*Zoological Studies* **55:** 26 (2016)
doi:10.6620/ZS.2016.55-26

Zoological Studies

# Gene Structure and Sequence Polymorphism of the Coat Color Gene, *Mc1r*, in the Black-Bellied Vole (*Eothenomys melanogaster*)

Yung-Chih Lai[1,2,*], Shiao-Wei Huang[1], and Hon-Tsen Yu[1,3]

[1]*Department of Life Science, National Taiwan University, Taipei 10617, Taiwan*

[2]*Integrative Stem Cell Center, China Medical University Hospital, China Medical University, Taichung 40402, Taiwan. E-mail: d2205002kimo@yahoo.com.tw*

[3]*Genome and Systems Biology Degree Program, National Taiwan University and Academia Sinica, Taipei 10617, Taiwan. E-mail: ayu@ntu.edu.tw*

*(Received 14 March 2015; Accepted 4 March 2016)*

**Yung-Chih Lai, Shiao-Wei Huang, and Hon-Tsen Yu (2016)** Color polymorphism is a long-standing issue in ecological and evolutionary biology. The black-bellied vole (*Eothenomys melanogaster*) with complete melanic and brown forms provides an outstanding opportunity to study the genetic polymorphism underpinning color variation. Mutations in the coding region of melanocortin 1 receptor (*Mc1r*) have been shown to cause color variation in a wide range of species. However, the contribution to color variation produced by the *Mc1r* regulatory regions have rarely been studied in wild animals. To this end, the *Mc1r* promoter sequence in black-bellied voles was cloned and characterized in this study. At least 11 distinct transcription initiation sites were identified using 5'-RACE. Furthermore, a candidate core promoter region in the upstream GC-rich sequence was identified based on key transcription factor binding motifs. The black-bellied vole *Mc1r* coding region was conserved with that found in the house mouse and demonstrated characteristics that are consistent with the structure of a G-protein coupled receptor, *e.g.* seven transmembrane domains. We found a negative association between coat color variations and polymorphisms of either regulatory or coding regions. This implies that *Mc1r* might reflect geographic cline rather than adaptive evolution. Although we found a negative association, the extra information we obtained in the *Mc1r* promoter of the black-bellied vole can be beneficial to other studies in exploring the association between regulatory mutations and adaptive phenotypes in wild animals.

**Key words:** *Mc1r*, Coat color, Black-bellied vole, *Eothenomys melanogaster*, 5'RACE.

## BACKGROUND

The striking color polymorphism in animals has held a longstanding attraction to ecologists and evolutionists. Many color variations reflect an adaptive function in a wide range of animals, such as rodents (Burtt 1981; Caro 2005). The rodent genus *Eothenomys* is distributed in the Oriental Region, and contains eight species (Luo et al. 2004). Within this genus, black-bellied voles (*Eothenomys melanogaster*) are endemic to Taiwan, occur in the motane areas (1500-2500 m in altitude) and prefer understory habitats in the mist forests (Yu 1993, 1994). The voles show a wide variation in dorsal coat color from brown to black across its geographic range (Fig. 1). In general, the gradation of color from dark on the dorsum to light on the ventrum, named countershading, is widespread in animals to avoid detection by predators (Rowland 2009). Consequently, the dorso-ventral differentiation of coat colors often features a demarcation line in the flanks between the dark dorsum and light ventrum, especially in mammals. In striking contrast, the dark black-bellied vole form shows uniformly dark dorsal and ventral coat colors (complete melanism) without a clear demarcation between dorsal and ventral regions (Fig. 1a). This pigmentation pattern

*Correspondence: Present address: Integrative Stem Cell Center, China Medical University Hospital, Taichung 40402, Taiwan. E-mail: d92b41001@ntu.edu.tw

appears to be similar to that found in melanic laboratory mouse strains, *e.g.* C57BL/6J. The complete melanic form is rarely found in other species of mammals. Moreover, the dark black-bellied vole form is the dominant color morph in some localities while in other places none of the voles are dark, providing an excellent model in which to explore the sequence variation of coat color genes in different morphs.

To date, a large number of genes have been suggested to affect vertebrate pigmentation. For example, in mice, more than 100 loci and 800 phenotypic alleles have been identified (Bennett and Lamoreux 2003). Among them, melanocortin-1 receptor (*Mc1r*), a critical regulator in pigmentation synthesis, has been shown to contribute to color polymorphism in numerous species (Majerus and Mundy 2003). The gene encodes a G-protein coupled receptor with seven transmembrane domains that is specifically expressed in melanocytes (melanin-producing cells). In mammalian melanocytes, there are two basic types of melanin: (1) eumelanins, which are black or brown, and (2) pheomelanins, which are yellow or red. Typically, coat color variation among mammals results from the mixing of these two types of pigments in different proportions. The expression of the *Mc1r* gene is regulated by

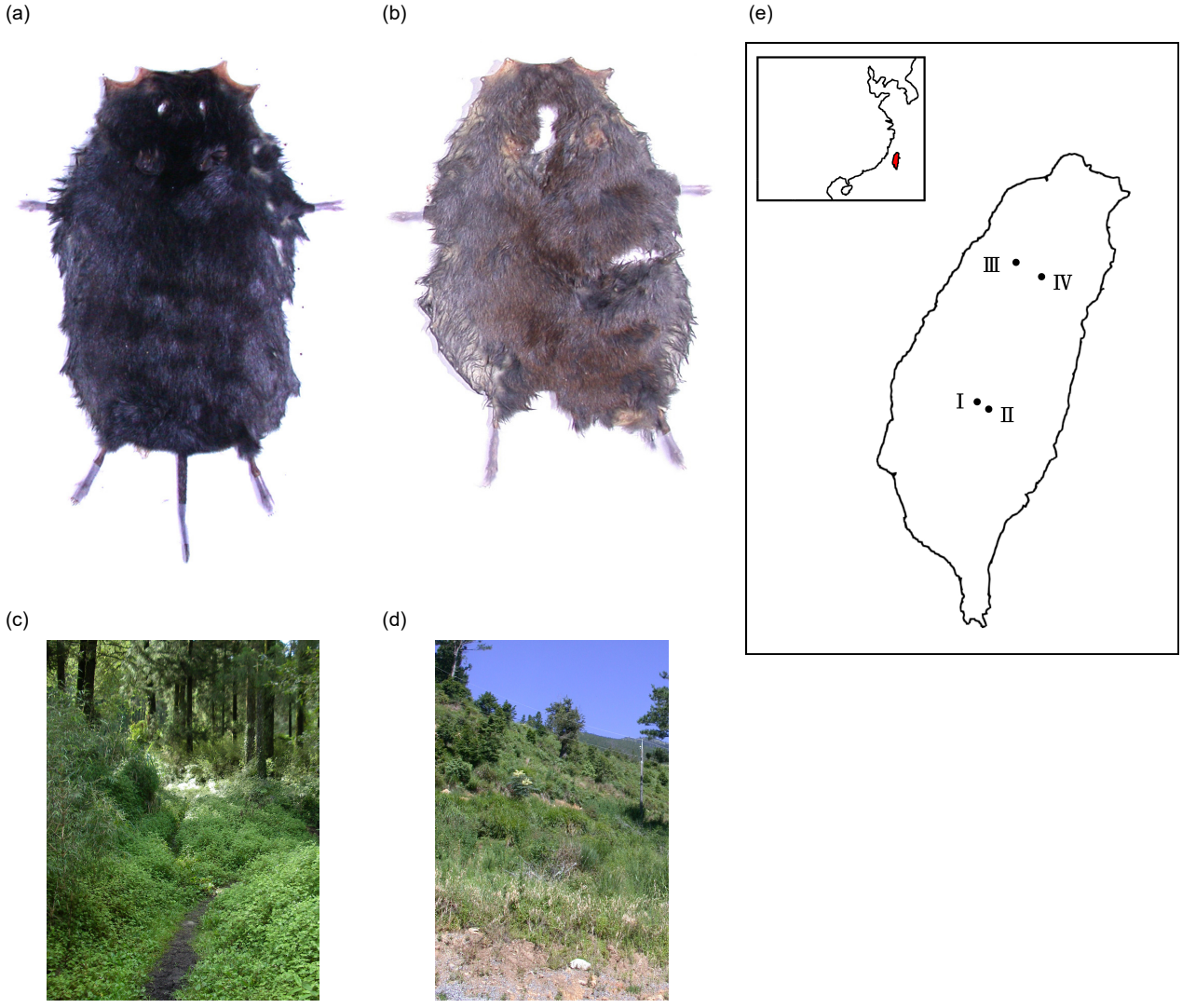


**Fig. 1.** Variation in coat color of black-bellied voles (*Eothenomys melanogaster*). (a) A black color form (Yu2045) sampled at Alishan, location I in the map (b) A brown color form (Yu2014) sampled at Wuling, location IV in the map (c) Alishan habitat (I in the map) with darker soil color (d) Wuling habitat (IV in the map) with lighter soil color (e) Map of Taiwan showing sites of sample collection: Alisan (I), Tataka (II), Guanwu (III), and Wuling (IV); the map inset shows the geographical location of Taiwan (red color) in East Asia.

the agonist, α-Melanocyte-stimulating hormone (α-MSH), and the antagonist, agouti protein. Binding of *Mc1r* by α-MSH results in an increase in the synthesis of eumelanins, whereas *Mc1r* bound by agouti protein will increase the production of pheomelanins (Hearing 2000). In addition, *Mc1r* coding mutations have been shown to be responsible for color differences in a wide range of species (Hoekstra et al. 2006), suggesting that it may be a good candidate gene to reveal the relationship of sequence mutation and color polymorphism in black-bellied voles.

One of the principal goals in evolutionary biology is to elucidate the genetic mechanisms underlying an adaptive trait. However, whether the evolution of an adaptive phenotype mainly results from mutations in the *cis*-regulatory elements (regulatory theory) or the coding regions remains controversial (Hoekstra and Coyne 2007; Wray 2007). Analysis of colors in vertebrates is currently the most promising approach to explore this question, because variation in adaptive morphological traits in animal colors dramatically demonstrates a clear relationship between phenotypes and *Mc1r* genotypes (review in Hoekstra 2006; Mundy 2005). However, so far efforts have been mostly focused on finding mutations in the *Mc1r* coding region, even though *Mc1r* cis-regulatory mutations may play a role in color variation (Rouzaud and Hearing 2005; Adachi et al. 2000). Since, the promoter region of the *Mc1r* gene has rarely been cloned in wild mammals, our analysis here can provide extra insights into the phenotype-genotype association underlying coat color variation. Therefore, the aim of the present study was to clone and characterize the black-bellied vole *Mc1r* gene promoter and coding regions. In addition, we also describe the color polymorphism of black-bellied voles in natural populations and its negative association with *Mc1r* polymorphism.

## MATERIALS AND METHODS

### Samples and measurements of coat color

A total of 37 black-bellied voles were collected from 4 locations in Taiwan (Table 1) that belong to two geographical regions in the Central Mountain Range (CMR) - the northern and southern regions, which are approximately 112km apart (Fig. 1e). Alishan (Al, N = 14) and Tataka (Ta, N = 8) are only 9 km apart and located in the southern part of the CMR (referred to as southern samples); Guanwu (Gu, N = 6) and Wuling (Wu, N = 9) are 23 km apart and located in the northern part of the CMR (referred to as northern samples). The average annual precipitation for Alishan, Tataka, Guanwu, and Wuling were 3798 mm, 3136 mm, 3598 mm, and 1473 mm, respectively. Voles were trapped from forested habitats of motane regions and brought to the lab for processing. Within 2-3 days after entering the lab, animals were sacrificed by cervical vertebral dislocation, and tissue samples were preserved from all individuals for DNA molecular analysis. Skin specimens were prepared as voucher specimens for coat color analysis. Because black-bellied voles exhibit coat color polymorphisms, we classified each specimen collected as melanic, intermediate, or brown (Fig. 1).

### Rapid amplification of 5' complementary DNA ends (5'RACE) and primer design

Poly $(A)^+$ RNAs from skin and tail tissues were extracted according to the Trizol protocol (Invitrogen). The 5'-UTR sequences and the transcription initiation sites of the *Mc1r* gene were determined using a 5' RACE kit (Roche). Initially, we amplified and obtained the exon sequences of the *Mc1r* gene using a pair of primers (mMc1r-1-for and mMc1r-2-rev) specific for the house mouse (*Mus musculus*) (Wada et al. 2005). Based on these exon sequences, we designed three gene-specific primers for 5'RACE: EmMc1r_SP1 (5'-CCAGACAGCAGATGAAGCAA-3') for synthesis of first strand cDNA using AMV reverse transcriptase, EmMc1r_SP2 (5'-GAGGCCATCTGGGATAGACA-3') for first PCR to amplify the adaptor-ligated cDNA, and EmMc1r_SP3 (5'-GGTAGCCAGTCCAAGGTGAG-3') for a second PCR reaction. Using these primers, we obtained 5'UTR sequences. To obtain the sequences encompassing 5'UTR and its entire exon, an additional primer pair, EmMc1r_F1 (5'-CTACGGGGGCTTTGAACAC-3') and EmMc1r_R1 (5'-TGGTCCCAGGCAGTTTGTG-3'), was designed to amplify a 1714 bp fragment in the voles.

### Genomic DNA extraction, PCR cloning, and sequencing

Vole genomic DNA was extracted from frozen muscle, liver or kidney tissues using commercial kits (Lamda). PCR amplification was performed

in thermal cyclers (ABI 2720 or Bio-Rad PTC 200) in 100 μl volume containing 1.0 unit of Phusion DNA polymerase (Finnzymes), 20 μl 5× Phusion HF buffer, 0.2 mM for each dNTP and 1.5 mM $MgCl_2$. Thermocycler conditions were as follows: (1) an initial denaturing step of 98°C for 30 s; (2) 30 cycles of the following: 10 s at 98°C, 10 s at 58-61°C, 2 min at 72°C; and (3) a 10 min

**Table 1.** Nucleotide polymorphisms at the *Mc1r* gene

| population | coat color | specimen | 88 | 89 | 374 | 750 | 913 | 1104 | 1279 | 1323 | 1518 | 1524 | 1607 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Guanwu (Gu) 33% 3598 mm | melanic | Yu2086 | C | C | G | G | C | T | A | C | G | G | A |
| | melanic | Yu2088 | . | . | . | . | . | . | . | . | G/A | . | . |
| | intermediate | Yu2091 | . | . | . | . | . | . | . | . | . | . | . |
| | intermediate | Yu2092 | . | . | . | . | . | . | . | . | . | . | . |
| | intermediate | Yu2093 | . | . | . | . | . | . | . | . | . | . | . |
| | intermediate | Yu2087 | . | . | . | . | C/G | . | . | . | . | . | . |
| Wuling (Wu) 0% 1473 mm | brown | Yu2036 | . | . | . | . | . | . | . | . | . | . | . |
| | brown | Yu2041 | . | . | . | . | . | . | . | T | . | . | . |
| | brown | Yu2037 | . | . | . | A | . | . | . | T | . | . | . |
| | brown | Yu2014 | . | . | . | G/A | . | . | . | T | . | . | . |
| | brown | Yu2040 | . | . | . | G/A | . | . | . | T | . | . | . |
| | brown | Yu2039 | . | . | . | . | . | . | . | C/T | . | . | . |
| | brown | Yu2044 | . | . | . | . | . | . | . | C/T | . | . | . |
| | brown | Yu2038 | . | . | . | G/A | . | . | . | C/T | . | . | . |
| | brown | Yu2042 | . | . | . | G/A | . | . | . | C/T | . | . | . |
| Tataka (Ta) 50% 3136 mm | melanic | Yu2022 | A | G | . | . | . | C | G | . | . | . | G |
| | melanic | Yu2072 | A | G | . | . | . | C | G | . | . | . | G |
| | melanic | Yu2071 | . | . | A/G | . | . | C | G | . | . | G/A | G |
| | melanic | Yu2064 | C/A | C/G | . | . | . | C | G | . | . | G/A | G |
| | intermediate | Yu2076 | . | . | . | . | . | C | G | . | . | A | G |
| | brown | Yu2082 | . | . | A | . | . | C | G | . | . | . | G |
| | brown | Yu2020 | A | G | . | . | . | C | G | . | . | . | G |
| | brown | Yu2066 | C/A | C/G | . | . | . | C | G | . | . | G/A | G |
| Alishan (Al) 64% 3798 mm | melanic | Yu2045 | . | . | A | . | . | C | G | . | . | . | G |
| | melanic | Yu2046 | . | . | A | . | . | C | G | . | . | . | G |
| | melanic | Yu2050 | . | . | A | . | . | C | G | . | . | . | G |
| | melanic | Yu2051 | . | . | A | . | . | C | G | . | . | . | G |
| | melanic | Yu2053 | . | . | A | . | . | C | G | . | . | . | G |
| | melanic | Yu2058 | . | . | A | . | . | C | G | . | . | . | G |
| | melanic | Yu2059 | . | . | A | . | . | C | G | . | . | . | G |
| | melanic | Yu2049 | . | . | . | . | . | C | G | . | . | . | G |
| | melanic | Yu2055 | . | . | A/G | . | . | C | G | . | . | . | G |
| | intermediate | Yu2056 | . | . | A | . | . | C | G | . | . | . | G |
| | intermediate | Yu2057 | . | . | A | . | . | C | G | . | . | . | G |
| | brown | Yu2048 | . | . | A | . | . | C | G | . | . | . | G |
| | brown | Yu2052 | . | . | A | . | . | C | G | . | . | . | G |
| | brown | Yu2054 | C/A | C/G | A/G | . | . | C | G | . | . | . | G |
| Gene structure | | | 5'UTR | | | Exon | | | | | | | 3'UTR |
| Amino acid | | | | | | | Arg Gly | | Ala Thr | | | | |
| Amino acid position | | | | | | 54 | 109 | 172 | 231 | 245 | 310 | 312 | |

Numbers in the top row correspond to nucleotide positions. Dots represent identity with respect to the first sequence. For heterozygous sites, the genotype is indicated. Gene structure, amino acid positions, and nonsynoymous changes are indicated at the bottom of the corresponding column. % indicates the proportion of the melanic form. The amount of annual precipitation is given in mm.

extension step at 72°C. The 1714 bp PCR product was purified using a QIAquick gel extraction kit (Qiagen). After A-tailing with 10mM dATP at 72°C for 30 min, TA cloning was performed using the pGEM-T vector system (Promega) according to the manufacturer's instructions. At least five clones were sequenced for each vole.

### Data analysis

Sequences were edited and aligned manually by Sequencher, v.4.1.4 (Genecodes). All *Mc1r* sequences generated in this study have been deposited in GenBank (accession numbers GU001524-GU001573). To define the boundaries of the coding region, *E. melanogaster Mc1r* sequences were aligned with those of the house mouse (*M. musculus*) and pocket mouse (*Chaetodipus intermedius*) by ClustalW. In addition, the transmembrane domains were independently predicted by four web tools: HMMTOP 2.0 (http://www.enzim.hu/hmmtop/), SOSUI (http://bp.nuap.nagoya-u.ac.jp/sosui/), TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0/), and TMpred (http://www.ch.embnet.org/software/TMPRED_form.html). Potential transcription factor binding sites were predicted by the Transcription Element Search System (TESS; http://www.cbil.upenn.edu/tess). A parsimony haplotype network based on *Mc1r* sequences from all alleles was constructed by TCS 1.21. Evidence for selection at *Mc1r* was based on Tajima's D statistic using the program DnaSP 4.10.9.

## RESULTS

### Color polymorphism

The coat colors on black-bellied voles are variable between individuals and populations (Table 1). For example, in Alishan, we found three coat colors and the melanic form was dominant. In Wuling, we only found the brown form. In addition, the most melanic population (Alishan) lives in the most humid environment (3798 mm annual precipitation), while the least melanic population (Wuling) inhabits the driest habitat (1473 mm annual precipitation). The relationship between coat color and annual precipitation corresponds to Gloger's rule, which states that individual endothermic animals are darker in humid habitats than those in drier environments (Gloger 1833; Zink and Remsen 1986). It implies that coat colors may, at least in part, reflect environmental variation.

### Characteristics of the 5'UTR and coding regions of the *Mc1r* gene

A region of 1714 bp of the *Mc1r* gene, including the entire coding region (954 bp), 5'UTR (588 bp), and 3'UTR (172 bp), was characterized for the black-bellied vole. The 5'UTR sequence shows 71.7% nucleotide identity with that of the house mouse (Adachi et al. 2000). Totally, eleven transcription initiation sites identified by the 5'-RACE were stretched over an ~500 bp GC-rich (57.38%) region upstream of the *Mc1r* coding region (Fig. 2, positions at 1, 199, 256, 271, 276, 293, 302, 305, 318, 378, and 474). No apparent transcriptional initiation site clusters were noted, in contrast to those reported in humans (Moro et al. 1999). We predict three and one putative transcription factor binding sites for SP-1 and AP-2, respectively, using TESS. Moreover, five CANNTG motifs, which have been experimentally validated in the house mouse *Mc1r* promoter were also discovered in black-bellied voles (Adachi et al. 2000). Neither the TATA nor CAAT boxes were found in this proximal region.

From start codon (ATG) to stop codon (TGA), the nucleotide positions 589-1542 (954 bp) corresponds to the gene exon sequence in the house mouse and pocket mouse. The exon sequence in the black-bellied vole shows 89.0% and 82.8% nucleotide identity to that of the house mouse (Wada et al. 2005) and the pocket mouse (Nachman et al. 2003), respectively. All four bioinformatics tools independently predict that the *Mc1r* coding region we sequenced contains seven putative transmembrane domains, a key feature of the G-protein coupled receptor.

### *Mc1r* sequence polymorphism

Eleven polymorphic sites were detected, three in the 5'UTR, seven in the exon, and one in the 3'UTR (Table 1). Within the exon, five of the substitutions were synonymous and two were nonsynonymous. The two nonsynonymous mutations (amino acid positions 109 and 231) occurred in the second extracellular region and third intracellular region, respectively. None of the eleven substitutions showed a significant association with coat color variation (Table 1), *i.e.*, we did not observe variants that were unique to melanic, intermediate, or brown coat color in

the four populations. This suggests that *Mc1r* is not a principal determinant underlying coat color polymorphism in the black-bellied vole, at least in the four populations we sampled.

### Phylogenetic relationships and testing for selection

Sequence variation at *Mc1r* was largely consistent with the phylogeographic relationship inferred from mitochondrial cytochrome B sequence variation (Chang 2007). Three polymorphic sites (Table 1: 1104, 1279, and 1607) support a split between the northern and southern vole populations, which correspond to the basal split inferred from mtDNA data. In addition, the genealogy of *Mc1r* haplotypes (Fig. 3) is similar in topology to those generated from mitochondrial cytochrome B for black-bellied voles (Chang 2007). Only two nonsynonymous substitutions were observed in the *Mc1r* sequence, one in a heterozygote (Table 1: specimen: Yu2087, nucleotide position 913, amino acid position 109), and the other separating the northern and southern vole populations (nucleotide position 1279, amino acid position 231). Although Tajima's D statistic was positive in our analysis, it did not significantly differ from zero (0.87142, $P > 0.10$), suggesting that the coat color related gene, *Mc1r*, is not subject to selection.

## DISCUSSION

Gene expression in most eukaryotes is under the control of regulatory elements, especially

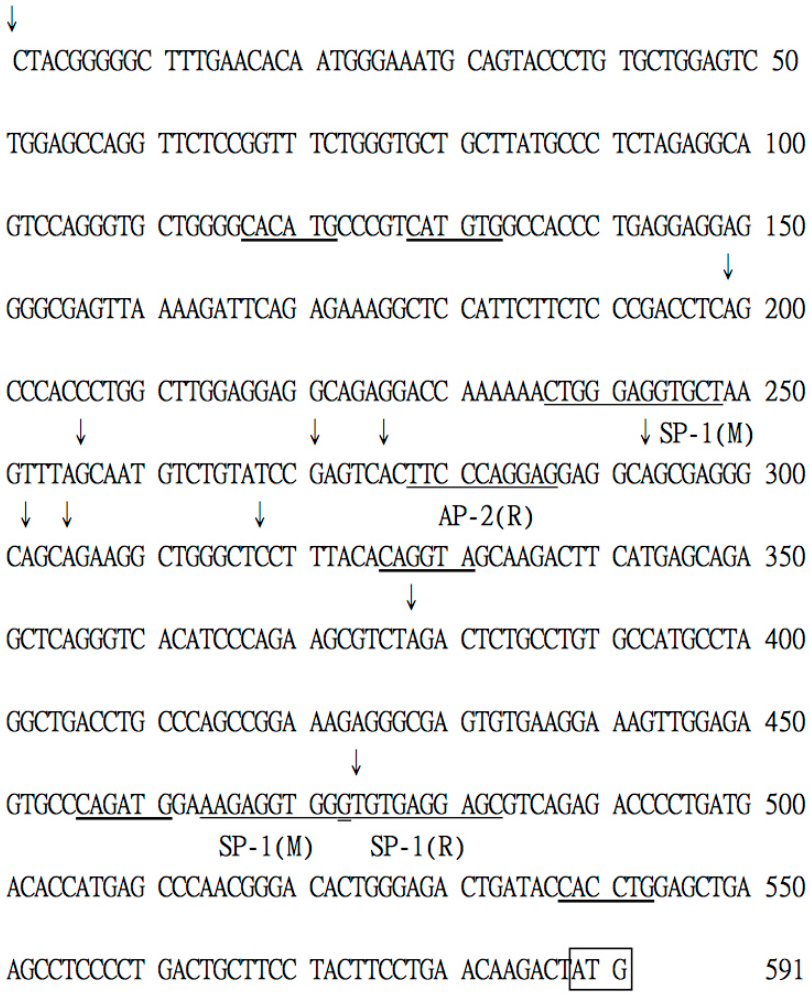


**Fig. 2.** The 5' UTR nucleotide sequence of the *Mc1r* gene. The maximum length of the 5'UTR is 588 bp. The first ATG codon is boxed. Transcriptional initiation sites are indicated by arrows. Consensus binding sites for SP-1 and AP-2 transcription factors are underlined. The parentheses refer to mouse (M) or rat (R). The CANNTG motifs based on reports from mouse experiments are heavily underlined.

promoter regions, which are usually located upstream of transcription initiation sites. We cloned and characterized the promoter region of *Mc1r* in black-bellied voles. Our results show that the ~ 500 bp GC-rich region upstream of the *Mc1r* coding region seems to be the main promoter. Not only are all of the detected transcription initiation sites located in this region, but also the features within this region (including GC-rich, TATA-less, and SP-1 transcription factor binding sites) are consistent with a G-protein-coupled receptor. In addition, the first two CANNTG motifs in the black-bellied vole are completely identical in sequence to the promoter region in the house mouse, as recognized by the transcription factor of the basic-helix-loop-leucine zipper (bHLH-Zip) protein family (Adachi et al. 2000). Therefore, we suspect that the two CANNTG motifs upstream from most of the transcription initiation sites are important candidate core promoters.

In contrast to most studies focused only on the coding region, we found non-association with color variations both in cis-regulatory and coding regions. Similar to our findings in the black-bellied vole, *Mc1r* mutations did not associate with color variations in other animals with a wide spectrum of color morphs (Goncalves et al. 2012; Hosoda et al. 2005; McRobie et al. 2014). It is likely that other genes, *e.g. Agouti* (Linnen et al. 2009) or *Tyrp1* (Schmutz et al. 2002), may be responsible for coat color variation for each of these species. In the future, more advanced approaches will be adopted to explore the molecular evolution that underlies these adaptive traits (Supplement Fig. 1). For example, Hoekstra et al. (2006) showed that a single gene can influence several traits. In contrast, a single phenotype also can be regulated by multiple genes (Steiner et al. 2007). Now, the strength of population genomics has grown and will enable us to explore the genetic evolution of color genes at the whole genome scale (*e.g.* Poelstra et al. 2014). Taking advantage of next-generation sequencing (NGS), scientists are able to explore the whole genome for various adaptive characteristics. Alternatively, causative enhancers of *Mc1r* could be too far away from the transcription initiation sites to be identified. At this time, we cannot exclude that *Mc1r* mutations may underlie coat color variation in populations we did not screen. For example, among pocket mice an association between coat color variation and *Mc1r* coding substitution has been found in only one of four natural populations (Hoekstra and Nachman 2003). Consequently, our data do not completely rule out the potential role of *Mc1r* in this context. However, the consistent patterns between *Mc1r*-based and cytochrome B-based phylogenies and the fact that no significant selection was detected by Tajima's D test suggest that the differences at the *Mc1r* locus, including regulative elements and the coding region, might reflect a geographic cline, as indicated by TCA analysis, rather than adaptive evolution. This means that genetic drift, rather than natural selection, could be the primary force driving *Mc1r* sequence polymorphism.

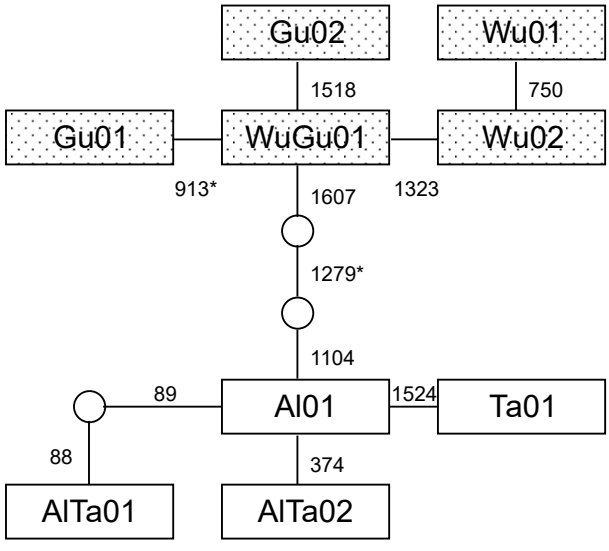


**Fig. 3.** Statistical parsimony network based on *Mc1r* haplotypes. Each line between haplotypes indicates one substitution, with the positions of the nucleotide substitutions noted above. * indicates a nonsyonymous substitution. Al, Ta, Gu, and Wu: alleles from the Alishan, Tataka, Guanwu, and Wuling populations, respectively. AlTa: alleles from the Alishan and Tataka populations. WuGu: alleles from the Guanwu and Wuling populations. The stippled background indicates the two northern populations, and the white background indicates the two southern populations.

Despite the negative association of *Mc1r* and pigmentation in the black-bellied vole, feature and promoter region sequence similarities with the house mouse provide may aid in obtaining regulatory sequence information for the *Mc1r* gene in other species. This could also be applied to previous studies that found no association between colors and *Mc1r* coding mutations (Cheviron et al. 2006; Herczeg et al. 2010; MacDougall-Shackleton et al. 2003; Skoglund and Hoglund 2010; Wlasiuk and Nachman 2007). Based on our results, one can re-evaluate the phenotype-genotype association in the regulatory region, especially for core promoters. In addition, the experiments we designed to determine whether

cis-regulatory mutations and/or coding mutations play a crucial role in adaptive traits have provided a promising avenue to confirm regulatory evolution. In the future, it might require more time to obtain sufficient evidence to confirm the hypothesis of regulatory evolution (Carroll 2005).

Finally, coat color variation in black-bellied voles may have some adaptive significance as we observed that more melanic forms were from wetter localities (see Table 1). This is in keeping with a long-recognized phenomena concerning mammalian coat color known as Gloger's Rule in which animals in more humid environments tend to be darker than their counterparts in drier habitats (Gloger 1833; Zink and Remsen 1986). Two interpretations have been given for Gloger's rule. One is that melanic forms are more cryptic against a darker background caused by higher vegetation density (*i.e.*, shade) and darker soil color (*i.e.*, saturated with moisture) as a result of increased levels of precipitation (Lai et al. 2008). An alternative explanation is that bacteria are more abundant and active in humid environments, and the melanic form with more eumelanin pigments resists bacterial degradation better than brown forms caused by pheomelanin pigments (Burtt and Ichida 2004). Both hypotheses require advanced experiments to confirm or deny their predictions in the future.

## CONCLUSIONS

The key pigment gene, *Mc1r*, has provided a vital proxy to explore molecular evolution based on its protein coding region for various wild animals. Further, we used the black-bellied vole as an example to clone and characterize its regulatory region. The multiple transcription initiation sites and candidate transcription factor binding sites in its promoter provide regulatory potential for animal color patterns. Both the promoter and coding regions we sequenced are consistent with the biogeographical distribution, suggesting other candidate color genes could play a role for the populations we investigated. Whether the adaptive phenotypes are mainly caused by regulatory or coding mutations can be clarified by surveying their homologs in more populations and species.

**Acknowledgments:** Special thanks are due to Y.-C. Chang and Dr. L.-K. Lin for providing valuable samples and invaluable assistance in the field. We also thank S.-H. Kuan, S.-Y. Jiang, S.-M. Yang, Dr. G.-C. Ma, Dr. H.-P. Lu, Dr. C.-M. Liao, and S.-T. Wang for assistance in the field. We wish to thank T.-L. Chou, Dr. W.-J. Liu, and Dr. J.-H. Lu for providing 5'RACE technical assistance. We would like to thank the Consulting Center for Statistics and Bioinformatics, NTU for helpful advice. Thanks very much for revising and improving an earlier draft of the manuscript by Dr. R. B. Widelitz, Dr. S. Y. V. Liu, Dr. M.-S. Shiao, Dr. B. R. Stein and T. Cheng. This study was supported by the Foundation for the Advancement of Outstanding Scholarship, and National Science Council (NSC 94-2311-B-002-017, 95-2311-B-002-003, and 96-2311-B-002-001 to H.-T. Yu).

## REFERENCES

Adachi S, Morii E, Kim DK, Ogihara H, Jippo T, Ito A, Lee YM, Kitamura Y. 2000. Involvement of mi-transcription factor in expression of α-melanocyte-stimulating hormone receptor in cultured mast cells of mice. J Immunol **164(2):**855-860.

Bennett DC, Lamoreux ML. 2003. The color loci of mice - a genetic century. Pigm Cell Res **16(4):**333-344.

Burtt EH. 1981. The adaptiveness of animal colors. Bioscience **31(10):**723-729.

Burtt EH, Ichida JM. 2004. Gloger's rule, feather-degrading bacteria, and color variation among song sparrows. Condor **106(3):**681-686.

Caro T. 2005. The adaptive significance of coloration in mammals. Bioscience **55(2):**125-136.

Carroll SB. 2005. Evolution at two levels: On genes and form. Plos Biology **3(7):**1159-1166.

Chang YC. 2007. Population divergence of Père David's red-backed vole (*Eothenomys melanogaster*) in Taiwan. Master thesis, Tunghai University, Taichung, Taiwan.

Cheviron ZA, Hackett SJ, Brumfield RT. 2006. Sequence variation in the coding region of the melanocortin-1 receptor gene (*MC1R*) is not associated with plumage variation in the blue-crowned manakin (*Lepidothrix coronata*). Proc Biol Sci **273(1594):**1613-1618. doi:10.1098/rspb.2006.3499.

Gloger C. 1833. Das abändern der vögel durch einfluss des klimas. A. Schulz, Breslau, German

Goncalves GL, Hoekstra HE, Freitas TRO. 2012. Striking coat colour variation in tuco-tucos (Rodentia: Ctenomyidae): a role for the melanocortin-1 receptor? Biological Journal of the Linnean Society **105(3):**665-680.

Hearing VJ. 2000. The melanosome: The perfect model for cellular responses to the environment. Pigm Cell Res **13:**23-34.

Herczeg G, Matsuba C, Merila J. 2010. Sequence variation in the melanocortin-1 receptor gene (*Mc1r*) does not explain variation in the degree of melanism in a widespread amphibian. Annales Zoologici Fennici **47(1):**37-45.

Hoekstra HE. 2006. Genetics, development and evolution of adaptive pigmentation in vertebrates. Heredity **97(3):**222-234.

Hoekstra HE, Coyne JA. 2007. The locus of evolution: Evo devo and the genetics of adaptation. Evolution **61(5):**995-1016.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. Science **313(5783):**101-104.

Hoekstra HE, Nachman MW. 2003. Different genes underlie adaptive melanism in different populations of rock pocket mice. Mol Ecol **12(5):**1185-1194.

Hosoda T, Sato JJ, Shimada T, Campbell KL, Suzuki H. 2005. Independent nonframeshift deletions in the *MC1R* gene are not associated with melanistic coat coloration in three mustelid lineages. J Hered **96(5):**607-613.

Lai YC, Shiroishi T, Moriwaki K, Motokawa M, Yu HT. 2008. Variation of coat color in house mice throughout Asia. J Zool **274(3):**270-276.

Linnen CR, Kingsley EP, Jensen JD, Hoekstra HE. 2009. On the origin and spread of an adaptive allele in deer mice. Science **325(5944):**1095-1098.

Luo J, Yang DM, Suzuki H, Wang YX, Chen WJ, Campbell KL, Zhang YP. 2004. Molecular phylogeny and biogeography of oriental voles: genus *Eothenomys* (Muridae, Mammalia). Mol Phylogenet Evol **33(2):**349-362.

MacDougall-Shackleton EA, Blanchard L, Gibbs HL. 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(10):**1675-1681.

Majerus MEN, Mundy NI. 2003. Mammalian melanism: natural selection in black and white. Trends Genet **19(11):**585-588.

McRobie HR, King LM, Fanutti C, Coussons PJ, Moncrief ND, Thomas APM. 2014. Melanocortin 1 receptor (*MC1R*) gene sequence variation and melanism in the gray (*Sciurus carolinensis*), fox (*Sciurus niger*), and red (*Sciurus vulgaris*) squirrel. J Hered **105(3):**423-428.

Moro O, Ideta R, Ifuku O. 1999. Characterization of the promoter region of the human melanocortin-l receptor (MC1R) gene. Biochem Biophys Res Commun **262(2):**452-460.

Mundy NI. 2005. A window on the genetics of evolution: *MC1R* and plumage colouration in birds. Proc R Soc B-Biol Sci **272(1573):**1633-1640.

Nachman MW, Hoekstra HE, D'Agostino SL. 2003. The genetic basis of adaptive melanism in pocket mice. Proc Natl Acad Sci U S A **100(9):**5268-5273.

Poelstra JW, Vijay N, Bossu CM, Lantz H, Ryll B, Muller I, Baglione V, Unneberg P, Wikelski M, Grabherr MG, Wolf JB. 2014. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. Science **344(6190):**1410-1414.

Rouzaud F, Hearing VJ. 2005. Regulatory elements of the melanocortin 1 receptor. Peptides **26(10):**1858-1870.

Rowland HM. 2009. From Abbott Thayer to the present day: what have we learned about the function of countershading? Philos Trans R Soc B Biol Sci **364(1516):**519-527. doi:10.1098/rstb.2008.0261.

Schmutz SM, Berryere TG, Goldfinch AD. 2002. *TYRP1* and *MC1R* genotypes and their effects on coat color in dogs. Mamm Genome **13(7):**380-387.

Skoglund P, Hoglund J. 2010. Sequence polymorphism in candidate genes for differences in winter plumage between Scottish and Scandinavian willow grouse (*Lagopus lagopus*). PLoS ONE **5(4):**e10334. doi:10.1371/journal.pone.0010334.

Steiner CC, Weber JN, Hoekstra HE. 2007. Adaptive variation in beach mice produced by two interacting pigmentation genes. Plos Biology **5(9):**1880-1889.

Wada A, Kunieda T, Nishimura M, Kakizoe-Ishida Y, Watanabe N, Ohkawa K, Tsudzuki M. 2005. A nucleotide substitution responsible for the tawny coat color mutation carried by the MSKR inbred strain of mice. J Hered **96(2):**145-149.

Wlasiuk G, Nachman MW. 2007. The genetics of adaptive coat color in gophers: coding variation at *Mc1r* is not responsible for dorsal color differences. J Hered **98(6):**567-574.

Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet **8(3):**206-216.

Yu HT. 1993. Natural-history of small mammals of subtropical montane areas in central Taiwan. J Zool **231(3):**403-422.

Yu HT. 1994. Distribution and abundance of small mammals along a subtropical elevational gradient in central Taiwan. J Zool **234(4):**577-600.

Zink RM, Remsen JVJ. 1986. Evolutionary processes and patterns of geographic variation in birds. Curr Ornithol **4:**1-69.

**Supplement Fig. 1.** Progression of approaches to analyze phenotype-genotype association for adaptive traits. (download)