PNAS RESEARCH ARTICLE | EVOLUTION

Genome-wide parallelism underlies contemporary adaptation in urban lizards

Kristin M. Winchell[a,b,c,1,2], Shane C. Campbell-Staton[a], Jonathan B. Losos[b,2], Liam J. Revell[c,d], Brian C. Verrelli[e], and Anthony J. Geneva[f]

Contributed by Jonathan B. Losos; received October 4, 2022; accepted December 8, 2022; reviewed by Rachael Bay and Diana J. Rennison

Urbanization drastically transforms landscapes, resulting in fragmentation, degradation, and the loss of local biodiversity. Yet, urban environments also offer opportunities to observe rapid evolutionary change in wild populations that survive and even thrive in these novel habitats. In many ways, cities represent replicated "natural experiments" in which geographically separated populations adaptively respond to similar selection pressures over rapid evolutionary timescales. Little is known, however, about the genetic basis of adaptive phenotypic differentiation in urban populations nor the extent to which phenotypic parallelism is reflected at the genomic level with signatures of parallel selection. Here, we analyzed the genomic underpinnings of parallel urban-associated phenotypic change in *Anolis cristatellus*, a small-bodied neotropical lizard found abundantly in both urbanized and forested environments. We show that phenotypic parallelism in response to parallel urban environmental change is underlain by genomic parallelism and identify candidate loci across the *Anolis* genome associated with this adaptive morphological divergence. Our findings point to polygenic selection on standing genetic variation as a key process to effectuate rapid morphological adaptation. Identified candidate loci represent several functions associated with skeletomuscular development, morphology, and human disease. Taken together, these results shed light on the genomic basis of complex morphological adaptations, provide insight into the role of contingency and determinism in adaptation to novel environments, and underscore the value of urban environments to address fundamental evolutionary questions.

urban evolution | rapid adaptation | parallelism | urbanization | *Anolis*

Significance

Urbanization drastically transforms landscapes worldwide, leading to altered eco-evolutionary dynamics. Many organisms are tolerant of, and even adapt to, these novel environments, presenting opportunities to study evolutionary change over rapid timescales. Here we provide a detailed investigation of the genomic basis of rapid adaptation in a species that thrives in urban environments. Integrating environmental, phenotypic, and genomic data, we demonstrate that populations exposed to similar environmental modification across distinct genetic clusters exhibit parallel phenotypic divergence underlain by parallel genomic divergence. We identify putative genomic targets of natural selection related to functionally relevant traits, thus helping to elucidate the mechanisms of rapid adaptive evolution of complex traits at the genomic level.

It is increasingly evident that humans influence ecological dynamics and evolutionary trajectories of organisms occupying human-dominated spaces (1–6). Abundant anthropogenic materials and structures combined with a deficiency of green spaces create novel biotic and abiotic conditions and complex socio-eco-evolutionary dynamics in cities (3, 7, 8). Studies have found wide-ranging functional, phenotypic, regulatory, and genomic consequences across diverse taxa (1), yet our understanding of evolutionary mechanisms in urban environments is nascent (9, 10). Among the central outstanding questions is to what extent phenotypic adaptations and parallelism are reflected at the genomic level (2, 9, 11). Consequently, the relative importance of contingency versus determinism in contemporary adaptation to urban environments remains underexplored.

While it is clear that urbanization is associated with phenotypic and genomic changes (1, 3), we still know very little about the genetic targets of selection underpinning adaptive urban trait shifts (12). Previous studies have identified genome-wide patterns of differentiation associated with urban environments (13–17). However, many urban genomics studies have focused primarily on nonadaptive differentiation or on genetic variation for which we do not understand the functional relevance, and in many cases, the phenotypic effects of identified signatures of selection at the genomic level remain unknown (12, 18). Some studies have highlighted specific genotype–phenotype associations, yet they focus largely on a priori candidate loci or inferred phenotypic associations via functional annotation (e.g., refs. 15, 17, and 19–21). Connecting environmental, phenotypic, and genomic changes is essential to understand the evolutionary processes shaping adaptations to novel environments. We address this knowledge gap by investigating genomic divergence associated with parallel urban environmental change and by identifying loci that may underlie adaptive urban phenotypes using the Puerto Rican crested anole (*Anolis cristatellus*).

Results

Background Population Structure. The Puerto Rican crested anole is a neotropical lizard that exploits urban niche space and has been well studied in both urban and

Author contributions: K.M.W., S.C.C.-S., L.J.R., and A.J.G. designed research; K.M.W. and A.J.G. performed research; K.M.W., S.C.C.-S., J.B.L., B.C.V., and A.J.G. analyzed data; and K.M.W., S.C.C.-S., J.B.L., L.J.R., B.C.V., and A.J.G. wrote the paper.

Reviewers: R.B., University of California, Davis; and D.J.R., University of California, San Diego.

The authors declare no competing interest.

[1]Present address: Department of Biology, New York University, New York, NY 10012.

[2]To whom correspondence may be addressed. Email: Losos@wustl.edu or Kristin.Winchell@NYU.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2216789120/-/DCSupplemental.

Published January 12, 2023.

nonurban environments, exhibiting repeated morphological and physiological adaptations (5, 14, 22). To evaluate genomic signatures of parallel urban adaptation, we sampled three paired urban–forest sites across Puerto Rico displaying parallel patterns of environmental divergence (Fig. 1 *A* and *B*; n = 96 individuals, n = 16 per population). We employed a custom semitargeted exon capture array to sequence coding regions (exons) of the *Anolis* genome. In a discriminant analysis of principal components (DAPC), urban and forest populations clustered by geographic region (San Juan, Arecibo, and Mayagüez; Fig. 1*C*) and not by habitat (urban versus forest), in line with evidence from mtDNA (5) and RNAseq data (14). Multiple lines of evidence indicate that nearby urban and forest populations share more genetic diversity than do lizards from urban or forest sites in different parts of the species' range (Fig. 1 *C* and *D* and *SI Appendix*, Figs. S1–S3). Analyses of relatedness (Fig. 1*D* and *SI Appendix*, Fig. S2) indicate that individuals within a population tend to be more closely related to each other compared to individuals within the paired site in the same municipality (ANOVA: $F_{df=1,718}$ =78.6, $P < 0.001$; *SI Appendix*, Fig. S2) or to individuals from any other population (ANOVA: $F_{df=1,4558}$ =2160.4, $P < 0.001$). These results provide robust support for urban populations arising repeatedly and independently across the island.

Urban-Associated Divergence. To search for outlier loci displaying genomic divergence between urban and forest habitats, we used two complementary approaches. We first employed a genotype–environment association test (GEA) to identify loci associated with urbanization across all three municipalities while accounting for underlying island-wide population structure (Fig. 2*A* and *SI Appendix*, Fig. S4). Additionally, we used a principal components-based genetic outlier analysis for each urban–forest pair (Fig. 2 *B–D*, "PCA analyses") to detect genomic regions of unusually high differentiation among individuals while accounting for population structure among sites. Genes containing at least one single nucleotide polymorphism (SNP) identified as outliers across all analyses were deemed the strongest candidates for selection in urban environments.

In total, we identified 91 total variants (out of 100 outlier SNPs identified by the PCA analyses and 1153 outlier SNPs identified by the GEA analysis) across three regions of the *Anolis* genome that met the following criteria: a 4.4-Mbp region of CHR1, a 4.4-Mbp region of CHR2, and a 34.5-kbp region of CHR4 (Fig. 2 *A–E* and *SI Appendix*, Fig. S5). Of these 91 variants, 33% were in focal exons (see *Methods*). The observed overlap between the GEA and PCA outlier SNPs is significantly greater than expected by chance (permutation test, $P < 0.001$, *SI Appendix*, Fig. S6) and represents 0.078% of the total number of SNPs tested (115,976). These patterns are consistent with selection in urban environments repeatedly targeting three specific genomic regions containing 33 genes across all three municipalities (Fig. 2*F*). The observed overlap between the GEA and PCA at the gene level is significantly greater than expected by chance (permutation test, $P = 0.006$, *SI Appendix*, Fig. S6).

In addition, several genomic regions display idiosyncratic genomic differentiation within a single municipality (Fig. 2 *B–D*), which may reflect regional environmental differences, divergent selection pressures among cities, or geographic variation in initial standing genetic variation of the urban populations. In fact, only 3% of the outlier SNPs identified in the PCA analyses were outliers in all three municipalities and 19% were outliers in 2 of the 3 municipalities (*SI Appendix*, Fig. S7). We found more overlapping candidate SNPs between Arecibo and San Juan than between the other regions, a result that may be explained by the

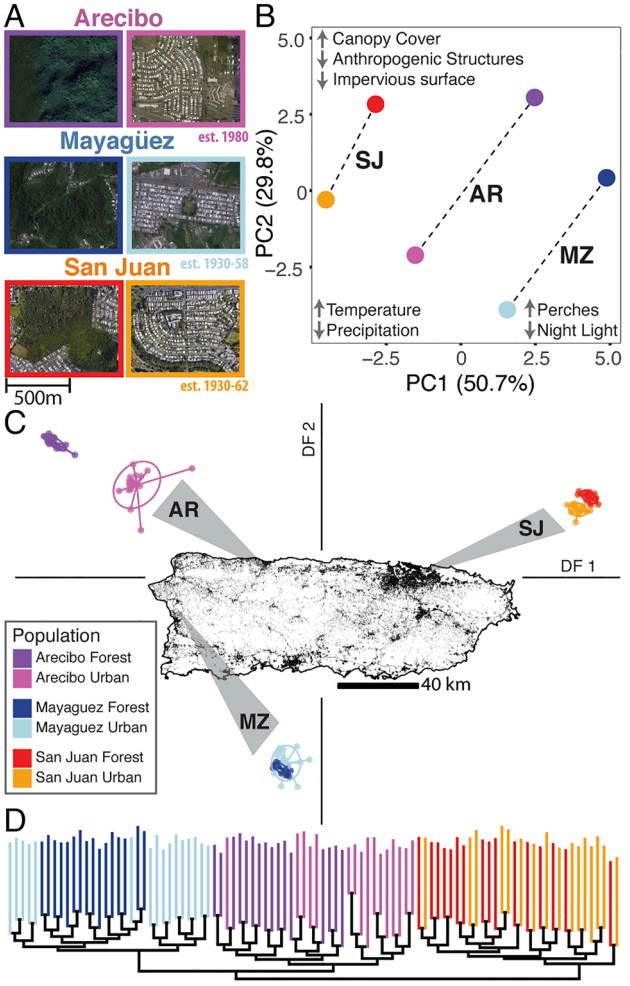

Fig. 1. Environmental and population divergence. We sampled paired urban–forest sites in three municipalities (regions) across the island of Puerto Rico. Population structure analyses support independent urban–forest pairs in each geographic region. Across all panels, colors correspond to municipality and site, as follows: Arecibo, urban: pink, forest: purple; Mayagüez, urban: light blue, forest: dark blue; San Juan: urban: orange, forest: red. (*A*) Satellite imagery of forest and urban sites sampled in each municipality (images: Google Earth and Maxar Technologies 2001). Estimated dates of urban establishment are indicated below the urban images. (*B*) Urban and forest habitats differ in parallel in multidimensional habitat space, with urban environments characterized by substantially reduced tree cover, extensive impervious surface cover, warmer and drier climate, artificial light at night, and abundant anthropogenic structures. Principal components analysis of habitat indicates parallel shifts across the three municipalities in multivariate habitat space between urban and forest sites. (*C*) DAPC of genomic variation overlaid on a map of Puerto Rico showing urbanization extent (23) in black. Individual samples are colored by site. Gray triangles indicate geographic locations of municipalities sampled. (*D*) Midpoint-rooted sample tree, with individual samples colored by site. Individuals from within each region (but not necessarily each habitat type within region) were more genetically similar to one another on average than to individuals from other regions.

environmental similarity of these two municipalities. Specifically, San Juan and Arecibo are cooler, receive less precipitation, and have more large trees contributing to more extensive canopy cover compared to Mayagüez (Fig. 1*B*). Taken together, these analyses highlight the complex interaction of determinism and contingency in shaping adaptive genomic divergence in urban environments.

The wide variety of environmental changes associated with urbanization present myriad opportunities for phenotypic adaptations that enable colonization and persistence in these novel habitats. For example, abundant anthropogenic food resources could affect dietary and metabolic processes (24); altered disease

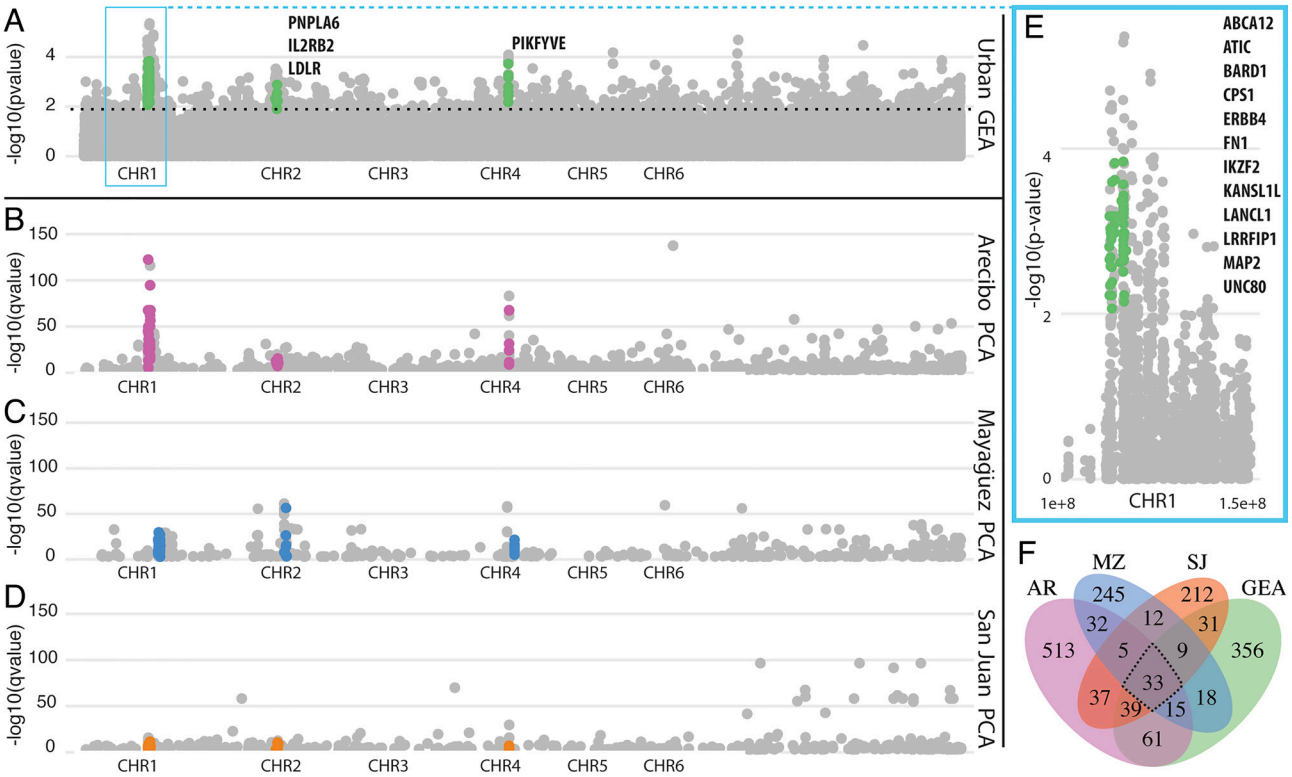

Fig. 2. Parallelism of urban-associated genomic changes. (*A*) Manhattan plot of SNPs identified by the urban genotype–environment association test (GEA), with the significance threshold indicated by the black dotted line and genes containing shared outlier SNPs listed next to the peaks for chromosomes 2 and 4. We complemented this analysis with three PCAs, one for each municipality: (*B*) Arecibo, (*C*) Mayagüez, and (*D*) San Juan. Colored points in each Manhattan plot are the 91 SNPs identified in all four tests, and all outlier SNPs are shown in *B*–*D* in gray. (*E*) The peak on chromosome 1 identified by the blue rectangle is shown in greater detail with genes containing shared outlier SNPs across the GEA and PCA analyses listed. (*F*) Venn diagram of overlap in genes containing outlier SNPs across the three municipalities in the PCA analyses and the GEA. The 33 urban-associated genes contained outlier SNPs in all four tests. Larger versions of all Manhattan plots are in *SI Appendix*, Fig. S5.

dynamics might impact immune function (25); novel resources could challenge cognitive abilities and behaviors (26); altered structural environments may influence locomotor morphology (27); and urban heat islands could challenge thermal and desiccation tolerances (14, 28).

We explored potential functions associated with the 33 urban-associated genes (Fig. 2*F* and *SI Appendix*, Table S8). We found a single significantly enriched term from the KEGG database (Glutathione metabolism), and a gene ontology analysis highlights immune functions, wound healing, and inflammatory responses, which may indicate habitat-specific differences that necessitate selection on injury recovery and immunocompetence. Urban wildlife often face disease dynamics and stressors different from those experienced by their nonurban counterparts, resulting in selection on immune function and stress response (25,29). Indeed, previous research has established that urban anoles exhibit elevated injury rates, including bone fractures, missing digits, and autotomized tails (30, 31) as well as increased parasite infections (32). In addition, several of these genes have been implicated in neural function and motor regulation [e.g., MAP2 (33), UNC80 (34), ZSWIM4 (35), and PNPLA6 (36)], metabolic function [e.g., LDLR (37), ATIC (38), and CPS1 (39)], skin development [e.g., ABCA12 (40)], and epithelial pigmentation [e.g., MREG (41)]. These results highlight a number of potential functional targets of selection in cities. However, explicit connections with higher levels of biological hierarchy are needed to understand the specific phenotypic consequences of parallel urban selection observed at the genomic level.

Genomic Underpinnings of Urban Phenotypes. We next conducted a series of genome-wide association studies (GWAS) to identify loci associated with morphological features that display evidence of adaptive differentiation in urban environments (Fig. 3 *A*–*E*). Uncovering the genetic basis of limb and toepad morphology has been a long-sought-after goal given the critical role of these traits in the adaptive radiation of this genus (22). Previous work has implicated gene expression in adaptive interspecific variation in limb development in *Anolis* (42, 43). We identified 2,908 genes in the *A. cristatellus* genome associated with variation in limb length and toepad morphology. These genes were significantly enriched for 126 Gene Ontology terms (*SI Appendix*, Tables S9-S10), including cellular components and biological processes primarily related to cellular function and the nervous system. Morphology-associated genes were also enriched for terms in the Human Phenotype Ontology, many of which are related to the formation and function of limb and toepad traits and locomotion and disease, such as "Abnormality of the musculoskeletal system," "Anatomical structural development," "Gait disturbance,""Abnormality of the musculature of the limbs," "Limb muscle weakness," "Abnormality of skeletal morphology," "Generalized abnormality of skin," and "Abnormality of movement." Our findings indicate a polygenic basis for intraspecific limb length variation and suggest that selection acts on multiple genomic targets to shape complex phenotypes over short evolutionary timescales.

To identify potential genomic targets of selection involved in urban limb and toepad morphology, we leveraged the rapid and consistent morphological shifts associated with urbanization in *A.*

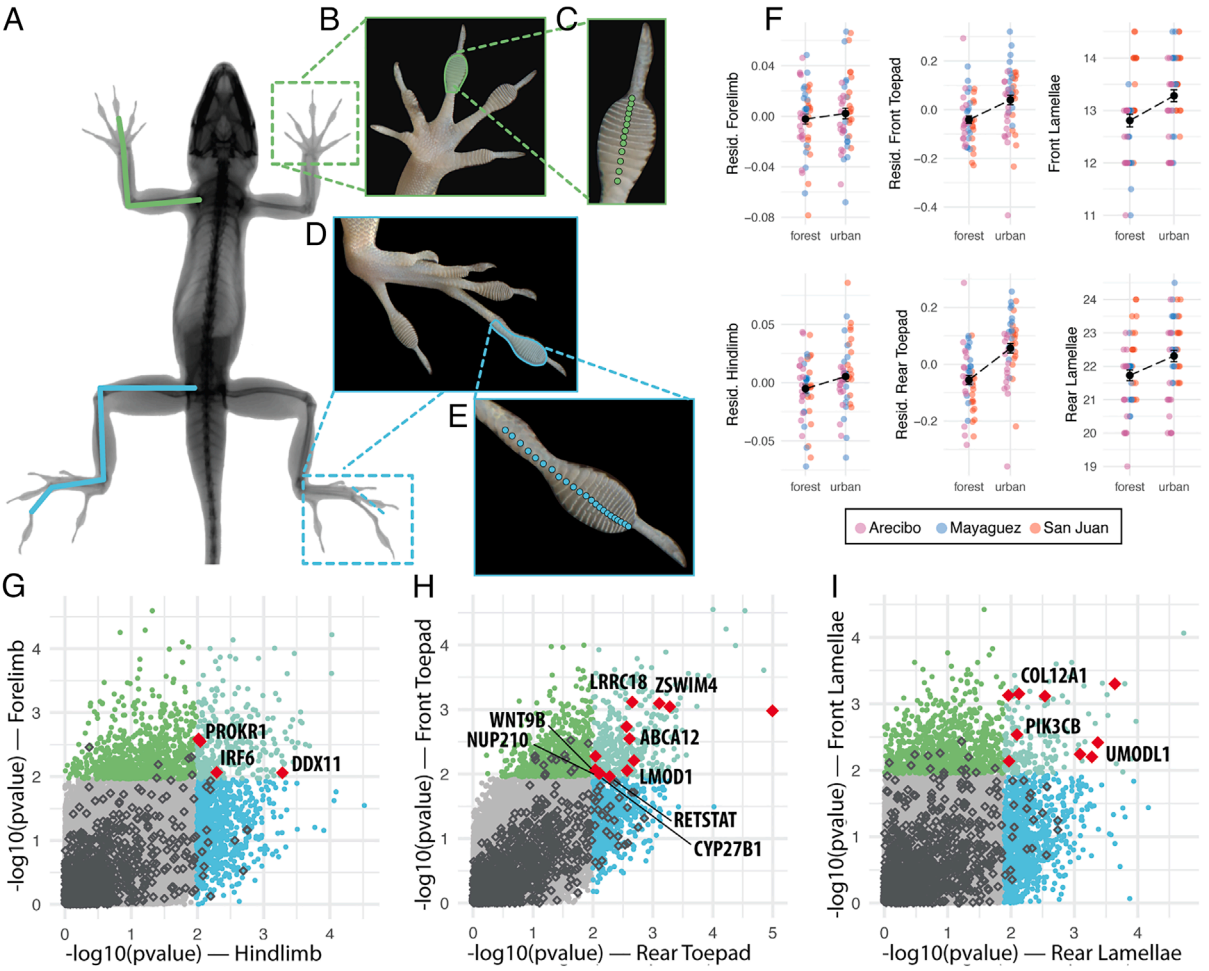

Fig. 3. Phenotypic parallelism and genomic underpinnings. We focused on six morphological traits with known urban-associated divergence: (*A*) forelimb and hindlimb lengths, (*B*) front toepad area, (*C*) front toepad lamella count, (*D*) rear toepad area, (*E*) rear toepad lamella count. (*F*) At the phenotypic level: mean and SE for each trait across all populations by habitat type (urban, forest) with individuals colored by municipality: pink, Arecibo; blue, Mayagüez; red, San Juan, with mean and SE by habitat shown in black. At the genomic level: overlap in outlier SNPs for each of the three traits between posterior (blue points) and anterior elements (green points) for each trait: (*G*) limbs, (*H*) toepad area, and (*I*) toepad lamellae; SNPs associated with both hindlimb and forelimb elements are indicated in the upper right quadrant (teal points). Outlier SNPs associated with urbanization (GEA analysis) are shown as hollow gray diamonds, with filled red diamonds indicating urban SNPs that also overlap with both forelimb and hindlimb morphological elements. Gene names correspond to one or more of the urban-morphology SNPs in the upper right quadrant (red diamond).

cristatellus. Previous studies have shown parallel divergence in limb length and toepad morphology across urban and forest populations of this species associated with differences in structural habitat (5, 44). Therefore, we measured forelimb and hindlimb lengths, toepad area, and lamella number. These data show that urban populations exhibit parallel increases in all six traits (Fig. 3*F* and *SI Appendix*, Fig. S11), consistent with previous studies. These phenotypic differences have been shown to translate into differences in locomotor performance (44, 45), supporting locomotor morphology as a likely target of adaptation in urban anoles.

We identified 154 loci as outliers in both habitat and phenotypic association tests. Of these variants, 20% were in focal exons (see *Methods*). The observed overlap between the GEA and GWAS outlier SNPs is significantly greater than expected by chance for all six morphological traits (forelimb: N = 10 SNPs, $P < 0.001$; hindlimb: N = 17 SNPs, $P < 0.001$; front toepad: N = 27 SNPs, $P < 0.001$; rear toepad: N = 42, $P < 0.001$; front lamellae: N = 26, $P < 0.001$; rear lamellae: N = 32, $P < 0.001$; *SI Appendix*, Fig. S6). These loci represent the strongest candidates underpinning adaptive urban phenotypes (Fig. 3 *G*–*I*). The large number of genomic targets highlights polygenic selection on standing genetic variation as a key process underlying rapid morphological adaptation to urban structural environments.

We further narrowed the candidate set to the 93 genes associated with both anterior and posterior elements of each trait (*SI Appendix*, Table S9, intersection at the SNP level in Fig. 3 *G*–*I*). These 93 candidate genes represent several functions associated with skeletomuscular development, morphology, and disease. Urban-limb-associated genes are involved in angiogenesis of peripheral limbs [PROKR1 (46)] and are implicated in diseases in humans and mice involving shortened and deformed limbs [IRF6 (47) and DDX11 (48)]. Genes associated with urban-toepad morphology are involved in the development of keratin, collagen, and skin [e.g., ABCA12 (40), CYP27B1 (49), and COL12A1 (50)], major components of the anole epidermis and scales (51), as well as smooth muscle contraction [LMOD1 (52)], which is involved in the conformation and release of toepads from surfaces (53). In addition, 82 of the candidate genes have Gene Card (54) entries, of which 22 include limb or limb bone terms in their phenotype and 22 reference skin phenotypes. Notable among these are several genes involved in bone formation, differentiation, elongation, and pathology of limbs in humans and mice:

BRD4 (55), CYP27B1 (56), FLNB (57), FN1 (58), and IRF6 (47). The presence of several disease-associated genes (previously identified in other vertebrates) among our candidate loci points to genomic targets of large—and potentially deleterious—phenotypic effects to bring about rapid morphological change.

Phenotypic Parallelism Mirrored at the Genomic Level. To the extent that cities are altered in similar ways, we might expect parallel selection pressures to result in parallel phenotypic adaptations across urban populations, which may be reflected at the genomic level (16, 21, 59). A handful of studies have documented phenotypic and, to a lesser extent, regulatory and genetic parallelism across urban populations, yet idiosyncrasy of adaptive responses is also common (3). We identified common genomic targets of selection for urban-associated morphological divergence across municipalities using two approaches. We first tested for polygenic parallelism by performing PCAs on a subset of the original dataset containing only outlier genomic regions ("local PCA"), an analysis which can provide insight into whether haplotypes are similarly diverging across urban–forest pairs. We then tested each allele for parallel responses to urbanization by comparing the effect size of habitat to the effect size of the interaction term (habitat x municipality) in a linear model for each SNP.

We found parallel shifts in the primary axis of genetic variation (eigenvector 1) associated with each trait across the three municipalities (Fig. 4 *A*–*C*; ANOVA habitat effect: forelimb $F_{df = 1, 90}$ = 88.4, $P < 0.001$; hindlimb $F_{df = 1, 90}$ = 59.3, $P < 0.001$; front lamellae $F_{df = 1, 90}$ = 100.0, $P < 0.001$; rear lamellae $F_{df = 1, 90}$ = 45.2, $P < 0.001$; front toepad area $F_{df = 1, 90}$ = 51.8, $P < 0.001$, and rear toepad area $F_{df = 1, 90}$ = 143.6, $P < 0.001$; full ANOVA results in *SI Appendix*, Fig. S12). Similarly, at the allelic level, we find that 88% of urban-morphology SNPs are diverging in parallel, with a greater habitat than interaction effect (Fig. 4 *D*–*F*). A greater habitat effect size compared to the interaction term effect size supports parallelism of the genotype, whereas a greater interaction term would suggest that the alleles differ between urban and forest pairs in different ways across municipalities. We confirmed the robustness of our method by comparing the difference between the habitat and interaction effects to the null expectation based on the background (nonoutlier) set of SNPs, finding that each trait had a significantly greater effect of habitat compared to neutral genetic variation (Fig. 4*G* and *SI Appendix*, Fig. S12) (two-sided *t* test; forelimb: t = 5.19, df = 9, $P = 0.0006$; hindlimb: t = 3.34, df = 16, $P = 0.004$; front toepad: t = 6.20, df = 26, $P = 1.5e^{-6}$; rear toepad: t = 5.51, df = 41, $P = 3.2^{-9}$; front lamellae: t = 12.64, df = 25, $P = 2.3e^{-12}$; rear lamellae: t = 8.88, df = 31, $P = 5.1e^{-10}$). Together, these results indicate that adaptive divergence associated with urban morphology is occurring via repeated selection on similar regions of the genome across the three geographic regions.

Discussion

This study provides a detailed investigation of the genomic basis of rapid adaptation in a species that thrives in urban environments, identifying putative genomic targets of natural selection related to functionally relevant phenotypes and helping to elucidate the mechanisms of rapid adaptive evolution of complex phenotypes at the genomic level. We found that populations of urban anole lizards exposed to similar environmental modification across distinct genetic clusters exhibited parallel signatures of selection associated with urbanization and urban-associated morphological divergence in coding regions of the genome. Our findings contribute uniquely to the growing field of urban evolutionary ecology and, more broadly, to our understanding of rapid and contemporary adaptation in three key ways.

First, of considerable interest in evolutionary ecology is the question of whether parallelism at phenotypic levels is mirrored at the genomic level. Here, we connect parallel environmental divergence with parallel phenotypic divergence underlain by parallel genomic divergence. Genomic parallelism has rarely been demonstrated in response to urbanization (3) and rarely connects parallelism at environmental, phenotypic, and genomic levels [with a few recent exceptions (14, 16, 21)]. Theory predicts that adaptive evolution in closely related populations is more likely to arise via parallel genomic change (60, 61). Indeed, we observe multiple crested anole populations using similar genomic regions in their adaptation to urban environments, whereas more distantly related species do not experience genomic parallelism across the adaptive radiation of *Anolis* (62). Our study supports the hypothesis that cities can act as replicated natural laboratories with respect to their selective pressures and evolutionary outcomes. Consequently, we may be able to predict population responses to urbanization based on genetic markers.

Second, an outstanding goal in the study of contemporary evolution is understanding the genomic basis of adaptation to novel environments. Adaptive traits may be shaped by gene expression variation as well as coding sequence variation or covariance contributed by both (63). For complex traits with many underlying genes, we might predict gene expression variation to be a more likely mechanism to accomplish rapid adaptive trait shifts, as they can be more subtle with respect to their effects on phenotype (63). On the other hand, while it is indeed possible that cryptic amino acid variation may be segregating within populations, these changes in coding regions tend to be more restrictive and more likely to be deleterious than adaptive with respect to function compared to regulatory variation (63). Evidence suggests that changes in gene expression underlie some urban adaptations, such as thermal tolerance (14) and insecticide resistance (64), whereas changes to coding regions underlie others such as harm avoidance (19, 65). We demonstrate that adaptive changes in complex morphological phenotypes can also be associated with changes in protein-coding genes. Understanding the genomic basis of adaptations to novel environments will shed light on the constraints on evolvability that facilitate or inhibit parallel adaptation across populations experiencing similar selective pressures, particularly since urban adaptation can be relatively rapid (6, 66).

Lastly, sequence-based models have shown that mutations in evolutionarily conserved genes are more likely to result in deleterious phenotypes and disease (67). We find that several loci associated with adaptive morphological changes are implicated in disease phenotypes in humans and other organisms, suggesting that the variation we have identified here underlying adaptive phenotypes may be deleterious in nonurban settings but beneficial in urban environments. This pattern may seem paradoxical, but it has also been shown in previous studies that genes most closely tied to functional relevance may also represent candidates for maximizing fitness across diverse environments, such as variation for immunity, diet and subsistence, and bone development linked to positive selection in humans (68–70). Our observation here with anoles opens the possibility that genes of high evolutionary conservation could also be involved in adaptation to urban environments and worth pursuing in the future. Consequently, we suggest that the study of rapid adaptation to novel environments, and specifically urban adaptation, should not focus solely on malleable gene regions but also on mutational targets with potentially large effects. Genetic variation resulting in large phenotypic effects may facilitate population shifts to alternative fitness peaks under rapid

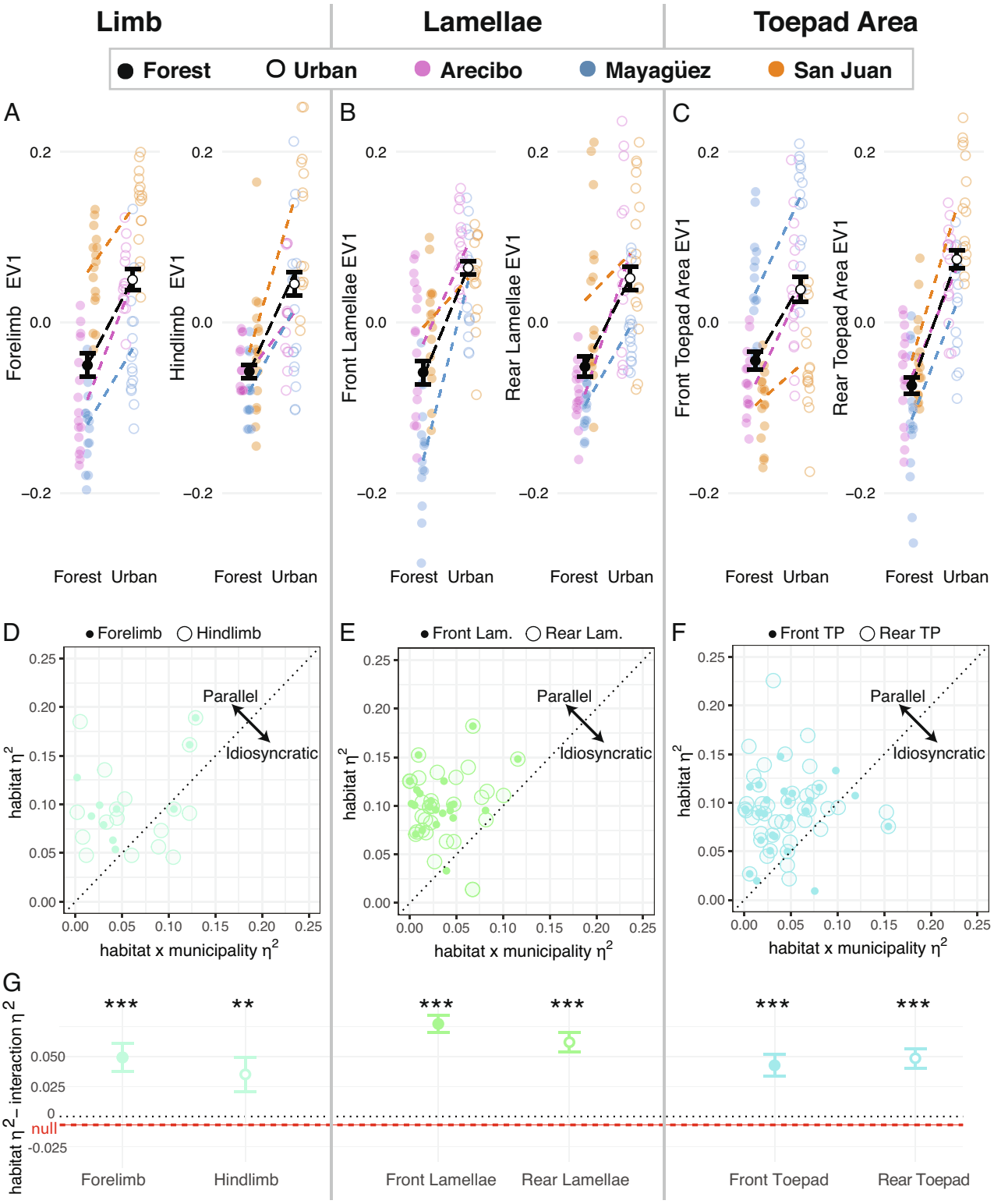

Fig. 4. Parallelism of genomic architecture of urban morphology. We performed local PCAs of outlier SNPs for each of the six morphological traits. The first axis of genomic variation summarized by each PCA (eigenvector 1) indicates parallel genomic change in the urban–forest pairs across the three municipalities for (*A*), forelimb length (FL) and hindlimb length (HL); (*B*) front lamellae (FLAM) and rear lamellae (RLAM), (*C*) front toepad area (FTP) and rear toepad area (RTP). In each plot *A*–*C* colored points indicate individuals colored by population, and black/white points indicate the mean and SE across all population pairs. We also examined allele-level divergence across the three urban–forest pairs, summarized by the effect sizes (partial eta, η^2) of habitat and the interaction of habitat by municipality, where a greater effect size of habitat versus the interaction effect (points above the black 1:1 dashed line) indicates a parallel response associated with urbanization, whereas a greater interaction effect (points below the black dashed line) indicate municipality specific, idiosyncratic divergence between urban and forest populations (e.g., local adaptation). Front (filled points) and rear (hollow points) elements for each trait are shown in each plot for (*D*) limb length, (*E*) lamellae, and (*F*) toepad area. (*G*) The mean and SE of the difference between the habitat and interaction effect sizes for each trait is compared to the null expectation (mean effect size of background SNPs; red dashed line). As in *D*–*F* the black dotted line indicates equal effect size of habitat (urban) and municipality specific divergence. Significance levels by the two-sided *t* test against the null expectation: $P = 0.01$**, $P < 0.001$***.

anthropogenic change and may play a much greater role in contemporary evolution than currently appreciated.

Materials and Methods

All R analyses were completed in R version 4.0.3 (2020-10-10).

Field Methods. *Anolis cristatellus* is a neotropical lizard native to the island of Puerto Rico, has an island-wide distribution, and is commonly encountered in both urban and forest environments. Deep mtDNA breaks exist between populations distributed across Puerto Rico with mitochondrial clades associated with Southern, Northeastern, and Northwestern regions of the island (71). Between 2012 and 2014, we captured adult male *A. cristatellus* from paired urban and forest sites in three municipalities across Puerto Rico (San Juan, Arecibo, and Mayagüez) as part of ongoing research on urban ecology and evolution. Lizards were captured as encountered (using floss lasso or by hand), without preferentially capturing lizards in specific microhabitats (e.g., on buildings versus vegetation). Although sites were sampled in different years, no site was sampled during more than one sampling period (*SI Appendix*, Fig. S13). Age and sex of lizards were determined based on snout–vent length (minimum SVL 45 mm) and secondary sexual characteristics (large dewlap, enlarged postanal scales, and enlarged tail base).

We collected a sample of the distal tail (~10 mm) from each lizard and preserved the tissue in 95% EtOH. Tissue samples were transported to the University of Massachusetts Boston and were stored at −80 °C. All lizards were transported to a field laboratory where we obtained skeletal X-ray (with a portable Kodex digital X-ray system) and high-resolution toepad scans (with an Epson flatbed scanner at 2,100 dpi). All lizards were returned to their point of capture following data collection. We selected 16 individuals from each population for inclusion in this study based on availability and quality of digital morphological data and tissues, without consideration of phenotype.

Morphological Measurements. Morphological traits were measured using ImageJ (72) using the ObjectJ plugin. Limb bones and SVL were measured three times each from digital X-ray. Replicate measurements for each skeletal element from both left and right limb elements were averaged, excluding any bones showing evidence of fractures, which impact bone length (both recent and healed fractures are visible on X-ray), as in previous studies (5, 27). No individuals were excluded from any analyses because of bilateral limb damage. Forelimbs were calculated as the sum of bone lengths for the third metacarpal, ulna, and humerus. Hindlimbs were calculated as the sum of bone lengths for the first phalanx of the fourth digit, fourth metatarsal, tibia, and femur. Toepad lamellae on the third forelimb digit and fourth hindlimb digit were counted three times in ImageJ; counts for right and left digits were averaged. The expanded toepad defined by the lamellae was traced three times, with replicate areas for both right and left toes averaged. Toepads that were damaged were excluded, and one individual was excluded from front toepad analyses and one from rear toepad analyses because of bilateral damage. Limb lengths and toepad area were size-adjusted by taking the residuals of the relationship between each natural-log–transformed trait and natural-log–transformed SVL.

Repeatability of measurement, estimated by the intraclass correlation coefficient in R with the function *ICCest* in R package "ICC" (73), was high for all traits (ICC = 0.97 for metacarpals and phalanges; ICC = 0.99 for all others). We evaluated normality of each trait using the Shapiro–Wilk test of normality, implemented in R base package "stats" with the function *Shapiro.test*. All traits were normally distributed (forelimb: $W = 0.99$, $P = 0.721$; hindlimb: $W = 0.98$, $P = 0.171$; front lamellae: $W = 0.98$, $P = 0.158$; rear lamellae: $W = 0.99$, $P = 0.361$; front toepad area: $W = 0.99$, $P = 0.361$; rear toepad area: $W = 0.99$, $P = 0.700$). Partial effect size (η_p^2) for habitat and the habitat by municipality interaction were calculated using the function *partial_eta_squared* in R package "rstatix" (74). All traits followed the same patterns reported previously (5, 27) (note that San Juan populations are a subset of the individuals from ref. 5). Urban lizards have longer forelimbs and hindlimbs, larger front and rear toepads, and front and rear toepads with more lamellae scales (Fig. 3*F*).

Evaluating Urbanization. Sites were selected nonrandomly for sampling based on apparent conformance to an urban–forest dichotomy, as well as for logistical reasons (such as investigator safety and access), with urban sites dominated by anthropogenic structures and impervious surfaces and forest sites characterized by extensive tree canopy cover and minimal human disturbance. Establishing when a site transitioned to a human-dominated "urban" habitat can be challenging given the age of the municipalities sampled: Arecibo was founded in 1616, Mayagüez in 1760, and San Juan in 1509; thus, the influence of urbanization may extend 250 to 500 y. However, at the local site scale, we estimate that the urban areas sampled range in minimum age (based on aerial imagery and landmark establishment) from approximately 1960 (Mayagüez, San Juan) to 1980 (Arecibo). With generation times in *A. cristatellus* commonly assumed to be approximately 12 mo (e.g., refs. 71 and 75), this time period is equivalent to at least 32 generations in Arecibo, 50 in San Juan, and 55 in Mayagüez. We quantitatively evaluated urbanization using landscape and climate data for each site, extracted in ArcGIS (ESRI 2020), followed by principal components analysis in R.

To quantify urbanization across our sites, we included 25 landscape variables to describe climatic and structural site variation. We included all 19 BIOCLIM (76) climate layers (v2.1) and light at night global radiance (light at night: NOAA Global Radiance Calibrated Nighttime Lights F16_20100111-20110731_rad_v4 GeoTIFF, https://ngdc.noaa.gov/eog/dmsp/download_radcal.html). Because both of these datasets are at a larger resolution (~1 km^2) and summarize variables that have diffuse effects across the landscape (e.g., effects of light at night are not constrained by site boundaries), we extracted site-level averages from each sampled area plus a 1-km buffer around the perimeter. We also included two higher-resolution (30 m^2) land-cover layers: impervious surface cover (77) and canopy cover (78), from which we extracted site-level averages within the boundaries of each sampled site. Lastly, to describe local-scale structural habitat, which has previously been shown to be relevant for urban anole morphology (5, 27, 79), we followed the procedure in Prado-Irwin et al. (79) to quantify perch availability, habitat openness, and anthropogenic perch presence. We obtained orthoimagery (80) of each site and placed a 200-m buffer around the centroid of each sampled area. Within each of these size-standardized sampling areas, we distributed 50 random points with a minimum distance of 5 m between each point in ArcGIS. We then counted the number of points that were located on a potential perch (any structure, vegetation or anthropogenic in nature) and counted how many points fell on an anthropogenic structure (building, fence, poles, etc.) as our measures of perch availability and anthropogenic perch availability. For any point that fell on a structure, the distance to the nearest perch was 0; for all others, we calculated the distance between the random point and the edge of the nearest structure.

We conducted a principal components analysis on the site-level averages for the 25 environmental variables to summarize urbanization across our six sites (*SI Appendix*, Fig. S13). The first three principal components captured 96.9% of variance. Higher values of PC1 indicated colder, wetter, and more variable climate (i.e., BIO1, BIO2, BIO6, BIO7, BIO9, BIO10, BIO11, BIO13, BIO15, and BIO16) as well as more perches and less light at night. Higher values of PC2 indicated colder and wetter climate (i.e., BIO3, BIO4, BIO5, BIO8, BIO12, BIO14, BIO17, and BIO19), less impervious surfaces and anthropogenic structures, less habitat openness, and more canopy cover. Higher values of PC3 indicated drier and warmer climates with more variable temperatures (i.e., BIO2, BIO3, BOI5, BIO7, BIO12, BIO13, BIO15, BIO16, and BIO18) as well as less light at night (*SI Appendix*, Fig. S13). Urban and forest sites diverged in parallel in climatic and structural habitat variation (Fig. 1). Urban and forest sites differed in PC1 ($\chi^2 = 7.01$, $P = 0.008$) and PC2 ($\chi^2 = 10.60$, $P = 0.001$), but not in PC3 ($\chi^2 = 3.02$, $P = 0.082$), across all sites (linear-mixed effects models, LRT).

Molecular Methods. We extracted whole genomic DNA from a total of 96 samples (n = 16 per population) of homogenized tail tissue using Wizard SV Genomic Purification Kits. We made the following modifications to the extraction protocol to improve DNA yield: Tissues were digested for 24 h with an additional 5 μL of proteinase K added after 12 h; the total elution volume was reduced to 70 μL, and samples were washed with dH20 and the flow through elution. We verified DNA concentration fluorometrically using a Qubit v2.0 and the presence of genomic DNA with minimal degradation with gel electrophoresis.

We designed a custom exon capture bait set to selectively target portions of the exome. We targeted only exons >120 bp (the length of the bait) and with GC content 40 to 70%. We identified exons to target in two complementary sets of focal and nonfocal exons. We developed a data pipeline to identify exomic regions to target for the "focal" set based on relevant Gene Ontology terms. We used AmiGo to search the annotated *Anolis carolinensis* genome (AnoCar2.0) (81,82) for the following keywords: water loss, therm*, temperature, stress, skelet*,

sensory, scale, ossify*, muscle, metabol* locomotion, limb, immun*, growth, feed, fear, epithel*, epiderm*, diet, dietary, desiccation, dehydrat*, color, cognit*, brain, bone, behavior; downloading the gene and exon data for all results via Ensembl and Biomart, resulting in a set of ~30,000 of the 200,000 exons in the *A. carolinensis* exome. We manually curated this list and ranked each sequence by relevance and priority for sequencing based on the gene ontology description. For example, exons that mapped to many or widely varying functions that were not part of our targeted search were excluded from the focal set. We identified a subset of 1,600 exons to target for sequencing as our "focal exons." To this list, we also added exons in the RARS gene region (chromosome 1:112550768-112574864), which was previously identified as a target of selection in urban heat islands in *A. cristatellus* (14).

We next identified all genes not represented by at least one exon in the focal exon set. We targeted the first exon from each of these genes meeting capture criteria (>120 bp, 40 to 70% GC), and randomly distributed the remaining probes throughout the remainder of the exome ("nonfocal exons"). Thus, we covered the entire exome but varied our sequencing strategy based on whether or not we expected the gene to be relevant for urban adaptation. Our final capture array targeted at least one exon per gene across the entire exome with more exons targeted in regions of high interest ("focal exons"). Mitochondrial genes were excluded from the capture design.

Bait design was performed by RAPiD Genomics. Probes were screened against the *A. carolinensis* genome (82) and a draft genome assembly of *A. sagrei* (83), and targets that would not be likely to capture because of GC content, would overcapture across the genome, or could not be mapped were removed. Specifically, probes were limited to those with no more than two hits at 85% identity and greater than 80 bp to either genome (*A. carolinensis* or *A. sagrei*) to keep a tight capture and return optimum results. After this filtering, the final probe set was designed to capture a total of 6,781 focal exons using 16,284 probes spanning an average of 82% of each exon. The remaining 40,715 probes were distributed across the nonfocal exon set, focusing on maximizing the number of genes hit and only placing two probes in large scaffolds with 62% of each exon covered by probes, on average. Although we specifically targeted RARS in our candidate set [previously identified as associated with urban thermal tolerance plasticity in this species (14)], probes to capture this gene did not pass filtering and were excluded from the final set.

Library preparation was performed by RAPiD Genomics for illumina sequencing utilizing their high-throughput workflow with proprietary chemistry. Briefly, DNA is sheared to a mean fragment length of 400 bp, fragments are end-repaired, followed by incorporation of unique dual-indexed illumina adapters and PCR enrichment. RAPiD Genomics probe set "RG_10801" was developed and synthesized based on the targets provided. These probes were hybridized to the libraries and enriched for the specified targets. Samples were sequenced using HiSeq 2×150 and sequenced approximately 2.7 million read pairs per sample.

We removed sequencing and sample barcode adapters as well as trimmed and filtered reads based on quality scores using Illumiprocessor (84) (v2.09), a wrapper for the read filtering program Trimmomatic (85) (v0.32). We created a nonredundant exome for *Anolis carolinensis* by removing duplicated exons from the *Anolis carolinensis* v2.1 exome (81) using CD-HIT-EST (86) (v4.7). We aligned our quality filtered reads to this nonredundant set of *A. carolinensis* exons using BWA (87) (v0.7.17-r1188). We called and filtered variants using GATK (88) following the GATK best practices (89, 90) with the exception of Base Quality Score Recalibration, which was not possible as there does not exist a reference variant set for *Anolis cristatellus*. We first marked duplicates and called haplotypes for each sample individually, then merged gVCFs for each regional population (Arecibo, Mayagüez, and San Juan), and jointly called genotypes using all individuals (both forest and urban individuals) from each region separately. We retained all sites and used a standard minimum confidence threshold for calling of 20. After genotyping, we merged resulting VCFs from each population for filtering. We first filtered SNPs using GATK VariantFiltration based on examination of empirical distributions extracted using the GATK VariantsToTable function. We used the following filtering expression "QUAL < 0.00 || MQ < 40.00 || SOR > 10.00 || QD < 2.000 || FS > 60.000 || MQRankSum < −12.50 || ReadPosRankSum < −8.00 || ReadPosRankSum > 8.00" and then jointly filtered both variant and invariant sites to remove sites with read depths less than 5 and greater than 60. We further filtered variants using VCFtools (91) (v0.1.15) for a minimum quality of 25 and for a maximum of 25% missing samples per site. The resulting filtered All Sites set contained a total of 7,736,725 called and aligned sites from 36,838 exons. Of these, 354,106 sites were variable (SNPs) drawn from 35,696 exons. We converted the resulting vcf file to the appropriate format for each analysis as follows. We annotated the sequence file using snpEff (92) and the ASU_Acar_v2.1 annotation (81). To convert from vcf to bed and ped formats, we used PLINK (93) and VCFtools (91). To convert from vcf to genepop formats, we used STACKS (94). To subset the vcf file by municipality (San Juan, Arecibo, and Mayagüez), we used bcftools (95).

Population Structure and Genetic Diversity. We investigated population structure, genetic diversity, and inferred phylogenetic relationships based on a total of 105,706 SNPs filtered using bcftools (95) to contain no missing sites, a minimum minor allele frequency of 0.01, and to remove sites with linkage greater than $r^2 = 0.2$ within 10-kb windows (retaining the site in an LD pair with the greater allele frequency). We investigated population structure via identity-by-state (IBS distance) and DAPC. We calculated IBS among all individuals across all six sites using PLINK (93) [DST: (IBS2 + 0.5*IBS1) / (N SNP pairs)] and tested whether IBS differed across sites with ANOVA (high relatedness across sites and municipalities might suggest dispersal events; *SI Appendix*, Fig. S2). We implemented DAPC with the R package "adegenet" (96,97) implemented with the function *dapc* (see *SI Appendix*, Fig. S1 for PCA results). Although *k*-means clustering implemented with the function *find.clusters* in the R package "adegenet" (96, 97) supports the existence of three distinct genetic clusters (equivalent to the municipality for each urban–forest pair), we used group identity based on our sampling (urban or forest from each of the three municipalities) and $k = 6$. We cross-validated the number of retained principal component axes in the DAPC with the function *xvalDapc*, which supported retaining 10 principal component axes. Discriminant functions 1 and 2 in the DAPC (Fig. 1*C* and *SI Appendix*, Fig. S1) show clear separation of genetic variation between geographic regions.

Additional methods similarly validate the existence of three independent urban–forest population pairs. The sample tree indicates that, on average, individuals from within each geographic region (but not necessarily each habitat type within a region) were more genetically similar to one another than to individuals from other geographic regions (Fig. 1*D*). Sequence alignment was performed with the "SNPhylo" pipeline (98) followed by tree model fitting and optimization with IQTree (99) with ModelFinder (100) and ascertainment bias for SNP data (-m TEST+ASC). The midpoint-rooted sample tree was visualized in R with "phytools" (101) and "phangorn" (102). We also estimated admixture coefficients using sparse Nonnegative Matrix Factorization algorithms with the function *snmf* in the R package "LEA" (103); three genetic clusters were most strongly supported (*SI Appendix*, Fig. S1).

In addition, we calculated traditional metrics of population divergence and relatedness. We calculated nucleotide diversity for each of the six sample sites as well as F_{ST} and D_{XY} between all pairs of sites using the Python scripts parseVCF.py and popgenWindows.py (https://github.com/simonhmartin/genomics_general). We excluded indels and included both variant and invariant sites in our analysis. We calculated summary statistics in 10-kb windows excluding any windows with fewer than 100 called sites (*SI Appendix*, Fig. S3). We used VCFtools (91) to calculate observed heterozygosity (--het), relatedness (--relatedness), Tajima's D (--TajimaD 100000), and unadjusted AJK statistic (*SI Appendix*, Figs. S2 and S3).

Signatures of Selection: Urbanization. We conducted selection scan analyses on 115,976 SNPs filtered to remove SNPs with more than 0.25 missing sites and a minor allele frequency threshold of 0.01. We employed two complementary methods to identify loci associated with urbanization: a differentiation outlier method (PCA) and a genetic–environment association method (GEA). PCA approaches are agnostic to environmental variables and, similar to F_{ST} outlier approaches, detect regions of unusually high differentiation among individuals while also (unlike F_{ST} methods) taking into account population structure without needing to specify group identity in advance (104, 105). PCA approaches are less powerful at detecting adaptive divergence when environmental differentiation is weakly correlated with population structure (106, 107). Environmental association methods, in contrast, tend to detect more loci of small effect (spanning a range of F_{ST} values, e.g., ref. 108) by identifying genomic associations with a specified environmental variable.

First, we performed a genome scan for selection in R using principal component analysis implemented in the "pcadapt" (109) package. We analyzed each geographic region separately (San Juan, Mayagüez, and Arecibo) to isolate genetic divergence between urban and forest pairs within each region (an analysis of all three paired populations in a single PCA identifies genomic variation primarily associated with geographic region and not habitat; *SI Appendix*, Fig. S1). We retained 6 to 8 principal components in each PCA based on the proportion variance captured in each PC for each population. We identified outlier SNPs ($\alpha = 0.001$) in each accounting for a false-discovery rate of 1% by calculating q-values with the *qvalue* function in R package "qvalue" (110). We found the intersection of outlier SNPs between all three geographic regions ("PCA outlier SNPs"). We repeated this intersection at the gene level based on the aligned *A. carolinensis* Ensembl gene ID, with outlier genes identified as containing at least one outlier SNP ("PCA outlier genes").

Second, we conducted a genotype association test with urbanization using a logistic linear mixed effects model (binomial family) implemented with the functions *fitNullModel* and *assocTestSingle* in the R package "GENESIS" (111). Our model used a 50-kb sliding window with a 10-kb slide. To account for population structure and regional variation, we incorporated a genetic relatedness matrix (estimated with functions *pcair* and *pcrelate* in "GENESIS") and municipality as covariates. We identified outlier SNPs as the smallest 1% of the distribution of *P*-values for the association test ("habitat outlier SNPs") and identified outlier genes as those containing at least one outlier SNP ("habitat outlier genes").

Signatures of Selection: Morphology. We examined genotype-trait associations for six composite traits (measurement described above): hindlimb length, forelimb length, toepad area (front and rear), and toepad lamella number (front and rear). We conducted a genotype association test with urbanization implemented with the functions *fitNullModel* and *assocTestSingle* in the R package "GENESIS" (111) for each of the six traits. Our model used a 50-kb sliding window with a 10-kb slide. To account for population structure and regional variation, we incorporated a genetic relatedness matrix (estimated with functions *pcair* and *pcrelate* in "GENESIS") and municipality as covariates. We identified SNPs as the smallest 1% of the distribution of p-values for each association test and identified outlier genes as those containing at least one outlier SNP in each analysis.

Common Signatures of Selection across Analyses. We identified a core set of urban genes by finding the intersection of genes containing at least one outlier SNP in the PCA analysis (across all three municipalities) and the GEA analysis ("urban-associated genes," N = 33). By using a combination of outlier detection approaches, we identify a conservative set of loci under selection in urban environments, although focusing on the overlap between approaches is likely to miss loci under weak selection (104,107). We also identified a set of urban morphological genes by finding the intersection of outlier SNPs in each GWAS analysis with the GEA analysis. We identified the subset of these loci for each morphological trait that was shared between anterior and posterior elements (e.g., forelimb and hindlimb) as candidate urban morphology genes. The genetic architecture of early limb development is conserved between hindlimbs and forelimbs in vertebrate taxa, although there are clear limb-specific programs that activate later in development to establish different morphologies between the two (112–115).

We tested the significance of the overlap in SNPs between the GEA and PCA with permutation and 1,000 iterations, randomly sampling SNPs without replacement four times for the GEA/PCA (once for each municipality and once for the GEA) and finding the intersection of SNPs across the four sets. Similarly, we tested the significance of the overlap in SNPs between each of the six GWAS and the GEA with permutation and 1,000 iterations, randomly sampling without replacement twice for each test (once for the trait, once for the GEA). We also tested the significance of the overlap in genes between the GEA and PCA analyses, again with a permutation test with 1,000 iterations, where we first randomly sampled SNPs and then calculated the intersection of genes in which those SNPs are found.

Functional Associations. We used the function *gost* in the R package "gprofileR2" (116) to perform a gene list functional enrichment on two sets of genes: urban-associated genes (n = 23 genes) and morphology-associated genes (any gene containing an outlier SNP identified by one of the six morphology association tests; n = 1,776 genes). We provided as a custom background gene set the full list of genes containing at least one SNP in our dataset (n = 6,389 genes).

Evaluating Repeatability. We investigated parallel genomic divergence between urban and forest populations with two approaches. First, we examined polygenic divergence associated with urbanization by performing a local principal components analysis on outlier genomic regions. PCAs were implemented with the function *snpgdsPCA* in the R package "SNPRelate" (117) on each of the seven sets of outlier SNPs (urban GEA, urban-morphology). This analysis can provide insight into whether haplotypes are similarly diverging across urban–forest pairs (16, 118, 119). We then used a linear model to determine the effect of habitat (urban or forest), municipality (Arecibo, Mayagüez, and San Juan), and their interaction on the primary axes of genomic variation in the outlier sets (i.e., PC1 and PC2). A significant habitat effect would indicate that divergence associated with the urban environment or the trait (depending on the model) is associated with urbanization, whereas a significant municipality effect indicates regional variation driving divergence associated with the trait (e.g., as in ref. 16).

Second, we investigated parallel divergence at the allele level by examining effect sizes (eta^2) of allele frequencies for all SNPs in our dataset (120). We used the *etasquared* function in the R package "rstatix." We then compared the effect size of the habitat effect versus the interaction effect of habitat x municipality, where a stronger interaction effect suggests greater variation by region and the converse supporting parallelism. We compared effect sizes for all outlier SNPs identified in our two urbanization analyses (GEA, intersection of all three PCA) as well as the outlier SNPs identified by the intersection of the urbanization GEA and each morphology test (urban morphology SNPs). We compared effect sizes to the effect sizes of the background set of SNPs (SNPs not identified as outliers in any test).

Data, Materials, and Software Availability. Raw sequence data (121) have been deposited in NCBI Sequence Read Archive (BioProject PRJNA872192). All datafiles and code needed to repeat analyses (122) have been deposited in Zenodo (PRJNA872192 DOI:10.5281/Zenodo.6636371).

ACKNOWLEDGEMENTS. This study was conducted under Permits #2012-IC-049, #2013-IC-033, #2014-IC-024, from the Puerto Rico Departamento de Recursos Naturales y Ambientales (DRNA), and in compliance with University of Massachusetts Institutional Animal Care and Use Committee (IACUC) protocol #2012001. This research was funded in part by grants from the NSF to L.J.R. (DEB 1354044), J.B.L. and A.J.G. (DEB 1927194), and L.J.R. and K.M.W. (DEB 1701706) and by the University of Massachusetts Boston Bollinger Memorial Research Grant. We thank the following people for helpful discussions on the manuscript development and methodology: L. Abueg, K. Aviles-Rodriguez, E. Carlen, J. Munshi-South, N. Rochette, K. Schliep, J. Zschau.

Author affiliations: [a]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544; [b]Department of Biology, Washington University, St. Louis, MO 63130; [c]Department of Biology, University of Massachusetts Boston, Boston, MA 02125; [d]Facultad de Ciencias, Universidad Católica de la Santísima Concepción, Concepción, Chile 4090541; [e]Center for Biological Data Science, Virginia Commonwealth University, Richmond, VA 23284; and [f]Department of Biology and Center for Computational and Integrative Biology, Rutgers University–Camden, NJ 08103

1. M. T. J. Johnson, J. Munshi-South, Evolution of life in urban environments. *Science* **358**, 6363 (2017).
2. J. S. Santangelo, "Urban environments as a framework to study parallel evolution" in *Urban Evolutionary Biology*, M. Szulkin, J. Munshi-South, A. Charmantier, Eds. (Oxford University Press, USA, 2020), pp. 36–53.
3. M. Szulkin, J. Munshi-South, A. Charmantier, Eds., *Urban Evolutionary Biology* (Oxford University Press, USA, 2020).
4. K. M. Winchell, R. G. Reynolds, S. R. Prado-Irwin, A. R. Puente-Rolón, L. J. Revell, Phenotypic shifts in urban areas in the tropical lizard *Anolis cristatellus*. *Evolution* **70**, 1009–1022 (2016).
5. A. P. Hendry, T. J. Farrugia, M. T. Kinnison, Human influences on rates of phenotypic change in wild animal populations. *Mol. Ecol.* **17**, 20–29 (2008).
6. M. Alberti, Global urban signatures of phenotypic change in animal and plant populations. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 8951–8956 (2017).
7. E. Shochat, P. S. Warren, S. H. Faeth, N. E. McIntyre, D. Hope, From patterns to emerging processes in mechanistic urban ecology. *Trends Ecol. Evol.* **21**, 186–191 (2006).
8. S. Des Roches *et al.*, Socio-eco-evolutionary dynamics in cities. *Evol. Appl.* **14**, 248–267 (2021).
9. L. R. Rivkin *et al.*, A roadmap for urban evolutionary ecology. *Evol. Appl.* **12**, 384–398 (2019).

10. L. S. Miles, E. J. Carlen, K. M. Winchell, M. T. Johnson, Urban evolution comes into its own: Emerging themes and future directions of a burgeoning field. *Evol. Appl.* **14**, 3–11 (2021).
11. B. C. Verrelli *et al.*, A global horizon scan for urban evolutionary ecology. *Trends Ecol. Evol.* **37**, 1006–1019 (2022).
12. C. Perrier, A. Caizergues, A. Charmantier, "Adaptation genomics in urban environments" in *Urban Evolutionary Biology*, M. Szulkin, J. Munshi-South, A. Charmantier, Eds. (Oxford University Press, USA, 2020), pp. 74–90.
13. M. Ravinet *et al.*, Signatures of human-commensalism in the house sparrow genome. *Proc. R. Soc. B.* **285**, 20181246 (2018).
14. S. C. Campbell-Staton *et al.*, Parallel selection on thermal physiology facilitates repeated adaptation of city lizards to urban heat islands. *Nat. Ecol. Evol.* **4**, 652–658 (2020).
15. J. C. Mueller *et al.*, Genes acting in synapses and neuron projections are early targets of selection during urban colonization. *Mol. Ecol.* **29**, 3403–3412 (2020).
16. P. Salmón *et al.*, Continent-wide genomic signatures of adaptation to urbanisation in a songbird across Europe. *Nat. Commun.* **12**, 1–14 (2021).
17. A. Harpak *et al.*, Genetic adaptation in New York city rats. *Genome Biol. Evol.* **13**, evaa247 (2021).
18. L. S. Miles, L. R. Rivkin, M. T. Johnson, J. Munshi-South, B. C. Verrelli, Gene flow and genetic drift in urban environments. *Mol. Ecol.* **28**, 4138–4151 (2019).
19. J. C. Müller, J. Partecke, B. J. Hatchwell, K. J. Gaston, K. L. Evans, Candidate gene polymorphisms for behavioural adaptations during urbanization in blackbirds. *Mol. Ecol.* **22**, 3629–3637 (2013).
20. S. E. Harris, J. Munshi-South, C. Obergfell, R. O'Neill, Signatures of rapid evolution in urban and rural transcriptomes of white-footed mice (*Peromyscus leucopus*) in the New York metropolitan area. *PLoS One* **8**, e74938 (2013).
21. J. S. Santangelo *et al.*, Global urban environmental change drives adaptation in white clover. *Science* **375**, 1275–1281 (2022).
22. J. Losos, "Lizards in an evolutionary tree" in *Lizards in an Evolutionary Tree* (University of California Press, 2009).
23. W. A. Gould, S. Martinuzzi, O. M. Ramos-González, Developed land cover of Puerto Rico. Scale 1: 260 000, IITF-RMAP-10 (US Department of Agriculture Forest Service, International Institute of Tropical Forestry, Río Piedras, PR, 2008).
24. C. Isaakson, F. Bonier, "Urban evolutionary physiology" in *Urban Evolutionary Biology*, M. Szulkin, J. Munshi-South, A. Charmantier, Eds. (Oxford University Press, USA, 2020), pp. 217–233.
25. C. A. Bradley, S. Altizer, Urbanization and the ecology of wildlife diseases. *Trends Ecol. Evol.* **22**, 95–102 (2007).
26. D. Sol, O. Lapiedra, S. Ducatez, "Cognition and adaptation to urban environments" in *Urban Evolutionary Biology*, M. Szulkin, J. Munshi-South, A. Charmantier, Eds. (Oxford University Press, USA, 2020), pp. 253–267.
27. K. M. Winchell, A. C. Battles, T. Y. Moore, "Terrestrial locomotor evolution in urban environments" in *Urban Evolutionary Biology*, M. Szulkin, J. Munshi-South, A. Charmantier, Eds. (Oxford University Press, USA, 2020), pp. 197–216.
28. S. E. Diamond, R. A. Martin, "Evolutionary consequences of the urban heat island" in *Urban Evolutionary Biology*, M. Szulkin, J. Munshi-South, A. Charmantier, Eds. (Oxford University Press, USA, 2020), pp. 91–110.
29. S. S. French, H. B. Fokidis, M. C. Moore, Variation in stress and innate immunity in the tree lizard (*Urosaurus ornatus*) across an urban–rural gradient. *J. Comp. Physiol. B.* **178**, 997–1005 (2008).
30. R. K. Tyler, K. M. Winchell, L. J. Revell, Tails of the city: Caudal autotomy in the tropical lizard, *Anolis cristatellus*, in urban and natural areas of Puerto Rico. *J. Herpetol.* **50**, 435–441 (2016).
31. K. M. Winchell, D. Briggs, L. J. Revell, The perils of city life: Patterns of injury and fluctuating asymmetry in urban lizards. *Biol. J. Linnean Soc.* **126**, 276–288 (2019).
32. C. J. Thawley *et al.*, Urbanization affects body size and parasitism but not thermal preferences in *Anolis* lizards. *J. Urban Ecol.* **5**, juy031 (2019).
33. B. Shafit-Zagardo, N. Kalcheva, Making sense of the multiple MAP-2 transcripts and their role in the neuron. *Mol. Neurobiol.* **16**, 149–162 (1998).
34. E. Yeh *et al.*, A putative cation channel, NCA-1, and a novel protein, UNC-80, transmit neuronal activity in *C. elegans*. *PLoS Biol.* **6**, e55 (2008).
35. D. J. Tischfield *et al.*, Loss of the neurodevelopmental gene Zswim6 alters striatal morphology and motor regulation. *Neurobiol. Dis.* **103**, 174–183 (2017).
36. Y. Song *et al.*, Knockdown of Pnpla6 protein results in motor neuron defects in zebrafish. *Dis. Models Mech.* **6**, 404–413 (2013).
37. G. W. Go, A. Mani, Low-density lipoprotein receptor (LDLR) family orchestrates cholesterol homeostasis. *Yale J. Boil. Med.* **85**, 19 (2012).
38. M. Boutchueng-Djidjou *et al.*, The last enzyme of the de novo purine synthesis pathway 5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase (ATIC) plays a central role in insulin signaling and the golgi/endosomes protein network*[S]. *Mol. Cell. Proteom.* **14**, 1079–1092 (2015).
39. S. Matsumoto *et al.*, Urea cycle disorders–update. *J. Hum. Genet.* **64**, 833–847 (2019).
40. M. Akiyama, The roles of ABCA12 in epidermal lipid barrier formation and keratinocyte differentiation. *Biochim. Biophys. Acta* **1841**, 435–440 (2014).
41. X. S. Wu, J. A. Martina, J. A. Hammer III, Melanoregulin is stably targeted to the melanosome membrane by palmitoylation. *Biochem. Biophys. Res. Commun.* **426**, 209–214 (2012).
42. T. J. Sanger, L. J. Revell, J. J. Gibson-Brown, J. B. Losos, Repeated modification of early limb morphogenesis programmes underlies the convergence of relative limb length in *Anolis* lizards. *Proc. R. Soc. B. Biol. Sci.* **279**, 739–748 (2012).
43. S. Park, C. R. Infante, L. C. Rivera-Davila, D. B. Menke, Conserved regulation of hoxc11 by pitx1 in *Anolis* lizards. *J. Exp. Zool. Part B. Mol. Dev. Evol.* **322**, 156–165 (2014).
44. K. M. Winchell, I. Maayan, J. R. Fredette, L. J. Revell, Linking locomotor performance to morphological shifts in urban lizards. *Proc. R. Soc. B.* **285**, 20180229 (2018).
45. J. J. Kolbe, A. C. Battles, K. J. Avilés-Rodríguez, City slickers: Poor performance does not deter *Anolis* lizards from using artificial substrates in human-modified habitats. *Funct. Ecol.* **30**, 1418–1429 (2016).
46. S. Meng *et al.*, TBX20 regulates angiogenesis through the prokineticin 2–prokineticin receptor 1 pathway. *Circulation* **138**, 913–928 (2018).
47. C. R. Ingraham *et al.*, Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (Irf6). *Nature Genetics* **38**, 1335–1340 (2006).
48. S. Eppley, R. J. Hopkin, B. Mendelsohn, A. M. Slavotinek, Clinical report: Warsaw breakage syndrome with small radii and fibulae. *Am. J. Med. Genet. Part A* **173**, 3075–3081 (2017).
49. M. Seifert, W. Tilgen, J. Reichrath, Expression of 25-hydroxyvitamin D-1α-hydroxylase (1αOHase, CYP27B1) splice variants in HaCaT keratinocytes and other skin cells: modulation by culture conditions and UV-B treatment in vitro. *Anticancer Research* **29**, 3659–3667 (2009).
50. M. Chiquet, D. E. Birk, C. G. Bönnemann, M. Koch, Collagen XII: Protecting bone and muscle integrity by organizing collagen fibrils. *Int. J. Biochem. Cell Biol.* **53**, 51–54 (2014).
51. L. Alibardi, Differentiation of the epidermis during scale formation in embryos of lizard. *J. Anatomy* **192**, 173–186 (1998).
52. D. Halim *et al.*, Loss of LMOD1 impairs smooth muscle cytocontractility and causes megacystis microcolon intestinal hypoperistalsis syndrome in humans and mice. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E2739–E2747 (2017).
53. L. McGregor, *Locomotor morphology of Anolis: Comparataive investigations of the design and function of the subdigital adhesive system* (University of Calgary, 1999).
54. M. Safran *et al.*, "The GeneCards suite chapter" in *Practical Guide to Life Science Databases* (Springer, 2022), pp. 27–56.
55. C. R. Paradise *et al.*, Brd4 is required for chondrocyte differentiation and endochondral ossification. *Bone* **154**, 116234 (2022).
56. R. P. Naja, O. Dardenne, A. Arabian, R. St, Arnaud, chondrocyte-specific modulation of Cyp27b1 expression supports a role for local synthesis of 1, 25-dihydroxyvitamin D3 in growth plate development. *Endocrinology* **150**, 4024–4032 (2009).
57. J. Lu *et al.*, Mutations cause chondrocyte defects in skeletal development. *Human Molecular Genetics* **16**, 1661–1675 (2007).
58. A. Bentmann *et al.*, Circulating fibronectin affects bone matrix, whereas osteoblast fibronectin modulates osteoblast function. *Journal of Bone and Mineral Research* **25**, 706–715 (2010).
59. B. J. Cosentino, J. P. Gibbs, Parallel evolution of urban–rural clines in melanism in a widespread mammal. *Sci. Rep.* **12**, 1–7 (2022).
60. K. R. Elmer, A. Meyer, Adaptation in the age of ecological genomics: Insights from parallelism and convergence. *Trends Ecol. Evol.* **26**, 298–306 (2011).
61. K. M. Lee, G. Coop, Population genomics perspectives on convergent adaptation. *Phil. Trans. R. Soc.* **374**, 20180236 (2019).
62. R. B. Corbett-Detig, S. L. Russell, R. Nielsen, J. Losos, Phenotypic convergence is not mirrored at the protein level in a lizard adaptive radiation. *Mol. Biol. Evol.* **37**, 1604–1614 (2020).
63. H. E. Hoekstra, J. A. Coyne, The locus of evolution: Evo devo and the genetics of adaptation. *Evolution* **61**, 995 (2007).
64. B. J. Cassone *et al.*, Gene expression divergence between malaria vector sibling species *Anopheles gambiae* and *An. coluzzii* from rural and urban Yaoundé Cameroon. *Mol. Ecol.* **23**, 2242–2259 (2014).
65. W. F. van Dongen, R. W. Robinson, M. A. Weston, R. A. Mulder, P. J. Guay, Variation at the DRD4 locus is associated with wariness and local site selection in urban black swans. *BMC Evol. Biol.* **15**, 1–11 (2015).
66. United Nations Center for Human Settlement (HABITAT). Cities in a globalizing world. Global report on human settlements 2001 (Earthscan Publications Ltd, UK and USA, 2001).
67. S. Subramanian, S. Kumar, Evolutionary anatomies of positions and types of disease-associated and neutral amino acid mutations in the human genome. *BMC Genomics* **7**, 306 (2006).
68. D. Gurdasani *et al.*, Uganda genome resource enables insights into population history and genomic discovery in Africa. *Cell* **179**, 984–1002.e36 (2019).
69. A. M. Hancock *et al.*, Colloquium paper: Human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 8924–8930 (2010).
70. D. A. Stover, G. Housman, A. C. Stone, M. S. Rosenberg, B. C. Verrelli, Evolutionary genetic signatures of selection on bone-related variation within human and chimpanzee populations. *Genes* **13**, 183 (2022).
71. R. G. Reynolds *et al.*, Archipelagic genetics in a widespread Caribbean anole. *J. Biogeography* **44**, 2631–2647 (2017).
72. W. S. Rasband, U. S. ImageJ. https://imagej.nih.gov/ij/ (National Institutes of Health Bethesda, Maryland, USA, 1997–2018).
73. M. E. Wolak, D. J. Fairbairn, Y. R. Paulsen, Guidelines for estimating repeatability. *Methods Ecol. Evol.* **3**, 129–137 (2012).
74. K. Alboukadel, Rstatix: Pipe-friendly framework for basic statistical tests. R package version 0.7.0. https://CRAN.R-project.org/package=rstatix (The Comprehensive R Archive Network, 2021).
75. J. Eales, R. S. Thorpe, A. Malhotra, Weak founder event signal in a recent introduction of Caribbean *Anolis*. *Mol. Ecol.* **17**, 1416–1426 (2008).
76. S. E. Fick, R. J. Hijmans, WorldClim 2: New 1km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
77. C. G. Homer, C. Huang, L. Yang, B. K. Wylie, M. J. Coan, Development of a 2001 National land cover database for the United States. *Photogramm. Eng. Remote Sens.* **70**, 829–840 (2004).
78. J. W. Coulston, Modeling percent tree canopy cover–A pilot study. *Photogramm. Eng. Remote Sens.* **78**, 715–727 (2012).
79. S. R. Prado-Irwin, L. J. Revell, K. M. Winchell, Variation in tail morphology across urban and forest populations of the crested anole (*Anolis cristatellus*). *Biol. J. Linn. Soc.* **128**, 632–644 (2019).
80. US Geological Survey, USGS high resolution orthoimagery for Puerto Rico and the U.S. Virgin Islands. (2012). https://www.sciencebase.gov/catalog/item/59efaa13e4b0220bbd99b87b. Last accessed 27 December 2022.
81. W. L. Eckalbar *et al.*, Genome reannotation of the lizard *Anolis carolinensis* based on 14 adult and embryonic deep transcriptomes. *BMC Genomics* **14**, 49 (2013).
82. J. Alföldi *et al.*, The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature* **477**, 587–591 (2011).
83. A. J. Geneva *et al.*, Chromosome-scale genome assembly of the brown anole (*Anolis sagrei*), a model species for evolution and ecology. *Commun. Biol.* **5**, 1126 (2022).
84. B. C. Faircloth, Illumiprocessor: A trimmomatic wrapper for parallel adapter and quality trimming. (2013) http://dx.doi.org/10.6079/J9ILL. Last accessed 27 December 2022.
85. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
86. L. Fu, B. Niu, Z. Zhu, S. Wu, W. Li, CD-HIT: Accelerated for clustering the next generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
87. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv, 1303.3997. https://doi.org/10.48550/arXiv.1303.3997. Last accessed 27 December 2022.

88. A. McKenna *et al.*, The genome analysis toolkit: A mapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–303 (2010).
89. M. DePristo *et al.*, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
90. G. A. Van der Auwera, B. D. O'Connor, "Genomics in the Cloud: Using docker" in *GATK, and WDL in Terra* (O'Reilly Media, ed. 1, 2020).
91. P. Danecek *et al.*, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
92. P. Cingolani, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92 (2012).
93. S. Purcell *et al.*, PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am. J. Hum. Genet.* **81** (2007).
94. J. Catchen, P. Hohenlohe, S. Bassham, A. Amores, W. Cresko, Stacks: An analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140 (2013).
95. P. Danecek *et al.*, Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
96. T. Jombart, Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405 (2008).
97. T. Jombart, I. Ahmed, Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070–3071 (2011).
98. T. H. Lee, H. Guo, X. Wang, C. Kim, A. H. Paterson, SNPhylo: A pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genomics* **15**, 162 (2014).
99. B. Q. Minh *et al.*, IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
100. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin, ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
101. L. J. Revell, phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).
102. K. P. Schliep, phangorn: Phylogenetic analysis in R. *Bioinformatics* **27**, 592–593 (2011).
103. E. Frichot, O. François, LEA: An R package for landscape and ecological association studies. *Methods Ecol. Evol.* **6**, 925–929 (2015).
104. S. Hoban *et al.*, Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *Am. Nat.* **188**, 379–397 (2016).
105. K. Luu, E. Bazin, M. G. Blum, Pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Mol. Ecol. Resources* **17**, 67–77 (2017).
106. T. Capblancq, K. Luu, M. G. Blum, E. Bazin, Evaluation of redundancy analysis to identify signatures of local adaptation. *Mol. Ecol. Resources* **18**, 1223–1233 (2018).
107. K. E. Lotterhos, M. C. Whitlock, The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Mol. Ecol.* **24**, 1031–1046 (2015).
108. C. J. Brauer, M. P. Hammer, L. B. Beheregaray, Riverscape genomics of a threatened fish across a hydroclimatically heterogeneous river basin. *Mol. Ecol.* **25**, 5093–5113 (2016).
109. F. Privé *et al.*, Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. *Mol. Biol. Evol.* **37**, 2153–2154 (2020).
110. J. D. Storey, A. J. Bass, A. Dabney, D. Robinson, Qvalue: Q-value estimation for false discovery rate control. R package version 2.22.0 (The Comprehensive R Archive Network, 2020).
111. S. M. Gogarten *et al.*, Genetic association testing using the GENESIS R/Bioconductor package. *Bioinformatics* **35**, 5346–5348 (2019).
112. M. Logan, SAGE profiling of the forelimb and hindlimb. *Genome Biol.* **3**, 1–3 (2002).
113. J. Melville *et al.*, Expression of a hindlimb-determining factor Pitx1 in the forelimb of the lizard *Pogona vitticeps* during morphogenesis. *Open Biol.* **6**, 160252 (2016).
114. C. Minguillon, J. Del Buono, M. P. Logan, Tbx5 and Tbx4 are not sufficient to determine limb-specific morphologies but have common roles in initiating limb outgrowth. *Dev. Cell* **8**, 75–84 (2005).
115. N. Shubin, C. Tabin, S. Carroll, Fossils, genes and the evolution of animal limbs. *Nature* **388**, 639–648 (1997).
116. L. Kolberg, U. Raudvere, Gprofiler 2: Interface to the "g:Profiler" toolset. R package version 0.2.0 (The Comprehensive R Archive Network, 2020).
117. Z. Xiuwen *et al.*, A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328 (2012).
118. H. Li, P. Ralph, Local PCA shows how the effect of population structure differs along the genome. *Genetics* **211**, 289–304 (2019).
119. G. Montejo-Kovacevich, *et al.*, Repeated genetic adaptation to high altitude in two tropical butterflies. *Nat. Commun.* **13**, 4676 (2022).
120. D. I. Bolnick, R. D. Barrett, K. B. Oke, D. J. Rennison, Y. E. Stuart, (Non) parallel evolution. *Annu. Rev. Ecol. Evol. Syst.* **49**, 303–330 (2018).
121. Geneva, A.J. & Winchell, K.M. Genome-wide parallelism underlies contemporary adaptation in urban environments. *NCBI*. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA872192. 22 August 2022.
122. Winchell, K.M. Analyses for: Genome-wide parallelism underlies contemporary adaptation in urban lizards. *Zenodo*. https://doi.org/10.5281/zenodo.6636371. 13 December 2022.