Evolution, 58(8), 2004, pp. 1794–1808

ADAPTIVE REPTILE COLOR VARIATION AND THE EVOLUTION OF THE *MC1R* GENE

ERICA BREE ROSENBLUM,[1,2] HOPI E. HOEKSTRA,[3,4] AND MICHAEL W. NACHMAN[3,5]
[1]*Museum of Vertebrate Zoology, University of California, Berkeley, 3101 Valley Life Sciences Building, Berkeley, California 94720*
[2]*E-mail: ericabr@socrates.berkeley.edu*
[3]*Department of Ecology and Evolutionary Biology, University of Arizona, BioSciences West Building, Tucson, Arizona 85721*
[5]*E-mail: nachman@u.arizona.edu*

Abstract.—The wealth of information on the genetics of pigmentation and the clear fitness consequences of many pigmentation phenotypes provide an opportunity to study the molecular basis of an ecologically important trait. The melanocortin-1 receptor (*Mc1r*) is responsible for intraspecific color variation in mammals and birds. Here, we study the molecular evolution of *Mc1r* and investigate its role in adaptive intraspecific color differences in reptiles. We sequenced the complete *Mc1r* locus in seven phylogenetically diverse squamate species with melanic or blanched forms associated with different colored substrates or thermal environments. We found that patterns of amino acid substitution across different regions of the receptor are similar to the patterns seen in mammals, suggesting comparable levels of constraint and probably a conserved function for *Mc1r* in mammals and reptiles. We also found high levels of silent-site heterozygosity in all species, consistent with a high mutation rate or large long-term effective population size. *Mc1r* polymorphisms were strongly associated with color differences in *Holbrookia maculata* and *Aspidoscelis inornata*. In *A. inornata*, several observations suggest that *Mc1r* mutations may contribute to differences in color: (1) a strong association is observed between one *Mc1r* amino acid substitution and dorsal color; (2) no significant population structure was detected among individuals from these populations at the mitochondrial ND4 gene; (3) the distribution of allele frequencies at *Mc1r* deviates from neutral expectations; and (4) patterns of linkage disequilibrium at *Mc1r* are consistent with recent selection. This study provides comparative data on a nuclear gene in reptiles and highlights the utility of a candidate-gene approach for understanding the evolution of genes involved in vertebrate adaptation.

Key words.—Adaptation, association study, melanocortin-1 receptor, molecular evolution, pigmentation, Squamata.

Received December 18, 2003. Accepted May 4, 2004.

Although squamate color has been studied in detail from ecological, physiological and systematic perspectives, little is known about the molecular basis of color evolution in this group. Reptile coloration has long been studied as an example of adaptive evolution (Cott 1940; Norris and Lowe 1964). Geographic variation in dorsal color is common in reptiles, and selective pressures can often be inferred. In some cases, color variation coincides with substrates of different color, while in other cases, color variation seems to be associated with different thermal environments. Overall body coloration is under strong natural selection in squamates because diurnal reptiles experience intense selection for substrate matching to diminish vulnerability to visual predators (Norris 1965; Kettlewell 1973). Color variation is also known to influence temperature regulation in a variety of ectotherms (Majerus 1998; Bittner et al. 2002). Advances in understanding the genetic basis of pigmentation in other vertebrate taxa provide an opportunity to investigate the genetic basis of color variation in reptiles.

A large number of genes involved in mammalian pigmentation have been isolated, and their functions have been well characterized (Bennett and Lamoreux 2003). One gene that has received much recent attention encodes the melanocortin-1 receptor (*Mc1r*). This G-protein coupled receptor is a key switch in a signal transduction pathway in melanocytes, the melanin-producing cells (Barsh 1996). In mammals and birds, with increased *Mc1r* activity production of eumelanin (dark, brown/black pigment) is enhanced; while decreased *Mc1r* activity results in the production of phaeomelanin (light, yellow/red pigment) (Robbins et al. 1993; Takeuchi et al. 1996a). Different mutations at *Mc1r* cause either light or dark phenotypes in several laboratory and domestic animals (e.g., mouse, Barsh 1996; dog, Newton et al. 2000; pig, Kijas et al. 1998; horse, Marklund et al. 1996; fox, Vage et al. 1997; chicken, Takeuchi et al. 1996b). In all experimentally verified cases, dominant mutations are associated with a gain of *Mc1r* function and result in dark color, while recessive mutations are associated with a reduction or loss of *Mc1r* function and result in light color. *Mc1r* has also been studied in natural populations and has been implicated in intraspecific color variation in both birds and mammals (e.g., bananaquit, Theron et al. 2001; lesser snow goose and arctic skua, Mundy et al. 2004; jaguar and jaguarundi, Eizirik et al. 2003; black bear, Ritland et al. 2001; pocket mouse, Nachman et al. 2003). In most of these examples, the different color morphs do not correspond with obvious differences in the environment, and thus the link to fitness is unclear. However, in pocket mice, melanic *Mc1r* mutants are found on dark lava fields while light animals are found on corresponding light-colored rocks, presumably as an adaptation for crypsis. Several other studies have failed to find an association between *Mc1r* genotype and intraspecific color variation (Kerns et al. 2003; MacDougall-Shackleton et al. 2003; Mundy and Kelly 2003). Because approaches that require detailed genomic knowledge or large-scale breeding experiments are not currently feasible with squamates, candidate-gene studies provide a useful way to investigate the molecular basis of reptile color variation.

Although reptile pigment cell morphology differs from that of mammals and birds, the function of melanin is remarkably conserved. Reptiles have three important cell layers for pig-

[4] Present address: Division of Biological Sciences, University of California, San Diego, 9500 Gilman Drive, MC 0116, La Jolla, California 92093; E-mail: hoekstra@ucsd.edu.

ment production. The layer closest to the epidermis consists of pigment-containing xanthophore cells that generate yellow or orange colors. The middle layer consists of iridophore cells that produce structural colors through the reflective properties of the cells. Finally, the deepest pigment cell layer produces melanin, and the overall darkness of the body is largely a consequence of the amount of melanin deposited by these melanophores (Bagnara and Hadley 1973; Morrison et al. 1995). Unlike mammals and birds, reptile melanophores are known only to produce eumelanin (Bagnara and Hadley 1973), so *Mc1r* activity might simply affect the amount of melanin produced (rather than switching between the two types of melanin). In spite of these differences, changes in the production and dispersion of melanin granules are ultimately responsible for changes in the dorsal color of reptiles (Hadley 1997).

In North America, there are a number of particularly striking cases of reptile color variation for which *Mc1r* is a promising candidate gene. A subset of populations in these taxa exhibits dramatic changes in the darkness of the dorsal portion of the body (Fig. 1). Seven species from four families (Anniellidae, Colubridae, Phrynosomatidae, and Teiidae) were chosen for this study to sample various phylogenetic levels and to represent diverse selective pressures (Table 1). In all cases, there is a geographically widespread wild-type phenotype and a geographically restricted variant phenotype (Fig. 1). The variant phenotypes are thought to be derived because other conspecific and congeneric populations exhibit wild-type color morphology. In some species, the derived phenotype is darker than the wild type and is referred to as a ''melanic'' form. In other species, the derived phenotype is lighter than the wild type and is referred to as a ''blanched'' form. In five species the different color morphs are associated with different-colored substrates, while in two species the different color morphs are thought to be associated with different thermal environments. *Phrynosoma platyrhinos* (desert horned lizard) and *Uta stansburiana* (common side-blotched lizard) have melanic forms on dark substrates at the Pisgah Lava flow in southern California (Norris and Lowe 1964). *Aspidoscelis inornata* (little striped whiptail, see Dixon 1967; formerly *Cnemidophorus inornatus*, Reeder et al. 2002), *Holbrookia maculata* (common lesser earless lizard, see Smith 1943), and *Sceloporus undulatus* (eastern fence lizard; see Lowe and Norris 1956) have blanched forms on the light-colored substrates of White Sands National Monument in southern New Mexico. *Sceloporus undulatus* also has melanic populations on the dark rocks at the Carrizozo Lava Flow in southern New Mexico (Lewis 1949). Substrate matching in these species is generally considered to be an adaptation for crypsis. *Anniella pulchra* (California legless lizard) has melanic populations along coastal California (Pearse and Pogson 2000). *Thamnophis sirtalis* (common gartersnake) has melanic populations on islands in Lake Erie (Lawson and King 1996) in which melanism appears to be controlled by a single recessive gene (King 2003). Melanism in *A. pulchra* and *T. sirtalis* is hypothesized to be a thermoregulatory adaptation to cool environments (Bittner et al. 2002).

Here, we evaluate patterns of *Mc1r* variation at three hierarchical levels. First, we compare interspecific patterns of *Mc1r* nucleotide evolution among reptiles to those previously documented among mammals. We evaluate whether patterns of constraint are consistent with *Mc1r* having a similar function in reptiles and in mammals. Second, we document patterns of variation within squamate species and present some of the first estimates of autosomal nucleotide diversity in reptiles. These patterns of variation provide a framework in which to interpret association studies. Finally, we compare intraspecific variation at *Mc1r* with color variation in each of seven species as a first step toward identifying the genetic basis of color variation in natural populations of squamates.

Materials and Methods

Sampling

A total of 130 individuals representing seven species were sampled (Table 1). For each taxon, two individuals were initially sampled from each color morph. For the three taxa in which an association was observed between *Mc1r* genotype and color phenotype (*A. inornata*, *H. maculata*, and *S. undulatus*, see Results), sample size was increased to 10 or more individuals per color morph (Table 1). Association studies (linkage disequilibrium mapping) can identify genes underlying a particular trait through nonrandom associations of genotypes and phenotypes. Population structure can confound this approach by creating spurious associations between phenotypes and causally unrelated genotypes. To minimize the confounding effect of population structure, samples were obtained for wild-type and derived populations in as close geographic proximity as possible. The average distance between color morphs in this study was approximately 70 km. Collecting localities are given in Table 1; a list of specimens is provided in Appendices 1 and 2 (see Appendices 1–4, available online at http://dx.doi.org/10.1554/03-741.1.s1).

Isolation of Reptile Mc1r

Whole genomic DNA was extracted from frozen tissue or blood with Qiagen (Valencia, CA) extraction kits. Primers were initially developed to amplify a portion of the reptile *Mc1r* gene by identifying conserved regions in several taxa for which *Mc1r* sequence has been reported (*Gallus gallus*, Genbank AY220305; *Tangara cucullata*, AF362606; *Homo sapiens*, AF514787; *Mus musculus*, NM_008559; *Chaetodipus intermedius*, AY259033; *Bos taurus*, AF445641; *Canis familiaris*, AF064455; *Takifugu poecilonotus* AB073678). These initial primers and the corresponding positions in the *Mc1r* sequence of *Mus* (NM_008559) were: F1: 5′-TGG GGC TGG TGA GCY TGG TG-3′ (site 137–156 *Mus*), F2: 5′-TAC TAC TTC ATC TGC TGC CTG GC-3′ (site 214–236 *Mus*), R1: 5′-CCC AGS AGG ATG GTG AGG GTG-3′ (site 737–715 *Mus*), R2: 5′-AAG GCR TAG ATG AGG GGG TC-3′ (site 893–874 *Mus*). *Mc1r* consists of a single exon (954 bp of coding region in *Mus*). An internal portion of the gene (~600 bp) was amplified, sequenced, and used to develop nested, species-specific genome walking primers. Using a Universal Genome Walker Kit (Clonetech, Palo Alto, CA), libraries were created by digesting genomic DNA with restriction enzymes EcoRV, DraI, PvuII, and StuI. Genome walking adaptors were then ligated onto the digested fragments. Using a

FIG. 1. Map of collecting localities and photographs comparing wild-type and derived color morphs of seven squamate species. Gray, black, and white circles represent collecting localities of wild-type, melanic, and blanched forms, respectively. Species collected at each locality are identified by genus and species initials.

TABLE 1. Taxa surveyed and sampling design. Numbers of individuals sampled for each color morph are presented for seven squamate species. Collection locality abbreviations used: JLTER, Jornada Long-Term Ecological Research Station; WSNM, White Sands National Monument; WSMR, White Sands Missile Range. Sample names are assigned based on genus and species initials, followed by ''W'' for wild type, ''M'' for melanic, or ''B'' for blanched, and a sequential numeral.

Species	Color morph (no. sampled)	Collection locality	Sample names	Distance between morphs (km)
Anniella pulchra California legless lizard	wild type (2)	Monterey and Kern Cos., CA	ApW1, ApW2	160
	melanic (2)	Spanish Bay, Monterey Co., CA	ApM1, ApM2	
Thamnophis sirtalis Common gartersnake	wild type (2)	Chaffey's Lock, Ontario, Canada	TsW1, TsW2	600
	melanic (2)	Pelee Island, Ontario, Canada	TsM1, TsM2	
Phrynosoma platyrhinos Desert horned lizard	wild type (2)	Barstow and Kelso Dunes, San Bernardino Co., CA	PpW1, PpW2	60
	melanic (2)	Pisgah Lava Flow, San Bernardino Co., CA	PpM1, PpM2	
Uta stansburiana Side-blotched lizard	wild type (2)	Adjacent to Pisgah Lava Flow, San Bernardino Co., CA	UsW1, UsW2	5
	melanic (2)	Pisgah Lava Flow, San Bernardino Co., CA	UsM1, UsM2	
Sceloporus undulatus Eastern fence lizard	wild type (15)	JLTER and WSMR, Dona Ana and Otero Cos., NM	SuW1–SuW15	30
	blanched (11)	WSNM and WSMR, Otero Co., NM	SuB1–SuB11	
	melanic (10)	WSMR, Otero and Lincoln Cos., NM	SuM1–SuM10	
Holbrookia maculata Lesser earless lizard	wild type (15)	JLTER and WSMR, Dona Ana and Otero Cos., NM	HmW1–HmW15	80
	blanched (23)	WSNM and WSMR, Otero Co., NM	HmB1–HmB23	
Aspidoscelis inornata Little striped whiptail	wild type (15)	JLTER and WSMR, Dona Ana and Otero Cos., NM	AiW1–AiW15	60
	blanched (24)	WSNM and WSMR, Otero Co., NM	AiB1–AiB24	

nested polymerase chain reaction (PCR) protocol with gene-specific primers developed for each species and adaptor primers provided by Clontech, regions up- and downstream of *Mc1r* were amplified. Appropriate size fragments were sequenced for each of the seven target species to obtain sequence from the entire coding region and some flanking regions.

Species-specific PCR primers were then developed in regions adjacent to *Mc1r* to amplify the entire coding region of the gene (online Appendix 3). Diploid PCR products were sequenced in both directions on an ABI377 or ABI3100 (Applied Biosystems, Foster City, CA) with species-specific sequencing primers in addition to PCR primers. Internal primers universal to all reptiles surveyed were also used for sequencing. Sequences were edited and aligned in Sequencher (Gene Codes Co., Ann Arbor, MI). Nucleotide and amino acid polymorphisms were viewed in MEGA (Kumar et al. 2001). We sequenced diploid PCR products directly and did not attempt to resolve haplotype phase experimentally. Heterozygous sites were identified by visual inspection of chromatograms and confirmed by sequence from both DNA strands. All *Mc1r* sequences are deposited in GenBank (accession numbers AY586032–AY586157 and AY586159–AY586162).

Interspecific Molecular Evolution

To investigate patterns of molecular evolution, *Mc1r* sequences representing one wild-type individual per species were aligned with published *Mc1r* sequences from other vertebrate species. Other species included two avian taxa (chicken, *G. gallus*; tanager, *T. cucullata*), five mammals (human, *H. sapiens*; house mouse, *Mus domesticus*; rock pocket mouse, *C. intermedius*; domestic cow, *B. taurus*; and domestic dog, *C. familiaris*), and one fish (pufferfish, *T. poecilonotus*). All alignments were performed using CLUSTAL V (Higgins 1994), and the nucleotide alignment was adjusted based on the amino acid alignment. A tree of the 15 vertebrate species was constructed from *Mc1r* sequences using PAUP* 4.02 (Swofford 1999). Hierarchical likelihood-ratio tests implemented in Modeltest 3.06 (Posada and Crandall 1998) were used to determine the best-fit model of nucleotide substitution, GTR+ Γ + i. Parsimony, neighbor-joining, and maximum likelihood algorithms all produced the same topology. Bootstrap analyses were conducted to determine confidence in the branching structure.

To compare rates of molecular evolution between mammals ($N = 5$) and reptiles ($N = 7$), we quantified nucleotide variation across the entire gene and in different functional compartments of the gene using DnaSP (Rozas and Rozas 1999). In these analyses, we created one sequence per individual by randomly choosing one base at each heterozygous site, when present. Nucleotide differences were also calculated across the entire *Mc1r* gene using a sliding window with a 50-bp window size and 20-bp interval. To compare rates of nonsynonymous and synonymous evolution, K_A/K_S ratios (Nei and Gojobori 1986) were estimated between pairs of species with PAML (Yang 1997).

Intraspecific Patterns of Variation

For each taxon we investigated the level and pattern of *Mc1r* variation within species. Numbers and frequencies of all polymorphisms were counted. Haplotype phase was inferred computationally using the HAP program (Eskin et al. 2003). Two measures of nucleotide variability, π (Nei and Li 1979) and θ (Watterson 1975), were calculated. Nucleotide diversity, π, is based on the average number of nucleotide differences between two sequences randomly drawn from a sample, and θ is based on the proportion of segregating sites

in a sample. Under neutral, equilibrium conditions, both π and θ estimate the parameter $4N_e\mu$ for autosomal loci, where N_e is the effective population size and μ is the neutral mutation rate. Departures from a neutral, equilibrium frequency distribution of polymorphisms were evaluated using Tajima's *D* (Tajima 1989), which is based on the normalized difference between π and θ. At equilibrium with respect to mutation and drift, the expectation for Tajima's *D* is zero; deviations from this expectation may be due to selection, demographic influences, or both. π, θ and Tajima's *D* were calculated using DnaSP.

For the three species with large sample sizes (*S. undulatus*, *H. maculata*, and *A. inornata*) several other analyses were performed. The neutral expectation of equal ratios of synonymous to nonsynonymous changes within and between species (McDonald and Kreitman 1991) was tested using Fisher's exact tests (Fisher 1934). Ratios of polymorphism within species to divergence between species were also compared with expectations under a neutral model using the HKA test (Hudson et al. 1987). In this test, *Mc1r* sequences were compared to sequences from the mitochondrial ND4 gene (*A. pulchra*, AY620747; *A. inornata*, AY620816; *H. maculata*, AY620749; *S. undulatus*, AY620748). Primers used to amplify and sequence ND4 and associated tRNAs were ND4: 5′-CAC CTA TGA CTA CCA AAA GCT CAT GTA GAA GC-3′ and LEU: 5′-CAT TAC TTT TAC TTG GAT TTG CAC CA-3′ (Arevalo et al. 1994). Significance of HKA tests were evaluated by coalescent simulation. The McDonald-Kreitman test is based on fixed differences between species and the HKA test is based on a single randomly chosen allele from each species (Hudson et al. 1987; McDonald and Kreitman 1991). For these tests, each species was compared to the closest taxon for which *Mc1r* data were available (see Results for details). Using DnaSP, haplotypes were also used to estimate the population recombination parameter, $C = 4N_ec$, where *c* is the recombination rate per nucleotide (Hudson et al. 1987).

Association Studies

We first amplified and sequenced *Mc1r* from two individuals (four alleles) of each color morph in each species to look for associations between genotype and color phenotype. In the four species in which no association was seen among these eight alleles, we did not sequence additional alleles. In the three species in which an association was seen (*S. undulatus*, *H. maculata*, and *A. inornata*), we sequenced *Mc1r* from a minimum of 10 additional individuals of each color morph (Table 1). Associations between single nucleotide polymorphisms and color were tested in 2×2 contingency tables using Fisher's exact tests.

Because population structure can lead to spurious associations, we estimated the degree of population structure by calculating F_{ST} (Wright 1951) for the two species showing a strong association between *Mc1r* and color. This was done for both *Mc1r* and for approximately 750 bp of the mitochondrial ND4 gene in *H. maculata* and *A. inornata*. To test for obvious departures from neutrality in the mitochondrial DNA (mtDNA) dataset, Tajima's *D* was calculated for each color morph. We then compared F_{ST} for *Mc1r* and ND4. All F_{ST} estimates were bootstrapped with 1000 permutations to obtain confidence intervals using Arlequin (Schneider et al. 2000). Although the individuals sampled were not the same for *Mc1r* and ND4, samples were adjusted so that similar numbers of individuals from the same collecting localities were used in the analysis. Mitochondrial sequences used for F_{ST} calculations are deposited in GenBank (AY620749–AY620816). The comparison between *Mc1r* and mtDNA provides a conservative test for population structure because mtDNA has an effective population size one-quarter that of autosomal genes and is therefore expected to show greater differentiation than *Mc1r* among populations under neutral conditions. For *A. inornata*, tests of linkage disequilibrium were conducted in DnaSP, and statistical significance was evaluated using a chi-square test with corrections for multiple comparisons.

RESULTS

Interspecific Molecular Evolution

The amino acid alignment of *Mc1r* from the seven squamate species and eight other vertebrates is shown in Figure 2. The nucleotide sequence alignment for the seven reptiles is provided in online Appendix 4. The length of *Mc1r* differs among these species; in the phrynosomatid lizards *H. maculata*, *P. platyrhinos*, *S. undulatus*, and *U. stansburiana*, *Mc1r* is 948 bp, while in *T. sirtalis* and *A. inornata Mc1r* is 945 bp. In *A. pulchra*, the first extracellular region of the gene is highly divergent from the other species and has two start codons that are in frame. However, an intervening stop codon suggests that only one start codon is functional, resulting in a gene length of 921 bp.

A neighbor-joining tree for 15 vertebrate *Mc1r* sequences is shown in Figure 2. The pufferfish *Mc1r* sequence was used to root this tree. The topology of the tree is consistent with the well-established phylogeny of vertebrate classes: mammals are the sister group to the clade birds + squamates, and birds are the sister group to the squamates, which are monophyletic. Lower-level relationships revealed by *Mc1r* are also generally consistent with other phylogenies of these taxa, although there are some minor differences (e.g., Estes and Pregill 1998; Murphy et al. 2001). Bootstrap analyses showed strong support ($> 95\%$) for Reptilia, Aves, and Mammalia and for a monophyletic clade of all sand lizards (*Holbrookia*, *Phrynosoma*, *Sceloporus*, and *Uta*).

Table 2 summarizes rates of evolution for different regions of *Mc1r* in both mammals and squamates. The extracellular regions of the *Mc1r*-encoded receptor are important for ligand binding and regulation of *Mc1r* activity, the hydrophobic transmembrane domains maintain the structural integrity of the protein, and the intracellular regions are important for G-protein signal transduction. Rates of nonsynonymous evolution (K_A) in squamates differ by more than five-fold in comparisons among different regions of the receptor: the first extracellular region evolves most quickly, followed by all other extracellular regions, all transmembrane regions, and finally all intracellular regions. In contrast, rates of synonymous substitution (K_S) vary by less than two-fold among these different regions of the receptor, suggesting only small differences (if any) in the underlying mutation rate. Thus,

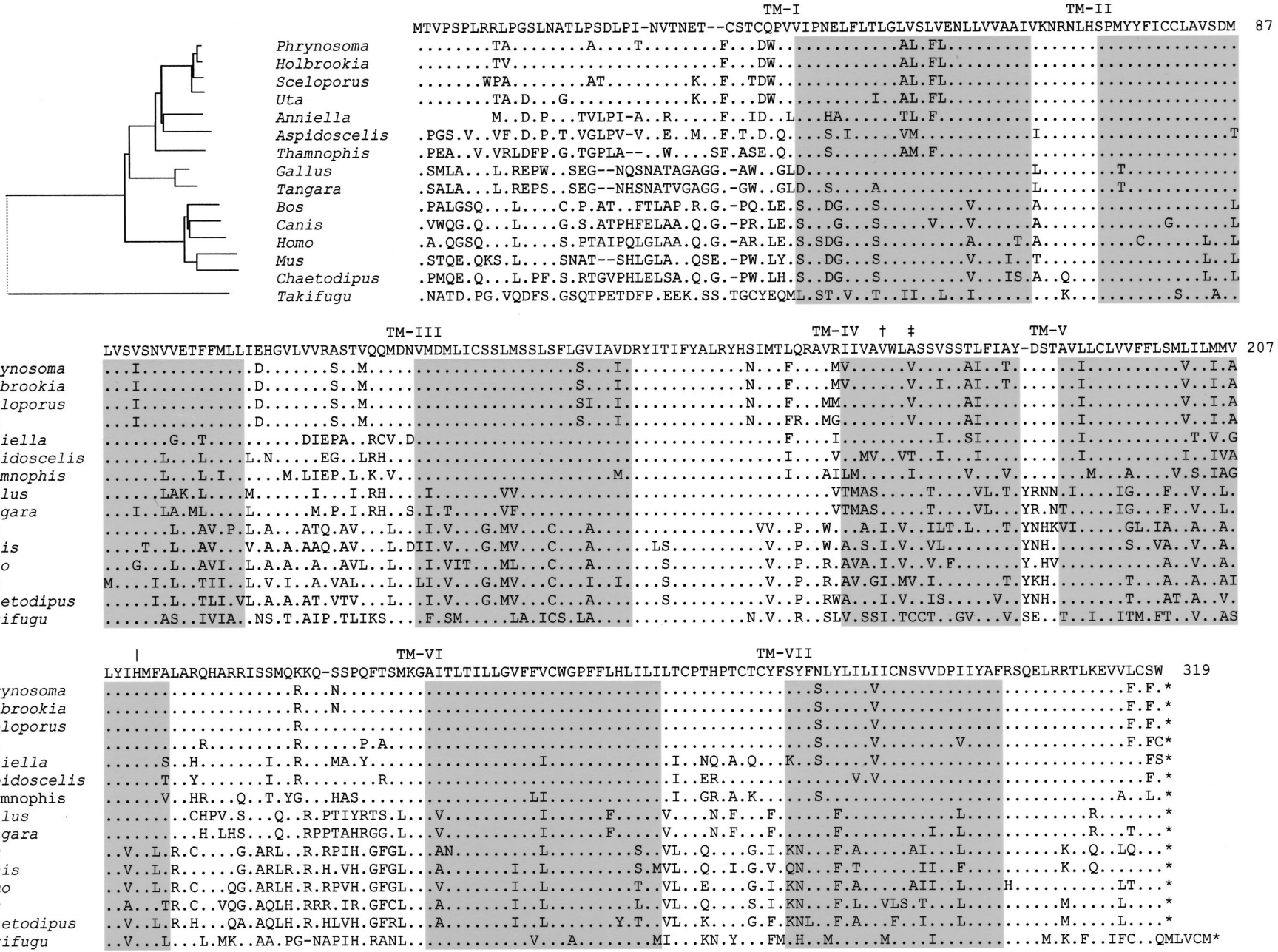

FIG. 2. *Mc1r* amino acid alignment for vertebrates. A neighbor-joining phylogenetic reconstruction is shown in the upper-left corner. Branch lengths are proportional to amount of sequence change. One individual from each of one fish, five mammal, two bird, and seven reptile species is identified in the left column by generic name. The consensus sequence is listed on top with single letter abbreviations for amino acids. Dots represent identity to the consensus, dashes represent gaps, and asterisks represent stop codons. Shaded areas represent transmembrane domains. Symbols indicate amino acid residues at which there is an association with color variation in the following species: *Holbrookia maculata* (†), *Aspidoscelis inornata* (‡), *Sceloporus undulatus* (|).

TABLE 2. Nonsynonymous substitutions (K_A), synonymous substitutions (K_S), and the ratio of nonsynonymous/synonymous substitutions (K_A/K_S) for different *Mc1r* protein regions for reptiles and mammals. K_A/K_S estimates are presented for the complete *Mc1r* gene and then for regions corresponding to different functional compartments: the first extracellular region, all other extracellular regions combined, all transmembrane regions, and finally all intracellular regions. Gene regions are presented in order of decreasing K_A/K_S.

Gene region	Length (bp)	Reptile K_A	Mammal K_A	Reptile K_S	Mammal K_S	Reptile K_A/K_S	Mammal K_A/K_S
Entire gene	984	0.104	0.123	0.495	0.601	0.210	0.205
First extracellular region	120	0.375	0.282	0.729	0.713	0.514	0.396
All other extracellular regions	105	0.223	0.142	0.539	0.649	0.414	0.219
All transmembrane regions	513	0.059	0.106	0.436	0.540	0.135	0.196
All intracellular regions	174	0.054	0.075	0.542	0.728	0.100	0.103

differences in the rates of protein evolution probably reflect differences in the level of constraint, as can be seen in the K_A/K_S values (Table 2). The overall level of constraint across the entire *Mc1r* gene is nearly identical in reptiles (K_A/K_S = 0.210) and mammals (K_A/K_S = 0.205) and is close to the average value for many genes in mammals (average K_A/K_S = 0.20, Li 1997). The rank order of K_A/K_S for the four different regions of the receptor (Table 2) is also the same in reptiles and mammals. Finally, in both groups, length variation is only seen in the first extracellular region, which also shows the highest K_A/K_S value. Sliding window analyses also reveal similarities in rates of protein evolution between mammals and reptiles across the *Mc1r* coding region (Fig. 3).

We looked for evidence of selection in interspecific comparisons by calculating *Mc1r* K_A/K_S ratios for each species pair separately. No K_A/K_S ratio is greater than one, indicating that no strong positive selection is detected in any of the mammalian or reptilian lineages. However, this approach is limited to finding extremely strong selection because K_A/K_S values are averaged across the gene.

Intraspecific Patterns of Variation

Levels of nucleotide polymorphism at *Mc1r* are summarized in Table 3 for each reptile species and for each color morph within each species, both for the entire gene and for synonymous and nonsynonymous sites separately. The average level of nucleotide diversity across all seven species is 0.287%; the lowest value is seen in *U. stansburiana* (π = 0.124%), and the highest value is seen in *T. sirtalis* (π = 0.423%). The average level of nucleotide diversity for synonymous sites for the seven reptile species in Table 3 is 0.99%, roughly 10 times larger than the average value seen in humans or in house mice (Li and Sadler 1991; Nachman 1997). Because population structure can lead to high levels of nucleotide diversity, we estimated π for synonymous sites in each color morph separately; in most cases, animals representing each color morph were sampled within close proximity of each other. The average value of π by color morph is 0.83%, which is still substantially larger than seen in mammals. If mutation rates are approximately similar in reptiles and in mammals, this suggests a larger long-term effective population size for the reptiles studied here compared to humans and mice. For example, if we assume $\mu = 2 \times 10^{-8}$ (Nachman and Crowell 2000), then $N_e = \pi/4\mu = 0.0083/(8 \times 10^{-8}) = 10^5$. To further explore the idea that N_e is large, we estimated the population recombination parameter, $C = 4N_ec$, for each color morph of *S. undulatus*, *H. maculata*, and *A. inornata*, the three species for which we have large samples

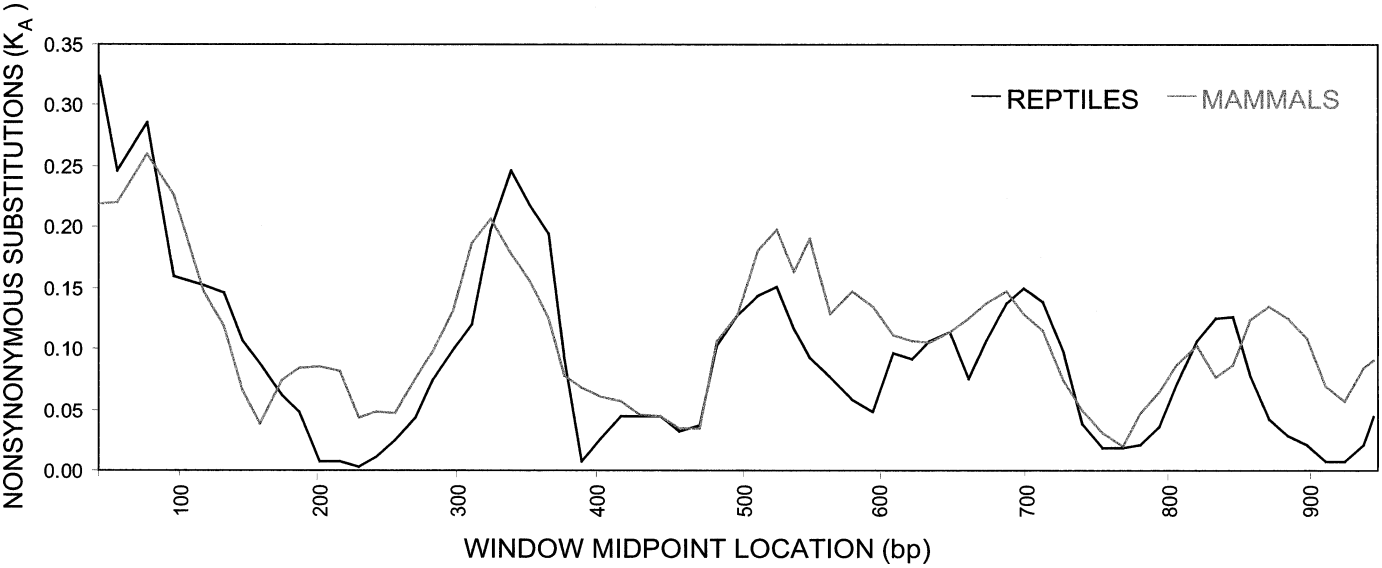

FIG. 3. Evolutionary constraint across *Mc1r* in reptiles and mammals. Sliding window comparison of nonsynonymous substitution rate (K_A) for mammals and reptiles across all regions of the *Mc1r* gene.

TABLE 3. Intraspecific patterns of variation at *Mc1r* within reptiles. Sample size is given as number of alleles and is therefore twice the number of individuals. For each species, estimates are first presented for all samples combined and then for each color morph separately. Number of segregating sites is presented for all groups. Estimators of nucleotide diversity (π) and nucleotide polymorphism (θ) are presented for all sites combined and for synonymous (S) and nonsynonymous (NS) sites separately. π and θ are per site estimates. The population recombination parameter (C) is estimated between adjacent sites only for species with large sample sizes. Asterisk indicates $P < 0.05$ for Tajima's D in *Aspidoscelis inornata*.

Species	Gene Length (bp)	No. Alleles	No. Seg. Sites	π (all sites)	π (S)	π (NS)	θ (all sites)	θ (S)	θ (NS)	C	Tajima's D
Anniella pulchra	921	8	9	0.00299	0.01006	0.00072	0.00377	0.01207	0.00111		−1.01760
Wild type		4	7	0.00398	0.01193	0.00144	0.00415	0.01220	0.00157		−0.38921
Melanic		4	1	0.00054	0.00223	0.00000	0.00059	0.00244	0.00000		−0.61237
Thamnophis sirtalis	945	8	10	0.00423	0.01625	0.00035	0.00408	0.01504	0.00054		0.18387
Wild type		4	7	0.00441	0.01805	0.00000	0.00404	0.01653	0.00000		0.89518
Melanic		4	7	0.00406	0.01445	0.00070	0.00404	0.01419	0.00077		0.03892
Phrynosoma platyrhinos	948	8	5	0.00226	0.00950	0.00000	0.00204	0.00855	0.00000		0.50437
Wild type		4	5	0.00334	0.01180	0.00070	0.00288	0.00965	0.00076		1.54057
Melanic		4	1	0.00070	0.00295	0.00000	0.00058	0.00242	0.00000		1.63299
Uta stansburiana	948	8	4	0.00124	0.00329	0.00060	0.00163	0.00507	0.00054		−1.02972
Wild type		4	4	0.00229	0.00657	0.00093	0.00230	0.00716	0.00076		−0.06501
Melanic		4	0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000		0.00000
Sceloporus undulatus	948	72	18	0.00418	0.01410	0.00112	0.00392	0.01105	0.00172	0.0417	0.20206
Wild type		30	16	0.00371	0.01214	0.00111	0.00426	0.01127	0.00210	0.2027	−0.43658
Melanic		20	5	0.00198	0.00839	0.00000	0.00149	0.00630	0.00000	>10	1.18636
Blanched		22	13	0.00470	0.01459	0.00165	0.00376	0.01100	0.00152	0.0165	0.88662
Holbrookia maculata	948	76	22	0.00419	0.01477	0.00089	0.00473	0.01722	0.00085	0.0094	−0.35106
Wild type		30	15	0.00360	0.01321	0.00060	0.00399	0.01343	0.00105	0.0285	−0.33137
Blanched		46	9	0.00165	0.00659	0.00012	0.00216	0.00810	0.00032	0.0043	−0.66872
Aspidoscelis inornata	945	78	4	0.00160	0.00439	0.00071	0.00086	0.00266	0.00028	0.0137	1.78487*
Wild type		30	3	0.00104	0.00428	0.00000	0.00080	0.00330	0.00000	>10	0.69721
Blanched		48	4	0.00076	0.00201	0.00036	0.00095	0.00295	0.00032	0.0000	−0.47217

(Table 3). Larger populations are expected to exhibit less linkage disequilibrium and thus higher values of *C*. These estimates are based on haplotypes that have been resolved computationally (rather than experimentally) and thus may contain some error. Nonetheless, the average value of $4N_ec$ from each color morph is 0.0994, a value considerably larger than in humans (e.g., Przeworski and Wall 2001). If we assume that recombination rates are similar in reptiles and mammals (roughly 10^{-8} per site), then patterns of linkage disequilibrium at *Mc1r* in our data suggest an average long-term population size of approximately 10^6 for each of the color morphs of *S. undulatus*, *H. maculata*, and *A. inornata* ($N_e = [0.0994]/[4 \times 10^{-8}] = 2.4 \times 10^6$). While these two estimates of N_e differ by a factor of ten and are necessarily rough, both values are substantially larger than seen in mammals.

Patterns of nucleotide variation within species are largely consistent with a neutral model of molecular evolution, with a few exceptions. The distribution of allele frequencies at *Mc1r* as reflected in Tajima's *D* is generally consistent with neutral expectations for each species (Table 3). With the exception of Tajima's *D* for the total sample of *A. inornata* (discussed below), none of the other values differ significantly from the neutral expectation of zero. We compared the ratio of nonsynonymous to synonymous changes within and between species for the three species with the largest sample sizes (*S. undulatus*, *H. maculata*, *A. inornata*) using the McDonald-Kreitman test. We conducted three tests: (1) polymorphism within *S. undulatus* compared to *S. undulatus–H. maculata* fixed differences; (2) polymorphism within *H. maculata* compared to *H. maculata–S. undulatus* fixed differences; and (3) polymorphism within *A. inornata* compared to *A. inornata–A. pulchra* fixed differences. The three intraspecific ratios are similar to one another and to the ratios observed between species, providing no evidence of selection on *Mc1r* (intraspecific nonsynonymous/synonymous:interspecific nonsynonymous/synonymous test 1 6/12:10/29, FET $P > 0.05$; test 2 3/19:10/29, FET $P > 0.05$; test 3 1/3:75/99, FET $P > 0.05$). These tests were also performed using divergence values corrected for multiple substitutions (Jukes and Cantor 1969), and the results were unchanged. We performed three HKA tests with the same set of taxa, comparing variation at *Mc1r* to ND4 within and between species (*Mc1r* polymorphism/divergence:ND4 polymorphism/divergence 1 22/52: 25/153, 2 18/52:44/153, 3 4/181:11/266). Tests 1 and 2 are not significant ($P > 0.05$ for each), however test 3, comparing polymorphism within *A. inornata* to *A. inornata–A. pulchra* divergence, is highly significant and reveals a lower ratio of polymorphism to divergence at *Mc1r* compared to ND4 (HKA $\chi^2 = 17.1$, $P < 0.01$). HKA tests for *A. inornata* are significant for the entire sample and for both blanched and wild-type samples. HKA tests were performed with and without corrections for multiple substitutions and the results were the same.

Association Studies

Comparisons between *Mc1r* genotype and color phenotype are shown in Figure 4 for each of the seven species. In the initial screen, no association between *Mc1r* and color was observed in *A. pulchra*, *P. platyrhinos*, *U. stansburiana*, or *T. sirtalis*, ruling out *Mc1r* as a strong contributor to the pigmentation differences in these species. In each of these species, *Mc1r* polymorphisms are observed, so the absence of an association is not due to lack of variation. In *S. undulatus*, *H. maculata*, and *A. inornata*, an association between *Mc1r* genotype and color was initially detected.

After increasing the sample size in *S. undulatus*, we observed a weak association between color and an amino acid polymorphism at site 208 (Fig. 4). Although this association is significant using a Fisher's exact test ($N = 36$, FET $P < 0.001$), homozygous histidine and heterozygous histidine/tyrosine individuals are found in both wild-type and blanched populations. This suggests that *Mc1r* does not play a simple role in *S. undulatus* color variation, if any role at all.

With increased sample sizes, associations between *Mc1r* genotype and color remained strong and highly significant for both *A. inornata* ($N = 39$, FET $P < 10^{-7}$) and *H. maculata* ($N = 38$, FET $P < 10^{-7}$). In *H. maculata*, one transition at nucleotide site 502 is strongly associated with color differences. At this position, 14 of 15 wild-type samples are homozygous G/G, while one is heterozygous A/G. Twenty-one of 23 blanched individuals are homozygous A/A and two are heterozygous A/G. This polymorphism leads to a conservative amino acid change at position 168, in the fourth transmembrane domain, of valine to isoleucine. This site is otherwise invariant in all reptiles (fixed for valine), but is fixed for isoleucine in mammals (Fig. 2).

In *A. inornata*, a polymorphism at nucleotide position 509 is strongly associated with color phenotype. All 15 wild-type individuals are homozygous C/C at this site, and 23 of 24 blanched samples are homozygous T/T or heterozygous T/C. Eighteen blanched individuals are homozygous for this change, and five are heterozygous. This C-to-T transition results in a polarity-changing amino acid substitution at amino acid position 170, in the fourth transmembrane domain, from threonine to isoleucine. This site is not conserved among other reptiles or among mammals (Fig. 2); in reptiles, alanine, valine, threonine, and isoleucine residues are observed in different species.

For the two species for which *Mc1r* genotypes were strongly associated with color, *A. inornata* and *H. maculata*, F_{ST} was calculated for both *Mc1r* and mtDNA (Table 4). Levels of mtDNA variation are presented in Table 5. For *H. maculata*, mtDNA F_{ST} calculations were based on 797 bp of ND4 for 15 wild-type and 24 blanched samples. In this taxon, sampling localities for which mtDNA data were not available were removed from the *Mc1r* dataset, leaving 11 wild-type and 12 blanched individuals for estimating *Mc1r* F_{ST}. For *H. maculata*, *Mc1r* $F_{ST} = 0.63$ and mtDNA $F_{ST} = 0.94$. For *A. inornata*, mtDNA F_{ST} calculations were based on a 779-bp fragment of ND4 sequenced for 11 wild-type and 18 blanched individuals. In this taxon, *Mc1r* $F_{ST} = 0.64$ and mtDNA $F_{ST} = 0.09$ with nonoverlapping confidence intervals. These results suggest that there may be substantial population structure in *H. maculata* but not in *A. inornata*.

Patterns of linkage disequilibrium at *Mc1r* in *A. inornata* are noteworthy. There is only a single haplotype associated with isoleucine at amino acid site 170, while there are at least four haplotypes associated with the alternative residue (thre-

Anniella pulchra

spec.	nuc.	a.a.	
	01235689	0	1
	87407171	5	0
	43683298	8	3
	GGCCCGGG	R	T
ApW1	.RYY.RR.	R/Q	T/I
ApW2	R.....	.	.
ApM1	....T..R	.	.
ApM2	....T...	.	.

Thamnophis sirtalis

	002346678	0
	259804693	1
	977155951	0
	GTSGCCTCT	A
TsW1	...A....Y	.
TsW2	.CG...CT.	.
TsM1	R........	A/H
TsM2	...RYYY..	.

Phrynosoma platyrhinos

	12588
	44467
	79970
	CCCCA
PpW1	.TTTG
PpW2	.....
PpM1	T....
PpM2	.....

Uta stansburiana

	3589	3
	0264	1
	9875	5
	CCGG	W
UsW1	...T	C
UsW2	YYR.	.
UsM1	....	.
UsM2	....	.

Sceloporus undulatus

spec.	nuc.	a.a.			
	*				*
	001223446666777889	0	0	1	2
	594181341129188340	2	3	0	0
	981689241223439696	0	3	7	8
	GCTGCGTCCGCGTCCGTA	S	S	V	H
SuW1	....	.	.	.	.
SuW2	......C.........C.	.	.	.	.
SuW3	...R.R..........CG	.	.	V/M	.
SuW4	C..A..Y.Y.......Y.	T	.	.	.
SuW5	C..R..Y..R...Y..C.	T	.	.	.
SuW6	...............CG	.	.	.	.
SuW7	...R............CG	.	.	.	.
SuW8	C..AS.C...........	T	.	.	.
SuW9	...A..C..A......C.	.	.	.	.
SuW10	.Y.A..Y..R..Y.Y.C.	.	T/L	.	.
SuW11	C.....C..A......C.	T	.	.	.
SuW12	..YR..CY.AY.....C.	.	.	.	Y/H
SuW13	...A..Y.....Y...Y.	.	.	.	.
SuW14	..YR..CY.RY.....Y.	.	.	.	Y/H
SuW15	......Y.........CR	.	.	.	.
SuB1	...A............CG	.	.	.	.
SuB2	..YR..YY.RY.....CR	.	.	.	Y/H
SuB3	..Y...CY.AY.....C.	.	.	.	Y/H
SuB4	.YYR..YY.RY.Y.Y.C.	.	T/L	.	Y/H
SuB5	..C...CT.AT.....C.	.	.	.	Y
SuB6	C.C...CT.AT.....C.	T	.	.	Y
SuB7	C.C...YT.AT....KC.	T	.	.	Y
SuB8	C.YR........Y.....	T	.	.	.
SuB9	..Y...YY.RY.....C.	.	.	.	Y/H
SuB10	..C...CT.AT.....C.	.	.	.	Y
SuB11	C..A..C..R........	T	.	.	.
SuM1	................YR	.	.	.	.
SuM2	...A..Y.........CR	.	.	.	.
SuM3	...R..Y.........CR	.	.	.	.
SuM4	...R.......R....CG	.	.	.	.
SuM5	...R..Y....R....YR	.	.	.	.
SuM6	...R............CG	.	.	.	.
SuM7	...R............CG	.	.	.	.
SuM8	...........R....CG	.	.	.	.
SuM9	......Y.........YR	.	.	.	.
SuM10	...R..Y.........CR	.	.	.	.

FET site 622: p <0.001

Holbrookia maculata

spec.	nuc.	a.a.		
	*		*	
	00011111234556666678899	0	1	3
	02912488109073489109O1	1	6	0
	69916109695290840119 60	0	8	0
	GTGGCCGCGCCGCCCCTTGGGA	V	V	S
HmW1	..........T....T..AAAC	.	.	N
HmW2	........RYT....T..A..C	.	.	.
HmW3	......R...Y...........	.	.	.
HmW4	..........T....T..A..C	.	.	.
HmW5	.........YYR...Y..R..C	.	V/I	.
HmW6	.........YT....T..A..C	.	.	.
HmW7	....	.	.	.
HmW8	R.........T....T..A..C	.	.	.
HmW9	..........T....T..A..C	.	.	.
HmW10	R.....R...T....T..A..C	.	.	.
HmW11	.Y.R.YR..........Y.A..	V/A	.	.
HmW12	..............T....AC	.	.	N
HmW13	..............Y......	.	.	.
HmW14	.Y........Y........NNN	V/A	.	.
HmW15	....	.	.	.
HmB1	..R....T...A..........	.	I	.
HmB2	..R....T...A..........	.	I	.
HmB3	.......T...AT.........	.	I	.
HmB4	.......T...AT.........	.	I	.
HmB5	..R....T...A..........	.	I	.
HmB6	...........AY...Y.....	.	I	.
HmB7	.......T...AY......NNN	.	I	.
HmB8	..R....T...A..........	.	I	.
HmB9	..A....T...A..........	.	I	.
HmB10	..A....T...A..........	.	I	.
HmB11	.......Y...AT.........	.	I	.
HmB12	..R....T...A..........	.	I	.
HmB13	....Y..Y...RYY...Y....	.	V/I	.
HmB14	.......T...AT...C.....	.	I	.
HmB15	..A....T...A..........	.	I	.
HmB16	..A....T...A..........	.	I	.
HmB17	.......T...AT.........	.	I	.
HmB18	.......T...AT.........	.	I	.
HmB19	.......T...AT.........	.	I	.
HmB20	.......T...AT.........	.	I	.
HmB21	..A....T...A..........	.	I	.
HmB22	.......T...RY.Y..Y....	.	V/I	.
HmB23	..A....T...A..........	.	I	.

FET site 502: p<10^-7

Aspidoscelis inornata

spec.	nuc.	a.a.
	*	*
	2557	1
	6046	7
	1995	0
	CCTT	T
AiW1	...Y	.
AiW2	..YY	.
AiW3	Y..Y	.
AiW4	...Y	.
AiW5	....	.
AiW6	..YY	.
AiW7	....	.
AiW8	...Y	.
AiW9	Y..C	.
AiW10	T..C	.
AiW11	....	.
AiW12	Y..Y	.
AiW13	..YY	.
AiW14	...Y	.
AiW15	....	.
AiB1	TT.C	I
AiB2	TT.C	I
AiB3	TY.C	I/T
AiB4	TT.C	I
AiB5	TT.C	I
AiB6	TT.C	I
AiB7	TT.C	I
AiB8	YYYC	I/T
AiB9	TT.C	I
AiB10	TT.C	I
AiB11	....	.
AiB12	TT.C	I
AiB13	TT.C	I
AiB14	TT.C	I
AiB15	TT.C	I
AiB16	TT.C	I
AiB17	TT.C	I
AiB18	TT.C	I
AiB19	TT.C	I
AiB20	TT.C	I
AiB21	YYYC	I/T
AiB22	TT.C	I
AiB23	YY.Y	I/T
AiB24	YY.Y	I/T

FET site 509: p<10^-7

FIG. 4. *Mc1r* nucleotide and amino acid polymorphism. Sample numbers are listed in the left column (spec.) for each species. All polymorphic sites are reported for nucleotide positions (nuc.) and amino acid positions (a.a.). Position information is read vertically. The consensus sequence is reported above individual sample data, and dots represent identity to the consensus. Asterisks identify sites for which significant associations between amino acid substitutions and color are observed. Results for Fisher's exact test (FET) testing for an association between amino acid and color variation are reported for asterisked sites at the bottom of each column. Sample names are comprised of genus and species initials, followed by ''W'' for wild type, ''M'' for melanic, or ''B'' for blanched, and a sequential numeral.

TABLE 4. F_{ST} for *Mc1r* and ND4 for *Holbrookia maculata* and *Aspidoscelis inornata*.

Species	*Mc1r* Sequence length (bp)	ND4 Sequence length (bp)	*Mc1r* Sample size	ND4 Sample size	*Mc1r* F_{ST}	ND4 F_{ST}
H. maculata	948	797	23	39	0.63067	0.94118
A. inornata	945	779	39	29	0.63901	0.09000

onine) at this position. This is reflected in the estimates of $4N_ec$ in Table 3: no recombination is observed within the blanched (derived) color morph, while considerable recombination is observed within the wild-type (ancestral) color morph. For the blanched alleles, nucleotide position 509 (which corresponds with amino acid site 170) is in statistically significant linkage disequilibrium with sites 261 ($P < 0.001$) and 549 ($P < 0.005$). No linkage disequilibrium is observed in the wild-type alleles.

DISCUSSION

We have documented interspecific and intraspecific patterns of molecular evolution at *Mc1r* and conducted association studies between *Mc1r* genotype and color phenotype in seven reptile species with distinct color morphs. Interspecific comparisons reveal similar patterns of functional constraint at *Mc1r* in reptiles and mammals. Intraspecific comparisons show abundant nucleotide variation at *Mc1r* across all seven squamates, and deviation from neutral expectations consistent with selection in *A. inornata*. Association studies reveal a strong association between genotype and color for *A. inornata*, and no evidence for population structure using mtDNA. We discuss each of these points in detail below.

Interspecific Molecular Evolution

At present we know very little about the genetic basis of pigmentation in reptiles, but there are several reasons for suspecting that at least some genes and pathways are conserved across vertebrates. Several pigmentation genes from mammals have clear orthologs in both birds and fish, and thus it is likely that they will also be present in reptiles (Kelsh et al. 1996; Camacho-Hubner et al. 2002; Logan et al. 2003). At the cellular level, similar cell types are found in mammals and reptiles and appear to have similar function; both groups possess pigment-producing cells derived from neural crest tissue that form pigment granules in specialized organelles (Bagnara and Hadley 1973). These overall similarities suggest that a candidate-gene approach using mammalian orthologs may be a useful way to study the genetic basis and evolution of pigmentation in reptiles. This study represents a first attempt to isolate pigmentation genes in reptiles and demonstrates the feasibility of this approach.

Patterns of molecular evolution suggest that reptilian and mammalian *Mc1r* display similar levels of constraint and therefore may share a conserved function. Sliding window analysis (Fig. 3) as well as measures of K_A/K_S for different regions of the receptor (Table 2) both reveal a close correspondence between rates of evolution across the *Mc1r* gene in mammals and reptiles. In both groups, transmembrane domains and intracellular regions, which are essential for signaling, are highly conserved. The first extracellular region is most variable (both in terms of amino acid substitutions and insertions/deletions), suggesting less functional constraint for this region. Overall rates of protein evolution relative to silent substitution rates (K_A/K_S) for *Mc1r* are nearly identical in mammals and reptiles, and we find no evidence for positive directional selection in this analysis.

It is important to recognize that the K_A/K_S values observed for *Mc1r*, which are very close to typical values for other genes and provide no evidence for positive selection, do not exclude the possibility that some changes at *Mc1r* are adaptive. For example, in mammals, overall K_A/K_S values are well below one in interspecific comparisons, yet there is strong evidence within one species of pocket mouse that several amino acid changes at *Mc1r* are responsible for an adaptive color difference (Nachman et al. 2003). Thus, while our data suggest that *Mc1r* does not display a long history of strong positive selection, this does not preclude the possibility that individual amino acid changes are responsible for adaptive color differences among populations within species or between closely related species.

TABLE 5. Patterns of ND4 variation for *Holbrookia maculata* and *Aspidoscelis inornata*. Gene length, number of individuals sampled, and number of segregating sites are presented. Nucleotide diversity (π) and nucleotide polymorphism (θ) are shown for all sites.

Species	Gene length (bp)	Sample size	Segregating sites	π (all sites)	θ (all sites)
H. maculata	797	39	29	0.01243	0.00772
Wild type		15	1	0.00017	0.00039
Blanched		24	8	0.00218	0.00269
Aspidoscelis inornata	779	29	11	0.00170	0.00360
Wild type		11	6	0.00159	0.00263
Blanched		18	6	0.00165	0.00224

Intraspecific Patterns of Variation

Mc1r displays substantial levels of nucleotide variation in all seven squamate species. Remarkably little is known of the amount or distribution of nucleotide variation within populations for nuclear genes outside of a few model organisms. In *Drosophila melanogaster*, the average level of silent-site heterozygosity is approximately 10^{-2}, and assuming per-site mutation rates of approximately 10^{-8}, suggests population sizes on the order of 10^5–10^6 (e.g., Moriyama and Powell 1996; Aquadro et al. 2001). In humans, the average level of nucleotide diversity ($\pi = 10^{-3}$) is at least one order of magnitude lower than in *D. melanogaster* and N_e has been repeatedly estimated at approximately 10^4, assuming mutation rates of 2×10^8. In both species, there is substantial variation in levels of heterozygosity among different genes. In humans, *Mc1r* has been studied in several different populations (Rana

et al. 1999; Harding et al. 2000), and the observed level of nucleotide diversity differs in different populations, ranging from 0.03% to 0.20%. Although nonmodel species have been less well studied, a recent estimate of nucleotide diversity for two passerine birds was approximately 0.25% (Primmer et al. 2002).

Here we provide some of the first estimates of nucleotide diversity for nuclear genes in reptiles and compare levels of variation among vertebrate species. Strikingly, the lowest levels of nucleotide variability observed at *Mc1r* in squamates are substantially higher than the highest levels observed in humans, mice, or birds. The average level of silent-site nucleotide diversity at *Mc1r* in squamates is approximately 1%, or 10 times the average value in humans. This could be due to pooling of genetically diverse populations, a high mutation rate, large long-term population size, or some combination of these factors. Comparing among the seven squamate species, higher levels of nucleotide diversity correspond to broader-scale geographic sampling (Table 1). However, high levels of nucleotide diversity are not entirely explained by population subdivision as π is still large when calculated for each color morph separately. Variation within color morphs does not appear to be structured by locality, although population structure in the recent past could have contributed to the high observed levels of variability. We are unaware of any direct measures of mutation rates at nuclear genes in reptiles, although a rough constancy of the molecular clock for several vertebrate genes (e.g., Kumar and Hedges 1998; Hedges and Kumar 2003) suggests that mutation rates may not be substantially different. If that is the case, the data presented here suggest that the squamates surveyed may have large long-term effective population sizes compared to mammals. This idea also receives support from the low levels of linkage disequilibrium observed in our data and correspondingly high levels of the population recombination parameter (Table 3). The possibility that squamates have large effective population sizes warrants further exploration as nuclear data from reptile populations become available.

In general, the distributions of allele frequencies at *Mc1r* in the different species are consistent with a neutral, equilibrium model of evolution. The distribution of Tajima's *D*-values in Table 3 is noteworthy for its similarity to the expected null distribution. Indeed, we have performed 21 tests and observed one that is significant, exactly the expected number of significant tests with an $\alpha = 0.05$. Given the small number of segregating sites and small sample sizes, caution is warranted in interpreting these results because there is little power to reject the null hypothesis using Tajima's *D* with small samples (Simonsen et al. 1995). Nonetheless, several of our samples are reasonably large (*S. undulatus* and *H. maculata*) and provide no evidence of an excess of either rare alleles or an excess of intermediate-frequency alleles, as may be expected following a population expansion or contraction, respectively. The ratio of nonsynonymous to synonymous changes within and between species is consistent with neutral expectations for *S. undulatus*, *H. maculata*, and *A. inornata* (McDonald and Kreitman 1991), as is the ratio of polymorphism to divergence between *Mc1r* and ND4 for *S. undulatus* and *H. maculata* (Hudson et al. 1987). The HKA test reveals a significantly reduced ratio of polymorphism to divergence at *Mc1r* compared to ND4 for *A. inornata*; however, this is not due solely to reduction of nucleotide variability among the blanched alleles. Thus, the only evidence for selection on *Mc1r* is seen in *A. inornata* and includes a significant skew in the frequency distribution of alleles. This observation is consistent with a functional role for *Mc1r* in the color differences in this species, as described below.

Association Studies

Association studies reveal significant associations between *Mc1r* genotype and color phenotype in three of the species studied: *S. undulatus*, *H. maculata*, and *A. inornata*. In *S. undulatus*, although the association is statistically significant, it is not particularly strong or consistent with a simple role for *Mc1r* (Fig. 4), and we did not pursue it further. In both *H. maculata* and *A. inornata*, the association is nearly perfect and highly significant (Fig. 4).

Results from association studies must be interpreted in comparison with patterns at putatively neutral loci to determine if the gene of interest is spuriously associated with the phenotype of interest as a consequence of population subdivision. We investigated this possibility by comparing F_{ST} for *Mc1r* with F_{ST} for the mitochondrial gene ND4 in each species. Although mtDNA may show nonneutral patterns of evolution (Nachman et al. 1996; Nielsen and Weinreich 1999), there is no a priori reason to suspect differential selection on ND4 among populations of different color morphs. For this data set, Tajima's *D*-values calculated for mtDNA within color morphs do not show significant departures from neutrality.

For *H. maculata*, F_{ST} at ND4 shows a highly subdivided population. Because mitochondria are maternally transmitted, estimates of population subdivision based on mtDNA can be elevated if females are more philopatric than males. Home range size for *H. maculata* males in New Mexico may be somewhat larger than for females (0.10 ha vs. 0.06 ha; Gennaro 1972); however, there is no evidence for dramatic male-biased dispersal in *H. maculata*. This suggests that *Mc1r* differences between color morphs in this species may be explained by population structure rather than a functional association between genotype and phenotype. Although *Mc1r* could still play a role in color variation in *H. maculata*, association studies on this geographic scale are unlikely to resolve the issue. Unfortunately, finer-scale sampling may not be possible because there are gaps in the distribution of this species across the landscape.

In contrast, F_{ST} calculated for ND4 in *A. inornata* reveals a nearly panmictic population, while F_{ST} for *Mc1r* shows significant differences among color morphs. At face value, this lends support to the hypothesis that *Mc1r* is involved in color variation in this species. However, mitochondrial and nuclear discordance can arise in several ways as has been documented in other species (e.g., Patton and Smith 1994). For example, it is possible that the two color morphs are well differentiated at most nuclear loci but that there has been recent introgression and replacement of one mitochondrial type into the other population. Sampling of additional nuclear genes will help resolve the nature of genetic differentiation between these color morphs.

In addition to the strong association between *Mc1r* and color and the absence of evident population structure in these samples of *A. inornata*, several other observations are consistent with recent selection on *Mc1r* and its role in *A. inornata* color variation. First, the amino acid substitution at position 170 is located in the fourth transmembrane domain. Although this is not a specific residue known to affect color in other species, transmembrane domains are generally highly conserved, and a number of transmembrane mutations are known to result in color change in other taxa (e.g., Ritland et al. 2001; Theron et al. 2001). Second, Tajima's *D* is significantly positive in this species, and this is the only significant deviation observed. Given the number of tests we conducted, this result should be treated with caution. However, the deviation is clearly due to the excess of intermediate-frequency sites when the two color morphs are pooled, including the single amino acid polymorphism and the sites linked to it (Fig. 4). Third, there is a striking difference in the level of linkage disequilibrium at *Mc1r* in the blanched color morph compared to the wild-type color morph (Table 3, Fig. 4). There is no significant linkage disequilibrium between any sites for the wild-type alleles; however, there is significant linkage disequilibrium between three nucleotide positions for the blanched sample. In fact, the isoleucine residue at site 170 is associated with a single haplotype (T-isoleucine-TC), while the threonine at this position is associated with at least four haplotypes. This pattern is consistent with recent selection driving the T-isoleucine-TC haplotype to near fixation in the blanched sample. Notably, the ancestral haplotype (T-threonine-TC) is present in the wild-type sample (e.g., CiW10, Fig. 4). The blanched phenotype is believed to be the derived phenotype, following the recent colonization of these lizards into the geographically restricted habitat with light-colored substrate.

Despite these observations, it is premature to claim that *Mc1r* is responsible for the observed color differences in *A. inornata*, for several reasons. First, we cannot rule out the possibility of significant population structure at most nuclear loci, with recent introgression of mtDNA across color morphs. Second, the association is not perfect. Five individuals are heterozygous at site 170 and present a blanched phenotype. Known *Mc1r* mutations in other animals can be either completely or partially dominant, and the dark phenotype is always dominant over the light phenotype. These five heterozygous *A. inornata* may be slightly darker than the other blanched animals, but this was not obvious at the time of collection. Spectrophotometric data will be used to resolve this question with increased sample sizes. Alternatively, pigmentation differences may be due to multiple genes, with *Mc1r* contributing only a portion to the observed phenotypic variation. Third, we cannot rule out the possibility that a linked gene is responsible for the color variation, although this seems unlikely given the low observed levels of linkage disequilibrium in general in these taxa. A strong test of the hypothesis that the I170T mutation at *Mc1r* is responsible for color differences in *A. inornata* will be to compare the activity of the wild-type and mutant receptor in an in vitro assay (e.g., Robbins et al. 1993; Vage et al. 1997).

Our results implicate *Mc1r* in color variation in only one of seven species surveyed. Although it is difficult to rule out the involvement of *Mc1r* in color variation in the other species, our data certainly raise the possibility that different genes may underlie similar phenotypes in different squamates. The data presented here suggest that a candidate-gene approach may continue to be useful for identifying the genetic basis of color variation in natural populations of reptiles.

Acknowledgments

We thank L. M. Chan, H. L Gibbs, and G. H. Pogson for their generosity in sharing tissue samples and L. M. Chan, J. M. Ray, and R. B. King for photograph contributions. We express our gratitude to J. Krenz for generous assistance throughout this project and to M. Saunders for consultation on data analysis. We thank the following for helpful comments on the manuscript: the Moritz, Wake, and Nachman labs; R. B. King; J. Hey; and two anonymous reviewers. Funding was provided by a National Science Foundation grant to MWN, a National Institutes of Health Postdoctoral Fellowship to HEH, and a National Science Foundation Doctoral Dissertation Improvement Grant to EBR.

Literature Cited

Aquadro, C. F., V. B. DuMont, and F. A. Reed. 2001. Genome-wide variation in the human and fruitfly: a comparison. Curr. Opin. Genet. Dev. 11:627–634.

Arevalo, E., S. K. Davis, and J. W. Sites. 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. Syst. Biol. 43:387–418.

Bagnara, J. T., and M. E. Hadley. 1973. Chromatophores and color change: the comparative physiology of animal pigmentation. Prentice-Hall, Inc., Englewood Cliffs, NJ.

Barsh, G. S. 1996. The genetics of pigmentation: from fancy genes to complex traits. Trends Genet. 12:299–305.

Bennett, D. C., and M. L. Lamoreux. 2003. The color loci of mice: a genetic century. Pigment Cell. Res. 16:333–344.

Bittner, T. D., R. B. King, and J. M. Kerfin. 2002. Effects of body size and melanism on the thermal biology of garter snakes (*Thamnophis sirtalis*). Copeia 2002:477–482.

Camacho-Hubner, A., C. Richard, and F. Beermann. 2002. Genomic structure and evolutionary conservation of the tyrosinase gene family from *Fugu*. Gene 285:59–68.

Cott, H. B. 1940. Adaptive coloration in animals. Methuen and Co., London.

Dixon, J. R. 1967. Aspects of the biology of the lizards of the White Sands, New Mexico. Contrib. Sci. 129:1–22.

Eizirik, E., N. Yuhki, W. E. Johnson, M. Menotti-Raymond, S. S. Hannah, and S. J. O'Brien. 2003. Molecular genetics and evolution of melanism in the cat family. Curr. Biol. 13:448–453.

Eskin, E., E. Halperin, and R. Karp. 2003. Large scale reconstruction of haplotypes from genotype data. Pp. 104–113 *in* Proceedings of the seventh Annual International Conference on Computational Molecular Biology, 10–13 April, Berlin.

Estes, R., and G. Pregill. 1998. Relationships of the lizard families. Stanford Univ. Press, Stanford, CA.

Fisher, R. A. 1934. Statistical methods for research workers. Oliver and Boyd, Edinburgh.

Gennaro, A. L. 1972. Home range and movements of *Holbrookia maculata maculata* in eastern New Mexico. Herpetologica 28: 165–168.

Hadley, M. E. 1997. Comparative aspects of non-mammalian pigmentation: regulation of pigment cells. Pigment Cell. Res. 10: 114.

Harding, R. M., E. Healy, A. J. Ray, N. S. Ellis, N. Flanagan, C. Todd, C. Dixon, A. Sajantila, I. J. Jackson, M. A. Birch-Machin, and J. L. Rees. 2000. Evidence for variable selective pressures at *MC1R*. Am. J. Hum. Genet. 66:1351–1361.

Hedges, S. B., and S. Kumar. 2003. Genomic clocks and evolutionary timescales. Trends Genet. 19:200–206.
Higgins, D. G. 1994. CLUSTAL V: Multiple alignment of DNA and protein sequences. Pp. 307–318 in A. M. Griffin and H. G. Griffin, eds. Methods in molecular biology: computer analysis of sequence data. II. Humana Press, Totowa NJ.
Hudson, R. R., M. Kreitman, and M. Aguade. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics 116: 153–160.
Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. Pp. 353–354 *in* H. N. Munro, ed. Mammalian protein metabolism. Academic Press, New York.
Kelsh, R. N., M. Brand, Y.-J. Jiang, C.-P. Heisenberg, S. Lin, P. Haffter, J. Odenthal, M. C. Mullins, F. J. M. Van Eeden, M. Furutani-Seiki, M. Granato, M. Hammerschmidt, D. A. Kane, R. M. Warga, D. Beuchle, L. Vogelsang, and C. Nuesslein-Volhard. 1996. Zebrafish pigmentation mutations and the processes of neural crest development. Development 123:369–389.
Kerns, J. A., M. Olivier, G. Lust, and G. S. Barsh. 2003. Exclusion of melanocortin-1 receptor (*Mc1r*) and *Agouti* as candidates for dominant black in dogs. J. Hered. 94:75–79.
Kettlewell, B. 1973. The evolution of melanism: the study of a recurring necessity. Clarendon Press, Oxford, U.K.
Kijas, J. M. H., R. Wales, A. Tornsten, P. Chardon, M. Moller, and L. Andersson. 1998. Melanocortin receptor 1 (*MC1R*) mutations and coat color in pigs. Genetics 150:1177–1185.
King, R. B. 2003. Mendelian inheritance of melanism in the garter snake *Thamnophis sirtalis*. Herpetologica 59:484–489.
Kumar, S., and S. B. Hedges. 1998. A molecular timescale for vertebrate evolution. Nature 392:917–920.
Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. BioInformatics 17:1244–1245.
Lawson, R., and R. B. King. 1996. Gene flow and melanism in Lake Erie garter snake populations. Biol. J. Linn. Soc. 59:1–19.
Lewis, T. H. 1949. Dark coloration in the reptiles of the Tularosa malpais, New Mexico. Copeia 3:181–184.
Li, W.-H. 1997. Molecular evolution. Sinauer, Sunderland, MA.
Li, W.-H., and L. A. Sadler. 1991. Low nucleotide diversity in man. Genetics 129:513–524.
Logan, D. W., R. J. Bryson-Richardson, K. E. Pagan, M. S. Taylor, P. D. Currie, and I. J. Jackson. 2003. The structure and evolution of the melanocortin and MCH receptors in fish and mammals. Genomics 81:184–191.
Lowe, C. H., and K. S. Norris. 1956. A subspecies of the lizard *Sceloporus undulatus* from the White Sands of New Mexico. Herpetologica 12:125–127.
MacDougall-Shackleton, E. A., L. Blanchard, and H. L. Gibbs. 2003. Unmelanized plumage patterns in old world leaf warblers do not correspond to sequence variation at the Melanocortin-1 receptor locus (*Mc1r*). Mol. Biol. Evol. 20:1675–1681.
Majerus, M. E. N. 1998. Melanism: evolution in action. Oxford Univ. Press, Oxford, U.K.
Marklund, L., M. Johansson-Moller, K. Sandberg, and L. Andersson. 1996. A missense mutation in the gene for melanocyte-stimulating hormone receptor (*MC1R*) is associated with the chestnut coat color in horses. Mammal. Genome 7:895–899.
McDonald, J. H., and M. Kreitman. 1991. Adaptive protein evolution at the ADH locus in *Drosophila*. Nature 351:652–654.
Moriyama, E. N., and J. R. Powell. 1996. Intraspecific nuclear DNA variation in *Drosophila*. Mol. Biol. Evol. 13:261–277.
Morrison, R. L., M. S. Rand, and S. K. Frost-Mason. 1995. Cellular basis of color differences in three morphs of the lizard *Sceloporus undulatus erythrocheilus*. Copeia 1995:397–408.
Mundy, N. I., and J. Kelly. 2003. Evolution of a pigmentation gene, the melanocortin-1 receptor, in primates. Am. J. Phys. Anthropol. 121:67–80.
Mundy, N. I., N. S. Badcock, T. Hart, K. Scribner, K. Janssen, and N. J. Nadeau. 2004. Conserved genetic basis of a quantitative plumage trait involved in mate choice. Science 303:1870–1873.
Murphy, W. J., E. Eizirik, S. J. O'Brien, O. Madsen, M. Scally, C. J. Douady, E. Teeling, O. A. Ryder, M. J. Stanhope, W. W. de Jong, and M. S. Springer. 2001. Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science 294: 2348–2351.
Nachman, M. W. 1997. Patterns of DNA variability at X-linked loci in *Mus domesticus*. Genetics 147:1303–1316.
Nachman, M. W., and S. L. Crowell. 2000. Estimate of the mutation rate per nucleotide in humans. Genetics 156:297–304.
Nachman, M. W., W. M. Brown, M. Stoneking, and C. F. Aquadro. 1996. Nonneutral mitochondrial DNA variation in humans and chimpanzees. Genetics 142:953–963.
Nachman, M. W., H. E. Hoekstra, and S. L. D'Agostino. 2003. The genetic basis of adaptive melanism in pocket mice. Proc. Natl. Acad. Sci. USA 100:5268–5273.
Nei, M., and T. Gojobori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3:418–426.
Nei, M., and W.-H. Li. 1979. Mathmatical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76:5269–5273.
Newton, J. M., A. L. Wilkie, L. He, S. A. Jordan, D. L. Metallinos, N. G. Holmes, I. J. Jackson, and G. S. Barsh. 2000. Melanocortin 1 receptor variation in the domestic dog. Mammal. Genome 11: 24–30.
Nielsen, R., and D. M. Weinreich. 1999. The age of nonsynonymous and synonymous mutations in animal mtDNA and implications for the mildly deleterious theory. Genetics. 153:497–506.
Norris, K. S. 1965. Color adaptation in desert reptiles and its thermal relationships. *In* W. W. Milstead, ed. Lizard ecology: a symposium. Univ. of Missouri Press, MO.
Norris, K. S., and C. H. Lowe. 1964. An analysis of background color-matching in amphibians and reptiles. Ecology 45:565–580.
Patton, J. L., and M. F. Smith. 1994. Paraphyly, polyphyly, and the nature of species boundaries in pocket gophers (genus *Thomomys*). Syst. Biol. 43:11–26.
Pearse, D. E., and G. H. Pogson. 2000. Parallel evolution of the melanic form of the California legless lizard, *Anniella pulchra*, inferred from mitochondrial DNA sequence variation. Evolution 54:1041–1046.
Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818.
Primmer, C. R., T. Borge, J. Lindell, and G.-P. Saetre. 2002. Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. Mol. Ecol. 11:603–612.
Przeworski, M., and J. D. Wall. 2001. Why is there so little linkage disequilibrium in humans? Genet. Res. Camb. 77:143–151.
Rana, B. K., D. Hewett-Emmett, L. Jin, B. H.-J. Chang, N. Sambuughin, M. Lin, S. Watkins, M. Bamshad, L. B. Jorde, M. Ramsay, T. Jenkins, and W.-H. Li. 1999. High polymorphism at the human melanocortin 1 receptor locus. Genetics 151: 1547–1557.
Reeder, T. W., C. J. Cole, and H. C. Dessauer. 2002. Phylogenetic relationships of whiptail lizards of the genus *Cnemidophorus* (Squamata: Teiidae): a test of monophyly, reevaluation of karyotypic evolution, and a review of hybrid origins. Am. Mus. Novit. 3365:1–61.
Ritland, K., C. Newton, and H. D. Marshall. 2001. Inheritance and population structure of the white-phased ''Kermode'' black bear. Curr. Biol. 11:1468–1472.
Robbins, L. S., J. H. Nadeau, K. R. Johnson, M. A. Kelly, L. Roselliirehfuss, E. Baack, K. G. Mountjoy, and R. D. Cone. 1993. Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. Cell 72: 827–834.
Rozas, J., and R. Rozas. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15:174–175.
Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin: a software for population genetics data analysis. Ver. 2.000. Genetics and Biometry Lab., Dept. of Anthropology, Univ. of Geneva.
Simonsen, K. L., G. A. Churchill, and C. F. Aquadro. 1995. Properties of statistical tests of neutrality for DNA polymorphism data. Genetics 141:413–429.

Smith, H. M. 1943. The White Sands earless lizard. Zool. Ser. Field Mus. Nat. Hist. 24:339–344.
Swofford, D. L. 1999. Phylogenetic analysis using parsimony (PAUP). Ver. 4.01b. Sinauer, Sunderland, MA.
Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.
Takeuchi, S., H. Suzuki, S. Hirose, Y. Masafumi, C. Sato, H. Yamamoto, and S. Takahashi. 1996a. Molecular cloning and sequence analysis of the chick melanocortin 1-receptor gene. Biochim. Biophys. Acta 1306:122–126.
Takeuchi, S., H. Suzuki, M. Yabuuchi, and S. Takahashi. 1996b. A possible involvement of melanocortin 1-receptor in regulating feather color pigmentation in the chicken. Biochim. Biophys. Acta 1308:164–168.
Theron, E., K. Hawkins, E. Bermingham, R. E. Ricklefs, and N. I. Mundy. 2001. The molecular basis of an avian plumage polymorphism in the wild: a melanocortin-1-receptor point mutation is perfectly associated with the melanic plumage morph of the bananaquit, *Coereba flaveola*. Curr. Biol. 11:550–557.
Vage, D. I., D. Lu, H. Klungland, S. Lien, S. Adalsteinsson, and R. D. Cone. 1997. A non-epistatic interaction of agouti and extension in the fox, *Vulpes vulpes*. Nat. Genet. 15:311–315.
Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. 7: 256–276.
Wright, S. 1951. The genetical structure of populations. Ann. Eugen. 15:323–354.
Yang, Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13:555–556.

Corresponding Editor: J. Hey