Genetic architecture and evolution of color variation in American black bears

Highlights
- The cinnamon morph of American black bears is due to a missense mutation in \textit{TYRP1}
- Bear \textit{TYRP1} variants are hypomorphs that impair protein localization to melanosomes
- The cinnamon variant arose 9,360 years ago and provides an adaptive advantage

Authors
Emily E. Puckett, Isis S. Davis, Dawn C. Harper, ..., Christopher B. Kaelin, Michael S. Marks, Gregory S. Barsh

Correspondence
emily.puckett@memphis.edu

In brief
Puckett et al. show that a single base-paired variant in \textit{TYRP1} results in the iconic cinnamon morph of the American black bear. They demonstrate that \textit{TYRP1} variants in black and grizzly bears are loss-of-function alleles and impair protein localization to melanosomes. The black bear allele arose 1,440 generations ago in the southwestern range.
SUMMARY

Color variation is a frequent evolutionary substrate for camouflage in small mammals, but the underlying genetics and evolutionary forces that drive color variation in natural populations of large mammals are mostly unexplained. The American black bear, *Ursus americanus* (*U. americanus*), exhibits a range of colors including the cinnamon morph, which has a similar color to the brown bear, *U. arctos*, and is found at high frequency in the American southwest. Reflectance and chemical melanin measurements showed little distinction between *U. arctos* and cinnamon *U. americanus* individuals. We used a genome-wide association for hair color as a quantitative trait in 151 *U. americanus* individuals and identified a single major locus (*p < 10^{-13}*). Additional genomic and functional studies identified a missense alteration (R153C) in Tyrosinase-related protein 1 (*TYRP1*) that likely affects binding of the zinc cofactor, impairs protein localization, and results in decreased pigment production. Population genetic analyses and demographic modeling indicated that the R153C variant arose 9.36 kya in a southwestern population where it likely provided a selective advantage, spreading both northwards and eastwards by gene flow. A different *TYRP1* allele, R114C, contributes to the characteristic brown color of *U. arctos* but is not fixed across the range.

INTRODUCTION

Variation in animal color is a long-standing system with which to investigate gene action and evolution. In mammals, nearly all pigment is produced by melanocytes via polymerization of dihydroxyindole derivatives to yield insoluble melamins that are transferred to overlying keratinocytes in skin and/or hair. Genetic variation in melanin biosynthesis or transfer gives rise to hair, eye, and skin color differences and can underlie positive selection.\(^1,2\) Mutations of melanin biosynthesis underlie a number of conditions associated with impaired fitness or a disease (e.g., human albinism), or unusual color morphs of large mammals that are specifically targeted for trophy hunting.\(^3\)

Identifying and understanding the genetic causes and consequences of mammalian color variation is of special interest in large mammals for two reasons. First, by contrast to small mammals in which color morphs are often selected as a means of predator avoidance, the evolutionary forces underlying color
Quantitative and chemical characterization of bear hair color

In addition to cinnamon, light-colored morphs of U. americanus have been described as blond, light brown, or chocolate; thus, the results presented reflect all of these color morphs. We measured reflectance of hair samples from 391 U. americanus and 33 U. arctos from across North America. Most U. americanus qualitatively described as black, based on photographs, exhibited reflectance values < 50 (median = 5), whereas animals scored as cinnamon had broadly ranging values (median = 67); overall, the distribution of reflectance varied continuously from 0.5 to 182 (possible range: 0–255; Figure 1A; Table S1). Reflectance in U. arctos also varied continuously (median = 48). Thus, we consider the reflectance measure a quantitative variable.

We determined if hair color of cinnamon U. americanus and U. arctos was caused by dilution of eumelanin or increased production of pheomelanin, using high-performance liquid chromatography (HPLC) of the different pigment types in awn hair samples. As depicted in Figures 1B and 1C, the major determinant of light-colored hair is reduced amounts of eumelanin in both bear species.

Allelic identification of TYRP1 variants

We sequenced the genomes of 24 U. americanus to >30x depth and an additional 166 animals to 1.3x depth, mapped to the...
reference genome, then imputed missing loci in the low-coverage samples using the high-coverage individuals as a reference panel. A genome-wide association study (GWAS) for quantitative hair reflectance values from 151 bears identified 120 single-nucleotide variants (SNVs) above a significance threshold of 10^{-8} (Figure 2A; Table S2). A single peak had the lowest p values and spanned the pigmentation gene TYRP1 (Figure 2B), in which we identified a significant (p = 1.99 × 10^{-13}) G to A SNV, HiC_scaffold_24:6724152g>a that predicts a missense alteration, p.Arg153Cys, referred to in what follows as TYRP1R153C. This alteration within the luminal domain of TYRP1, a type I transmembrane protein, is within a critical region near the zinc cofactor binding site and thus is likely deleterious (Figure 2C). Nine additional SNVs had lower p values and were 130 kb downstream of the TYRP1 transcription start site (Figure 2B); however, after accounting for TYRP1R153C genotypes as a covariate, the significance of these sites fell below the genome-wide threshold. Further, haplotype analysis identified low diversity in the Nevada population, thus we used these samples as input into HAPLOVIEW and identified a single 97 kb haplotype carrying TYRP1R153C (Figures 2B and S1). Recombination is apparent in populations outside of the southwest, suggesting that TYRP1R153C is derived from a single mutational event.

We explored coding sequence variation in TYRP1 and 12 additional pigmentation genes in U. americanus, U. arctos, and U. maritimus (polar bear). For U. americanus, we identified 46 missense or nonsense polymorphisms within 11 genes (Table S3); among these, only the TYRP1R153C variant was correlated with color (Figure S2A) and predicted to be deleterious. We identified 19 missense variants between U. americanus and U. arctos and 33 missense or nonsense variants between U. arctos and U. maritimus (Table S3). The U. americanus R153C variant was not observed in U. arctos or U. maritimus; however, a different TYRP1 variant was identified in 60% of U. arctos individuals, HiC_scaffold_24:6725331g>a, that predicts a p.Arg114Cys substitution (TYRP1R114C). Both mutations occur in the luminal domain of TYRP1, which contains 10 evolutionarily conserved cysteine (Cys) residues that form 5 disulfide bonds; acquisition of an unpaired Cys residue in TYRP1R153C and TYRP1R114C is likely to interfere with native disulfide bonding and protein folding (Table S3; Figures 2C and S2).

We also examined TYRP1 variation at orthologous positions in other species. In humans, loss of function for TYRP1 causes a rare form of albinism, rufous OCA3, observed mostly in individuals of African or Puerto Rican ancestry and characterized by reddish skin and hair and frequent visual abnormalities. Among 15 individuals with clinical albinism and mutations of TYRP1, R114C and R153C were each observed once as compound heterozygotes. The R153C variant is found at a relatively high frequency in individuals of European ancestry, 7.91 × 10^{-4}, but is classified as a variant of uncertain significance. In laboratory rats, the R114C allele is likely responsible for a spontaneous brown mutation that is fixed in the Brown Norway strain.

Functional characterization

To test whether the TYRP1 R153C and R114C variants affect function, we assessed their impact on pigmentation following expression in mouse melan-b cells, which are derived from TYRP1^{+/+} mutant mice and are hypopigmented due to the absence of functional or immunohistochemically detectable TYRP1. Wild-type (WT) mouse or human TYRP1 can rescue pigmentation in melan-b cells. We introduced the TYRP1 R153C or R114C variants into their orthologous positions in
mouse and human TYRP1 and expressed WT or variant TYRP1, or mCherry-syntaxin 13 (mCh-STX13) as a negative control, from recombinant retroviruses in melan-b cells. Using a colorimetric melanin content assay (Figure 3A), WT human TYRP1 (hTYRP1) rescued pigmentation relative to control (untransfected or mCh-STX13-transfected) melan-b cells to nearly the same level as WT (melan-Ink4a) melanocytes. By contrast, the R153C and R114C variants rescued pigmentation in melan-b cells—respectively, only 52.0% and 14.1%, as well as WT hTYRP1, and 46.0% and 6.8%, as well as WT mouse TYRP1. This correlates with melanin quantification in bear hair detected by HPLC (Figure 1B), and it indicates that R153C and R114C are causal variants for reduced pigmentation.

To test if the defect in pigmentation reflected impaired localization of the variants to melanosomes, we examined melan-b cells transfected with the different constructs by bright-field microscopy to visualize pigment granules and by immunofluorescence microscopy for TYRP1. The majority of melan-b cells transfected with WT hTYRP1 are densely pigmented. By contrast, most cells expressing hTYRP1R153C or hTYRP1R114C were either non-pigmented or lightly pigmented; some of the latter contained large aggregates of pigment, perhaps reflecting melanosome autophagy (Figure 3B). By immunofluorescence microscopy, only background labeling for TYRP1 was detected in melan-b cells that were untransduced (\(/-\); Figure 3G) or transduced with mCh-STX13 (Figure 3F). WT hTYRP1 was detected predominantly in “rings” around pigment granules detected by bright-field microscopy (Figures 3C [arrowheads] and 3H), representing localization to the limiting membrane of mature melanosomes.21 By contrast, hTYRP1R114C primarily localized to punctate and diffuse structures; the puncta did not overlap with pigment granules when present (Figures 3E...
Figure 3. Functional characterization of TYRP1 alleles

(A) TYRP1-deficient melan-b cells that were untransduced (−) or stably expressing the WT, R153C, or R114C variants of human (h) or mouse (m) TYRP1 or mCh-STX13 as a control were analyzed by quantitative melanin content assay and normalized to protein content. Data represent percent normalized melanin content relative to that of WT melan-Ink4a cells from six experiments (colored dots), each performed in duplicate, and analyzed by two-way ANOVA and Tukey’s multiple comparison test.

(B) Stable melan-b transductants expressing WT, R153C, or R114C hTYRP1 variants were analyzed by bright-field microscopy, and individual cells were characterized as densely pigmented, lightly pigmented, light with dense aggregates, or non-pigmented. Data from three experiments with 200 of each cell type per experiment were analyzed by a mixed effect analysis with Tukey’s multiple comparison test relative to WT hTYRP1.

(C–G) Indicated untransduced (−; G) or stable melan-b transductants (C–F) were analyzed by immunofluorescence microscopy for TYRP1 (left, green) and bright-field microscopy (BF) for pigment granules (middle; right, overlay with pigment granules pseudocolored magenta. Insets, 7× magnification of boxed regions.

(legend continued on next page)
and 3H). hTYRP1R153C was mostly detected in a mixture of ring-like and diffuse/punctate patterns (Figures 3D and 3H). Since melanosome size and shape are determined prior to the onset of pigmentation, TYRP1 would appear ring-like on both pigmented and non-pigmented melanosomes; thus, the reduced ring structures for hTYRP1R153C and hTYRP1R114C indicate that they do not localize properly to melanosomes. Considering only the cohort of cells that were densely pigmented, detectable pigment granules overlapped significantly less with TYRP1 in cells expressing hTYRP1R153C or hTYRP1R114C than WT hTYRP1 (Figure 3I). Given that the TYRP1 lumenal domain is required for proper melanosome localization, we speculate that eumelanin dilution in bears may be explained by protein misfolding due to substitution of Cys for arginine (Arg) at residues 114 or 153, leading to enhanced protein degradation and impaired trafficking to and/or retention within melanosomes.

### Figure 4. Spatial distribution of TYRP1R153C across the American black bear range

Quantitative and spatial assessment of TYRP1R153C in Ursus americanus. Inset: hair color reflectance was compared with the SNP genotype (n = 317). We used t tests to compare reflectance values between genotypes: homozygous ancestral and heterozygous, p = 2.02 × 10⁻³; heterozygous and homologous derived, p = 1.41 × 10⁻⁴. These results indicate the allele acts with semidominance. The geospatial pattern of R153C across the range (n = 906) where color denotes genotype (black, homozygous ancestral GG; brown, heterozygous GA; gold, homozygous derived AA), with the species range shown in dark gray.

For U. americanus, examining hair reflectance as a function of TYRP1R153C genotype in 317 individuals revealed semidominant gene action, with a median reflectance value for heterozygous G/A (Arg/Cys) individuals of 50.3 that was intermediate between values for homozygous G/G (Arg/Arg) and A/A (Cys/Cys) individuals of 4.7 and 94.0, respectively (Figure 4, inset).

We determined the spatial distribution of TYRP1R153C in 906 samples across the species’ geographic range. A high frequency of the derived allele was apparent in the southwest, declining northward along the Rocky Mountains into Southeast Alaska and Yukon Territory (Figure 4). By contrast, a low frequency of the derived allele was apparent in the eastern lineage, primarily limited to the Great Lakes region and a single sample in Connecticut, USA. The presence of the derived allele in Missouri, Arkansas, and Oklahoma, in the USA, was concordant with the known translocation history of bears from Minnesota.
into Arkansas in the 1960s and the high proportion of genetic diversity from the Great Lakes region now found in the Central Interior Highlands. Spatial distribution of $TYRP1^{R153C}$ was consistent with the phenotypic distribution of non-black $U. americanus$, as estimated from a survey of 40,000 animals in the 1980s. Notably, $TYRP1^{R153C}$ was associated with bears qualitatively described as cinnamon, chocolate, or light brown, and it therefore accounts for the majority of color diversity among $U. americanus$.

**Demography**

Our $TYRP1$ haplotype tree (Figure S2A) indicated that introgression of $U. arctos$ alleles into $U. americanus$ did not provide the genetic material for light coloration in the latter species. We investigated the relationship of the $TYRP1^{R153C}$ mutation to $U. americanus$ demographic history. First, we used runtc to estimate the first coalescent (a proxy for allele age) of the variant along the scaffold and obtained a time of 9.36 kya. Second, we ran MSMC-IM from whole-genome resequencing to estimate the allele age of $TYRP1^{R153C}$. Our working model of historic lineage divergence (purple arrows), approximate geographic region of $TYRP1^{R153C}$ mutation (asterisks, *), and migration across the species’ geographic range (solid black arrows; dashed black arrow indicating recent translocation). See also Figure S3.
(WGS) of six individuals, which estimated that divergence between the western and eastern lineages began 100 kya with cessation of gene flow 22 kya (Figures 5A and 5B). Taken together with additional coalescent and range expansion analyses (Figure S3), our results suggest that R153C arose in a western lineage population of *U. americanus* and spread via gene flow across the range (Figure 5C), including contemporary eastward movement too young to be captured using these coalescent estimates.

**Selection analyses**

We used population genetic modeling and simulations to investigate whether clinal variation of R153C allele frequency was more likely to be explained by selection or genetic drift. We compared models with a range of dominance (*h*) and selection (*s*) coefficients. As depicted in Figure 6, the clinal pattern of allele frequencies was most consistent with weak selection (*0.005 < s < 0.01*) and was unlikely to be explained by genetic drift (*s = 0*). Despite the long haplotype carrying the derived TYRP1<sup>R153C</sup> allele, there was no statistical evidence for a selective sweep (Figures S4 and S5), likely due to the combination of weak selection and low effective population size within the southwest population. In humans, SLC24A5 drives the main axis of geographic color variation in Caucasians and has a large *s* of 0.16.28 Although other genes, TYRP1 among them, have weaker coefficients of 0.01–0.05.28,29

Two hypotheses about the selective forces acting upon the cinnamon morph of *U. americanus* have been proposed. First, lighter coloration aids in thermoregulation of bears in the hotter and/or drier climates of the southwest; second, the cinnamon morph is mimetic with *U. arctos* where these species are sympatric. We used BayEnv to test for correlation between R153C allele frequency and climate, geography, or the presence/absence of *U. arctos* prior to anthropogenic extirpation, but we identified no notable Bayes factors (Table S4). Although this approach cannot rigorously exclude thermoregulation or mimicry as a force driving selection, it is possible that an additional, as yet untested mechanism underlies weak selection such as crypsis in open canopy environments. For example, individuals color matched within their environment has been suggested to decrease infanticide and/or predation in the Giant Panda.31

**Implications for infraspecific taxonomy in American black bears**

Historically, brown coat color has been used as a defining phenotype for delimiting 4 of the 16 *U. americanus* subspecies: *U. a. amblyceps* (southwest/southern Rocky Mountains), *U. a. cinnamomum* (northern Rocky Mountains), *U. a. luteolus* (Louisiana), and *U. a. machetes* (western Mexico).32–34 Further, the original description of *U. luteolus* from 1821 notes the cinnamon morph ranging from Virginia to Louisiana, where it is not presently found. Our results indicate that the geographic distribution of brown coats is being driven by recent and ongoing gene flow, suggesting that cinnamon morphs will increase in frequency within the eastern lineage in the future. We suggest that the young age of the causative variant and rapid spread via gene flow make this a poor character trait for infraspecific delimitation.

**Conclusion**

Here, we show that the cinnamon morph of the American black bear shares phenotypic similarity with brown bears in their coat coloration, and we demonstrate that eumelanin dilution causes this similarity. We identified two independently evolved and species-specific alleles within *TYRP1* that result in Arg to Cys...
residue changes that disrupt protein localization to the melanosome. Our functional assay shows that the U. arctos R114C change has a greater effect on pigment production in melan-b cells than the U. americanus R153C, yet R114C has not fixed within the species (Figure S2B), suggesting it may modify color without being the initial causative locus. By contrast, the U. americanus R153C is responsible for most color variation within the species. Neither variant appears detrimental to fitness in bears unlike in humans with OCA3; instead, R153C appears to be under weak positive selection. Crypsis as an adaptive mechanism has generally explained why prey species and ambush predators color match within their environments; here, we suggest crypsis as a broader adaptive mechanism for large-bodied species.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
  - Sample collection
  - Hair analyses
  - Whole genome sequencing and variant calling
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Chemical analysis of hairs
  - GWAS and Haplotype Characterization
  - Candidate gene analysis
  - Genotyping TYRP1R153C
  - Functional Validation of TYRP1 Alleles
  - Allele age
  - Demography
  - Selection tests
  - Gene flow simulation

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2022.11.042.

ACKNOWLEDGMENTS

We thank Omar Skalli for microscopy training, Jay Puckett for assistance with maps, and two anonymous reviewers for comments that improved the manuscript. E.E.P., I.S.D., and M.D.P. were supported by funds from the University of Memphis. C.B.K. and G.S.B. were supported by NIH R01 AR067925 and the HudsonAlpha Institute. M.S.M. and D.C.H. were supported by NIH R01 AR076241. A.P. was supported by NIH R01 AR067925 and the HudsonAlpha Institute. M.S.M. and C.B.K. (supporting); writing – original draft, E.E.P., G.S.B., and M.S.M. (equal) and G.B. (supporting); writing – review & editing, all authors.

REFERENCES

1. Hoekstra, H.E., Hirschmann, R.J., Bundey, R.A., Insel, P.A., and Crossland, J.P. (2006). A single amino acid mutation contributes to adaptive beach mouse color pattern. Science 313, 101–104. https://doi.org/10.1126/science.1126121.
2. Basu Mallick, C., Iliescu, F.M., Mois, M., Hill, S., Tamang, R., Chauvey, G., Goto, K., Ho, S.Y.W., Gallego Romero, I., Crivellaro, F., et al. (2013). The light skin allele of SLC24A5 in South Asians and Europeans shares identity by descent. PLoS Genet. 9, e1003912. https://doi.org/10.1371/journal.pgen.1003912.
3. Benjamin-Fink, N., and Reilly, B.K. (2017). Conservation implications of wildlife translocations; the state’s ability to act as conservation units for wildebeest populations in South Africa. Glob. Ecol. Conserv. 12, 46–58. https://doi.org/10.1016/j.gecco.2017.08.008.
4. Wikipedia (2021). The cinnamon bear. https://en.wikipedia.org/wiki/The_Cinnamon_Bear.
5. Rounds, R.C. (1987). Distribution and analysis of colourmorphs of the black bear (Ursus americanus). J. Biogeogr. 14, 521–538.
6. Pedersen, M.W., De Sanctis, B., Sarem, N.F., Sikora, M., Puckett, E.E., Gu, Z., Moon, K.L., Kapp, J.D., Vinner, L., Vardanyan, Z., et al. (2021). Environmental genomics of Late Pleistocene black bears and giant short-faced bears. Curr. Biol. 31, 2728–2736.e6. https://doi.org/10.1016/j.cub.2021.04.027.
7. Puckett, E.E., Etter, P.D., Johnson, E.A., and Eggert, L.S. (2015). Phylogeographic analyses of American black bears (Ursus americanus) suggest four glacial refugia and complex patterns of post-glacial admixture. Mol. Biol. Evol. 32, 2338–2350.
8. Ito, S., Del Bino, S., Hirobe, T., and Wakamatsu, K. (2020). Improved HPLC conditions to determine eumelanin and pheomelanin contents in biological samples using an ion pair reagent. Int. J. Mol. Sci. 21, 5134.
9. Ito, S., Nakaniishi, Y., Valenzuela, R.K., Brillant, M.H., Kolbe, L., and Wakamatsu, K. (2011). Usefulness of alkaline hydrogen peroxide oxidation to analyze eumelanin and pheomelanin in various tissue samples: application to chemical analysis of human hair melamin. Pigment Cell Melanoma Res. 24, 605–613. https://doi.org/10.1111/j.1755-146X.2011.00864.x.
10. Srivastava, A., Kumar Sarasvani, V., Fiddes, I., Sheehan, S.M., Seger, R.L., Bater, M.E., Neptune-Bear, S., Lindqvist, C., and Korstanje, R. (2019). Genome assembly and gene expression in the American black bear provides new insights into the renal response to hibernation. DNA Res. 26, 37–44. https://doi.org/10.1093/dnares/dsy036.
11. Rubinacci, S., Ribeiro, D.M., Hofmeister, R.J., and Delaneau, O. (2021). Efficient phasing and imputation of low-coverage sequencing data using large reference panels. Nat. Genet. 53, 120–126. https://doi.org/10.1038/s41588-020-00756-0.
81. Hutchinson, M., Xu, T., Houlder, D., Nix, H., and McMahon, J. (2009). ANUCLIM 6.0 User’s Guide (Australian National University, Fenner School of Environment and Society).

82. Haroldson, M.A., Clapham, M., Costello, C.M., Gunther, K.A., Kendall, K.C., Miller, S.D., Pigeon, K.E., Proctor, M.F., Rode, K.D., Servheen, C., et al. (2021). Brown bear (Ursus arctos; North America). In Bears of the World, V.P.a.M. Melletti, ed. (Cambridge University Press).

83. Lischer, H.E.L., and Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics 28, 298–299. https://doi.org/10.1093/bioinformatics/btr642.

84. Haller, B.C., and Messer, P.W. (2019). SLiM 3: forward genetic simulations beyond the Wright–Fisher model. Mol. Biol. Evol. 36, 632–637. https://doi.org/10.1093/molbev/msy228.

85. Malaney, J.L., Lackey, C.W., Beckmann, J.P., and Matocq, M.D. (2018). Natural rewilding of the Great Basin: genetic consequences of recolonization by black bears (Ursus americanus). Divers. Distrib. 24, 168–178. https://doi.org/10.1111/ddi.12666.
## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Antibodies          |        |            |
| TA99/Mel-5          | American Type Culture Collection | HB-8704 |
| Alexafluor 488-labeled donkey anti-mouse IgG | Jackson Immunoresearch | 715-545-151 |
| Bacterial and virus strains |
| E.coli-Top10        | Invitrogen | C66411 |
| Biological samples  |
| U. arctos- DNA sample | This study | GYE906 |
| U. arctos- DNA sample | This study | GYE922 |
| U. arctos- DNA sample | This study | GYE953 |
| U. arctos- DNA sample | This study | AK17500 |
| U. arctos- DNA sample | This study | AK17512 |
| U. arctos- DNA sample | This study | AK17578 |
| U. arctos- DNA sample | Liu et al. | RF01 |
| U. arctos- DNA sample | Liu et al. | SJ501 |
| U. arctos- DNA sample | Liu et al. | OFS01 |
| U. arctos- DNA sample | Benazzo et al. | ALP1 |
| U. arctos- DNA sample | Cahill et al. | AKAdmiralty1 |
| U. arctos- DNA sample | Miller et al. | AKAdmiralty2 |
| U. arctos- DNA sample | Cahill et al. | AKAdmiralty3 |
| U. arctos- DNA sample | Miller et al. | AKBaranof1 |
| U. arctos- DNA sample | Liu et al. | AKBaranof2 |
| U. arctos- DNA sample | Liu et al. | AKChichagof1 |
| U. arctos- DNA sample | Liu et al. | AKChichagof2 |
| U. arctos- DNA sample | Liu et al. | AKChichagof3 |
| U. arctos- DNA sample | Liu et al. | AKChichagof4 |
| U. arctos- DNA sample | Cahill et al. | AKChichagof5 |
| U. arctos- DNA sample | Cahill et al. | AKDenali1 |
| U. arctos- DNA sample | Miller et al. | AKNenai |
| U. arctos- DNA sample | Benazzo et al. | APN2 |
| U. arctos- DNA sample | Benazzo et al. | GRE2 |
| U. arctos- DNA sample | Endo et al. | JPhc1 |
| U. arctos- DNA sample | Endo et al. | JPhc2 |
| U. arctos- DNA sample | Endo et al. | JPhc3 |
| U. arctos- DNA sample | Endo et al. | JPhc4 |
| U. arctos- DNA sample | Endo et al. | JPhs1 |
| U. arctos- DNA sample | Endo et al. | JPhs2 |
| U. arctos- DNA sample | Liu et al. | MTgmp |
| U. maritimus- DNA sample | Miller et al. | Dalarna, SWE |
| U. maritimus- DNA sample | Miller et al. | PB1 |
| U. americanus- DNA sample | This study | PB2 |
| U. americanus- DNA sample | This study | AK17023 |
| U. americanus- DNA sample | This study | AK17047 |
| U. americanus- DNA sample | This study | AK17117 |

(Continued on next page)
| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| U. americanus- DNA sample | This study | AZ12 |
| U. americanus- DNA sample | This study | ID10 |
| U. americanus- DNA sample | This study | MI334 |
| U. americanus- DNA sample | This study | MI335 |
| U. americanus- DNA sample | This study | MN_6083 |
| U. americanus- DNA sample | This study | MS3783 |
| U. americanus- DNA sample | This study | NC00417 |
| U. americanus- DNA sample | This study | NC0056 |
| U. americanus- DNA sample | This study | NVb83 |
| U. americanus- DNA sample | This study | NVb99 |
| U. americanus- DNA sample | This study | NVg5 |
| U. americanus- DNA sample | This study | WV17_01 |
| U. americanus- DNA sample | This study | HA1 |
| U. americanus- DNA sample | This study | HA2 |
| U. americanus- DNA sample | This study | HA3 |
| U. americanus- DNA sample | This study | HA4 |
| U. americanus- DNA sample | This study | HA5 |
| U. americanus- DNA sample | This study | HA6 |
| U. americanus- DNA sample | This study | HA7 |
| U. americanus- DNA sample | This study | HA8 |
| U. americanus- DNA sample | This study | HA9 |

Chemicals, peptides, and recombinant proteins

| REAGENT | SOURCE | IDENTIFIER |
|---------|--------|------------|
| BamHI | NEB | R0136L |
| HindIII | NEB | R0104L |
| XholI | NEB | R0146L |
| SnaBI | NEB | R0130S |
| Lipofectamine 2000 | Invitrogen | 11668-019 |
| Hygromycin B | Research Products International | H75020-1.0 |
| RPMI 1640 medium | Gibco | 11875-085 |
| FBS | R&D Systems | S1150 |
| PBS | Corning | 21-031-CM |
| Tris-HCl | Fisher | BP153-1 |
| EDTA | Sigma | E5134-50G |
| dithiothreitol | Sigma | 3483-12-3 |
| DMSO | Sigma | D2660 |
| LB Broth | Fisher | BP1426-2 |
| Ampicillin | Fisher | BP1760-25 |
| Agar | Fisher | BP1423-2 |
| Agarose | GibcoBRL | 15510-027 |
| 2-log DNA Ladder | NEB | N3200L |
| GelGreen Nucleic Acid Stain | Biotium | 41005 |
| Tris-EDTA pH 8.0 | Ambion | AM9858 |
| Terrific Broth | Amresco | J869-5KG |
| Glucose | Sigma | G7528-250G |
| Sodium Hydroxide | Sigma | S8045 |
| SDS | Invitrogen | 15525-017 |
| Potassium Acetate | Sigma | P1190-100G |
| Glacial Acetic Acid | Fisher | BP1185-500 |

(Continued on next page)
### Critical commercial assays

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Prime Time Gene Expression (for qPCR) | Integrated DNA Technologies (IDT) | 10007065 |
| Complete Protease Inhibitor Cocktail Tablets | Roche | P8849 |
| Prolong Gold Antifade Mountant | Invitrogen | P36931 |
| Q5 High-Fidelity DNA Polymerase | NEB | M0491L |
| QiAquick Gel Extraction Kit | Qiagen | 28704 |
| BCA Protein Assay Kit | Pierce | 23225 |

### Deposited data

| DNA Sequencing | This study | SRA: PRJNA867575 |
|----------------|------------|------------------|
| TYRP1 mouse WT sequence | Browning and Browning[^41] | NM_031202.3 |
| Crystal structure of the human TYRP1 luminal domain | NCBI Protein Data Base | PDB: 5M8L |

### Experimental models: Cell lines

- melan-b | Dorothy Bennett’s laboratory | Gabriel et al.[^12] |

### Oligonucleotides

| pBMN vector R153C primer; forward: GGGCCCTGGATA TGGCAAAGTGCACAACT CACCCTTTATTTGT | IDT | custom oligo synthesis |
|-----------------------------|--------|------------------|
| pBMN vector R153C primer; reverse: ACAATAAAAGGG TGAGTTGTGCACCTTGCC ATATCCAGGGCCC | IDT | custom oligo synthesis |
| pBMN vector R114C primer; forward: GACACAACGTGTGGACGTGCTGTCCTGGCTGGAGAGGAGCTGC | IDT | custom oligo synthesis |
| pBMN vector R114C primer; reverse: GCAGCTCCTCTC AGGCCAGGAGCAGCAGT CCCACAGTGGTC | IDT | custom oligo synthesis |
| pBMN flanking primer; forward: CCTCTAGAAGTC GGATCCCTTTTTAGTGC AATTCCTGCAGG | IDT | custom oligo synthesis |
| pBMN flanking primer; reverse: GTAATCAAAGTTGG TTTCTGCTATCCCCTCAAGTCTAGGTCAG | IDT | custom oligo synthesis |
| qPCR TYRP1 primer; forward: CTTGGAAGTGCAGG AGAAACC | This paper | N/A |
| qPCR TYRP1 primer; reverse: CTGGTGCAGCAGAC ACAAAC | This paper | N/A |
| pPCR TYRP1 reference allele probe: FAM-ATGG CGAAGGCAAAATTC-BHQ1 | This paper | N/A |
| pPCR TYRP1 alternative allele probe: TET-ATGGCGGA GTGCAAAAT-BHQ1 | This paper | N/A |

(Continued on next page)
## REAGENT or RESOURCE SOURCE

| **REAGENT or RESOURCE** | **SOURCE** | **IDENTIFIER** |
|-------------------------|------------|----------------|
| Recombinant DNA         |            |                |
| pCDNA3-TRP1             | Gift from Walter Storkus, Univ. of Pittsburgh | N/A |
| pBMN-I-hygro-mCherry-hSTX13 | Marks laboratory | Taylor et al. [43] |
| pBMN-I-hygro            | Gift from Andrew Peden, Univ. of Sheffield | Liu et al. [35] |

## Software and algorithms

| **Software and algorithms** | **Source** | **Identifier** |
|----------------------------|------------|----------------|
| BWA v0.7.17                | Altschul et al. [44] | https://github.com/lh3/bwa |
| GATK v4.1.8.0              | Zhang et al. [45] | https://gatk.broadinstitute.org/hc/en-us |
| SAMTOOLS v1.9              | Li et al. [36] | https://sourceforge.net/projects/samtools/files/samtools/1.9/ |
| BCFTOOLS v1.9              | Ng and Henikoff [17] | https://sourceforge.net/projects/samtools/files/samtools/1.9/ |
| BEAGLE v5.1                | Choi et al. [38] | https://faculty.washington.edu/browning/beagle/b5_1.html |
| GLIMPSE 1.0.0              | Rubinacci et al. [11] | https://github.com/odelaneau/ GLIMPSE/pkgs/container/ GLIMPSE%2Fglimpse |
| GEMMA v0.98.1              | Zhou and Stephens [12] | https://github.com/genetics-statistics/GEMMA |
| HAPLOVIEW 4.2              | Barrett et al. [14] | https://www.broadinstitute.org/haploview/downloads |
| Haplostrips                | Marnetto and Huerta-Sánchez [13] | https://bitbucket.org/dmarnetto/haplostrips/src/master/ |
| PROVEAN                   | Adzhubei et al. [49] | https://www.jcvi.org/research/provean/downloads |
| PolyPhen2                  | Geneious Prime [50] | http://genetics.bwh.harvard.edu/pph2/ |
| SIFT                      | R Core Team [51] | https://sift.bii.a-star.edu.sg/index.html |
| GENEIOUSPRIME             | Setty et al. [52] | https://www.geneious.com/prime/ |
| Leica LAS-X               | Leica Microsystems | https://www.leica-microsystems.com/products/microscope-software/p/leica-las-x-ls/ |

## Microfluidics

| **Microfluidics** | **Source** | **Identifier** |
|-------------------|------------|----------------|
| Bio-Vision Technologies | N/A | https://bio-vision.com/ |
| FIJI – ImageJ | Bennett et al. [53] | https://imagej.net/software/fiji/ |
| Adobe Photoshop v. 24.0.0 | Adobe.com | https://www.adobe.com/products/photoshop/ |
| Graphpad Prism v. 9.4.1 | GraphPad.com | https://www.graphpad.com/updates/prism-900-release-notes |
| Pymol              | Schrödinger LLC | https://pymol.org/ |
| program R v4.0.3   | Bennett et al. [54] | https://www.r-project.org/ |
| Runtc              | Platt et al. [55] | https://github.com/jaredgk/runtc |
| STACKS v2.53       | Morita et al. [56] | https://catchenlab.life.illinois.edu/stacks/ |
| PLINK v1.9         | Dennis et al. [57] | https://github.com/mbi-genetics/plink/ |
| MSMC2              | Meng et al. [58] | https://github.com/mbi-genetics/msmc2 |
| SNPable            | Bowman et al. [59] | https://lh3lh3.users.sourceforge.net/snpable.shtml |
| Rangeexpansion     | Delevoye et al. [59] | https://github.com/BenjaminPeter/rangeexpansion |

(Continued on next page)
We extracted DNA from blood or tissue from 906 *U. americanus* individuals for whole genome sequencing (WGS) or qPCR genotyping of *TYRP1R153C*. DNA was extracted from an additional six *U. arctos* individuals (muscle samples) for WGS. Hair color was measured from 391 *U. americanus* and 33 *U. arctos* individuals. Animals were not maintained in the lab, see below for details.

**METHOD DETAILS**

**Sample collection**
We partnered with state, provincial, and federal wildlife agencies, and university partners in North America which collected hair, tissue (i.e., muscle or blood), and photographs from bears that were hunter harvested, vehicle killed, or live animals captured for other research studies or management purposes. Not all samples contained all three sampling elements (hair, tissue, and photographs). Photographs of bears confirmed all samples were adults; cubs may turn from brown to black following a molt, although black to brown transitions have not been observed (personal observation DL Garshelis). Samples were collected from 2015–2020 and georeferenced. The qPCR analysis (see below) incorporated previously collected bear samples that were also georeferenced but without phenotypic data.

**Hair analyses**
We mounted one awn hair from 391 individuals onto slides, then captured an image from the middle of the sample under 20x magnification using an Olympus BX63 microscope at the Integrated Microscopy Center at the University of Memphis. We viewed all
We sequenced 15 U. americanus and six U. arctos individuals to ~30x depth (key resources table). Specifically, 350bp insert libraries were prepared using the NEB Next Ultra II DNA kit, then sequenced (150bp paired-end) on an Illumina NextGen by Novogene (Chula Vista, CA). We sequenced nine additional U. americanus individuals to 50x depth using Illumina HiSeqX at HudsonAlpha (Huntsville, AL). We also downloaded 26 U. arctos and two U. maritimus WGS from the NCBI SRA to use in downstream analyses (key resources table). For all species, we mapped reads to the U. americanus reference genome10 using BWA-MEM v0.7.175 with default parameters. We marked duplicates, then called variants for each sample using the GenotypeGVCFs function in GATK v4.1.8.09 where we set the heterozygosity parameter to 5.0x10^-4. For each species, we combined the samples using CombineGVCFs before joint genotyping using GenotypeGVCFs in GATK. We filtered the dataset using BCFTOOLS v1.955 by requiring minimum and maximum depth (i ’MIN (FMT/DP) > 4 & MAX (FMT/DP) < 300’), minimum quality score (-i FMT/GQ >= 30), and biallelic SNPs (-m 2 -M 2 -v snps). The total number of SNPs for U. americanus was 8,885,511. We phased the data from each species separately using BEAGLE v5.1.58,41

Low coverage

We sequenced 166 U. americanus animals to an average of 1.3x depth. Specifically, we followed the manufacturer’s instructions for the Illumina DNA Prep Kit except that we diluted the reactions to 20% volume and started with 20 ng of genomic DNA. Two pooled libraries were constructed and each was sequenced on one lane of an Illumina 4000 at Novogene (Davis, CA). We mapped reads to the U. americanus reference and called variants as described above. The draft U. americanus (2n = 74) genome was assembled into 94,016 scaffolds; we analyzed the longest 36 scaffolds (and removed the X chromosome, HiC_scaffold_1). We used our high coverage WGS samples as a reference panel for joint imputation and phasing of the low coverage data using GLIMPSE.11

QUANTIFICATION AND STATISTICAL ANALYSIS

Chemical analysis of hairs

We quantified the amount of eumelanin and pheomelanin in 13 U. americanus and three U. arctos individuals. Samples were selected based on both geographic location and visual inspection of hair color under a compound microscope to select for observed variation (U. americanus: black = 9; brown = 4; U. arctos: dark brown = 2, brown = 1). We followed the melanin determination protocols.8,9 Eumelanin and pheomelanin concentrations are reported as the concentrations of PTCA (pyrrole-2,3,5-tricarboxylic acid) and TTCA (thiazole-2,4,5-tricarboxylic acid) which are, respectively, the specific degradation products for each melanin. We plotted the figure in Figures 1B and 1C and report ANOVA results in the figure legend.

GWAS and Haplotype Characterization

To identify loci associated with the brown phenotype, we ran GEMMA v0.98.1.12 Specifically, we combined our high and imputed low coverage datasets, then used the quantitative hair reflectance values (green channel of RGB measurement) as the phenotype (n = 151). We ran a principal components analysis (PCA) on the relatedness matrix created within GEMMA, and used the first two PC axes as covariates to account for population structure in the analysis. Our GWAS resulted in 23 genomic regions with P-values lower than 10^-8, although a single peak stood out and contained a known pigmentation gene,

Candidate gene analysis

To evaluate deleterious protein coding variation in bears, we downloaded coding sequences (CDS) for 13 candidate coat color genes (ASIP, CBD103, EDNRB, KIT, KITLG, MC1R, MLPH, OCA2, PMEL, SLC24A5, SLC45A2, TYR, TYRP1, and TYRP2) from the dog genome (Canis familiaris; CanFam3.1). Exons were queried against genome assemblies of U. americanus, U. arctos, and U. maritimus15,43,35 using the blastn and dc-megablast versions of NCBI BLAST+.44,45 We extracted the resulting sequences from the phased genome assemblies using SAMTOOLS v1.945 and BCFTOOLS and concatenated exons to form haplotypes. We translated each haplotype and identified missense and nonsense polymorphisms within U. americanus, between U. americanus and four U. arctos samples, and between U. arctos and U. maritimus. We tested each variant for deleterious effects using PROVEAN, PolyPhen2, and SIFT47–49 (Table S3). For any variant where two of the programs predicted a deleterious variant, we identified the individuals and coat color phenotypes of those animals to assess if there was concordance between U. americanus animals with
brown coats and the alternate allele. Only the previously identified TYRP1<sup>R153C</sup> showed such a concordance. We constructed a neighbor joining tree from the haplotypes (with introns) using GENEIOUSPRIME<sup>50</sup> to visualize the gene tree.

To understand the location of the arginine to cysteine amino acid changes within the native TYRP1 protein, we input the X-ray crystal structure of the human TYRP1 luminal domain (PDB ID: 5M8L, ref Lai et al. <sup>15</sup>) into PyMol v2.5, displayed it as a ribbon diagram to emphasize alpha helices, beta sheets, and loops, and then highlighted the following features: the location of R114 and R153 amino acids (expressed in atom format), cysteines and corresponding disulfide bonds throughout the protein, and the zinc cofactors.

**Genotyping TYRP1<sup>R153C</sup>**

We genotyped 906 <i>U. americanus</i> samples for the TYRP1<sup>R153C</sup> variant using a qPCR assay.

- The primers included (5’ to 3’):
  - F: CCTTGAAGTCAAGGAGAAACC
  - R: CTGGTGCAATAGCAAAAC
- The probes included (5’ to 3’):
  - reference allele probe: FAM-ATGGCGAAGGCACAAATC-BHQ1
  - alternate allele probe: TET-ATGGCGAAGTAGCACAAT-BHQ1
- Primers and probes were ordered from IDT (Coralville, IA). Reaction conditions had final primer and probe concentrations of 1μM and 0.2 μM, respectively. The thermocycler conditions in a BioRAD CFX96 included: 95°C for 3min, and 40 cycles of 95°C for 15sec then 57.9°C for 30sec. For samples where both hair color phenotypes and R153C genotypes were available (n = 317), we tested if reflectance was significantly different between genotypes using two-sided t-tests in R v3.6.0<sup>51</sup> and report results in the legend of Figure 4.

**Functional Validation of TYRP1 Alleles**

To test if the TYRP1 variants we identified in American black (R153C) and brown (R114C) bears caused a change in pigmentation phenotype, we first used site-specific mutagenesis to generate recombinant retroviruses to co-express wildtype (WT) or variant versions of human and mouse TYRP1 with a drug resistance marker for stable cell line selection. Specifically, we first subcloned a WT TYRP1 cDNA based on the human sequence (pCDNA3-TRP1, a gift to MS Marks from W Storkus, University of Pittsburgh, Pittsburgh, PA, USA) into the pBWMN-I-Hygro retroviral vector.<sup>52</sup> The R153C and R114C variants were generated individually by overlapping PCR using synthetic oligonucleotides bearing the desired mutations (denoted by asterisks [*] in the sequences below).

- The oligos for R153C were (5’ to 3’):
  - F: GGGCCCTCGATAGCCAGAA*GGACAACTACCCCTTATTATG
  - R: ACAATAAAGGGTAGTTGTGCA*CCTTGCATATCCAGGCCCCCC
  - The oligos for R114C were (5’ to 3’):
  - F: GACACAAAAGTGGGAGAC*GTCTCGGCTGGAGGAGGACTGC
  - R: GCAGCTCCTCCTGGCAGGACA*GACAGTCCACAGTGTGTC
  - The flanking oligos for both products were (5’ to 3’):
  - F: CCTCTAGACTGCGGGCTCATATTTAAATTTGAGGACAG
  - R: GAATCAAAAGTTGTCTTGCACATTCCGATCAAGTCTCCGACAG
- The R153C mutation was generated by combining a 707bp and a 386bp product to obtain a 1,050bp product. The R114C mutation was generated by combining a 503bp and a 503bp product to obtain a 1,050bp product. The PCR products bearing the mutations were cloned individually as 1,013bp BamHI fragments into the BamHI site in the pBWN-I-Hygro vector. The R153C and R114C variants were generated individually by overlapping PCR using synthetic oligonucleotides bearing the desired mutations (denoted by asterisks [*] in the sequences below).

- The primers included (5’ to 3’):
  - R: CTGGTCGCAATGACAAAC
  - F: CCTTGAAGTCAGGAGAAACC
- The probes included (5’ to 3’):
  - reference allele probe: TET-ATGGCGAAGTAGCACAAT-BHQ1
  - alternate allele probe: FAM-ATGGCGAAGGCACAAATC-BHQ1
- Primers and probes were ordered from IDT (Coralville, IA). Reaction conditions had final primer and probe concentrations of 1μM and 0.2 μM, respectively. The thermocycler conditions in a BioRAD CFX96 included: 95°C for 3min, and 40 cycles of 95°C for 15sec then 57.9°C for 30sec. For samples where both hair color phenotypes and R153C genotypes were available (n = 317), we tested if reflectance was significantly different between genotypes using two-sided t-tests in R v3.6.0<sup>51</sup> and report results in the legend of Figure 4.

The WT mouse TYRP1 cDNA sequence (NCBI accession NM_031202.3) was also used to generate the mutant forms. The three versions (WT, R153C, and R114C) were synthesized individually by Synbio Technologies as XhoI/HpaI fragments (of 2270bp) and cloned into the XhoI/SnaBI site of the pBWMN-I-Hygro vector. The three sequences were confirmed by Sanger sequencing. The generated mutations were confirmed by Sanger sequencing.

The WT mouse TYRP1 cDNA sequence was used to generate the mutant forms. The three versions (WT, R153C, and R114C) were synthesized individually by Synbio Technologies as XhoI/HpaI fragments (of 2270bp) and cloned into the XhoI/SnaBI site of the pBWMN-I-Hygro vector. The three sequences were confirmed by Sanger sequencing.

Immortalized melan-b melanocytes derived from mice homozygous for the TYRP1<sup>R153C</sup> allele were cultured as described.<sup>53,54</sup> Melan-b cells were infected with recombinant retroviruses encoding TYRP1 variants (described above). Briefly, Plat-E cells<sup>55</sup> were cultured and transfected with the retroviral vectors described above, empty pBWN-I-hygro, or pBWN-I-hygro expressing mCherry-tagged human STX13<sup>56</sup> using Lipofectamine 2000. The medium was replaced with fresh melanocyte medium the following day and cell supernatants were harvested 48-72 h post-transfection as described.<sup>57</sup> The medium, containing recombinant retroviruses, was used directly to infect melan-b cells as described<sup>58,59</sup> except that cells were seeded the previous day in a 10-cm dish. Two days later the medium was supplemented with 500 mg ml<sup>1</sup> hygromycin B to select for infected cells. Antibiotic-resistant cells were used directly as polyclonal cell populations for experiments. Two separate stable lines per construct were selected and included in the analyses.

We performed spectrophotometric assays or melanin and protein content as described.<sup>58,59</sup> Briefly, melanocytes in a confluent 10-cm dish were harvested by trypsinization. Cells were counted, and duplicate samples of 10<sup>6</sup> cells for each experimental condition per experiment were washed once with RPMI + 10% FBS and twice with ice-cold PBS, and then resuspended in 50 mM Tris-HCl, pH 7.4, 2 mM EDTA, 150 mM NaCl, 1 mM diithiothreitol, and cOmplete Protease Inhibitor Cocktail Tablets (Roche). Cells were probe-sonicated on ice using the Sonic Dismembrator Model 100 (Fisher Scientific), and sonicates were fractionated into supernatant and pellet fractions by centrifugation at 20,000 x g for 15 min at 4 °C. Protein concentrations in supernatants were estimated using the DC
colorimetric protein assay (Pierce) with a standard curve derived using differing concentrations of BSA. Melanin-containing pellets were washed in 500 µl 50% ethanol, 50% ether, collected by centrifugation at 20,000 x g, and resuspended in 1 ml of 2 M NaOH, 20% DMSO. Melanin was solubilized by heating at 60°C for 40–60 min and quantified using spectrophotometry to measure Optical Density (OD) at 492 nm. OD values were normalized to total protein content in each lysate. The assay was performed on six different days using samples collected from two different sets of stable cell lines per construct.

We performed imaging and analysis as described. Briefly, melanocytes were seeded onto Matrigel-(Corning) coated coverslips 48–72 h before fixation. Cells were fixed with 2% formaldehyde in PBS for 20 min, washed briefly with PBS, then labeled with TA99/Mel-5 antibody to TYRP1 (American Type Culture Collection, Rockville, MD) diluted in blocking buffer (PBS, 0.1% BSA, 0.02% saponin) for 1 h at room temperature (RT). After a 15-min wash in PBS, cells were incubated with Alexafluor 488-labeled donkey anti-mouse IgG (Jackson Immunoresearch) diluted in blocking buffer for 30 min at RT. Samples were washed for 15 min in PBS, mounted with Prolong Gold Antifade Mountant (Invitrogen), and analyzed by epifluorescence microscopy on a DMI 6000B microscope (Leica Biosystems) equipped with a 63x plan apochromat (Plan Apo) objective (1.4NA) and a Hamamatsu Photronics ORCA-Flash4.0 sCMOS digital camera. Both fluorescence and bright field images were acquired as a z-series with 0.19-µm steps and were deconvolved using the blind deconvolution algorithm of Microvolution software (Bio-Vision Technologies) and ImageJ. Final images were composed using Adobe Photoshop software. Images were quantified using ImageJ.

To quantify the percentages of cells with varying levels of pigmentation, bright field images of individual cells within randomly chosen fields were characterized as non-pigmented (no visual pigmentation or fewer than 10 pigment granules per cell), densely pigmented (easily visualized pigmentation throughout the cell with numerous granules), lightly pigmented (signal above background but either very low signal or fewer than 40% of the total number of granules for wild-type cells), or lightly pigmented with prominent pigment aggregates (structures larger than 1 micron in diameter, typically one or two per cell). Data from three experiments with 200 cells per sample type per experiment were analyzed by mixed effect analysis and Tukey’s multiple comparison test.

To quantify the TYRP1 staining pattern, cells in randomly chosen fields were characterized as having predominantly ring-like structures (easily detectable “donuts” of less than 1 micron in diameter), diffuse and/or punctate staining (hazy cytoplasmic fluorescence and/or small puncta with no visible lumen), or a combination of both. Cells were analyzed blinded and without observing pigmentation status. Data from four experiments with 150 cells per sample type per experiment were analyzed by two-way ANOVA (because there were multiple variables) with Dunnett’s multiple comparison test.

The area of overlap between TYRP1 and pigmented melanosomes in the cell periphery of melanocytes was quantified using a thresholding approach in ImageJ on deconvolved fluorescence images as detailed. Briefly, images were cropped to individual representative cells that had sufficient pigmentation to allow for thresholding. Both the bright field image and the corresponding deconvolved TYRP1 immunofluorescence image were subjected to the following manipulations. Background was subtracted using the Background Subtract function with a rolling ball radius of 2 microns, and then an identical mask corresponding to the perinuclear region (which is typically very densely labeled by TYRP1) was cropped out of both images. The Adjust>Threshold function was used to generate binary images that closely resembled the original images. Structures larger than five pixels were quantified using the Analyze Particles function, the area of overlap between the two images was calculated by using the Math>Multiply function, and values were calculated relative to the total area of thresholded bright field objects. Data from at least 30 cells of each cell type from three independent experiments were analyzed by Kruskal-Wallis test (because the data were not Gaussian) with Dunn’s multiple comparison test.

All statistical analyses were performed using Graphpad Prism 9 for MacOS.

**Allele age**

We estimated the age of TYRP1<sup>R153C</sup> using runtc, a program that estimates the time of first coalescence to the nearest variant on a chromosome. To increase sample size for this analysis, we used the combined high and imputed low coverage data for HiC_scaffold_24. We reran the GLIMPSE pipeline on the low coverage samples where the high coverage reference panel was not first filtered for MAF, thereby leaving more singletons in the reference dataset for coalescence with the focal allele. We estimated the coalescence time for all variants on the scaffold with all range wide samples together, as well as separately in the western lineage, Southeast Alaska (SEAK), and eastern populations separately, using a mutation rate of 1x10<sup>-8</sup> and the recombination rate from the dog genome, 0.97cM Mb<sup>-1</sup>. We report all and lineage specific distributions and variant estimates in Figure S3A.

**Demography**

We inferred the demographic history of U. americanus, then tested if spatial patterns of TYRP1<sup>R153C</sup> alleles were due to neutral or adaptive processes given the demographic. We first assessed U. americanus population structure using PCA. We aligned a previously generated RAD-Seq dataset† to the U. americanus reference genome and called SNPs using STACKS v2.53, including filters for one SNP per RAD-tag, >80% genotyping success per locus, and MAF > 0.05; this resulted in 37,422 SNPs. We pulled down those loci from the high coverage WGS samples to increase the sample size (n = 117). We ran a PCA analyses in PLINK v1.9, using the RAD-seq+hcWGS and observed a concordant pattern to the previous analysis; thus, we felt confident with analyzing lineages and populations as previously determined.

We estimated change in effective population size (N) through time using MSMC2. For this analysis we used scaffolds 2–37, and made individual mask files with the bamCaller.py script, then used the phased datasets to make genome-wide masks with SNPable. We generated input scripts using generate_multithetsep.py. Output was converted to years and number of individuals using a mutation
rate of $1 \times 10^{-871}$ and a generation time of 6.5 years.\textsuperscript{72} We ran MSMC2 on three samples (six haplotypes) from both the eastern and western lineages, and the admixed SEAK population. Historic $N_e$ was approximately 30,000. From 100–10kya $N_e$ declined to 3,000 in the western lineage, yet remained constant in the eastern lineage. $N_e$ also declined in Alaska from 50–10kya to approximately 8,000. MSMC2 plots indicate $N_e$ increased since the Last Glacial Maximum (18–22kya). This demographic history provides an explanation for the low haplotype diversity in Nevada as the ancestral and causative alleles sit on highly homogeneous haplotypes (Figures S1 and S4). Using the output from MSMC2, we further estimated the cross-coalescent model and input results into MSMC-IM\textsuperscript{77} to estimate the patterns and timing of population divergence (using MSMC2 estimates), and rate of gene flow over time (using PSMC\textsuperscript{1} estimates on two haplotypes for the Idaho and Arizona samples). Estimates may be inaccurate in the most recent time lags of MSMC2 analyses warranting caution when estimating contemporary $N_e$. Our data appear to have a limit for inferences younger than 1–5kya.

While our demographic model infers broad patterns of lineage relatedness, it does not identify the directionality of range expansion; thus, we used the RAD-seq+hcWGS data to understand patterns of range expansion via the directionality index ($\Psi$).\textsuperscript{72} We removed geographic sites with fewer than three samples for a total of 111 samples across 23 populations. We input genotypes and geographic data into the rangeExpansion package in R, then calculated $\Psi$ for all population pairs using the get.all.psi function. We determined the significance of each pairwise $\Psi$ estimate by calculating the standard error of the upper triangle of the matrix excluding the diagonal, then calculating a Z-score for each pairwise comparison. To visualize the range expansion patterns, we plotted pairwise values where $Z > |5|$ within the western and eastern lineages, and between the two lineages and SEAK (Figures S3B–S3D).

### Selection tests

We calculated $\pi$ ($\rho$) using VCFTOOLS\textsuperscript{74} in 1kb sliding windows with a 100bp step size from the imputed low coverage WGS data using samples from the Nevada (western lineage; $n = 38$), SEAK ($n = 95$), and Appalachian Mountains (representative of eastern lineage; $n = 38$) populations. We further calculated $F_{ST}$ between black and brown animals within either the Nevada or SEAK populations using the same sliding window scheme.

We tested for a selective sweep surrounding TYRP1 using both integrated haplotype score (iHS) and extended haplotype homozygosity (EHH)\textsuperscript{75} implemented in selscan v2.\textsuperscript{76,77} The iHS analysis was run on scaffold 24 of the *U. americanus* reference genome within the Nevada population, then normalized based on allele frequency using the norm function within selscan v2. No sites within the haplotype block carrying TYRP1\textsuperscript{R153C} had iHS scores greater than $|2|$. We ran the EHH analysis in both the Nevada and SEAK populations using the –keep-low-freq flag. Further, we identified 15 sites on each of the 36 long scaffolds (or 540 sites each) that matched the derived allele frequency of R153C within each test population and ran EHH on each of those sites. We compared the decay curves of R153C to these allele frequency matched background loci as a test if the decay in haplotype homozygosity was within the range of the background (Figure S5). While decay curves between the ancestral and derived alleles suggested extended haplotype homozygosity for the derived R153C, the background matched samples showed the overall scores were not out of range for the background sites. Both the low $N_e$ within the Nevada population and low selection coefficient complicate the analysis of selection, as simulation studies suggest low power to detect selection under these conditions.\textsuperscript{76} Moreover, we chose not to use cross-population sweep tests due to the decrease in power with increasing divergence, and our estimates that the eastern and western lineage populations diverged $\sim$15.8k generations ago.\textsuperscript{78}

We further tested for possible drivers of selection using BayEnv2.\textsuperscript{30,79} Using the combined RAD-Seq, hcWGS, and lcWGS data at the RAD loci, we clustered the data into 24 geographic areas, identified a centroid longitude and latitude coordinate, then extracted values of 11 climate variables and altitude for each coordinate. The climate variables from the CliMond dataset\textsuperscript{80,81} included: bio1 (annual mean temperature), bio3 (isothermality), bio4 (temperature seasonality), bio5 (max temperature of warmest week), bio6 (min temperature of coldest week), bio10 (mean temperature of warmest quarter), bio11 (mean temperature of coldest quarter), bio12 (annual precipitation), bio19 (precipitation of coldest quarter), and bio20 (annual mean radiation). Beyond the climate variables, altitude, latitude, and longitude were all tested as selective forces. Finally, we used the presence or absence of *U. arctos* on the landscape as a factor. For each geographic location, we scored the site with the presence of *U. arctos* before range contraction due to human activity based on maps of the expected species distribution.\textsuperscript{82} We prepared the covariance matrix using PGDSpider,\textsuperscript{83} and the environmental matrix according to the BayEnv2 manual. We ran 100,000 iterations of the model and analyzed the Bayes Factors (Table S4) for the alternative hypothesis that the variable of interest was a selective force on allele frequency of TYRP1\textsuperscript{R153C}.

### Gene flow simulation

To test if a model of gene flow without selection could produce the spatial gradient of declining frequency of the alternative allele along a south-to-north axis is the western range, we used the forward genetic simulation program SLiM.\textsuperscript{84} We modeled four populations, broadly representative of Nevada, Arizona/New Mexico, Idaho, and SEAK. These population were selected as we have at least one high coverage genome from each population (key resources table). We reran MSMC-IM on one genome from each population for all pairwise combinations to estimate historic gene flow rates to use as input in our model (Figure S6).

Our model seeds a new variant (i.e., R153C) at a frequency of 1% into the Nevada population at generation 0, then tracks allele frequency within each population for 1,440 generations such that we get a simulated estimate to compare to observed data. Our data suggests that R153C has a dominance coefficient ($h$) greater than 0.5, but not 1; thus, we varied $h$ from 0 to 1 using 0.1 increments. Further, we ran each model with a selection coefficient ($\theta$) of either 0, 0.005, or 0.01. MSMC-IM estimates bidirectional gene flow, thus for each population pair we use the peak estimated rate (Figure S6) since the Last Glacial Maximum (LGM). We ran three
different models varying the selection coefficients (s = 0, 0.005, or 0.01). From generations 980–1,310, population size increased in each population at a per generation rate of 10^{-6} which was a rate approximated from our MSMC results. At generation 1,390 (i.e., 50 generations prior to the end of the simulation and meant to represent anthropogenic influences on bear populations), we imposed population specific bottlenecks given estimates from the literature. Specifically, the Nevada population was reduced to 100, and the Arizona/New Mexico, Idaho, and SEAK populations were reduced to 5,000, 7,000, and 4,000 respectively.7 We ran 1,000 iterations of each model. We recorded the number of simulations in which the derived allele went extinct in all populations or was still polymorphic following 1,440 generations (no simulations observed the derived allele fixing in all populations; reported in the legend of Figure 6). For simulations where the locus remained polymorphic, we averaged the frequency within each population across simulations, and compared to the observed frequency from our qPCR assay and contour plots of percent of black animals in the population.

This simulation with 33 different scenarios (h by s combinations) identified that both increasing values of h or s resulted in increased allele frequencies and decreased probability of allele extirpation from the landscape. We reran the simulation with an h of 0.75 for the final model (Figure 6).