytochrome bc_1 complex and resulted in a reversion to a more common character state in vertebrates.

Introduction

Cytochrome *b* (cyt-*b*) is the central catalytic subunit of the *Q*-cycle and the only membrane-bound molecule in the cytochrome bc_1 complex. The structure and function of cyt-*b* has been determined in great detail and in many cases down to the importance of individual amino acid residues (Degli Esposti et al. 1993; Zhang et al. 1998). Patterns of cyt-*b* nucleotide and amino acid substitution also have received much attention (e.g., Irwin, Kocher, and Wilson 1991; DeWalt et al. 1993; Graybeal 1993; Xia, Hafner, and Sudman 1996; Griffiths 1997). Such studies have evaluated cyt-*b* in terms of variation at many levels, such as among sites (e.g., codon positions), among base exchange types (i.e., transitions and transversions), and among regions or functional domains of the protein (i.e., intermembrane, transmembrane, and matrix).

Past studies of the molecular evolution of cyt-*b* indicate that the functional domains evolve at different rates of amino acid replacement (e.g., Irwin, Kocher, and Wilson 1991; DeWalt et al. 1993; Griffiths 1997). The common hypothesis for this phenomenon is that the selection responsible for maintaining the *Q*-cycle mechanism greatly influences rates of amino acid replacement. However, this hypothesis has yet to be critically evaluated in terms of evolutionary changes in the physicochemical properties that determine the phenotype of the protein. If the functional properties of the domains are being maintained by selection, very few amino acid replacements would be expected in those regions with the most residues implicated in the *Q*-cycle mechanism.

There are many physicochemical properties associated with amino acid residues (e.g., Sneath 1966; Woese et al. 1966; Alff-Steinberger 1969; Grantham 1974; Kyte and Doolittle 1982). Several of these have been shown to correlate with the evolution of the genetic code, the amino acid composition of polypeptides, and rates of amino acid replacement (e.g., Grantham 1974; Hughes, Ota, and Nei 1990; Haig and Hurst 1991; Xia and Li 1998). As a result, similar codons generally specify amino acids with similar physicochemical properties (Sonneborn 1965; Epstein 1966; Goldberg and Wittes 1966; Alff-Steinberger 1969). Furthermore, these properties contribute to a vast array of structural and functional characteristics of proteins that can be influenced directly by natural selection.

We evaluated the physicochemical changes that result from the molecular evolution of the cyt-*b* functional domains in pocket gophers and cetartiodactyls. Six amino acid properties shown to correlate with rates of amino acid replacement (Xia and Li 1998) were considered: composition of the side chain, polarity, and molecular volume (Grantham 1974), as well as polar requirement (Woese et al. 1966), hydropathy (Kyte and Doolittle

Key words: codon composition, physicochemical properties, amino acid replacement, cytochrome *b* functional domains, Geomyidae, Cetartiodactyla.

Address for correspondence and reprints: David A. McClellan, Department of Zoology, 574 WIDB, Brigham Young University, Provo, Utah 84602. E-mail: david_mcclellan@byu.edu.

Mol. Biol. Evol. 18(6):917–925. 2001

1982), and isoelectric point (Alff-Steinberger 1969). To evaluate the selective influences affecting the molecular evolution of cyt-*b*, we first calculated the goodness of fit between an observed distribution of physicochemical changes and an expected distribution based on codon composition (briefly outlined in McCracken et al. 1999) to measure how well the data fit neutral expectations. Next, we compared the inferred numbers of conservative, moderate, radical, and very radical physicochemical changes relative to the number of possible evolutionary pathways using a normal distribution to obtain relative measures (*z*-scores) of the selective influence on each amino acid property.

The goodness-of-fit score (GF-score) and the set of three *z*-scores can be interpreted as a measure of the selective influences on each gene region, with the magnitude of the GF-score indicative of the intensity of selection and the *z*-scores indicative of the direction of selection. For example, an observed distribution with a GF-score > 7.815 (df $= 3$, $\alpha = 0.05$) and a *z*-score > 1.645 ($\alpha = 0.05$) between the conservative and moderate magnitude classes with the conservative class being greater can be said to have experienced negative selection that favors conservative change. Alternatively, an observed distribution with a GF-score > 7.815 and a *z*-score > 1.645 between radical and very radical magnitude classes with the very radical class being greater can be said to have experienced positive selection that favors very radical change at those sites of amino acid replacement.

Materials and Methods

Pocket Gopher Phylogenetics

Cytochrome *b* sequences for 15 species of pocket gophers (Geomyidae: Rodentia) were downloaded from the National Center for Biotechnology Information. Phylogenetic analysis was performed using the general time-reversible (GTR) model of maximum-likelihood in PAUP*, version 4.0b1 (Swofford 1998), with maximum-likelihood estimates of substitution rates, nucleotide frequencies, proportion of invariable sites, and the gamma distribution shape parameter (α) producing the tree in figure 1*A.* This tree topology is consistent with trees constructed by Hafner et al. (1994) and Xia, Hafner, and Sudman (1996) using cytochrome oxidase subunit I (COI) and cyt-*b.*

Cetartiodactyl Phylogenetics

Complete cytochrome *b* sequences for representatives of eight artiodactyl families and four cetaceans were downloaded from the National Center for Biotechnology Information. Phylogenetic analysis was performed using methods similar to those listed above for pocket gophers, with the exception that the tree search was constrained using the basic relationships published by Nikaido, Rooney, and Okada (1999) based on the inferred occurrence of SINE and LINE insertions, which have been shown to be valuable phylogenetic markers (e.g., Nikaido, Rooney, and Okada 1999; Shedlock and Okada 2000; Shedlock, Milinkovitch, and Okada 2000). This analysis produced the tree in figure 1*B.*

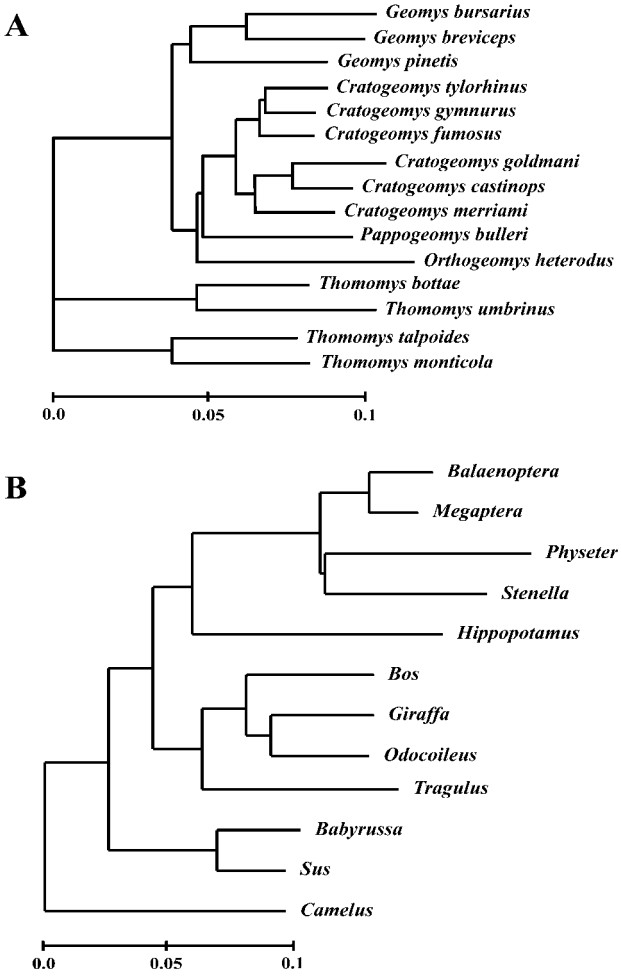

FIG. 1.—Phylogenetic relationships of (*A*) pocket gophers and (*B*) cetartiodactyls.

Substitution Analysis

Ancestral character state reconstructions were optimized using the GTR maximum-likelihood model, with the same parameter estimates used to construct the tree. The sequence at each branch terminus on the tree was compared with its immediate ancestral sequence to estimate the number and types of substitutions (similar to Li 1993; see also DeWalt et al. 1993; Xia, Hafner, and Sudman 1996; Xia 1998; McClellan 2000).

Information pertaining to each inferred nonsynonymous substitution was recorded, including codon position, type of base exchange (transition or transversion), and exact location in the protein. Substitutions were categorized further by the functional domain in which they occurred (Degli Esposti et al. 1993; Zhang et al. 1998). The magnitudes of physicochemical changes in composition, polarity, and molecular volume (Grantham 1974), as well as polar requirement (Woese et al. 1966), hydropathy (Kyte and Doolittle 1982), and isoelectric point (Alff-Steinberger 1969), resulting from each inferred nonsynonymous substitution were also recorded.

Expected Distribution of Physicochemical Changes

Distributions of changes in physicochemical properties of inferred amino acid replacements were compared with distributions calculated based on the assumption of completely random amino acid replacement (McCracken et al. 1999; Xia 2000) expected under the condition of selective neutrality. Expected frequencies per physicochemical property magnitude class were calculated by

$$F_i = \frac{\sum_{j=1}^{60} (r_{ij} n_j)}{\sum_{i=1}^{m} \sum_{j=1}^{60} (r_{ij} n_j)}, \quad (1)$$

where F_i is the expected relative frequency of amino acid replacements of magnitude class i ($i = 1, 2, \ldots, m$, where m is the greatest magnitude class in which the condition $F_i > 0$ prevails), r_{ij} is the number of possible evolutionary pathways of physicochemical property magnitude class i across all possible nonsynonymous substitutions of codon j ($j = 1, 2, \ldots, 60$), and n_j is the average number of times that codon j occurs in the extant DNA sequences. Equation (1) thus describes an average distribution of possible physicochemical differences. Thus, as selection decreases, the observed distribution of physicochemical differences of inferred amino acid replacements is predicted to take on a shape similar to the expected distribution.

The calculation of this expected distribution differs from that presented in McCracken et al. (1999) in that the magnitude class frequencies are not transformed by the nonsynonymous substitution frequencies expected under the codon degeneracy model (McClellan 2000). This is a null model of amino acid replacement and, as it applies to the present study, assumes completely random replacement. Thus, modification of the expected distribution of physicochemical changes with regard to patterns of codon degeneracy and an inferred transition bias is unnecessary.

The distributions of physicochemical changes inferred from the phylogenetic analysis of the pocket gopher cyt-*b* DNA sequences were compared with the expected distribution calculated with equation (1) using a χ^2 GF test of four magnitude categories: conservative, moderate, radical, and very radical. These categories are similar to those used by Wyckoff, Wang, and Wu (2000). Good fit ($P > 0.05$) between expected and observed distributions indicates that observed evolution of the DNA sequences is consistent with the assumption of selective neutrality. This is not to say, however, that the evolution of the DNA sequences is not influenced by selection. Good fit means only that the preponderance of sites experiencing amino acid replacements are accumulating changes as predicted by the model.

Detecting the Causes of Deviation from Expectations

The method developed by Hughes, Ota, and Nei (1990) extends Nei and Gojobori's (1986) method of estimating rates of synonymous and nonsynonymous nucleotide substitution. Nei and Gojobori (1986) categorized each nucleotide site as either synonymous, nonsynonymous, or a fraction thereof. Hughes, Ota, and Nei (1990) took this one step further by dividing each nonsynonymous site (or fractional nonsynonymous site) into either conservative, radical, or a fraction thereof based on the extent to which the possible evolutionary pathways result in either conservative or radical changes relative to a certain amino acid property. The number of inferred conservative substitutions is then divided by the mean number of conservative sites, and the number of inferred radical substitutions is divided by the mean number of radical sites. If these two ratios are equal, the amino acid replacements are said to take place at random relative to that particular property. If the ratio for more conservative replacements is greater than the ratio for more radical replacements, however, the property is being conserved, as would be expected under conditions of purifying selection. If the conservative ratio is less than the radical ratio, the replacements are said to promote radical changes, as would be expected under conditions of positive selection (Hughes, Ota, and Nei 1990).

One questionable aspect of this line of thought is the fractional assignment of a single nucleotide site as part conservative and part radical. We prefer to consider potential evolution by the number of possible nonsynonymous evolutionary pathways available to a given codon (Xia 1998; Xia and Li 1998). Another questionable aspect of the Hughes, Ota, and Nei (1990) model is the consideration of just two categories of amino acid replacement (conservative and radical) based on the qualitative properties described by Taylor (1986). We prefer to use four categories of replacement with respect to quantifiable amino acid properties (as in Wyckoff, Wang, and Wu 2000). Considering four categories (conservative, p_1; moderate, p_2; radical, p_3; and very radical, p_4) allows a more detailed analysis of more dynamic evolutionary trends. Each ratio is calculated by

$$p_i = \frac{D_i}{\sum_{j=1}^{60} r_{ij} n_j}, \quad (2)$$

where r_{ij} and n_j are as defined previously, and D_i is the number of inferred nonsynonymous substitutions resulting in a physicochemical change of magnitude class i ($i = 1, 2, 3, 4$, depending on its relative conservative or radical nature). The standard error (SE) for each proportion is calculated by

$$\mathrm{SE} = \sqrt{\frac{p_i(1 - p_i)}{\sum_{j=1}^{60} r_{ij} n_j}} \quad (3)$$

and evaluated using a standard z test for comparing two binomial proportions.

Results and Discussion

Cytochrome *b* Functional Domain Evolution

The cyt-*b* protein can be divided into three functional domains: intermembrane, matrix, and transmem-

Table 1
Rates (mean ± SE) of Cytochrome *b* Gene Evolution

Rates[a]	Matrix	Transmembrane	Intermembrane	Cytochrome *b*
Geomyids				
r_{syn}	0.390 ± 0.035	0.451 ± 0.020	0.448 ± 0.028	0.439 ± 0.015
r_{ns}	0.262 ± 0.055	0.338 ± 0.033	0.076 ± 0.026	0.253 ± 0.022
p_{invar}	0.785 ± 0051	0.790 ± 0.028	0.933 ± 0.024	0.829 ± 0.019
ns/site	2.278 ± 1.994	2.400 ± 1.829	1.539 ± 1.330	2.267 ± 1.811
Cetartiodactyls				
r_{syn}	0.609 ± 0.015	0.979 ± 0.004	0.943 ± 0.007	0.907 ± 0.009
r_{ns}	0.635 ± 0.025	0.895 ± 0.016	0.248 ± 0.022	0.672 ± 0.024
p_{invar}	0.651 ± 0.025	0.643 ± 0.025	0.886 ± 0.016	0.712 ± 0.023
ns/site	1.909 ± 1.269	2.960 ± 2.251	2.583 ± 2.065	2.706 ± 2.096

[a] r_{syn} = synonymous substitutions per nucleotide site; r_{ns} = nonsynonymous substitutions per amino acid site; p_{invar} = proportion of invariable amino acid sites; ns/site = nonsynonymous substitutions per variable site.

brane. The intermembrane domain, which extends between the inner and outer mitochondrial membranes, evolves significantly more slowly than the other two functional domains (Irwin, Kocher, and Wilson 1991; Griffiths 1997). The standard hypothesis for this phenomenon is that evolutionary constraints are imposed by the function of the Q_o redox center (Irwin, Kocher, and Wilson 1991; DeWalt et al. 1993; Griffiths 1997). The intermembrane contains the greatest proportion of residues implicated in this function (Degli Esposti et al. 1993) and consists of 105 amino acid residues in mammals (Zhang et al. 1998), with about 29% of these either completely or mostly conserved in metazoans (Degli Esposti et al. 1993).

The matrix domain, which is located entirely within the inner surface of the inner mitochondrial membrane, is composed of 65 amino acid residues and has a relatively high proportion of polar and basic amino acid residues (Griffiths 1997). There are relatively few residues in the matrix domain that are conserved or have been implicated in the proton-input function of a Q_i redox center (Degli Esposti et al. 1993). Most regions of the matrix domain have no known function (Irwin, Kocher, and Wilson 1991) or conserved properties in metazoans (Degli Esposti et al. 1993), and thus variable sites within this domain are expected to evolve in a nearly neutral manner relative to the other two functional domains.

The transmembrane domain, which consists of that portion of the cyt-*b* protein that transverses the inner mitochondrial membrane, is composed of 209 amino acid residues in mammals (Zhang et al. 1998) and is characterized by its hydrophobic properties (Irwin, Kocher, and Wilson 1991; Griffiths 1997). Furthermore, most of the observed amino acid replacements in the transmembrane region are among the hydrophobic amino acids leucine, isoleucine, and valine (Irwin, Kocher, and Wilson 1991; Kornegay et al. 1993). In addition, about 19% of the residues that compose the transmembrane domain are completely or nearly conserved in the vast majority of metazoans (Degli Esposti et al. 1993). Many of these residues have been implicated in the function of heme ligation, redox activity, or structural stability.

Summary of Cyt-*b* Data

Rates of synonymous substitution, r_{syn}, did not differ significantly between domains (table 1), with the exception being the matrix domain in cetartiodactyls. Rates of nonsynonymous substitution, r_{ns}, however, did differ significantly. Both the matrix and the transmembrane domains showed amino acid replacements significantly more frequently than the intermembrane domain in both geomyids and cetartiodactyls. The proportion of invariable amino acid sites, p_{invar}, corresponded with the pattern of rates of nonsynonymous substitution among the functional domains. The intermembrane domain exhibited a greater proportion of invariable sites than either the matrix or the transmembrane domain in both geomyids and cetartiodactyls, indicating that purifying selection is much greater in the intermembrane than in the other two domains. The number of nonsynonymous substitutions per variable site, however, did not appear to be greatly affected by selection in either group. These results suggest that variable amino acid sites in the intermembrane domain are influenced by much less selective constraint than the invariable sites, which experience an extreme form of purifying selection.

Comparisons of the relative proportions of the several amino acid residues that have undergone replacement in the cyt-*b* protein were indicative of the relative patterns of nonsynonymous substitution. For example, in pocket gophers, residues with a positive hydropathy (Kyte and Doolittle 1982) constituted about 32% of the matrix domain and 34% of the intermembrane domain, but 59% of the transmembrane domain. Furthermore, amino acid replacements at these sites constituted about 35%, 25%, and 88%, respectively, of all inferred replacements in these domains. Similar relationships between replacement sites and the magnitude of the physicochemical property were observed for composition (amino acid replacements resulting in no change in composition of the side chain: 24% in the matrix, 25% in the intermembrane, and 88% in the transmembrane), polarity (changes in polarity $\leq$ 6.2: 35%, 25%, and 86%, respectively), and polar requirement (changes in polar requirement $\leq$ 5.6: 35%, 25%, and 86%, respectively) and, to a lesser extent, for molecular volume (changes in mo-

lecular volume > 105: 35%, 38%, and 69%, respectively). The relationship between the magnitude of changes in these amino acid properties and the general rate of amino acid replacement underscores the importance of partitioning gene sequence data by functional domain when analyzing the molecular evolution of cyt-*b*.

Evolution Relative to Individual Amino Acid Properties

With regard to changes in physicochemical properties, the evolution of cyt-*b* in the two mammalian groups was similar (figs. 2 and 3). The matrix domain exhibited a good fit to expected magnitude distributions ($P > 0.05$) in relation to every property considered. The transmembrane domain, however, exhibited a poor fit ($P < 0.05$) to every expected distribution. Intermembrane domain evolution was not nearly as conservative across taxonomic groups. Pocket gophers exhibited a good fit to expected distributions for every amino acid property, whereas cetartiodactyls exhibited a good fit for polarity and isoelectric point, but not composition of the side chain, molecular volume, polar requirement, or hydropathy. The good fit of the pocket gopher data may be somewhat artificial, being the result of the extremely low rate of nonsynonymous substitution in the intermembrane domain (table 1).

Despite the good fit of the matrix domain to expected magnitude distributions, statistical comparisons between magnitude classes (conservative, moderate, radical, and very radical) revealed selection that has constrained evolution such that conservative and moderate changes predominate with regard to molecular volume in pocket gophers and with regard to all properties but hydropathy in cetartiodactyls. This trend may reflect a general evolutionary maintenance of some amino acid properties in the matrix domain. Furthermore, although the low GF scores denote an evolutionary profile that is not significantly different than neutral expectations (a conclusion that may correspond to the domain's general lack of function), overall conservation of some amino acid properties may nonetheless be important to the overall evolution of the cyt-*b* protein in mammals.

The relatively high rates of nonsynonymous substitution in the transmembrane domain (table 1) are not necessarily indicative of low levels of selection. The poor fit of the resultant amino acid replacements to neutral expectations supports this assertion, especially when compared with the GF scores calculated for the matrix domain, which exhibits similar rates of substitution. Poor GF scores with high rates of substitution in the transmembrane domain probably are due to the effect of codon composition.

Each codon possesses an inherent suite of possible amino acid changes precluding all others. As a result, some codons are inherently conservative by nature, whereas others are inherently radical. Codons that encode phenylalanine, leucine, isoleucine, methionine, and valine (codons with thymine at the second position) are examples of inherently conservative amino acids. Of the 104 possible evolutionary pathways for these 16 codons, only 12 result in nonconservative (i.e., moderate, radical, or very radical) changes with regard to composition of the side chain, 19 with regard to polarity, 16 with regard to molecular volume, 12 with regard to polar requirement, 26 with regard to hydropathy, and 5 with regard to isoelectric point. Collectively, these codons make up more than 52% of the transmembrane domain in both pocket gophers and cetartiodactyls. Phenylalanine, leucine, isoleucine, methionine, and valine are each hydrophobic and thus act to anchor the transmembrane domain into position through the membrane. The codons for these amino acids were the targets of a disproportionate number of nonsynonymous substitutions (72.5%) in the transmembrane domain, not because there are necessarily more mutations at these sites, but most probably because mutations at these sites are more likely to result in conservative changes, which in turn are more likely to become fixed. The correspondence between the magnitude of changes in amino acid properties and the general rate of amino acid replacement noted in the previous section may be the result of the genetic code playing a role that has been called ''a guardrail set near the sharp ridge of the protein landscape'' (Aita, Urata, and Husimi 2000). The parameters of the genetic code may constrain evolution in some regions, yet promote high, nearly neutral rates of evolution in others, and in both cases the effects of negative selection tend to amplify the results, in this case maintaining a rate of nonsynonymous substitution in the transmembrane domain that is comparable to the nonsynonymous rate of the matrix domain, which is affected by selection relatively very little (figs. 2 and 3).

As stated above, the pocket gopher intermembrane domain exhibited good fit to neutral expectations; the only possible exception was isoelectric point (fig. 2). Just one of the 17 amino acid replacements in this domain was ''very radical,'' whereas the rest were conservative. Although these numbers were insufficient to constitute a lack of global fit, the *z*-scores comparing magnitude classes indicated a significant difference between conservative and moderate categories and between radical and very radical categories. This may be interpreted as a state of general negative selection, with enough positive selection for the replacement glutamine (Q) → arginine (R) at amino acid position 137 to become fixed in a single taxon (*Geomys bursarius*) contrary to neutral expectations and conserved cyt-*b* primary structure (Degli Esposti et al. 1993).

The evolution of the cetartiodactyl intermembrane domain exhibited good fit for polarity and isoelectric point, although the latter appeared somewhat constrained. Composition, molecular volume, polar requirement, and hydropathy exhibited poor fits to expected distributions, indicating that there have been other influences on the evolution of these properties. Comparisons of observed magnitude classes suggested that conservative changes in molecular volume, as well as moderate changes in both composition and hydropathy, have been promoted by selection in this functional domain (fig. 3). Analysis of inferred changes in polar requirement suggested that not only has there been negative selection

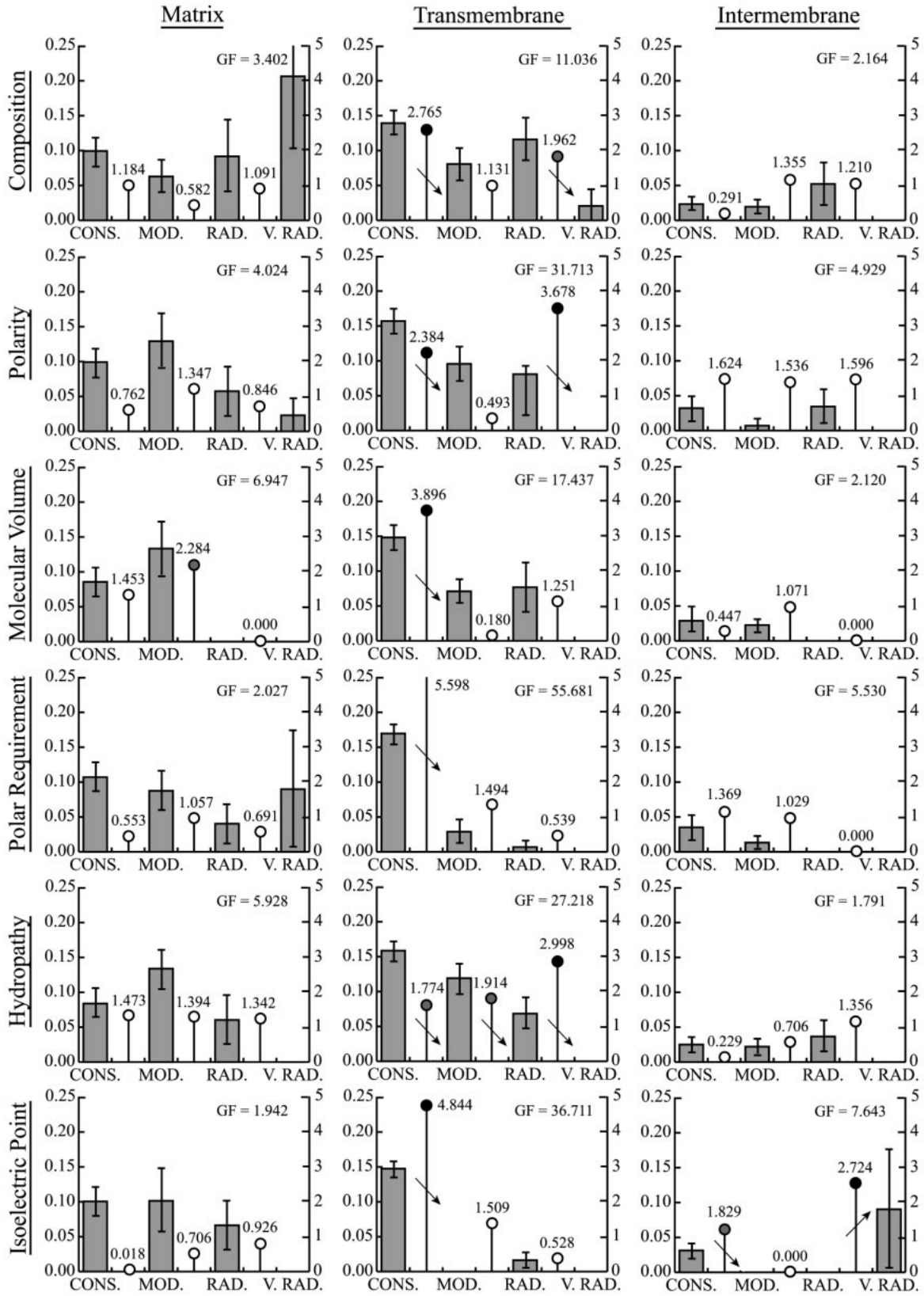

FIG. 2.—Proportions of pocket gopher amino acid replacements per evolutionary pathway for the three functional domains of the cytochrome *b* protein and six physicochemical properties shown to correlate with rates of amino acid replacement. The x-axis gives the number of amino acid replacements per magnitude class (CONS. = conservative changes; MOD. = moderate changes; RAD. = radical changes; V. RAD. = very radical changes), indicated by the bars. Associated standard errors are also indicated. The y-axis gives the value of the *z*-score, indicated by the lines and circles. Levels of significance are also indicated: white circles, $P > 0.05$; gray circles, $P < 0.05$; black circles, $P < 0.01$. Arrows indicate direction of significant difference.

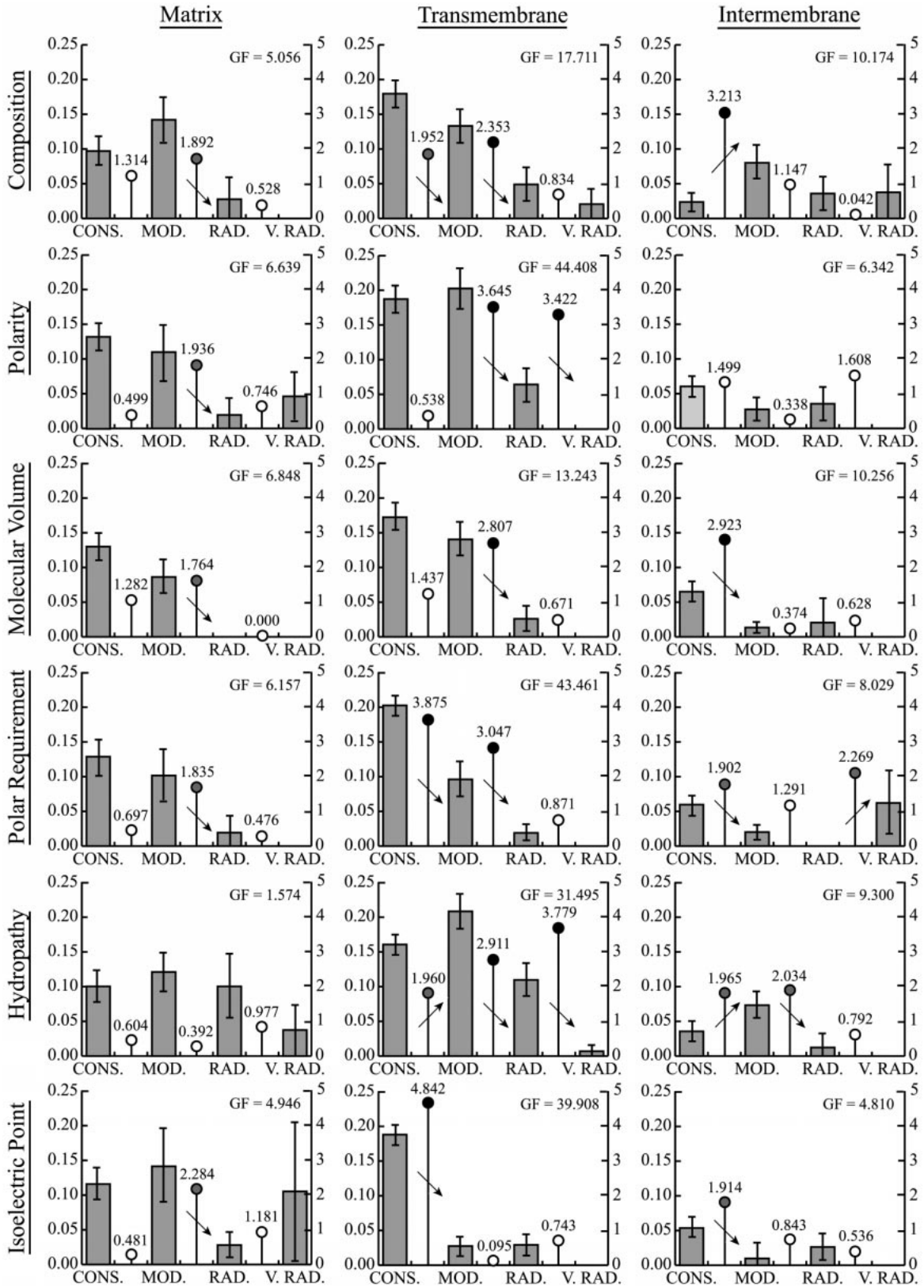

FIG. 3.—Proportions of cetartiodactyl amino acid replacements per evolutionary pathway for the three functional domains of the cytochrome *b* protein and six physicochemical properties shown to correlate with rates of amino acid replacement. The x-axis gives the number of amino acid replacements per magnitude class (CONS. = conservative changes; MOD. = moderate changes; RAD. = radical changes; V. RAD. = very radical changes), indicated by the bars. Associated standard errors are also indicated. The y-axis gives the value of the *z*-score, indicated by the lines and circles. Levels of significance are also indicated: white circles, $P > 0.05$; gray circles, $P < 0.05$; black circles, $P < 0.01$. Arrows indicate direction of significant difference.

promoting conservative changes, but there has also been enough positive selection to promote the same ''very radical'' amino acid replacement, aspartic acid (D) → threonine (T) at amino acid site 159, in two separate lineages (*Camelus bactrianus* and, most notably, all cetaceans). Although this amino acid residue has not been implicated in the *Q*-cycle mechanism (Degli Esposti et al. 1993), this change represents a large local decrease in polar requirement and may be of substantial evolutionary importance. The location of this residue is near the amino-terminal end in the second helix of the large *cd*-loop, which interacts directly with the other external surface domains of the bc_1 complex. The effect of such a replacement has yet to be determined. However, a T at this position is a common characteristic shared by many vertebrates, whereas D is shared by relatively few (Degli Esposti et al. 1993).

Conclusions

Our results support previous hypotheses for different evolutionary rates of substitution in the cyt-*b* functional domains and propose mechanisms for their maintenance. The good fit of amino acid replacements to neutral expectations in the cyt-*b* matrix domain reflects this region's lack of function. Although the transmembrane domain shows a nonsynonymous substitution rate comparable to the matrix domain, transmembrane amino acid replacements do not fit neutral expectations for every amino acid property considered. Our analyses suggest that this phenomenon results from interaction between codon composition and purifying selection. As a result, purifying selection serves to ''amplify'' transmembrane amino acid replacement rates because the most plentiful amino acids are predisposed to conservative change. Not surprisingly, the intermembrane domain, which has the greatest number of amino acids implicated in the *Q*-cycle mechanism of the electron transport chain, evolves most slowly. Amino acid replacements in the pocket gopher intermembrane domain fit neutral expectations very well, whereas the cetartiodactyl intermembrane domain does not fit the model for four of the six amino acid properties. Positive selection resulting in ''very radical'' amino acid replacements was detected in both groups.

Acknowledgments

Nora J. Besansky, David W. Foltz, Mark S. Hafner, James V. Remsen, Frederick H. Sheldon, and two anonymous reviewers made helpful comments on the manuscript. Special thanks go to Norihiro Okada and Masami Hasegawa for office space and resources. This research was supported by grants from the Louisiana State University Museum of Natural Science, the American Museum of Natural History Theodore Roosevelt Fund, the American Society of Mammalogists, and the Japan Society for the Promotion of Science Monbusho Fund.

LITERATURE CITED

AITA, T., S. URATA, and Y. HUSIMI. 2000. From amino acid landscape to protein landscape: analysis of genetic codes in terms of fitness landscape. J. Mol. Evol. **50**:313–323.

ALFF-STEINBERGER, C. 1969. The genetic code and error transmission. Proc. Natl. Acad. Sci. USA **64**:584–591.

DEGLI ESPOSTI, M., S. DE VRIES, M. CRIMI, A. GHELLI, T. PATERNELLO, and A. MEYER. 1993. Mitochondrial cytochrome b: evolution and structure of the protein. Biochem. Biophys. Acta **1143**:243–271.

DEWALT, T. S., P. D. SUDMAN, M. S. HAFNER, and S. K. DAVIS. 1993. Phylogenetic relationships of pocket gophers (Cratogeomys and Pappogeomys) based on mitochondrial DNA cytochrome *b* sequences. Mol. Phylogenet. Evol. **2**: 193–204.

EPSTEIN, C. J. 1966. Role of the amino-acid 'code' and of selection for conformation in the evolution of proteins. Nature **210**:25–28.

GOLDBERG, A. L., and R. E. WITTES. 1966. Genetic code: aspects of organization. Science **153**:420–424.

GRANTHAM, R. 1974. Amino acid difference formula to help explain protein evolution. Science **185**:862–864.

GRAYBEAL, A. 1993. The phylogenetic utility of cytochrome *b*: lessons from bufonid frogs. Mol. Phylogenet. Evol. **2**: 256–269.

GRIFFITHS, C. S. 1997. Correlation of functional domains and rates of nucleotide substitution in cytochrome *b*. Mol. Phylogenet. Evol. **7**:352–365.

HAFNER, M. S., P. D. SUDMAN, F. X. VILLABLANCA, T. A. SPRADLING, J. W. DEMASTES, and S. A. NADLER. 1994. Disparate rates of molecular evolution in cospeciating hosts and parasites. Science **265**:1087–1090.

HAIG, D., and L. D. HURST. 1991. A quantitative measure of error minimization in the genetic code. J. Mol. Evol. **33**: 412–417.

HUGHES, A. L., T. OTA, and M. NEI. 1990. Positive Darwinian selection promotes charge profile diversity in the antigen-binding cleft of class I major-histocompatibility-complex molecules. Mol. Biol. Evol. **7**:515–524.

IRWIN, D. M., T. D. KOCHER, and A. C. WILSON. 1991. Evolution of the cytochrome *b* gene of mammals. J. Mol. Evol. **32**:128–144.

KORNEGAY, J. R., T. D. KOCHER, L. A. WILLIAMS, and A. C. WILSON. 1993. Pathways of lysozyme evolution inferred from the sequences of cytochrome *b* in birds. J. Mol. Evol. **37**:367–379.

KYTE, J., and R. F. DOOLITTLE. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.

LI, W.-H. 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J. Mol. Evol. **36**: 96–99.

MCCLELLAN, D. A. 2000. The codon-degeneracy model of molecular evolution. J. Mol. Evol. **50**:131–140.

MCCRACKEN, K. G., J. HARSHMAN, D. A. MCCLELLAN, and A. D. AFTON. 1999. Data set incongruence and correlated character evolution: an example of functional convergence in the hind-limbs of stifftail diving ducks. Syst. Biol. **48**:683–714.

NEI, M., and T. GOJOBORI. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. **3**:418–426.

NIKAIDO, M., A. P. ROONEY, and N. OKADA. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: hippopotamuses are the closest extant relatives of whales. Proc. Natl. Acad. Sci. USA **96**:10261–10266.

SHEDLOCK, A. M., M. C. MILINKOVITCH, and N. OKADA. 2000. SINE evolution, missing data, and the origin of whales. Syst. Biol. **49**:1–10.

SHEDLOCK, A. M., and N. OKADA. 2000. SINE insertions: powerful tools for molecular systematics. BioEssays **22**:148–160.

SNEATH, P. H. A. 1966. Relations between chemical structure and biological activity. J. Theor. Biol. **12**:157–195.

SONNEBORN, T. M. 1965. Degeneracy of the genetic code: extent, nature, and genetic implications. Pp. 377–397 *in* V. BRYSON and H. J. VOGEL, eds. Evolving genes and proteins. Academic Press, New York.

SWOFFORD, D. L. 1998. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer, Sunderland, Mass.

TAYLOR, W. R. 1986. The classification of amino acid conservation. J. Theor. Biol. **119**:205–218.

WOESE, C. R., D. H. DUGRE, S. A. DUGRE, M. KONDO, and W. C. SAXINGER. 1966. On the fundamental nature and evolution of the genetic code. Cold Spring Harb. Symp. Quant. Biol. **31**:723–736.

WYCKOFF, G. J., W. WANG, and C.-I. WU. 2000. Rapid evolution of male reproductive genes in the descent of man. Nature **403**:304–308.

XIA, X. 1998. The rate heterogeneity of nonsynonymous substitutions in mammalian mitochondrial genes. Mol. Biol. Evol. **15**:336–344.

———. 2000. Phylogenetic relationships among horseshoe crab species: effect of substitution models on phylogenetic analysis. Syst. Biol. **49**:87–100.

XIA, X., M. S. HAFNER, and P. D. SUDMAN. 1996. On transition bias in mitochondrial genes of pocket gophers. J. Mol. Evol. **43**:32–40.

XIA, X., and W.-H. LI. 1998. What amino acid properties affect protein evolution? J. Mol. Evol. **47**:557–564.

ZHANG, Z., L. HUANG, V. M. SHULMEISTER, Y.-I. CHI, K. K. KIM, L.-W. HUNG, A. R. CROFTS, E. A. BERRY, and S.-H. KIM. 1998. Electron transfer by domain movement in cytochrome bc_1. Nature **392**:677–684.

DAVID IRWIN, reviewing editor

Accepted February 5, 2001