Leveraging publicly available data to facilitate urban ecology and evolution

Shawn Arreguin[1], Joseph F. Walker[1*] Natalie L.R. Love[2*]

[1]Department of Biological Sciences, the University of Illinois at Chicago, Illinois, 60607, USA
[2]Negaunee Institute for Plant Conservation Science and Action, Chicago Botanic Garden, Glencoe, Illinois 60022 USA
Authors for correspondence: nlove@chicagobotanic.org and jfw52@uic.edu
*Denotes equal contribution
ORCID IDs:
Shawn Arreguin: https://orcid.org/0009-0003-9302-9979
Joseph F. Walker: https://orcid.org/0000-0003-2928-8899
Natalie L.R. Love: https://orcid.org/0000-0002-5013-5478

Abstract

The intensifying pace of urbanization has prompted researchers to investigate its ecological and evolutionary consequences more deeply, yet logistic and monetary challenges can impose a barrier to research. Publicly available databases offer scientists from broad backgrounds an opportunity to circumvent those barriers, enhancing participation in urban ecology and evolution. In conjunction with computational advances, these datasets allow research to be done *in silico* - entirely with a computer connected to the internet. Throughout this review, we highlight how *in silico* approaches allow researchers to test their own hypotheses, thereby radically advancing our understanding of evolutionary and ecological processes that shape life in cities. By collating presently used databases, covering emerging tools and resources, and presenting biases and limitations of these data, we hope that this review will act as a guide for interested researchers._Opportunities to take part in this exciting field of research will continue to grow as new databases and associated tools are developed.

Keywords: Open Source; In silico; Urbanization; Databases; Inclusive Research

1) Introduction

The development of large scale community science databases, the emerging emphasis on open data initiatives, and advances in computational capacity have ushered in a new era of research. Increasingly, scientific inquiry can be conducted *in silico* - entirely with a computer connected to the internet. This approach leverages existing, publicly available data without the need for original data collection. *In silico* projects remove cost barriers imposed by generating new data, making them an excellent resource for both student learning and obtaining preliminary data. This facilitates broad accessibility in academic settings and allows standalone big-data projects to be conducted cost-effectively. Using novel and creative approaches, *in silico* research has already begun to yield new and unique insights into ecological and evolutionary processes.

Public repositories of biological data are diverse, data-rich and importantly, freely accessible online. They span a wide range of taxonomic, ecological, and spatial scales, as well as containing many data types. For example, repositories can comprise community science occurrence (e.g., iNaturalist, eBird), remotely sensed (e.g., Landsat and Sentinel satellite missions), genetic (e.g., Dryad, MacroPopGen, GenBank), trait (e.g., TRY, Aminote, EltonTrait), and open access data from published articles, among others. These datasets are an increasingly valuable resource, especially when facing uncertain funding climates. However, datasets are often scattered across disparate internet repositories, and researchers may need to synthesize data found across several sources to conduct a study. Thus, there is a need to summarize and collate resources as well as provide examples of existing studies to facilitate future *in silico* research.

Publicly available community science generated data tends to be clustered in urban areas (Di Cecco et al., 2021). As a result, this data is particularly useful for conducting studies of ecological and evolutionary processes in urban landscapes. Urbanization causes widespread land cover and environmental change (Grimm et al., 2008), resulting in areas characterized by several interrelated factors, including increased temperatures (e.g., urban heat islands), high levels of pollution, fragmented habitats, and habitat loss. These changes create novel environments, inducing strong selective pressures that shape evolutionary trajectories and ecological interactions. For example, urbanization can drive rapid adaptive evolution (Cheptou et al., 2008; Nacci et al., 2010; Winchell, Campbell-Staton, et al., 2023) or shifts in phenology (D. Li et al., 2021; Merckx et al., 2021; Olsen et al., 2020), creating downstream effects on ecosystem-level processes. While ubiquitous globally, urbanization is not uniform, and its impacts on biological processes may differ spatially, temporally, and taxonomically. The pace of urbanization is rapidly expanding, with urban land area predicted to more than quadruple by 2050 from levels in 2000 (Angel et al., 2011). Given these patterns of urbanization, it is critical to assess these effects at broad spatial scales and across levels of biological organization, thereby clarifying general patterns and improving the predictability of responses to urbanization.

Traditional fieldwork in urban landscapes presents several challenges. Privately owned land, theft or vandalism of field equipment (Stadnicki et al., 2024), and high urban heterogeneity (Cadenasso et al., 2007) can impede gathering enough data to make city-wide generalizations. These factors are further exacerbated when investigating trends over multiple cities (Magle et al., 2019). Repositories, such as iNaturalist, contain biological data collected across diverse land ownership and land use classes in urban areas (Di Cecco et al., 2021). Many datasets are also available at broad spatial and temporal scales, facilitating multi-city comparisons and investigations into change through time (Shultz et al., 2020). When combined with freely accessible predictor variables (e.g., climate, land cover, and sociodemographics), publicly available biological data provide a powerful platform to readily overcome existing challenges and conduct hypothesis-driven research entirely *in silico*.

This paper aims to provide a thorough guide for conducting urban ecology and evolution (UEE) research *in silico*. This is done by: (1) summarizing the existing data repositories currently available and highlighting the research questions that have been addressed by innovative approaches that leverage these data across multiple levels of biological organization (Table 1, Table 2, Figure 1), (2) highlighting emerging resources and computational tools that will maximize the utility of publicly available data for research *in silico,* and (3) discussing biases and limitations in these data and demonstrating how this data can be used to enhance participation in UEE research. By outlining both resources and example studies, this paper aims to provide a roadmap for future researchers addressing ecological and evolutionary dynamics in cities *in silico*.

2) Data sources and examples of novel studies leveraging publicly available data

In this section, we summarize publicly available data sources and highlight specific studies that have leveraged these resources to answer novel questions in urban landscapes, often leveraging urbananization gradients (Figure 1). Importantly, we focus on research that has been conducted mostly or entirely *in silico*. For clarity, this section is arranged by biological level of organization. This is not meant to be a comprehensive review; rather, the studies chosen highlight innovative approaches and emphasize the utility of these data for advancing the field of UEE. Our review focuses on studies that cover all biological levels and are entirely reliant on publicly available datasets. When few qualifying studies were found, we supplemented subsections with hybrid approaches.

To facilitate future *in silico* research, we present two summary tables to assist in study design. **Table 1** summarizes research questions that have been addressed and the sources of biological data used. **Table 2** compiles commonly used non-biological data that often serve as covariates or predictor variables, such as measures of urban intensity (e.g., percent of a surface that is impervious). Together, these resources are used to illustrate the depth and breadth of questions and data that can be synthesized for UEE studies.

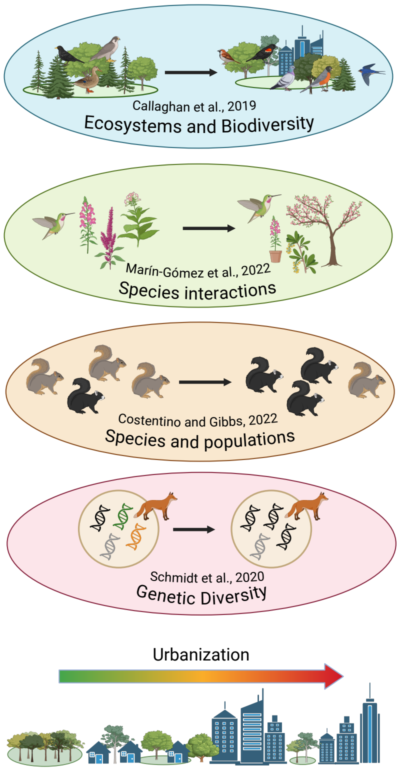

Figure 1. From genes to ecosystems, examples of publicly available datasets that have enabled urban ecology and evolution studies to be conducted at multiple levels of biological organization. Here, we illustrate examples of studies that have assessed the effects of urbanization using publicly available datasets. The schematic depicts the dominant pattern that emerges with increasing urbanization. The corresponding research questions addressed by these studies can be found in **Table 1**.

Biological Level	**Subject**	**Questions**	**Biological Data Sources**	**Examples**
Ecosystems and Biodiversity	Community assemblage and biodiversity distribution	• How does species diversity vary across urbanization gradients?[1] • Are patterns of urban biodiversity related to socioeconomic or sociodemographic variables or historic redlining?[2] • Where are the biodiversity hotspots on urban landscapes?[3]	iNaturalist, Google Street View, eBird, Birds of the World, Norfolk Bat Survey, Global Biodiversity Information Facility (GBIF), Australian Department of the Environment	Beninde et al., 2023[3]; Border et al., 2017[1,3]; Callaghan, Bino, et al., 2019[1]; Estien et al., 2024[2]; Fitch, Wilson, et al., 2019[1]; Ives et al., 2016[3]; Kinnunen et al., 2024[1,2]; Leong and Trautwein, 2019[2]; E. Li et al., 2019[1,5]; Prudic et al., 2018[4]; RužanoviĆ and StankoviĆ, 2023[1]; Velasquez-Camacho et al., 2024[3]
	Urban Tolerance	• Does urbanization result in biotic homogenization of species composition?[1] • Does urbanization act as a filter to select for tolerant species?[2] • Is there phylogenetic signal in urban tolerance?[3] • Which species traits confer urban tolerance?[4] • Is urban tolerance generalizable across cities?[5] • **How does urban tolerance vary by taxonomic group?** • **Does urbanization filter for reproductive strategies in plants?**	Amniote Life-History database, AnAge, EtlonTrait, PanTHERIA, Previously published phylogeny, eBird, GBIF, iNaturalist, AVONET, BirdBase, Birds of the World, Bird Life	Hensley et al., 2019[1,2,4,5]; Winchell et al., 2020[3,4]; Callaghan, Major, et al., 2019[2,4]; Neate-Clegg et al., 2023[2,3,4]; Kinnunen et al., 2022[1,3,5]; Santini et al., 2019[3,4]; Curti et al., 2024[5]
	Ecosystem function	• How does urbanization interact with other stressors (e.g., drought) to affect urban vegetation health and resilience?[1] • Are patterns of vegetation health and resilience related to socioeconomic or sociodemographic variables?[2]	NASA (National Aeronautics and Space Administration), ESA (European Space Agency), Copernicus	Dong et al., 2023[1,2]; Leisenheimer et al., 2024[1];
	Phenological Change	• How does urbanization affect the timing and duration of community-level phenological events?[1] • Does urbanization result in phenological shifts?[2] • **Does tree species diversity explain growing season length in urban forests?** • **Do phenological shifts affect species-interactions in urban areas?**	iNaturalist, pan-European Phenology Project (PEP), national Phenology Network (NPN), GBIF, Previously published phylogeny, Open Tree of Life	D. Li et al., 2021[1]; Owen et al., 2024[2]; Wohlfahrt et al., 2019[1,2]; Merckx et al., 2021[2]; Li et al., 2017[1]
Species Interactions	Plant-animal and animal-animal interactions	• How does urbanization affect plant-pollinator networks?[1] • How does urbanization affect plant-herbivore interactions?[2] • How does urbanization affect predation?[3] • Do interactions change through time?[4] • **Do urban gardens provide pollinator resources to native insects?**	iNaturalist, eBird, GBIF	Marín-Gómez et al., 2022[1]; Meineke et al., 2019[2,4]; Putman et al., 2021[3]
Species and Populations	Adaptations	• Is there evidence of heat stress adaptation among urban populations?[1] • How does urbanization affect body size?[2] • How does urbanization affect color patterning?[3]	iNaturalist, VertNet, National Ecological Observatory Network (NEON), North American Census of Small Mammals (NACSM)	Cosentino and Gibbs, 2022[3]; Fukano et al., 2023[1,3]; Hantak et al., 2021[2,4.]
	Plastic responses	• Do traits respond plastically to urbanization?[1] • **Do traits exhibit a greater degree of plasticity in urban environments?**	iNaturalist, TRY, Nocturna, Swedish and Finnish Butterfly Monitoring Scheme, Finnish Biodiversity Info Facility, BugGuide	Batool et al., 2024[1]; Ibsen et al., 2023[1]; Merckx et al., 2021[1]
	Genetic Diversity	• How does urbanization affect population genetic parameters like genetic diversity, N_e, F_{ST}, allelic richness, etc.)[1] • Have population genetic parameters changed through time in urban areas?[2] • Is genetic diversity related to socioeconomic or sociodemographic variables?[3] • **How does urbanization affect inbreeding in wild populations?**	Dryad, GenBank, DataOne, BOLD, Sequence Read Archive	Millette et al., 2020[1,2]; Miraldo et al., 2016[1]; Schmidt et al., 2020, 2021[1]; Schmidt and Garroway, 2022[3]; Habrich et al., 2021[1]; Karachaliou et al., 2025[1]

Table 1. This table summarizes the types of research questions that have been addressed across various levels of biological organization and provides representative examples. Questions in bold deserve greater exploration using publicly available datasets. It also summarizes the resources used to obtain the biological data analyzed in the studies. Often, these data serve as response variables in models evaluating the effect of urbanization on ecological and evolutionary processes.

Predictor Variable Category	**Sources**	**Examples**
Urbanization	Ocean Health Index Cumulative Human Impacts[1], National Land Cover Database Impervious Surface (NLCD)[2], Global Man-made Impervious Surface (GMIS)[3], 2010 Global Tree Cover Dataset[4], VIIRS Nighttime Light[5], Human Population Density[6], Human Footprint Index[7], Anthropogenic biomes or Anthromes[8], History database of the Global Environment (HYDE)[9], National Park Service Noise Pollution[10], US National Transportation Noise Map[11], Copernicus Land Monitoring Service (CLMS) impervious surface[12], European Settlement Map[13], Gridded Population of the World[14], Defense Meteorological Satellite Program Operational Linescan System[15]; OpenPopGrid[16], Australian Bureau of Statistics[17], OpenDataSoft[18]; World Urban Areas[19]	Batool et al., 2024[6]; Callaghan, Major, et al., 2019[5]; Cosentino and Gibbs, 2022[3,4]; Curti et al., 2024[2,5,10]; Estien et al., 2024[2]; Fitch, Wilson, et al., 2019[2]; Hantak et al., 2021[6]; Karachaliou et al., 2025; Leong and Trautwein, 2019[2]; D. Li et al., 2021[6]; E. Li et al., 2019[2, 6, 11]; Meineke et al., 2019[6]; Millette et al., 2020[9]; Miraldo et al., 2016[8]; Owen et al., 2024[2]; Putman et al., 2021[2]; Schmidt et al., 2020, 2021[7]; Schmidt and Garroway, 2022[7]; Wohlfahrt et al., 2019[12,13]; Habrich et al., 2021[14]; Li et al., 2017[15], Border et al., 2017[16], Neate-Clegg et al., 2023[5,14,18], Santini et al., 2019[19]
Climate (current and projected)	WorldClim[1], CHELSA[2], ClimateNA[3], PRISM[4], NASA NEX-DCP30[5], NASA NEX-GDDP[6], Basin Climate Characterization Model (BCM)[7], Environment Canda[8], TerraClimate, Palmer Drought Severity Index[9]	Batool et al., 2024[8]; Beninde et al., 2023[1,2,3]; Cosentino and Gibbs, 2022[1]; Hantak et al., 2021[4]; D. Li et al., 2021[4,5,7]; Meineke et al., 2020[4]; Owen et al., 2024[1]; Winchell et al., 2020[1]; Dong et al., 2019[4,9]
Land Use and Land Cover	National Land Cover Database (NLCD)[1], MODIS[2], Landsat[3], CORINE Land Cover[4], USGS Hydrography Dataset[5]; Copernicus Land Monitoring Service (CLMS)[6], Global Roads Open Access Dataset (gROADS)[7], Highway Performance Monitoring System[8], Mapping Inequality database (historical HOLC maps)[9]	Callaghan, Major, et al., 2019[2,3]; Callaghan, Bino, et al., 2019[1]; Estien et al., 2025[9]; E. Li et al., 2019[1, 3, 5]; RužanoviĆ and StankoviĆ, 2023[4]; Wohlfahrt et al., 2019[1,6], Habrich et al., 2021[7,8], RužanoviĆ and StankoviĆ, 2023[4]
Socio-demographic and Socio-economic	United States Census Bureau[1], Statistics Canada[2], CalEnviroScreen[3], CEJST, IPUMS[4]	Batool et al., 2024[2]; Kinnunen et al., 2025, 2022[1]; E. Li et al., 2019[1,3]; Schmidt et al., 2021, 2020[1,2]; Schmidt and Garroway, 2022[1,2,4]

Table 2. This table summarizes the data sources for common predictor variables used in models assessing urban ecological and evolutionary processes. It also lists the specific studies that the data sources were used in. Studies may use more than one data source to capture different aspects of urbanization.

2.1) Ecosystems and biodiversity distribution

2.1.1) **In silico** *insights into community assemblage and biodiversity distribution*

Occurrence datasets (e.g., iNaturalist, eBird, GBIF) have been used to examine how biodiversity varies along urbanization gradients. With these datasets, researchers can test whether urbanization homogenizes species assemblages within cities and identify biodiversity hotspots in urban landscapes (**Table 1**). These studies span a wide range of taxonomic clades, including birds (Callaghan, Bino, et al., 2019; Hensley et al., 2019; Kinnunen et al., 2024), insects (Fitch, Wilson, et al., 2019; Prudic et al., 2018; Ružanović & Mičetić Stanković, 2023), mammals (Border et al., 2017; Santini et al., 2019), and plants (Ives et al., 2016; Velasquez-Camacho et al., 2024). Several studies have also compared patterns of biodiversity and community assembly across clades (Beninde et al., 2023; Estien et al., 2024; Leong & Trautwein, 2019; E. Li et al., 2019; Curti et al., 2024). This work has helped uncover the effects of urbanization on species diversity and richness, which may depend on various factors, such as local landscape structure or evolutionary history. (Callaghan, Bino, et al., 2019) leveraged eBird to obtain over four million occurrence records throughout the contiguous USA, finding that bird richness and diversity were highest in urban greenspaces relative to natural green areas or medium/high developed areas. They attributed this response to increased habitat heterogeneity within urban greenspaces relative to either natural or developed areas, suggesting habitat structure influences urban bird diversity. Fitch, Wilson, et al., (2019) repurposed data from previous studies (Fitch, Glaum, et al., 2019; Jamieson et al., 2019) to find that urbanization positively influences introduced bee species, but not native species in Michigan, USA. These findings provide insights into how origin may predict species responses to urbanization.

Recent work has revealed that historical housing policies in the USA, such as redlining, the systematic exclusion of minority neighborhoods from receiving the same loans as wealthier, often white neighborhoods, continue to influence ecological and evolutionary processes in cities (Schell et al., 2020). For example, Estien et al., (2024) integrated over 120,000 iNaturalist observations, digitized HOLC (home-owners loan corporation) historical maps, and data from the National Land Cover Database (NLCD) to assess whether species richness and community assembly of birds, insects, arachnids, reptiles, and amphibians differ among historical HOLC grades that range from grade A ("best" or "greenlined" neighborhoods) to D (hazardous or redlined). Greenlined neighborhoods had higher species richness than redlined neighborhoods, indicating that the legacy of discriminatory housing practices shapes contemporary patterns of urban biodiversity. The authors hypothesized that this pattern can be attributed to spatial differences in environmental quality caused by systematic disinvestment in redlined neighborhoods.

2.1.2) Assessing urban tolerance across cities

Urban environments impose novel conditions that constrain which species can survive in cities. Species are often divided between those that succeed in urban environments – urban exploiters or urban adapters – and those that are excluded and threatened by urban expansion – urban avoiders (Blair, 1996, 2001; Faeth et al., 2011). Urban adapters are hypothesized to possess exaptations that facilitate urban success or tolerance (Winchell et al., 2023).

In silico approaches have combined occurrence data (e.g., GBIF, iNaturalist, and eBird) with trait databases to test hypotheses about urban tolerance, such as the Functional Biotic Homogenization hypothesis (Olden et al., 2004). This hypothesizes that within taxonomic groups, species traits are predicted to converge in response to urban conditions. Such analyses can uncover which species are most tolerant, which traits are most beneficial, and whether urban tolerance patterns are city-specific or generalizable across urban areas.

Winchell et al. (2020) used a phylogenetic framework to understand the role evolutionary history plays in urban tolerance among Caribbean anoles (*Anolis* spp.). Using GBIF and iNaturalist occurrence data, they found strong phylogenetic signals of urban tolerance in the *Anolis* genus. Additionally, locomotor traits such as rear lamellae number and hindlimb length were positively associated with urban tolerance. Warm temperature specialists were more likely to be urban tolerant, indicating that the urban heat island effect may impose an ecological filter for urban species composition. Previously, Callaghan, Major, et al., (2019) found a similar pattern in birds. The authors classified eBird observations using a continuous urbanization index (**Table 2**). The results showed generalism in feeding and breeding habitat, clutch size, and residual brain and body size were important for urban bird tolerance. Neate-Clegg et al., (2023) also found that select functional traits are associated with urban tolerance on a global scale. Latitude and human population density modulated this association for some traits. For example, the relationship between urban tolerance and body mass changed with latitude.

City-specific tolerance has been found as well, especially when accounting for city-level characteristics, such as the socioeconomic status and structure of the built environment. Hensley et al., (2019), using previously published bird counts, and Kinnunen et al., (2022), using eBird observations, combined bird counts and species-level trait data (Amniote Life-History database, BirdLife, Birds of the World) to investigate patterns of city-specific tolerance. Both studies found that urban tolerance of bird species is influenced by inter-city variability, resulting in differential trait-based filtering. Traits that varied across cities include body size, clutch size, longevity, diet guild, habitat preference, and migratory status.

Our ability to understand urban filters will continue to increase as publicly available datasets grow. Moreover, *in silico* approaches can be used to compare cities around the globe, explaining how cultural, historical, and structural characteristics of cities influence species filters. Further research can also test whether these conclusions broadly extend to other taxonomic groups.

2.1.3) Understanding urban ecosystem function with publicly available data

Freely available remotely-sensed satellite data (e.g., Landsat, MODIS, Sentinel, etc.) allows researchers to study landscape-scale responses of ecosystems to urbanization (Cârlan et al., 2020; Dong et al., 2019; Khikmah et al., 2024; Leisenheimer et al., 2024). Satellite programs collect imagery with high spatial resolutions (e.g., 10-30 meters) and temporal frequencies (e.g., twice monthly), covering a global extent that can span decades. For example, Landsat, a satellite program developed by NASA in the USA, has been collecting 30 meter resolution imagery at a global scale for more than 50 years (Crawford et al., 2023). These satellites are capable of capturing multispectral imagery, such as near-infrared light (NIR). NIR is reflected by healthy, photosynthetically active vegetation, and can be used to calculate the normalized difference vegetation index (NDVI) by using red reflectance (see Huang et al., 2021 for a review and recommendations for using NDVI in scientific studies) among other vegetation indices. NDVI can be used to monitor vegetative responses, including plant stress. When combined with climate data, researchers can begin to understand how altered environmental conditions, like heat or drought, influence responses of urban vegetation, especially along sociodemographic and economic gradients. Dong et al., (2019) used this metric, via MODIS data, to assess the sensitivity of urban vegetation to drought over a 20-year period in Los Angeles, California, USA. They found that vegetation was more sensitive to drought in economically disadvantaged communities, leading to less healthy vegetation and greater exposure to heatwaves under drought conditions.

Historically, analyzing remotely-sensed satellite data imposed a computational burden that limited its accessibility. With cloud-based computing such as Google Earth Engine (GEE), remote sensing workflows have undergone radical transformations. At present, GEE provides access to >80 petabytes of publicly available data for noncommercial use and affords researchers the computational power to analyze it (Cardille et al., 2024). This makes an internet connection the sole constraint for these projects. With tools such as the R package *rgee* (Aybar, 2023), researchers can integrate GEE remote sensing data and functionality into an R-based workflow via a user-friendly interface. These new tools enable *in silico* multi-city comparisons, providing novel opportunities to assess whether cities differ in vegetation response to climate, drought, or urbanization and identify the potential drivers of such differences.

2.1.4) Phenological changes reflected in publicly available data

Mounting evidence suggests that phenology - the timing of life history events such as autumn senescence, seasonal migration, or flowering period - is sensitive to temperature and is an important bioindicator of climatic shifts (Fitter & Fitter, 2002; Walther et al., 2002; Visser & Both, 2005). Cities tend to be warmer than surrounding rural areas due to the absorption and maintenance of heat by impermeable surfaces, a phenomenon known as the urban heat island effect. These increased temperatures advance phenology and lengthen periods of activity in urban areas relative to rural landscapes (D. Li et al., 2021; Merckx et al., 2021; Wohlfahrt et al., 2019). Phenological data may be sourced from iNaturalist (D. Li et al., 2021), digitized museum collections (e.g., herbarium specimens; Willis et al., 2017), dedicated phenological monitoring networks (Merckx et al., 2021; Wohlfahrt et al., 2019), and remotely-sensed data sources (Li et al., 2017). For example, Wohlfahrt et al., (2019) demonstrated urbanization in Europe leads to earlier leaf out, flowering, and fruiting as well as delayed senescence by leveraging data from the Pan European Phenological Project. The use of public data demonstrated a consistent effect of the urban heat island on plant phenology.

Although plants have been the focal group for most urban phenological studies using publicly available data, occurrence data also exist for a broad range of other groups (Olsen et al., 2020). For example, Owen et al., (2024) used iNaturalist records to document earlier activity of the invasive spotted lanternfly (*Lycorma delicatula*) in response to increased temperatures in cities. Their results suggest that the urban heat island effect could facilitate species movements into cities in colder climates. Similarly, Merckx et al., (2021) used community science data from insect monitoring networks in Europe to investigate the phenology of two lepidopteran species. The authors found that urban populations had longer periods of flight activity relative to their rural counterparts. Follow-up experiments confirmed that these shifts were driven by changes in life-history plasticity in both species. Research expanding beyond plants will facilitate our understanding of how urbanization affects ecological interactions. As suggested by Bell et al. (2021), occurrence datasets such as iNaturalist and museum collections are well-suited to fill this knowledge gap.

2.1.5) Uncovering urban species interactions using **in-silico** *approaches*

The high habitat heterogeneity in urban environments can influence species behaviors and interactions (Diamond & Martin, 2021; Irwin et al., 2020). Occurrence data have been widely used to document spatial patterns of biodiversity (e.g., using the spatial coordinates of these records; Beninde et al., 2023), but their associated photographic data remain an underutilized research asset (Pernat et al., 2022). These photographs, sometimes termed “secondary data”, are a rich source of information that can be creatively leveraged to quantify ecological interactions in cities. Importantly, they overcome the challenges of assembling sufficiently large datasets for organisms that are difficult to monitor (Putman et al., 2021).

Marín-Gómez et al., (2022) developed a framework for these studies by compiling a dataset of hummingbird-plant interactions across an urbanization gradient. The authors leveraged 742 photographs available through iNaturalist and eBird to construct plant-pollinator networks. The findings demonstrated that in Mexico City, MX, the hummingbird-plant networks were more generalized compared to the surrounding natural areas. They also found that in urban settings, non-native plant species and non-bird pollinated plants were important network components. Thereby, emphasizing the important role that urban pollinator gardens play in increasing functional connectivity between plants and pollinators.

Image data has also been used to infer how species interactions change through time. Meineke & Davies, (2019) used 576 herbarium records to document changes in plant-herbivore interactions over a 100-year time period by assessing leaf herbivore damage. They found that although rising temperatures increased herbivory, urbanization disrupted this trend, potentially due to a lower abundance of herbivores in cities. Subsequent work has provided a framework for studying long-term plant-herbivore dynamics *in silico* by replicating the findings using digitized herbarium specimens (Meineke et al., 2020).

Photographic data has also been used to study animal-animal interactions. Putman et al., (2021) used images from iNaturalist to study predation and parasitism of the elusive Southern alligator lizard in California, USA. From the images, they quantified the risk of predation, using lizard tail breaks as a proxy, and parasitism, based on the presence and abundance of ticks. Their results demonstrated that urbanization has opposing effects of predation and parasitism on lizards. Urbanization correlates with increased predation, attributed to increased house cat abundance, but also decreased parasitism, likely due to the lower abundance of viable tick hosts in cities.

In sum, these studies demonstrate the power of crowd-sourced image data and how secondary data may be used for understanding species interactions. This relatively untapped resource allows large datasets to be obtained that would otherwise be arduous for an individual researcher to collect. *In silico* approaches that leverage secondary photographic data will be further facilitated by machine-learning techniques that are specialized for large-scale image analyses (see section 3.2).

2.2) Species and populations

2.2.1) Publicly available data reveal urban evolutionary adaptations

Research assessing evolutionary processes in urban landscapes has seen a marked increase in the last 15 years (Diamond & Martin, 2021). This body of work has uncovered how trait evolution is shaped by urban conditions, such as elevated levels of impervious surfaces, increased temperature, changes in resource availability, and pollution (Cheptou et al. 2008; Campbell-Staton et al., 2021; Oziolor et al., 2014). Increasingly, multi-city comparisons have become a microcosm for studying parallel evolution (Reid et al., 2016; Santangelo et al., 2022; Winchell et al., 2023). A typical multi-city project would involve significant logistical challenges and financial costs, but the abundance of photographic data from community science initiatives alleviates these issues. This makes *in silico* approaches an advantageous avenue to understand how urbanization may elicit adaptive morphological changes.

Cosentino & Gibbs (2022) demonstrated the power of *in silico* work by asking how urbanization influences melanization patterns in eastern gray squirrels. This is a known Mendelian trait in eastern grey squirrels, allowing for genotyping by photographic data. The authors compiled >60,000 iNaturalist photos across 43 cities, and community scientists scored squirrel coat color using Zooniverse, a platform to crowd source data collection for scientific research. Melanism increased with the percent impervious surfaces, especially in large, heavily forested, northern cities. The authors theorize that the inferred parallel shifts in melanism are in response to similar selective pressures, such as predation or selection on correlated traits.

Similarly, Fukano et al., (2023) combined iNaturalist photos of the flowering plant *Oxalis corniculata* with the percent impervious surface as a metric of urbanization. The authors found higher proportions of red individuals in urban areas, supporting their hypothesis that red leaf color is adaptive within an urban context. Further experimentation supported this, as red coloration potentially helps plants cope with the urban heat island effect.

Other digitized repositories have extended UEE studies beyond the traits that may be captured in photographs. In one such example, Hantak et al., (2021) used publicly available mammalian datasets to test Bergmann's rule, which states that in cooler climates populations tend to have larger body sizes (Bergmann, 1847). A dataset of over 100,000 size measurement records was assembled using VertNet, NEON, and NACSM, and correlated with human population density. The authors found that, despite the urban heat island effect making cities warmer (Phelan et al., 2015), Bergmann's rule did not apply. Their results provide further avenues for study, such as whether patterns in body size are better explained by access to novel food sources.

In summary, these studies demonstrate the effectiveness *of in silico* approaches for compiling multi-city datasets of photographic data and identifying broad trends of parallel evolution. Integrating digitized traits and other specimen data will undoubtedly provide further avenues for study and help clarify how urban environments drive evolutionary change.

2.2.2) Inferring plastic responses to urbanization using publicly available data

Phenotypic plasticity, the ability for a single genotype to express differing phenotypes due to environmental conditions, can facilitate success in urban areas by buffering organisms against stressful urban conditions (Bressler et al., 2020). However, plasticity may also be maladaptive, leading to rapid adaptation or local extinction (Gomulkiewicz & Holt, 1995; Chevin et al., 2010; Ghalambor et al., 2015; Diamond & Martin, 2021). Studying plasticity (adaptive or maladaptive) in urban organisms can inform researchers of the genetic mechanisms underlying population persistence in novel environments.

In silico approaches allow scientists to ask whether urban environmental cues elicit known plastic responses (**Table 1**). For example, Batool et al., (2024) tested how temperature differences, due to the urban heat island and latitude, affect cuticular melanin of monarch larvae, a known plastic trait. Using iNaturalist and BugGuide, they quantified melanism in larvae and correlated it with human population density and latitude. Larvae at warmer, lower latitudes had decreased levels of melanin. Urbanization, however, did not elicit a significant effect on melanism, indicating that increased temperatures due to the urban heat island effect were insufficient to trigger a plastic response.

Trait databases such as the TRY Plant Trait Database (**Table 1**) have also proven valuable for studying plasticity. Using a partially *in silico* approach, Ibsen et al., (2023) compared water use and carbon gain strategies between 30 species of trees in urban Los Angeles and in their respective native habitats. Urban trees were irrigated, and trait data were collected *in situ,* while trait data for unirrigated trees in their native range were obtained from TRY. To collect such a trait dense dataset would require global fieldwork and be logistically challenging; however, the TRY database allowed Ibsen et al., (2023) to adapt several studies into a robust comparative analysis. Their analysis revealed a decoupling of the relationship between carbon gain and water use among urban trees relative to those in their native habitat. The effects were magnified in arid locations, presumably as a result of high resource availability.

Although these examples show that plasticity may be inferred from *in silico* datasets, caution should be taken to avoid conflating plasticity and evolutionary adaptation. The outlined examples involved researchers using datasets of species with known plastic responses to make their inference (Ibsen et al., 2023; Batool et al., 2024), and this is key to designing a robust study using publicly available data.

2.2.3) Investigating how urbanization influences genetic diversity **in silico**

Urbanization is hypothesized to reduce genetic variation due to habitat loss and increased fragmentation, which decreases effective population size and disrupts gene flow (Johnson & Munshi-South, 2017). Understanding the influence urbanization has on genetic diversity is imperative for developing conservation strategies and predicting the persistence of urban populations.

Traditionally, samples are collected and sequenced, which can be laborious, time consuming, expensive, and logistically infeasible. These challenges can be amplified for multi-city or multi-species studies. However, there is a rich history of depositing genetic data into repositories such as GenBank, DataOne, or MacroPopGen, making *in silico* approaches a powerful alternative (**Table 1**).

For example, Schmidt et al., (2020) compiled a dataset of microsatellites to assess whether urbanization and human land use negatively affect population genetic parameters. DataOne and Dryad were used to obtain data from 66 species of North American mammals and non-migratory birds. Using these microsatellites, they calculated effective population size, gene diversity, allelic richness, and fixation index (F_{st}). These metrics were then correlated with measures of urbanization, including urban vs. rural land use classifications, human population density (U.S. Census Bureau and Statistics Canada), and the Human Footprint index. With increasing urbanization, mammals exhibited consistent and significant decreases in effective population size, allelic richness, and gene diversity, as well as an increase in population differentiation. Similar results were found when Miraldo et al., (2016) used publicly available amphibian *cytochrome B* (cytB) and mammalian *cytochrome c oxidase subunit 1* (CO1) sequences to analyze genetic diversity at a global scale. However, they did not find that human disturbance negatively affected genetic diversity using the cytB sequences in mammals. Conversely, Millette et al., (2020) found that genetic diversity at the bird and mammalian CO1 gene was not affected by increased human density. *In silico a*pproaches using microsatellites have also been applied to amphibians and reptiles (Schmidt & Garroway, 2021) and marine fishes (Karachaliou et al., 2025), showing neutral and negative effects, respectively.

Habrich et al., (2021) reused microsatellite data from MacroPopGen (**Table 1**) to ask how road networks and human population density (**Table 2**) affect the genetic diversity of North American terrestrial mammals. They found genera-specific effects that are weakly modulated by life history traits and behaviors. For example, genera with large home range sizes, such as Caribou *(Rangifer*), were most negatively affected by increasing road density. Furthermore, the authors found that human density has a greater negative effect on genetic diversity than road impacts do, corroborating the results of Schmidt et al., (2020).

Public data also allows researchers to explore how past and present social policies and practices impact the genetic landscape of cities. For example, historical redlining influences neighborhood demographics and the distribution of wealth across urban areas. Although outlawed today, the effects of this practice remain evident. Schmidt and Garroway (2022) analyzed microsatellite data from DataOne for 40 terrestrial animal taxa. They tested whether socioeconomic legacies influence population genetic metrics, and found that wildlife in predominantly white neighborhoods maintained slightly larger effective population sizes, higher genetic diversity, and lower genetic differentiation. These results highlight how past and contemporary societal processes contribute to neighborhood-specific results (Schell et al., 2020).

Collectively, the studies outlined demonstrate that urbanization and its influences on population dynamics may be inferred using *in silico* approaches. Consistent patterns emerge across studies, yielding critical insights that can inform city planning more effectively.

3) Resources for use in Urban Ecology and Evolution

3.1) Emerging and Underutilized Datasets

In this section, we present a range of emerging and underutilized resources that show promise for studying UEE. Like all resources, the extent to which they may be leveraged varies. Importantly, scientists and the general public can contribute to these repositories by depositing their own original data, allowing data to be repurposed and fully utilized in new and creative ways. Larger repositories and *in silico* approaches offer the ability to combine and synthesize data, producing new insights into the broader patterns of urbanization. In addition to the datasets and repositories covered here, The Opportunistic Database of Biodiversity Databases (https://earthskysea.org/biodiversity-databases/) and The Urban Eco-Evo Trait Mapping Toolkit (Savage et al., 2024 - https://www.urbanevoecotools.org/) document databases that may have relevance to UEE.

3.1.1) Digitized Museum Specimens

Natural history collections house a plethora of information relevant to UEE (Shultz et al., 2020), but access to physical specimens can be costly and logistically challenging. The rapid digitization of collections allows access to specimens entirely *in silico*, furthering research that has already leveraged this data source for UEE (e.g., Meineke & Davies, 2019; Meineke et al., 2020; Pearson et al., 2020; Hantak et al., 2021).

Digitization efforts go beyond images. Digital 3D models of specimens are becoming a prevalent data type, adding another dimension to current research. Many 3D structures can be accessed on Morphosource, a data repository that houses 3D and 2D media of physical objects (e.g., bones, whole specimens, cultural artifacts). Through Morphosource, researchers can access media from research bodies such as oVert and interact with files in the browser.

iDigBio, a repository aggregating digitization efforts across predominantly U.S. based institutions, facilitates access to digitized specimens. The Global Registry of Scientific Collections (GRSciColl), through GBIF, helps connect iDigBio with global digitized collections. Together, these platforms increase the accessibility of digitized data for UEE research.

3.1.2) Spatial Resources

In large urban systems where the vast majority of land is privately owned, spatial resources are especially valuable. Google Street View (GSV) has images for all roadways in an urban area, often generated every few years. This data may be used to analyze the distribution of urban vegetation and changes in the urban landscape over time. GSV has already shown potential in investigating diversity in urban areas (Velasquez-Camacho et al., 2024; Ye et al., 2025), surveying plants growing in private yards (Ringland et al., 2021), and monitoring invasive species (Closa-Tarres et al., 2025). With the rapid development of AI tools for image recognition, this resource may become more prominent in UEE.

Another valuable resource is the ArcGIS Living Atlas of the World, which contains a plethora of maps (e.g., iNaturalist observations and World Traffic Service) and deep-learning packages (e.g., Tree Point Classification and Building Footprint Extraction) that may be leveraged for urban research. While a membership is required, many institutions and universities have licenses for student use, making this widely accessible in an academic setting.

3.1.3) Ecological Survey Data

Several repositories of ecological survey data have been created to standardize data re-use, while also bypassing the need for extensive and time consuming literature reviews. Movebank collects, archives, and provides visualization for animal tracking data (Kays et al., 2021), including data from urban systems that may be used to study how urban organisms navigate through cities. For example, the Chicago Black Crown Night Heron project has uploaded tracking data of urban herons.

Similarly, sPlotOpen provides a centralized resource for archiving vegetation plot surveys (Sabatini et al., 2021). Although sPlotOpen is mainly composed of surveys in natural or semi-natural sites, the addition of vegetation plot surveys will assist in UEE pursuits. For example, sPlotOpen has been used to investigate how species co-occurrence can be used for ecological predictions (Siefert et al., 2024), a concept relevant to urban systems.

Arctos (Cicero et al., 2024) is another integrative and unified platform that provides online access to physical data housed at museums and cultural institutions. Data is contributed by diverse organizations (e.g., NEON, LTER, and Museums) spanning many disciplines (e.g., geology, anthropology, botany, and mammalogy). For example, Arctos provides access to specimen measurements, mark/recapture data, species sighting surveys, documents (e.g., field notes), and media, among others. The collated data provide a broad foundation for *in silico* UEE studies.

3.1.4) Genetic Data

Aside from microsatellite and mitochondrial DNA, publicly available genetic data have been underutilized in UEE, likely due to a lack of geolocated sequence information. Presently available sequence data can be accessed through the National Center for Biotechnology Information (NCBI), which houses data ranging from genes to genomes. Unfortunately, filtering for applicable sequence data is difficult due to inconsistent documentation of metadata (e.g., location information). The Global Genome Biodiversity Network (GGBN) remediates this issue for sequenced museum collections by linking the NCBI accession number with the museum collection record and any associated information (location, collector, date, etc.). Preserved, unsequenced DNA can also be found through GGBN. Although unsequenced, this resource may help overcome barriers associated with organism collection and DNA extraction, connecting researchers with museums that hold samples of interest.

Repository or Dataset	**Data Type or Information**	**Citation/Link**
Digitized Natural History Collections: iDigBio	Digitizied photographs, scans, CT-scans, etc., of preserved natural history specimens	https://www.idigbio.org/
Global Registry of Scientific Collections	A repository of information that connects researchers to data	https://scientific-collections.gbif.org/
MorphoSource	A repository of 3D and 2D media of natural history specimens	https://www.morphosource.org/
Google Street View Images	Photographs of Roadways	https://developers.google.com/maps/documentation/streetview/get-api-key
ArcGIS Living Atlas of the World	Paid application with access to maps and deep-learning packages	https://livingatlas.arcgis.com/en/home/

Movebank	A database to store, collect, analyze, visualize, and archive tracking data of animals.	Kays et al., 2021
sPlotOpen	Vegetation plot surveys across landscapes	Sabatini et al., 2021 https://www.idiv.de/research/projects/splot/splotopen-splot/
Arctos	Biological data includes specimen measurements, mark/recapture data, species sighting surveys, etc.	Cicero et al., 2024
National Center for Biotechnology Information	Repositories for genetic data, encompassing genes, transcriptomes, protein sequences, and genomes.	https://www.ncbi.nlm.nih.gov/
Global Genome Biodiversity Network	Facilitates sharing of physically archived DNA and tissue samples. Connects sequence data derived from collections to relevant metadata.	https://www.ggbn.org/ggbn_portal/

Table 3. This table presents emerging and underutilized datasets that may prove useful to the field of urban ecology and evolution. It also summarizes the information housed in each repository and the corresponding link address or citation.

3.2) Utilizing Artificial Intelligence for UEE Research

Given the large number of publicly available photographic datasets (e.g., iNaturalist records), most applied methods of Artificial Intelligence in UEE have utilized image processing to automatically extract relevant information from photographs. Convolutional neural networks (CNNs) have been the most prominent in accomplishing this. Tools utilizing CNNs have been developed to extract trait data from photographic museum collections of plants (LeafMachine2 by Weaver & Smith, 2023; Love et al., 2021; Davis et al., 2020; Pearson et al., 2020; Meineke et al., 2020) and animal skeletons (Skelevision; Weeks et al., 2023) as well as phenological data from iNaturalist photographs (PhenoVision by Dinnage et al., 2025; Reeb et al., 2022; Stewart et al., 2025). These tools provide detailed workflows, open source code, and trained models, allowing researchers to leverage them for their own studies. Similarly, tools have been developed to analyze Google Street View (GSV) images using CNNs and existing plant identification software (Closa-Tarres et al., 2025; Velasquez-Camacho et al., 2024; Ringland et al., 2021). Automated image selection may allow this technology to be more widely used for UEE research, as most studies choose images manually or based on prior information (but see Velasquez-Camacho et al., 2024). When building new tools, developers should ensure their code is well documented and open source so that researchers can modify programs for specific use cases. Although only recently explored, artificial intelligence methods have already proven to be a powerful tool for investigating patterns in UEE.

4) Dealing with biases and limitations of publicly available datasets

Although publicly available datasets provide an invaluable resource for ecology and evolution, they are not without limitations. Several studies highlighted in this review have outlined how this data can be biased, incomplete, or lacking key pieces of metadata. For example, the distribution of iNaturalist occurrence data can be geographically biased within cities (Carlen et al., 2024), or incomplete metadata can result in unusable genetic data (Schmidt et al., 2020). Data sourced from GenBank may also require quality inspection. Sequences may be assigned to the wrong taxa or a synonym, contaminated, or contain sequencing errors (X. Li et al., 2018; van den Burg & Vieites, 2023). Acknowledging these limitations and ensuring they are addressed is an essential aspect of performing reliable *in silico* research.

4.1) Spatial biases in sampling

Spatial sampling biases are thoroughly documented across publicly available datasets (see Beck et al., 2014; Di Cecco et al., 2021; Johnston et al., 2022; Backstrom et al., 2025). These biases are often amplified in urban areas, as sociological variables influence sampling accessibility and effort. For example, wealthier, predominantly white neighborhoods tend to be better sampled, likely due to a history of unjust land-use policies (Carlen et al., 2024).

Rather than reiterating where biases occur, here we present resources and frameworks for identifying, evaluating, and mitigating biases in community science datasets. Backstrom et al., (2025) formulated a method to estimate spatial and temporal biases, their interactions, and their influence on model inference. Bowler et al., (2025) proposed treating sampling biases in species-level abundance data as a “missing data” problem. In their review, they present several frameworks for identifying the types of bias, the implications for those biases, and the statistical methods to deal with them. Community science datasets can also be spatially subsampled, decreasing the overall sample size but reducing unevenness in the data (Beck et al., 2014). Fink et al., (2023) present a Double Machine

Learning method to analyze abundance trends while accounting for observer bias on eBird data. Furthermore, Bird et al., (2014) reviewed a range of statistical methods, from linear models to machine learning, that can account for biases in community science biodiversity datasets. In total, these tools provide researchers with an array of approaches to circumvent biases, making *in silico* studies more reliable.

4.2) Data access issues

Open data initiatives and centralized repositories have facilitated easier access to data, yet barriers to data reuse still exist. For example, Schmidt et al., (2020) excluded 36 of the 313 initially found microsatellite datasets from their analysis due to missing or improperly deposited metadata (e.g., spatial information, population delineation). Some of these issues were resolved by contacting the research groups that generated the data, but this quickly becomes infeasible as datasets are added. It is imperative to standardize archival procedures to ensure that data can be effectively reused. A recent set of guidelines by Leigh et al., (2024) provides best practices for archiving genomic data, and the FAIR principles (Findable, Accessible, Interoperable, and Reusable) offer broader guidance and are applicable to all scientific data types (Wilkinson et al., 2016). The adoption of standard archival procedures will create more operable data repositories, greatly facilitating the novel insights generated from *in silico* studies.

5) Enhancing Participation in Urban Ecology and Evolution with Publicly Available Datasets

Publicly available datasets have the power to broaden participation in UEE by removing financial, structural, and logistical barriers. Because these studies only require an internet connection, they are accessible to all students. This accessibility is especially relevant to non-traditional students, who are often balancing employment, family, and other responsibilities that hinder their engagement in traditional lab work. Expanding *in silico* UEE approaches can improve participation, by promoting a more equitable and inclusive research environment, enabling the participation of individuals historically underrepresented in science.

5.1) Improving the undergraduate research experience

Publicly available data are particularly well suited for course-based undergraduate research experiences (CUREs), which expose students to authentic scientific pursuits that generate novel findings. These experiences can produce similar student outcomes to traditional undergraduate research experiences, which have been shown to improve retention and persistence among students from underrepresented groups (Shapiro et al., 2015; Eagan et al., 2013; Hunter et al., 2007; Lopatto, 2007; Russell et al., 2007). Because acquiring original data for these programs can be logistically and financially challenging, publicly available databases provide a free and accessible resource. This makes them a low-cost alternative for instructors seeking to obtain data for research in the classroom (see Arreguin et al., 2025 for a CURE utilizing publicly available genetic data). As outlined throughout this review, this data has relevance across all levels of biological organization, making these datasets suitable for a broad range of classes or research programs. Moreover, the large size of these datasets exposes students to valuable data science and programming experience. This is highly translational across fields, and early training can support and influence students' future careers.

Practical examples are the CUREs developed by Gastreich (2021), centered around assessing urban bird diversity using eBird data, and Heard et al., (2025), which investigates the impacts of urbanization on lizards using iNaturalist. Such courses can offer a useful example for instructors and be readily adapted and modified to explore UEE across spatial scales and taxa.

5.2) Facilitating graduate and faculty research

Beyond undergraduate training, graduate students and research faculty can successfully employ publicly available data for their own research goals. The low-cost nature of these data make for an excellent resource when gathering preliminary data for research and grant proposals or performing standalone research. The use of publicly available datasets can alleviate barriers associated with original data collection. Individuals with disabilities may find aspects of field or laboratory based studies to be challenging, impeding their ability to participate in UEE projects. Furthermore, the high percentage of private land ownership in cities often hinders urban field collections, and the monetary cost associated with data generation may be prohibitively expensive. Publicly available datasets allow researchers to perform rigorous, hypothesis driven research that contributes meaningfully to the field of UEE.

6) Concluding Remarks

In this review, we have collated presently used and developing datasets, along with potentially useful tools, with the intention that this work will act as a guide for utilizing *in silico* approaches. It is important to note that this work is only possible thanks to databases that make data freely available and researchers who work hard to ensure that their data is easily accessible. As tools and datasets are developed, it is imperative that data remains publicly accessible and code is open-source. In doing so, we allow scientists to adjust tools as needed and use data to its full extent, furthering the range of questions that can be answered.

Data Accessibility

Datasets highlighted in this manuscript are accessible via the in-text tables and the appendix.

Conflicts of Interests Statements

The authors declare no conflicts of interests

Author Contributions Statement

Shawn Arreguin: Conceptualization, writing - original draft, writing - review and editing, visualizations (supporting). Joseph F. Walker: Conceptualization, writing - original draft, writing - review and editing. Natalie L.R. Love: Conceptualization, visualizations (lead), writing - original draft, writing - review and editing.

Acknowledgements

We would like to thank Erin Grady for helping us find relevant literature. Roberta Mason-Gamer, Beatriz Dybas de Natividade, Hector Palma, Alexa Tyszka, Jiale Zhang, Jordan Kressman, Karolis Ramanaskus, Eric Bretz, Timi

Okeseni, and Linda Mansour all provided valuable comments and suggestions to improve our manuscript and pointed us to relevant databases. This manuscript is supported by start-up funding through the Walker Lab at the University of Illinois Chicago.

References

Angel, S., Parent, J., Civco, D. L., Blei, A., & Potere, D. (2011). The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050. *Progress in Planning*, *75*(2), 53–107. https://doi.org/10.1016/j.progress.2011.04.001

Arreguin, S., Walker-Hale, N., Casagrande, M. R., Bishop, S., Pugacewicz, E., Ramizuddin, M., Savitzky, C., Morales, N., Dhan, A., Natividade, B. D. da, Madrid, V., Simmons, C., Diaz, A., Ahmed, D., Ho, Z., Pellecer, J. N., Diaz, S., Mathew, T., Mirshed, T., … Walker, J. F. (2025). *Going green: Recycling transcriptomes to infer evolutionary relationships, gene duplication, gene tree conflict, and patterns of molecular evolution in the Apocynaceae* (p. 2025.06.20.660724). bioRxiv. https://doi.org/10.1101/2025.06.20.660724

Aybar et al., (2020). rgee: An R package for interacting with Google Earth Engine. Journal of Open Source Software, 5(51), 2272, https://doi.org/10.21105/joss.02272

Backstrom, L. J., Callaghan, C. T., Worthington, H., Fuller, R. A., & Johnston, A. (2025). Estimating sampling biases in citizen science datasets. Ibis, 167(1), 73–87. https://doi.org/10.1111/ibi.13343

Batool, R., Jenan, J., Schuster, M., & Murray, R. L. (2024). How do latitude and urban heat islands affect larval melanization in monarch butterflies (Danaus plexippus) ? Canadian Journal of Zoology, 102(3), 315–321. https://doi.org/10.1139/cjz-2023-0099

Beck, J., Böller, M., Erhardt, A., & Schwanghart, W. (2014). Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics*, *19*, 10–15. https://doi.org/10.1016/j.ecoinf.2013.11.002

Bell, K. C., Adams, B. J., Brown, B. V., Ludt, W. B., Pauly, G. B., Shultz, A. J., & Vendetti, J. E. (2021). Museums Are Critical Infrastructure for Studying Urban Biodiversity. *BioScience*, *71*(5), 433. https://doi.org/10.1093/biosci/biab038

Beninde, J., Delaney, T. W., Gonzalez, G., & Shaffer, H. B. (2023). Harnessing iNaturalist to quantify hotspots of urban biodiversity: The Los Angeles case study. Frontiers in Ecology and Evolution, 11, 983371. https://doi.org/10.3389/fevo.2023.983371

Bergmann, C. (1847). About the relationships between heat conservation and body size of animals. Goett. Stud. (original Ger.) 1, 595–708.

Bird, T. J., Bates, A. E., Lefcheck, J. S., Hill, N. A., Thomson, R. J., Edgar, G. J., Stuart-Smith, R. D., Wotherspoon, S., Krkosek, M., Stuart-Smith, J. F., Pecl, G. T., Barrett, N., & Frusher, S. (2014). Statistical solutions for error and bias in global citizen science datasets. Biological Conservation, 173, 144–154. https://doi.org/10.1016/j.biocon.2013.07.037

Blair, R. B. (1996). Land Use and Avian Species Diversity Along an Urban Gradient. Ecological Applications, 6(2), 506–519. https://doi.org/10.2307/2269387

Blair, R. B. (2001). Birds and Butterflies Along Urban Gradients in Two Ecoregions of the United States: Is Urbanization Creating a Homogeneous Fauna? In J. L. Lockwood & M. L. McKinney (Eds.), Biotic Homogenization (pp. 33–56). Springer US. https://doi.org/10.1007/978-1-4615-1261-5_3

Border, J. A., Newson, S. E., White, D. C. J., & Gillings, S. (2017). Predicting the likely impact of urbanisation on bat populations using citizen science data, a case study for Norfolk, UK. Landscape and Urban Planning, 162, 44–55. https://doi.org/10.1016/j.landurbplan.2017.02.005

Bowler, D. E., Boyd, R. J., Callaghan, C. T., Robinson, R. A., Isaac, N. J. B., & Pocock, M. J. O. (2025). Treating gaps and biases in biodiversity data as a missing data problem. Biological Reviews of the Cambridge Philosophical Society, 100(1), 50–67. https://doi.org/10.1111/brv.13127

Bressler, S. A., Diamant, E. S., Tingley, M. W., & Yeh, P. J. (2020). Nests in the cities: Adaptive and non-adaptive phenotypic plasticity and convergence in an urban bird. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1941), 20202122. https://doi.org/10.1098/rspb.2020.2122

Cadenasso, M. L., Pickett, S. T. A., & Schwarz, K. (2007). Spatial heterogeneity in urban ecosystems: Reconceptualizing land cover and a framework for classification. *Frontiers in Ecology and the Environment*, *5*(2), 80–88. https://doi.org/10.1890/1540-9295(2007)5%255B80:SHIUER%255D2.0.CO;2

Callaghan, C. T., Bino, G., Major, R. E., Martin, J. M., Lyons, M. B., & Kingsford, R. T. (2019). Heterogeneous urban green areas are bird diversity hotspots: Insights using continental-scale citizen science data. Landscape Ecology, 34(6), 1231–1246. https://doi.org/10.1007/s10980-019-00851-6

Callaghan, C. T., Major, R. E., Wilshire, J. H., Martin, J. M., Kingsford, R. T., & Cornwell, W. K. (2019). Generalists are the most urban-tolerant of birds: A phylogenetically controlled analysis of ecological and life history traits using a novel continuous measure of bird responses to urbanization. Oikos, 128(6), 845–858. https://doi.org/10.1111/oik.06158

Campbell-Staton, S. C., Velotta, J. P., & Winchell, K. M. (2021). Selection on adaptive and maladaptive gene expression plasticity during thermal adaptation to urban heat islands. *Nature Communications*, *12*(1), 6195. https://doi.org/10.1038/s41467-021-26334-4

Cardille, J. A., Crowley, M. A., Saah, D., & Clinton, N. E. (Eds.). (2024). Cloud-Based Remote Sensing with Google Earth Engine: Fundamentals and Applications. Springer International Publishing. https://doi.org/10.1007/978-3-031-26588-4

Cârlan, I., Mihai, B.-A., Nistor, C., & Große-Stoltenberg, A. (2020). Identifying urban vegetation stress factors based on open access remote sensing imagery and field observations. Ecological Informatics, 55, 101032. https://doi.org/10.1016/j.ecoinf.2019.101032

Carlen, E. J., Estien, C. O., Caspi, T., Perkins, D., Goldstein, B. R., Kreling, S. E. S., Hentati, Y., Williams, T. D., Stanton, L. A., Des Roches, S., Johnson, R. F., Young, A. N., Cooper, C. B., & Schell, C. J. (2024). A framework for contextualizing social-ecological biases in contributory science data. People and Nature, 6(2), 377–390. https://doi.org/10.1002/pan3.10592

Cheptou, P.-O., Carrue, O., Rouifed, S., & Cantarel, A. (2008). Rapid evolution of seed dispersal in an urban environment in the weed Crepis sancta. Proceedings of the National Academy of Sciences, 105(10), 3796–3799. https://doi.org/10.1073/pnas.0708446105

Chevin, L.-M., Lande, R., & Mace, G. M. (2010). Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory. PLOS Biology, 8(4), e1000357.

https://doi.org/10.1371/journal.pbio.1000357

Cicero, C., Koo, M. S., Braker, E., Abbott, J., Bloom, D., Campbell, M., Cook, J. A., Demboski, J. R., Doll, A. C., Frederick, L. M., Linn, A. J., Mayfield-Meyer, T. J., McDonald, D. L., Nachman, M. W., Olson, L. E., Roberts, D., Sikes, D. S., Witt, C. C., & Wommack, E. A. (2024). Arctos: Community-driven innovations for managing natural and cultural history collections. PLOS ONE, 19(5), e0296478. https://doi.org/10.1371/journal.pone.0296478

Closa-Tarres, A., Rojano, F., & Strager, M. P. (2025). Geospatial Data and Google Street View Images for Monitoring Kudzu Vines in Small and Dispersed Areas. Earth, 6(2), 40. https://doi.org/10.3390/earth6020040

Cosentino, B. J., & Gibbs, J. P. (2022). Parallel evolution of urban–rural clines in melanism in a widespread mammal. *Scientific Reports*, *12*, 1752. https://doi.org/10.1038/s41598-022-05746-2

Crawford, C. J., Roy, D. P., Arab, S., Barnes, C., Vermote, E., Hulley, G., Gerace, A., Choate, M., Engebretson, C., Micijevic, E., Schmidt, G., Anderson, C., Anderson, M., Bouchard, M., Cook, B., Dittmeier, R., Howard, D., Jenkerson, C., Kim, M., … Zahn, S. (2023). The 50-year Landsat collection 2 archive. Science of Remote Sensing, 8, 100103. https://doi.org/10.1016/j.srs.2023.100103

Curti, J. N., Barton, M., Flores, R. G., Lechner, M., Lipman, A., Montgomery, G. A., Park, A. Y., Rochel, K., & Tingley, M. W. (2024). Using unstructured crowd-sourced data to evaluate urban tolerance of terrestrial native animal species within a California Mega-City. *PLOS ONE*, *19*(5), e0295476. https://doi.org/10.1371/journal.pone.0295476

Davis, C. C., Champ, J., Park, D. S., Breckheimer, I., Lyra, G. M., Xie, J., Joly, A., Tarapore, D., Ellison, A. M., & Bonnet, P. (2020). A New Method for Counting Reproductive Structures in Digitized Herbarium Specimens Using Mask R-CNN. *Frontiers in Plant Science*, *11*, 1129. https://doi.org/10.3389/fpls.2020.01129

Di Cecco, G. J., Barve, V., Belitz, M. W., Stucky, B. J., Guralnick, R. P., & Hurlbert, A. H. (2021). Observing the Observers: How Participants Contribute Data to iNaturalist and Implications for

Biodiversity Science. BioScience, 71(11), 1179–1188. https://doi.org/10.1093/biosci/biab093

Diamond, S. E., & Martin, R. A. (2021). Evolution in Cities. Annual Review of Ecology, Evolution, and Systematics, 52(Volume 52, 2021), 519–540. https://doi.org/10.1146/annurev-ecolsys-012021-021402

Dinnage, R., Grady, E., Neal, N., Deck, J., Denny, E., Walls, R., Seltzer, C., Guralnick, R., & Li, D. (2025). PhenoVision: A framework for automating and delivering research-ready plant phenology data from field images. *Methods in Ecology and Evolution*, *16*(8), 1763–1780. https://doi.org/10.1111/2041-210X.70081

Dong, C., MacDonald, G., Okin, G. S., & Gillespie, T. W. (2019). Quantifying Drought Sensitivity of Mediterranean Climate Vegetation to Recent Warming: A Case Study in Southern California. Remote Sensing, 11(24), 2902. https://doi.org/10.3390/rs11242902

Eagan, M. K., Hurtado, S., Chang, M. J., Garcia, G. A., Herrera, F. A., & Garibay, J. C. (2013). Making a Difference in Science Education: The Impact of Undergraduate Research Programs. 50(4), 656–850. https://doi.org/10.3102/0002831213482038

Estien, C. O., Fidino, M., Wilkinson, C. E., Morello-Frosch, R., & Schell, C. J. (2024). Historical redlining is associated with disparities in wildlife biodiversity in four California cities. Proceedings of the National Academy of Sciences, 121(25), e2321441121. https://doi.org/10.1073/pnas.2321441121

Faeth, S. H., Bang, C., & Saari, S. (2011). Urban biodiversity: Patterns and mechanisms. Annals of the New York Academy of Sciences, 1223(1), 69–81. https://doi.org/10.1111/j.1749-6632.2010.05925.x

Fink, D., Johnston, A., Strimas-Mackey, M., Auer, T., Hochachka, W. M., Ligocki, S., Oldham Jaromczyk, L., Robinson, O., Wood, C., Kelling, S., & Rodewald, A. D. (2023). A Double machine learning trend model for citizen science data. *Methods in Ecology and Evolution*, *14*(9), 2435–2448. https://doi.org/10.1111/2041-210X.14186

Fitch, G., Glaum, P., Simao, M.-C., Vaidya, C., Matthijs, J., Iuliano, B., & Perfecto, I. (2019). Changes in adult sex ratio in wild bee communities are linked to urbanization. Scientific Reports, 9, 3767. https://doi.org/10.1038/s41598-019-39601-8

Fitch, G., Wilson, C. J., Glaum, P., Vaidya, C., Simao, M.-C., & Jamieson, M. A. (2019). Does urbanization favour exotic bee species? Implications for the conservation of native bees in cities. Biology Letters, 15(12), 20190574. https://doi.org/10.1098/rsbl.2019.0574

Fitter, A. H., & Fitter, R. S. R. (2002). Rapid Changes in Flowering Time in British Plants | Science. Science, 296(5573), 1689–1691. https://doi.org/10.1126/science.1071617

Fukano, Y., Yamori, W., Misu, H., Sato, M. P., Shirasawa, K., Tachiki, Y., & Uchida, K. (2023). From green to red: Urban heat stress drives leaf color evolution. *Science Advances*, *9*(42), eabq3542. https://doi.org/10.1126/sciadv.abq3542

Gastreich, K. R. (2021). *Assessing Urban Biodiversity With the eBird Citizen Science Project: A Course-Based Undergraduate Research Experience (CURE) Module*. https://doi.org/10.24918/cs.2020.18

Ghalambor, C. K., Hoke, K. L., Ruell, E. W., Fischer, E. K., Reznick, D. N., & Hughes, K. A. (2015). Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. Nature, 525(7569), 372–375. https://doi.org/10.1038/nature15256

Gomulkiewicz, R., & Holt, R. D. (1995). When does Evolution by Natural Selection Prevent Extinction? Evolution, 49(1), 201–207. https://doi.org/10.2307/2410305

Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X., & Briggs, J. M. (2008). Global Change and the Ecology of Cities. Science, 319(5864), 756–760. https://doi.org/10.1126/science.1150195

Habrich, A. K., Lawrence, E. R., & Fraser, D. J. (2021). Varying genetic imprints of road networks and human density in North American mammal populations. Evolutionary Applications, 14(6), 1659. https://doi.org/10.1111/eva.13232

Hantak, M. M., McLean, B. S., Li, D., & Guralnick, R. P. (2021). Mammalian body size is determined by interactions between climate, urbanization, and ecological traits. Communications Biology, 4(1), 1–10. https://doi.org/10.1038/s42003-021-02505-3

Heard, M. J., Putman, B. J., Riskin, S. H., & Thawley, C. J. (2025). Collecting Novel Data from Inaturalist Photos to Teach Students About the Impacts of Urbanization on Animal Behavior & Ecological Interactions. The American Biology Teacher, 87(4–5), 286–291. https://doi.org/10.1525/abt.2025.87.4-5.286

Hensley, C. B., Trisos, C. H., Warren, P. S., MacFarland, J., Blumenshine, S., Reece, J., & Katti, M. (2019). Effects of Urbanization on Native Bird Species in Three Southwestern US Cities. Frontiers in Ecology and Evolution, 7. https://doi.org/10.3389/fevo.2019.00071

Huang, S., Tang, L., Hupy, J. P., Wang, Y., & Shao, G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. Journal of Forestry Research, 32(1), 1–6. https://doi.org/10.1007/s11676-020-01155-1

Hunter, A.-B., Laursen, S. L., & Seymour, E. (2007). Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development. Science Education, 91(1), 36–74. https://doi.org/10.1002/sce.20173

Ibsen, P. C., Santiago, L. S., Shiflett, S. A., Chandler, M., & Jenerette, G. D. (2023). Irrigated urban trees exhibit greater functional trait plasticity compared to natural stands. Biology Letters, 19(1), 20220448. https://doi.org/10.1098/rsbl.2022.0448

Irwin, R. E., Youngsteadt, E., Warren, P. S., & Bronstein, J. L. (2020). The Evolutionary Ecology of Mutualisms in Urban Landscapes. In M. Szulkin, J. Munshi-South, & A. Charmantier (Eds.), Urban Evolutionary Biology (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780198836841.003.0007

Ives, C. D., Lentini, P. E., Threlfall, C. G., Ikin, K., Shanahan, D. F., Garrard, G. E., Bekessy, S. A., Fuller, R. A., Mumaw, L., Rayner, L., Rowe, R., Valentine, L. E., & Kendal, D. (2016). Cities are hotspots for threatened species. Global Ecology and Biogeography, 25(1), 117–126. https://doi.org/10.1111/geb.12404

Jamieson, M. A., Carper, A. L., Wilson, C. J., Scott, V. L., & Gibbs, J. (2019). Geographic Biases in Bee Research Limits Understanding of Species Distribution and Response to Anthropogenic

Disturbance. Frontiers in Ecology and Evolution, 7. https://doi.org/10.3389/fevo.2019.00194

Johnson, M. T. J., & Munshi-South, J. (2017). Evolution of life in urban environments. Science, 358(6363), eaam8327. https://doi.org/10.1126/science.aam8327

Johnston, A., Matechou, E., & Dennis, E. B. (2023). Outstanding challenges and future directions for biodiversity monitoring using citizen science data. *Methods in Ecology and Evolution*, *14*(1), 103–116. https://doi.org/10.1111/2041-210X.13834

Karachaliou, E., Schmidt, C., de Greef, E., Docker, M. F., & Garroway, C. J. (2025). Urbanisation Is Associated With Reduced Genetic Diversity in Marine Fish Populations. Molecular Ecology, 34(7), e17711. https://doi.org/10.1111/mec.17711

Kays, R., Davidson, S. C., Berger, M., Bohrer, G., Fiedler, W., Flack, A., Hirt, J., Hahn, C., Gauggel, D., Russell, B., Kölzsch, A., Lohr, A., Partecke, J., Quetting, M., Safi, K., Scharf, A., Schneider, G., Lang, I., Schaeuffelhut, F., … Wikelski, M. (2022). The Movebank system for studying global animal movement and demography. Methods in Ecology and Evolution, 13(2), 419–431. https://doi.org/10.1111/2041-210X.13767

Khikmah, F., Sebald, C., Metzner, M., & Schwieger, V. (2024). Modelling Vegetation Health and Its Relation to Climate Conditions Using Copernicus Data in the City of Constance. Remote Sensing, 16(4), 691. https://doi.org/10.3390/rs16040691

Kinnunen, R. P., Fraser, K. C., Schmidt, C., & Garroway, C. J. (2022). The socioeconomic status of cities covaries with avian life-history strategies. Ecosphere, 13(2), e3918. https://doi.org/10.1002/ecs2.3918

Kinnunen, R. P., Fraser, K. C., Schmidt, C., & Garroway, C. J. (2025). Structural and socioeconomic features of cities predict migratory bird species richness. *Journal of Avian Biology*, *2025*(1), e03189. https://doi.org/10.1111/jav.03189

Leigh, D. M., Vandergast, A. G., Hunter, M. E., Crandall, E. D., Funk, W. C., Garroway, C. J., Hoban, S., Oyler-McCance, S. J., Rellstab, C., Segelbacher, G., Schmidt, C., Vázquez-Domínguez, E., & Paz-Vinas, I. (2024). Best practices for genetic and genomic data archiving. *Nature Ecology &*

Evolution. https://doi.org/10.1038/s41559-024-02423-7

Leisenheimer, L., Wellmann, T., Jänicke, C., & Haase, D. (2024). Monitoring drought impacts on street trees using remote sensing—Disentangling temporal and species-specific response patterns with Sentinel-2 imagery. Ecological Informatics, 102659. https://doi.org/10.1016/j.ecoinf.2024.102659

Leong, M., & Trautwein, M. (2019). A citizen science approach to evaluating US cities for biotic homogenization. PeerJ, 7, e6879. https://doi.org/10.7717/peerj.6879

Li, D., Barve, N., Brenskelle, L., Earl, K., Barve, V., Belitz, M. W., Doby, J., Hantak, M. M., Oswald, J. A., Stucky, B. J., Walters, M., & Guralnick, R. P. (2021). Climate, urbanization, and species traits interactively drive flowering duration. Global Change Biology, 27(4), 892–903. https://doi.org/10.1111/gcb.15461

Li, E., Parker, S. S., Pauly, G. B., Randall, J. M., Brown, B. V., & Cohen, B. S. (2019). An Urban Biodiversity Assessment Framework That Combines an Urban Habitat Classification Scheme and Citizen Science Data. *Frontiers in Ecology and Evolution*, *7*. https://doi.org/10.3389/fevo.2019.00277

Li, X., Shen, X., Chen, X., Xiang, D., Murphy, R. W., & Shen, Y. (2018). Detection of Potential Problematic Cytb Gene Sequences of Fishes in GenBank. Frontiers in Genetics, 9. https://doi.org/10.3389/fgene.2018.00030

Li, X., Zhou, Y., Asrar, G. R., Mao, J., Li, X., & Li, W. (2017). Response of vegetation phenology to urbanization in the conterminous United States. Global Change Biology, 23(7), 2818–2830. https://doi.org/10.1111/gcb.13562

Lopatto, D. (2007). Undergraduate Research Experiences Support Science Career Decisions and Active Learning. CBE—Life Sciences Education, 6(4), 297–306. https://doi.org/10.1187/cbe.07-06-0039

Love, N. L. R., Bonnet, P., Goëau, H., Joly, A., & Mazer, S. J. (2021). Machine Learning Undercounts Reproductive Organs on Herbarium Specimens but Accurately Derives Their Quantitative Phenological Status: A Case Study of Streptanthus tortuosus. *Plants*, *10*(11), Article 11. https://doi.org/10.3390/plants10112471

Magle, S. B., Fidino, M., Lehrer, E. W., Gallo, T., Mulligan, M. P., Ríos, M. J., Ahlers, A. A., Angstmann, J., Belaire, A., Dugelby, B., Gramza, A., Hartley, L., MacDougall, B., Ryan, T., Salsbury, C., Sander, H., Schell, C., Simon, K., St Onge, S., & Drake, D. (2019). Advancing urban wildlife research through a multi-city collaboration. *Frontiers in Ecology and the Environment*, *17*(4), 232–239. https://doi.org/10.1002/fee.2030

Marín-Gómez, O. H., Rodríguez Flores, C., & Arizmendi, M. del C. (2022). Assessing ecological interactions in urban areas using citizen science data: Insights from hummingbird–plant meta-networks in a tropical megacity. *Urban Forestry & Urban Greening*, *74*, 127658. https://doi.org/10.1016/j.ufug.2022.127658

Meineke, E. K., & Davies, T. J. (2019). Museum specimens provide novel insights into changing plant–herbivore interactions. Philosophical Transactions of the Royal Society B: Biological Sciences, 374(1763), 20170393. https://doi.org/10.1098/rstb.2017.0393

Meineke, E. K., Tomasi, C., Yuan, S., & Pryer, K. M. (2020). Applying machine learning to investigate long-term insect–plant interactions preserved on digitized herbarium specimens. Applications in Plant Sciences, 8(6). https://doi.org/10.1002/aps3.11369

Merckx, T., Nielsen, M. E., Heliölä, J., Kuussaari, M., Pettersson, L. B., Pöyry, J., Tiainen, J., Gotthard, K., & Kivelä, S. M. (2021). Urbanization extends flight phenology and leads to local adaptation of seasonal plasticity in Lepidoptera. Proceedings of the National Academy of Sciences, 118(40), e2106006118. https://doi.org/10.1073/pnas.2106006118

Millette, K. L., Fugère, V., Debyser, C., Greiner, A., Chain, F. J. J., & Gonzalez, A. (2020). No consistent effects of humans on animal genetic diversity worldwide. Ecology Letters, 23(1), 55–67. https://doi.org/10.1111/ele.13394

Miraldo, A., Li, S., Borregaard, M. K., Flórez-Rodríguez, A., Gopalakrishnan, S., Rizvanovic, M., Wang, Z., Rahbek, C., Marske, K. A., & Nogués-Bravo, D. (2016). An Anthropocene map of genetic diversity. Science, 353(6307), 1532–1535. https://doi.org/10.1126/science.aaf4381

Nacci, D. E., Champlin, D., & Jayaraman, S. (2010). Adaptation of the Estuarine Fish Fundulus heteroclitus (Atlantic Killifish) to Polychlorinated Biphenyls (PCBs). Estuaries and Coasts, 33(4), 853–864. https://doi.org/10.1007/s12237-009-9257-6

Neate-Clegg, M. H. C., Tonelli, B. A., Youngflesh, C., Wu, J. X., Montgomery, G. A., Şekercioğlu, Ç. H., & Tingley, M. W. (2023). Traits shaping urban tolerance in birds differ around the world. Current Biology, 33(9), 1677-1688.e6. https://doi.org/10.1016/j.cub.2023.03.024

Olden, J. D., Poff, N. L., Douglas, M. R., Douglas, M. E., & Fausch, K. D. (2004). Ecological and evolutionary consequences of biotic homogenization. Trends in Ecology & Evolution, 19(1), 18–24. https://doi.org/10.1016/j.tree.2003.09.010

Olsen, K., Holm, T. E., Pape, T., & Simonsen, T. J. (2020). Natural history museum collection and citizen science data show advancing phenology of Danish hoverflies (Insecta: Diptera, Syrphidae) with increasing annual temperature. PLOS ONE, 15(5), e0232980. https://doi.org/10.1371/journal.pone.0232980

Owen, H. L., Meng, F., & Winchell, K. M. (2024). Urbanization and environmental variation drive phenological changes in the spotted lanternfly, Lycorma delicatula (Hemiptera: Fulgoridae). Biological Journal of the Linnean Society, 143(4), blae099. https://doi.org/10.1093/biolinnean/blae099

Oziolor, E. M., Bigorgne, E., Aguilar, L., Usenko, S., & Matson, C. W. (2014). Evolved resistance to PCB- and PAH-induced cardiac teratogenesis, and reduced CYP1A activity in Gulf killifish (*Fundulus grandis*) populations from the Houston Ship Channel, Texas. *Aquatic Toxicology*, *150*, 210–219. https://doi.org/10.1016/j.aquatox.2014.03.012

Pearson, K. D., Nelson, G., Aronson, M. F. J., Bonnet, P., Brenskelle, L., Davis, C. C., Denny, E. G., Ellwood, E. R., Goëau, H., Heberling, J. M., Joly, A., Lorieul, T., Mazer, S. J., Meineke, E. K., Stucky, B. J., Sweeney, P., White, A. E., & Soltis, P. S. (2020). Machine Learning Using Digitized Herbarium Specimens to Advance Phenological Research. BioScience, 70(7), 610–620. https://doi.org/10.1093/biosci/biaa044

Pernat, N., August, T., Groom, Q. J., Memedemin, D., & Reyserhove, L. (2022). An iNaturalist-

Pl@ntNet-workflow to identify plant-pollinator interactions – a case study of Isodontia mexicana. https://doi.org/10.37044/osf.io/em3rk

Phelan, P. E., Kaloush, K., Miner, M., Golden, J., Phelan, B., Iii, H. S., & Taylor, R. A. (2015). Urban Heat Island: Mechanisms, Implications, and Possible Remedies. *Annual Review of Environment and Resources*, *40*(Volume 40, 2015), 285–307. https://doi.org/10.1146/annurev-environ-102014-021155

Prudic, K., Oliver, J., Brown, B., & Long, E. (2018). Comparisons of Citizen Science Data-Gathering Approaches to Evaluate Urban Butterfly Diversity. Insects, 9(4), 186. https://doi.org/10.3390/insects9040186

Putman, B. J., Williams, R., Li, E., & Pauly, G. B. (2021). The power of community science to quantify ecological interactions in cities. Scientific Reports, 11(1), 3069. https://doi.org/10.1038/s41598-021-82491-y

Reeb, R. A., Aziz, N., Lapp, S. M., Kitzes, J., Heberling, J. M., & Kuebbing, S. E. (2022). Using Convolutional Neural Networks to Efficiently Extract Immense Phenological Data From Community Science Images. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.787407

Reid, N. M., Proestou, D. A., Clark, B. W., Warren, W. C., Colbourne, J. K., Shaw, J. R., Karchner, S. I., Hahn, M. E., Nacci, D., Oleksiak, M. F., Crawford, D. L., & Whitehead, A. (2016). The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science*, *354*(6317), 1305–1308. https://doi.org/10.1126/science.aah4993

Ringland, J., Bohm, M., Baek, S.-R., & Eichhorn, M. (2021). Automated survey of selected common plant species in Thai homegardens using Google Street View imagery and a deep neural network. Earth Science Informatics, 14(1), 179–191. https://doi.org/10.1007/s12145-020-00557-3

Russell, S. H., Hancock, M. P., & McCullough, J. (2007). Benefits of Undergraduate Research Experiences. Science, 316(5824), 548–549. https://doi.org/10.1126/science.1140384

Ružanović, L., & Mičetić Stanković, V. (2023). Urban Beetle Diversity in Natural History Collections—

A Hundred-Year Perspective. Diversity, 15(12), 1224. https://doi.org/10.3390/d15121224

Sabatini, F. M., Lenoir, J., Hattab, T., Arnst, E. A., Chytrý, M., Dengler, J., De Ruffray, P., Hennekens, S. M., Jandt, U., Jansen, F., Jiménez-Alfaro, B., Kattge, J., Levesley, A., Pillar, V. D., Purschke, O., Sandel, B., Sultana, F., Aavik, T., Aćić, S., … Bruelheide, H. (2021). sPlotOpen – An environmentally balanced, open-access, global dataset of vegetation plots. Global Ecology and Biogeography, 30(9), 1740–1764. https://doi.org/10.1111/geb.13346

Santangelo, J. S., Ness, R. W., Cohan, B., Fitzpatrick, C. R., Innes, S. G., Koch, S., Miles, L. S., Munim, S., Peres-Neto, P. R., Prashad, C., Tong, A. T., Aguirre, W. E., Akinwole, P. O., Alberti, M., Álvarez, J., Anderson, J. T., Anderson, J. J., Ando, Y., Andrew, N. R., … Johnson, M. T. J. (2022). Global urban environmental change drives adaptation in white clover. *Science*, *375*(6586), 1275–1281. https://doi.org/10.1126/science.abk0989

Santini, L., González-Suárez, M., Russo, D., Gonzalez-Voyer, A., von Hardenberg, A., & Ancillotto, L. (2019). One strategy does not fit all: Determinants of urban adaptation in mammals. Ecology Letters, 22(2), 365–376. https://doi.org/10.1111/ele.13199

Savage, A. M., Willmott, M. J., Moreno-García, P., Jagiello, Z., Li, D., Malesis, A., Miles, L. S., Román-Palacios, C., Salazar-Valenzuela, D., Verrelli, B. C., Winchell, K. M., Alberti, M., Bonilla-Bedoya, S., Carlen, E., Falvey, C., Johnson, L., Martin, E., Kuzyo, H., Marzluff, J., … Gotanda, K. M. (2024). Online toolkits for collaborative and inclusive global research in urban evolutionary ecology. *Ecology and Evolution*, *14*(6), e11633. https://doi.org/10.1002/ece3.11633

Schell, C. J., Dyson, K., Fuentes, T. L., Des Roches, S., Harris, N. C., Miller, D. S., Woelfle-Erskine, C. A., & Lambert, M. R. (2020). The ecological and evolutionary consequences of systemic racism in urban environments. Science, 369(6510), eaay4497. https://doi.org/10.1126/science.aay4497

Schmidt, C., Domaratzki, M., Kinnunen, R. P., Bowman, J., & Garroway, C. J. (2020). Continent-wide effects of urbanization on bird and mammal genetic diversity. Proceedings of the Royal Society B: Biological Sciences, 287(1920), 20192497. https://doi.org/10.1098/rspb.2019.2497

Schmidt, C., & Garroway, C. J. (2021). The population genetics of urban and rural amphibians in North

America. Molecular Ecology, 30(16), 3918–3929. https://doi.org/10.1111/mec.16005

Schmidt, C., & Garroway, C. J. (2022). Systemic racism alters wildlife genetic diversity. Proceedings of the National Academy of Sciences, 119(43), e2102860119. https://doi.org/10.1073/pnas.2102860119

Shapiro, C., Moberg-Parker, J., Toma, S., Ayon, C., Zimmerman, H., Roth-Johnson, E. A., Hancock, S. P., Levis-Fitzgerald, M., & Sanders, E. R. (2015). Comparing the Impact of Course-Based and Apprentice-Based Research Experiences in a Life Science Laboratory Curriculum. Journal of Microbiology & Biology Education, 16(2), 186–197. https://doi.org/10.1128/jmbe.v16i2.1045

Shultz, A. J., Adams, B. J., Bell, K. C., Ludt, W. B., Pauly, G. B., & Vendetti, J. E. (2020). Natural history collections are critical resources for contemporary and future studies of urban evolution. Evolutionary Applications, 14(1), 233–247. https://doi.org/10.1111/eva.13045

Siefert, A., Laughlin, D. C., & Sabatini, F. M. (2024). You shall know a species by the company it keeps: Leveraging co-occurrence data to improve ecological prediction. Journal of Vegetation Science, 35(6), e13314. https://doi.org/10.1111/jvs.13314

Stadnicki, I., Corsini, M., & Szulkin, M. (2024). Application of criminology in urban ecology and evolution: Routine Activity Theory and field equipment disappearance dynamics. *Ecological Indicators*, *165*, 112095. https://doi.org/10.1016/j.ecolind.2024.112095

Stewart, R. D., Bard, N. W., van der Bank, M., & Davies, T. J. (2025). Leveraging machine learning and citizen science data to describe flowering phenology across South Africa. *PLANTS, PEOPLE, PLANET*, *n/a*(n/a). https://doi.org/10.1002/ppp3.70059

van den Burg, M. P., & Vieites, D. R. (2023). Bird genetic databases need improved curation and error reporting to NCBI. Ibis, 165(2), 472–481. https://doi.org/10.1111/ibi.13143

Velasquez-Camacho, L., Merontausta, E., Etxegarai, M., & de-Miguel, S. (2024a). Assessing urban forest biodiversity through automatic taxonomic identification of street trees from citizen science applications and remote-sensing imagery. International Journal of Applied Earth Observation and Geoinformation, 128, 103735. https://doi.org/10.1016/j.jag.2024.103735

Visser, M. E., & Both, C. (2005). Shifts in phenology due to global climate change: The need for a yardstick. Proceedings of the Royal Society B: Biological Sciences, 272(1581), 2561–2569. https://doi.org/10.1098/rspb.2005.3356

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J.-M., Hoegh-Guldberg, O., & Bairlein, F. (2002). Ecological responses to recent climate change. Nature, 416(6879), 389–395. https://doi.org/10.1038/416389a

Weaver, W. N., & Smith, S. A. (2023). From leaves to labels: Building modular machine learning networks for rapid herbarium specimen analysis with LeafMachine2. *Applications in Plant Sciences*, *11*(5), e11548. https://doi.org/10.1002/aps3.11548

Weeks, B. C., Zhou, Z., O'Brien, B. K., Darling, R., Dean, M., Dias, T., Hassena, G., Zhang, M., & Fouhey, D. F. (2023). A deep neural network for high-throughput measurement of functional traits on museum skeletal specimens. *Methods in Ecology and Evolution*, *14*(2), 347–359. https://doi.org/10.1111/2041-210X.13864

Willis, C. G., Ellwood, E. R., Primack, R. B., Davis, C. C., Pearson, K. D., Gallinat, A. S., Yost, J. M., Nelson, G., Mazer, S. J., Rossington, N. L., Sparks, T. H., & Soltis, P. S. (2017). Old Plants, New Tricks: Phenological Research Using Herbarium Specimens. Trends in Ecology & Evolution, 32(7), 531–546. https://doi.org/10.1016/j.tree.2017.03.015

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

Winchell, K. M., Campbell-Staton, S. C., Losos, J. B., Revell, L. J., Verrelli, B. C., & Geneva, A. J. (2023). Genome-wide parallelism underlies contemporary adaptation in urban lizards. Proceedings of the National Academy of Sciences, 120(3), e2216789120. https://doi.org/10.1073/pnas.2216789120

Winchell, K. M., Losos, J. B., & Verrelli, B. C. (2023). Urban evolutionary ecology brings exaptation back into focus. Trends in Ecology & Evolution, 38(8), 719–726. https://doi.org/10.1016/j.tree.2023.03.006

Winchell, K. M., Schliep, K. P., Mahler, D. L., & Revell, L. J. (2020). Phylogenetic signal and evolutionary correlates of urban tolerance in a widespread neotropical lizard clade*. Evolution, 74(7), 1274–1288. https://doi.org/10.1111/evo.13947

Wohlfahrt, G., Tomelleri, E., & Hammerle, A. (2019). The urban imprint on plant phenology. Nature Ecology & Evolution. https://doi.org/10.1038/s41559-019-1017-9

Ye, Y., Qiu, H., Zhou, X., & Li, X. (2025). Urban landscape pattern affects plant diversity in green spaces: Evidence from street view imagery. Global Ecology and Conservation, 57, e03395. https://doi.org/10.1016/j.gecco.2024.e03395

Appendix 1

Biological Datasets covered	
iNaturalist	https://www.inaturalist.org
eBird	https://ebird.org/home
BugGuide	https://bugguide.net
Dryad	https://datadryad.org/
DataOne	https://www.dataone.org/
Birds of the World	https://birdsoftheworld.org/bow/home
iSpot	https://www.ispotnature.org/
ForgID	https://www.frogid.net.au/
GBIF	https://www.gbif.org/
Norfolk Bat Survey	http://www.nbis.org.uk/
NASA (Aeronautics and Space Administration)	https://data.nasa.gov/
ESA (European Space Agency)	https://earth.esa.int/eogateway/catalog
Copernicus	https://dataspace.copernicus.eu/
Pan-European Phenology Project	http://www.pep725.eu/
National Phenology Network	https://www.usanpn.org/
Open Tree of Life	https://tree.opentreeoflife.org/
Amniote Life-History database	https://doi.org/10.6084/m9.figshare.c.3308127.v1
AnAge	https://genomics.senescence.info/species/index.html
EtlonTrait	https://opentraits.org/datasets/elton-traits.html
PanTHERIA	https://doi.org/10.6084/m9.figshare.c.3301274.v1
VertNet	https://www.vertnet.org/
National Ecological Observatory Network	https://data.neonscience.org/
BirdTree	https://birdtree.org/
TRY	https://www.try-db.org/TryWeb/Home.php
Finnish Moth Monitoring Scheme	https://demo.gbif.org/dataset/8fd3f044-b7f5-4ff0-9d78-9b1881f4942f

Swedish and Finnish Butterfly Monitoring Scheme	https://www.gbif.org/fr/dataset/be77e203-486c-4651-91b9-8347968b728c https://www.gbif.org/dataset/181eab51-9399-4baa-a0df-8de01a3acf19
Finnish Biodiversity Info Facility	https://laji.fi/en
BOLD	https://boldsystems.org/
AVONET	https://figshare.com/s/b990722d72a26b5bfead
BirdBase	https://doi.org/10.6084/m9.figshare.27051040
Bird Life	https://datazone.birdlife.org/
Austrian Department of the Environment	https://www.dcceew.gov.au/environment/environmental-information-data/databases-applications/snes
Predictor Variables Covered	
Ocean Health Index Cumulative Human Impacts	https://oceanhealthindex.org/resources/data/
National Land Cover Databases	https://www.usgs.gov/centers/eros/science/nlcd-product-suite
Global Man-made Impervious Surface (GMIS)	https://www.earthdata.nasa.gov/data/catalog/sedac-ciesin-sedac-ulandsat-gmis-v1-1.0
2010 Global Tree Cover Dataset	https://www.earthdata.nasa.gov/data/catalog/lpcloud-gfcc30tc-003
VIIRS Nighttime Light	https://eogdata.mines.edu/products/vnl/
Human Population Density	https://doi.org/10.6084/m9.figshare.c.3890191
Human Footprint Index	https://wcshumanfootprint.org/data-access
Anthropogenic biomes or Anthromes	https://anthroecology.org/anthromes/
History database of the Global Environment (HYDE)	https://landuse.sites.uu.nl/datasets/
National Park Service Noise Pollution	https://irma.nps.gov/DataStore/Reference/Profile/2217356
US National Transportation Noise Map	https://maps.dot.gov/BTS/NationalTransportationNoiseMap/
Copernicus Land Monitoring Service (CLMS)	https://land.copernicus.eu/en
European Settlement Map	https://www.eea.europa.eu/en/datahub/datahubitem-view/42783031-34d5-4b38-b9ba-a0783723053c

Gridded Population of the World	https://www.earthdata.nasa.gov/data/catalog/sedac-ciesin-sedac-gpwv4-popdens-r11-4.11
WorldClim	https://www.worldclim.org/
CHELSA	https://www.chelsa-climate.org/
ClimateNA	https://climatena.ca/
PRISM	https://prism.oregonstate.edu/
NASA NEX-DCP30	https://www.nccs.nasa.gov/services/data-collections/land-based-products/nex-dcp30
NASA NEX-GDDP	https://www.nccs.nasa.gov/services/data-collections/land-based-products/nex-gddp
Basin Climate Characterization Model (BCM)	https://www.usgs.gov/centers/california-water-science-center/science/basin-characterization-model-bcm#overview
Environment Canda	https://climate.weather.gc.ca/
TerraClimate	https://www.climatologylab.org/terraclimate.html
MODIS	https://modis.gsfc.nasa.gov/
Landsat	https://landsat.gsfc.nasa.gov/
CORINE Land Cover	https://land.copernicus.eu/en/products/corine-land-cover
USGS Hydrography Dataset	https://www.usgs.gov/3d-hydrography-program/access-3dhp-data-products
Global Roads Open Access Dataset	https://www.earthdata.nasa.gov/data/catalog/sedac-ciesin-sedac-groads-v1-1.0
Highway Performance Monitoring System	https://www.fhwa.dot.gov/policyinformation/hpms.cfm
Mapping Inequality database (historical HOLC maps)	https://dsl.richmond.edu/panorama/redlining/
United States Census Bureau	https://data.census.gov/
Statistics Canada	https://www.statcan.gc.ca/en/start
CalEnviroScreen	https://oehha.ca.gov/calenviroscreen
Climate and Economic Justice Screening Tool (CEJST)	https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/B6ULET (archived due to website removal)
IPUMS	https://www.ipums.org/
Meteorological Satellite Program Operational Linescan System	https://www.ncei.noaa.gov/products/dmsp-operational-linescan-system
OpenPopGrid	https://openpopgrid.geodata.soton.ac.uk/

Australian Bureau of Statistics	https://www.abs.gov.au/
OpenDataSoft	https://public.opendatasoft.com/explore/dataset/geonames-all-cities-with-a-population-1000/table/?disjunctive.cou_name_en&sort=name
World Urban Areas 2012	https://purl.stanford.edu/yk247bg4748
Emerging and Underutilized datasets	
iDigBio	https://www.idigbio.org/
The Global Registry of Scientific Collections	https://scientific-collections.gbif.org/
ArcGIS Living Atlas of the World	https://livingatlas.arcgis.com/en/home/
Google Street View	https://developers.google.com/maps/documentation/streetview/get-api-key
Movebank	https://www.movebank.org/cms/movebank-main
sPlotOpen	https://idata.idiv.de/ddm/Data/ShowData/3474?version=55
Arctos	https://arctosdb.org/
National Center for Biotechnology Information	https://www.ncbi.nlm.nih.gov/
Global Genome Biodiversity Network	https://www.ggbn.org/ggbn_portal/
MorphoSource	https://www.morphosource.org/
Other helpful resources	
Opportunistic Database of Biodiversity Databases	https://earthskysea.org/biodiversity-databases/
Urban Eco-Evo Toolkit, Trait Mapping Database	https://www.urbanevoecotools.org/