

## A TEST OF THE GENETIC SPECIES CONCEPT: CYTOCHROME-*b* SEQUENCES AND MAMMALS

ROBERT D. BRADLEY\* AND ROBERT J. BAKER

*Department of Biological Sciences and Museum of Texas Tech University, Lubbock, TX 79409*

Levels of sequence variation in mitochondrial cytochrome-*b* gene were examined to ascertain if this molecule can provide a reference point in making decisions concerning species-level distinctions. DNA-sequence data from 4 genera of rodents (*Neotoma*, *Reithrodontomys*, *Peromyscus*, and *Sigmodon*) and 7 genera of bats (*Artibeus*, *Carollia*, *Chiroderma*, *Dermanura*, *Glossophaga*, *Rhinophylla*, and *Uroderma*), including recognized sister species, were examined to develop hypotheses for evaluating levels of sequence variation. Several patterns associated with DNA-sequence variation emerged from this study. Specifically, genetic distance values <2% were indicative of intraspecific variation; values between 2 and 11% had a high probability of being indicative of conspecific populations or valid species and merit additional study concerning specific status; and values >11% were indicative of specific recognition. It appears that genetic distance values may be useful for determination of species boundaries under the framework of the Genetic Species Concept.

Key words: cytochrome-*b* sequences, Genetic Species Concept, mammals, molecular systematics

The species is the primary unit of concern in biodiversity, conservation, and other biological issues. Recent estimates of the number of species range from 10 to 100 million (Wilson 1998). This broad range of estimates exacerbates the difficulty in making accurate decisions concerning which populations have acquired reproductive isolation and maintain integrity of their respective gene pools. Such decisions are rarely accompanied with all sufficient biological information concerning  $\geq 2$  putative species (i.e., sympatric distributions, overlapping breeding seasons, fixed levels of genetic differentiation, complete fossil record, geological events and history, and time since divergence). Rather, in the process of determining the validity of putative species, systematists generally rely on indirect information in the form of some character system(s) (e.g., variation in size and shape of morphologic characters, cytogenetics, allozymes, DNA sequences). In re-

ality, or at least in practice, the end product is a working definition of the species concept based on that particular set of characters.

To date, biologists have developed  $\geq 22$  species concepts (Mayden 1997) to describe and categorize biological diversity. This voluminous literature is accompanied by an ever-growing literature that discusses and tests merits and flaws of each concept. Although there usually is considerable debate among systematists as to the exact boundaries and utility of these working “species concepts” for various types of characters, most lists of species (classifications) are products of the application of combinations of these working “species concepts.”

One of the lesser-known concepts, the Genetic Species Concept (sensu Dobzhansky 1950; Mayr 1969; Simpson 1943), is a measurement of genetic differences used to infer reproductive isolation and evolutionary independence (Mayden 1997). One criticism of the Genetic Species Concept is in

\* Correspondent: rbradley@ttacs.ttu.edu

determining the magnitude of genetic variation required to distinguish between 2 putative species. The methodology outlined in our study addresses this criticism and permits implementation of the Genetic Species Concept. Although this methodology, and concomitantly the Genetic Species Concept, cannot be universally applied without knowledge of levels of genetic divergence from closely related species and from sister species, it can serve as a useful tool for taxonomists and systematists.

Over the last 10 years, there has been a large increase in the quantity of DNA sequences being generated for mammalian taxa. The plethora of DNA sequence-data now available provides an opportunity to critically test the utility of the Genetic Species Concept. This increase in DNA-sequence data has been accompanied by observations that this variation possibly is partitioned in a predictable pattern across closely related species (Avice and Walker 1999; Johns and Avice 1998). This predictable pattern is based on the magnitude of genetic differentiation and may or may not be independent of time since divergence. What is critical is that as with morphology, allozymes, karyology, etc., a certain magnitude of divergence in sequence identity usually is associated with the completion of speciation.

We examined the magnitude of sequence variation in the mitochondrial cytochrome-*b* gene to ascertain its utility in providing an additional reference point (to those described above) in making decisions concerning species-level distinction. The significance of this gene to biodiversity issues is not trivial if the conclusions of Avice and Walker (1999) and Johns and Avice (1998) are correct. Those authors concluded that the cytochrome-*b* gene showed a high level of congruence with species limits based on classical taxonomic studies. They further concluded that there may be approximately twice as many species of vertebrates as previously described under the Biological Species Concept of Mayr (1963). Exploration

of mammalian biodiversity by comparing the Biological Species Concept of Mayr (1963), Phylogenetic Species Concept (Wiley 1981), mtDNA phylogroups (Avice and Walker 1999), and other species concepts such as the Genetic Species Concept should provide valuable information for understanding mammalian species of the world (Wilson and Reeder 1993).

Our goal was to test the hypothesis whether levels of sequence divergences can be used to identify species-level differentiation under the framework of the Genetic Species Concept. We examined DNA-sequence variation in the cytochrome-*b* gene from 4 genera of rodents (*Neotoma*, *Reithrodontomys*, *Peromyscus*, and *Sigmodon*) and 7 genera of bats (*Artibeus*, *Carollia*, *Chiroderma*, *Dermanura*, *Glossophaga*, *Rhinophylla*, and *Uroderma*). Our intention was to use DNA-sequence variation from recognized sister species of mammals to develop a working hypothesis for cytochrome-*b* data. Ultimately, an understanding of the level and magnitude of variation that accompanies well-documented end points of speciation and values that are typical of conspecific populations can be used to identify situations where reevaluation using other data (morphology, allozymes, chromosomes, etc.) is merited.

#### MATERIALS AND METHODS

To properly test the Genetic Species Concept, it is paramount that 5 criteria are met for the hypothesis to be rigorous and reliable. First, phylogenetic principles and methods must be used to independently test results of genetic distance analyses. Second, a stable and constantly evolving protein coding gene (e.g., cytochrome *b*) would be preferred to avoid affects of gaps, unequal sequence evolution, and rate heterogeneity. However, other DNA markers (e.g., mitochondrial 12S gene—Hooper and Van Den Bussche 2001; Jolley et al. 2000) may prove valuable. The cytochrome-*b* gene, which has its greatest application in understanding relationships of congeners (although its value at the generic rank and above is well demonstrated; Smith and Patton 1999), has been so intensely

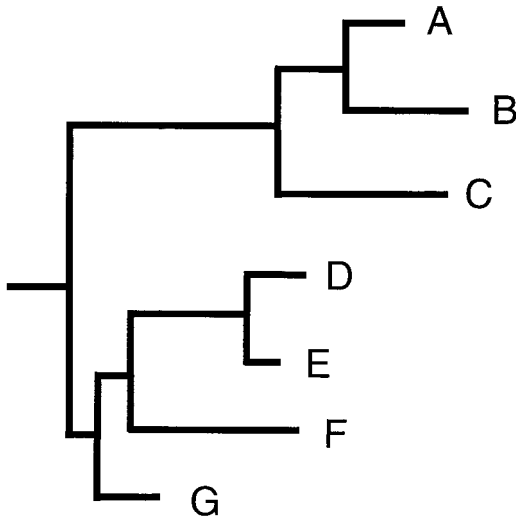


FIG. 1.—Hypothetical tree depicting relationships of taxa A–G. Two pairs of sister taxa are depicted (A and B, D and E); all other comparisons require multiple speciation events and do not reflect sister taxon relationships.

used in mammalian systematics that it has become the system of choice by many systematists. Third, known species must be present within the dataset to act as “bench marks” for evaluating different levels of sequence divergence. Fourth, the dataset must involve only endpoints that include a single speciation event in the gene tree (Fig. 1). In this gene tree, there are 2 examples (A and B, D and E) where data fit our criteria. The other comparisons involve multiple speciation events and are not appropriate for this analysis. Fifth, a dataset containing as many of the available sister taxa as possible would strengthen the analysis. If this 5th criterion is not met, then DNA-sequence divergences for some taxa will be overestimated.

DNA sequences from the complete mitochondrial cytochrome-*b* gene (1,140 bp in bats and 1,143 bp in rodents) were generated in previous or ongoing studies (Baker et al. 1994; Bell et al. 2001; Bradley et al. 2000; Carroll et al., in press; Edwards and Bradley 2001; Edwards et al. 2001; F. G. Hoffmann and R. J. Baker, in litt.; Hoffmann and Baker 2001; Peppers and Bradley 2000; Peppers et al., in press; Tiemann-Boege et al. 2000; Van Den Bussche et al. 1998; Wright et al. 1999) and were deposited with GenBank. References for accession numbers and studies from which they were generated are available

from the authors. These genera were selected for the following reasons. First, they were groups for which we have published phylogenies that have passed the critique and examination of other expert systematists familiar with each genus. Second, we have expertise and familiarity with the phylogenetic relationships of these taxa. Third, independent datasets (i.e., chromosomes, allozymes, or DNA markers) corroborated the phylogenies we generated initially. Fourth, those genera represented a cross-section of mammalian taxonomy and diversity (Chiroptera and Rodentia).

In the initial studies, nucleotides were treated as unordered discrete characters (A, C, G, and T), and a variety of data analyses (parsimony, genetic distance, and maximum likelihood) were used using PAUP\* (Swofford 2000). Nodal support was examined using bootstrap (Felsenstein 1985), and Bremer support values (Bremer 1994) were calculated using PAUP\* and auto-decay programs (Eriksson 1997). Phylogenetic relationships and subsequent taxonomic or classification decisions were based on congruencies among the various analyses.

In this study, for purposes of continuity and comparison across several taxonomic ranks, genetic distances were calculated using the Kimura 2-parameter model of nucleotide substitution (Kimura 1980) assuming minimum evolution. Cytochrome-*b* distance values were calculated at 5 taxonomic levels: intrapopulation, intrasubspecific, intraspecific, intrageneric, and sister species. For calculations involving higher taxonomic levels, pairwise comparisons of the previous level (lower) were removed. For example, in comparisons of intraspecific samples, pairwise distances of samples representing the same subspecies were not included; only pairwise distances of those samples to other subspecies were used. Average genetic distances were calculated in cases where >1 set of pairwise comparisons was available for a taxonomic level.

## RESULTS

Genetic distances were obtained using DNA sequences generated in previous systematic studies (listed in “Materials and Methods”) and the Kimura 2-parameter model of evolution (Kimura 1980). Collecting records and taxonomic histories were used to identify appropriate compari-

TABLE 1.—Average sequence-divergence values for intrapopulational, intrasubspecific, intraspecific, intrageneric, and sister-species comparisons. Values were calculated using a Kimura 2-parameter model of evolution and are given as percentages. Ranges were calculated using the lowest and highest mean value within each set of comparisons.

Comparison	Chiroptera			Rodentia			All		
	$\bar{X}$	Range	<i>n</i>	$\bar{X}$	Range	<i>n</i>	$\bar{X}$	Range	<i>n</i>
Intrapopulational	1.28	0.00–3.83	24	0.21	0.00–0.53	4	1.13	0.00–3.83	28
Intrasubspecific	1.23	0.00–4.21	11	0.85	0.00–1.87	11	1.04	0.00–4.21	22
Intraspecific	3.00	0.09–8.70	14	2.09	0.00–6.29	18	2.49	0.00–8.70	32
Sister species	6.83	2.50–16.42	12	9.55	2.70–19.23	11	8.13	2.50–19.23	23
Intrageneric	10.45	2.51–19.83	7	13.23	2.23–21.97	4	11.46	2.23–21.97	11

sons among samples to examine intrapopulational, intrasubspecific, intraspecific, and intrageneric DNA-sequence variation. Topologies generated from parsimony, genetic distance, and likelihood methods obtained from the original studies were used to identify appropriate comparisons for the estimation of DNA-sequence variation between sister taxa.

#### *Variation within a Species*

*Intrapopulational variation.*—Genetic variation at the intrapopulational level was assessed by examining samples collected from the same geographic locality. Twenty-four populations of bats (11 species) and 4 populations of rodents (4 species) were used to evaluate DNA-sequence variation within populations. The average genetic distance for all populations ( $n = 28$ ) was 1.13 (range, 0–3.83; Table 1, Fig. 2).

*Intrasubspecific variation.*—Samples representative of a single subspecies (as mapped in Hall 1981) were examined to estimate the magnitude of genetic variation within subspecific categories. If multiple samples (of a subspecies) were from the same population, the pairwise genetic distances for within the population were removed from the calculations. Eleven subspecies of bats and 11 subspecies of rodents were used to evaluate DNA sequence variation within a subspecies. The average genetic distance for all subspecies ( $n = 22$ ) was 1.04 (range, 0–4.21; Table 1, Fig. 3).

*Intraspecific variation.*—This category

examined genetic variation within a single species. Species boundaries were determined from the results of more detailed studies identified in the “Materials and Methods.” This category included 14 species of bats and 18 species of rodents. The average genetic distance for all species ( $n = 32$ ) was 2.49 (range, 0–8.70; Table 1, Fig. 4).

#### *Variation within a Genus*

*Sister-species variation.*—Twenty-three pairs of sister species for bats ( $n = 12$ ) and rodents ( $n = 11$ ) were used to evaluate DNA-sequence variation between sister species. Sister species were determined from the more detailed studies listed in the “Materials and Methods.” The average genetic distance for all sister species comparisons ( $n = 23$ ) was 8.13 (range, 2.50–19.23; Table 1, Fig. 5).

*Intrageneric variation.*—This category examined genetic variation within a genus and included 7 genera of bats and 4 genera of rodents. The average genetic distance for all genera ( $n = 11$ ) was 11.46 (range, 2.23–21.97; Table 1, Fig. 6).

#### DISCUSSION

Our goal was to assess levels of DNA-sequence divergence under the framework of the Genetic Species Concept. Specifically, how predictive are sequence-divergence values, obtained from the cytochrome-*b* gene, in identifying species-level differentiation in mammals? Determination of how

**Intrapopulation Variation**

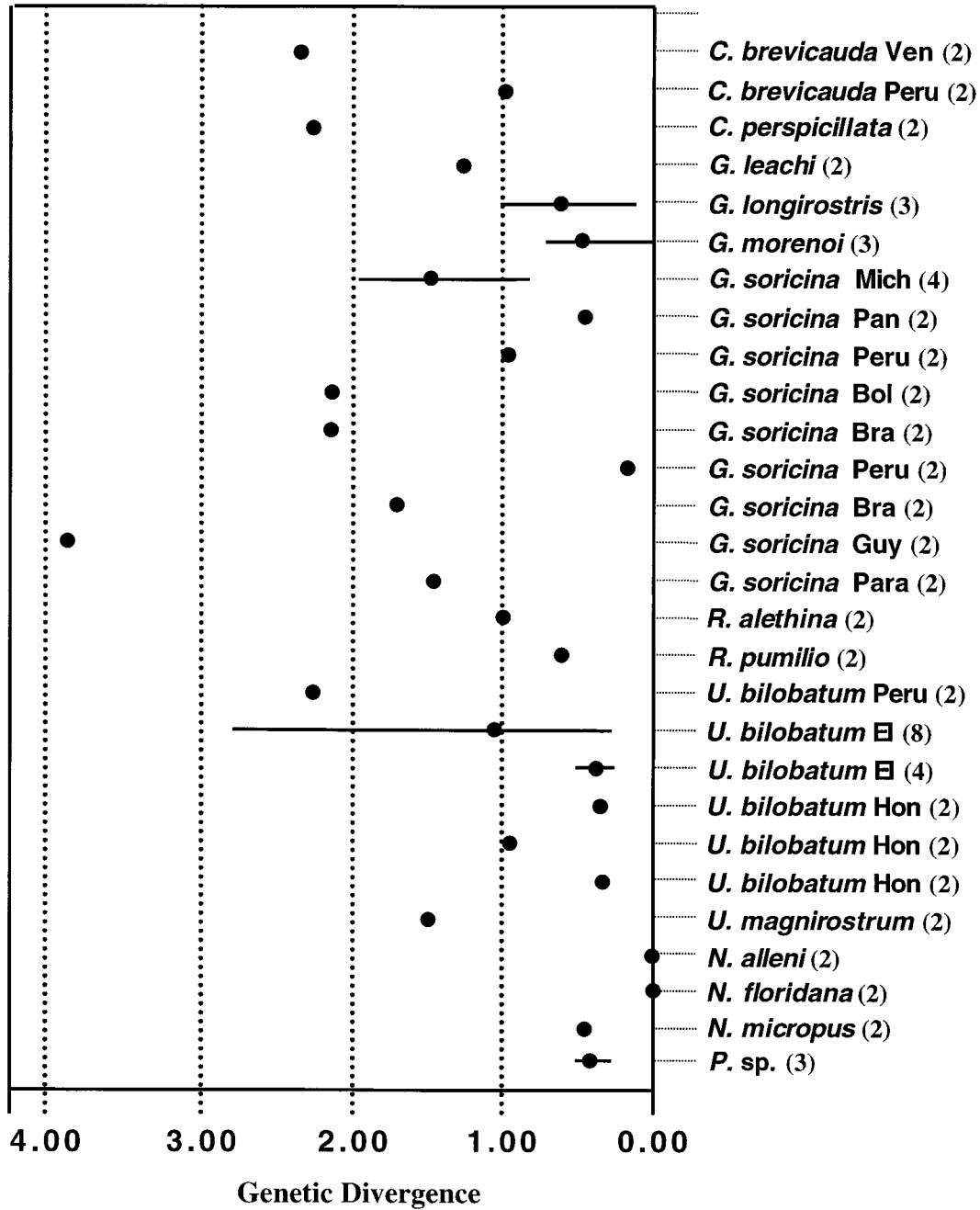


FIG. 2.—Intrapopulation comparisons of Kimura 2-parameter genetic distances for 24 populations of bats and 4 populations of rodents. Average genetic distances are shown as filled circles, bars represent the range of values obtained within a population, and sample sizes are shown in parentheses.

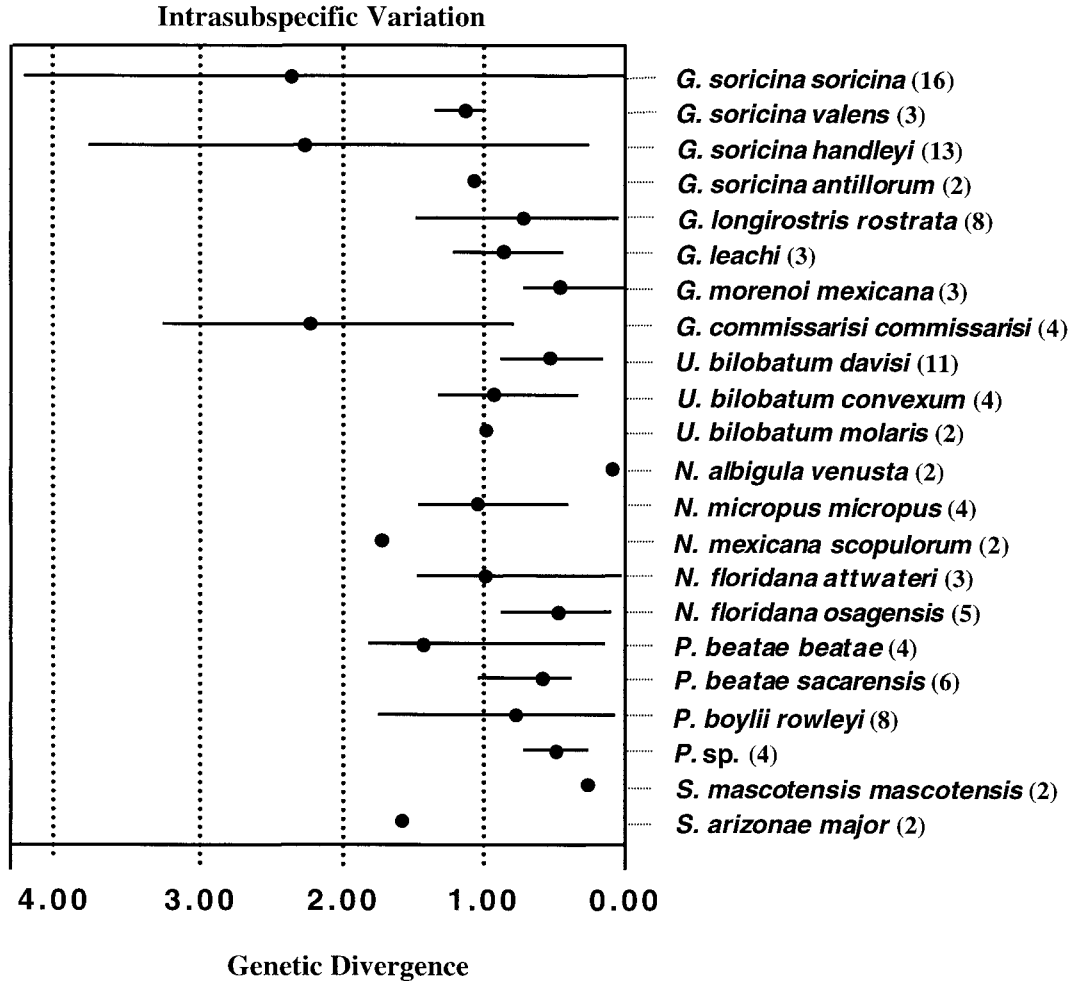


FIG. 3.—Intrasubspecific comparisons of Kimura 2-parameter genetic distances for 11 subspecies of bats and 11 subspecies of rodents. Average genetic distances are shown as filled circles, bars represent the range of values obtained within a subspecies, and sample sizes are shown in parentheses. Genetic variation produced from comparisons of individuals from the same population has been removed.

much genetic variation is present in typical populations and subspecific groups and between conspecific taxa is critical to addressing this question. An overview of these patterns of variation obtained in our study is presented in Fig. 7.

#### *Variation within a Species*

*Intrapopulational variation.*—Genetic distance values ranged from 0.00 to 3.83% ( $\bar{X} = 1.13$ ) with bats ( $\bar{X} = 1.28$ ) possessing about 6 times the amount of genetic diver-

gence as rodents ( $\bar{X} = 0.21$ ). It is not clear whether this is a result of sample size (bats,  $n = 24$ ; rodents,  $n = 4$ ), the volant nature of bats, or other biological differences between bats and rodents. For example, emigration and immigration among bat populations could increase intrapopulational variation. Our data indicate that the highest level of genetic divergence present in any of the populations studied herein is <4% with most populations possessing divergence values of about 1% (Fig. 7).

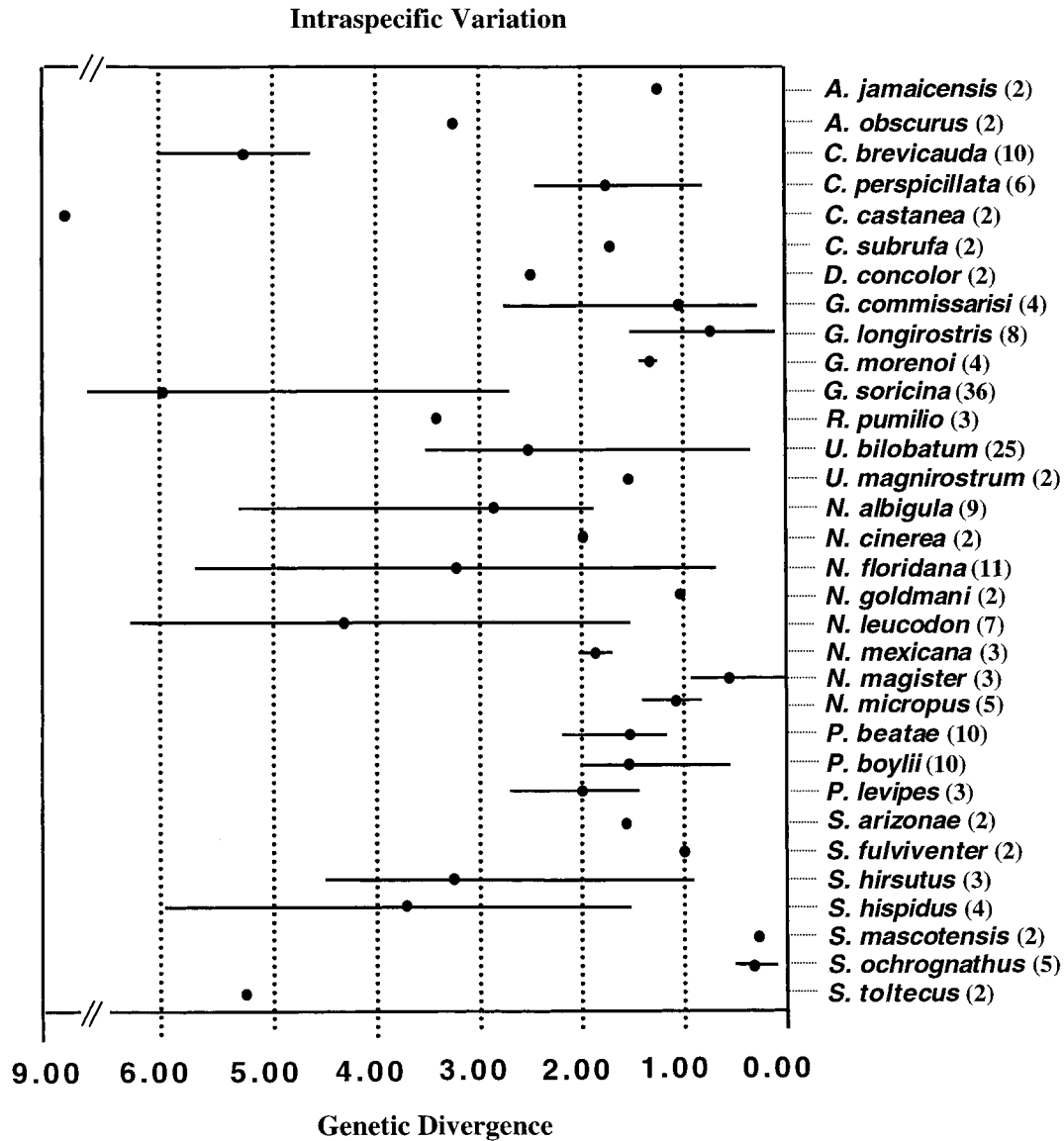


FIG. 4.—Intraspecific comparisons of Kimura 2-parameter genetic distances for 14 subspecies of bats and 18 subspecies of rodents. Average genetic distances are shown as filled circles, bars represent the range of values obtained within a species, and sample sizes are shown in parentheses. Genetic variation produced from comparisons of individuals from the same population or subspecies has been removed.

*Intrasubspecific variation.*—For species that contain  $\geq 2$  recognized subspecies, average genetic distance values within subspecific taxa were 0.09–2.34% with an average of 1.04%. Bats ( $\bar{X} = 1.23$ ,  $n = 11$ ) and rodents ( $\bar{X} = 0.85$ ,  $n = 11$ ) possessed

similar values, which were similar to those obtained in the intrapopulation comparisons and may suggest that an underestimation of this comparison was a result of insufficient geographic coverage in our study. Based on the data obtained herein

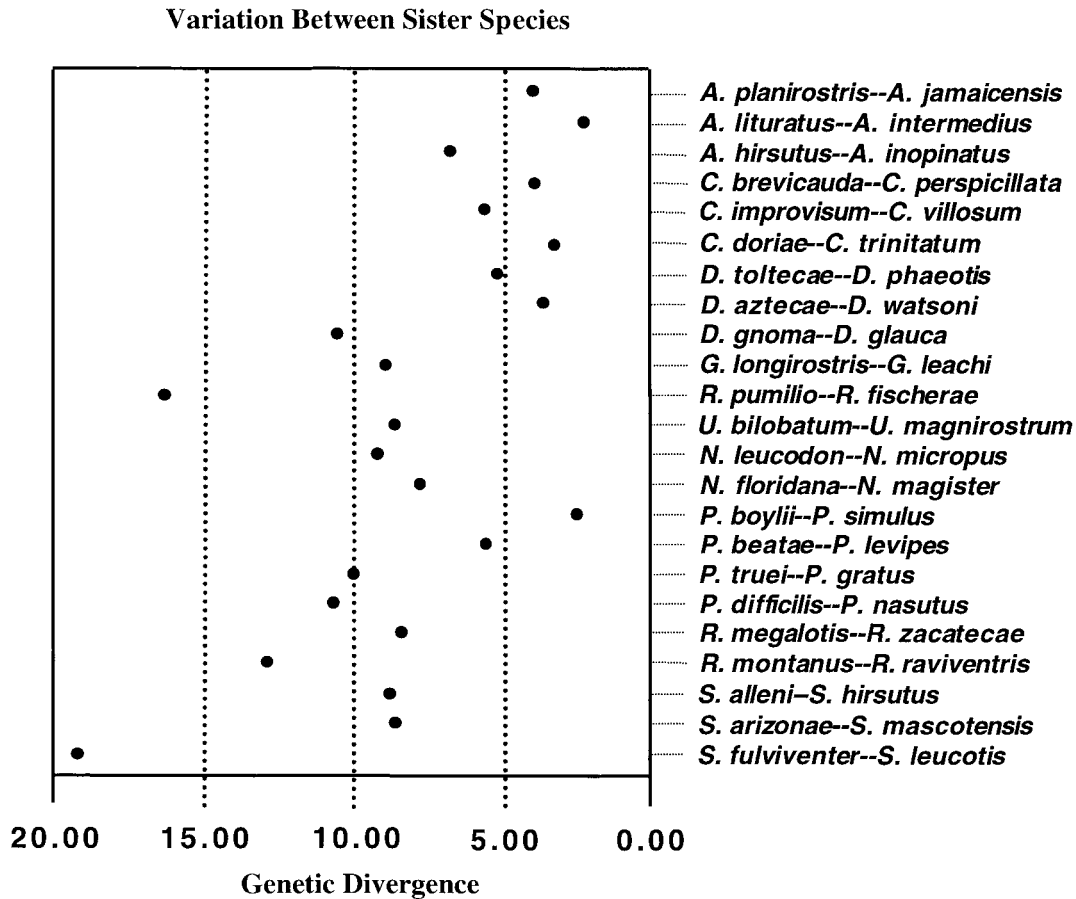


FIG. 5.—Sister taxa comparisons of Kimura 2-parameter genetic distances for 10 pairs of sister species of bats and 11 pairs of sister species of rodents. Average genetic distances are shown as filled circles. Genetic variation produced from comparisons of individuals from the same population, subspecies, or species has been removed.

(Fig. 7), subspecies as recognized in the taxa examined herein, have values  $>2.34\%$  ( $\bar{X} = 1.0\%$ ).

*Intraspecific variation.*—Average genetic distance values were 0.09–8.70% with an average of 2.49% (Fig. 7). Bats ( $\bar{X} = 3.00$ ,  $n = 14$ ) possessed slightly higher values than rodents ( $\bar{X} = 2.09$ ,  $n = 18$ ). Among the intraspecific comparisons, *C. castanea* had the highest value (8.70) whereas all other taxa had values  $<5.91$ . Only 2 samples of *C. castanea* were examined, and these values exceeded most of the sister species comparisons. It may be that these 2 samples represent 2 species as defined by

the Genetic Species Concept. This is an example of where additional studies are needed to determine if 2 biological species exist.

#### *Variation within a Genus*

*Variation between sister species.*—If the cytochrome-*b* gene contains information that can be used to identify the completion of biological speciation, it will be resolved by understanding the magnitude and pattern of sequence divergence between sister species. Several factors affect values that distinguish sister species: first, the ideal value which is the magnitude of variation that distinguishes the new species from each other

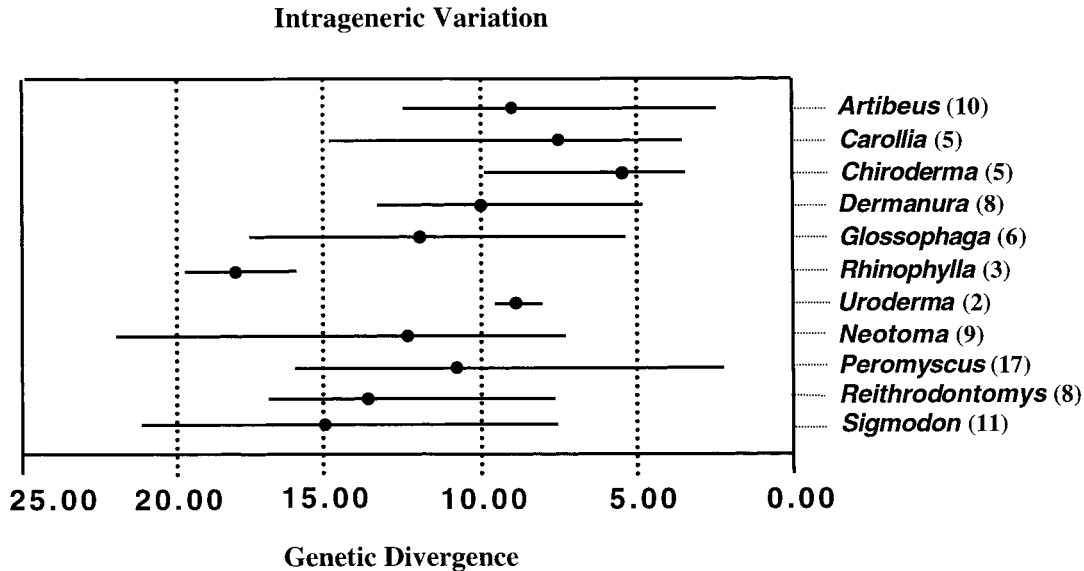


FIG. 6.—Intragenetic comparisons of Kimura 2-parameter genetic distances for 6 genera of bats and 4 genera of rodents. Average genetic distances are shown as filled circles, bars represent the range of values obtained within a genus, and sample sizes are shown in parentheses. Genetic variation produced from comparisons of individuals from the same population, subspecies, or species has been removed.

at the time of speciation (ideal value); second, the amount of within species variation within each of the 2 new species; and third, the amount of sequence divergence that has accumulated between the 2 species since the completion of the speciation process. The last value, of course, is affected by time since divergence and rate of phyletic change within each taxon. Additionally the very nature of phylograms and decisions involving the accuracy of identifying sister species makes it difficult to generate a large sample of values of sister-species comparisons. For example, Johns and Avise (1998) intentionally midpoint rooted their phylogeny because they did not have access to all critical taxa for appropriate sister-species comparisons and because they did not believe that they were qualified to address the pertinent outgroup taxa. This approach artificially inflates intragenetic values beyond those that were present between sister species at the conclusion of the speciation process.

In our study, average genetic distance values were 2.50–19.23% with an average of 8.13%. Rodents ( $\bar{X} = 9.55$ ,  $n = 11$ ) typically had higher values than did bats ( $\bar{X} = 6.83$ ,  $n = 12$ ). Obviously, this value is affected by time since divergence (recent ancestry producing lower values), and therefore it is difficult to develop a simple predictive range. However, most sister taxa differed by 4–11% (Fig. 7).

*Intragenetic variation.*—These values were calculated by including all possible pairwise comparisons of species within a genus. Average genetic distance values were 2.23–21.97% with an average of 11.46%. Values for bats ( $\bar{X} = 10.45$ ,  $n = 7$ ) and rodents ( $\bar{X} = 13.23$ ,  $n = 4$ ) depicted that rodent species typically had higher values between species than did bats; although 2 species of bats (*R. pumilio* and *R. alethina*) possessed one of the highest values (19.83%) of all taxa examined. It appears that members of the same genus typically

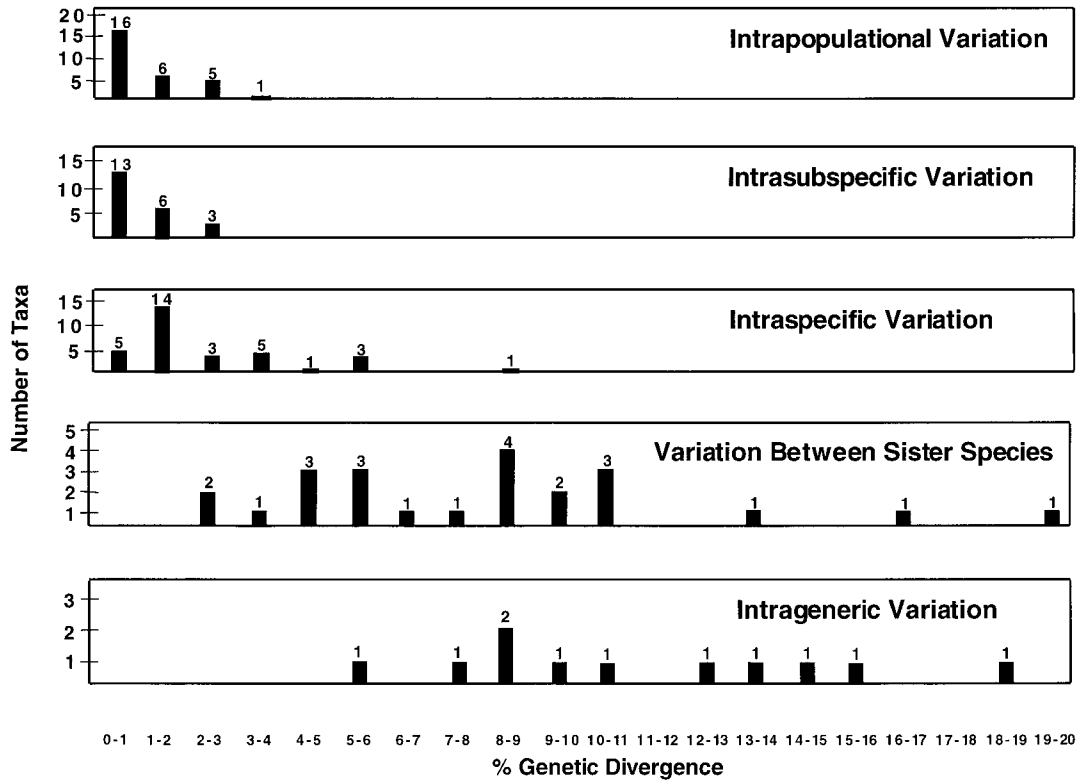


FIG. 7.—Distribution of average genetic distances (Kimura 2-parameter) obtained from comparisons of samples from the same population (intrapopulational), samples representing the same subspecies (intrasubspecific), samples representing the same species (intraspecific), samples representing sister species (sister species), and samples representing the same genus (intrageneric). Bars represent the number of taxa belonging to the same category of genetic divergence. Categories were arbitrarily defined in 1% intervals.

possess values (Fig. 7) of about 2–22% ( $\bar{X}$  = 12%).

*Application of genetic distances.*—We examined several independent datasets to evaluate the utility of genetic distance values derived from DNA sequences as a method of identifying species or species boundaries. This methodology stems from the observation that taxa identified as conspecific in phylogenetic analyses often possess levels of sequence divergences that equal or exceed values of other pairwise comparisons of valid species within the clade of interest.

Obviously, many factors can affect the results of our study. Some of these are biological and some are methodological; the

first 2 are methodological and the final 3 are biological. First, sampling strategy or lack of appropriate sampling may influence average sequence divergence at any level of comparison (intrapopulational, intrasubspecific, intraspecific, or intrageneric). For example, sampling of family units or closely related individuals would greatly underestimate the average value for intrapopulational comparisons. Alternatively, a few samples from isolated populations or populations representing geographic extremes of a species' distribution could result in the overestimation of average genetic divergences for within species comparisons. However, unique populations or populations of concern would be identified. Sec-

ond, the concept of the genus varies extensively in the groups of organisms we studied. For example, *Peromyscus* may contain as many as 64 or as few as 53 species depending on the interpretation (Carleton 1989). Relative to our goal, intrageneric distance values are shown only as a reference point relative to values for sister species. How the genus is defined will affect the variation but not the conclusions relative to sister taxa. Third, breeding structure and reproductive life history may serve to increase or decrease average genetic divergences. For example, migratory taxa that breed in one region and give birth in another would potentially possess greater genetic divergence values than philopatric taxa. Fourth, taxa with rigid social and breeding structures would be expected to show lower levels of genetic variation at the intrapopulational but higher levels for interpopulational comparisons. Fifth, parameters such as lineage sorting and retention of ancestral polymorphisms (Avice 1994; Edwards and Beerli 2000) would be expected to influence genetic divergences depending on the level of comparison. For example, levels of sequence divergences may reflect taxonomic divergence, or alternatively, they simply may track lineage sorting or retention of ancestral polymorphisms (sequences in this case).

An example of the utility of this methodology and logic is demonstrated in the 2 studies involving molecular and phylogenetic divergence among species of cotton rats in the genus *Sigmodon* (Peppers and Bradley 2000; Peppers et al., in press), which are summarized in Fig. 8. Early in these studies, we obtained samples from several populations of *S. hispidus* in North, Central, and South America. Those samples included 3 subspecies (*S. h. hirsutus*, *S. h. hispidus*, and *S. h. toltecus*) and were intended to represent genetic variants within the nominal taxon. Among-sample comparisons of those subspecies revealed exceptionally high levels of sequence divergence relative to comparisons to other species in

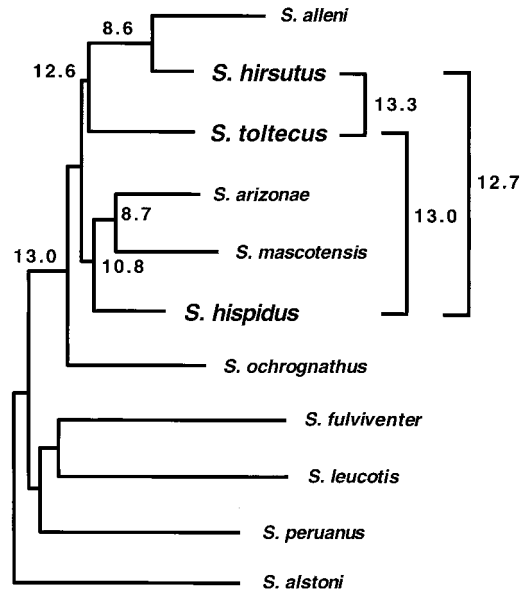


FIG. 8.—Neighbor-joining tree constructed using Kimura 2-parameter genetic distances for members of the rodent genus *Sigmodon*. Until recently (Peppers and Bradley 2000), taxa in large bold type were considered conspecific members of *S. hispidus*. Average genetic divergence values are provided for select taxa. The high levels of sequence divergence correspond to the paraphyletic nature of these 3 samples.

the genus (*S. alleni*, *S. arizonae*, *S. mascotensis*, and *S. ochrognathus*). Additionally, all preliminary analyses (parsimony, genetic distance, and likelihood) portrayed those samples as paraphyletic. As more samples were included in the study, it became obvious that *S. hispidus* was a composite species and that minimally *S. h. hirsutus* and *S. h. toltecus* should be elevated to species status. It was the greater than expected levels of sequence divergence between *S. hirsutus*, *S. hispidus*, and *S. toltecus* compared with other recognized species in the genus that initially drew our attention to the presence of these cryptic species.

Likewise, before our study, *Neotoma albigula*, *Peromyscus boylii*, and *Reithrodontomys megalotis* were identified as possessing greater genetic variation within a taxon

than was present between other taxa currently recognized as congeners (Bell et al. 2001; Bradley et al. 2000; Edwards et al. 2001; Tiemann-Boege et al. 2000). In those datasets, there were instances where members of conspecific taxa were paraphyletic and possessed sequence-divergence values greater than typically distinguishing sister species within that genus. Two scenarios could explain those observations. First, lineage sorting of haplotypes before speciation (Avice 1994) could produce paraphyletic patterns. Second, currently recognized conspecific taxa actually represent >1 biological species. Consequently, based on the above data and congruency with other datasets, those authors recommended or supported the splitting of *N. albigula* into 2 species (*N. albigula* and *N. leucodon*), *P. boyllii* into 3 species (*P. beatae*, *P. boyllii*, and *P. levipes*), and *R. megalotis* into 2 species (*R. megalotis* and *R. zacatecae*). Of course, species identification is at best tentative when examining nonsympatric populations, and we view these decisions for recognition of new species as hypotheses for further investigation. Further investigation should help us understand the magnitude of variation in the cytochrome-*b* gene and how it can be used in studies of biodiversity.

Perhaps the greatest utility of sequence-divergence values is simply in the identification of taxa and questions for further investigation. For example, in our study *C. castanea* possessed greater genetic variation ( $\bar{X} = 8.70$ ), within what currently is viewed as a single species than is documented for intrageneric comparisons of *Carollia* ( $\bar{X} = 7.87$ ). Similarly, *G. soricina* and *C. brevicauda* exhibited intraspecific divergence values (5.91 and 5.24, respectively) that were larger than those obtained for other bat species ( $\bar{X} = 3.00$ ). We view these values as evidence that multiple species may be present within each of these taxa and further investigations are warranted. Although we expect that similar trends would be established in other nonmammalian

organisms, we leave that conformation or rejection to expert systematists in the appropriate discipline.

*Relevance to the Genetic Species Concept.*—Our observations support the hypothesis that values of DNA-sequence divergence could be used as an additional data source for establishment of an appropriate “measure of taxonomic rank.” Specifically, sequence-divergence values may be used to support the credence of the Genetic Species Concept. This is especially true when sequence-divergence values are used in conjunction with cladistic methodology. We propose use of sequence-divergence values (phenetic data) only when the relationships identified in the neighbor-joining analyses are identical to clades produced from parsimony or likelihood analyses (cladistic) of the same DNA sequences. This has the advantage of evaluating relationships with 2 alternative types of analyses (phenetic and cladistic), and invoking 2 species concepts (genetic—Dobzhansky 1950; Mayr 1969; Simpson 1943; phylogenetic—Cracraft 1983).

Historically, the Genetic Species Concept has been criticized because of its inability to establish ranges of genetic variation, sufficient sampling, etc. The ease with which DNA sequences are being generated addresses this criticism to the same extent that variation in morphology is used to detect species boundaries based on phenotype. Having genetic distance values that indicate (within 95% confidence limits) which populations are conspecific or where the speciation process has been completed would be ideal. Unfortunately, these types of data are not yet available. Therefore, the best scenario is a working criterion of genetic distances based on available empirical data. These categories can then serve as testable hypotheses that can be accepted, rejected, or modified as additional data become available.

Using this logic, we view genetic distance values (<2%), for the cytochrome-*b* gene, as having a high probability of indi-

cating intraspecific variation. Values  $>2\%$  may be indicative of conspecific populations or valid species. The higher the value, the greater the probability that 2 biological species may be represented. If data presented in this paper are predictive, then values from 7% to 11%, the burden of proof should shift toward documentation that only a single species is involved. Therefore, values  $<2\%$  would equal intraspecific variation; values between 2% and 11% would merit additional study concerning specific status, and values  $>11\%$  would be indicative of specific recognition. We would not expect these values to apply for other loci given the variability and rate of evolution between different genes. However, as more DNA sequences become available, it will be of interest to test for predictable variation at additional loci. It also will be of interest to see if other organisms with different life history strategies (i.e., birds, reptiles, fish, various invertebrates and plants) have similar relationships between genetic variation in the cytochrome-*b* gene and completion of speciation or if it will be necessary to develop a set of predictors independently for taxonomic groups.

Considering the extent to which genetic variation fluctuates in all natural processes and the complex nature of the speciation process, undoubtedly there will be examples where speciation will be completed with  $<2\%$  genetic differentiation in the cytochrome-*b* gene as well as instances where conspecific populations exist with  $>11\%$  genetic divergence. We predict that these examples will be uncommon and that these will provide additional insights into the nature of the speciation process.

Finally, a word of caution concerning use of mitochondrial genome as an indicator of genetic isolation. There will be examples where mitochondrial capture will result in misleading information about phylogenetic relationships. Examples include mule deer (*Odocoileus hemionus*) and white-tailed deer (*O. virginianus*)—Carr et al. 1986) and other instances of hybridization (Jones et al.

1995). Data from nuclear genes can help resolve examples where this is a problem. However, data on which our conclusions are based (cytochrome-*b* sequences) suggest that there is substantial information in the mitochondrial genome concerning the number of biological species in nature.

#### ACKNOWLEDGMENTS

We thank D. Hillis, F. Hoffmann, J. Wickliffe, C. Edwards, F. Mendez-Harclerode, and D. Carroll for helpful comments on various versions of this manuscript. We also thank the many individuals and institutions that provided tissue loans for the taxa used in the original studies and referenced in this study. This study was supported in part by the National Institutes of Health (DHHS A141435-01 to R. D. Bradley), the National Science Foundation (BSR 91-07-143 to R. J. Baker), and funds appropriated by the state of Texas to develop a biological database and bioinformatics program.

#### LITERATURE CITED

- AVISE, J. C. 1994. Molecular markers, natural history and evolution. Chapman and Hall, New York.
- AVISE, J. C., AND D. WALKER. 1999. Species realities and numbers in sexual vertebrates: perspectives from an asexually transmitted genome. Proceedings of the National Academy of Sciences of the United States of America 96:992–995.
- BAKER, R. J., V. A. TADDEI, J. L. HUDGEONS, AND R. A. VAN DEN BUSSCHE. 1994. Systematic relationships within *Chiroderma* (Chiropter: Phyllostomidae) based on cytochrome-*b* sequence variation. Journal of Mammalogy 75:321–327.
- BELL, D. M., ET AL. 2001. Patterns of karyotypic megaevolution in *Reithrodontomys*: evidence from a cytochrome-*b* phylogenetic hypothesis. Journal of Mammalogy 82:81–91.
- BRADLEY, R. D., I. TIEMANN-BOEGE, C. W. KILPATRICK, AND D. J. SCHMIDLY. 2000. Taxonomic status of *Peromyscus boylii sacarensis*: inferences from DNA sequences of the mitochondrial cytochrome-*b* gene. Journal of Mammalogy 81:875–884.
- BREMER, K. 1994. Branch support and tree stability. Cladistics 10:295–304.
- CARLETON, M. D. 1989. Systematics and evolution. Pp. 7–141 in Advances in the study of *Peromyscus* (Rodentia) (G. L. Kirkland, Jr. and J. N. Layne, eds.). Texas Tech University Press, Lubbock.
- CARR, S. M., S. W. BALLINGER, J. N. DERR, L. H. BLANKENSHIP, AND J. W. BICKHAM. 1986. Mitochondrial DNA analysis and hybridization between sympatric white-tailed deer and mule deer in West Texas. Proceedings of the National Academy of Sciences of the United States of America 83:9576–9580.
- CARROLL, D. S., L. L. PEPPERS, C. JONES, AND R. D. BRADLEY. In press. *Sigmodon ochrognathus* is

- monotypic: evidence from DNA sequences. *The Southwestern Naturalist*.
- CRACRAFT, J. 1983. Species concepts and speciation analysis. Pp. 159–187 in *Current ornithology* (R. F. Johnston, ed.). Plenum Press, New York.
- DOBZHANSKY, T. 1950. Mendelian populations and their evolution. *The American Naturalist* 74:312–321.
- EDWARDS, C. W., AND R. D. BRADLEY. 2001. Molecular phylogenetics of the *Neotoma floridana* species group. *Journal of Mammalogy* 82:791–798.
- EDWARDS, C. W., C. F. FULHORST, AND R. D. BRADLEY. 2001. Molecular phylogenetics of the *Neotoma albigula* species group: further evidence of a paraphyletic assemblage. *Journal of Mammalogy* 82:267–279.
- EDWARDS, S. V., AND P. BEERLI. 2000. Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54:1939–1854.
- ERIKSSON, T. 1997. Autodecay, version 3.03. Botaniska Institutionen, Stockholm University, Stockholm, Sweden.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791.
- HALL, E. R. 1981. *The mammals of North America*. 2nd ed. John Wiley & Sons, New York.
- HOFFMANN, F. G., AND R. J. BAKER. 2001. Systematics of bats of the genus *Glossophaga* (Chiroptera: Phyllostomidae) and phylogeography in *G. soricina* based on the cytochrome-*b* gene. *Journal of Mammalogy* 82:1092–1101.
- HOOFFER, S. R., AND R. A. VAN DEN BUSSCHE. 2001. Phylogenetic relationships of plecotine bats and allies based on mitochondrial ribosomal sequences. *Journal of Mammalogy* 82:131–137.
- JOHNS, G. C., AND J. C. AVISE. 1998. A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. *Molecular Biology and Evolution* 15:1481–1490.
- JOLLEY, T. W., R. L. HONEYCUTT, AND R. D. BRADLEY. 2000. Evolution of the 12S rRNA gene in pocket gophers (genus: *Geomys*). *Journal of Mammalogy* 81:1025–1034.
- JONES, J. K., R. D. BRADLEY, AND R. J. BAKER. 1995. Hybrid pocket gophers and some thoughts on the relationships of natural hybrids to the rules of nomenclature and the endangered species concept. *Journal of Mammalogy* 76:43–49.
- KIMURA, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16:111–120.
- MAYDEN, R. L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. Pp. 381–424 in *Species: the units of biodiversity* (M. F. Claridge, H. A. Dawah, and M. R. Wilson, eds.). Chapman and Hall, New York.
- MAYR, E. 1963. *Animal species and evolution*. The Belnap Press of Harvard University, Cambridge, Massachusetts.
- MAYR, E. 1969. *Principles of systematic zoology*. McGraw-Hill, New York.
- PEPPERS, L. L., AND R. D. BRADLEY. 2000. Molecular systematics of the genus *Sigmodon*. *Journal of Mammalogy* 81:332–343.
- PEPPERS, L. L., D. S. CARROLL, AND R. D. BRADLEY. In press. Molecular systematics of the genus *Sigmodon*: evidence from the mitochondrial cytochrome-*b* gene. *Journal of Mammalogy*.
- SIMPSON, G. G. 1943. Criteria for genera, species, and subspecies in zoology and paleontology. *Annals of the New York Academy of Science* 44:145–178.
- SMITH, M. F., AND J. L. PATTON. 1999. Phylogenetic relationships and the radiation of sigmodontine rodents in South America: evidence from cytochrome *b*. *Journal of Mammalogian Evolution* 6:89–128.
- SWOFFORD, D. L. 2000. PAUP\*: phylogenetic analysis using parsimony (\* and other methods). Version 4. Sinauer Associates, Inc., Publishers, Sunderland, Massachusetts.
- TIEMANN-BOEGE, I., C. W. KILPATRICK, D. J. SCHMIDLY, AND R. D. BRADLEY. 2000. Molecular phylogenetics of the *Peromyscus boylii* species group. *Molecular Phylogenetics and Evolution* 16:366–378.
- VAN DEN BUSSCHE, R. A., J. L. HUDGEONS, AND R. J. BAKER. 1998. Phylogenetic accuracy, stability, and congruence: relationships within and among the New World bat genera *Artibeus*, *Dermanura*, and *Koopmania*. Pp. 59–71 in *Bat biology and conservation* (T. H. Kunz and P. A. Racey, eds.). Smithsonian Institution Press, Washington, D.C.
- WILEY, E. O. 1981. *Phylogenetics: the theory and practice of phylogenetic systematics*. Wiley Interscience, New York.
- WILSON, D. E., AND D. M. REEDER. 1993. *Mammal species of the world*. 2nd ed. Smithsonian Institution Press, Washington, D.C.
- WILSON, E. O. 1998. *Consilience the unity of knowledge*. Alfred A. Knopf, Inc., New York.
- WRIGHT, A. J., R. A. VAN DEN BUSSCHE, B. K. LIM, M. D. ENGSTROM, AND R. J. BAKER. 1999. Systematics of the genera *Carollia* and *Rhinophylla* based on the cytochrome-*b* gene. *Journal of Mammalogy* 80:1202–1213.

Submitted 21 April 2001. Accepted 20 June 2001.

Associate Editor was David M. Leslie, Jr.