

Main Manuscript for

Genome-wide parallelism underlies contemporary adaptation in urban lizards

Kristin M. Winchell^{1-4*}, Shane C. Campbell-Staton², Jonathan B. Losos^{3*}, Liam J. Revell^{4,5}, Brian C. Verrelli⁶, Anthony J. Geneva⁷

¹Department of Biology, New York University, New York, NY 10012 (current affiliation)

²Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544

³Department of Biology, Washington University, St. Louis, MO 63130

⁴Department of Biology, University of Massachusetts Boston, Boston, MA 02125

⁵Facultad de Ciencias, Universidad Católica de la Santísima Concepción, Concepción, Chile

⁶Center for Biological Data Science, Virginia Commonwealth University, Richmond, VA 23284

⁷Department of Biology and Center for Computational and Integrative Biology, Rutgers University–Camden, NJ, 08103

*Correspondence to: Losos@wustl.edu & Kristin.Winchell@NYU.edu

Classification: Biological Sciences, Evolution

Keywords: urban evolution, rapid adaptation, parallelism, urbanization, *Anolis*

Author Contributions: KMW and LJR conceived of the project with input from SCCS and AJG. KMW conducted fieldwork with support from LJR, and KMW conducted molecular work with support from AJG and LJR. AJG performed sequence alignment, filtering, and variant calling. KMW conducted statistical analyses with support from AJG and SCCS. All authors participated in the development and writing of the manuscript.

Additional Information: Supplementary Information is available for this paper. Correspondence and requests for materials should be addressed to K. M. Winchell (Kristin.winchell@NYU.edu).

Data Availability: All raw sequence data have been deposited to the NCBI Sequence Read Archive under BioProject PRJNA872192. All data needed to repeat analyses have been archived on Zenodo with R-markdown scripts and PDFs with doi: 10.5281/zenodo.6636371 and are available as supplementary files during review.

1 ABSTRACT

2 Urbanization drastically transforms landscapes, resulting in fragmentation, degradation,
3 and the loss of local biodiversity¹. Yet urban environments also offer opportunities to observe
4 rapid evolutionary change in wild populations that survive and even thrive in these novel
5 habitats. In many ways, cities represent replicated “natural experiments” in which
6 geographically separated populations adaptively respond to similar selection pressures over
7 rapid evolutionary timescales². Little is known, however, about the genetic basis of adaptive
8 phenotypic differentiation in urban populations nor the extent to which phenotypic parallelism
9 is reflected at the genomic level with signatures of parallel selection³. Here we analyzed the
10 genomic underpinnings of parallel urban-associated phenotypic change in *Anolis cristatellus*, a
11 small-bodied neotropical lizard found abundantly in both urbanized and forested environments.
12 We show that phenotypic parallelism in response to parallel urban environmental change is
13 underlain by genomic parallelism and identify candidate loci across the *Anolis* genome
14 associated with this adaptive morphological divergence. Our findings point to polygenic
15 selection on standing genetic variation as a key process to effectuate rapid morphological
16 adaptation. Identified candidate loci represent several functions associated with
17 skeletomuscular development, morphology, and human disease. Taken together, these results
18 shed light on the genomic basis of complex morphological adaptations, provide insight into the
19 role of contingency and determinism in adaptation to novel environments, and underscore the
20 value of urban environments to address fundamental evolutionary questions.

22 SIGNIFICANCE

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

33 INTRODUCTION

34 It is increasingly evident that humans influence ecological dynamics and evolutionary
35 trajectories of organisms occupying human-dominated spaces²⁻⁷. Abundant anthropogenic
36 materials and structures combined with a deficiency of green spaces create novel biotic and
37 abiotic conditions and complex socio-eco-evolutionary dynamics in cities^{1,4,9}. Studies have
38 found wide-ranging functional, phenotypic, regulatory, and genomic consequences across
39 diverse taxa², yet our understanding of evolutionary mechanisms in urban environments is
40 nascent¹⁰⁻¹¹. Among the central outstanding questions is to what extent phenotypic adaptations
41 and parallelism are reflected at the genomic level^{3,10,12}. Consequently, the relative importance
42 of contingency versus determinism in contemporary adaptation to urban environments remains
43 underexplored.

While it is clear that urbanization is associated with phenotypic and genomic changes^{2,4}, we still know very little about the genetic targets of selection underpinning adaptive urban trait shifts¹³. Previous studies have identified genome-wide patterns of genomic differentiation associated with urban environments¹⁴⁻¹⁸. However, many urban genomics studies have focused primarily on non-adaptive 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^{13,19}. 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.^{16,18,20-22}). 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 nonurban environments, exhibiting repeated morphological and physiological adaptations^{5,15,23}. 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-B; n=96 individuals, n=16 per population). We employed a custom semi-targeted 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. 1C) and not by habitat (urban versus forest), in line with evidence from mtDNA⁵ and RNAseq data¹⁵. 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. 1C-D, Supplemental Figs. S1-S3). Analyses of relatedness (Fig. 1D, Supplemental 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: $F_{df=1, 718}=78.6$, $p<0.001$; Supplemental 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. 2A; Supplemental Fig. S4). Additionally, we used a principal components-based genetic outlier analysis for each urban-forest pair (Fig. 2B-D) to detect genomic regions of unusually high differentiation among individuals while accounting for population structure among sites. Genes

87 containing at least one single nucleotide polymorphism (SNP) identified as outliers across all
88 analyses were deemed the strongest candidates for selection in urban environments.

89 In total, we identified 91 total variants (out of 100 outlier SNPs identified by the PCA
90 analysis and 1153 outlier SNPs identified by the GEA analysis) across three regions of the *Anolis*
91 genome that met these criteria: a 4.4 Mbp region of CHR1, a 4.4 Mbp region of CHR2, and a
92 34.5 kbp region of CHR4 (Fig. 2A-E, Supplemental Fig. S5). Of these 91 variants, 33% were in
93 focal exons (see Methods). The observed overlap between the GEA and PCA outlier SNPs is
94 significantly greater than expected by chance (permutation test, $p < 0.001$, Supplemental Fig. S6)
95 and represents 0.078% of the total number of SNPs tested (115,976). These patterns are
96 consistent with selection in urban environments repeatedly targeting three specific genomic
97 regions containing 33 genes across all three municipalities (Fig. 2F). The observed overlap
98 between the GEA and PCA at the gene level is significantly greater than expected by chance
99 (permutation test, $p = 0.006$, Supplemental Fig. S6).

100 In addition, several genomic regions display idiosyncratic genomic differentiation within
101 a single municipality (Fig. 2B-D) which may reflect regional environmental differences,
102 divergent selection pressures among cities, or geographic variation in initial standing genetic
103 variation of the urban populations. In fact, only 3% of the outlier SNPs identified in the PCA
104 analyses were outliers in all three municipalities and 19% were outliers in 2 out of the 3
105 municipalities (Supplemental Fig. S7). We found more overlapping candidate SNPs between
106 Arecibo and San Juan than between the other regions, a result that may be explained by the
107 environmental similarity of these two municipalities. Specifically, San Juan and Arecibo are
108 cooler, receive less precipitation, and have more large trees contributing to more extensive
109 canopy cover compared to Mayagüez (Fig. 1B). Taken together, these analyses highlight the
110 complex interaction of determinism and contingency in shaping adaptive genomic divergence in
111 urban environments.

112 The wide variety of environmental changes associated with urbanization present myriad
113 opportunities for phenotypic adaptations that enable colonization and persistence in these
114 novel habitats. For example: abundant anthropogenic food resources could affect dietary and
115 metabolic processes²⁴; altered disease dynamics might impact immune function²⁵; novel
116 resources could challenge cognitive abilities and behaviors²⁶; altered structural environments
117 may influence locomotor morphology²⁷; and urban heat islands could challenge thermal and
118 desiccation tolerances^{15,28}.

119 We explored potential functions associated with the 33 urban-associated genes (Fig. 2F,
120 Supplemental Table S8). We found a single significantly enriched term from the KEGG database
121 (Glutathione metabolism), and a gene ontology analysis highlights immune functions, wound
122 healing, and inflammatory responses, which may indicate habitat-specific differences that
123 necessitate selection on injury recovery and immunocompetence. Urban wildlife often face
124 disease dynamics and stressors different from those experienced by their non-urban
125 counterparts, resulting in selection on immune function and stress response^{25,29}. Indeed,
126 previous research has established that urban anoles exhibit elevated injury rates, including
127 bone fractures, missing digits, and autotomized tails³⁰⁻³¹ as well as increased parasite
128 infections³². In addition, several of these genes have been implicated in neural function and
129 motor regulation (e.g., MAP2³³, UNC80³⁴, ZSWIM4³⁵, PNPLA6³⁶), metabolic function (e.g.,
130 LDLR³⁷, ATIC³⁸, CPS1³⁹), skin development (ABCA12⁴⁰), and epithelial pigmentation (MREG⁴¹).

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 genus²³. 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 (Supplemental 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. 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 fore- and hind- limb lengths, toepad area, and lamella number. These data show that urban populations exhibit parallel increases in all six traits (Fig. 3F, Supplemental Fig. S11), consistent with previous studies. These phenotypic differences have been shown to translate into differences in locomotor performance⁴⁴⁻⁴⁵, 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, $p<0.001$; Supplemental Fig. S6). These loci represent the strongest candidates underpinning adaptive urban phenotypes (Fig. 3G-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 fore- and hind-elements of each trait (Supplemental Table S9, intersection at the SNP level in Fig. 3G-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 (PROKR1⁴⁶) and are implicated in diseases in humans and mice involving shortened and deformed limbs (IRF6⁴⁷ and DDX11⁴⁸). Genes associated with urban-toepad morphology are involved in the development of keratin, collagen, and skin (e.g., ABCA12⁴⁰, CYP27B1⁴⁹, COL12A1⁵⁰), major components of the anole epidermis and scales⁵¹, as well as smooth muscle contraction (LMOD1⁵²), which is involved in toepad conformation and release from surfaces⁵³. In addition, 82 of the candidate genes have Gene Card⁵⁴ 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: BRD4⁵⁵, CYP27B1⁵⁶, FLNB⁵⁷, FN1⁵⁸, and IRF6⁴⁷. 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^{17,22,59}. 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 common³. 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. 4A-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 Supplemental 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. 4D-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 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 (non-outlier) set of SNPs, finding each trait had a significantly greater effect of habitat compared to neutral genetic variation (Fig. 4G; Supplemental 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: $t=5.51$, $df=41$, $p=3.2e^{-9}$; 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.

Firstly, 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 urbanization³ and rarely connects parallelism at environmental, phenotypic, and genomic levels (with a few recent exceptions^{15,17,22}). Theory predicts that adaptive evolution in closely related populations is more likely to arise via parallel genomic change⁶⁰⁻⁶¹. 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.

Secondly, 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 both⁶³. 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 phenotype⁶³. 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 variation⁶³. Evidence suggests that changes in gene expression underlie some urban adaptations, such as thermal tolerance¹⁵ and insecticide resistance⁶⁴, whereas changes to coding regions underlie others such as harm avoidance^{20,65}. 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 rapid⁶⁻⁷.

Lastly, sequence-based models have shown that mutations in evolutionarily conserved genes are more likely to result in deleterious phenotypes and disease⁶⁶. We find that several loci associated with adaptive morphological changes are implicated in disease phenotypes in humans and other organisms, suggesting the variation we have identified here underlying adaptive phenotypes may be deleterious in non-urban 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 humans⁶⁷⁻⁶⁹. 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 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 island⁷¹. Between 2012-2014 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 (Supplemental Fig. S13). Age and sex of lizards was determined based on snout-vent-length (minimum SVL 45mm) and secondary sexual characteristics (large dewlap, enlarged postanal scales, enlarged tail base).

We collected a sample of the distal tail (~10mm) 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 xrays (with a portable Kodex digital xray system) and high-resolution toepad scans (with an Epson flatbed scanner at 2100dpi). 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 ImageJ⁷² using the ObjectJ plugin. Limb bones and snout-vent length (SVL) were measured three times each from digital xrays. 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 xray), 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 counts 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 intraclass correlation coefficient in R with the function *ICCest* in R package 'ICC'⁷³, was high for all traits (ICC=0.97 for metacarpals, 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. 3F).

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 250-500 years. 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 months (e.g., Refs.75,76), 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 BIOCLIM⁷⁷ 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 (~1km²) 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 1km buffer around the perimeter. We also included two higher resolution (30m²) land-cover layers: impervious surface cover⁷⁸ and canopy cover⁷⁹, 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^{5,27,80}, we followed the procedure in Prado-Irwin et al.⁸⁰ to quantify perch availability, habitat openness, and anthropogenic perch presence. We obtained orthoimagery⁸¹ of each site and placed a 200m 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 5m 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 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 (Supplemental Fig. S13). The first three principal components captured 96.9% of variance. Higher values of PC1 indicated colder, wetter, and more variable climate (BIO1, BIO2, BIO6, BIO7, BIO9, BIO10, BIO11, BIO13, BIO15, BIO16) as well as more perches and less light at night. Higher values of PC2 indicated colder and wetter climate (BIO3, BIO4, BIO5, BIO8, BIO12, BIO14, BIO17, 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 (BIO2, BIO3, BIO5, BIO7, BIO12, BIO13, BIO15, BIO16, BIO18) as well as less light at night (Supplemental 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 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 hours with an additional 5μL of proteinase K added after 12 hours, the total elution volume was reduced to 70 μL and samples were washed with dH2O and the flow through elution. We verified DNA concentration fluorometrically using a Qubit v2.0 and 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 >120bp (the length of the bait) and with GC content 40-70%. We identified exons to target in two complementary sets of focal and non-focal 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)⁸²⁻⁸³ 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*¹⁵.

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 (>120bp, 40-70% GC), and randomly distributed the remaining probes throughout the remainder of the exome (“non-focal 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* genome⁸³ and a draft genome assembly of *A. sagrei*⁸⁴, 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 non-focal 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 species¹⁵), 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 400bp, 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 2x150, 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 Illumiprocessor⁸⁵ (v2.09) a wrapper for the read filtering program Trimmomatic⁸⁶ (v0.32). We created a non-redundant exome for *Anolis carolinensis* by removing duplicated exons from the *Anolis carolinensis* v2.1 exome⁸² using CD-HIT-EST⁸⁷ (v4.7). We aligned our quality filtered reads to this non-redundant set of *A. carolinensis* exons using BWA⁸⁸ (v0.7.17-r1188). We called and filtered variants using GATK⁸⁹ following the GATK best practices⁹⁰⁻⁹¹ 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” 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 VCFtools⁹² (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 snpEff⁹³ and the ASU_Acar_v2.1 annotation⁸². To convert from vcf to bed and ped formats, we used PLINK⁹⁴ and VCFtools⁹². To convert from vcf to genepop formats, we used STACKS⁹⁵. To subset the vcf file by municipality (San Juan, Arecibo, Mayagüez) we used bcftools⁹⁶.

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 bcftools⁹⁶ 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 10kb windows (retaining the site in an LD pair with the greater allele frequency). We investigated population structure via identity-by-state (IBS distance) and discriminant analysis of principal components (DAPC). We calculated IBS among all individuals across all six sites using PLINK⁹⁴ ($DST: (IBS2 + 0.5*IBS1) / (N \text{ SNP pairs})$), and tested if IBS differed across sites with ANOVA (high relatedness across sites and municipalities might suggest dispersal events; Supplemental Fig. S2). We implemented DAPC with the R package ‘adegenet’⁹⁷⁻⁹⁸ implemented with the function *dapc* (see Supplemental Fig. S1 for PCA results). Although *k*-means clustering implemented with the function *find.clusters* in the R package ‘adegenet’⁹⁷⁻⁹⁸ 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 function *xvalDapc*, which supported retaining 10 principal component axes. Discriminant functions 1 and 2 in the DAPC (Fig. 1C, Supplemental 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. 1D). Sequence alignment was performed with the ‘SNPhylo’ pipeline⁹⁹ followed by tree model fitting and optimization with IQTree¹⁰⁰ with ModelFinder¹⁰¹ and ascertainment bias for SNP data (-m TEST+ASC). The midpoint rooted sample tree was visualized in R with ‘phytools’¹⁰² and ‘phangorn’¹⁰³. We also estimated admixture coefficients using sparse Non-Negative Matrix

Factorization algorithms with the function *snmf* in the R package ‘LEA’¹⁰⁴; three genetic clusters were most strongly supported (Supplemental 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 10kb windows excluding any windows with fewer than 100 called sites (Supplemental Fig. S3). We used VCFtools⁹² to calculate observed heterozygosity (--het), relatedness (--relatedness), Tajima’s D (--TajimaD 100000), and unadjusted AJK statistic (Supplemental Figs. S2-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^{105–106}. PCA approaches are less powerful at detecting adaptive divergence when environmental differentiation is weakly correlated with population structure^{107–108}. Environmental association methods, in contrast, tend to detect more loci of small effect (spanning a range of F_{ST} values, e.g. ref. 109) 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’¹¹⁰ package. We analyzed each geographic region separately (San Juan, Mayagüez, 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; Supplemental Fig. S1). We retained 6–8 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 false-discovery rate of 1% by calculating q-values with the *qvalue* function in R package ‘qvalue’¹¹¹. 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’¹¹². Our model used a 50kb sliding window with a 10kb 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’¹¹² for each of the six traits. Our model used a 50kb sliding window with a 10kb 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^{105,108}. 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^{113–116}.

We tested the significance of the overlap in SNPs between the GEA and PCA with permutation and 1000 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 1000 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 1000 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 ‘gprofileR2’¹¹² 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=1776 genes). We provided as a custom background gene set the full list of genes containing at least one SNP in our dataset (n=6389 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’¹¹⁸ 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^{17,119-120}. We then used a linear model to determine the effect of habitat (urban or forest), municipality (Arecibo, Mayagüez, 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 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. 17).

Second, we investigated parallel divergence at the allele level by examining effect sizes (η^2) of allele frequencies for all SNPs in our dataset¹²¹. 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).

Acknowledgements and funding: 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 National Science Foundation to LJR (DEB 1354044), JBL & AJG (DEB 1927194), and LJR & KMW (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.

References

1. Shochat, E., Warren, P.S., Faeth, S.H., McIntyre, N.E., & Hope, D. From patterns to emerging processes in mechanistic urban ecology. *Trends in Ecology and Evolution* **21**, 186–191 (2006).
2. Johnson, M.T.J., & Munshi-South, J. Evolution of life in urban environments. *Science* **358**, 6363 (2017).
3. Santangelo, J.S., et al. Urban environments as a framework to study parallel evolution. in *Urban Evolutionary Biology* (eds Szulkin, M., Munshi-South, J., & Charmantier, A.) 36-53 (2020).
4. Szulkin, M., Munshi-South, J., & Charmantier, A. (Eds.). *Urban Evolutionary Biology*. Oxford University Press, USA. (2020).
5. Winchell, K.M., Reynolds, R. G., Prado-Irwin, S.R., Puente-Rolón, A.R., & Revell, L.J. Phenotypic shifts in urban areas in the tropical lizard *Anolis cristatellus*. *Evolution* **70**, 1009-1022 (2016).
6. Hendry, A.P., Farrugia, T.J., & Kinnison, M.T. Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology* **17**, 20-29 (2008).

- 615 7. Alberti, M., et al. Global urban signatures of phenotypic change in animal and plant
616 populations. *Proceedings of the National Academy of Sciences* **114**, 8951-8956 (2017).
- 617 8. United Nations Center for Human Settlement (HABITAT). Cities in a Globalizing World.
618 Global Report on Human Settlements 2001. (2001).
- 619 9. Des Roches, S., et al. Socio-Eco-Evolutionary Dynamics in Cities. *Evolutionary Applications*
620 **14**, 248–267 (2021).
- 621 10. Rivkin, L.R., et al. A roadmap for urban evolutionary ecology. *Evolutionary Applications*,
622 **12**, 384–398 (2019).
- 623 11. Miles, L.S., Carlen, E. J., Winchell, K.M., & Johnson, M.T. Urban evolution comes into its
624 own: Emerging themes and future directions of a burgeoning field. *Evolutionary*
625 *Applications* **14**, 3-11 (2021).
- 626 12. Verrelli, B.C. et al. A global horizon scan for urban evolutionary ecology. *Trends in Ecology*
627 *& Evolution* **37**, 1006-1019 (2022).
- 628 13. Perrier, C., Caizergues, A. and Charmantier, A., Adaptation genomics in urban
629 environments. . in *Urban Evolutionary Biology* (eds Szulkin, M., Munshi-South, J., &
630 Charmantier, A.) 74-90 (2020).
- 631 14. Ravinet, M., et al. Signatures of human-commensalism in the house sparrow genome.
632 *Proceedings of the Royal Society B* **285**, 20181246 (2018).
- 633 15. Campbell-Staton, S.C., et al. Parallel selection on thermal physiology facilitates repeated
634 adaptation of city lizards to urban heat islands. *Nature Ecology & Evolution* **4**, 652-658
635 (2020).
- 636 16. Mueller, J.C., et al. Genes acting in synapses and neuron projections are early targets of
637 selection during urban colonization. *Molecular Ecology* **29**, 3403-3412 (2020).
- 638 17. Salmón, P., et al. Continent-wide genomic signatures of adaptation to urbanisation in a
639 songbird across Europe. *Nature Communications* **12**, 1-14 (2021).
- 640 18. Harpak, A., et al. Genetic adaptation in New York city rats. *Genome Biology and Evolution*,
641 **13**, evaa247 (2021).
- 642 19. Miles, L.S., Rivkin, L.R., Johnson, M.T., Munshi-South, J., & Verrelli, B.C. Gene flow and
643 genetic drift in urban environments. *Molecular Ecology* **28**, 4138-4151 (2019).
- 644 20. Müller, J.C., Partecke, J., Hatchwell, B.J., Gaston, K.J., & Evans, K.L. Candidate gene
645 polymorphisms for behavioural adaptations during urbanization in blackbirds. *Molecular*
646 *Ecology* **22**, 3629-3637 (2013).
- 647 21. Harris, S. E., Munshi-South, J., Obergfell, C., & O'Neill, R. Signatures of rapid evolution in
648 urban and rural transcriptomes of white-footed mice (*Peromyscus leucopus*) in the New
649 York metropolitan area. *PLoS One* **8**, e74938 (2013).
- 650 22. Santangelo, J.S., et al. Global urban environmental change drives adaptation in white
651 clover. *Science* **375**, 1275-1281 (2022).
- 652 23. Losos, J. Lizards in an evolutionary tree. In *Lizards in an Evolutionary Tree*. University of
653 California Press. (2009).
- 654 24. Isaakson, C. & Bonier, F. Urban Evolutionary Physiology. in *Urban Evolutionary Biology*
655 (eds Szulkin, M., Munshi-South, J., & Charmantier, A.) 217-233 (2020).
- 656 25. Bradley, C.A., & Altizer, S. Urbanization and the ecology of wildlife diseases. *Trends in*
657 *Ecology & Evolution* **22**, 95-102 (2007).

- 658 26. Sol, D., Lapiedra, O., & Ducatez, S. Cognition and adaptation to urban environments. in
659 *Urban Evolutionary Biology* (eds Szulkin, M., Munshi-South, J., & Charmantier, A.) 253-267
660 (2020).
- 661 27. Winchell, K.M., Battles, A.C., & Moore, T. Y. Terrestrial locomotor evolution in urban
662 environments. in *Urban Evolutionary Biology* (eds Szulkin, M., Munshi-South, J., &
663 Charmantier, A.) 197-216 (2020).
- 664 28. Diamond, S.E., & Martin, R.A. Evolutionary consequences of the urban heat island. in
665 *Urban Evolutionary Biology* (eds Szulkin, M., Munshi-South, J., & Charmantier, A.) 91-110
666 (2020).
- 667 29. French, S.S., Fokidis, H.B., & Moore, M.C. Variation in stress and innate immunity in the
668 tree lizard (*Urosaurus ornatus*) across an urban–rural gradient. *Journal of Comparative*
669 *Physiology B* **178**, 997-1005 (2008).
- 670 30. Tyler, R.K., Winchell, K.M., & Revell, L.J. Tails of the city: caudal autotomy in the tropical
671 lizard, *Anolis cristatellus*, in urban and natural areas of Puerto Rico. *Journal of Herpetology*
672 **50**, 435-441 (2016).
- 673 31. Winchell, K.M., Briggs, D., & Revell, L.J. The perils of city life: patterns of injury and
674 fluctuating asymmetry in urban lizards. *Biological Journal of the Linnean Society* **126**, 276-
675 288 (2019).
- 676 32. Thawley, C.J., et al. Urbanization affects body size and parasitism but not thermal
677 preferences in *Anolis* lizards. *Journal of Urban Ecology* **5**, juy031 (2019).
- 678 33. Shafit-Zagardo, B., & Kalcheva, N. Making sense of the multiple MAP-2 transcripts and
679 their role in the neuron. *Molecular Neurobiology* **16**, 149-162 (1998).
- 680 34. Yeh, E., et al. A putative cation channel, NCA-1, and a novel protein, UNC-80, transmit
681 neuronal activity in *C. elegans*. *PLoS biology* **6**, e55 (2008).
- 682 35. Tischfield, D.J., et al. Loss of the neurodevelopmental gene *Zswim6* alters striatal
683 morphology and motor regulation. *Neurobiology of Disease* **103**, 174-183 (2017).
- 684 36. Song, Y., et al. Knockdown of *Pnpla6* protein results in motor neuron defects in zebrafish.
685 *Disease Models & Mechanisms* **6**, 404-413 (2013).
- 686 37. Go, G.W., & Mani, A. Low-density lipoprotein receptor (LDLR) family orchestrates
687 cholesterol homeostasis. *The Yale journal of biology and medicine* **85**, 19 (2012).
- 688 38. Boutchueng-Djidjou, M., et al. The Last Enzyme of the De Novo Purine Synthesis Pathway
689 5-aminoimidazole-4-carboxamide Ribonucleotide Formyltransferase/IMP Cyclohydrolase
690 (ATIC) Plays a Central Role in Insulin Signaling and the Golgi/Endosomes Protein
691 Network*[S]. *Molecular & Cellular Proteomics* **14**, 1079-1092 (2015).
- 692 39. Matsumoto, S., et al. Urea cycle disorders—update. *Journal of human genetics* **64**, 833-
693 847 (2019).
- 694 40. Akiyama, M. The roles of ABCA12 in epidermal lipid barrier formation and keratinocyte
695 differentiation. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*
696 **1841**, 435-440 (2014).
- 697 41. Wu, X.S., Martina, J.A., & Hammer III, J.A. Melanoregulin is stably targeted to the
698 melanosome membrane by palmitoylation. *Biochemical and biophysical research*
699 *communications* **426**, 209-214 (2012).

42. Sanger, T.J., Revell, L.J., Gibson-Brown, J.J., & Losos, J.B. Repeated modification of early limb morphogenesis programmes underlies the convergence of relative limb length in *Anolis* lizards. *Proceedings of the Royal Society B: Biological Sciences* **279**, 739-748 (2012).
43. Park, S., Infante, C.R., Rivera-Davila, L.C., & Menke, D.B. Conserved regulation of *hoxc11* by *pitx1* in *Anolis* lizards. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **322**, 156-165 (2014).
44. Winchell, K.M., Maayan, I., Fredette, J.R., & Revell, L.J. Linking locomotor performance to morphological shifts in urban lizards. *Proceedings of the Royal Society B* **285**, 20180229 (2018).
45. Kolbe, J.J., Battles, A.C., & Avilés-Rodríguez, K.J. City slickers: poor performance does not deter *Anolis* lizards from using artificial substrates in human-modified habitats. *Functional Ecology* **30**, 1418-1429 (2016).
46. Meng, S., et al. TBX20 Regulates Angiogenesis Through the Prokineticin 2–Prokineticin Receptor 1 Pathway. *Circulation* **138**, 913-928 (2018).
47. Ingraham, C.R., et al. Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (*Irf6*). *Nature Genetics* **38**, 1335-1340 (2006).
48. Eppley, S., Hopkin, R. J., Mendelsohn, B., & Slavotinek, A.M. Clinical Report: Warsaw Breakage Syndrome with small radii and fibulae. *American Journal of Medical Genetics Part A* **173**, 3075-3081 (2017).
49. Seifert, M., Tilgen, W., & Reichrath, J. Expression of 25-hydroxyvitamin D-1 α -hydroxylase (1 α OHase, 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. Chiquet, M., Birk, D.E., Bönnemann, C. G., & Koch, M. Collagen XII: Protecting bone and muscle integrity by organizing collagen fibrils. *The International Journal of Biochemistry & Cell Biology* **53**, 51-54 (2014).
51. Alibardi, L. Differentiation of the epidermis during scale formation in embryos of lizard. *The Journal of Anatomy* **192**, 173-186 (1998).
52. Halim, D., et al. Loss of LMOD1 impairs smooth muscle cytocontractility and causes megacystis microcolon intestinal hypoperistalsis syndrome in humans and mice. *Proceedings of the National Academy of Sciences* **114**, E2739-E2747 (2017).
53. McGregor, L. *Locomotor morphology of Anolis: Comparative investigations of the design and function of the subdigital adhesive system*. University of Calgary. (1999).
54. Safran, M., Rosen, N., Twik, M., BarShir, R., Iny Stein, T., Dahary, D., Fishilevich, S., & Lancet, D. *The GeneCards Suite Chapter, Practical Guide to Life Science Databases*, 27-56 (2022).
55. Paradise, C.R., Galvan, M.L., Pichurin, O., Jerez, S., Kubrova, E., Dehghani, S.S., Carrasco, M.E., Thaler, R., Larson, A.N., van Wijnen, A.J. & Dudakovic, A. *Brd4* is required for chondrocyte differentiation and endochondral ossification. *Bone* **154**, 116234 (2022).
56. Naja, R.P., Dardenne, O., Arabian, A., & St. Arnaud, R. 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. Lu, J., Lian, G., Lenkinski, R., De Grand, A., Vaid, R.R., Bryce, T., Stasenko, M., Boskey, A., Walsh, C. & Sheen, V. Filamin B mutations cause chondrocyte defects in skeletal development. *Human Molecular Genetics* **16**,1661-1675 (2007).
58. Bentmann, A., Kawelke, N., Moss, D., Zentgraf, H., Bala, Y., Berger, I., Gasser, J.A. & Nakchbandi, I.A. Circulating fibronectin affects bone matrix, whereas osteoblast fibronectin modulates osteoblast function. *Journal of Bone and Mineral Research* **25**, 706-715 (2010).
59. Cosentino, B.J., & Gibbs, J.P. Parallel evolution of urban–rural clines in melanism in a widespread mammal. *Scientific Reports* **12**, 1-7 (2022).
60. Elmer, K.R. & Meyer, A. Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends in Ecology & Evolution* **26**, 298-306 (2011).
61. Lee, K.M., & Coop, G. Population genomics perspectives on convergent adaptation. *Phil. Trans. R. Soc. B* **374**, 20180236 (2019).
62. Corbett-Detig, R.B., Russell, S. L., Nielsen, R., & Losos, J. Phenotypic Convergence Is Not Mirrored at the Protein Level in a Lizard Adaptive Radiation. *Molecular Biology and Evolution* **37**, 1604-1614 (2020).
63. Hoekstra, H.E., & Coyne, J.A.. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* **61**, 995 (2007).
64. Cassone, B.J., et al. Gene expression divergence between malaria vector sibling species *Anopheles gambiae* and *An. coluzzii* from rural and urban Yaoundé Cameroon. *Molecular Ecology* **23**, 2242-2259 (2014).
65. van Dongen, W.F., Robinson, R.W., Weston, M.A., Mulder, R.A., & Guay, P.J. Variation at the DRD4 locus is associated with wariness and local site selection in urban black swans. *BMC Evolutionary Biology* **15**, 1-11 (2015).
66. Subramanian S., & Kumar S. Evolutionary anatomies of positions and types of disease-associated and neutral amino acid mutations in the human genome. *BMC Genomics* **7**, 306 (2006).
67. Gurdasani, D., et al. Uganda Genome Resource Enables Insights into Population History and Genomic Discovery in Africa. *Cell* **179**, 984–1002.e36 (2019).
68. Hancock, A.M., et al. Colloquium Paper: Human Adaptations to Diet, Subsistence, and Ecoregion Are Due to Subtle Shifts in Allele Frequency. *PNAS* **107**, 8924–8930 (2010).
69. Stover, D.A., Housman, G., Stone, A.C., Rosenberg, M.S., & Verrelli, B.C. Evolutionary genetic signatures of selection on bone-related variation within human and chimpanzee populations. *Genes* **13**, 183 (2022).
70. Gould, W.A., Martinuzzi, S., & 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).

Supplemental References (Methods):

71. Reynolds, R.G., et al. Archipelagic genetics in a widespread Caribbean anole. *Journal of Biogeography* **44**, 2631-2647 (2017).
72. Rasband, W.S. ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, <https://imagej.nih.gov/ij/>, 1997-2018.

- 785 73. Wolak, M.E., Fairbairn, D.J., & Paulsen, Y.R. Guidelines for Estimating Repeatability.
786 *Methods in Ecology and Evolution* **3**, 129-137 (2012).
- 787 74. Alboukadel, K. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package
788 version 0.7.0. <https://CRAN.R-project.org/package=rstatix> (2021).
- 789 75. Eales, J., Thorpe, R.S., & Malhotra, A. Weak founder event signal in a recent introduction
790 of Caribbean *Anolis*. *Molecular Ecology* **17**, 1416-1426 (2008).
- 791 76. Reynolds, R.G., Strickland, T.R., Kolbe, J.J., Falk, B.K., Perry, G., Revell, L.J., Losos, J.B.
792 Archipelagic genetics in a widespread Caribbean anole. *Journal of Biogeography* **44**, 2631-
793 2647 (2017).
- 794 77. Fick, S.E., & Hijmans, R.J. WorldClim 2: new 1km spatial resolution climate surfaces for
795 global land areas. *International Journal of Climatology* **37**, 4302-4315 (2017).
- 796 78. Homer, C.G., Huang, C., Yang, L., Wylie, B.K., & Coan, M.J., Development of a 2001
797 National Land Cover Database for the United States: Photogrammetric Engineering and
798 Remote Sensing, v. 70, no. 7, p. 829–840.
- 799 79. Coulston, J.W., et al. Modeling percent tree canopy cover—A pilot study:
800 Photogrammetric Engineering and Remote Sensing, v. 78, no. 7, p. 715–727.
- 801 80. Prado-Irwin, S.R., Revell, L.J., & Winchell, K.M. Variation in tail morphology across urban
802 and forest populations of the crested anole (*Anolis cristatellus*). *Biological Journal of the*
803 *Linnean Society* **128**, 632-644 (2019).
- 804 81. US Geological Survey. *USGS High Resolution Orthoimagery for Puerto Rico and the U.S.*
805 *Virgin Islands*. Available at: <https://www.sciencebase.gov/catalog/item/59efa>
806 [a13e4b0220bbd99b87b](https://www.sciencebase.gov/catalog/item/59efa) (2012).
- 807 82. Eckalbar, W.L. et al. Genome reannotation of the lizard *Anolis carolinensis* based on 14
808 adult and embryonic deep transcriptomes. *BMC Genomics* **14**, 49 (2013).
- 809 83. Alföldi, J., et al. The genome of the green anole lizard and a comparative analysis with
810 birds and mammals. *Nature* **477**, 587-591 (2011).
- 811 84. Geneva, A.J., et al. Chromosome-scale genome assembly of the brown anole (*Anolis*
812 *sagrei*), a model species for evolution and ecology. *Communications Biology* **5**, 1126
813 (2022).
- 814 85. Faircloth, B.C. illumiprocessor: a trimmomatic wrapper for parallel adapter and quality
815 trimming. (2013).
- 816 86. Bolger, A.M., Lohse, M., & Usadel, B. Trimmomatic: A flexible trimmer for Illumina
817 Sequence Data. *Bioinformatics* **btu170** (2014).
- 818 87. Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. CD-HIT: accelerated for clustering the next
819 generation sequencing data. *Bioinformatics* **28**, 3150-3152 (2012).
- 820 88. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
821 *arXiv*, 1303.3997.
- 822 89. McKenna, A., et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing
823 next-generation DNA sequencing data. *Genome Res* **20**, 1297-303 (2010).
- 824 90. DePristo, M., et al. A framework for variation discovery and genotyping using next-
825 generation DNA sequencing data. *Nat Genet* **43**, 491-498 (2011).
- 826 91. Van der Auwera, G.A., & O'Connor, B.D. Genomics in the Cloud: Using Docker, GATK, and
827 WDL in Terra (1st Edition). O'Reilly Media. (2020).
- 828 92. Danecek, P., et al. The Variant Call Format and VCFtools, *Bioinformatics* (2011).

93. Cingolani, P., et al. 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). 2012 Apr-Jun;6(2):80-92. PMID: 22728672
94. Purcell, S., et al. PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics* **81** (2007).
95. Catchen, J., Hohenlohe, P., Bassham, S., Amores, A., & Cresko, W. *Stacks: an analysis tool set for population genomics*. *Molecular Ecology* (2013).
96. Danecek, P., et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
97. Jombart, T. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403-1405 (2008).
98. Jombart, T., & Ahmed, I. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* (2011).
99. Lee, T.H., Guo, H., Wang, X., Kim, C., & Paterson, A.H. SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genomics* **15** (2014).
100. Minh, B.Q., et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530-1534 (2020).
101. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., & Jermin, L.S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
102. Revell, L.J. phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217-223 (2012).
103. Schliep, K.P. phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592-593 (2011).
104. Frichot, E., & François, O. LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution* **6**, 925-929 (2015).
105. Hoban, S., et al. Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *American Naturalist* **188**, 379–397 (2016).
106. Luu, K., Bazin, E., & Blum, M. G. pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources* **17**, 67-77 (2017).
107. Capblancq, T., Luu, K., Blum, M. G., & Bazin, E. Evaluation of redundancy analysis to identify signatures of local adaptation. *Molecular Ecology Resources* **18**, 1223-1233 (2018).
108. Lotterhos, K.E., & Whitlock, M.C. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology* **24**, 1031-1046 (2015).
109. Brauer, C.J., Hammer, M.P., & Beheregaray, L.B. Riverscape genomics of a threatened fish across a hydroclimatically heterogeneous river basin. *Molecular Ecology* **25**, 5093-5113 (2016).
110. Privé, F., et al. Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. *Molecular Biology and Evolution* (2020).
111. Storey, J.D., Bass, A.J., Dabney, A., & Robinson, D. qvalue: Q-value estimation for false discovery rate control. R package version 2.22.0. (2020).

- 872 112. Gogarten, S.M., et al. Genetic association testing using the GENESIS R/Bioconductor
873 package. *Bioinformatics* (2019).
- 874 113. Logan, Malcolm. SAGE profiling of the forelimb and hindlimb. *Genome Biology* **3.3**, 1-3
875 (2002).
- 876 114. Melville, J., Hunjan, S., McLean, F., Mantziou, G., Boysen, K. & Parry, L.J. Expression of a
877 hindlimb-determining factor Pitx1 in the forelimb of the lizard *Pogona vitticeps* during
878 morphogenesis. *Open Biology* **6**, 160252 (2016).
- 879 115. Minguillon, C., Del Buono, J. & Logan, M.P. Tbx5 and Tbx4 are not sufficient to determine
880 limb-specific morphologies but have common roles in initiating limb outgrowth.
881 *Developmental Cell* **8**, 75-84 (2005).
- 882 116. Shubin, N., Tabin, C. & Carroll, S. Fossils, genes and the evolution of animal limbs. *Nature*
883 **388**, 639–648 (1997).
- 884 117. Kolberg, L., & Raudvere, U. gprofiler2: Interface to the 'g:Profiler' Toolset. R package
885 version 0.2.0 (2020).
- 886 118. Xiuwen, Z., et al. A High-performance Computing Toolset for Relatedness and Principal
887 Component Analysis of SNP Data. *Bioinformatics* (2012).
- 888 119. Li, H., & Ralph, P. Local PCA shows how the effect of population structure differs along the
889 genome. *Genetics* **211**, 289-304 (2019).
- 890 120. Montejo-Kovacevich, G., et al. Repeated genetic adaptation to high altitude in two tropical
891 butterflies. *bioRxiv* (2021).
- 892 121. Bolnick, D. I., Barrett, R. D., Oke, K. B., Rennison, D. J. & Stuart, Y. E. (Non) parallel
893 evolution. *Annu. Rev. Ecol. Evol. Syst.* **49** 303–330 (2018).

Figure Legends

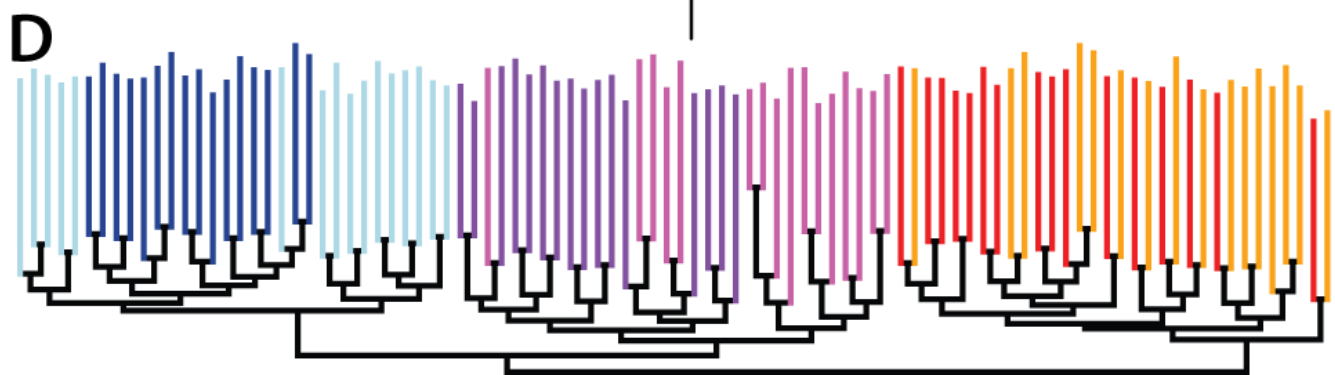
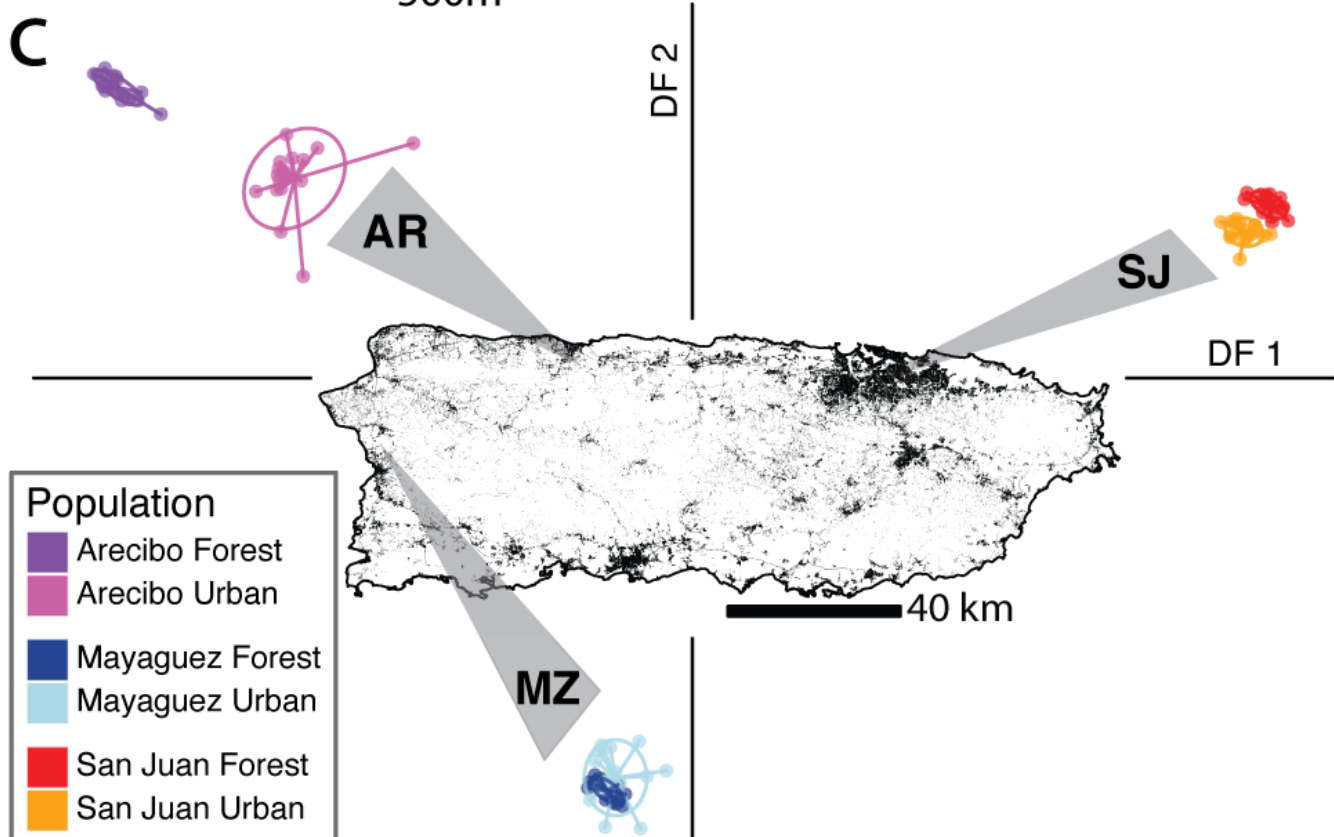
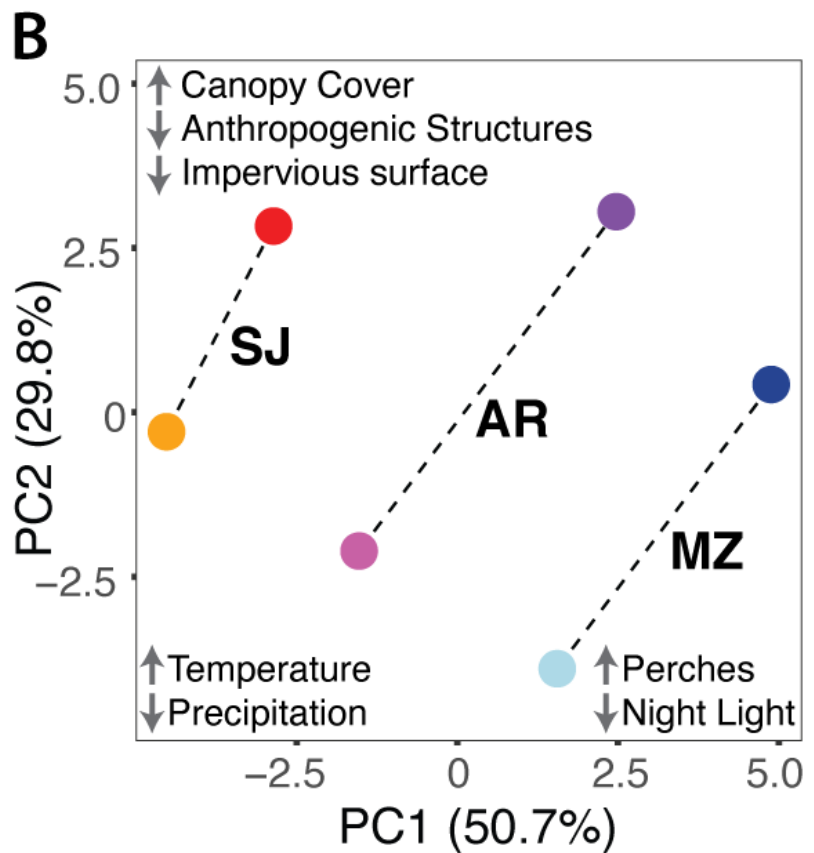
Fig. 1 – Environmental and population divergence. We sampled paired urban-forest sites in three municipalities (regions) across the island of Puerto Rico. Population structure analyses support independent urban-forest pairs in each geographic region. Across all panels, colors correspond to municipality 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 & Maxar Technologies 2001). Estimated dates of urban establishment are indicated below the urban images. (B) Urban and forest habitats differ in parallel in multi-dimensional 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 extent⁷⁰ 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.

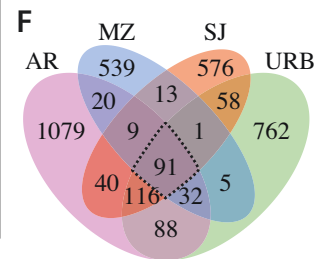
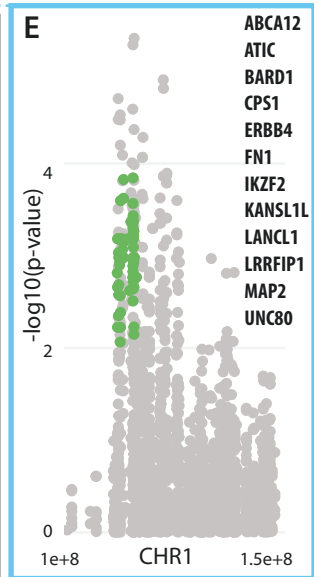
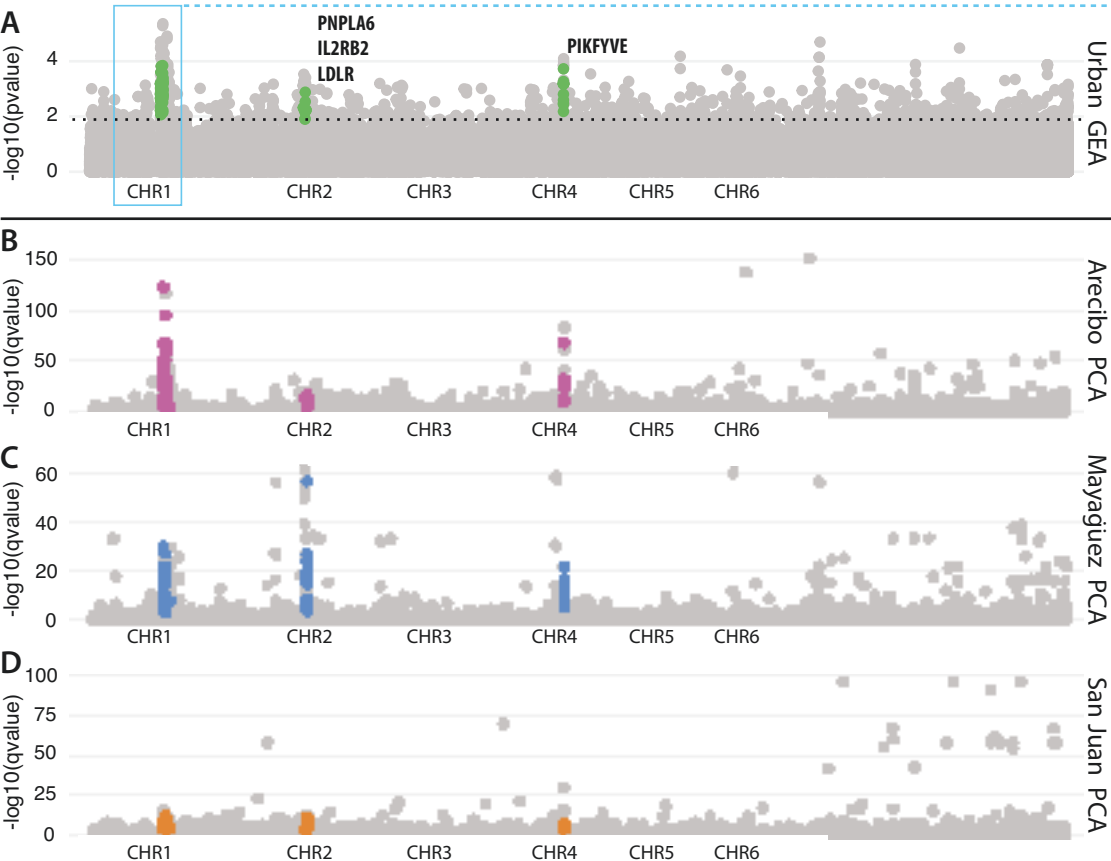
Fig. 2 – Parallelism of urban associated genomic changes. (A) Manhattan plot of SNPs identified by the urban genotype-environment association test (GEA), with significance threshold indicated by black dotted line and genes containing shared outlier SNPs listed next to the peaks for chromosomes 2 and 4. We complemented this analysis with three PCAs, one for each municipality: (B) Arecibo, (C) Mayagüez, (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 genotype-environment association test. The 33 urban-associated genes contained outlier SNPs in all four tests. Larger versions of all Manhattan plots are in Supplemental Fig. S5.

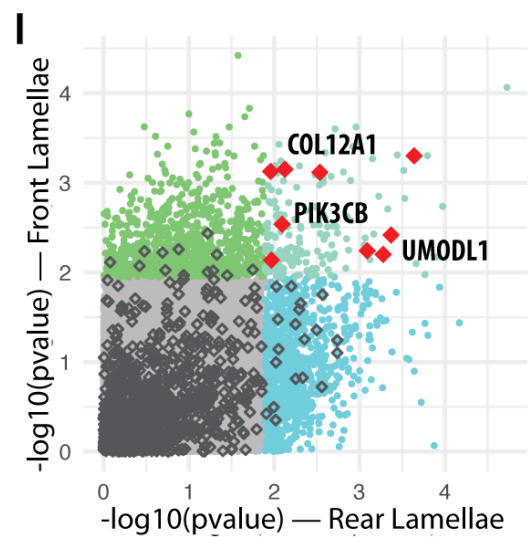
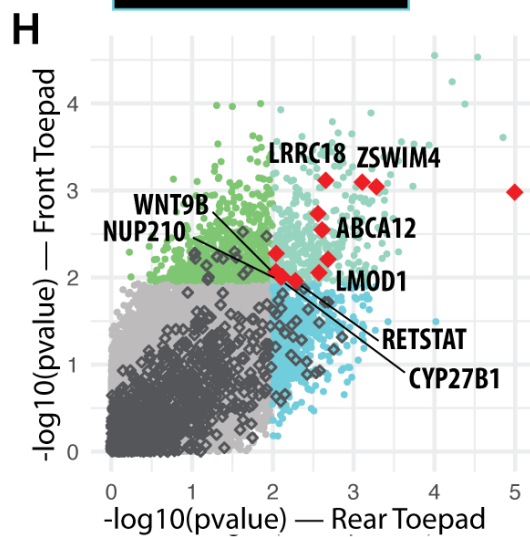
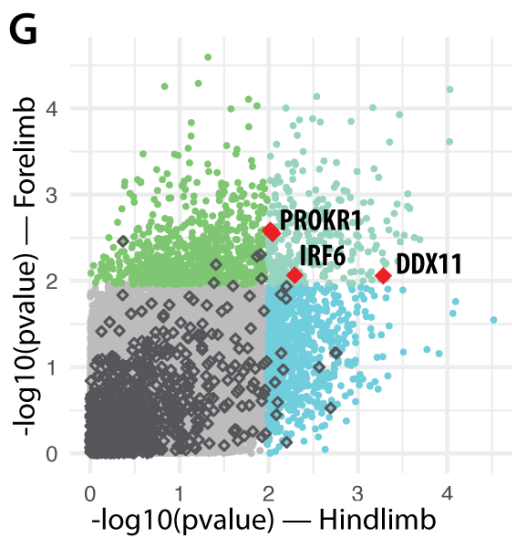
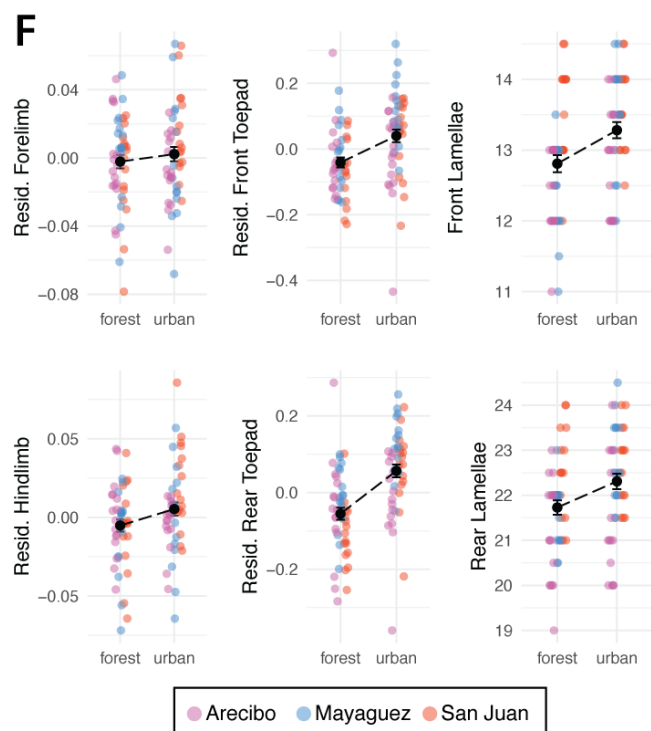
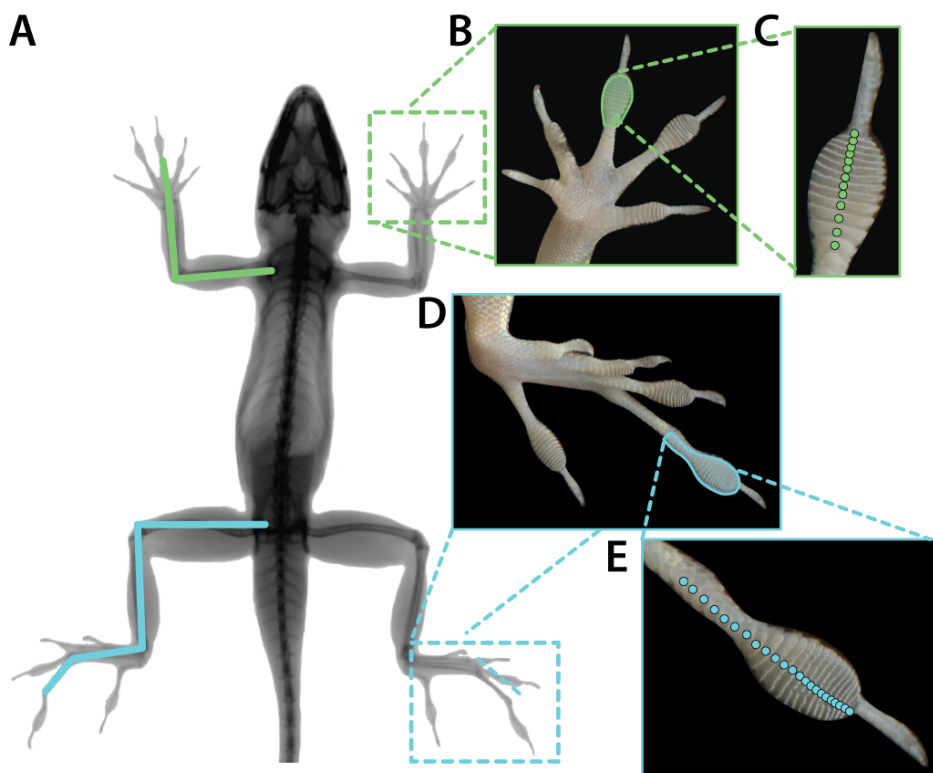
Fig. 3 — Phenotypic parallelism and genomic underpinnings. We focused on six morphological traits with known urban-associated divergence: (A) forelimb and hindlimb lengths, (B) front toepad area, (C) front toepad lamella count, (D) rear toepad area, (E) rear toepad lamella count. (F) At the phenotypic level: mean and standard error 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 standard error by habitat shown in black. At the genomic level: overlap in outlier SNPs for each of the three traits between hind- (blue points) and forelimb elements (green points) for each trait: (G) limbs, (H) toepad area, (I) toepad lamellae; SNPs associated with both hind- 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 fore- and hindlimb morphological elements. Gene names correspond to one or more of the urban-morphology SNPs in the upper right quadrant (red diamond).

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 standard error 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) indicate 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 standard error 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 two-sided t-test against the null expectation: $p=0.01$ **, $p<0.001$ ***.







Limb

Lamellae

Toepad Area

● Forest

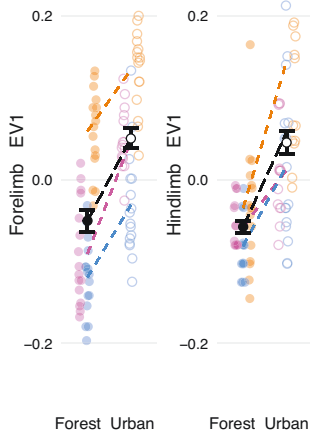
○ Urban

● Arecibo

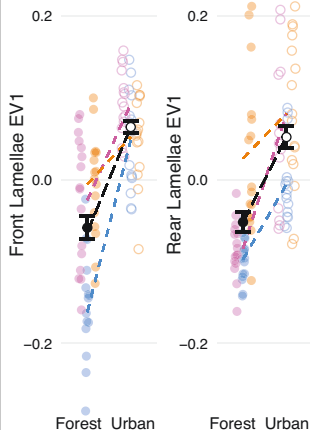
● Mayagüez

● San Juan

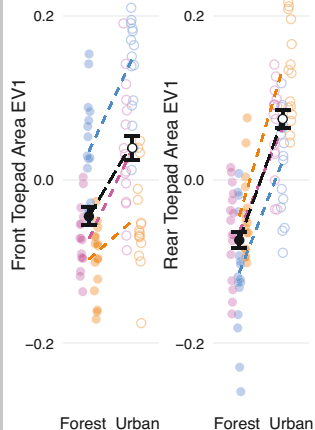
A



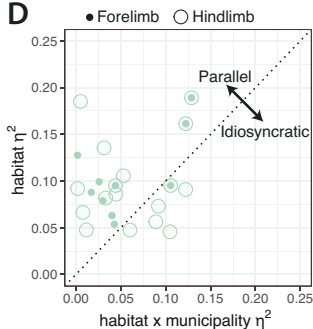
B



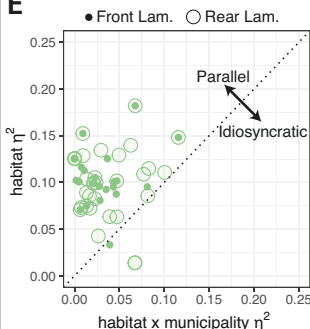
C



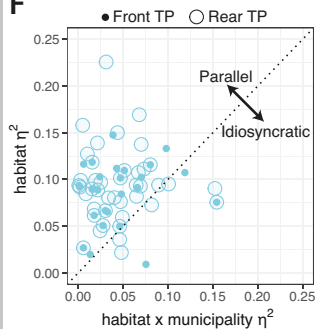
D



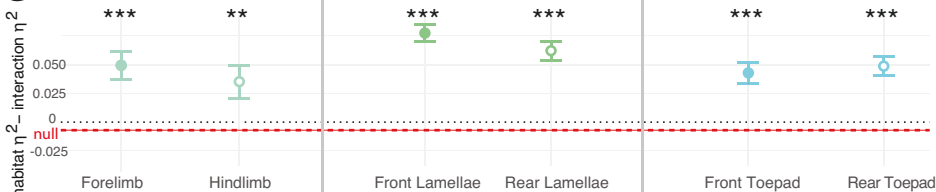
E



F



G



Supporting Information for

Genome-wide parallelism underlies contemporary adaptation in urban lizards

Kristin M. Winchell, Shane C. Campbell-Staton, Jonathan B. Losos, Liam J. Revell, Brian C. Verrelli, Anthony J. Geneva

Kristin Winchell & Jonathan Losos
Email: kristin.winchell@nyu.edu & losos@wustl.edu

This PDF file includes:

Figures S1-S7, S11-S13
Tables S8-S10

Other supporting materials for this manuscript include the following:

Raw sequence data: NCBI Sequence Read Archive, BioProject PRJNA872192
Data and associated code needed to repeat analyses: Zenodo, doi:
[10.5281/zenodo.6636371](https://doi.org/10.5281/zenodo.6636371)

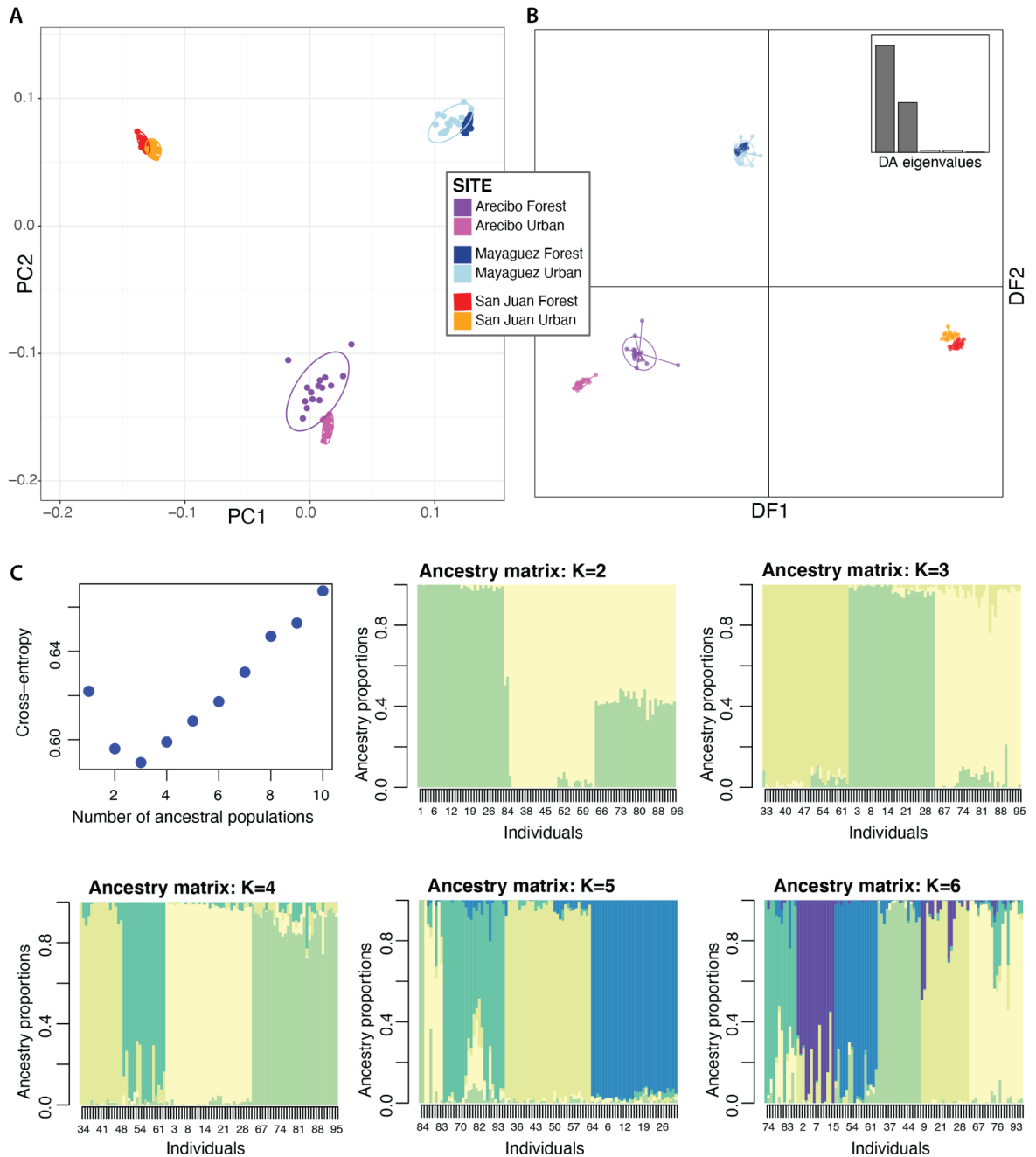


Fig. S1: Population Structure Analyses. (A) Principal components of genetic variation PC1 and PC2 shows three distinct clusters by region — Arecibo (AR), Mayagüez (MZ), and San Juan (SJ). All 96 samples are plotted with 95% confidence ellipses by site (one urban, one forest per region). (B) Discriminant analysis of principal components (DAPC) shows population structure of samples with 95% confidence ellipses. (C) Ancestry analysis using sparse nonnegative matrix factorization (snmf) suggests 3 ancestral populations. Cross-entropy values for 1-10 possible ancestral populations and ancestry matrix per individual for K=2-6.

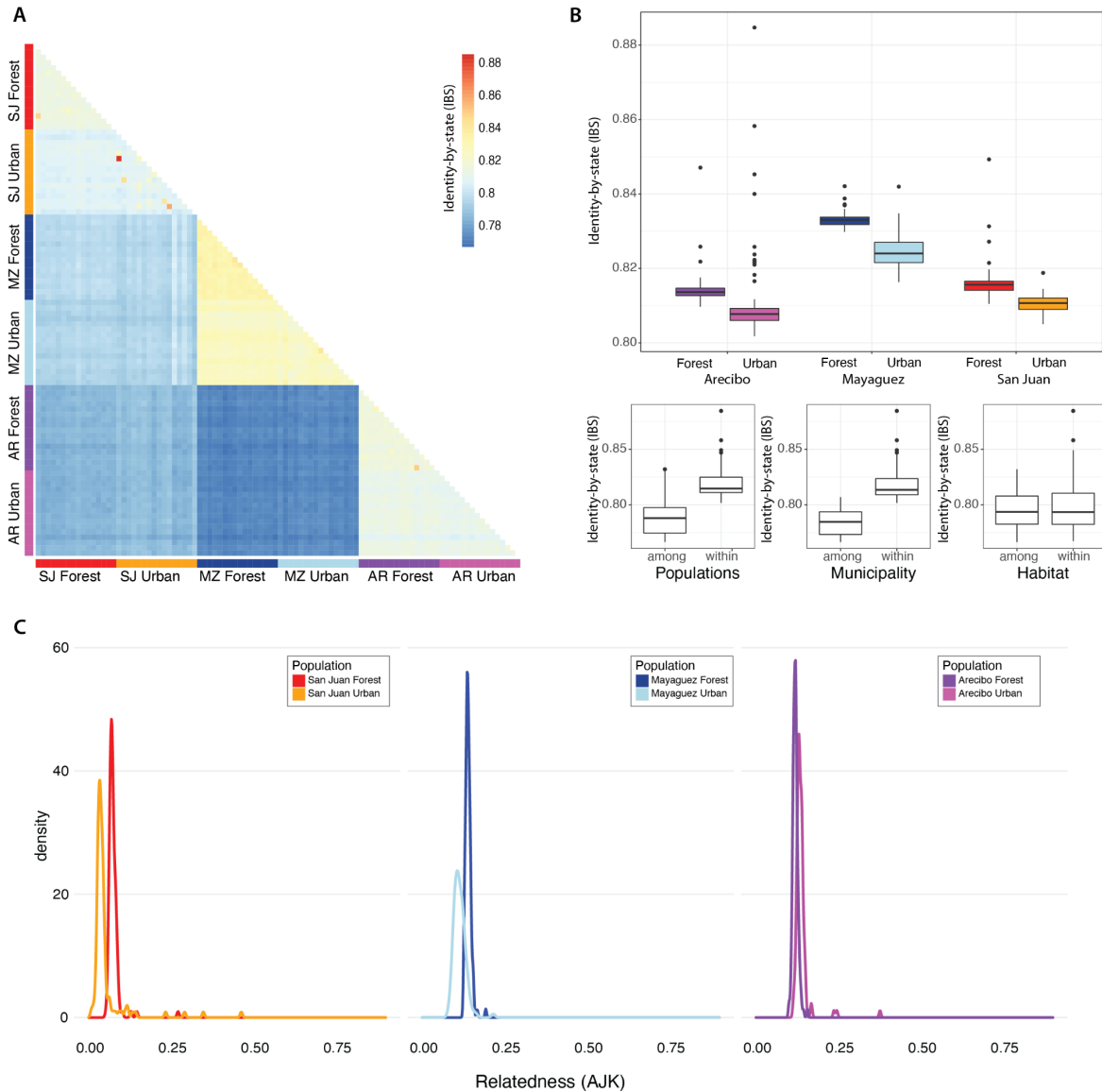


Fig. S2: Relatedness Analyses. (A) Pairwise IBS distances across all individuals. (B) Boxplots of pairwise IBS by habitats and municipalities. Individuals are more closely related to each other in forest vs urban populations in each region (top). Across all individuals, (bottom left) individuals from the same population are more closely related to each other than to individuals from different populations, and (bottom middle) individuals from the same municipality are more closely related to each other than to individuals from different populations, but (bottom right) individuals from the same habitat type (urban or forest) are not more closely related to each other than to individuals from the other habitat. In each, the center line represents the median, box limits represent upper and lower quartiles, whiskers represent 1.5x interquartile range, and points represent outliers. (C) Unadjusted AJK statistic based on the method of Yang et al. (2010), calculated in vcftools; relatedness is slightly higher in forest populations.

Yang et al. Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* **42**, 565-569 (2010).



Fig. S3: Genetic Diversity Analyses. (A) Genome-wide average pairwise F_{ST} (upper triangle) and d_{xy} (lower triangle). Darker colors for each represent higher values (note: scales differ for each metric). (B) Genome-wide average and standard error for Tajima's D in forest and urban populations in each of the three municipalities. In each urban-forest pair, genome-wide Tajima's D is elevated in the urban population compared to the forest. (C) Genome-wide genetic diversity, population averages: observed heterozygosity (H_o), IBS distance, allele diversity (π), unadjusted AJK statistic of relatedness, and Tajima's D. We found that urban populations exhibited slightly elevated heterozygosity compared to their forest counterparts (ANOVA: $F_{df=1, 90}=15.12$, $p<0.001$), and slightly lower Tajima's D (ANOVA: $F_{df=1, 45844}=159.9$, $p<0.001$), although Tajima's D was positive for both urban and forest populations. Genome-wide measures of F_{ST} and d_{xy} mirrored these findings (parts A & B).

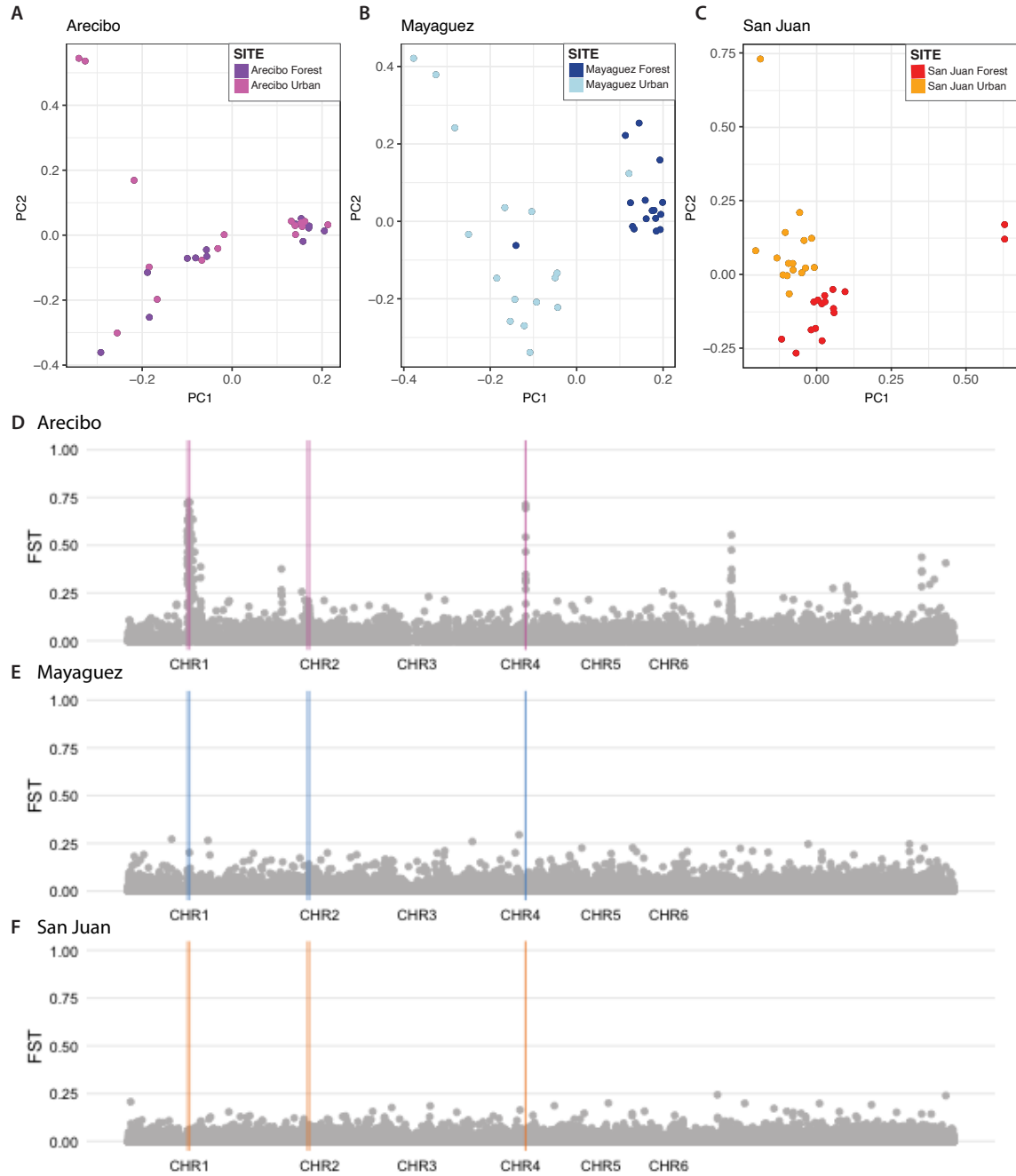


Fig. S4: Urban-forest divergence. (A) Principal components of genetic variation PC1 and PC2 for each municipality separately. (B) F_{ST} Manhattan plots for each municipality, with outlier regions identified in the main text highlighted with vertical bars. Colors are consistent with municipality coloration throughout the text.

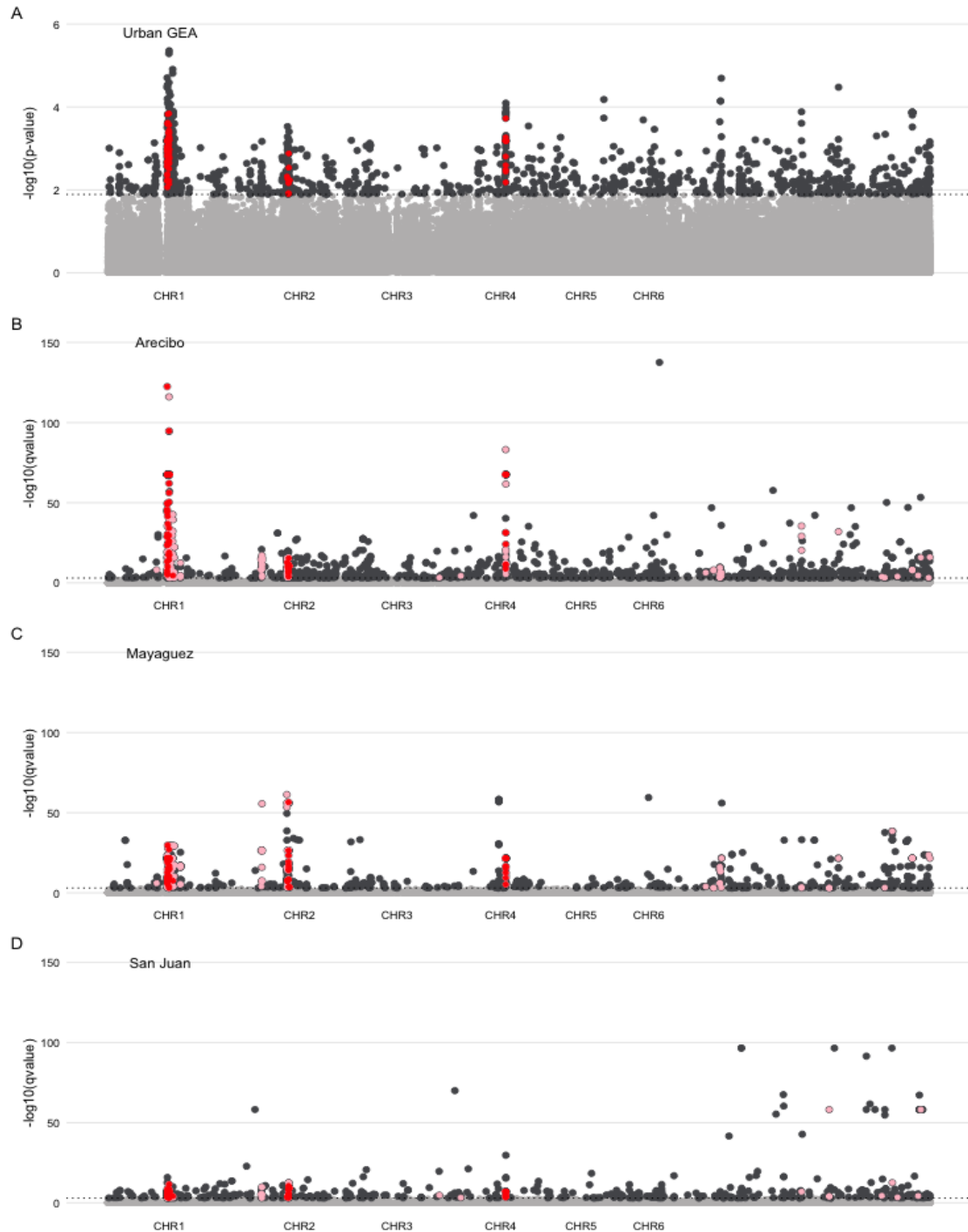


Fig. S5: Manhattan Plots. (A) Manhattan plot for urbanization GEA analysis (Fig. 2A). Dark colored dots are outliers (above dashed black line), with red dots indicating outliers identified by both the GEA analysis and each municipality-specific PCA analysis. (B-D) Manhattan plots for urbanization PCA analyses (Fig. 2B-D) for each municipality: (B) Arecibo, (C) Mayagüez, (D) San Juan. Dark colored dots are outliers (above dashed black line), with pink dots indicating outliers shared among at least one other municipality and red dots indicating outliers shared among all three municipalities.

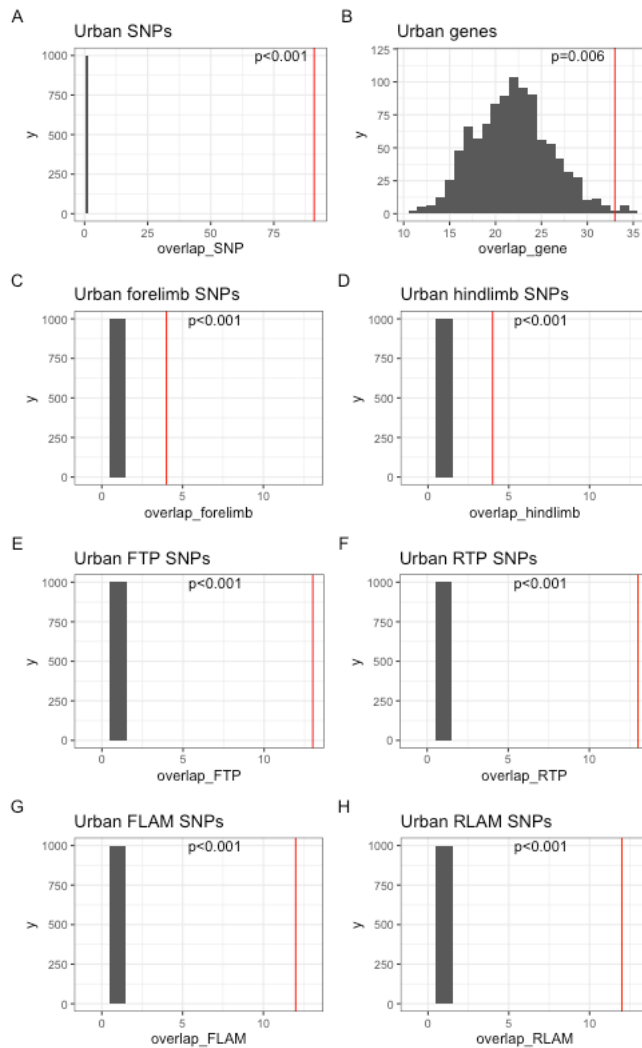


Figure S6: Permutation testing for SNP overlap. We compared the observed number of SNPs identified in multiple analyses (indicated by red lines) against a null distribution generated by 1000 permutations, randomly sampling SNPs without replacement and finding the intersection among sets. (A) Intersection of SNPs identified by each municipality specific PCA and the urban GEA. (B) Intersection of the gene sets in which the SNPs from A are found. (C) Intersection of forelimb SNPs and urban GEA. (D) Intersection of hindlimb SNPs and urban GEA. (E) Intersection of front toepad area SNPs and urban GEA. (F) Intersection of rear toepad area SNPs and urban GEA. (G) Intersection of front lamella number SNPs and urban GEA. (H) Intersection of rear lamella number SNPs and urban GEA.

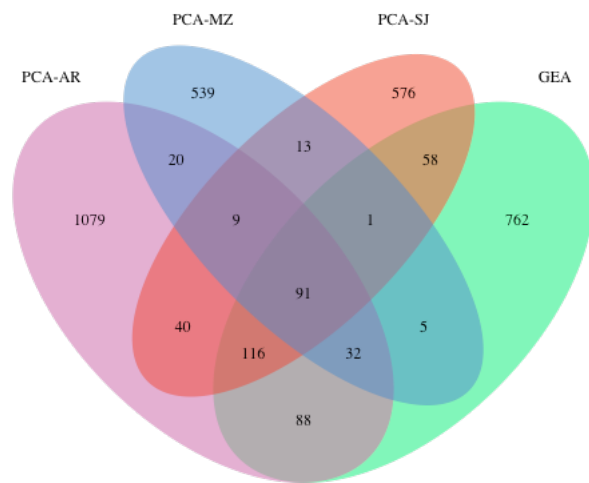


Figure S7: Parallelism at SNP level for urban-associated divergence. Number of SNPs identified by each analysis (PCA and GEA) and the overlap across municipalities (AR: Arecibo, MZ: Mayagüez, SJ: San Juan).

Table S8: Core Urbanization Genes. Core urbanization genes, with chromosome number or scaffold ID (CHR), basepair range of the gene (BP RANGE), Ensembl Gene ID, Ensembl Gene Name, and number of outlier SNPs from each analysis (PCA: AR – Arecibo, MZ – Mayagüez, SJ – San Juan; GEA analysis of urbanization). Core urbanization genes are genes containing at least one outlier SNPs as identified by all three regional PCAs and the urban GEA analysis. Red gene names indicate orthologous genes.

CHROM	BP RANGE	Ensembl Gene ID	Ensembl Gene Name	AR	MZ	SJ	URB
1	113066381-113117937	ENSACAG000000010914	MAP2	3	4	3	3
1	113141429-113340095	ENSACAG000000010788	UNC80	16	11	13	15
1	113358033-113391188	ENSACAG000000010727	KANSL1L	3	5	7	2
1	113528903-113590569	ENSACAG000000010566	LANCL1	6	6	6	8
1	113615520-113785329	ENSACAG000000010170	CPS1	16	14	17	18
1	114201462-114850987	ENSACAG000000009921	ERBB4	14	13	8	12
1	115423904-115538818	ENSACAG000000025864	IKZF2	1	1	1	1
1	115857510-116174022	ENSACAG000000042550	SPAG16	7	7	4	6
1	116420144-116492139	ENSACAG000000015438	BARD1	4	2	1	4
1	116556912-116732080	ENSACAG000000015468	ABCA12	31	33	29	36
1	116783469-116837241	ENSACAG000000015527	ATIC	9	10	3	8
1	116843550-116944630	ENSACAG000000015637	FN1	25	24	18	29
1	117149424-117196652	ENSACAG000000015921	MREG	4	2	2	4
1	117399045-117547961	ENSACAG000000015939	LRRFIP1	1	1	1	1
1	119408228-119789786	ENSACAG000000010229	AGAP1	1	1	1	1
1	124499114-124520464	ENSACAG000000001866	LOC103279484	1	1	2	2
1	126980620-127122615	ENSACAG000000000663	PCSK2	2	2	1	2
2	28469915-28548128	ENSACAG000000008734	PDE4A	3	1	1	3
2	28997144-29017162	ENSACAG000000009140	LOC100567117	5	4	1	5
2	29078684-29117860	ENSACAG000000009232	ZSWIM4	9	5	3	6
2	76778852-76819589	ENSACAG000000015657	IL12RB2	1	1	2	1
2	77827500-77878712	ENSACAG000000017092	LOC100567765	2	1	3	2
2	78813223-78867986	ENSACAG000000016939	PRKCSH	1	1	1	2
2	79330174-79462728	ENSACAG000000016655	DOCK6	5	8	6	5
2	79628928-79656337	ENSACAG000000016589	LDLR	29	20	13	6
2	80010007-80078980	ENSACAG000000016134	PNPLA6	6	2	2	2
2	81152464-81247292	ENSACAG000000015126	PZP	4	3	7	5
4	90376544-90468847	ENSACAG000000009917	PIKFYVE	26	27	15	25
4	90486024-90513339	ENSACAG000000009873	IDH1	3	2	1	3
GL343208.1	3302944-3443557	ENSACAG000000004721	NBEAL1	7	6	2	5
GL343282.1	19405-32643	ENSACAG000000009091	ALDH2	3	2	1	4
GL343343.1	240591-330691	ENSACAG000000009797	PLEKHM3	1	1	1	5
GL343645.1	51343-136417	ENSACAG000000009409	PCNT	1	1	1	1

Table S9: Candidate Urban-morphology Genes. We identified 93 candidate urban morphological genes (genes containing outliers identified by the urbanization GEA as well as both the front and rear trait elements for at least one of the three traits). Chromosome number or scaffold ID (CHROM), basepair range of the gene (BP RANGE), Ensembl Gene ID, Ensembl Gene Name, and number of outlier SNPs from each analysis per gene (URB: GEA analysis of urbanization; and GWAS analyses: HL: hindlimb length, FL: forelimb length, RTP: rear toepad area, FTP: front toepad area, RLAM: rear lamella number, FLAM: front lamella number). Red gene names indicate orthologous genes.

CHR	BP RANGE	Gene ID	Gene	URB	HL	FL	RTP	FTP	RLAM	FLAM
1	3335193-3403703	ENSACAG00000006398	PRKCI	1	0	0	1	2	1	1
1	43599498-43616646	ENSACAG00000003234	THAP9	1	0	1	1	0	4	1
1	83128470-83269448	ENSACAG00000005053	SEMASB	1	0	0	2	2	0	0
1	116556912-116732080	ENSACAG000000015468	ABCA12	36	0	0	1	1	1	0
1	116843550-116944630	ENSACAG000000015637	FN1	29	0	0	1	3	0	1
1	120462472-120516154	ENSACAG000000006033	VP54	3	0	0	1	1	1	0
1	122896101-122921363	ENSACAG000000005446	MGH3	4	1	1	0	1	0	1
1	126820405-126830700	ENSACAG000000008824	PROKR1	4	2	2	0	0	0	0
1	136748188-136960833	ENSACAG000000014770	LRRN4	1	1	1	0	0	0	0
1	140105345-140139690	ENSACAG000000003314								
1	201542946-201667059	ENSACAG000000011955	FAM124A	1	1	1	1	0	0	0
2	23000209-23278514	ENSACAG00000005078	COL12A1	4	0	0	1	0	7	6
2	29078684-29117860	ENSACAG000000013431	ADGRV1	1	0	0	1	2	1	1
2	66873290-66891936	ENSACAG000000009232	ZSVIM4	6	1	0	2	1	0	0
2	75535339-75691925	ENSACAG000000013105	CYP27B1	1	0	0	1	1	0	0
2	79718148-79790752	ENSACAG000000014282	ABCA1	1	2	1	1	0	0	1
2	79978742-79997835	ENSACAG000000016589	LDLR	6	1	3	0	0	4	0
2	80010007-80078980	ENSACAG000000016567	RASAL3	1	0	1	0	0	1	1
2	81432089-81698195	ENSACAG000000016313	MCOLN1	1	1	1	0	0	1	0
2	87950355-88105843	ENSACAG000000016134	PNPLA6	1	0	0	0	3	0	1
2	103323245-103369334	ENSACAG000000017477	novel gene	1	1	1	0	0	0	2
2	105955241-106078933	ENSACAG000000011487	L1CAM	1	0	0	2	1	0	0
2	163591805-163688957	ENSACAG000000017270	BRD4	1	0	0	0	1	3	1
2	187094323-187121094	ENSACAG000000017636	ARHGAP44	1	0	0	0	0	1	1
3	25697211-25825544	ENSACAG00000002031	FLNB	1	5	1	2	1	1	0
3	27860158-27929176	ENSACAG000000014113	FAM107A	1	0	0	0	0	1	1
3	48040000-48546399	ENSACAG000000004981	PAX7	2	1	1	0	0	0	0
3	87412874-87495042	ENSACAG000000002884	PIK3CB	2	0	0	0	0	1	1
3	95605387-95769701	ENSACAG000000028908	CDH23	2	0	1	1	1	0	0
3	138003141-138047548	ENSACAG000000024802	TBC1D4	1	1	2	0	0	0	0
3	181727196-181753348	ENSACAG000000012756	NUP210	1	0	0	1	1	1	0
3	187423022-187445584	ENSACAG000000011590	REP52	1	0	0	0	0	2	1
3	663978-807616	ENSACAG000000003165	UNCODL1	3	0	0	0	0	3	2
4	38209186-38261159	ENSACAG000000004319	SPART	1	0	0	5	5	1	0
4	76942903-77076711	ENSACAG000000002246	CRYL1	1	0	0	0	0	2	1
4		ENSACAG000000011144	TG	1	0	1	2	1	0	1
4		ENSACAG000000000201	NPC1	1	0	0	0	0	1	1
4		ENSACAG000000038653	novel gene	1	1	1	2	1	4	1
4		ENSACAG000000004665								
4	85612001-85635151	ENSACAG000000001823	KIF23	1	4	1	0	0	0	0
4	88728141-88858955	ENSACAG000000005200	ABCA4	4	0	1	2	1	1	1
4	123042437-123184648	ENSACAG000000016163	ITPR3	1	1	1	0	1	0	0
4	126156825-126169373	ENSACAG000000005931	IRF6	1	1	1	0	0	0	0
4	133593968-133618008	ENSACAG000000005019	LIMOD1	3	1	1	4	3	0	0
4	141903703-142470323	ENSACAG000000013632	PTPRH	4	0	0	5	4	0	2
4	145057504-145120505	ENSACAG000000016540	ZMYND8	2	0	0	0	0	2	1
5	447399-516161	ENSACAG000000005073	PRKCC	3	0	0	1	1	0	0
5	945998-966444	ENSACAG000000005702	LOC100568046	1	0	0	0	0	2	1
5	11289221-11322225	ENSACAG000000000039	DDX11	1	1	1	0	0	0	0
5	30977072-30992532	ENSACAG000000012771	NR2C1	3	0	0	1	1	0	0
5	37193204-37285367	ENSACAG000000014205	OTOG	1	1	3	0	1	0	0
5	64040556-64076520	ENSACAG000000002182	TUBGCP6	1	1	1	0	0	0	0
6	2963444-3125588	ENSACAG000000006816	novel gene	2	5	4	3	6	6	5
6	932237-9322560	ENSACAG000000008018	PFKP, PITRM1	9	0	0	0	0	6	3
6		ENSACAG000000029008								
6	44616494-44708953	ENSACAG000000002542	ITPRD1	2	0	0	2	2	0	0
6	58867960-58872334	ENSACAG000000009218	POH1	1	2	3	0	0	3	2
6	58912473-58940401	ENSACAG000000009079	INTS10	1	0	0	0	0	1	1
6	65623132-65627641	ENSACAG000000016690	WNT9B	1	0	0	1	2	0	1
6	72701257-72759015	ENSACAG000000017953	LOC100555613	1	3	2	1	0	0	0
6	74419348-74477010	ENSACAG000000007044	TBX21	1	0	0	1	1	0	0
GL343203.1	2588388-2892130	ENSACAG000000002067	HMCN1	3	4	2	0	0	2	5
GL343203.1	3318791-3365752	ENSACAG000000002460	NIBAN1	6	0	0	1	1	0	0
GL343208.1	1342864-2038258	ENSACAG000000004428	PARD3B	5	0	0	0	1	0	0
GL343208.1	3302944-3443557	ENSACAG000000004721	NBEAL1	5	0	0	1	1	0	0
GL343212.1	2647035-2675158	ENSACAG000000009638	NAB2	1	0	2	6	1	0	1
GL343212.1	2853070-2891851	ENSACAG000000009385	TNS2	1	0	0	2	2	0	1
GL343220.1	241427-249490	ENSACAG000000017518	novel gene	1	1	1	1	1	0	0
GL343220.1	946844-947821	ENSACAG000000008699	novel gene	1	0	0	1	1	2	1
GL343231.1	1506676-1673061	ENSACAG000000014758	THADA	1	1	3	1	0	0	0
GL343233.1	2164936-2331969	ENSACAG000000001757	SH2D4B	1	0	0	0	0	1	1
GL343238.1	2559-12147	ENSACAG000000011330	RETSAT	2	0	0	2	1	0	0
GL343252.1	191448-192470	ENSACAG000000021036	novel gene	1	1	4	0	0	2	0
GL343253.1	733594-735179	ENSACAG000000022303	novel gene	1	0	0	0	0	3	1
GL343279.1	786036-862059	ENSACAG000000015586	MYH14	1	1	2	1	0	0	0
GL343282.1	533744-611760	ENSACAG0000000042073	novel gene	3	0	1	5	6	0	0
GL343297.1	672977-673986	ENSACAG000000011623	LRRIC18	1	2	1	1	1	0	0
GL343325.1	929627-959076	ENSACAG000000000647	AP5M1	1	1	0	1	1	0	0
GL343326.1	73295-131733	ENSACAG000000011574	POLQ	1	0	0	1	0	2	2
GL343362.1	544682-549131	ENSACAG000000017042	KLHL23	1	1	1	1	0	1	0
GL343379.1	43592-206094	ENSACAG000000010585	BAC2B	1	0	0	0	0	1	2
GL343391.1	491033-632328	ENSACAG000000007222	EPHA1	1	1	5	2	1	0	0
GL343392.1	395779-479391	ENSACAG000000004741	SV2B	4	1	1	1	0	0	1
GL343471.1	197754-225168	ENSACAG000000010023	novel gene	1	0	0	1	1	0	0
GL343482.1	544756-556377	ENSACAG000000005955	MYORG	1	1	0	1	1	0	0
GL343491.1	580892-596715	ENSACAG000000001287	MMPI4	1	2	1	0	1	1	0
GL343500.1	60011-65863	ENSACAG000000001421	RAI1	1	1	1	1	0	1	5
GL343704.1	87040-112621	ENSACAG000000001793	FDT1	1	1	1	0	1	0	0
GL343722.1	27216-40256	ENSACAG000000002300	NLRX1	1	0	0	0	1	1	1
GL343737.1	4180-12257	ENSACAG000000001916	ZC3H15	1	0	0	1	1	0	0
GL343740.1	65377-197640	ENSACAG000000010989	ADAMTS13	3	2	1	0	1	3	1
GL343747.1	103538-154118	ENSACAG000000003845	TRPM7	1	1	2	1	0	3	1
GL343760.1	84568-93753	ENSACAG000000013878	TAS1R3	1	0	3	1	1	0	0
GL343806.1	57854-171679	ENSACAG000000005908	FBN1	1	1	0	0	1	0	0
GL343981.1	57201-70918	ENSACAG000000016435	HIV	1	0	0	2	1	0	0
GL344241.1	17761-44408	ENSACAG000000017776	NCAN	1	0	0	1	1	1	0

Table S10: GO Analysis. Functional enrichment of morphology-associated genes with associated p-value and data source (Gene Ontology, Human Phenotype Ontology). Terms related to locomotor morphology are highlighted.

Term Name	p-value	Source	Term Name	p-value	Source
cell periphery	0.0004	GO:CC	Abnormality of facial musculature	0.0165	HP
movement of cell or subcellular component	0.0005	GO:BP	Abnormality of the skeletal system	0.0165	HP
locomotion	0.0007	GO:BP	Hepatic failure	0.0165	HP
cell migration	0.0023	GO:BP	Abnormality of the musculature of the lower limbs	0.0165	HP
localization of cell	0.0035	GO:BP	Abnormal abdomen morphology	0.0165	HP
localization	0.0035	GO:BP	Behavioral abnormality	0.0171	HP
cell motility	0.0035	GO:BP	Abnormality of movement	0.0171	HP
anatomical structure morphogenesis	0.0035	GO:BP	Pediatric onset	0.0174	HP
Abnormality of the face	0.0036	HP	Abnormal blood monovalent inorganic cation concentration	0.019	HP
HP root	0.0036	HP	Diabetes mellitus	0.0199	HP
Abnormality of the nervous system	0.0036	HP	Abnormality of the gastrointestinal tract	0.023	HP
Phenotypic abnormality	0.0036	HP	Hyperlordosis	0.023	HP
Abnormality of the musculoskeletal system	0.0036	HP	Abnormality of the musculature	0.023	HP
Clinical modifier	0.004	HP	Glucose intolerance	0.023	HP
Dermatological manifestations of systemic disorders	0.004	HP	Abnormality of the eye	0.0234	HP
Abnormal nervous system physiology	0.004	HP	Abnormal retinal morphology	0.0234	HP
Abnormality of head or neck	0.004	HP	Abnormal cranial nerve morphology	0.0247	HP
Abnormality of the head	0.0044	HP	Cholelithiasis	0.0247	HP
Onset	0.0044	HP	Abnormal glucose homeostasis	0.0247	HP
Abnormality of the endocrine system	0.0048	HP	Limb-girdle muscle weakness	0.0248	HP
neurogenesis	0.0048	GO:BP	Abnormal cry	0.0255	HP
generation of neurons	0.0052	GO:BP	Abnormality of the kidney	0.0255	HP
Proximal muscle weakness	0.0065	HP	Lower limb spasticity	0.0255	HP
Abnormal circulating metabolite concentration	0.0065	HP	multicellular organism development	0.0256	GO:BP
Abnormality of the genitourinary system	0.0065	HP	Decreased liver function	0.0267	HP
Functional motor deficit	0.0099	HP	Abnormality of hepatobiliary system physiology	0.0282	HP
basement membrane	0.0101	GO:CC	EMG: myopathic abnormalities	0.0282	HP
plasma membrane	0.0101	GO:CC	neuron differentiation	0.0294	GO:BP
anatomical structure development	0.0101	GO:BP	Ophthalmoparesis	0.03	HP
Abnormality of the urinary system	0.011	HP	Neurodevelopmental abnormality	0.0301	HP
Abnormality of metabolism/homeostasis	0.0113	HP	Abnormality of the biliary system	0.0306	HP
Abnormality of facial soft tissue	0.0115	HP	Weakness of facial musculature	0.031	HP
Abnormal cranial nerve physiology	0.0115	HP	Intellectual disability	0.0312	HP
Abnormality of the seventh cranial nerve	0.0115	HP	cell morphogenesis	0.0342	GO:BP
Mode of inheritance	0.0118	HP	Abnormal choroid morphology	0.0354	HP
Abnormality of higher mental function	0.0118	HP	Pain	0.0361	HP
multicellular organismal process	0.0125	GO:BP	cell-cell junction	0.0363	GO:CC
Abnormal posterior eye segment morphology	0.0139	HP	Abnormal eye physiology	0.0371	HP
Abnormality of the gallbladder	0.0139	HP	Abnormality of thyroid physiology	0.0382	HP
Gait disturbance	0.0139	HP	Abnormality of the upper urinary tract	0.0394	HP
Cranial nerve paralysis	0.0139	HP	EMG abnormality	0.0398	HP
Abnormal fundus morphology	0.0139	HP	Abnormality of peripheral nerves	0.0398	HP
Abnormality of the musculature of the limbs	0.0139	HP	Constitutional symptom	0.0401	HP
Facial palsy	0.0139	HP	Waddling gait	0.0431	HP
Abnormality of the digestive system	0.0139	HP	Wide nasal bridge	0.0431	HP
Abnormality of cardiovascular system morphology	0.0139	HP	regulation of signal transduction	0.0451	GO:BP
Abnormal circulating protein concentration	0.0139	HP	Autosomal dominant inheritance	0.0452	HP
Clinical course	0.0147	HP	tissue development	0.0475	GO:BP
Abnormality of the thyroid gland	0.0149	HP	system development	0.0475	GO:BP
Growth abnormality	0.0149	HP	Bipolar affective disorder	0.0486	HP
Abnormal gallbladder physiology	0.0149	HP	Abnormality of the optic nerve	0.0492	HP
Abnormality of the peritoneum	0.0149	HP	Adult onset	0.0494	HP
Motor polyneuropathy	0.0149	HP	Abnormal nerve conduction velocity	0.0495	HP
Cholecystitis	0.0149	HP	Abnormal skeletal muscle morphology	0.0495	HP
Autosomal recessive inheritance	0.0152	HP	Hypothyroidism	0.0495	HP
Abnormal muscle physiology	0.0155	HP	Aplasia/Hypoplasia affecting the fundus	0.0495	HP
Jaundice	0.0155	HP	Spinal cord compression	0.0495	HP
Abnormality of the calf musculature	0.0155	HP	Hepatomegaly	0.0495	HP
nervous system development	0.0155	GO:BP	Abnormality of the skin	0.0499	HP
Abnormality of skeletal morphology	0.0165	HP			
Elevated hepatic transaminase	0.0165	HP			
Abnormal cervical spine morphology	0.0165	HP			
Limb muscle weakness	0.0165	HP			
Generalized abnormality of skin	0.0165	HP			
Abnormal nervous system morphology	0.0165	HP			
Abnormal enzyme/coenzyme activity	0.0165	HP			
Abnormality of the cardiovascular system	0.0165	HP			

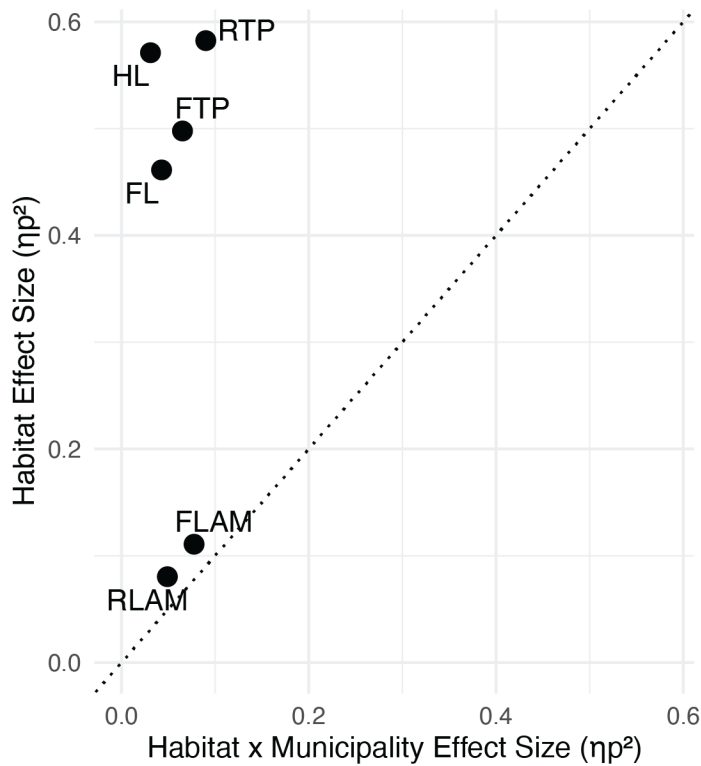
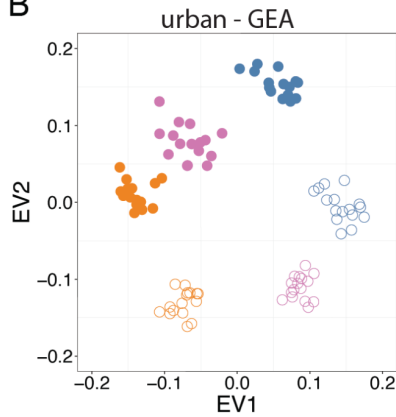


Fig. S11: Phenotypic Parallelism. Parallel morphological shifts in data set analyzed here, $n=16$ per population, for six traits: hindlimb and forelimb lengths (HL, FL), front and rear toepad areas (FTP, RTP), and front and rear number of lamellae (FLAM, RLAM). Effect size (partial eta, ηp^2) for each trait for the habitat effect (urban v forest) and interaction effect of habitat x municipality. Traits were natural-log transformed and natural-log transformed body size (snout-vent length) was included as a covariate in models for limb length and toepad areas. Dotted line is 1:1; points above indicate consistent effects of urbanization across municipalities.

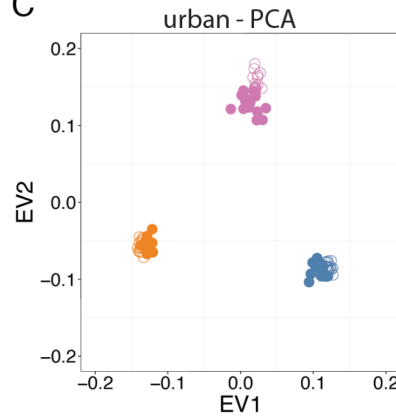
A

trait	habitat	municipality	interaction (habitat*municipality)
background	$F(df=1,90) = 1.46, p = 0.230; \eta^2 = 0.02$	$F(df=2,90) = 8738.20, p < \mathbf{0.001}; \eta^2 = 0.99$	$F(df=2,90) = 7.43, p = \mathbf{0.001}; \eta^2 = 0.14$
urban - GEA	$F(df=1,90) = 705.23, p < \mathbf{0.001}; \eta^2 = 0.89$	$F(df=2,90) = 830.49, p < \mathbf{0.001}; \eta^2 = 0.95$	$F(df=2,90) = 39.36, p < \mathbf{0.001}; \eta^2 = 0.47$
urban - PCA	$F(df=1,90) = 6.31, p = \mathbf{0.014}; \eta^2 = 0.07$	$F(df=2,90) = 10912.40, p < \mathbf{0.001}; \eta^2 = 1.00$	$F(df=2,90) = 20.51, p < \mathbf{0.001}; \eta^2 = 0.31$
hindlimb	$F(df=1,90) = 59.31, p < \mathbf{0.001}; \eta^2 = 0.40$	$F(df=2,90) = 13.57, p < \mathbf{0.001}; \eta^2 = 0.23$	$F(df=2,90) = 4.79, p = \mathbf{0.011}; \eta^2 = 0.10$
forelimb	$F(df=1,90) = 88.39, p < \mathbf{0.001}; \eta^2 = 0.50$	$F(df=2,90) = 90.21, p < \mathbf{0.001}; \eta^2 = 0.67$	$F(df=2,90) = 3.28, p = \mathbf{0.042}; \eta^2 = 0.07$
rear lamellae	$F(df=1,90) = 44.98, p < \mathbf{0.001}; \eta^2 = 0.33$	$F(df=2,90) = 15.42, p < \mathbf{0.001}; \eta^2 = 0.26$	$F(df=2,90) = 4.53, p = \mathbf{0.013}; \eta^2 = 0.09$
front lamellae	$F(df=1,90) = 100.38, p < \mathbf{0.001}; \eta^2 = 0.53$	$F(df=2,90) = 18.47, p < \mathbf{0.001}; \eta^2 = 0.29$	$F(df=2,90) = 11.29, p < \mathbf{0.001}; \eta^2 = 0.20$
rear toepad area	$F(df=1,90) = 151.05, p < \mathbf{0.001}; \eta^2 = 0.63$	$F(df=2,90) = 17.66, p < \mathbf{0.001}; \eta^2 = 0.28$	$F(df=2,90) = 2.19, p = 0.112; \eta^2 = 0.05$
front toepad area	$F(df=1,90) = 52.84, p < \mathbf{0.001}; \eta^2 = 0.37$	$F(df=2,90) = 60.26, p < \mathbf{0.001}; \eta^2 = 0.57$	$F(df=2,90) = 2.95, p = 0.057; \eta^2 = 0.06$

B



C



D

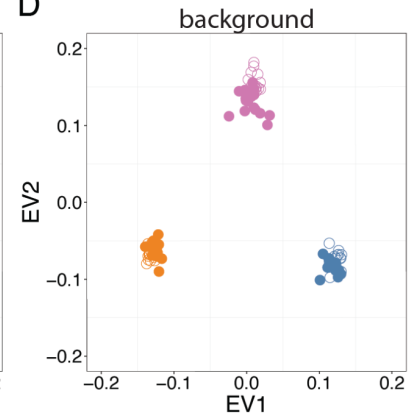


Fig. S12: Genomic Parallelism. (A) We performed local PCA of outlier SNPs (outliers in each analysis) and background SNPs (not outliers in either of the urbanization association tests - PCA and GEA) as a test of polygenic parallelism. We then conducted ANOVA for the first axis of genetic variation in each (eigenvector 1) with the following formula: $EV1 \sim \text{habitat} * \text{municipality}$, and evaluated the effect size of habitat versus the interaction effect. Results are for each ANOVA, with partial eta-squared (effect size) reported for each term. Significant p-values ($p < 0.05$) are bolded. (B) Parallelism of genomic architecture of urban-associated SNPs. Local PCA of outlier SNPs for the two tests for urban-association: (B) urban GEA, (C) PCA by municipality, and (D) background SNPs (non-outliers in either test).

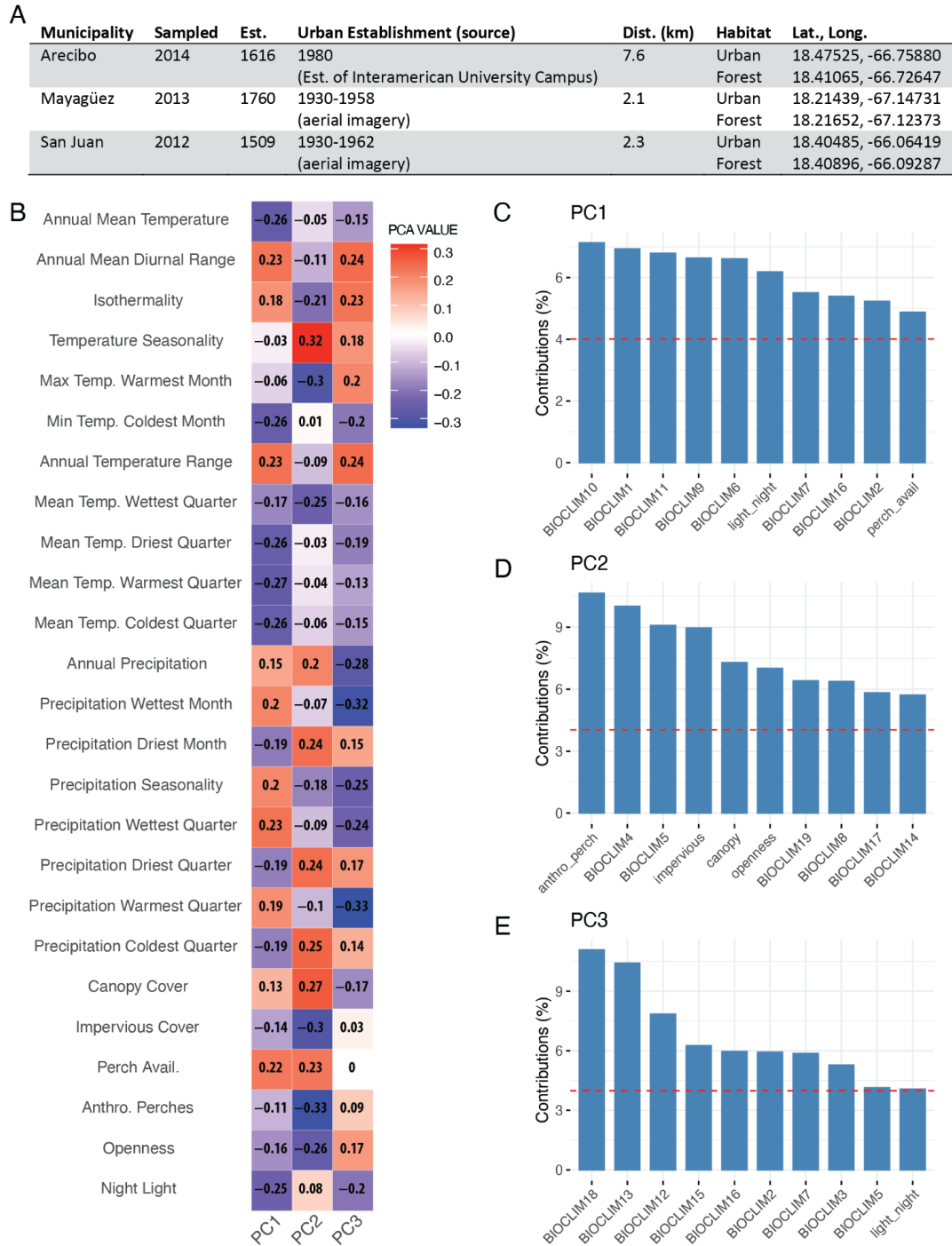


Fig. S13: Urban Environment Quantification. (A) Paired urban and forest sample sites in Puerto Rico from three municipalities, with sampling year, founding date of the municipality (Est.), estimated date urban habitat was established with source of estimate, distance between urban and forest pairs within each municipality, habitat type, and latitude and longitude of each site. (B) Correlation plot of environmental PCA loadings; cells shaded by the loading value, with negative loadings darker blue and positive loadings brighter red. Top 10 contributions to (C) PC1, (D) PC2, and (E) PC3 with average expected contribution indicated by dashed red line.