A community-science approach identifies genetic variants associated with three color morphs in ball pythons (*Python regius*)

Autumn R. Brown[1], Kaylee Comai[1], Dominic Mannino[1], Haily McCullough[1], Yamini Donekal[1], Hunter C. Meyers[1], The BIO306W Consortium[^1], Chiron W. Graves[1*], and Hannah S. Seidel[1*]

[1]Department of Biology, Eastern Michigan University, Ypsilanti, MI, USA

* Corresponding author:

E-mail: hseidel@emich.edu (HSS), cgraves6@emich.edu (CWG)

[^]The BIO306W Consortium comprises students in an undergraduate laboratory course at Eastern Michigan University. Students are listed in the Acknowledgments.

Abstract

Color morphs in ball pythons (*Python regius*) provide a unique and largely untapped resource for understanding the genetics of coloration in reptiles. Here we use a community-science approach to investigate the genetics of three color morphs affecting production of the pigment melanin. These morphs—Albino, Lavender Albino, and Ultramel—show a loss of melanin in the skin and eyes, ranging from severe (Albino) to moderate (Lavender Albino) to mild (Ultramel). To identify genetic variants causing each morph, we recruited shed skins of pet ball pythons via social media, extracted DNA from the skins, and searched for putative loss-of-function variants in homologs of genes controlling melanin production in other vertebrates. We report that the Albino morph is associated with missense and non-coding variants in the gene *TYR*. The Lavender Albino morph is associated with a deletion in the gene *OCA2*. The Ultramel morph is associated with a missense variant and a putative deletion in the gene *TYRP1*. Our study is one of the first to identify genetic variants associated with color morphs in ball pythons and shows that pet samples recruited from the community can provide a resource for genetic studies in this species.

Data availability: All relevant data are within the paper and its Supporting Information files. DNA sequences are available from GenBank, under accession numbers MZ269492-MZ269502.

Funding information: This work was supported by a Faculty Research Fellowship and James H. Brickley Award from Eastern Michigan University to HSS and an Undergraduate Research Stimulus Program Award and a Don Brown and Meta Hellwig Undergraduate Research Award from Eastern Michigan University to ARB. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Color patterns are distinctive and beautiful features of many animal species. Among the many functions of color are to camouflage animals to their surroundings, protect tissues from ultraviolet radiation, warn predators of poisons, and provide signals for mating and social communication [1]. Color patterns have also been targeted by artificial selection to create novel and fanciful color patterns in domestic animals [2].

Colors are produced through a combination of chemical pigments and physical structures (structural colors). Pigments absorb light, whereas structural colors are created when light is reflected by the nanoscale geometry of a tissue. Common pigments include brown-to-black melanin, present throughout the animal kingdom, and yellow-to-red carotenoids and pteridines, common in birds, reptiles, and lower vertebrates [3]. Color-producing structures include ordered keratin matrices in bird feathers [4], chitin layers in butterfly wings [5], and purine crystals in the skin of fish, amphibians, and reptiles [6,7]. These color-producing structures can produce a variety of colors, including iridescent colors, depending on the wavelengths of light they reflect [8].

The genetics and development of color patterns in vertebrates have been studied most extensively in mammals. Mammals largely rely on a single type of pigment (melanin), produced in the skin by a single type of cell (melanocyte) [9]. This system provides mammals with skin and hair colors ranging from black to reddish-brown, depending on the chemical structure of the melanin [10].

Skin colors outside mammals are more diverse and include bright colors of all hues. This increased complexity is produced through a combination of structural colors and melanin and non-melanin pigment, and it relies on multiple types of specialized color-producing cells in the skin [11–15]. Color cell development has been characterized to some extent in fish [16–18], and genes controlling the use of non-melanin pigments have been identified in a few species of fish and birds [19–25]. The genetics and development of non-mammalian color patterns is less well understood in other vertebrates, particularly in reptiles [although see 26–29].

A unique resource for understanding the genetics of color patterns in reptiles is the ball python (*P. regius*). Ball pythons are native to sub-Saharan Africa, but have become common as pets in the United States [30]. Wild ball pythons exhibit a mottled color pattern, consisting of brown-to-black melanin and red-to-yellow (non-melanin) pigments in the skin (Fig 1). Captive-bred ball pythons, by contrast, include many variants of the normal color pattern [26,31–33]. These variants, referred to as ‘color morphs’, include animals having reduced melanin (e.g. Albino), increased melanin (e.g. Cinnamon), reduced red-to-yellow coloration (e.g. Axanthic), or complex changes in the placement of color patches on the skin (e.g. Spider, Clown, Genetic Stripe, and Enchi). Many color morphs are heritable and show simple dominant or recessive patterns of inheritance. These inheritance patterns are consistent with single-gene causality, but only a single genetic variant associated with a ball python color morph (Piebald) has been identified to date [26]. Ball pythons therefore represent a tractable yet largely untapped resource for understanding the genetics of coloration in reptiles. An additional feature of ball pythons convenient for genetic studies is that DNA samples can be obtained non-invasively from shed skin [34].

The goal of the current study was to perform proof-of-concept experiments to identify the genetic causes of color morphs in ball pythons, using shed skins of pet ball pythons recruited via social media. We focused on three morphs for which candidate genes could be readily identified: Albino, Lavender Albino, and Ultramel. These morphs show a loss of brown-to-black coloration in the skin and eyes, characteristic of a defect in melanin production (Fig 1). These morphs are recessive and non-allelic (i.e. crosses between these morphs yield offspring with normal coloration). Their loss of melanin ranges from severe (Albino) to moderate (Lavender Albino) to mild (Ultramel). This range of phenotypes mirrors the range of phenotypes observed for loss-of-function variants in genes required for melanin production in other vertebrates (Table 1). This

similarity provided a list of candidate genes that we predicted might harbor loss-of-function variants causing the Albino, Lavender Albino, and Ultramel color morphs in ball pythons. Our study demonstrates the feasibility of using community-sourced samples for genetic studies in ball pythons and lays the groundwork for future investigations of reptile-specific color patterns in this species.

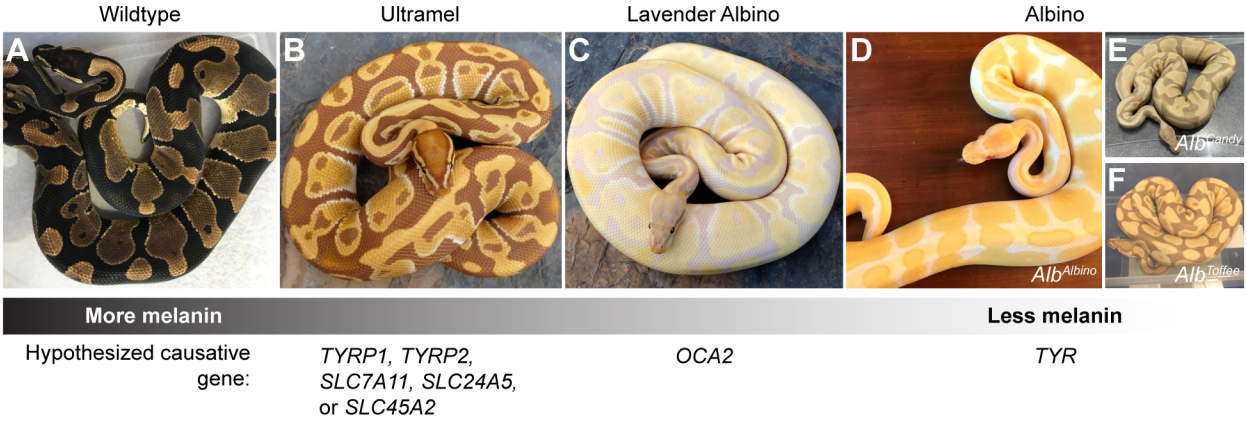

Fig 1. The Albino, Lavender Albino, and Ultramel color morphs have reduced brown-to-black coloration, characteristic of a loss of melanin. Red-to-yellow coloration is unaffected in these morphs. Hypothesized causative genes represent genes in which loss-of-function variants in other vertebrates produce similar phenotypes (Table 1). (A) Wildtype. (B) Ultramel. (C) Lavender Albino. (D-F) Albino. (D) Albino animal described as an Alb^{Albino} homozygote. (E) Albino animal described as an Alb^{Candy} homozygote. (F) Albino animal described as an Alb^{Toffee} homozygote. Photo credits, Ryan Young of Molecular Reptile, Chiron Graves, Phil Barclay, Michael Freedman of The Florida Reptile Ranch.

98 **Table 1. Loss-of-function phenotypes of melanogenesis genes across vertebrates.**

Gene	Species with identified loss-of-function variants		Role of protein	Loss-of-function phenotype
TYR	Fish	Japanese carp [35], Japanese rice fish [36]	Rate-limiting enzyme in melanin synthesis pathway	Severe loss of melanin
	Frogs and snakes	Japanese wild frogs (three species) [37], Japanese rat snake [38]		
	Birds	Chicken [39–41]		
	Ungulates	Water buffalo [42], Cattle [43], Red deer [44], Asinara donkey [45]		
	Rodents	House mouse [46], Wistar rat [47], Domestic guinea pig [48]		
	Primates	Crab-eating macaque [49], Capuchin monkey [50], Hamadryas baboon [51], Human [reviewed in 52]		
	Other mammals	Humpback whale [53], Ferret [54], American mink [55], Domestic cat [56–58], Silver fox [59]		
OCA2	Fish	Mexican cavefish [60], Lake Malawi cichlid [61]	Cl^- channel regulating melanosome pH	Moderate loss of melanin
	Snakes	Corn snake [62]		
	Mammals	Domestic dog [63], Bama miniature pig [64], House mouse [65], Human [reviewed in 52]		
TYRP1	Fish	Zebrafish [66]	Enzymes contributing to melanin synthesis	Mild loss of melanin
	Birds	Saker falcon [67]		
	Mammals	Chinese indigenous pig [68], Liangshan pig [69], Domestic cat [58,70], Soay sheep [71], Valais Red sheep [72], American mink [73], Human [reviewed in 52]		
TYRP2	Mammals	House mouse [74,75]		
SLC7A11	Mammals	House mouse [76]	Solute transporters	
SLC24A5	Fish	Zebrafish [77]		
	Mammals	Domestic dog [78], Horse [79], Human [reviewed in 52]		
SLC45A2	Fish	Japanese rice fish [80]		
	Birds	Chicken and Japanese quail [81]		
	Mammals	Horse [82], House mouse [83], Bengal tiger [84], Western lowland gorilla [85], Human [reviewed in 52]		

99

Results

Obtaining a reference sequence for melanogenesis genes in a wildtype ball python

Genes required for melanin production have been identified in humans and other vertebrates (Table 1). These genes encode enzymes that synthesize melanin (*TYR*, *TYRP1*, *TYRP2*) [86], a chloride channel required for maintaining the pH of melanosomes (*OCA2*) [87], and transporters thought to import solutes into the cell or into organelles (*SLC7A11*, *SLC24A5*, and *SLC45A2*) [76,88,89]. These genes are highly conserved among vertebrates and occur in single copy in most vertebrate genomes. Loss-of-function variants in these genes in humans cause a genetic disorder known as oculocutaneous albinism, which is characterized by loss of melanin in the skin, hair, and eyes [52]. This loss of melanin ranges from severe to mild, depending on the causative gene and exact genetics variants therein [52]. Similar phenotypes occur in other animals, where the loss of melanin extends to feathers and scales (Table 1).

To obtain a reference sequence for melanogenesis genes in ball python, we amplified and sequenced the coding regions of *TYR*, *TYRP1*, *TYRP2*, *OCA2*, *SLC7A11*, *SLC24A5*, and *SLC45A2* from a single ball python having normal coloration (henceforth 'wildtype'). Primers for amplification were designed against the genome of Burmese python (*Python bivittatus*), the closest relative of ball python for which genome sequence was available [90]. Comparison of ball python sequences to sequences from Burmese python revealed 97.2-99.2% nucleotide identity within coding regions and 96.9-98.5% nucleotide identity within flanking non-coding regions (Table 2). This analysis provided a reference sequence for genes in ball python that might harbor loss-of-function variants causing the Albino, Lavender Albino, and Ultramel color morphs.

Table 2. Comparison of ball python and Burmese python genomic and protein sequences for melanogenesis genes.

Region	Gene	Total nucleotide sequence aligned (bp)	Nucleotide identity (%)	Amino acid identity (%)
Coding	*TYR*	1,591	98.5	99.6
	TYRP1	1,584	99.2	99.6
	TYRP2	1,473	98.3	99.1
	OCA2	2,577	98.9	99.7
	SLC7A11	1,506	98.7	99.7
	SLC24A5	1,379	98.2	99.5
	SLC45A2	1,557	97.2	98.9
Non-coding	*TYR*	577	96.9	(not applicable)
	TYRP1	1,160	97.0	
	TYRP2	1,060	97.9	
	OCA2	2,154	98.5	
	SLC7A11	3,407	97.9	
	SLC24A5	1,335	97.9	
	SLC45A2	999	97.0	

The Albino color morph is associated with three haplotypes of *TYR*

The Albino color morph in ball pythons is characterized by an absence or near absence of melanin in the skin and eyes. This phenotype resembles the phenotype caused by loss of function of *TYR* in other vertebrates. *TYR* encodes the enzyme catalyzing the rate-limiting step of melanin production [86]. Loss of this enzymatic activity in other vertebrates causes a severe loss of melanin, which is typically more severe than the loss of melanin caused by loss of function of other melanogenesis genes (Table 1). We therefore hypothesized that the Albino color morph in ball pythons was caused by loss-of-function variants in *TYR*.

To test whether the Albino color morph was associated with variants in *TYR*, we performed a small-scale association study using 50 Albinos and 56 Non-Albinos, recruited from a total of 18 states in the United States. Polymorphic sites were identified through pilot sequencing of a subset of animals (see Methods). Six polymorphic sites were selected for the association study, two near the 5' end of *TYR* and four near the 3' end (Fig 2A). Each animal was genotyped at each site, and haplotypes were reconstructed using PHASE [91,92]. Association between haplotype and color morph was tested using the case-control test of PHASE. As a negative control, we also tested for association between haplotype and color morph at *TYRP1*, *TYRP2*, and *OCA2*, using two polymorphic sites per gene (S1 Table). This analysis revealed a significant association between haplotype and color morph at *TYR* ($p = 0.04$, Bonferroni corrected), but no association for the other three genes ($p > 0.05$, Bonferroni corrected). The association for TYR was driven by two features of the haplotype distribution: (i) diversity of *TYR* haplotypes was reduced among Albinos compared to Non-Albinos (Fig 2A), and (ii) two of the three *TYR* haplotypes found in Albinos were rare among Non-Albinos (Fig 2A). These results demonstrate an association between the Albino color morph and variants in TYR; they further show that all Albino animals were homozygous or compound heterozygous for any of three haplotypes of *TYR*.

We hypothesized that each of the three haplotypes of *TYR* found among Albinos might carry a distinct loss-of-function variant in the gene. To search for such variants, we selected one Albino animal homozygous for each haplotype and sequenced the *TYR* coding regions and adjacent splice sites in these animals. We found that one of these animals was homozygous for an aspartic acid-to-glycine variant (*D394G*) in the third coding region of *TYR*. A second animal was homozygous for a proline-to-leucine variant (*P384L*), also in the third coding region of *TYR*. Both variants alter conserved residues (Fig 2B), and the *P384L* variant occurs at the same site as a similar variant (*P384A*) associated with oculocutaneous albinism in humans [93]. The third animal carried no coding variants and no splice-site variants compared to wildtype. These results show that two of the three *TYR* haplotypes found among Albinos carried missense variants that are likely disruptive for TYR protein function. We term these *TYR* haplotypes TYR^{D394G} and TYR^{P384L}. We term the third Albino haplotype, which lacked coding or splice-site variants compared to wildtype, TYR^{Albino}.

We hypothesized that *D394G*, *P384L*, and an unidentified variant on the TYR^{Albino} haplotype were causative for the Albino color morph. This hypothesis predicts that animals homozygous or compound heterozygous for these variants will be Albino. Heterozygotes and non-carriers are predicted to be Non-Albino. To test this prediction, we genotyped our full panel of 50 Albinos and 56 Non-Albinos for the missense variants *D394G* and *P384L*. Consistent with our prediction, we found that animals homozygous or compound heterozygous for *D394G*, *P384L*, or the TYR^{Albino} haplotype were exclusively Albino (Fig 2C). Heterozygotes and non-carriers were exclusively Non-Albino (Fig 2C). The most common haplotype among Albinos was TYR^{Albino} (haplotype frequency of 71%), followed by TYR^{P384L} (16%) and TYR^{D394G} (13%). We conclude that *D394G*, *P384L*, and an unidentified variant on the TYR^{Albino} haplotype are likely causative for the Albino color morph. Any combination of these variants produces the Albino phenotype.

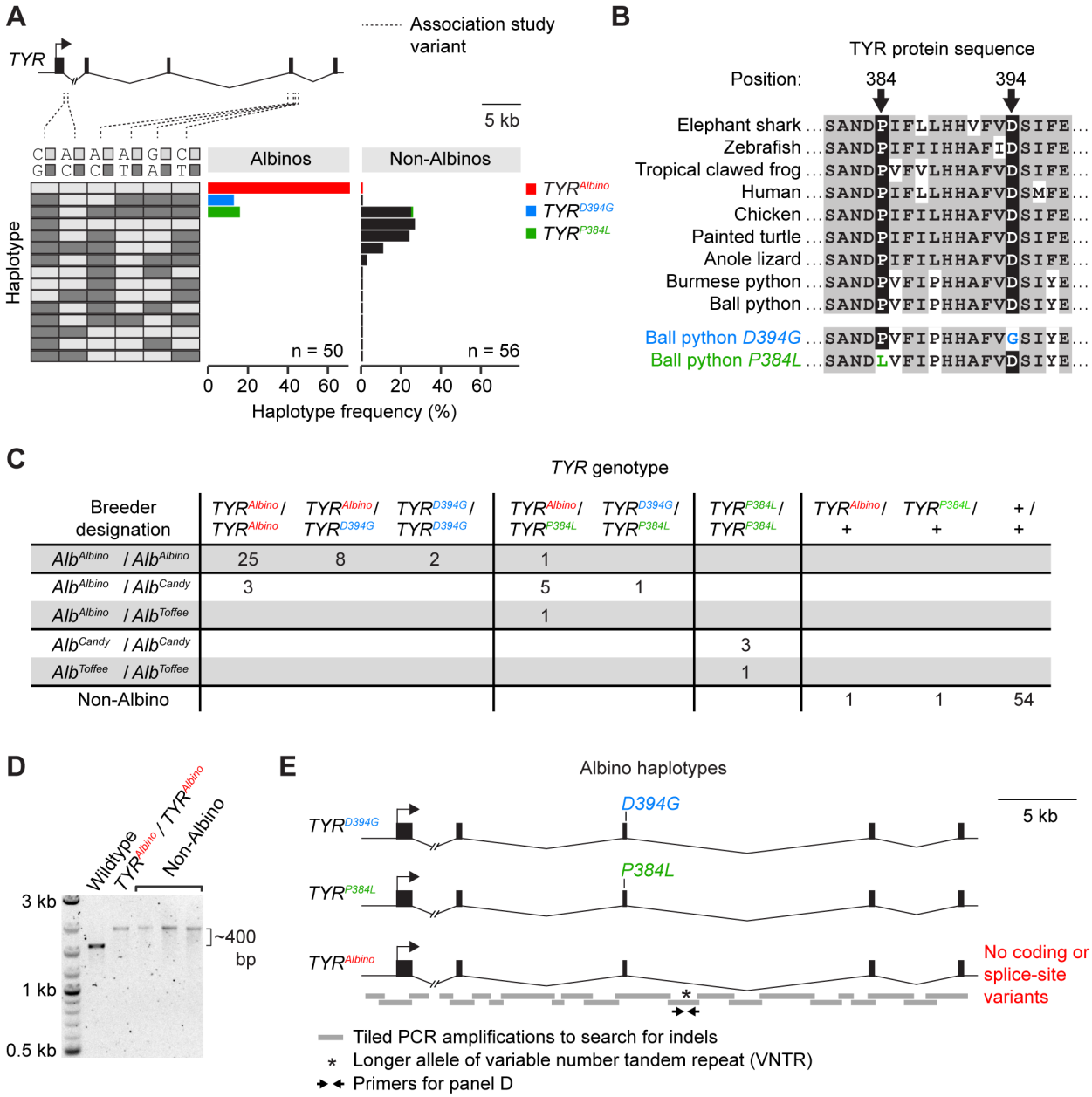

Fig 2. The Albino color morph is associated with three haplotypes of *TYR*. (A) *TYR* haplotype frequencies among Albinos and Non-Albinos. (B) Alignment of *TYR* protein sequences surrounding the missense variants *D394G* and *P384L*. (C) *TYR* genotypes and breeder designations of animals used in the association study. +, any *TYR* haplotype found exclusively among Non-Albinos. (D) PCR amplification of a genomic fragment containing the variable number tandem repeat (VNTR) shown in E. Non-Albinos are examples of Non-Albinos homozygous for the longer allele of the VNTR. (E) Schematic of the three *TYR* haplotypes found in Albinos. The TYR^{Albino} haplotype contains no coding variants and no splice-site variants compared to wildtype. (A, E) Hash mark, discontinuity in the Burmese python reference genome.

The *TYR*^{Albino} haplotype lacks an obvious loss-of-function variant

The lack of coding or splice-site variants on the *TYR*^{Albino} haplotype led us to hypothesize that this haplotype carried a loss-of-function variant in a non-coding region. To search for such variants, we examined the *TYR* promoter. We sequenced ~2 kb immediately upstream of the *TYR* start codon in a *TYR*^{Albino} homozygote and compared this sequence to wildtype. We found that the *TYR*^{Albino} haplotype differed from wildtype at a total three sites (three single-base substitutions). In all three cases, the *TYR*^{Albino} haplotype was homozygous for the allele shared with Burmese python (i.e. the ancestral allele). This result suggests that these variants are not causative for the Albino color morph. We conclude that the causative variant in the *TYR*^{Albino} haplotype does not reside within the sequenced region, ~2 kb upstream of the *TYR* start codon.

We hypothesized that the *TYR*^{Albino} haplotype might contain a large insertion or deletion (indel) in an intron that disrupts gene function. Intronic indels can disrupt slicing and have been found to disrupt the function of melanogenesis genes in other species [e.g. 62,73]. To search for large intronic indels, we tiled PCR amplicons across *TYR* introns and compared amplicon sizes between a *TYR*^{Albino} homozygote and a wildtype animal. These amplicons tiled across a total of ~38 kb of intronic sequence, which included part of intron 1 and all of introns 2, 3, and 4 (Fig 2E). (Tiling across intron 1 was incomplete because this intron contains a discontinuity in the Burmese python reference genome. Amplification across this discontinuity was unsuccessful in ball python, even in wildtype.) Size differences between the *TYR*^{Albino} homozygote and wildtype were assessed via standard agarose gel electrophoresis, which we estimate was sensitive enough to reveal indels larger than ~50-100 bp, depending on amplicon size (which ranged from ~0.7-3.5 kb). This analysis revealed one size difference (Fig 2D). Further sequencing revealed that this size difference was caused by a variable number tandem repeat (VNTR) that was ~400 bp larger in the *TYR*^{Albino} homozygote than in wildtype. Genotyping of the 56 Non-Albinos for this indel revealed that eight of these animals were homozygous for the larger allele of the VNTR (Fig 2D). This finding suggests that the larger allele of the VNTR is not causative for the Albino color morph. We propose that the *TYR*^{Albino} haplotype harbors a loss-of-function variant other than a large intronic indel or outside the genomic regions analyzed here.

TYR^{P384L} is associated with reduced phenotype severity compared to *TYR*^{D394G} and *TYR*^{Albino}

Ball python breeders describe the Albino color morph as having three alleles: *Alb*^{Albino}, *Alb*^{Candy}, and *Alb*^{Toffee}. Evidence for this view comes in part from variation in the coloration of Albino animals. In some Albino animals, the brown-to-black coloration observed in wildtype is entirely absent, and skin patches appear white (Fig 1D). These Albinos are typically considered to be *Alb*^{Albino} homozygotes. Other Albinos show a less severe phenotype, where skin patches are faintly beige or lavender instead of white (Figs 1E and 1F). These Albinos are typically considered to carry one or more copies of *Alb*^{Candy} or *Alb*^{Toffee}.

We hypothesized that the Albino alleles recognized by breeders might correspond to the *TYR* haplotypes identified through sequencing (*TYR*^{D394G}, *TYR*^{P384L}, and *TYR*^{Albino}). To test this hypothesis, we examined the breeder designations of the 50 Albinos in our panel. We found that the *Alb*^{Candy} or *Alb*^{Toffee} designations typically corresponded the *TYR*^{P384L} haplotype (Fig 2C). The *Alb*^{Albino} designation typically corresponded to the other two haplotypes (*TYR*^{D394G} and *TYR*^{Albino}, Fig 2C). This correspondence was imperfect, and exceptions existed (e.g. three animals designated as *Alb*^{Albino} / *Alb*^{Candy} compound heterozygotes did not carry the *TYR*^{P384L} haplotype, Fig 2C). We conclude that the *Alb*^{Candy} and *Alb*^{Toffee} designations typically (but not exclusively) represent the *TYR*^{P384L} haplotype. The *Alb*^{Albino} designation typically (but not exclusively)

represents the TYR^{D394G} or TYR^{Albino} haplotype. The association between the Alb^{Candy} and Alb^{Toffee} designations and the TYR^{P384L} haplotype suggests that this haplotype may confer a slightly less severe phenotype than do TYR^{D394G} and TYR^{Albino}.

The Lavender Albino color morph is associated with a deletion in *OCA2*

The Lavender Albino color morph is characterized by skin patches that are lavender instead of brown or black. This phenotype is thought to arise from melanin levels that are dramatically reduced but not entirely eliminated. Similar phenotypes have been observed in other vertebrates upon loss of function of *OCA2* (Table 1). *OCA2* encodes a chloride channel required for maintaining the pH of melanosomes [87,94]. When the OCA2 protein is absent or non-functional, the enzymes that synthesize melanin are less active, and only small amounts of melanin are produced [95,96]. The resulting phenotype is typically less severe than the loss-of-function phenotype of *TYR*, but more severe than the loss-of-function phenotypes of other melanogenesis genes (Table 1). We therefore hypothesized that the Lavender Albino color morph was caused by loss-of-function variants in *OCA2*.

To search for loss-of-function variants in *OCA2*, we selected a single Lavender Albino animal and sequenced 23 of the 24 coding regions of *OCA2* in this animal. Comparison of these sequences to wildtype revealed no coding variants and no splice-site variants. We attempted to repeat this analysis for the remaining coding region (coding region 18), but we were unable to amplify this coding region from the Lavender Albino animal. Further test amplifications revealed that the Lavender Albino animal was homozygous for a 1,514-bp deletion spanning coding region 18 (Figs 3A and 3B). This deletion removes 36 amino acids from the protein and likely introduces a frameshift into the transcript, given that the coding regions downstream of the deletion are out of frame compared to coding regions upstream of the deletion (Fig 3E). This frameshifted transcript is predicted to produce a protein lacking six of the 12 transmembrane helices present in wildtype OCA2. Truncations occurring at similar positions in the OCA2 protein have been identified in humans and are associated with oculocutaneous albinism [97]. We conclude that the deletion of *OCA2* coding region 18 likely disrupts protein function and is a strong candidate for the cause of the Lavender Albino color morph.

The Lavender Albino color morph is considered by breeders to have a single allele. We therefore predicted that the *OCA2* deletion would be shared by other Lavender Albino animals. We predicted that Non-Lavender Albinos would be heterozygous for the deletion or non-carriers. To test this prediction, we genotyped the *OCA2* deletion in 13 additional Lavender Albino animals. We also genotyped 76 Non-Lavender Albinos and one animal described as heterozygous for the Lavender Albino color morph. We found that all 13 Lavender Albinos were homozygous for the deletion (Figs 3C and 3D). All 76 Non-Lavender Albinos were non-carriers. The animal reported to be heterozygous for Lavender Albino was heterozygous for the deletion. These findings support the conclusion that the *OCA2* deletion is causative for the Lavender Albino color morph (Fig 3E).

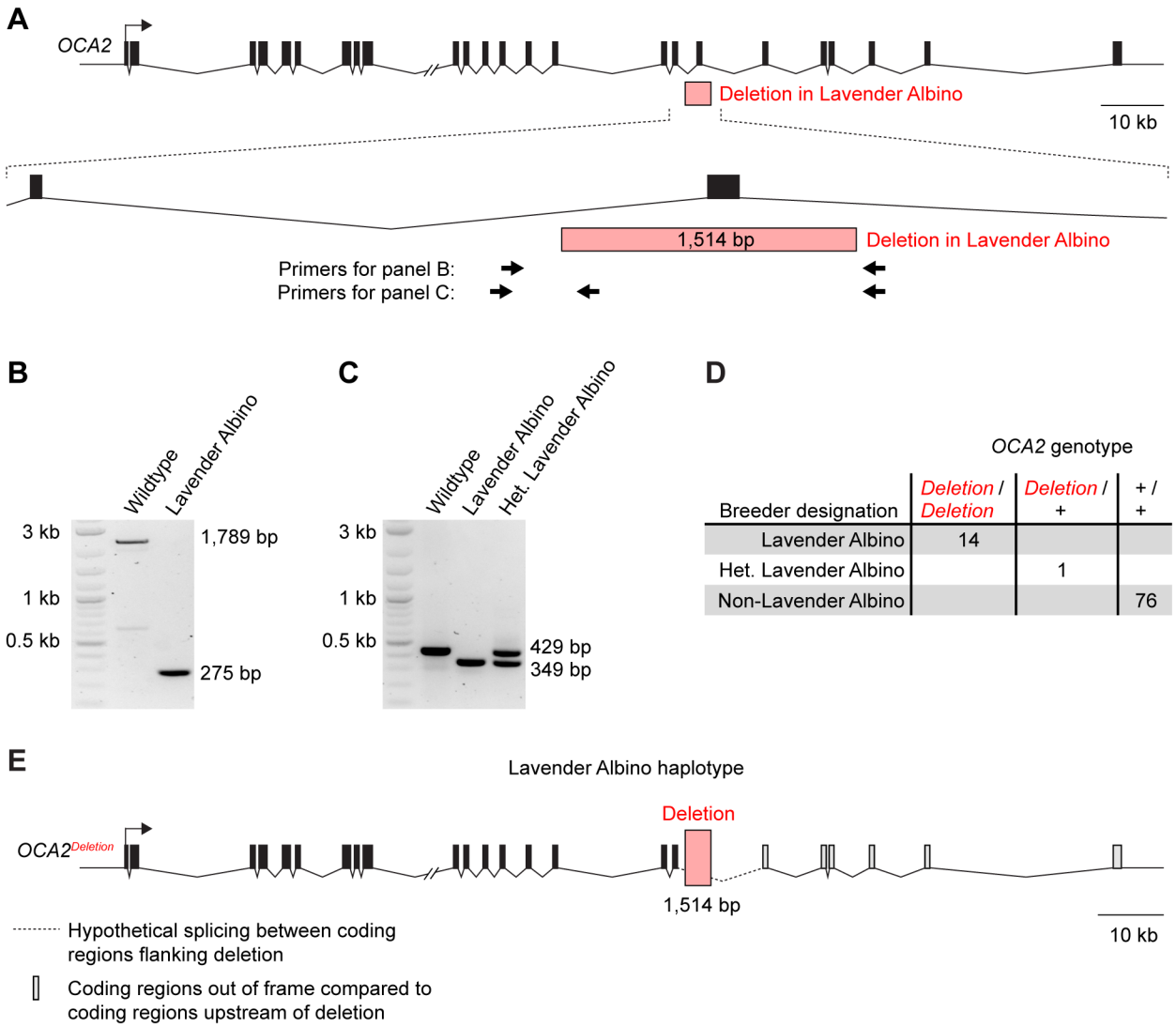

Fig 3. The Lavender Albino color morph is associated with a deletion in *OCA2*. (A) Schematic of the *OCA2* gene. (B) PCR amplification demonstrating the deletion in Lavender Albino. (C) PCR amplification used for genotyping the deletion in the animals in panel D. (D) Genotypes of 14 Lavender Albinos, 76 Non-Lavender Albinos, and one animal described as heterozygous for Lavender Albino. This set of animals includes the original Lavender Albino animal in which the OCA2 deletion was identified. (E) Schematic of the *OCA2* haplotype found in Lavender Albinos. (A, E) Hash mark, discontinuity in the Burmese python reference genome.

The Ultramel color morph is associated with a missense variant and a putative deletion in *TYRP1*

The Ultramel color morph is characterized by skin patches that are tan or light brown, rather than dark brown or black. This phenotype suggests a mild loss of melanin. Similar phenotypes in other vertebrates have been observed upon loss of function of any five genes: *TYRP1*, *TYRP2*, *SLC7A11*, *SLC24A5*, and *SLC45A2* (Table 1). *TYRP1* and *TYRP2* encode enzymes involved in synthesizing melanin [86]. Their enzymatic activities are partially reductant with other enzymes in the melanin synthesis pathway, thus explaining their milder loss-of-function phenotypes compared to *TYR*. *SLC7A11* encodes a transporter responsible for importing cystine into the cell [98]. Cystine is a precursor to some forms of melanin, and hence absence of the transporter alters melanin levels [76]. *SLC24A5* and *SLC45A2* encode a K^+-dependent Na^+-Ca^{2+} exchanger [88] and a putative sugar transporter [89], respectively. Loss of either of these proteins causes a reduction in melanin, through mechanisms that may involve defects in regulation of melanosome pH [83,88,99]. We hypothesized that the Ultramel color morph was caused by loss-of-function variants in one of these genes.

To search for loss-of-function variants in *TYRP1*, *TYRP2*, *SLC7A11*, *SLC24A5*, and *SLC45A2*, we selected a single Ultramel animal and sequenced the coding regions of each of these genes in this animal. Comparison of these sequences to wildtype revealed a single coding variant: an arginine-to-histidine substitution (*R305H*) in coding region 4 of *TYRP1*. The Ultramel animal was homozygous for this variant. The arginine residue at this site is conserved across vertebrates (Fig 4A) and is also conserved in *TYR*, which is paralogous to *TYRP1* [100]. An arginine-to-histidine substitution at the homologous site in *TYR* has been reported in humans and is associated with oculocutaneous albinism [101–103], suggesting that a histidine at this site is disruptive to protein function. The *R305H* variant is therefore a good candidate for the cause of the Ultramel color morph.

The Ultramel color morph is considered by breeders to have a single allele. We therefore predicted that the *R305H* variant would be shared by other Ultramels. We predicted that Non-Ultramels would be heterozygous or non-carriers. To test this prediction, we genotyped the *R305H* variant in 10 additional Ultramels and 78 Non-Ultramels. We found that five of the 10 Ultramels were homozygous for the *R305H* variant (Fig 4B). Of the remaining Ultramels, two were heterozygous for the *R305H* variant, and three did not carry it. By contrast, none of the 78 Non-Ultramels carried the *R305H* variant. This pattern is consistent with the *R305H* variant causing the Ultramel phenotype in some Ultramels but not others.

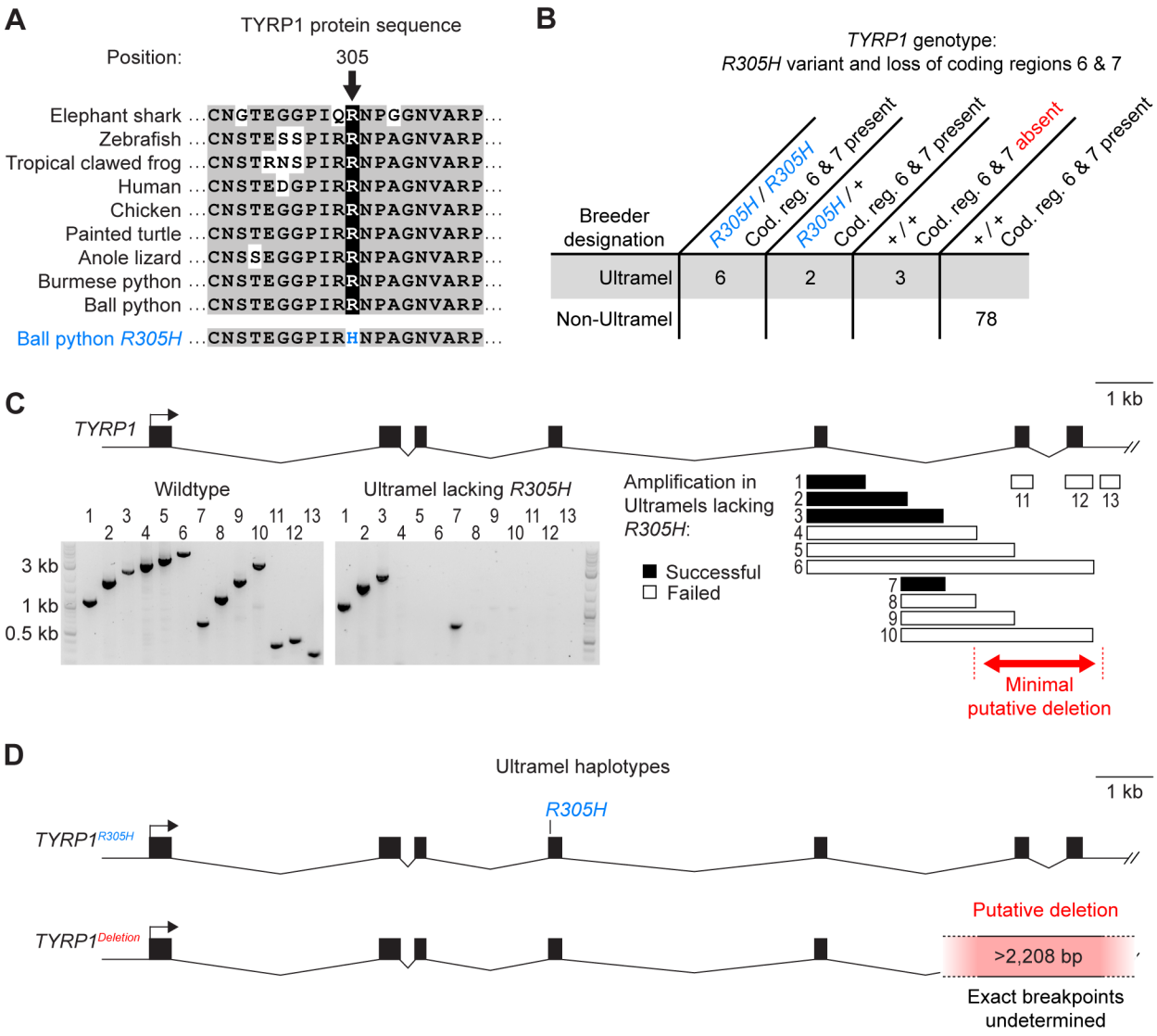

Fig 4. The Ultramel color morph is associated with a missense variant and a putative deletion in *TYRP1*. (A) Alignment of *TYRP1* protein sequence surrounding missense variant *R305H*. (B) Genotypes of 11 Ultramels and 78 Non-Ultramels. This set of animals includes the original Ultramel animal in which the *R305H* variant was identified. Ultramels heterozygous for *R305H* are presumed to be heterozygous for the putative deletion of coding regions 6 and 7; testing these animals for the putative deletion was not possible because the putative deletion cannot be detected in the heterozygous state. (C) Top, schematic of the *TYRP1* gene. Bottom left, PCR amplifications demonstrating the putative deletion in the Ultramels lacking *R305H*. Bottom right, alignment of PCR amplicons to the *TYRP1* gene. (D) Schematic of the two *TYRP1* haplotypes found in Ultramels. (A, D) Hash mark, discontinuity in the Burmese python reference genome.

We hypothesized that the *R305H* variant represents one of two loss-of-function alleles of *TYRP1.* Under this scenario, Ultramels lacking the *R305H* variant are predicted to be homozygous for an alternate loss-of-function allele. Ultramels heterozygous for *R305H* are predicted to be compound heterozygotes. To test this prediction, we amplified and sequenced coding regions 1 through 5 of *TYRP1* in one of the three Ultramels lacking the *R305H* variant. We found no coding variants and no splice-site variants compared to wildtype. We attempted to repeat this analysis for coding regions 6 and 7 of *TYRP1*, but we were unable to amplify these coding regions from this animal (Fig 4C). Identical results were observed for the other two Ultramels lacking the *R305H* variant (S5 Table). Further test amplifications indicated that coding regions 6 and 7 were missing from the genomes of all three animals (Fig 4C, S5 Table). Loss of these coding regions was specific to Ultramels lacking the *R305H* variant and did not occur in any of the Ultramels carrying the *R305H* variant, nor in any of the 78 Non-Ultramels (Fig 4B). A simple explanation of this pattern is that the Ultramels lacking the *R305H* variant were homozygous for a deletion spanning coding regions 6 and 7 of *TYRP1*. This putative deletion removes at least 2,208 bp from the genome and 117 amino acids from the TYRP1 protein, including the second of two zinc-binding domains. (The exact size and breakpoints of the putative deletion remain undetermined because the deletion extends into a discontinuity in the Burmese python reference genome.) Truncations occurring at similar positions in the TYRP1 protein have been identified in humans and are associated with oculocutaneous albinism [104]. We conclude that the Ultramel color morph is likely caused by variants in *TYRP1* and has two alleles: missense variant *R305H* and a putative deletion of coding regions 6 and 7 (Fig 4D).

Discussion

The goal of our study was to use pet samples recruited from the community to identify the genetic causes of the Albino, Lavender Albino, and Ultramel color morphs. We succeeded in recruiting 11 or more animals for each morph, along with a larger number of animals not belonging to these morphs. This sample size, albeit small, was sufficient to identify putatively causal variants for each morph (Fig 5). The Albino color morph was associated with three alleles of TYR: missense variant *D394G*, missense variant *P384L*, and haplotype TYR^{Albino}, which lacks coding or splice-site variants compared to wildtype. The Lavender Albino color morph was associated with a single allele of *OCA2*, a 1,514-bp deletion that removes *OCA2* coding region 18. The Ultramel color morph was associated with two alleles of *TYRP1*: missense variant *R305H* and a putative deletion that removes *TYRP1* coding regions 6 and 7. Due to the small sample size of our study, we cannot exclude the possibility of additional loss-of-function alleles of these genes. However, such alleles are expected to be at low frequency in the ball python population, given their absence from our sample. These findings are consistent with genetic data from other vertebrates (Table 1), indicating that the loss-of-function phenotypes of these genes range from severe (*TYR*) to moderate (*OCA2*) to mild (*TYPR1*). Our study demonstrates that pet ball pythons are a tractable resource for genetic analysis of coloration in reptiles, at least for color morphs having obvious candidate genes.

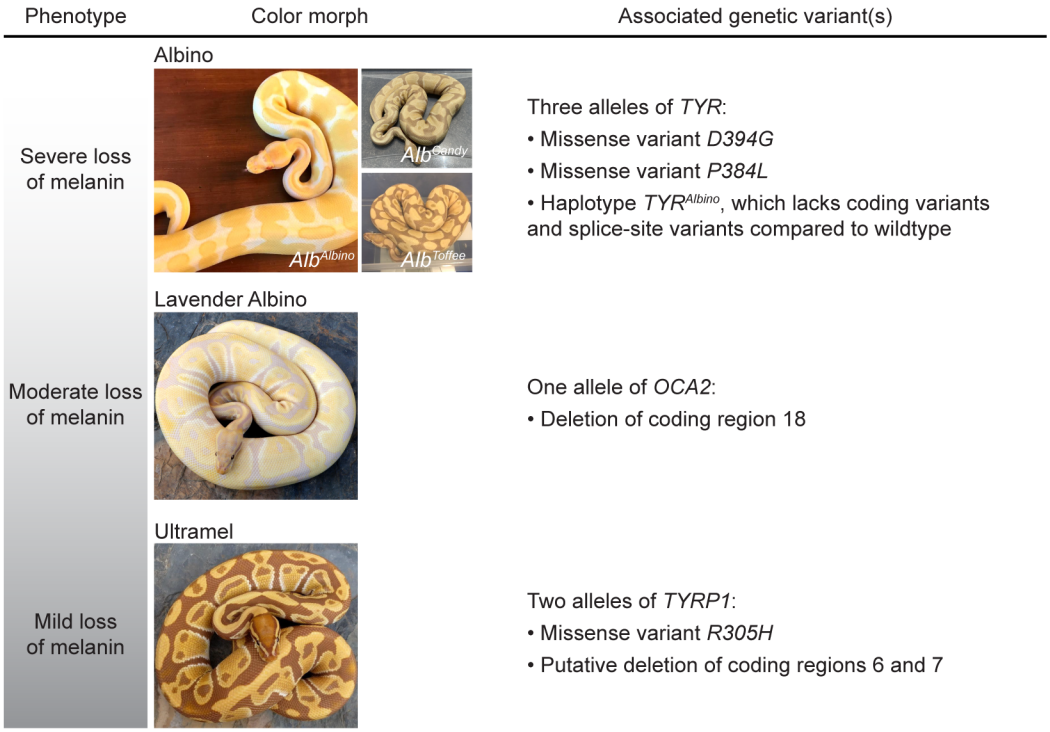

Fig 5. Summary of genetic variants associated with the Albino, Lavender Albino, and Ultramel color morphs. Photo credits, Ryan Young of Molecular Reptile, Chiron Graves, Phil Barclay, Michael Freedman of The Florida Reptile Ranch.

Molecular functions of genetic variants associated with color morphs

The missense variants and deletions found in *TYR*, *TYRP1*, and *OCA2* are likely hypomorphic or null alleles. The deletion in *OCA2* likely results in a frameshifted transcript that, if not degraded by nonsense-mediated decay, encodes a protein lacking six of OCA2's 12 transmembrane helices [94]. This truncated protein is therefore unlikely to retain the Cl^- channel activity observed for wildtype OCA2 [87]. *TYR* and *TYPR1* encode globular enzymes that require copper or zinc as co-factors [86]. Variants *P384L* and *D394G* in *TYR* reside in the second of TYR's two copper-binding domains and may therefore alter the ability of the TYR to bind copper. Variant *R305H* in *TYRP1* occurs at a site that normally forms salt bridges with residues located elsewhere in the peptide chain [86]. An arginine-to-histidine substitution at this site likely disrupts these salt bridges and may therefore interfere with proper protein folding. The putative deletion in *TYRP1* removes the C-terminal ~20% of the protein. This deleted region includes the second of TYRP1's two zinc-binding domains and several alpha helices residing to the enzyme's hydrophobic core [86]. Truncated TYRP1 protein is therefore unlikely to fold properly, nor is it likely to properly associate with zinc.

The TYR^{Albino} haplotype is the most common *TYR* haplotype found among Albinos, but it lacks an obvious loss-of-function variant. This haplotype contains no coding variants, no splice-site variants, no derived variants within 2 kb upstream of the start codon, and no large indels in three of four introns (with the exception of a longer allele of a VNTR, which is not specific to the TYR^{Albino} haplotype). We propose that the TYR^{Albino} haplotype contains a loss-of-function variant

not detectable by methods used in the current study. Examples include regulatory variants farther upstream of the start codon, substitutions in introns that disrupt slicing, or large indels or rearrangements involving the first intron. Cryptic loss-of-function variants in melanogenesis genes are thought to be relatively common in humans, where ~10-20% of oculocutaneous albinism patients are heterozygous for pathogenic variants in a known melanogenesis gene but lack a coding or splice-site variant on the opposite allele [105–107].

Multiple alleles of the Albino color morph

Prior to the current study, the Albino color morph was recognized by breeders as having three alleles: Alb^{Albino}, Alb^{Candy}, and Alb^{Toffee}. Alb^{Albino} was thought to confer a more severe phenotype than Alb^{Candy} and Alb^{Toffee} (Figs 1D-F). Our results confirm that the Albino color morph has three molecular alleles (TYR^{D394G}, TYR^{P384L}, and TYR^{Albino}), but these alleles do not correspond to the alleles recognized by breeders. The molecular allele TYR^{P384L} typically corresponds to the breeder designations Alb^{Candy} and Alb^{Toffee}, whereas the molecular alleles TYR^{D394G} and TYR^{Albino} typically correspond to the breeder designation Alb^{Albino} (Fig 2C). This correspondence is imperfect and exceptions exist (Fig 2C).

We speculate that the TYR^{P384L} haplotype was discovered twice and was named Alb^{Candy} by one breeder and Alb^{Toffee} by another breeder. We speculate that TYR^{D394G} and TYR^{Albino} confer similar phenotypes and were not previously recognized by breeders as distinct. The less severe phenotype associated with TYR^{P384L} may reflect the *P384L* variant being less disruptive to gene function compared with variants on the other two alleles. Alternatively, the less severe phenotype associated with TYR^{P384L} may reflect genetic linkage to modifiers in other genes. In either case, our study demonstrates that the Albino allele designations current in use by breeders are not an accurate reflection of molecular genotype. Renaming of the Albino alleles is warranted, although owners and breeders may be resistant to renaming due to cultural attachment to existing allele names.

Multiple alleles of the Ultramel color morph

Our finding of multiple *TYRP1* alleles for the Ultramel color morph was unexpected. This morph was not previously described as having multiple alleles. We speculate that one of the two alleles associated with the Ultramel color morph may represent an allele originally associated with a morph known as Caramel Albino. Multiple lineages of Caramel Albinos have been described, and their coloration similar to Ultramels. Caramel Albinos have been disfavored among owners and breeders because of spinal kinking and reduced female fertility. Caramel Albinos were not included in our study because we were unable to recruit any Caramel Albino samples. We speculate that the Caramel Albino morph may be allelic with Ultramel, and that one of two *TYRP1* alleles associated with Ultramel may represent an allele originally described as Caramel Albino. Alternately, Ultramel may be distinct from Caramel Albino, and the two *TYRP1* alleles may represent two distinct origins of the Ultramel morph.

Prospects for genetic testing

Many breeders of ball pythons wish to identify heterozygotes of recessive color morphs. Currently the only tool for identifying heterozygotes is test breeding, which is slow and laborious. The results of the current study will enable simple genetic testing for Albino, Lavender Albino, and Ultramel, which will allow heterozygotes to be identified prior to reproductive maturity. Testing for Lavender Albino can be performed by genotyping the *OCA2* deletion. Testing for Albino can be performed by sequencing the *TYR* missense variants *D394G* and *P384L*, and by genotyping variants that distinguish the TYR^{Albino} haplotype from other *TYR* haplotypes. Detection of *D394G* or *P384L* can be considered diagnostic for Albino because these variants are likely causative for the Albino color morph. Detection of the TYR^{Albino} haplotype is less diagnostic because the

causative variant on *TYR*^{Albino} haplotype remains unknown. Testing for one of the two Ultramel alleles can be performed by sequencing the *TYRP1* missense variant *R305H*. The other Ultramel allele—the putative deletion of *TYRP1* coding regions 6 and 7—cannot currently be detected in heterozygotes using simple methods because the breakpoints of this putative deletion are unknown.

Use of Burmese python reference genome

One challenge for genetic studies in ball python is the absence of a reference genome. To fill this gap, we relied on the genome of Burmese python [90], which is a scaffold-level assembly. Contigs in this assembly are small, and genes are sometimes fragmented across scaffolds. This fragmentation limited our analysis of the *TYR*^{Albino} haplotype and the putative deletion in *TYRP1*. For *TYR*^{Albino}, we were unable to assess whether the *TYR*^{Albino} haplotype contained a large indel or rearrangement in the first intron of *TYR*, due to a discontinuity in the Burmese python genome in this region. For the putative deletion in *TYRP1*, we were unable to map the exact breakpoints of the deletion, due to a discontinuity in the Burmese python genome immediately downstream of *TYRP1*. As genomic resources for Burmese python and ball python expand, we expect to identify a putatively causative variant on the *TYR*^{Albino} haplotype and to map the breakpoints of the putative deletion in *TYRP1.* This information will increase the prospects for genetic testing for these alleles.

Ball pythons as a genetic system

The market for pet ball pythons is huge, and many owners prefer animals with novel color patterns. This demand has led breeders to propagate genetic variants affecting coloration [31,33]. This effort has fortuitously created a collection of "mutants" useful for understanding the genetics of coloration in reptiles. Examples include ball python morphs with altered red-to-yellow coloration (e.g. Axanthic) and morphs in which the normal mottled color pattern is converted to dorsal strips (e.g. Clown, Genetic Stripe, Super Stripe). Red-to-yellow pigments are largely uncharacterized in reptiles [although see 28,29], and morphs affecting these pigments may provide insight into the metabolism and storage of these pigments. Morphs with stripes are reminiscent of evolutionary changes in color patterning across snake species [108] and may provide insights into the developmental mechanisms by which these changes evolve. Future studies of ball python color morphs will be aided by our groundwork showing that ball python samples can be recruited effectively from pet owners, and that genetic analyses in ball python can be scaffolded using the genome of Burmese python as a reference. Similar groundwork has also been provided by a recent study characterizing the ball python morph known as Piebald [26]. We expect that a continued community-science approach will be effective in developing ball pythons into a resource for understanding the genetics of reptile coloration.

Materials and methods

Recruitment of ball python sheds

Ball python sheds were recruited from pet owners and breeders by placing announcements in Twitter, Reddit, Instagram, and Facebook, and by contacting sellers having active listings on Morph Market (www.morphmarket.com). Contributors were instructed to allow sheds to dry (if wet) and to ship sheds via standard first-class mail. Contributors sending multiple sheds were instructed to package sheds individually in plastic bags during shipping. Contributors were not provided monetary compensation for sheds, although some contributors were given pre-paid shipping envelopes to cover shipping costs. Contributors were thanked via social media

473 whenever possible. Upon receipt, sheds were stored at -20°C to kill any insect larvae infesting
474 the sheds.
475 To maximize genetic diversity within each category of morph, we limited our sample of
476 animals of any one morph to animals contributed by different contributors. Exceptions were made
477 if animals contributed by the same contributor had been obtained from different breeders. The
478 goal of this strategy was to reduce the number of animals that were close relatives (e.g. siblings
479 or parent-offspring pairs). Obtaining full pedigree information was not possible because most
480 owners lacked this information (e.g. “I bought my animal at a pet store” or “I got my animal from
481 a friend who was moving away”). Our discovery of multiple alleles for the Albino and Ultramel
482 color morphs supports the idea that animals in our sample were derived from multiple lineages.
483 However, we cannot exclude the possibility that some animals in our sample may have been
484 close relatives.
485 The total set of animals comprised 50 Albinos, 14 Lavender Albinos, 11 Ultramels, one
486 animal described as heterozygous for Lavender Albino, and 46 animals described as having
487 normal coloration (i.e. wildtype) or belonging to color morphs other than Albino, Lavender Albino,
488 or Ultramel. Phenotypes of Albinos, Lavender Albinos, and Ultramels were confirmed by
489 examining shed skins for a reduction of brown-to-black coloration. The single animal described
490 as heterozygous for Lavender Albino was included to demonstrate the three-primer assay for
491 genotyping the OCA2 deletion. Heterozygotes for Albino or Ultramel and additional heterozygous
492 for Lavender Albino were excluded from our study because we have found that animals described
493 as heterozygous for recessive traits by breeders and owners are sometimes not heterozygous for
494 these traits (e.g. due to mis-attribution of paternity).

495 Performing experiments in an undergraduate laboratory 496 course

497 The majority of experiments and analyses described in this study were performed by
498 undergraduate students as part of a laboratory course at Eastern Michigan University (BIO306W).
499 This practice required that our experimental design rely on simple techniques, namely PCR and
500 Sanger sequencing. To avoid student errors in these techniques, we implemented the following
501 precautions. First, students never extracted DNA from more than one animal within the same
502 laboratory period. Second, students performed negative and positive control reactions for all PCR
503 amplifications. Data from students having incorrect controls were excluded from analysis. Third,
504 all sequence analyses were performed independently by three or more students. When the results
505 of all students did not all agree, sequences were re-analyzed by the instructor (HSS).

506 Annotation of melanogenesis genes in Burmese python

507 Our analyses in ball python used the genome of Burmese python as a reference. We
508 therefore required high-confidence gene annotations in Burmese python for the genes used in
509 our study (*TYR*, *TYRP1*, *TYRP2*, *OCA2*, *SLC7A11*, *SLC24A5*, and *SLC45A2*). Preliminary
510 inspection of existing gene annotations in Burmese python suggested that many gene annotations
511 contained errors. A main cause of these errors was that the Burmese python genome
512 (*Python_molurus_bivittatus-5.0.2*) is a scaffold-level assembly; many genes were split across
513 scaffolds. We therefore curated new gene annotations in Burmese python, using conservation of
514 gene structure across species. Alignment of protein sequences for each gene from corn snake,
515 anole lizard, chicken, and/or mouse revealed that gene structure was highly conserved for all
516 seven genes: Genes in each species contained the same number of coding regions, and the
517 coding-region boundaries relative to protein sequences were perfectly conserved across species,
518 with the exception of one slight difference in mouse for one boundary of *SLC45A2* (S2 Table).
519 We therefore felt confident that conservation of coding-region boundaries could be used to curate

new gene annotations in Burmese python. While it remains theoretically possible that our new gene annotations contain minor errors, our confidence in these gene annotations is high, given the high conservation of gene structure across species.

To curate new gene annotations, we performed TBLASTN searches of the Burmese python genome (*Python_molurus_bivittatus-5.0.2*) using protein sequences from anole lizard or corn snake as the query. Hits were examined manually. Gene annotations were built to match the coding-region boundaries conserved across species. Details for each gene annotation are given below.

TYR

The Burmese python genome was queried using TYR protein sequence from anole lizard (*XP_003219419.1*). The N-terminus of this query (coding region 1) hit Burmese python transcript *XM_007438960.1* on scaffold 4418. The C-terminus of this query (coding regions 2 to 5) hit Burmese python transcript *XM_007436041.2* on scaffold 3103. TYR was annotated as the union of these transcripts. The 3' boundary of coding region 1 was adjusted to match the boundary conserved across species (…TTC TCT TCA TGG CAA-3').

TYRP1

The Burmese python genome was queried using TYRP1 protein sequence from corn snake (*XP_034266320.1*). The N-terminus of this query (coding regions 1 to 5) hit Burmese python transcript *XM_007426971.3* on scaffold 801. The C-terminus of this query (coding regions 6 and 7) hit unannotated regions on this same scaffold. *TYRP1* was annotated as the union of transcript *XM_007426971.3* and coding regions 6 and 7. The boundaries of coding regions 6 and 7 were built to match the boundaries conserved across species (coding region 6, 5'-ATA TCT CAA CAT ACC....AAG TTC AGT GGC CAT-3'; coding region 7, 5'-CAC AAG CTC TCC ATG....CAG TCA GAT GTG TGA-3').

TYRP2

The Burmese python genome was queried using TYRP2 protein sequence from corn snake (*XP_034273310.1*). The N-terminus of this query (coding regions 1 to 3) hit unannotated regions on Burmese python scaffold 4970. The C-terminus of this query (coding regions 4 to 8) hit Burmese python transcript *XM_025174219.1* on this same scaffold. *TYRP2* was annotated as the union of coding regions 1 to 3 and C-terminal five coding regions of transcript *XM_025174219.1*. Transcript *XM_025174219.1* contains a sixth coding region that does not match *TYRP2* and was therefore excluded from the gene annotation. The boundaries of coding regions 1 to 3 were built to match the boundaries conserved across species (coding region 1, 5'-ATG GCC TTC CTG CTG....GTT GCC AAT GCA CAG-3'; coding region 2, 5'-GAC ATT TTG CTG GCT....GAG ATA CTC TAT TAG-3'; coding region 3, 5'- GAC CAG GCC GTC CCT....GAA AGA GAT CTG CAG-3').

OCA2

The Burmese python genome was queried using OCA2 protein sequence from corn snake (*XP_034287267.1*). The N-terminus of this query (coding regions 1 to 9) hit Burmese python transcript *XM_025173964.1* on scaffold 4704. The C-terminus of this query (coding regions 10 to 24) hit Burmese python transcript *XM_007433276.1* on scaffold 2194. *OCA2* was annotated as the union of these transcripts. The 5' boundary of coding region 10 was adjusted to match the boundary conserved across species (5'-ATT GTC CAC AGG ACA…).

SLC7A11

The Burmese python genome was queried using SLC7A11 protein sequence from corn snake (*XP_034257397.1*). All regions of the query hit Burmese python transcript *XM_007430906.3* on scaffold 1543. *SLC7A11* was annotated as transcript *XM_007430906.3*, with no further adjustments.

SLC24A5

The Burmese python genome was queried using SLC24A5 protein sequence from corn snake (*XP_034290605.1*) extended at its N-terminus to an upstream start codon present in the parent transcript (*XM_034434714.1*). The N-terminus of the query (coding region 1) hit an unannotated region on Burmese python scaffold 3984. The C-terminus of this query (coding regions 2 to 9) hit Burmese python transcript *XM_007438007.2* on this same scaffold. *SLC24A5* was annotated as the union of coding region 1 and transcript *XM_007438007.2*. The boundaries of coding region 1 were built to match the boundaries conserved across species (5'-ATG CAG CCT GCC GAG....TCC GCG AGG ATC CCG -3'). The 5' boundary of coding region 2 was adjusted to match the boundary conserved across species (5'-AGA ACG AAA CCC GCT...).

SLC45A2

The Burmese python genome was queried using *SLC45A2* protein sequence from corn snake (*XP_034298386.1*). The N-terminal region of this query (coding regions 1 and 2) hit Burmese python transcript *XM_007432459.2* on scaffold 1939. The C-terminal region of this query (coding regions 5 to 7) hit Burmese python transcript *XM_007437721.2* on scaffold 3858. *SLC45A2* was annotated as the union of these transcripts.

DNA extraction

Sheds were rinsed in tap water to remove dirt and debris. Sheds were air dried and lysed overnight at ~60˚C in ~1 ml lysis buffer (100 mM Tris-HCl pH 8.0, 100 mM EDTA, 2% sodium dodecyl sulfate, 3 mM $CaCl_2$, 2 mg/ml Proteinase K) per ~8 cm^2 piece of shed. Lysate was separated from visible fragments of undigested shed and further cleared by centrifugation at 13,000 x g for 2 min. To precipitate protein, ammonium acetate was added to supernatant to a final concentration of 1.875 M. Samples were incubated on ice for 5-10 min and centrifuged at 13,000 x g for 3-5 min at 4˚C. Supernatant was mixed with an equal volume of magnetic bead mixture (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 1.6 M NaCl, 0.2% Tween-20, 11% polyethylene glycol, 0.04% washed SpeedBeads [Sigma #GE45152105050250]), and samples shaken for 5-10 min. Beads were separated from supernatant using a magnet, washed twice in 0.2 ml 70% ethanol for 2 min, and air dried for ~1 min. DNA was eluted from beads in TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) at 65°C for >5 min.

Primer design and PCR

Primers were designed against the genome of Burmese python or against genomic sequences from ball python obtained in an earlier step of the study. Primers were designed using Primer3 [109], using default parameters and a target annealing temperature of 60°C. Amplification was first tested at 57°C, to allow for occasional divergence between ball python and Burmese python genomic sequences. In some cases, annealing temperatures were later adjusted to 52°C, 60°C, or 61°C, to obtain stronger product or to eliminate background bands.

Genomic fragments were amplified using OneTaq polymerase (NEB #M0480) or Q5 polymerase (NEB #M0491). Genotyping assays described below used OneTaq, unless otherwise specified. OneTaq reactions consisted of 1X OneTaq Standard Reaction Buffer, 200 µM dNTPs, 0.2 µM of each primer, and 0.025 U/µl OneTaq polymerase. OneTaq thermocycling conditions were as follows: 94°C for 2 min; 30-35 cycles of 94°C for 30 sec, 52-61°C for 30 sec, and 68°C for 1-4 min; and 68°C for 5 min. Q5 reactions consisted of 1X Q5 Reaction Buffer, 200 µM dNTPs, 0.5 µM of each primer, and 0.02 U/µl Q5 polymerase. Q5 thermocycling conditions were as follows: 98°C for 30 sec; 30-35 cycles of 98°C for 10 sec, 58-61°C for 15 sec, and 72°C for 1.5-3 min; and 72°C for 5 min. Reactions used 10-100 ng template DNA per 20 µl volume.

613 # Sanger sequencing

614 PCR products were purified for Sanger sequencing using magnetic beads or gel
615 extraction. For magnetic-bead purification, PCR reactions were mixed with three volumes of
616 magnetic-bead mixture (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 1.6-2.5 M NaCl, 0.2% Tween-20,
617 11-20% polyethylene glycol, 0.04% washed SpeedBeads [Sigma #GE45152105050250]), and
618 agitated for 5 min. Beads were separated from supernatant using a magnet, washed twice in 0.2
619 ml 80% ethanol for >30 sec, and air-dried for 30 sec. PCR products were eluted from beads in 10
620 mM Tris-HCl pH 8.0 for >3 min at 65°C. Gel extraction was performed using QIAquick Gel
621 Extraction Kit (Qiagen #28704), according to the manufacturer guidelines. Sanger sequencing
622 was performed by Eton Bioscience Inc (etonbio.com).

623 # Sequencing of coding regions and comparison to Burmese
624 python

625 Coding regions of melanogenesis genes were amplified and sequenced using primers
626 given in S3 Table. Chromatograms were trimmed using SnapGene Viewer
627 (snapgene.com/snapgene-viewer) and aligned to one another or to genomic sequences from
628 Burmese python using ApE (jorgensen.biology.utah.edu/wayned/ape). Alignments were
629 examined manually to identify divergent and polymorphic sites. Sequence identity between ball
630 python and Burmese python was calculated across alignable sequence, excluding indels.

631 # Association study genotyping

632 Variants for the association study were identified through amplification and pilot
633 sequencing of genomic fragments from five animals (four Non-Albinos and one Albino). For *TYR*,
634 *TYRP1*, and *TYRP2*, we identified one or more genomic fragments containing multiple
635 polymorphic sites within the same amplicon. These genomic fragments were selected for the
636 association study. For *OCA2*, we did not identify any genomic fragments containing multiple
637 polymorphic sites; we therefore selected two genomic fragments each containing a single
638 polymorphic site. Divergence between the Albino animal and the other four animals was not a
639 criterion for inclusion of variants in the association study. Variants were genotyped by amplifying
640 and sequencing genomic fragments containing the variants. Locations of variants and primers
641 used for amplifying and sequencing variants are given in S1 Table. Genotypes are provided in S6
642 Table.

643 # Haplotype reconstruction

644 Haplotypes were reconstructed using PHASE version 2.1 [91,92]. Parameters were set to
645 200 iterations, a thinning interval of 2, and a burn-in parameter of 100. These settings produced
646 identical or nearly identical output for seven runs seeded with different random numbers; thus,
647 these settings met the criteria for effective choice of parameter settings, according to the PHASE
648 documentation). The case-control permutation test was performed by comparing Albinos to Non-
649 Albinos.

650 # Genotyping assays for *TYR*

651 *TYR* missense variants *D394G* and *P384L* were genotyped by amplifying and sequencing
652 a genomic fragment containing *TYR* coding region 3. This fragment was amplified using primers
653 13F (5'-ACT TTC AGG TGG GCA GCA G-3') and 13R (5'-GCT GAC AAC TAA AAT CTC TGC
654 AA-3') and an annealing temperature of 52°C. The amplicon was sequenced using primer 13F.
655 Genotypes are provided in S6 Table.

656 *TYR* promoter and intronic regions are amplified using primers given in S4 Table.
657 Promoter fragments were sequenced in full, and intronic fragments were sequenced from one
658 end, to confirm that the correct region of the genome had been amplified. Sequencing primers
659 are given in S4 Table. Size differences among intronic fragments were assessed by separating
660 fragments on a 1.25% agarose gel. Gels were run long enough for the shortest ladder band (100
661 bp) to migrate ~8 cm from its starting position.
662 Allele sizes of the variable number tandem repeat (VNTR) in *TYR* were genotyped by
663 amplifying a genomic fragment located in *TYR* intron 3, using primers 214F (5'-TCT CAC CTG
664 ATG GCA CAT TC-3') and 209R (5'-GTG CCC ACC CTG ATG TTA TT-3') and an annealing
665 temperature of 60°C. Amplicon sizes were analyzed as for intronic fragments, described above.

666 Genotyping assays for *OCA2*

667 The *OCA2* deletion was initially identified by amplifying and sequencing a genomic
668 fragment spanning *OCA2* coding region 18. This fragment was amplified using primers 218F (5'-
669 ACC CCG TAG CCT CTT CAA AT-3') and 166R (5'-TGG GTG GCA AAC AAT CAT AA-3'), an
670 annealing temperature of 60°C, and Q5 polymerase. Amplicons were sequenced using both
671 primers.
672 The *OCA2* deletion was genotyped after its initial identification using a three-primer PCR
673 assay. This assayed used one forward primer and two reverse primers. The forward primer and
674 one of the reverse primers were located outside the deletion: 217F (5'-GGA GAG AGA ATC CAA
675 CCC TTG -3') and 166R (5'-TGG GTG GCA AAC AAT CAT AA-3'). The second reverse primer
676 was located within the deletion: 188R (5'-CAA AGA CCA TTG TCC ATT TCC-3'). The annealing
677 temperature was 57°C. This assay produces a 429-bp product for the wildtype allele and a 349-
678 bp product for the deletion allele. Heterozygotes produce both products. Genotypes are provided
679 in S6 Table.

680 Genotyping assays for *TYRP1*

681 *TYRP1* missense variant *R305H* was genotyped by amplifying and sequencing a genomic
682 fragment spanning *TYRP1* coding region 4. This fragment was amplified using primers 18F (5'-
683 GCT CTT TTC TCT AAG TCT GAC CTC -3') and 18R (5'-TCT TGT CCC ACA AAA GGA TTT -
684 3') and an annealing temperature of 57°C. The amplicon was sequenced using primer 18F.
685 The putative deletion of *TYRP1* coding regions 6 and 7 was identified using primers given
686 in S5 Table. The putative deletion was genotyped after its initial identification by amplifying a
687 genomic fragment of *TYRP1* spanning coding regions 6 and 7. This fragment was amplified using
688 primers 20F (5'-GCA TTG TTT TAT CAG CCA TGA A-3') and 21R (5'-GGA ATT GAG ACA AAT
689 CCT TGG-3') and an annealing temperature of 57°C. Genotypes are provided in S6 Table.

690 Protein sequence alignment

691 Protein sequences were aligned using Clustal Omega [110], using default parameters.

Acknowledgements

We thank Matt Rockman and Katy Greenwald for advice on haplotype reconstruction; Bob Winning, Anne Casper, David Kass, and two anonymous reviewers for comments on the manuscript; and the Educational Course Support program of New England BioLabs for reagents used in undergraduate teaching labs. We thank the following individuals for contributing ball python sheds: Adam and Nicole Schmid; Alycia Butler; Amanda Hall; Andelyn Czajka; Brad Carter of Driftless Reptiles; Bryan Rivera; Chiron Graves; Chun Ku of Dynasty Reptiles; Dale Porcher; Daniel Ross; David Wolf of Tornado Alley Reptiles; David Burstein; Dawn and Kelsi Greene of Super Natural Balls; Dayna Plehn; Debby Brauer; Epic Vibrant Balls; Eric Chung of Chung Reptiles; Erin Burt; George Straub; Haily McCullough; Jaden Christensen; Jake Lewis; Jamie Palazzo of New Day Reptiles; Jeff Kearns; Jeff Linton; Jodi Wilkowski; Joe Myers; John Cordone of Blue Water Reptiles; Jordan Noland; Justin Kobylka of J. Kobylka Reptiles; Lindsay VanOrman; Lisa Huis; Manuel San Juan; Mark Bilger; Lynnet Melton; Maryann Barbon; Mia Hynes; Michael, Lisa, and Bodie Cole of Ballroom Pythons South; Morgan Evans and Michael Kitto of MK Pythons; Morgan Shelton; Paul and Amber Fiorito of Vivid Scales; Pets 'n' Things of Saline, MI; Rachel Voyt; Royal Black Balls; Ryan Boyd and Brittney Delacruz; Ryan Young of Molecular Reptile; Sergio McDole; Stephanie S. Crisp; Steve P. of Prime Pets; Zac Parpart; and several anonymous contributors. We thank the following individuals for providing images of color morphs: Beth Woodyard; Cat Church; Cormier Jason; Christine Miller; Darin Taylor; Daniel Hatcher; David C. Callahan; Donald Grinstead; Elijah Snyder; Innovative Ectotherms; Jake Lewis; Jessica Allison; Jessica Van Riper; Justin Kobylka; James Thompson; Justin Revington; Mariette van Vuuren; Mark Hopkins; Mark Smith; Matthew Lopez; Michael Freedman of The Florida Reptile Ranch; Morgan Evans; Phil Barclay, Robert Cooper; Ron Heisenberg; Ryan Young; Selectively Bred Serpents; Seb Des Légendes Celtiques; and Thananan Jivaramonaikul.

Most of the data in this study were collected by undergraduates enrolled in a laboratory course at Eastern Michigan University. These students constitute the BIO306W Consortium. These students were Alexandra Ernst, Alia Frederick, Alissa Zoltowski, Amber Northcutt, Andraya Ackerman, Anna Pathammavong, Annette Miller, Ashly Matzek, Asra Akhlaq, Aubrey Martin, Bailey Knight, Benjamin Huff, Beth Wasserman, Brian Donald Condron, Caleb Sommer, Cassandra Rigor, Charles Southwell, Chase Chitwood, Chelsea Brown, Christina Roka, Ciarra Womack, Clay McKenzie, Daniela Nappo, Darby Fracassa, Deirdre McCarter, Dhruvalkumar Patel, Dominic Paoletti, Drake Dzierwa, Erica R. Geml, Erin Bissett, Ezekiel Butcher, Garrett Chance, Garry Lewis, Genesis Garmendia, Geo Pullockaran, Hajer Musa Abuzir, Haley Praski, Hanan Alroaini, Iqra Akhlaq, Ismael Yasin, Janelle Aethyr, Janelle Janisse, Jayce Alee Perysian, Jemar Rooks, Jonathan Chang, Jonathan Harris, Joseph H. Oberlitner, Joshua Mason, Juwan Taylor, Kailynn Sparks, Karissa Urban, Karli Siefman, Kealy Szymanski, Kelsy Roque, Keyan Marshall, Khaled Ali, Karleigh Hassenzahl, Kylie Powrie, Lauren Colone, Lissette Rosas, Manoj Perumallapalli, Mariam Samir, Maryam Nimer, Maya Mackey, Megan McNulty, Mel Roberts, Micaela Schempf, Molly Cook, Myah Kelly, Nahiel Sukar, Natalie Diaz, Natasha MacKay, Nathan Barnett, Nathaniel Gonzalez, Noura Taybeh, Pablo De la Vega, Rida Ali, Ronnie Bryans, Ryan Elliott, Saja Hussein, Samantha Glowacki, Samuel Teener, Sarah Holtzen, Sarah Schmidt, Shanti Bernstein, Shelan Mizuree, Smarpita Singh, Stevie Zabrosky, Taia Broadbent, Tommiea Robertson, Tyler Schallhorn, Verginio Persicone, William Soder, Wolfgang Ebersole, and Yvette Campbell.

References

1. Cuthill IC, Allen WL, Arbuckle K, Caspers B, Chaplin G, Hauber ME, et al. The biology of color. Science. 2017;357: eaan0221. doi:10.1126/science.aan0221

2. Cieslak M, Reissmann M, Hofreiter M, Ludwig A. Colours of domestication. Biol Rev. 2011;86: 885–899. doi:10.1111/j.1469-185X.2011.00177.x

3. Toews DPL, Hofmeister NR, Taylor SA. The Evolution and Genetics of Carotenoid Processing in Animals. Trends Genet. 2017;33: 171–182. doi:10.1016/j.tig.2017.01.002

4. Prum RO. Anatomy, physics, and evolution of avian structural colors. Bird Coloration. Cambridge, MA: Harvard University Press; 2006.

5. Ghiradella H. Structure and Development of Iridescent Lepidopteran Scales - the Papilionidae as a Showcase Family. Ann Entomol Soc Am. 1985;78: 252–264. doi:10.1093/aesa/78.2.252

6. Bagnara JT, Taylor JD, Hadley ME. The dermal chromatophore unit. J Cell Biol. 1968;38: 67–79. doi:10.1083/jcb.38.1.67

7. Saenko SV, Teyssier J, van der Marel D, Milinkovitch MC. Precise colocalization of interacting structural and pigmentary elements generates extensive color pattern variation in Phelsuma lizards. BMC Biol. 2013;11: 105. doi:10.1186/1741-7007-11-105

8. Kinoshita S, Yoshioka S. Structural colors in nature: The role of regularity and irregularity in the structure. ChemPhysChem. 2005;6: 1442–1459. doi:10.1002/cphc.200500007

9. Slominski A, Tobin DJ, Shibahara S, Wortsman J. Melanin pigmentation in mammalian skin and its hormonal regulation. Physiol Rev. 2004;84: 1155–1228. doi:10.1152/physrev.00044.2003

10. Meredith P, Sarna T. The physical and chemical properties of eumelanin. Pigm Cell Res. 2006;19: 572–594. doi:10.1111/j.1600-0749.2006.00345.x

11. Bagnara JT, Fernandez PJ, Fujii R. On the blue coloration of vertebrates. Pigm Cell Res. 2007;20: 14–26. doi:10.1111/j.1600-0749.2006.00360.x

12. Olsson M, Stuart-Fox D, Ballen C. Genetics and evolution of colour patterns in reptiles. Semin Cell Dev Biol. 2013;24: 529–541. doi:10.1016/j.semcdb.2013.04.001

13. Shawkey MD, D'Alba L. Interactions between colour-producing mechanisms and their effects on the integumentary colour palette. Philos Trans R Soc B-Biol Sci. 2017;372: 20160536. doi:10.1098/rstb.2016.0536

14. Bagnara JT, Matsumoto J. Comparative Anatomy and Physiology of Pigment Cells in Nonmammalian Tissues. The Pigmentary System. John Wiley & Sons, Ltd; 2006. pp. 11–59. doi:10.1002/9780470987100.ch2

15. Singh AP, Nuesslein-Volhard C. Zebrafish Stripes as a Model for Vertebrate Colour Pattern Formation. Curr Biol. 2015;25: R81–R92. doi:10.1016/j.cub.2014.11.013

16. Huang D, Lewis VM, Foster TN, Toomey MB, Corbo JC, Parichy DM. Development and genetics of red coloration in the zebrafish relative Danio albolineatus. eLife. 2021;10: e70253. doi:10.7554/eLife.70253

17. Kimura T, Nagao Y, Hashimoto H, Yamamoto-Shiraishi Y, Yamamoto S, Yabe T, et al. Leucophores are similar to xanthophores in their specification and differentiation processes in medaka. Proc Natl Acad Sci U S A. 2014;111: 7343–7348. doi:10.1073/pnas.1311254111

18. Nusslein-Volhard C, Singh AP. How fish color their skin: A paradigm for development and evolution of adult patterns Multipotency, plasticity, and cell competition regulate proliferation and spreading of pigment cells in Zebrafish coloration. Bioessays. 2017;39: 1600231. doi:10.1002/bies.201600231

19. Cooke TF, Fischer CR, Wu P, Jiang T-X, Xie KT, Kuo J, et al. Genetic Mapping and Biochemical Basis of Yellow Feather Pigmentation in Budgerigars. Cell. 2017;171: 427-+. doi:10.1016/j.cell.2017.08.016

20. Gazda MA, Araujo PM, Lopes RJ, Toomey MB, Andrade P, Afonso S, et al. A genetic mechanism for sexual dichromatism in birds. Science. 2020;368: 1270-+. doi:10.1126/science.aba0803

21. Kwon YM, Vranken N, Hoge C, Lichak MR, Francis KX, Camacho-Garcia J, et al. Genomic consequences of domestication of the Siamese fighting fish. 2021 Apr p. 2021.04.29.442030. doi:10.1101/2021.04.29.442030

22. Lopes RJ, Johnson JD, Toomey MB, Ferreira MS, Araujo PM, Melo-Ferreira J, et al. Genetic Basis for Red Coloration in Birds. Curr Biol. 2016;26: 1427–1434. doi:10.1016/j.cub.2016.03.076

23. Mundy NI, Stapley J, Bennison C, Tucker R, Twyman H, Kim K-W, et al. Red Carotenoid Coloration in the Zebra Finch Is Controlled by a Cytochrome P450 Gene Cluster. Curr Biol. 2016;26: 1435–1440. doi:10.1016/j.cub.2016.04.047

24. Toomey MB, Lopes RJ, Araujo PM, Johnson JD, Gazda MA, Afonso S, et al. High-density lipoprotein receptor SCARB1 is required for carotenoid coloration in birds. Proc Natl Acad Sci U S A. 2017;114: 5219–5224. doi:10.1073/pnas.1700751114

25. Zhang W, Wang H, Brandt DYC, Hu B, Sheng J, Wang M, et al. The genetic architecture of phenotypic diversity in the betta fish (Betta splendens). bioRxiv. 2021; 2021.05.10.443352. doi:10.1101/2021.05.10.443352

26. Garcia-Elfring A, Roffey HL, Hendry AP, Barrett RDH. A nonsense mutation in TFEC is the likely cause of the recessive piebald phenotype in ball pythons (Python regius). 2020 Nov p. 2020.10.30.362970. doi:10.1101/2020.10.30.362970

27. Guo L, Bloom J, Sykes S, Huang E, Kashif Z, Pham E, et al. Genetics of white color and iridophoroma in “Lemon Frost” leopard geckos. PLoS Genet. 2021;17: e1009580. doi:10.1371/journal.pgen.1009580

28. McLean CA, Lutz A, Rankin KJ, Stuart-Fox D, Moussalli A. Revealing the Biochemical and Genetic Basis of Color Variation in a Polymorphic Lizard. Mol Biol Evol. 2017;34: 1924–1935. doi:10.1093/molbev/msx136

29. Ullate-Agote A, Burgelin I, Debry A, Langrez C, Montange F, Peraldi R, et al. Genome mapping of a LYST mutation in corn snakes indicates that vertebrate chromatophore vesicles are lysosome-related organelles. Proc Natl Acad Sci U S A. 2020;117: 26307–26317. doi:10.1073/pnas.2003724117

30. Bale R. Ball python exports raise concerns as demand for the popular pet grows. National Geographic. 2020.

31. Broghammer S. Python regius: Atlas of Colour Morphs Keeping and Breeding. NTV Natur und Tier-Verlag; 2019.

32. Irizarry KJL, Bryden RL. In Silico Analysis of Gene Expression Network Components Underlying Pigmentation Phenotypes in the Python Identified Evolutionarily Conserved Clusters of Transcription Factor Binding Sites. Adv Bioinformatics. 2016;2016: 1286510. doi:10.1155/2016/1286510

33. McCurley K. Complete Ball Python, A Comprehensive Guide to Care, Breeding, and Genetic Mutations: Kevin McCurley: 9780976733409: Amazon.com: Books. ECO / Serpent's Tale NHBD; 2005.

34. Burbrink FT, Castoe TA. Molecular Phylogeography of Snakes. Snakes: Ecology and Conservation. Cornell University Press; 2011. pp. 38–77. doi:10.7591/9780801459092-006

35. Liu Q, Qi Y, Liang Q, Song J, Liu J, Li W, et al. Targeted disruption of tyrosinase causes melanin reduction in Carassius auratus cuvieri and its hybrid progeny. Sci China-Life Sci. 2019;62: 1194–1202. doi:10.1007/s11427-018-9404-7

36. Koga A, Wakamatsu Y, Kurosawa J, Hori H. Oculocutaneous albinism in the i(6) mutant of the medaka fish is associated with a deletion in the tyrosinase gene. Pigm Cell Res. 1999;12: 252–258. doi:10.1111/j.1600-0749.1999.tb00758.x

37. Miura I, Tagami M, Fujitani T, Ogata M. Spontaneous tyrosinase mutations identified in albinos of three wild frog species. Genes Genet Syst. 2017;92: 189–196. doi:10.1266/ggs.16-00061

38. Iwanishi S, Zaitsu S, Shibata H, Nitasaka E. An albino mutant of the Japanese rat snake (Elaphe climacophora) carries a nonsense mutation in the tyrosinase gene. Genes Genet Syst. 2018;93: 163–167. doi:10.1266/ggs.18-00021

39. Tobita-Teramoto T, Jang GY, Kino K, Salter DW, Brumbaugh J, Akiyama T. Autosomal albino chicken mutation (ca/ca) deletes hexanucleotide (-deltaGACTGG817) at a copper-binding site of the tyrosinase gene. Poult Sci. 2000;79: 46–50. doi:10.1093/ps/79.1.46

40. Chang C-M, Coville J-L, Coquerelle G, Gourichon D, Oulmouden A, Tixier-Boichard M. Complete association between a retroviral insertion in the tyrosinase gene and the recessive white mutation in chickens. BMC Genomics. 2006;7: 19. doi:10.1186/1471-2164-7-19

849 41. Chang CM, Furet JP, Coville JL, Coquerelle G, Gourichon D, Tixier-Boichard M.
850 Quantitative effects of an intronic retroviral insertion on the transcription of the tyrosinase
851 gene in recessive white chickens. Anim Genet. 2007;38: 162–167. doi:10.1111/j.1365-
852 2052.2007.01581.x

853 42. Florisbal Dame MC, Xavier GM, Oliveira-Filho JP, Borges AS, Oliveira HN, Riet-Correa F,
854 et al. A nonsense mutation in the tyrosinase gene causes albinism in water buffalo. BMC
855 Genet. 2012;13: 62. doi:10.1186/1471-2156-13-62

856 43. Schmutz SM, Berryere TG, Ciobanu DC, Mileham AJ, Schmidtz BH, Fredholm M. A form
857 of albinism in cattle is caused by a tyrosinase frameshift mutation. Mamm Genome.
858 2004;15: 62–67. doi:10.1007/s00335-002-2249-5

859 44. Reiner G, Tramberend K, Nietfeld F, Volmer K, Wurmser C, Fries R, et al. A genome-wide
860 scan study identifies a single nucleotide substitution in the tyrosinase gene associated with
861 white coat colour in a red deer (Cervus elaphus) population. BMC Genet. 2020;21.
862 doi:10.1186/s12863-020-0814-0

863 45. Utzeri VJ, Bertolini F, Ribani A, Schiavo G, Dall'Olio S, Fontanesi L. The albinism of the
864 feral Asinara white donkeys (Equus asinus) is determined by a missense mutation in a
865 highly conserved position of the tyrosinase (TYR) gene deduced protein. Anim Genet.
866 2016;47: 120–124. doi:10.1111/age.12386

867 46. Yokoyama T, Silversides D, Waymire K, Kwon B, Takeuchi T, Overbeek P. Conserved
868 Cysteine to Serine Mutation in Tyrosinase Is Responsible for the Classical Albino Mutation
869 in Laboratory Mice. Nucleic Acids Res. 1990;18: 7293–7298. doi:10.1093/nar/18.24.7293

870 47. Blaszczyk WM, Arning L, Hoffmann KP, Epplen JT. A Tyrosinase missense mutation
871 causes albinism in the Wistar rat. Pigm Cell Res. 2005;18: 144–145. doi:10.1111/j.1600-
872 0749.2005.00227.x

873 48. Yu F, Jiao S, Lai W, Liu Z, Zhu M, Zhu W, et al. Conserved aspartate-to-glycine mutation
874 in tyrosinase is associated with albino phenotype in domestic guinea pigs (Cavia porcellus).
875 Anim Genet. 2018;49: 354–355. doi:10.1111/age.12683

876 49. Kim Y-H, Park S-J, Choe S-H, Lee J-R, Cho H-M, Kim S-U, et al. Identification and
877 characterization of the tyrosinase gene (TYR) and its transcript variants (TYR_1 and
878 TYR_2) in the crab-eating macaque (Macaca fascicularis). Gene. 2017;630: 21–27.
879 doi:10.1016/j.gene.2017.07.047

880 50. Galante Rocha de Vasconcelos FT, Hauzman E, Henriques LD, Kilpp Goulart PR, Galvao
881 O de F, Sano RY, et al. A novel nonsense mutation in the tyrosinase gene is related to the
882 albinism in a capuchin monkey (Sapajus apella). BMC Genet. 2017;18: 39.
883 doi:10.1186/s12863-017-0504-8

884 51. Koga A, Hisakawa C, Yoshizawa M. Baboon bearing resemblance in pigmentation pattern
885 to Siamese cat carries a missense mutation in the tyrosinase gene. Genome. 2020;63.
886 doi:10.1139/gen-2020-0003

887 52. Montoliu L, Gronskov K, Wei A-H, Martinez-Garcia M, Fernandez A, Arveiler B, et al.
888 Increasing the complexity: new genes and new types of albinism. Pigment Cell Melanoma
889 Res. 2014;27. doi:10.1111/pcmr.12167

890 53. Polanowski AM, Robinson-Laverick SM, Paton D, Jarman SN. Variation in the Tyrosinase
891 Gene Associated with a White Humpback Whale (Megaptera novaeangliae). J Hered.
892 2012;103: 130–133. doi:10.1093/jhered/esr108

893 54. Blaszczyk WM, Distler C, Dekomien G, Arning L, Hoffmann K-P, Epplen JT. Identification
894 of a tyrosinase (TYR) exon 4 deletion in albino ferrets (Mustela putorius furo). Anim Genet.
895 2007;38: 421–423. doi:10.1111/j.1365-2052.2007.01619.x

896 55. Anistoroaei R, Fredholm M, Christensen K, Leeb T. Albinism in the American mink
897 (Neovison vison) is associated with a tyrosinase nonsense mutation. Anim Genet. 2008;39:
898 645–648. doi:10.1111/j.1365-2052.2008.01788.x

899 56. Imes DL, Geary LA, Grahn RA, Lyons LA. Albinism in the domestic cat (Felis catus) is
900 associated with a tyrosinase (TYR) mutation. Anim Genet. 2006;37: 175–178.
901 doi:10.1111/j.1365-2052.2005.01409.x

902 57. Lyons LA, Imes DL, Rah HC, Grahn RA. Tyrosinase mutations associated with Siamese
903 and Burmese patterns in the domestic cat (Felis catus). Anim Genet. 2005;36: 119–126.
904 doi:10.1111/j.1365-2052.2005.01253.x

905 58. Schmidt-Kuntzel A, Eizirik E, O'Brien SJ, Menotti-Raymond M. Tyrosinase and tyrosinase
906 related protein 1 alleles specify domestic cat coat color phenotypes of the albino and brown
907 loci. J Hered. 2005;96: 289–301. doi:10.1093/jhered/esi066

908 59. Yan S, Zhao D, Hu M, Tan X, Lai W, Kang J, et al. A single base insertion in the tyrosinase
909 gene is associated with albino phenotype in silver foxes (Vulpes vulpes). Anim Genet.
910 2019;50: 550–550. doi:10.1111/age.12816

911 60. Gross JB, Wilkens H. Albinism in phylogenetically and geographically distinct populations
912 of Astyanax cavefish arises through the same loss-of-function Oca2 allele. Heredity.
913 2013;111: 122–130. doi:10.1038/hdy.2013.26

914 61. Kratochwil CF, Urban S, Meyer A. Genome of the Malawi golden cichlid fish
915 (Melanochromis auratus) reveals exon loss of oca2 in an amelanistic morph. Pigment Cell
916 Melanoma Res. 2019;32: 719–723. doi:10.1111/pcmr.12799

917 62. Saenko SV, Lamichhaney S, Barrio AM, Rafati N, Andersson L, Milinkovitch MC.
918 Amelanism in the corn snake is associated with the insertion of an LTR-retrotransposon in
919 the OCA2 gene. Sci Rep. 2015;5: 17118. doi:10.1038/srep17118

920 63. Caduff M, Bauer A, Jagannathan V, Leeb T. OCA2 splice site variant in German Spitz dogs
921 with oculocutaneous albinism. PLoS One. 2017;12: e0185944.
922 doi:10.1371/journal.pone.0185944

923 64. Zhang Y, Hong Q, Cao C, Yang L, Li Y, Hai T, et al. A novel porcine model reproduces
924 human oculocutaneous albinism type. Cell Discov. 2019;5: 48. doi:10.1038/s41421-019-
925 0117-7

926 65. Shoji H, Kiniwa Y, Okuyama R, Yang M, Higuchi K, Mori M. A nonsense nucleotide
927 substitution in the oculocutaneous albinism II gene underlies the original pink-eyed dilution
928 allele (Oca2(p)) in mice. Exp Anim. 2015;64: 171–179. doi:10.1538/expanim.14-0075

929 66. Krauss J, Geiger-Rudolph S, Koch I, Nuesslein-Volhard C, Irion U. A dominant mutation in
930 tyrp1A leads to melanophore death in zebrafish. Pigment Cell Melanoma Res. 2014;27:
931 827–830. doi:10.1111/pcmr.12272

932 67. Cortimiglia C, Castiglioni B, Pizzi F, Stella A, Capra E. Involvement of tyrosinase-related
933 protein 1 gene in the light brown plumage phenotype of Falco cherrug. Anim Genet.
934 2017;48: 125–126. doi:10.1111/age.12506

935 68. Ren J, Mao H, Zhang Z, Xiao S, Ding N, Huang L. A 6-bp deletion in the TYRP1 gene
936 causes the brown colouration phenotype in Chinese indigenous pigs. Heredity. 2011;106:
937 862–868. doi:10.1038/hdy.2010.129

938 69. Wu X, Zhang Y, Shen L, Du J, Luo J, Liu C, et al. A 6-bp deletion in exon 8 and two
939 mutations in introns of TYRP1 are associated with blond coat color in Liangshan pigs.
940 Gene. 2016;578: 132–136. doi:10.1016/j.gene.2015.12.011

941 70. Lyons LA, Foe IT, Rah HC, Grahn RA. Chocolate coated cats: TYRP1 mutations for brown
942 color domestic cats. Mamm Genome. 2005;16: 356–366. doi:10.1007/s00335-004-2455-4

943 71. Gratten J, Beraldi D, Lowder BV, McRae AF, Visscher PM, Pemberton JM, et al.
944 Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-
945 colour polymorphism in a free-living population of Soay sheep. Proc R Soc B-Biol Sci.
946 2007;274: 619–626. doi:10.1098/rspb.2006.3762

947 72. Paris JM, Letko A, Hafliger IM, Ammann P, Flury C, Drogemuller C. Identification of two
948 TYRP1 loss-of-function alleles in Valais Red sheep. Anim Genet. 2019;50: 778–782.
949 doi:10.1111/age.12863

950 73. Cirera S, Markakis MN, Kristiansen T, Vissenberg K, Fredholm M, Christensen K, et al. A
951 large insertion in intron 2 of the TYRP1 gene associated with American Palomino
952 phenotype in American mink. Mamm Genome. 2016;27: 135–143. doi:10.1007/s00335-
953 016-9620-4

954 74. Budd P, Jackson I. Structure of the Mouse Tyrosinase-Related Protein-2 Dopachrome
955 Tautomerase (tyrp2/Dct) Gene and Sequence of 2 Novel Slaty Alleles. Genomics. 1995;29:
956 35–43. doi:10.1006/geno.1995.1212

957 75. Jackson I, Chambers D, Tsukamoto K, Copeland N, Gilbert D, Jenkins N, et al. A 2nd
958 Tyrosinase-Related Protein, Trp-2, Maps to and Is Mutated at the Mouse Slaty Locus.
959 Embo J. 1992;11: 527–535. doi:10.1002/j.1460-2075.1992.tb05083.x

960 76. Chintala S, Li W, Lamoreux ML, Ito S, Wakamatsu K, Sviderskaya EV, et al. Slc7a11 gene
961 controls production of pheomelanin pigment and proliferation of cultured cells. Proc Natl
962 Acad Sci U S A. 2005;102: 10964–10969. doi:10.1073/pnas.0502856102

963 77. Lamason RL, Mohideen M, Mest JR, Wong AC, Norton HL, Aros MC, et al. SLC24A5, a
964 putative cation exchanger, affects pigmentation in zebrafish and humans. Science.
965 2005;310: 1782–1786. doi:10.1126/science.1116238

966 78. Winkler PA, Gornik KR, Ramsey DT, Dubielzig RR, Venta PJ, Petersen-Jones SM, et al. A
967 Partial Gene Deletion of SLC45A2 Causes Oculocutaneous Albinism in Doberman
968 Pinscher Dogs. PLoS One. 2014;9: e92127. doi:10.1371/journal.pone.0092127

969 79. Mack M, Kowalski E, Grahn R, Bras D, Penedo MCT, Bellone R. Two Variants in SLC24A5
970 Are Associated with "Tiger-Eye" Iris Pigmentation in Puerto Rican Paso Fino Horses. G3-
971 Genes Genomes Genet. 2017;7: 2799–2806. doi:10.1534/g3.117.043786

972 80. Fukamachi S, Shimada A, Shima A. Mutations in the gene encoding B, a novel transporter
973 protein, reduce melanin content in medaka. Nature Genet. 2001;28: 381–385.
974 doi:10.1038/ng584

975 81. Gunnarsson U, Hellstrom AR, Tixier-Boichard M, Minvielle F, Bed'hom B, Ito S, et al.
976 Mutations in SLC45A2 cause plumage color variation in chicken and Japanese quail.
977 Genetics. 2007;175: 867–877. doi:10.1534/genetics.106.063107

978 82. Mariat D, Taourit S, Guerin G. A mutation in the MATP gene causes the cream coat colour
979 in the horse. Genet Sel Evol. 2003;35: 119–133. doi:10.1051/gse:2002039

980 83. Newton JM, Cohen-Barak O, Hagiwara N, Gardner JM, Davisson MT, King RA, et al.
981 Mutations in the human orthologue of the mouse underwhite gene (uw) underlie a new form
982 of oculocutaneous albinism, OCA4. Am J Hum Genet. 2001;69: 981–988.
983 doi:10.1086/324340

984 84. Xu X, Dong G-X, Hu X-S, Miao L, Zhang X-L, Zhang D-L, et al. The Genetic Basis of White
985 Tigers. Curr Biol. 2013;23: 1031–1035. doi:10.1016/j.cub.2013.04.054

986 85. Prado-Martinez J, Hernando-Herraez I, Lorente-Galdos B, Dabad M, Ramirez O, Baeza-
987 Delgado C, et al. The genome sequencing of an albino Western lowland gorilla reveals
988 inbreeding in the wild. BMC Genomics. 2013;14: 363. doi:10.1186/1471-2164-14-363

989 86. Lai X, Wichers HJ, Soler-Lopez M, Dijkstra BW. Structure and Function of Human
990 Tyrosinase and Tyrosinase-Related Proteins. Chem-Eur J. 2018;24: 47–55.
991 doi:10.1002/chem.201704410

992 87. Bellono NW, Escobar IE, Lefkovith AJ, Marks MS, Oancea E. An intracellular anion channel
993 critical for pigmentation. eLife. 2014;3: e04543. doi:10.7554/eLife.04543

994 88. Ginger RS, Askew SE, Ogborne RM, Wilson S, Ferdinando D, Dadd T, et al. SLC24A5
995 encodes a trans-golgi network protein with potassium-dependent sodium-calcium
996 exchange activity that regulates human epidermal melanogenesis. J Biol Chem. 2008;283:
997 5486–5495. doi:10.1074/jbc.M707521200

998 89. Vitavska O, Wieczorek H. The SLC45 gene family of putative sugar transporters. Mol Asp
999 Med. 2013;34: 655–660. doi:10.1016/j.mam.2012.05.014

1000 90. Castoe TA, de Koning APJ, Hall KT, Card DC, Schield DR, Fujita MK, et al. The Burmese
1001 python genome reveals the molecular basis for extreme adaptation in snakes. Proc Natl
1002 Acad Sci U S A. 2013;110: 20645–20650. doi:10.1073/pnas.1314475110

1003 91. Stephens M, Scheet P. Accounting for decay of linkage disequilibrium in haplotype
1004 inference and missing-data imputation. Am J Hum Genet. 2005;76: 449–462.
1005 doi:10.1086/428594

1006 92. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction
1007 from population data. Am J Hum Genet. 2001;68: 978–989. doi:10.1086/319501

1008 93. Simeonov DR, Wang X, Wang C, Sergeev Y, Dolinska M, Bower M, et al. DNA Variations
1009 in Oculocutaneous Albinism: An Updated Mutation List and Current Outstanding Issues in
1010 Molecular Diagnostics. Hum Mutat. 2013;34: 827–835. doi:10.1002/humu.22315

1011 94. Gardner J, Nakatsu Y, Gondo Y, Lee S, Lyon M, King R, et al. The Mouse Pink-Eyed
1012 Dilution Gene - Association with Human Prader-Willi and Angelman Syndromes. Science.
1013 1992;257: 1121–1124. doi:10.1126/science.257.5073.1121

1014 95. Ni-Komatsu L, Orlow SJ. Heterologous expression of tyrosinase recapitulates the
1015 misprocessing and mistrafficking in oculocutaneous albinism type 2: Effects of altering
1016 intracellular pH and pink-eyed dilution gene expression. Exp Eye Res. 2006;82: 519–528.
1017 doi:10.1016/j.exer.2005.08.013

1018 96. Puri N, Gardner JM, Brilliant MH. Aberrant pH of melanosomes in pink-eyed dilution (p)
1019 mutant melanocytes. J Invest Dermatol. 2000;115: 607–613. doi:10.1046/j.1523-
1020 1747.2000.00108.x

1021 97. Lee S, Nicholls R, Bundey S, Laxova R, Musarella M, Spritz R. Mutations of the P-Gene in
1022 Oculocutaneous Albinism, Ocular Albinism, and Prader-Willi-Syndrome Plus Albinism. N
1023 Engl J Med. 1994;330: 529–534. doi:10.1056/NEJM199402243300803

1024 98. Sato H, Tamba M, Ishii T, Bannai S. Cloning and expression of a plasma membrane
1025 cystine/glutamate exchange transporter composed of two distinct proteins. J Biol Chem.
1026 1999;274: 11455–11458. doi:10.1074/jbc.274.17.11455

1027 99. Bin B-H, Bhin J, Yang SH, Shin M, Nam Y-J, Choi D-H, et al. Membrane-Associated
1028 Transporter Protein (MATP) Regulates Melanosomal pH and Influences Tyrosinase
1029 Activity. PLoS One. 2015;10: e0129273. doi:10.1371/journal.pone.0129273

1030 100. Sturm R, Osullivan B, Box N, Smith A, Smit S, Puttick E, et al. Chromosomal Structure of
1031 the Human Tyrp1 and Tyrp2 Loci and Comparison of the Tyrosinase-Related Protein Cone
1032 Family. Genomics. 1995;29: 24–34. doi:10.1006/geno.1995.1211

1033 101. Gershonibaruch R, Rosenmann A, Droetto S, Holmes S, Tripathi R, Spritz R. Mutations of
1034 the Tyrosinase Gene in Patients with Oculocutaneous Albinism from Various Ethnic-
1035 Groups in Israel. Am J Hum Genet. 1994;54: 586–594.

1036 102. Tripathi RK, Strunk KM, Giebel LB, Weleber RG, Spritz RA. Tyrosinase gene mutations in
1037 type I (tyrosinase-deficient) oculocutaneous albinism define two clusters of missense
1038 substitutions. Am J Med Genet. 1992;43: 865–871. doi:10.1002/ajmg.1320430523

1039 103. Tsai CH, Tsai FJ, Wu JY, Lin SP, Chang JG, Yang CF, et al. Insertion/deletion mutations
1040 of type I oculocutaneous albinism in chinese patients from Taiwan. Hum Mutat. 1999;14:
1041 542. doi:10.1002/(SICI)1098-1004(199912)14:6<542::AID-HUMU14>3.0.CO;2-3

1042 104. Forshew T, Khaliq S, Tee L, Smith U, Johnson CA, Mehdi SQ, et al. Identification of novel
1043 TYR and TYRP1 mutations in oculocutaneous albinism. Clin Genet. 2005;68: 182–184.
1044 doi:10.1111/j.1399-0004.2005.00460.x

1045 105. Lasseaux E, Plaisant C, Michaud V, Pennamen P, Trimouille A, Gaston L, et al. Molecular
1046 characterization of a series of 990 index patients with albinism. Pigment Cell Melanoma
1047 Res. 2018;31: 466–474. doi:10.1111/pcmr.12688

1048 106. Rooryck C, Morice-Picard F, Elcioglu NH, Lacombe D, Taieb A, Arveiler B. Molecular
1049 diagnosis of oculocutaneous albinism: new mutations in the OCA1-4 genes and practical
1050 aspects. Pigment Cell Melanoma Res. 2008;21: 583–587. doi:10.1111/j.1755-
1051 148X.2008.00496.x

1052 107. Rooryck C, Morice F, Lacombe D, Taieb A, Arveiler B. Genetic basis of oculocutaneous
1053 albinism. Expert Review of Dermatology. 2009;4: 611–622. doi:10.1586/edm.09.53

1054 108. Allen WL, Baddeley R, Scott-Samuel NE, Cuthill IC. The evolution and function of pattern
1055 diversity in snakes. Behav Ecol. 2013;24: 1237–1250. doi:10.1093/beheco/art058

1056 109. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3-
1057 new capabilities and interfaces. Nucleic Acids Res. 2012;40: e115. doi:10.1093/nar/gks596

1058 110. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation
1059 of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol.
1060 2011;7: 539. doi:10.1038/msb.2011.75

1061

1062 # Supporting information

1063 **S1 Fig. Raw gel images.** (A) Raw image of the gel displayed in Fig 2. (B) Raw image of the gel
1064 displayed in Fig 3. (C) Raw image of the gel displayed in Fig 4. Brightness and contrast settings
1065 have not been adjusted in these images. x, experiment unrelated to the current study.

1066 **S1 Table. Variants used in the Albino association study and primers to genotype these**
1067 **variants.**

1068 **S2 Table. Conservation of gene structure of melanogenesis genes.**

1069 **S3 Table. Primers to amplify and sequence coding regions of melanogenesis genes.**

1070 **S4 Table. Primers to amplify non-coding regions of *TYR*.**

1071 **S5 Table. Primers to investigate the putative deletion in *TYRP1*.**

1072 **S6 Table. Genotypes.**