Title

The Genetic Basis of Adaptation following Plastic Changes in Coloration in a Novel Environment

Journal

Current Biology, 28(18)

ISSN

0960-9822

Authors

Corl, Ammon

Bi, Ke

Luke, Claudia

et al.

Publication Date

2018-09-01

DOI

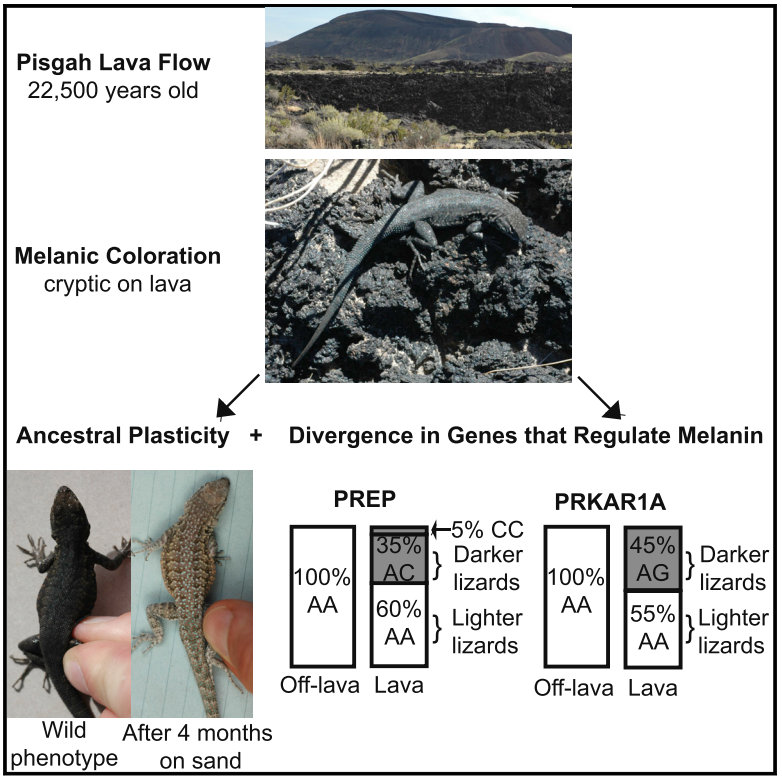
10.1016/j.cub.2018.06.075

Report

Current Biology

The Genetic Basis of Adaptation following Plastic Changes in Coloration in a Novel Environment

Graphical Abstract



Highlights

- Side-blotched lizards on a lava flow have melanic coloration that promotes crypsis
- Ancestral plasticity in coloration promotes background matching to the lava habitat
- Adaptive divergence in two genes that regulate melanin leads to darker lizards
- Plasticity can help survival in new habitats, where selection then refines phenotypes

Authors

Ammon Corl, Ke Bi, Claudia Luke, Akshara Sree Challa, Aaron James Stern, Barry Sinervo, Rasmus Nielsen

Correspondence

corlammon@gmail.com

In Brief

Plasticity has been hypothesized to precede adaptation to new environments. Corl et al. show that melanic lizards from a lava flow have both ancestral plasticity in coloration and divergence in two genes that increase pigmentation. Plasticity aided in the initial colonization of the new habitat, where natural selection then refined the phenotype.

Corl et al., 2018, Current Biology *28*, 2970–2977
September 24, 2018
https://doi.org/10.1016/j.cub.2018.06.075

CellPress

CellPress

Current Biology
Report

The Genetic Basis of Adaptation following Plastic Changes in Coloration in a Novel Environment

Ammon Corl,[1,2,7,*] Ke Bi,[1,3] Claudia Luke,[4] Akshara Sree Challa,[1] Aaron James Stern,[5,6] Barry Sinervo,[2] and Rasmus Nielsen[5]

[1]Museum of Vertebrate Zoology, University of California, Berkeley, 3101 Valley Life Sciences Building, Berkeley, CA 94720, USA
[2]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Santa Cruz, CA 95064, USA
[3]Computational Genomics Resource Laboratory (CGRL), California Institute for Quantitative Biosciences (QB3), University of California, Berkeley, Berkeley, CA 94720, USA
[4]Center for Environmental Inquiry, Sonoma State University, 1801 East Cotati Ave, Rohnert Park, CA 94928, USA
[5]Department of Integrative Biology, University of California, Berkeley, 3040 Valley Life Sciences Building # 3140, Berkeley, CA 94720-3140, USA
[6]Graduate Group in Computational Biology, University of California, Berkeley, Berkeley, CA 94720, USA
[7]Lead Contact
*Correspondence: corlammon@gmail.com
https://doi.org/10.1016/j.cub.2018.06.075

SUMMARY

Phenotypic plasticity has been hypothesized to precede and facilitate adaptation to novel environments [1–8], but examples of plasticity preceding adaptation in wild populations are rare (but see [9, 10]). We studied a population of side-blotched lizards, *Uta stansburiana*, living on a lava flow that formed 22,500 years ago [11] to understand the origin of their novel melanic phenotype that makes them cryptic on the black lava. We found that lizards living on and off of the lava flow exhibited phenotypic plasticity in coloration but also appeared to have heritable differences in pigmentation. We sequenced the exomes of 104 individuals and identified two known regulators of melanin production, PREP and PRKAR1A, which had markedly increased levels of divergence between lizards living on and off the lava flow. The derived variants in PREP and PRKAR1A were only found in the lava population and were associated with increased pigmentation levels in an experimental cohort of hatchling lizards. Simulations suggest that the derived variants in the PREP and PRKAR1A genes arose recently and were under strong positive selection in the lava population. Overall, our results suggest that ancestral plasticity for coloration facilitated initial survival in the lava environment and was followed by genetic changes that modified the phenotype in the direction of the induced plastic response, possibly through *de novo* mutations. These observations provide a detailed example supporting the hypothesis that plasticity aids in the initial colonization of a novel habitat, with natural selection subsequently refining the phenotype with genetic adaptations to the new environment.

RESULTS

Side-blotched lizards living on the Pisgah lava flow (Figure S1) exhibit melanic coloration, which promotes crypsis when the lizards are on the black lava (Figure 1A). Nearby populations living on light-colored soils are not melanic (Figure 1B). Phenotypic plasticity substantially contributes to the melanic phenotype, because Pisgah lizards housed on light-colored sand develop a substantially lighter phenotype (Figures 1C and 1D).

We quantified how the coloration of lizards changed in response to their background environment and observed that lizards from both the Pisgah and off-lava populations exhibited plastic phenotypes (Figure 1E). First, lizards housed for a year on lava were significantly darker than lizards housed for a year on sand for both the Pisgah ($F_{1,\ 6}$ = 13.4, p = 0.006) and off-lava ($F_{1,\ 7}$ = 164.1, p = 0.000003) populations (Figure 1E, Time 0). Second, the coloration of the lizards changed when the substrate was switched. Lizards from both populations housed on lava became significantly lighter when switched to sand (Time 10 versus 59, Pisgah: X^2_1 = 6.07, p = 0.0138; off-lava: X^2_1 = 13.38, p = 0.00025). Lizards from both populations housed on sand became significantly darker when switched to lava (Time 10 versus 59, Pisgah: X^2_1 = 5.21, p = 0.0224; off-lava: X^2_1 = 6.36, p = 0.0117). The two populations did not have different reaction norms, because there was no significant interaction between Population and Time (Time 10 versus 59) for either the lizards housed for a year on lava (X^2_1 = 0.02, p = 0.88) or sand (X^2_1 = 0.05, p = 0.83). Although substantial plasticity was detected, Pisgah lizards were significantly darker than off-lava lizards after being housed on lava for a year (Time 0, $F_{1,\ 8}$ = 9.6, p = 0.015). Pisgah and off-lava lizards housed on sand for a year did not differ significantly in coloration ($F_{1,\ 5}$ = 0.2, p = 0.67).

We hypothesized that the melanic phenotype could have a genetic component because Pisgah lizards can develop darker coloration than off-lava lizards on lava substrate (Figure 1E, Time 0). Therefore, we gathered genomic data to identify genes with high amounts of genetic differentiation between the Pisgah and off-lava populations. Individuals from the two populations

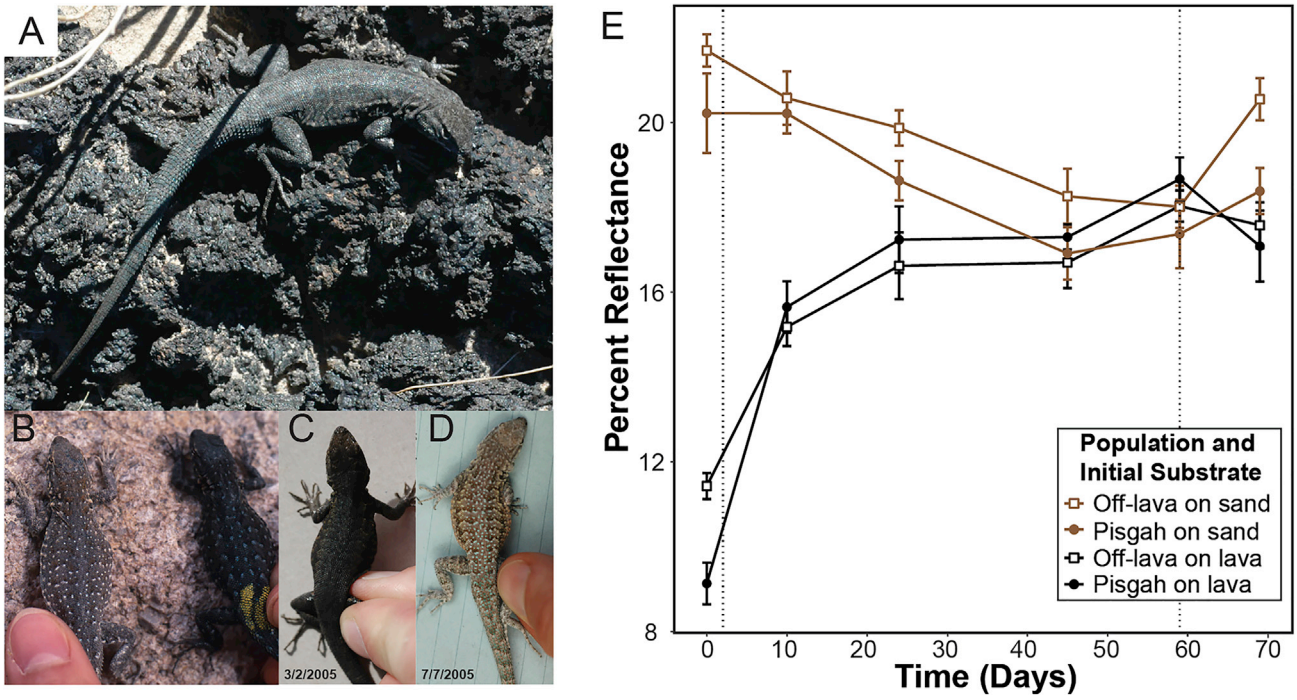


Figure 1. Coloration of the Pisgah Lava Flow and Off-Lava *U. stansburiana* in the Wild and after Being Housed on Lava or Sand Substrates

(A) A Pisgah male on a lava rock.

(B) A comparison of an off-lava lizard (left) to a Pisgah lizard (right, with yellow paint) on a rock at the off-lava site.

(C) A Pisgah male photographed 5 days after collection in the field.

(D) The same lizard as in (C) but after being housed for 4 months in the lab on light-colored sand.

(E) Plasticity in coloration of lizards housed on differently colored substrates. The dorsal coloration of the lizards was measured by the amount of light reflecting from the lizard relative to a white standard (i.e., percent reflectance), with lower values indicating darker lizards. Lizards were housed for a year on either sand or lava before their coloration was measured at Time 0. The substrate was switched on day 2 (vertical dotted line) and then back to the original substrate on day 59 (vertical dotted line). Line and point color correspond to the substrate that the lizards were housed on at Time 0 and for the year prior to the experiment with brown for sand and black for lava rock. Filled circles, Pisgah lizards; open squares, off-lava lizards. Points are mean values, and error bars are the SEM with the between-subject variability removed.

See also Figure S1.

could be distinguished on PC1 of a principle component analysis of SNPs, with no overlap between the Pisgah and off-lava populations (Figure 2A). Although there was some population structure, the overall genetic differentiation was low, because PC1 only explained 2.9% of the data. In addition, the mean F_{ST} between the two populations was 0.0215, indicating that most of the genetic variation was within rather than between populations. The joint site frequency spectrum for the Pisgah and off-lava populations showed that the allele frequencies of most SNPs were highly correlated and that there were no sites with fixed differences between the populations (Figure 2B). However, multiple sites did deviate from the general correlation and had divergent allele frequencies between the Pisgah and off-lava populations (outlying points, Figure 2B).

We used an association test to identify the genes with the most significant allele frequency differences between the populations. Among the most divergent genes were two genes with functions that suggested that they could be good candidates for influencing coloration (Table S1). One of these genes was prolyl endopeptidase (PREP), which is known to digest alpha melanocyte stimulating hormone (α-MSH), a hormone that stimulates the production of melanin [12]. The PREP gene was the 5th most highly differentiated gene between the populations and contained the 5th, 11th, and 13th most highly differentiated SNPs (Table S1), which were in introns (5th and 11th SNPs) or were a synonymous change (13th SNP). The other candidate gene was protein kinase cAMP-dependent type I regulatory subunit alpha (PRKAR1A), because variation in this gene has been associated with hyperpigmentation in humans [13–15]. The PRKAR1A gene was the 18th most highly differentiated gene between the populations and contained the 20th most highly differentiated SNP (Table S1), which was in an intron in the 5′ UTR. The top SNPs in the PREP and PRKAR1A genes were among the most differentiated sites as measured by F_{ST} (Figure 2C), were only found in the Pisgah population, and were at high frequencies compared to other SNPs restricted to the Pisgah population (Figure 2D). In addition to being missing from the off-lava population, the top associated SNPs in the PREP and PRKAR1A genes were not found in individuals from nine additional populations surrounding Pisgah, including an independently formed lava flow at Amboy Crater (Figure S2; Table S2).

We tested whether variation in PREP and PRKAR1A was associated with the dorsal coloration of lizards by studying an experimental cohort of newly hatched lizards that was raised in a

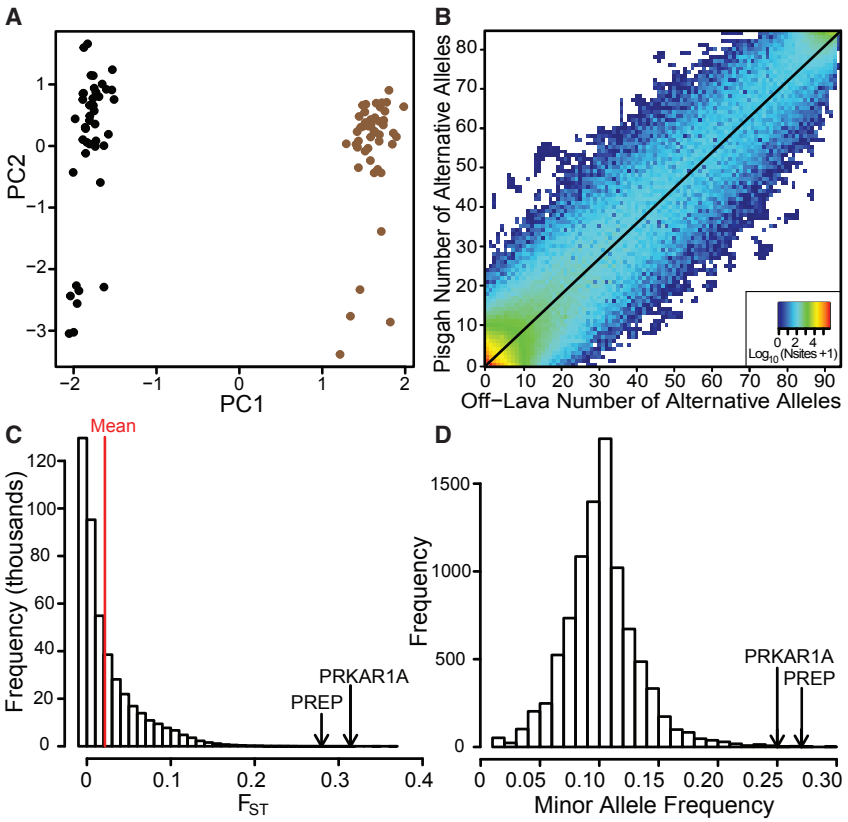


Figure 2. Assessments of the Genetic Differences between the Lava and Off-Lava Populations

(A) Principle component analysis of the genetic data (SNPs) for individuals from the Pisgah lava flow (black points) and off-lava (brown points) populations.

(B) The joint site frequency spectrum for the two populations. The number of alternative alleles relative to our reference sequence that were observed in our sample of 47 off-lava and 42 Pisgah individuals are plotted on the x and y axes, respectively. The color of the points indicates the number of sites (specifically log10 [number of sites +1]), from few sites (darker blue) to many sites (red).

(C) Histogram of the per-site estimates of F_{ST}. The mean F_{ST} is highlighted by the red line.

(D) Histogram of polymorphic sites that were only variable within the Pisgah population. The minor allele frequencies of these variants are plotted on the x axis. Arrows in (C) and (D) highlight the polymorphic sites in the PREP and PRKAR1A genes that were highly differentiated between the populations. In (B), PREP has 23 alternative alleles in Pisgah, PRKAR1A has 21 alternative alleles in Pisgah, and both genes have no alternative alleles in the off-lava population.

See also Figure S2 and Tables S1 and S2.

common environment (Figures 3A and 3B). We found that hatchling lizards with the derived C allele in the PREP gene were significantly darker than lizards homozygous for the ancestral A allele ($X^2_1 = 3.93$, p = 0.0237). We found that hatchling lizards with the derived G allele in the PRKAR1A gene were significantly darker than lizards homozygous for the ancestral A allele ($X^2_1 = 3.75$, p = 0.0263).

We utilized simulations to explore what levels of selection could lead to the observed frequency of the derived PREP and PRKAR1A alleles. The derived alleles in PREP and PRKAR1A were both, coincidentally, at a frequency of 0.23 in the Pisgah population (Table S2) but were not in linkage disequilibrium (p = 1, Table S3). We inferred the joint demographic history of the Pisgah and off-lava populations to serve as a demographic model under which we could estimate the selection coefficients and ages of the alleles. A demographic model with no migration estimated that the population divergence time was 5,200 generations ago, whereas the divergence time was 35,200 generations ago under a model with migration (Figure S3). Under demographic models with (Figure 3C) and without (Figure 3D) migration, we found that for both genes, the derived alleles were likely to have arisen recently (1,000 generations ago in both models) and have been subject to strong positive selection in the Pisgah population (s = 0.007 and 0.008, respectively).

DISCUSSION

Theories about plasticity and adaptation have a long history. In 1896, Baldwin proposed that plasticity is important for initial survival in new environments, where natural selection will subsequently adapt organisms in the direction of the induced plastic response by altering levels of plasticity or by changing the mean phenotype without changing plasticity [4, 5]. More recently, the "plasticity-first" hypothesis proposes that novel traits arise initially from variants produced via plasticity in a new environment, which are then subsequently refined by adaptive genetic evolution [1, 2, 8]. In other words, plasticity precedes and facilitates adaptation by natural selection, with "genes as followers" [2, 8]. Our data suggest that the origin of the melanic coloration of the Pisgah lizards follows many of the predictions of these theories. First, off-lava population lizards can change their coloration to match their environment, so plasticity in coloration evolved before the origin of the Pisgah lava flow (Figure 1E). Therefore, the first lizards to colonize Pisgah would have been able to achieve a substantial degree of background matching to the lava environment through ancestral plasticity in coloration, just as plasticity strongly contributes to the ability of Pisgah lizards to background match today (Figures 1C versus 1D, and 1E). Second, our experiments suggested that genetic changes to coloration have occurred in the Pisgah lizards, because they were darker than the off-lava lizards after being on a lava substrate for a year (Figure 1E, Time 0). A separate study also found evidence supporting genetic changes, because neonate Pisgah lizards from controlled genetic crosses were significantly darker than off-lava lizards despite all eggs being incubated in a common environment [16]. Our genomic data revealed that Pisgah lizards do have genetic changes in coloration, because variants in the PREP and PRKAR1A genes associated with darker phenotypes were found at substantially

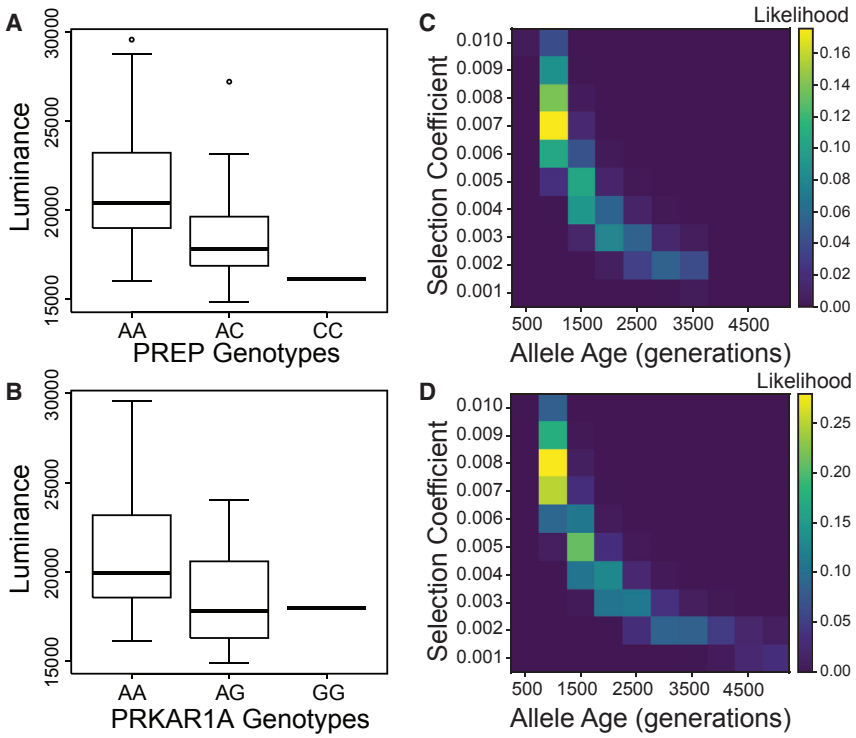


Figure 3. The Effects of PREP and PRKAR1A Alleles on Coloration and Their Inferred Selective History

(A and B) The coloration of hatchlings in relation to their genotypes for the alleles in the PREP (A) and PRKAR1A (B) genes with high levels of population differentiation. Lower values of luminance indicate darker lizards. For both genes, the ancestral allele is an A. The C allele in PREP and the G allele in PRKAR1A were only found in the Pisgah population. The boxplots for each genotype show the median (thick line in the box) and the 25th and 75th percentiles (bottom and top of the box); the whiskers extend to the furthest data point within 1.5 times the interquartile range, and outliers are shown as points.

(C and D) Estimates of the allele age and selection coefficient for the derived Pisgah allele of the PREP and PRKAR1A genes. Estimates come from demographic simulations with migration (C) and without migration (D) between the Pisgah and off-lava populations. The estimates of the allele age and selection coefficient are the same for both the PREP and PRKAR1A genes, because the derived alleles in these genes were both coincidentally at the same frequency in the Pisgah population.

See also Figure S3 and Tables S2 and S3.

higher frequencies in the Pisgah population. Overall, our data provide one of the most detailed examples of a "Baldwin effect," wherein plastic changes in coloration facilitated initial survival in the novel Pisgah environment by increasing crypsis, and subsequent genetic adaptations allowed the lizards to become even darker than their initial plastic responses.

Examples of the Baldwin effect or plasticity-first hypothesis have generally lacked an understanding of the genetic basis of the traits under study [1, 2, 17]. Our study helps fill this knowledge gap by providing a detailed mechanistic understanding of how the Pisgah lizards have adapted to the lava environment. The PREP and PRKAR1A genes are compelling candidates for adaptive changes in coloration because both are involved in a major pathway for signaling melanin production (Figure 4). The production of eumelanin (black and brown pigments) is stimulated by α-MSH binding to the melanocortin 1 receptor (MC1R), which signals for increased eumelanin production via the synthesis of cyclic adenosine monophosphate (cAMP) in the cell [18]. The PREP protein digests α-MSH to a truncated form, which may impair its ability to bind to MC1R [12], and inhibiting PREP activity can elevate the concentration of α-MSH [19]. Thus, the PREP gene could affect levels of signaling for melanin production via altered levels of α-MSH. The cAMP-dependent protein kinase A (PKA) protein is the primary mediator of cAMP signaling within the cell and is a tetramer composed of two regulatory subunits and two catalytic subunits [13]. There are four major regulatory subunit isoforms, with the PRKAR1A protein being one of the main regulatory subunit isoforms that control cAMP signaling [13]. Therefore, the PRKAR1A gene can affect melanin production by regulation of the PKA protein. Mutations in PRKAR1A can give rise to a hyperpigmentation phenotype in humans characterized by spotty skin [13–15].

PREP and PRKAR1A are a new pair of candidate genes for understanding adaptive changes in coloration in wild populations. Previous studies have observed that mutations in MC1R often underlie adaptive differences in coloration in wild populations [20–23] but that other genes involved in melanin production can also play a role [24, 25]. A prior study of *U. stansburiana* at Pisgah compared MC1R sequences of two individuals from both on and off the lava flow but did not detect fixed differences between the populations [26]. We also found no strong population differentiation in MC1R with our larger sample of individuals. Mutations in MC1R often result in canalized differences in phenotype [20, 22, 23, 27], which may explain why MC1R was not a target of selection at Pisgah. The ability to change coloration could be generally advantageous given that both Pisgah and off-lava lizards exhibited substantial plasticity (Figure 1E). The heterogeneous environment at Pisgah of light sand distributed among black rocks (Figure S1A) could make the ability to background match particularly advantageous. Thus, natural selection may have favored genetic changes that altered coloration while maintaining plasticity. PREP and PRKAR1A could be promising candidate genes for other systems where natural selection has acted on a plastic color phenotype.

The identification of candidate genes improves our understanding of the adaptive history of the Pisgah lizards. Variants in the PREP and PRKAR1A genes were correlated with the degree of dorsal darkness of hatchling lizards (Figure 3), which provided a critical validation that these variants have a phenotypic effect. This validation often cannot be done, because the genetic variants hypothesized to affect a plastic trait are fixed

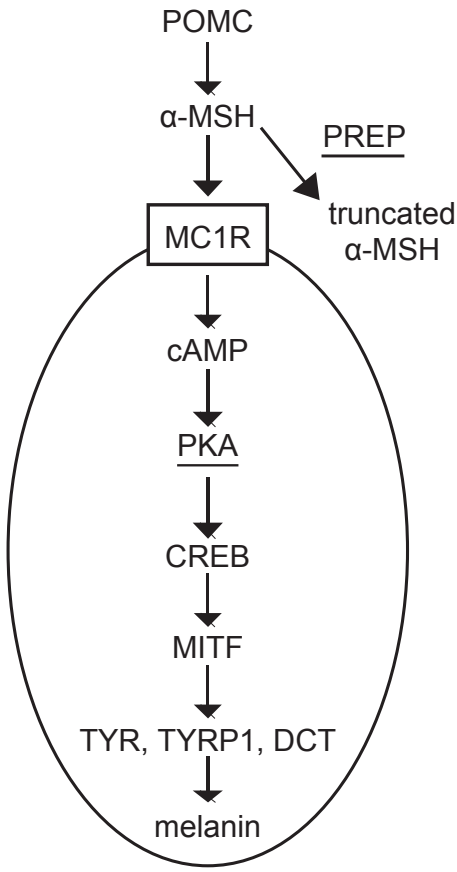


Figure 4. Pathway for Signaling Melanin Production [12, 18]
Reactions within a cell are in the oval, and parts of the pathway involving our candidate genes are underlined. Ultraviolet radiation stimulates the production of proopiomelanocortin (POMC) by the pituitary gland or in keratinocytes in the epidermis. POMC can be cleaved into alpha melanocyte stimulating hormone (α-MSH), which binds to the melanocortin 1 receptor (MC1R) in the cell membrane, leading to increased production of cyclic adenosine monophosphate (cAMP). Increased cAMP levels affect cAMP-dependent protein kinase A (PKA), which phosphorylates the cAMP response element-binding protein (CREB), in turn activating the melanogenesis-associated transcription factor (MITF). MITF is a transcription factor that regulates the expression of three key enzymes that help produce melanin: tyrosinase (TYR), tyrosinase-related protein 1 (TYRP1), and dopachrome tautomerase (DCT). The prolyl endopeptidase (PREP) protein digests α-MSH [12] and the protein kinase cAMP-dependent type I regulatory subunit alpha (PRKAR1A) protein is a regulatory subunit of the tetrameric PKA protein [13]. See also Table S1.

among populations, so their phenotypic effects cannot be studied *in situ* [1]. The derived variants in PREP and PRKAR1A did not change the protein sequence, so the effects on coloration that we observed likely result from changes in gene expression, as often occurs in the adaptation of plastic traits [9, 28]. Our demographic models suggest that variants in the PREP and PRKAR1A genes were likely to have arisen recently and experienced strong positive selection. The generation time of the Pisgah lizards is unknown, but even with a generation time substantially greater than the minimum of 1–2 years (Sinervo personal observation), our simulations suggest that the PREP and PRKAR1A alleles arose thousands of years after the formation of the lava flow. Another source of uncertainty is our lack of knowledge of the mutation rate of lizards, which could affect the timescale of our estimates. However, this error is unlikely to be extreme, as the demographic models showed good concordance with the known age of the lava flow. The split times inferred in our models (no migration: 5,200 generations; with migration: 35,200 generations, Figure S3) are compatible with the origin of the Pisgah lava flow 22,500 ± 1,300 years ago [11], especially if the actual population history was isolation with occasional gene flow, rather than the two disparate demographic scenarios that we examined. Thus, although there is uncertainty in the exact timing, the models suggest that the PREP and PRKAR1A alleles arose recently, within the last few thousand years.

Our genetic data also offered some unexpected insights into adaptation following plastic responses in a novel environment. The plasticity-first hypothesis proposes that there is genetic variation in plastic responses that is hidden until exposed in a novel environment and that this variation can be fuel for subsequent adaptive change [1, 2, 29]. However, adaptation to the Pisgah environment does not appear to arise from standing genetic variation because the derived variants in our candidate genes were restricted to the Pisgah population. We were unable to detect the derived allele in the PREP and PRKAR1A genes in the off-lava population, where we sequenced 57 diploid individuals (a sample of 114 chromosomes), nor in nine other populations from the surrounding area, where we sequenced 76 individuals (152 chromosomes). Thus, if the derived variants did originate before the lava flow formed, they must be at very low frequencies in the populations surrounding Pisgah. If there was standing genetic variation in the PREP and PRKAR1A genes in ancestral off-lava populations, then these variants were likely to be selected in the similar environment at the Amboy Crater lava flow, given that this kind of convergent selection on standing variants has been observed in sticklebacks [30, 31]. However, we did not detect the derived variants in the PREP and PRKAR1A genes in the Amboy Crater population (Table S2). Thus, the available evidence suggests that the PREP and PRKAR1A alleles likely arose as a result of *de novo* mutations in the Pisgah population. Therefore, plasticity in coloration preceded adaptive genetic changes in coloration at Pisgah, but adaptation did not result from standing genetic variation in the degree of plasticity.

Further insights into plasticity and adaptation can be developed on the foundation established by our study. First, it is unknown how *U. stansburiana* utilize information from their background environment to plastically alter their coloration, but knowing that levels of α-MSH could play a role may offer an important clue for future research. A possible mechanism is that visual signals are neuronally processed to affect the production in the pituitary of proopiomelanocortin, which is the precursor protein to α-MSH. Second, there is an opportunity for field studies of selection aimed at better understanding the adaptation of a highly plastic phenotype. Field studies are needed to evaluate whether plasticity typically has a positive role in adaptation, whether it is largely non-adaptive (e.g., [32]), or whether it accelerates or decelerates selection [33]. Third, many theories of plasticity posit that natural selection will adapt organisms by modifying the degree of phenotypic plasticity [1, 2, 5–7, 29], but rarely is anything known about the genetic basis of

the traits or the mechanism by which plastic responses are modified. The PREP and PRKAR1A genes are good candidates for being involved in plastic responses because both genes help regulate the production of melanin (Figure 4). In future studies, we will quantify the plastic responses of animals with different PREP and PRKAR1A genotypes to test whether variants in these genes alter levels of plasticity. This will illuminate whether plasticity was a target of selection as *U. stansburiana* adapted to the novel environment at Pisgah.

We found that the melanic coloration of Pisgah lizards results from both ancestral phenotypic plasticity and adaptive genetic changes that evolved after the formation of the lava flow, possibly through *de novo* mutations. Therefore, selection is still able to refine highly plastic phenotypes that match organisms to their environment. The genetic changes observed in PREP and PRKAR1A modified coloration in the direction of the induced plastic response, as first predicted by Baldwin [4, 5]. Thus, plasticity and adaptation can work together to promote survival in new environments. Overall, studying the plasticity and genetic basis of traits can offer important insights into how novel phenotypes evolve.

STAR★METHODS

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

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Tests of phenotypic plasticity
  - Genetic comparison of lava and off-lava lizards
  - Population survey of candidate gene variants
  - Testing for associations between genotypes and phenotypes
- METHOD DETAILS
  - Tests of phenotypic plasticity
  - Transcriptome sequencing and transcriptome assembly
  - Library preparation, sequence-capture, and sequencing
  - Read cleanup, mapping, and analysis
  - Population genetic analyses
  - Sequencing candidate genes
  - Scoring alleles in candidate genes
  - Estimating the demographic history and fitting models of selection on candidate genes
  - Testing for associations between genotypes and phenotypes
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Tests of phenotypic plasticity
  - Population genetic analyses
  - Testing for associations between genotypes and phenotypes
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes three figures and three tables and can be found with this article online at https://doi.org/10.1016/j.cub.2018.06.075.

A video abstract is available at https://doi.org/10.1016/j.cub.2018.06.075#mmc3.

ACKNOWLEDGMENTS

This work was funded by a University of California Catalyst Grant to R.N., an NSF dissertation improvement grant (DEB-0408172) to A.C. and B.S., and an NSF Emerging Frontier grant (EF-1241848) to B.S. This work used the Vincent J. Coates Genomics Sequencing Laboratory at U.C. Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668, S10RR027303, and S10 OD018174. This work used the Extreme Science and Engineering Discovery Environment (XSEDE) resource Blacklight (allocation # IBN100014), supported by National Science Foundation grant number ACI-1548562 [34]. Thanks to Tyler Linderoth, Sonal Singhal, and Lydia Smith for help in the lab; to Thorfinn Korneliussen for help with data analysis; and to Rauri Bowie, Jim McGuire, and the U.C. Berkeley Museum of Vertebrate Zoology for logistical support.

AUTHOR CONTRIBUTIONS

Conceptualization, A.C., C.L., and R.N.; Investigation, A.C., K.B., C.L., A.S.C., and B.S.; Formal Analysis, A.C., K.B., A.J.S., and R.N.; Writing – Original Draft, A.C.; Writing – Review & Editing, all authors; Funding Acquisition, A.C., B.S., and R.N.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 10, 2018
Revised: June 20, 2018
Accepted: June 29, 2018
Published: September 6, 2018

REFERENCES

1. Levis, N.A., and Pfennig, D.W. (2016). Evaluating "plasticity-first" evolution in nature: key criteria and empirical approaches. Trends Ecol. Evol. *31*, 563–574.
2. West-Eberhard, M.J. (2003). Developmental plasticity and evolution (New York: Oxford University Press).
3. Charmantier, A., McCleery, R.H., Cole, L.R., Perrins, C., Kruuk, L.E.B., and Sheldon, B.C. (2008). Adaptive phenotypic plasticity in response to climate change in a wild bird population. Science *320*, 800–803.
4. Baldwin, J.M. (1896). A new factor in evolution. Am. Nat. *30*, 441–451.
5. Crispo, E. (2007). The Baldwin effect and genetic assimilation: revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. Evolution *61*, 2469–2479.
6. Waddington, C.H. (1953). Genetic assimilation of an acquired character. Evolution *7*, 118–126.
7. Waddington, C.H. (1942). Canalization of development and the inheritance of acquired characters. Nature *150*, 563–565.
8. Schwander, T., and Leimar, O. (2011). Genes as leaders and followers in evolution. Trends Ecol. Evol. *26*, 143–151.
9. Scoville, A.G., and Pfrender, M.E. (2010). Phenotypic plasticity facilitates recurrent rapid adaptation to introduced predators. Proc. Natl. Acad. Sci. USA *107*, 4260–4263.
10. Bock, D.G., Kantar, M.B., Caseys, C., Matthey-Doret, R., and Rieseberg, L.H. (2018). Evolution of invasiveness by genetic accommodation. Nat Ecol Evol *2*, 991–999.
11. Phillips, F.M. (2003). Cosmogenic ^{36}Cl ages of Quaternary basalt flows in the Mojave Desert, California, USA. Geomorphology *53*, 199–208.
12. Perroud, B., Alvarado, R.J., Espinal, G.M., Morado, A.R., Phinney, B.S., and Warden, C.H. (2009). In vivo multiplex quantitative analysis of 3 forms of alpha melanocyte stimulating hormone in pituitary of prolyl endopeptidase deficient mice. Mol. Brain *2*, 14–19.
13. Bossis, I., and Stratakis, C.A. (2004). Minireview: PRKAR1A: normal and abnormal functions. Endocrinology *145*, 5452–5458.
14. Kirschner, L.S., Sandrini, F., Monbo, J., Lin, J.P., Carney, J.A., and Stratakis, C.A. (2000). Genetic heterogeneity and spectrum of mutations

of the PRKAR1A gene in patients with the carney complex. Hum. Mol. Genet. *9*, 3037–3046.

15. Kirschner, L.S., Carney, J.A., Pack, S.D., Taymans, S.E., Giatzakis, C., Cho, Y.S., Cho-Chung, Y.S., and Stratakis, C.A. (2000). Mutations of the gene encoding the protein kinase A type I-alpha regulatory subunit in patients with the Carney complex. Nat. Genet. *26*, 89–92. https://doi.org/10.1038/79238.
16. Corl, A., Lancaster, L.T., and Sinervo, B. (2012). Rapid formation of reproductive isolation between two populations of side-blotched lizards, *Uta stansburiana*. Copeia *2012*, 593–602.
17. Badyaev, A.V. (2009). Evolutionary significance of phenotypic accommodation in novel environments: an empirical test of the Baldwin effect. Philos. Trans. R. Soc. Lond. B Biol. Sci. *364*, 1125–1141.
18. D'Mello, S.A., Finlay, G.J., Baguley, B.C., and Askarian-Amiri, M.E. (2016). Signaling pathways in melanogenesis. Int. J. Mol. Sci. *17*, E1144.
19. Bellemère, G., Morain, P., Vaudry, H., and Jégou, S. (2003). Effect of S 17092, a novel prolyl endopeptidase inhibitor, on substance P and α-melanocyte-stimulating hormone breakdown in the rat brain. J. Neurochem. *84*, 919–929.
20. 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.
21. Hoekstra, H.E. (2006). Genetics, development and evolution of adaptive pigmentation in vertebrates. Heredity (Edinb) *97*, 222–234.
22. Rosenblum, E.B., Römpler, H., Schöneberg, T., and Hoekstra, H.E. (2010). Molecular and functional basis of phenotypic convergence in white lizards at White Sands. Proc. Natl. Acad. Sci. USA *107*, 2113–2117.
23. Nachman, M.W., Hoekstra, H.E., and D'Agostino, S.L. (2003). The genetic basis of adaptive melanism in pocket mice. Proc. Natl. Acad. Sci. USA *100*, 5268–5273.
24. Linnen, C.R., Poh, Y.-P., Peterson, B.K., Barrett, R.D.H., Larson, J.G., Jensen, J.D., and Hoekstra, H.E. (2013). Adaptive evolution of multiple traits through multiple mutations at a single gene. Science *339*, 1312–1316.
25. Hubbard, J.K., Uy, J.A.C., Hauber, M.E., Hoekstra, H.E., and Safran, R.J. (2010). Vertebrate pigmentation: from underlying genes to adaptive function. Trends Genet. *26*, 231–239.
26. Rosenblum, E.B., Hoekstra, H.E., and Nachman, M.W. (2004). Adaptive reptile color variation and the evolution of the Mc1r gene. Evolution *58*, 1794–1808.
27. Rosenblum, E.B. (2005). The role of phenotypic plasticity in color variation of Tularosa Basin lizards. Copeia *2005*, 586–596.
28. Levis, N.A., Serrato-Capuchina, A., and Pfennig, D.W. (2017). Genetic accommodation in the wild: evolution of gene expression plasticity during character displacement. J. Evol. Biol. *30*, 1712–1723.
29. West-Eberhard, M.J. (2005). Developmental plasticity and the origin of species differences. Proc. Natl. Acad. Sci. USA *102* (*Suppl 1*), 6543–6549.
30. Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G., Jr., Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Schluter, D., and Kingsley, D.M. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science *307*, 1928–1933.
31. Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S., et al.; Broad Institute Genome Sequencing Platform & Whole Genome Assembly Team (2012). The genomic basis of adaptive evolution in threespine sticklebacks. Nature *484*, 55–61.
32. Ghalambor, C.K., Hoke, K.L., Ruell, E.W., Fischer, E.K., Reznick, D.N., and Hughes, K.A. (2015). Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. Nature *525*, 372–375.
33. Paenke, I., Sendhoff, B., and Kawecki, T.J. (2007). Influence of plasticity and learning on evolution under directional selection. Am. Nat. *170*, E47–E58.
34. Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., et al. (2014). XSEDE: Accelerating Scientific Discovery. Comput. Sci. Eng. *16*, 62–74.
35. Singhal, S. (2013). *De novo* transcriptomic analyses for non-model organisms: an evaluation of methods across a multi-species data set. Mol. Ecol. Resour. *13*, 403–416.
36. Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C., and Good, J.M. (2012). Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. BMC Genomics *13*, 403.
37. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.
38. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetJournal. https://doi.org/10.14806/ej.17.1.200.
39. Petersen, K.R., Streett, D.A., Gerritsen, A.T., Hunter, S.S., and Settles, M.L. (2015). Super deduper, fast PCR duplicate detection in fastq files. BCB '15 (New York, New York, USA: ACM Press), pp. 491–492.
40. Magoč, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics *27*, 2957–2963.
41. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.
42. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. *19*, 455–477.
43. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.
44. Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics *22*, 1658–1659.
45. Huang, X., and Madan, A. (1999). CAP3: A DNA sequence assembly program. Genome Res. *9*, 868–877.
46. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.
47. Bi, K., Linderoth, T., Vanderpool, D., Good, J.M., Nielsen, R., and Moritz, C. (2013). Unlocking the vault: next-generation museum population genomics. Mol. Ecol. *22*, 6018–6032.
48. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics *15*, 356.
49. Korneliussen, T.S., and Moltke, I. (2015). NgsRelate: a software tool for estimating pairwise relatedness from next-generation sequencing data. Bioinformatics *31*, 4009–4011.
50. Fumagalli, M., Vieira, F.G., Linderoth, T., and Nielsen, R. (2014). ngsTools: methods for population genetics analyses from next-generation sequencing data. Bioinformatics *30*, 1486–1487.
51. Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V.C., and Foll, M. (2013). Robust demographic inference from genomic and SNP data. PLoS Genet. *9*, e1003905–e1003917.
52. Troscianko, J., and Stevens, M. (2015). Image calibration and analysis toolbox - a free software suite for objectively measuring reflectance, colour and pattern. Methods Ecol. Evol. *6*, 1320–1331.
53. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675.
54. Wickham, H. (2009). ggplot2: Elegant graphics for data analysis (Springer-Verlag).
55. R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. Available at: http://www.R-project.org/.
56. Chang, W. (2012). R Graphics Cookbook. Available at: www.cookbook-r.com/Graphs/Plotting_means_and_error_bars_(ggplot2)/.

57. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. J. Stat. Softw. *67*, https://doi.org/10.18637/jss.v067.i01.

58. Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. *29*, 644–652.

59. Slater, G.S., and Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics *6*, 31.

60. Smit, A., Hubley, R., and Green, P. (2013). RepeatMasker Open-4.0. 2013-2015. Available at: http://www.repeatmasker.org.

61. Corl, A., Davis, A.R., Kuchta, S.R., Comendant, T., and Sinervo, B. (2010). Alternative mating strategies and the evolution of sexual size dimorphism in the side-blotched lizard, *Uta stansburiana*: a population-level comparative analysis. Evolution *64*, 79–96.

62. Corl, A., Davis, A.R., Kuchta, S.R., and Sinervo, B. (2010). Selective loss of polymorphic mating types is associated with rapid phenotypic evolution during morphic speciation. Proc. Natl. Acad. Sci. USA *107*, 4254–4259.

63. Luke, C.A. (1989). Color as a phenotypically plastic character in the side-blotched lizard, *Uta stansburiana* (University of California Berkeley).

64. Miller, S.A., Dykes, D.D., and Polesky, H.F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. *16*, 1215.

65. Rohland, N., and Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. *22*, 939–946.

66. Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc *2010*, pdb.prot5448.

67. NimbleGen (2013). SeqCap EZ library SR user's guide (Roche NimbleGen, Inc), pp. 1–60.

68. Hodges, E., Rooks, M., Xuan, Z., Bhattacharjee, A., Benjamin Gordon, D., Brizuela, L., Richard McCombie, W., and Hannon, G.J. (2009). Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. Nat. Protoc. *4*, 960–974.

69. Nielsen, R., Korneliussen, T., Albrechtsen, A., Li, Y., and Wang, J. (2012). SNP calling, genotype calling, and sample allele frequency estimation from New-Generation Sequencing data. PLoS ONE *7*, e37558.

70. Nielsen, R., Paul, J.S., Albrechtsen, A., and Song, Y.S. (2011). Genotype and SNP calling from next-generation sequencing data. Nat. Rev. Genet. *12*, 443–451.

71. Kim, S.Y., Lohmueller, K.E., Albrechtsen, A., Li, Y., Korneliussen, T., Tian, G., Grarup, N., Jiang, T., Andersen, G., Witte, D., et al. (2011). Estimation of allele frequency and association mapping using next-generation sequencing data. BMC Bioinformatics *12*, 231.

72. Crawford, J.E., and Nielsen, R. (2013). Detecting adaptive trait loci in nonmodel systems: divergence or admixture mapping? Mol. Ecol. *22*, 6131–6148.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Exon-capture dataset of 52 Pisgah Lava Flow and 52 off-lava *Uta stanburiana*	This paper	NCBI SRA (BioProject ID: PRJNA476148, SRA accession: SRP150551) https://www.ncbi.nlm.nih.gov/sra/SRP150551
Anolis carolinensis transcript for the PREP gene	ENSEMBL	ENSACAT00000003886.2
Anolis carolinensis transcript for the PRKAR1A gene	ENSEMBL	ENSACAT00000013273.3
Oligonucleotides		
PRKAR1A forward primer, PRKAR1A-F1: TGAAAGGTGATACTGGGATCAACACC	This paper	N/A
PRKAR1A reverse primer, PRKAR1A-R1: CTCACTGCCGGTACTCATAGAAGC	This paper	N/A
PRKAR1A reverse primer, PRKAR1A-R2: TCTCCAGCCTTTCAAAGTATTCCCG	This paper	N/A
PREP forward primer, PREP-F1: CCACCATCCTCCATTTCCTTCAGTC	This paper	N/A
PREP forward primer, PREP-F2: TCATGTGGCACTCCAGAAGTTTGG	This paper	N/A
PREP reverse primer, PREP-R2: TCACCATCTTCACTGAATGCATAACCT	This paper	N/A
Software and Algorithms		
Workflow designed for transcriptome-based exon capture data analyses	Ke Bi, GitHub site for CGRL-QB3-UCBerkeley, [35, 36]	https://github.com/CGRL-QB3-UCBerkeley/denovoTargetCapturePopGen
Trimmomatic	[37]	http://www.usadellab.org/cms/?page=trimmomatic
Cutadapt	[38]	https://cutadapt.readthedocs.io/en/stable/installation.html
Super Deduper	[39]	https://github.com/ibest/HTStream/tree/master/hts_SuperDeduper
FLASH	[40]	http://ccb.jhu.edu/software/FLASH/
Bowtie2	[41]	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Pipelines for assembling an "exome reference" sequence, to have a more complete reference sequence that includes both the original targeted sequences as well as flanking sequences to the targeted sequences.	Ke Bi, GitHub site for CGRL-QB3-UCBerkeley	https://github.com/CGRL-QB3-UCBerkeley/seqCapture
SPAdes genome assembler	[42]	http://cab.spbu.ru/software/spades/
BLASTn	[43]	https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=Download
cd-hit-est	[44]	http://weizhongli-lab.org/cd-hit/
cap3	[45]	http://seq.cs.iastate.edu/cap3.html
Novoalign	Novocraft Technologies	http://www.novocraft.com/products/novoalign/
Picard	The Broad Institute	http://broadinstitute.github.io/picard/
GATK	[46]	https://software.broadinstitute.org/gatk/download/
SNPcleaner v2.2.4	Tyler Linderoth Github site [47]	https://github.com/tplinderoth/ngsQC/tree/master/snpCleaner
ANGSD	[48]	http://www.popgen.dk/angsd/index.php/ANGSD
NgsRelate	[49]	www.popgen.dk/software/index.php/NgsRelate
ngsTools	[50]	https://github.com/mfumagalli/ngsTools
FastSimCoal2	[51]	http://cmpg.unibe.ch/software/fastsimcoal2/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Image Calibration and Analysis Toolbox	[52]	http://www.jolyon.co.uk/myresearch/image-analysis/image-analysis-tools/
ImageJ	[53]	https://imagej.nih.gov/ij/
ggplot2 plot R package	[54]	https://CRAN.R-project.org/package=ggplot2
R software for statistical computing	[55]	https://www.r-project.org/
function summarySEwithin.R from the R Graphic Cookbook	[56]	www.cookbook-r.com/Graphs/Plotting_means_and_error_bars_(ggplot2)/
lme R package lme4	[57]	https://CRAN.R-project.org/package=lme4
Trinity	[58]	https://github.com/trinityrnaseq/trinityrnaseq/wiki
BLASTx	[43]	https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=Download
Exonerate	[59]	https://www.ebi.ac.uk/about/vertebrate-genomics/software/exonerate
RepeatMasker	[60]	http://www.repeatmasker.org/
Geneious 8.1.9	Biomatters Ltd.	https://www.geneious.com/

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Ammon Corl (corlammon@gmail.com).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Tests of phenotypic plasticity

Fifteen lava and eighteen off-lava side-blotched lizards (*Uta stansburiana*) were collected from the field. There were four experimental groups: 8 Pisgah lizards (3 females, 5 males) housed on lava gravel, 7 Pisgah lizards (3 females, 4 males) housed on sand, 10 off-lava lizards (7 females, 3 males) housed on lava gravel, and 8 off-lava lizards (7 females, 1 male) housed on sand. All lizards in the experiment were maintained in captivity under ultraviolet and Vitalights on a cycle of 12 hours dark and 12 hours light. A heat tape was placed under one end of each terrarium to provide a thermal gradient. Lizards were fed waxworms and crickets dusted with calcium carbonate and Superpreen bird vitamins.

Genetic comparison of lava and off-lava lizards

We used tissue samples of *U. stansburiana* collected in a prior study [16] from 23 Pisgah males, 29 Pisgah females, 23 off-lava males, and 29 off-lava females.

Population survey of candidate gene variants

We used tissue samples of *U. stansburiana* collected in a prior phylogeographic study [61, 62]. Information on the populations surveyed and the numbers of samples per population can be found in Figure S2 and Table S2.

Testing for associations between genotypes and phenotypes

We utilized specimens from a genetic-crossing experiment performed with Pisgah and off-lava *U. stansburiana* [16]. Protocols were approved by the U.C. Santa Cruz Institutional Animal Care and Use Committee (protocol SINE00.02-1). We gathered data on 26 male and 35 female hatchlings (61 individuals total)

METHOD DETAILS

Tests of phenotypic plasticity

Measurements of phenotypic plasticity in the Pisgah and off-lava lizard populations were conducted in 1987-1988 [63]. The Pisgah lizards were collected 1.5 miles south of the lava flow's edge in an area of lava rock interspersed with pockets of windblown sand. The off-lava site was 1.5 miles north of the flow in low hills with sandy alluvial soil. Fifteen lava and eighteen off-lava lizards were collected. Lizards from both sites were housed on lava gravel and sand to measure their ability to change coloration in response to soil color. Thus, four experimental groups were created with 8 Pisgah lizards housed on lava gravel, 7 Pisgah lizards housed on sand, 10 off-lava lizards housed on lava gravel, and 8 off-lava lizards housed on sand. Color measurements were first taken after at least one year on their substrate (day 0). Two days later, lizards were switched to the alternate soil type and color measurements were taken again 8,

22, 43, and 57 days following the switch. The soil substrate was switched back to the initial soil type 57 days after the first switch and additional color measurements were taken ten days later.

All color measurements were taken at 35°C between 11:00 and 14:00 hours. Skin reflectance was measured by a CE 395 Fast Scanning Spectral Scanner that was sensitive to reflected wavelengths of light from 350 to 750 nm. A spot optics lens and a receptacle for exchangeable metal plates were placed in front of the scanner. Light was received by the scanner through a 0.65 mm diameter hole in the plate. Reflectance was measured by pressing lizards lightly against the back side of the plate with their mid-dorsal region exposed through the hole. For each lizard reflectance taken, a white standard (a 100% rag cotton white card) was also measured for a comparison with total possible reflected light. Sunlight or the light from a 300-Watt slide projector bulb was used for illumination as both produced similar reflectance curves. Percent reflectance was calculated as the area of the reflectance curve of the lizard divided by that of the white standard. The area under each curve was measured by a Tektronix 212 portable oscilloscope connected to the spectral scanner. The oscilloscope screen was photographed, the negatives were then projected, the curves traced by hand, and an image analyzer was used to calculate total reflectance as the number of pixels captured on the screen.

Effective sample sizes of the lizards in the experiment varied due to a film development error and lizard mortality. The full sample size of 8 Pisgah lizards on lava, 7 Pisgah lizards on sand, 10 off-lava lizards on lava, and 8 off-lava lizards on sand was available for days 10, 45, and 59 of the experiment. On day 0, the film development error reduced the sample sizes of the lizards to 3 Pisgah on lava, 5 Pisgah on sand, 7 off-lava on lava, and 2 off-lava on sand. On day 24, the sample sizes were 8 Pisgah on lava, 7 Pisgah on sand, 9 off-lava on lava, and 7 off-lava on sand due to the film development error. On day 69, due to lizard mortality, the samples sizes were 7 Pisgah on lava, 6 Pisgah on sand, 9 off-lava on lava, and 8 off-lava on sand.

We graphed the change in coloration over time using the ggplot2 plot package [54] in R [55] and the function summarySEwithin.R from the R Graphic Cookbook [56], which removes the between-subject variability when calculating standard errors. We used an ANOVA in R to test if dorsal coloration differed among pairs of treatments on day 0 of the experiment. We tested for plastic responses of dorsal coloration following the initial substrate switch using a linear mixed effects model as implemented in the R package lme4 [57]. We tested whether variation in coloration was correlated with a fixed effect of time to test for plasticity. Specifically, we compared the dorsal coloration at time point 10 (eight days after the substrate was switched) and time point 59 (57 days after the substrate was switched) independently for each of the four treatments. This comparison underestimates the extent of plasticity, because even more extreme phenotypic differences were observed at time 0, before the substrate was switched (see Figure 1E). However, we conservatively compared time points 10 versus 59, because of the reduced sample size at time 0. Although not shown, similar results were obtained by comparing time points 0 versus 59. Our model included random intercepts for the individual lizards. We compared models with and without the effect of time with a likelihood ratio test. We tested if the populations differed in their reaction norms for each of the two treatments (housed initially on lava for one year and then switched to sand, housed initially on sand for one year and then switched to lava). We did this by constructing a model with fixed effects of population, time (Time 10 versus 59), time x population, and random intercepts for the individual lizards to explain variation in dorsal coloration. We then used a likelihood ratio test to compare models with and without the time x population interaction term.

Transcriptome sequencing and transcriptome assembly

We developed our exon-capture system by sequencing transcriptomes for eleven male *U. stansburiana*, which were collected at a focal study population near Los Banos, CA. To gain a broad sample of gene expression, we sequenced three different tissues (throat skin, brain, and testes) and used samples from males collected both before and during the spring breeding season. Libraries were constructed with a TruSeq RNA Sample Kit and then sequenced on a HiSeq 2000 with 100 bp paired-end reads. After sequencing, the reads were filtered following Singhal [35] before *de novo* assembly using Trinity [58]. Data were annotated using the programs BLASTx [43] and Exonerate [59] along with seven vertebrate genomes from Ensembl (*Anolis carolinensis*, human, mouse, chicken, zebra finch, *Xenopus tropicalis,* and zebrafish). We filtered the resulting annotated transcripts so that we retained only the longest isoform per gene, all sequences had between 35%–70% GC content, we removed sequence matching *A. carolinensis* ribosomal RNA, and we removed repetitive elements, low complexity regions, and short repeats using RepeatMasker [60]. We truncated the sequences for each gene so that there were no more than 250 base pairs on either side of the annotated coding region. In addition to our sequences obtained from the transcriptomes, we added in six *U. stansburiana* sequences from GenBank and three *A. carolinensis* sequences, which was done to include sequences for some functionally interesting genes that were not in the transcriptome assembly. Our final set of sequences included approximately 29.16 million bases of sequence and 16,345 contigs annotated to unique genes. These sequences were sent to Roche NimbleGen, where custom capture probes matching the lizard sequences were synthesized in their SeqCap EZ Developer Library.

Library preparation, sequence-capture, and sequencing

Tissue samples used in this project were collected in 2005 on the Pisgah Lava Flow (34.76157°N, −116.36576°W) and 7 km off the lava flow (34.81902°N, −116.33904°W). Animals were collected under a California Department of Fish and Game permit to B. Sinervo (#801060-01). The lizards were originally collected for a genetic crossing experiment between the lava and off-lava sites that tested for reproductive isolation between the populations [16]. We used samples from 52 individuals from each population, which consisted of 23 Pisgah males, 29 Pisgah females, 23 off-lava males, and 29 off-lava females.

DNA was extracted using a salt extraction protocol modified from [64], which included an overnight digestion with 50 μg of RNaseA to allow accurate quantification of the resulting DNA with a NanoDrop spectrophotometer. A 100 μL dilution at 16 ng/μL was made for

each sample in a 0.5 mL Diagenode Bioruptor tube. The concentration of the dilution was checked on a Nanodrop and adjusted if necessary. Samples were sonicated with a Bioruptor UCD-200 (Diagenode) for 15 minutes on the low setting with the machine set to 30 s on and 30 s off. The size range of the sonicated samples was checked by running 2 μL of the sample on an agarose gel. Samples were resonicated if the majority of DNA was not between 200 and 600 bp. Then 95 μL of the sonicated DNA was size-selected to remove fragments below 200 bp using a 0.8x SPRI bead cleanup [65], with the bead solution having 20% PEG and 2.5M NaCl.

Libraries for each sample were constructed following the protocol described in Meyer and Kircher [66] with two minor modifications. First, reactions were cleaned with a 1.3 fold volume of Seramag beads. Second, 4 μL of template DNA was used for the indexing PCR reactions and samples were dual-indexed for subsequent pooling. Cycling of individual PCR reactions varied from nine to fifteen cycles, depending on the amplification success of the individual samples, but most were amplified with only nine or ten cycles. Following library construction, two pools of DNA were constructed, with each pool containing 26 Pisgah and 26 off-lava samples. To ensure equal representation of each sample in the pool, a 30 μL dilution at 15 ng/μL was made for each library and the concentration of the dilution was checked on a Nanodrop and adjusted if necessary. Then 10 μL of each library dilution was taken and added to the pool. The concentrations of the two resulting pools were quantified using a Qubit fluorometer (Life Technologies).

Sequence-capture of each of the two library pools was conducted following the SeqCap EZ Library SR User's Guide v4.1 [67] with the following modifications. We added 1.3 μg of DNA from a library pool to its designated capture reaction. We added 3.25 μg of human COT-1 DNA and 3.25 μg of chicken Hybloc DNA to each capture. Barcode specific blocking oligos were added for each P5 and P7 adaptor in the pool. The libraries were hybridized for 71 hours. After hybridization, we verified target enrichment efficiency with qPCR analysis of the pre- and post-captured library pools, following Hodges et al. [68]. Finally, equal amounts of DNA from the two pools were combined and then this pool of all the individuals was sequenced on two lanes of a HiSeq4000 with 150 bp paired-end reads.

Read cleanup, mapping, and analysis

Our sequencing generated 179 gigabases of data in total. We processed this data following a workflow designed for transcriptome-based exon capture data analyses (https://github.com/CGRL-QB3-UCBerkeley/denovoTargetCapturePopGen). Specifically, raw data were cleaned using methods detailed in Bi et al. [36] and Singhal [35]. Raw fastq reads were filtered using Trimmomatic [37] and Cutadapt [38] to remove low quality reads and adaptor contamination. Exact duplicates were eliminated using Super Deduper [39]. We used FLASH [40] to merge overlapping paired-end reads. In addition, we used Bowtie2 [41] to map reads to potential contaminant sources (human and *E. coli* genome sequences) so that contaminants could be removed.

We assembled an "exome reference" sequence, to have a more complete reference sequence that included both the original targeted sequences as well as flanking sequences to the targeted sequences. We used the SPAdes genome assembler version 3.10.1 [42] to independently assemble the cleaned reads from four of the off-lava libraries that had high amounts of data. We used kmers of 21, 33, 55, and 77 for each individual assembly process. We then used BLASTn [43] searches to identify the targeted genes in each assembly, with the assemblies needing to be at least a 90% match to the target. For each individual assembly, we used cd-hit-est [44] and cap3 [45] to cluster and assemble the multiple contigs matching each targeted transcript into less-redundant assemblies. We then combined the assemblies for the four individuals by a similar methodology. Finally, we generated an exome reference sequence where all non-redundant and discrete contigs (exons and their flanking sequences) that were derived from the same target were joined together with Ns based on their relative blast hit positions to the transcriptome reference. Pipelines for processing the data can be found at https://github.com/CGRL-QB3-UCBerkeley/seqCapture.

We mapped our cleaned reads for each individual to the exome reference using Novoalign (novocraft.com) with a maximum allowable alignment score set to 180. Only reads that were mapped uniquely to the reference were retained. We used Picard (http://broadinstitute.github.io/picard/) to add read groups and GATK [46] to perform re-alignment around indels. On average, 68% of the reads mapped to the targeted regions. Across the 104 individuals in our study, the average coverage of the targeted regions was 21x (range = 16-28x), with 99.3% of the targeted regions covered by at least one read and at least 20x coverage for 44% of the sites. After mapping, we filtered the data using the script SNPcleaner v2.2.4 (https://github.com/tplinderoth/ngsQC/tree/master/snpCleaner) following the protocol specified in Bi et al. [47], so that we only considered sites in which at least 73 (70%) of the individuals in our dataset had a coverage of at least 2x.

Population genetic analyses

We performed a variety of analyses using the program ANGSD [48], which conducts population genetic analyses from allele frequencies estimated from genotype likelihoods, and thereby takes genotype uncertainties into account [69, 70]. We used the program NgsRelate [49] as implemented in the program ANGSD to infer relatedness coefficients among individuals in lava and off-lava populations. This was done so that we could identify individuals with high relatedness coefficients, which do not provide independent genetic information from one another. We excluded ten Pisgah and five off-lava individuals so that all relatedness coefficients among the individuals in the dataset were low (i.e., below 0.1, where a coefficient of 0 indicates unrelated individuals and a coefficient of 1 indicates identical twins). We performed a principle components analysis on the genetic data using the program ngsTools [50] to test whether individuals from each population group together. Using ANGSD, we estimated the joint site frequency spectrum for the lava and off-lava populations to compare the allele frequencies of the two populations across all variable sites. We also used ANGSD to obtain estimates of allele frequencies, F_{ST} for each variable site, and the average F_{ST} across all sites.

Instead of using F_{ST} to identify divergent single nucleotide polymorphisms (SNPs), we used the association test based on allele frequencies implemented in ANGSD. This test implements the association mapping method that Kim et al. [71] designed for next-generation sequencing data, which takes uncertainty in genotype calls into account. The resulting *P*-value can in this context be interpreted as a *P*-value associated with the null hypothesis that the two populations are not diverged from each other. A small *P*-value indicates strong evidence of differences in allele frequencies between the two populations. We only considered sites with data for at least 52 of the 89 total individuals, a minor allele frequency greater than 0.05, that were estimated to be polymorphic with a *P*-value of 0.000001, and that did not reject Hardy-Weinberg equilibrium or exhibit strand bias at a *P*-value of 0.000001. We additionally filtered the reads in ANGSD so that we only considered reads with a minimum base quality of 20, and pairs of reads where both mapped concordantly. We regarded the genes with the strongest population associations as candidates for adaptive divergence. A simulation study has shown that when the average F_{ST} between populations is less than 0.1, there is a high probability (> 0.9) that a single locally adapted locus will be among the top ten most divergent loci among the populations [72]. However, we did not know *a priori* how many genes have adaptively diverged between the two populations, so we researched the function of the top twenty most divergent genes. Functional information was obtained from GeneCards (https://www.genecards.org) and literature searches, which we used to identify candidate genes that could affect the color phenotype.

Sequencing candidate genes

We designed primers around the SNP with the top association in each of the candidate genes using our exome reference sequence. We amplified the PRKAR1A gene using the forward primer PRKAR1A-F1: TGAAAGGTGATACTGGGATCAACACC and either of two reverse primers, which were PRKAR1A-R1: CTCACTGCCGGTACTCATAGAAGC and PRKAR1A-R2: TCTCCAGCCTTT CAAAGTATTCCCG. We amplified the PREP gene using one of two forward primers, which were PREP-F1: CCACCATCCTCCATTT CCTTCAGTC and PREP-F2: TCATGTGGCACTCCAGAAGTTTGG along with a single reverse primer PREP-R2: TCACCATCTT CACTGAATGCATAACCT.

All PCR reactions were 10 μL in total volume and contained 1X of PCR buffer, 2.5 mM $MgCl_2$, 0.1 mM of each dNTP, 0.5 μM of each primer, 0.5 U Taq polymerase (Invitrogen), and 20 ng of genomic DNA. A touchdown protocol was used in which thermocycling began with an initial denature at 94°C for 5 minutes, then had 20 cycles of 94°C for 30 s, an annealing temperature that began at 68°C for 30 s, but that decreased by 0.5°C per cycle, and an elongation at 72°C for 1 minute. Following the touchdown phase were 20 additional cycles where the annealing temperature was kept constant at 58°C. The PCR ended with a final hold at 72°C for 5 minutes. The above protocol was used for all primer combinations except PREP-F2 with PREP-R2, for which the annealing temperature began at 66°C and went down to 56°C.

We used ExoSAP-IT (USB, Cleveland, OH) to clean the PCR products. Sequencing was performed with a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and 1 μL of the purified PCR product. Sequencing reactions were cleaned with Sephadex and sequenced on an ABI 3730 sequencer (Applied Biosystems, Foster City, CA, USA). Sequences were aligned and edited using Geneious 8.1.9 (Biomatters Ltd.).

Scoring alleles in candidate genes

We used the Sanger sequencing protocol described above to score variants within two candidate genes in nine additional populations in southern California and central Arizona (Figure S2, Table S2). These populations were closely related to the Pisgah population in a previous phylogeographic study [61, 62] or were geographically close to the Pisgah population. One of these populations was the Amboy Crater, which is an independently formed volcanic crater that formed 70,000 ± 500 years ago [11]. The lava habitats are quite similar between the Pisgah Lava Flow and the Amboy Crater, so we sequenced a larger number of individuals (34) from Amboy than other populations to test if there might have been similar selection on the candidate genes in the two lava environments.

We Sanger sequenced the individuals used for exon-capture as well as additional samples from the Pisgah and off-lava populations to better measure the allele frequency of the SNPs within our two candidate genes (see Table S2). We also used this data in a Fisher's exact test in R to test for a correlation between the genotypes of our focal SNPs in each candidate gene (Table S3), which would indicate linkage disequilibrium between the two genes.

We downloaded the transcripts for the PREP gene (ENSACAT00000003886.2) and the PRKAR1A gene (ENSACAT00000013273.3) for *Anolis carolinensis* from ENSEMBL. We used Geneious to align the individual *A. carolinensis* exons to our *U. stansburiana* exon-capture reference sequences for PREP and PRKAR1A. We used the intron/exon boundaries for *A. carolinesis* as guides for determining whether the divergent SNPs in PREP and PRKAR1A were in coding or non-coding regions of the gene.

Estimating the demographic history and fitting models of selection on candidate genes

The joint site frequency spectrum (jSFS) that we estimated from our exon-capture data (see above) summarizes the correlation structure of the genetic variation between the Pisgah and off-lava populations, which can be used to infer their demographic history. We used FastSimCoal2 [51] to estimate the demographic history of these two populations given the observed jSFS. The demographic history consists of estimates of the population sizes, migration rates, and the split time between the two populations. We performed inference under two possible models: A) a population split with no subsequent migration, and B) a population split with a constant rate of subsequent migration. We assumed a constant population size in the ancestral population and constant population size in both populations after the split. We used the default per nucleotide mutation rate of 2.5×10^{-8} that is often

used for humans, because a mutation rate estimate for lizards was not available. For both models, we optimized parameters 100 separate times and picked the optimization run with the highest composite likelihood.

We inferred the age and selection coefficients of the focal SNPs in our candidate genes by simulating allele frequency changes over time (i.e., allele trajectories) under the most likely demographic history inferred in FastSimCoal2 for the models with and without migration. We allowed the alleles to be subject to divergent selection in the Pisgah and off-lava demes. We determined the probability that the simulated present-day allele frequencies in the Pisgah and off-lava demes fell within a 95% confidence interval (CI) of the population allele frequencies. We calculated these CIs using the Clopper-Pearson method, assuming the sample frequency is distributed binomially according to the population frequency. For example, if the CIs in the Pisgah and off-lava populations were [0.10-0.30] and [0-0.10] and a simulation had frequencies of 0.20 and 0.00, then we "accepted" the simulation and used it to increment our estimate of the likelihood; contrastingly, if the simulation had frequencies of 0.40 and 0.00, then we "rejected" it. For the demographic model with migration, we performed 10,000 simulations at each point in a 10x10x10 grid of various combinations of the values of the age of the allele, the selection coefficient in the Pisgah population, and the selection coefficient in the off-lava population. We did the same kind of simulations for the demographic model without migration except that we did not simulate the selection coefficient in the off-lava population. These simulations allowed us to identify the parameters where the present day allele frequencies fell in the 95% CI at the highest rate.

Testing for associations between genotypes and phenotypes

We utilized specimens from a genetic-crossing experiment performed with Pisgah and off-lava *U. stansburiana* [16], to test whether variation in our candidate genes was associated with differences in coloration. In the genetic-crossing experiment, male lizards were housed with a female from their own population and a female from outside their population to measure within and between population fertility. This was done for both Pisgah and off-lava males. The eggs resulting from the genetic crosses were kept in a temperature and humidity controlled incubation chamber until the hatchlings emerged (see [16] for further details). Protocols were approved by the U.C. Santa Cruz Institutional Animal Care and Use Committee (protocol SINE00.02-1).

The focal variants in our candidate genes were only observed in the Pisgah population, so we restricted our analysis to hatchlings that had at least one parent from the Pisgah population. We gathered data on 61 hatchlings, which were genotyped by sequencing each candidate gene as described above. Hatchlings came from 18 different Pisgah females and 5 different off-lava females. Three of the off-lava females were mated to the same Pisgah sires as three of the Pisgah females, given the genetic-crossing design. The other two off-lava females were mated to a Pisgah sire that did not father offspring with a Pisgah female. To maximize sample sizes, we included data from four Pisgah females that were gravid with eggs when caught in the wild and then mated to another male in the lab, which were only used in [16] to examine fertility of wild-mated individuals.

To quantify the coloration of the hatchlings, we photographed them with a Nikon D7100 camera with a 60 mm macro lens. The camera was suspended at a constant height above the specimens on a Kaiser RTX camera stand and the specimens were illuminated with a 650 W light. The photos were taken in raw format with a fixed shutter speed and aperture (f5, 1/200, ISO 100). The images were inspected to make sure that the photos of the specimens were not saturated. The specimens were photographed next to a Kodak Gray Card Plus color standard and we used the 3.1% reflectance black patch and the 18% reflectance gray patch to standardize the photos. We used the Image Calibration and Analysis Toolbox [52] in ImageJ 1.5i [53] to calculate the average luminance (i.e., lightness/darkness) of the area of each lizard's back from just in front of the forelimbs to just in front of the hind limbs. We summed the luminance measures from the red, blue, and green channels to provide a metric of overall luminance.

We tested for an association between hatchling genotypes and phenotypes using a linear mixed model as implemented in the R package lme4 [57]. To explain variation in hatchling coloration, our full model included the genotype of the hatchling for the two candidate genes as fixed factors and included the identity of the hatchling's dam and sire, which were included as random effects. In our analysis, we grouped individuals homozygous for the alternative derived allele with individuals heterozygous for the alternative derived allele. This was done because there was only a single hatchling that was homozygous for the derived allele for each of the two candidate genes, which gave us little power for estimating phenotypic effects for this class of individuals. We used a likelihood ratio test to compare the full model to a null model where the genotype data for one of the candidate genes was excluded. Variants contributing to the Pisgah melanic phenotype were predicted to make the lizards darker in coloration, so we used a one-tailed test.

QUANTIFICATION AND STATISTICAL ANALYSIS

Tests of phenotypic plasticity

We used an ANOVA in R [55] to test if dorsal coloration differed among pairs of treatments on day 0 of the experiment. We tested for plastic responses of dorsal coloration following a switch in the lava or sand substrate using a linear mixed effects model as implemented in the R package lme4 [57]. P values less than 0.05 were considered significant.

Population genetic analyses

We used the association test based on allele frequencies implemented in ANGSD [48] to identify divergent single nucleotide polymorphisms (SNPs). This test implements the association mapping method that Kim et al. [71] designed for next-generation sequencing data, which takes uncertainty in genotype calls into account. The resulting *P*-value can in this context be interpreted as a *P*-value associated with the null hypothesis that the two populations are not diverged from each other. A small *P*-value indicates

strong evidence of differences in allele frequencies between the two populations. We researched the function of the top twenty most divergent genes (Table S1).

Testing for associations between genotypes and phenotypes

We tested for an association between hatchling genotypes and phenotypes using a linear mixed model as implemented in the R package lme4 [57]. P values less than 0.05 were considered significant. Variants contributing to the Pisgah melanic phenotype were predicted to make the lizards darker in coloration, so we used a one-tailed test. We only gathered data for hatchlings that had at least one parent from the Pisgah population, because the focal variants in our candidate genes were only observed in the Pisgah population. In the genetic crossing experiment that generated the hatchlings [16], a subset of the eggs had yolk removed for potential analyses of hormones. We restricted our analyses of coloration to the hatchlings that came from unmanipulated eggs to avoid any potentially confounding effects of the manipulation.

Further information on all the statistical tests can be found in the Method Details section.

DATA AND SOFTWARE AVAILABILITY

The exon-capture dataset generated for this study is available from the NCBI SRA (BioProject ID: PRJNA476148; SRA accession: SRP150551; https://www.ncbi.nlm.nih.gov/sra/SRP150551).