How does dispersal shape the genetic patterns of animal populations in European cities? A simulation approach

Paul Savary[1], Cécile Tannier[2], Jean-Christophe Foltête[2], Marc Bourgeois[3], Gilles Vuidel[2], Aurélie Khimoun[4], Hervé Moal[5] & Stéphane Garnier[4]

[1] Department of Biology, Concordia University – Montréal (QC), Canada

[2] UMR 6049 ThéMA, Université de Franche-Comté - CNRS – Besançon, France

[3] UMR 5600 Environnement Ville Société, Université Jean Moulin Lyon 3 - CNRS – Lyon, France

[4] UMR 6282 Biogéosciences, Université de Bourgogne - CNRS – Dijon, France

[5] ARP-Astrance – Paris, France

Correspondence: paul.savary@concordia.ca

Abstract

Context and objectives

Although urbanization is a major driver of biodiversity erosion, it does not affect all species equally. The neutral genetic structure of populations in a given species is affected by both genetic drift and gene flow processes. In cities, the size of animal populations determines drift and can depend on multiple processes, whereas gene flow essentially depends on the ability of species to disperse across urban areas. Considering this, we tested whether variations in dispersal constraints alone could explain the variability of neutral genetic patterns commonly observed in urban areas. Besides, we assessed how the spatial distribution of urban green spaces (UGS) and peri-urban forests acts on these patterns.

Methods

We simulated multi-generational genetic processes in virtual populations of animal species occupying either UGS or forest areas (both considered as a virtual species habitat) within and around 325 European cities. We used three dispersal cost scenarios determining the ability of species to cross the least favorable land cover types, while maintaining population sizes constant among scenarios. We then assessed genetic diversity and genetic differentiation patterns for each city and each habitat types across the three cost scenarios.

Results

Overall, as dispersal across the least favorable land cover types was more constrained, genetic diversity decreased and genetic differentiation increased. Across scenarios, the scale and strength of the relationship between genetic differentiation and dispersal cost-distances varied substantially, alike previously observed empirical genetic patterns. Forest areas contributed more to habitat connectivity than UGS, due to their larger area and mostly peri-urban location. Hence, population-level genetic diversity was higher in forests than in UGS and genetic differentiation was higher between UGS populations than between forest populations. However, interface habitat patches allowing individuals to move between different habitat types seemed to locally buffer these contrasts by promoting gene flow.

Discussion and conclusion

Our results showed that variations in spatial patterns of dispersal, and thus gene flow, could explain the variability of empirically observed genetic patterns in urban contexts. Besides, the largest habitat areas and biodiversity sources are likely to be found in areas surrounding city centers. This should encourage urban planners to pay attention to the areas promoting dispersal movements between urban habitats (e.g., UGS) and peri-urban habitats (e.g., forests), rather than among urban habitats, when managing urban biodiversity.

Keywords: urban ecology; ecological networks; gene flow; biodiversity conservation; green infrastructures

Introduction

Most of humankind currently lives in cities and urban areas are predicted to cover three times the area they covered in 2000 by 2030 (Seto et al., 2012), and four times by 2050 (Angel et al., 2011). Urbanization is an important component of the anthropogenic pressures triggering the erosion and spatial redistribution of biodiversity (Díaz et al., 2019). Indeed, it favors the spread of invasive species, concentrates pollution, and urban lifestyles are important drivers of natural resource over-exploitation and greenhouse gas emissions (McDonald et al., 2020). Last but not least, it is a direct cause of habitat destruction and fragmentation (Beninde et al., 2015). As a result, many empirical studies have evidenced negative effects of urbanization on biodiversity (Aronson et al., 2014; Piano et al., 2020). However, this relationship is complex and variable across taxa, biological organization levels (ecosystems, species, genes), and across cities (Fidino et al., 2020). Urban ecology research is therefore needed and particularly crucial if humanity is willing to limit its impact on biodiversity and to maintain the ecosystem services it provides (Verrelli et al., 2022).

Urban ecology studies have shown that species are not all affected in the same way by urbanization (Aronson et al., 2014; Blair, 1996; Fanelli et al., 2022; Fidino et al., 2020). While some species are mostly, if not only, present in cities because they are specialist of anthropized environments (urban adapters), others cannot survive in these areas (urban avoiders). Some more ubiquitous species are present in both urban and non-urban areas. These urban tolerant species are reliable indicators of urban influences on population dynamics along rural-to-urban gradients. They may also be the most affected by the impact of urban planning on environmental conditions across such gradients.

Species-specific responses to urbanization not only affect species diversity patterns, but also explain the variability of genetic patterns observed at the intra-specific level along rural-to-urban gradients. Although rapid genetic adaptations to urbanization have been observed in several species (Santangelo et al., 2022), urbanization also shapes neutral genetic patterns in a significant manner (Fusco et al., 2021; Miles et al., 2019). For instance, urban populations of human pests can exhibit higher levels of genetic diversity than non-urban ones (Miles et al., 2018). In contrast, Khimoun et al. (2020) and Schoville et al. (2013) did not detect any significant difference in genetic diversity nor any patterns of isolation by distance pattern when studying ant and butterfly populations, respectively, in urban and non-urban settings. Similarly, the relationship between amphibian genetic diversity or differentiation and the degree of urbanization of several North American cities was not significant in the study by Schmidt and Garroway (2021). On the contrary, Delaney et al. (2010) showed that urbanization decreased genetic diversity and increased genetic differentiation in three lizard species and one bird species, mainly due to higher road density in urban areas. Likewise, Stillfried et al. (2017) showed that urban populations of wild boars exhibited lower genetic diversity levels than suburban ones in Berlin.

The variability of the neutral genetic patterns observed in urban settings stems from the demographic dynamics determining the intensity of both genetic drift and gene flow (Frankham et al., 2004; Miles et al., 2019; Munshi-South and Richardson, 2020). On the one hand, genetic drift can lead to allele loss, especially in small-sized populations. On the other hand, gene flow following successful dispersal events between populations leads to genetic exchanges. This can compensate for the diversity loss due to drift and decrease the resulting genetic differentiation between populations. Consequently, both the size of urban populations and the permeability of urban environments to individual dispersal drive genetic diversity and genetic differentiation, because they determine genetic drift and gene flow. As such, the genetic patterns described above span the whole range of patterns that are theoretically expected from variations in the respective intensity of drift and gene flow across urban and non-urban areas (Frankham et al., 2004; Hutchison and Templeton, 1999). Under strong genetic drift and low gene flow, a low genetic diversity is expected. Besides, isolation by distance patterns should not be significant, due to the prevailing role of local stochasticity on pairwise ge-

netic differentiation irrespective of pairwise distances among populations. On the contrary, under moderate genetic drift and gene flow, a significant isolation by distance pattern and a higher genetic diversity level are expected. Finally, at intermediate levels, isolation by distance patterns are expected, at least temporarily, until a pairwise distance threshold beyond which drift becomes the prevailing stochastic driver of genetic differentiation. However, because these predicted genetic patterns depend on the respective intensities of both drift and gene flow, disentangling the influence of these two processes is a complex task.

In urban tolerant species, genetic drift could explain most of the variability observed in genetic patterns given that some of these species only form small populations in urban areas (e.g., Lourenço et al. (2017)), thereby exacerbating drift effects, whereas others can maintain large populations (e.g., Miles et al. (2018)). Yet, gene flow, which is mainly driven by the ability of urban tolerant species to disperse across urban fabric, could also be the prevailing driver of genetic responses to urbanization. Accordingly, assessing how dispersal limitations and resulting gene flow reductions shape genetic patterns is crucial for several reasons. First, many biodiversity conservation programs rely on habitat connectivity conservation and restoration for maintaining genetic diversity in urban areas. They assume that dispersal limitation is the main cause of biodiversity loss. Besides, they often focus on Urban Green Spaces (thereafter referred to as UGS) as biodiversity sources, although their spatial location within cities could convert them into sink patches (Lepczyk et al., 2017; Pulliam, 1988; Verrelli et al., 2022), thus compromising the efficiency of habitat and ecological corridor restoration in urban cores. Determining to which point species movements in these areas are sufficient for preventing diversity loss is therefore needed for estimating the potential benefit of such measures. For that purpose, simulations have been commonly recommended in landscape genetics (Balkenhol et al., 2016; Munshi-South and Richardson, 2020), particularly for studying how neutral genetic patterns emerge from the interplay of drift and gene flow processes, independently of adaptive processes. They also proved to be efficient for reproducing empirically observed genetic patterns in urban areas (Rochat et al., 2017).

Considering this, our research objective in this study was to answer the following question: how does dispersal limitation explain the variability of genetic patterns in urban tolerant species? To this end, we simulated neutral genetic patterns resulting from multi-generational gene flow between populations of urban tolerant species located in both UGS and forest areas within and around 325 European cities. Using scenarios introducing variations in the ability of three virtual animal species to disperse across urban fabric, we assessed how this ability affects genetic patterns, independently from any other process. Across scenarios, genetic drift intensity was constant for a given set of populations. The simulated variations in dispersal, for each city, could thus affect genetic patterns by modifying the respective intensities of drift and gene flow. We also compared the connectivity of UGS and forest areas, and contrasted the genetic diversity and differentiation levels observed in these habitats to shed light on how the connectivity and spatial configuration of habitats drive genetic responses to dispersal scenarios in cities.

Material and methods

To answer our research questions, we needed to assess the genetic responses to urbanization of urban tolerant species having different abilities to cross artificial areas, while being equivalent in terms of population density and dispersal abilities to cross favorable areas. We also needed to assess these responses at the level of entire urban areas and to replicate the analyses to ensure that our results were not merely due to the specificity of a single study area. To meet these conditions, we implemented a simulation approach and applied it to 325 European urban areas, hereafter referred to as cities.

For each city, we first modeled the connectivity of habitat patch networks for forest-dwelling species occupying both UGS and forest patches according to three dispersal scenarios, using a graph-based approach.

Then, we simulated drift and gene flow processes in populations occupying both types of habitats, and analyzed the resulting genetic pattern at both the within-population (genetic diversity) and between-population (genetic differentiation) levels. We provide details in the following sub-sections.

Habitat connectivity analyses

Land cover data

We used urban land cover data from the Urban Atlas 2018 database of the Copernicus European agency. These land cover data are available for 788 functional urban areas, *sensu* Organisation for Economic Co-operation and Development (OECD, Moreno-Monroy et al. (2021)), counting at least 50,000 inhabitants in 38 member or partner states of the European Union. These vector land cover data include 27 land cover types at a relatively fine spatial resolution (Minimum Mapping Unit: 0.25 ha inside urban core areas, and 1 ha in surrounding rural areas). They were reclassified into 10 land cover types (Table S1) and rasterized at a spatial resolution of 5 m. The two "Forests" and "UGS" land cover types were considered for delineating habitat patches in the analyzes. "Forests" tended to be peripheral (i.e., peri-urban) whereas "UGS" were mostly located within city cores, although in each city, some patches did not conform to this pattern. UGS included public green areas predominantly dedicated to recreational use (e.g., gardens, parks), as well as suburban natural areas that are used and managed as urban parks. We obtained the spatial coordinates of the center of every city from the Open Street Map database using Nominatim and the `jsonlite` R package (Ooms, 2014). In most cases, it coincided with the city hall, which is commonly used as the center of cities (Lemoy and Caruso, 2020)(Figure 1).

We delineated the cities under study by standardizing their proportion of artificial areas in order to assess the influence of dispersal on genetic patterns in areas having the same degree of urbanization, although differing in the configuration of their urban fabric and natural areas. Delineating study areas of the same spatial extent would have led the differences among cities to mainly reflect the effect of varying densities of artificial areas (i.e., proportion of sealed area in the study area). In contrast, we here assessed species genetic responses in areas which differ in the location and shape of their urban fabric. To that end, we calculated the proportion of artificial areas in disks of increasing radius centered on the city center. These radius ranged from 5 km to 40 km, with steps of 500 m. The target proportion of artificial areas was fixed at $20\% \pm 1\%$ after preliminary analyses. This threshold allowed us to maximize the number of cities for which a radius encompassing a fixed proportion of artificial areas could be found between 5 and 40 km and to minimize the variance of this radius. Cities were included in our sample when at least 95 % of the delineated disk was covered by the Urban Atlas land cover data (Figures 1A and 1B). When required, we completed the remaining peripheral sectors with Corine Land Cover data, matching their typology with our land cover classification, as indicated in Table S2. Our method selected 325 cities and mainly excluded coastal cities and very small or very large cities for which the proportion of artificial areas was never below or above 20 %, respectively, within the range of considered radii.

Dispersal cost scenarios

We wanted to assess whether genetic responses were affected by the capacity of species to cross the least favorable areas when dispersing from one habitat patch to another. We therefore made three dispersal cost scenarios, consisting of cost values associated with land cover types and representing the cost of species movements across pixels of each land cover type (Table 1). These costs were similar to the costs used by Sahraoui et al. (2017) and Tannier et al. (2016) for modeling habitat connectivity in urban areas for forest species such as rodents (e.g., *Mustela putorius*) or birds (e.g., *Picus canus*). They were minimal (1) in habitat areas and higher in land cover types supposed to highly affect forest species movements, such as grasslands (10), agricultural areas (50) or wetlands (100). These costs were constant across scenarios. In contrast, the costs associated with water, roads and artificial land cover types increased from scenario 1 to 3. They were

5

equal to 90, 900 and 9,000 for water and roads, and to 100, 1,000 and 10,000 for artificial areas in scenarios 1, 2 and 3, respectively. This reflects the fact that built-up areas are more difficult to cross than unfavorable yet open areas such as roads or water bodies. A cost scenario based on similar assumptions has been empirically validated by Balbi et al. (2019) in an urban context. The cost variations in the least favorable land cover types for a forest species allowed us to distinguish three virtual species with different dispersal behaviors in urban areas. Nonetheless, the total cost they could endure during dispersal was kept equal for the three virtual species. This means that they had similar absolute capacities to disperse whatever the cost scenario. Yet, the spatial paths they followed when dispersing could vary substantially from one scenario to another depending on the configuration of land cover types. Using these cost scenarios, we thus aimed to isolate the effect of variations in spatial patterns of dispersal on genetic patterns across rural-to-urban gradients.

Land cover types	Scenario 1	Scenario 2	Scenario 3
Forests	1	1	1
Urban Green Spaces	1	1	1
Grasslands	10	10	10
Semi-natural areas	10	10	10
Agricultural areas	50	50	50
Other open areas	50	50	50
Wetlands	100	100	100
Water	90	900	9000
Roads	90	900	9000
Artificial areas	100	1000	10000

Table 1. Dispersal cost values associated with each land cover type according to the scenarios considered. Scenarios 1 to 3 represent an increasing aversion of forest species for dispersal movements across water, roads, and artificial areas.

Landscape graphs

We modeled habitat connectivity with landscape graphs using the Graphab 2.8 software program (Foltête et al., 2021). Landscape graphs represent habitat patch networks as sets of habitat patches (nodes) connected by potential dispersal paths (links)(Urban and Keitt, 2001). We built them using the Urban Atlas land cover data and the three dispersal cost scenarios (one graph per scenario). Each forest or UGS habitat patch above 0.25 ha was a node of the graphs. Although a single type of habitat patch (node) is most often considered in graph-based connectivity analyzes, the special feature of our analyzes was the distinction between forests and UGS, considered as two distinct types of nodes in subsequent analyzes (Savary et al., in prep.).

We computed least-cost paths between every pair of habitat patches. This simple method estimates dispersal paths by finding the ones connecting patches while minimizing the total cost accumulated when crossing pixels with specified cost values. Despite the known limitations of this method (Zeller et al., 2012), it has proved to be relevant in empirical studies measuring species movements in urban areas (e.g., Balbi et al. (2019)). For each scenario, we obtained a set of potential dispersal paths and the accumulated cost along them, i.e., the cost-distance (Figure 1C). We then created three minimum planar graphs, *sensu* Fall et al. (2007), in every urban area. In these graphs, links correspond to least-cost paths, connect neighbor patches, and are weighted by the corresponding cost-distances.

We used these graphs to assess the contribution of each habitat patch to the connectivity of the habitat network and identify whether the distribution of forest and UGS patches could drive potential source-sink dynamics. To that purpose, we computed the 'Flux' (F) connectivity metric for each patch to estimate the amount of habitat that is reachable from the focal patch. It considers the area of the other patches and the

6

cost-distances associated with the shortest path to these other patches on the graph, as follows:

$$F_i = \sum_{j=1; j \neq i}^{n} a_j e^{-\alpha \times d_{ij}}$$

with i the focal patch index, j the index of every other patch among the n habitat patches, d_{ij} the cost-distance between patches i and j, and a_j the area of each patch j.

α was set such that the probability of covering a cost-distance equivalent to an average path of 5 km (d_{5km}) is equal to 0.05, i.e., $p(d_{5km}) = e^{-\alpha d_{5km}} = 0.05$. This distance can be considered as the maximum dispersal distance of forest species with medium dispersal capacities (Sahraoui et al., 2017). To obtain d_{5km}, we considered the set of links for assessing the relationship between (1) the length in metric units of the least-cost paths not crossing the most unfavorable areas and (2) the corresponding cost-distances, using log-log linear regressions (Tournant et al., 2013). We then converted 5 km into cost units and computed the average value for each urban area and each scenario. d_{5km} averaged 20,000 cost-distance units. By using the same d_{5km} and α values for all three dispersal scenarios in this analysis and for parameterizing the genetic simulations (see below), we ensured that the virtual species we considered had similar absolute capacities to disperse whatever the cost scenario.

When computing the F metric, we distinguished cases where patches i and j were respectively either (i) both forest patches ($F_{Forest \leftrightarrow Forest}$), (ii) forest and UGS patches ($F_{Forest \leftrightarrow UGS}$), (iii) UGS and forest patches ($F_{UGS \leftrightarrow Forest}$) or (iv) both UGS patches ($F_{UGS \leftrightarrow UGS}$). For each urban area and scenario, forest patches associated with $F_{Forest \leftrightarrow UGS}$ values in the upper quartile of the distribution were considered as "Forest Interface" patches. Similarly, UGS patches associated with $F_{UGS \leftrightarrow Forest}$ values in the upper quartile of the distribution were considered as "UGS Interface" patches. Indeed, these habitat patches are the most important ones for the connectivity between several types of habitats. We expect them to be mostly, though not strictly, located at the periphery of city centers. They could therefore play a significant role in potential source-sink dynamics. Their connectivity could also drive genetic diversity transfers from less to more anthropized habitat areas, and conversely, e.g., when UGS Interface patches are connected to other UGS located closer to the city center.

Finally, we computed the Equivalent Connectivity (EC) metric for estimating the connectivity of the whole habitat patch network. This metric represents the area of the unique patch that would provide species with the same amount of reachable habitat as the whole habitat patch network, given its degree of subdivision and the resistance of the matrix (Saura et al., 2011). We used the following formula (see also Figure S1):

$$EC = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j e^{-\alpha \times d_{ij}}}$$

Given that we distinguished two types of habitat patches, we could estimate the contribution to EC of (i) the forest patches and the connections between them ($EC_{Forest.Forest}$, both i and j are forest patches), (ii) the UGS patches and the connections between them ($EC_{UGS.UGS}$, both i and j are UGS patches) and (iii) the connections between forest and UGS patches, weighted by their respective areas ($EC_{Forest.UGS}$, i is a forest patch and j a UGS patch, or conversely). To make these values comparable among cities, we standardized them by the total area of each city. We also assessed the relative value of the three EC components here considered, standardizing each of them by their sum.

Genetic simulations

Population size and location

7

The large number of patches in each city prevented us from simulating a population in all habitat patches, due to limitations in computational capacities. Furthermore, we can reasonably assume that habitat patches are not all occupied in actual metapopulations. Nevertheless, we wanted the number of populations to reflect the subdivision of habitat areas. Therefore, for each type of habitat h, the number of populations $Npop_{h,c}$ in city c ranged from 10 to 400 and was proportional to the logarithm of the number of patches of that type in the city c ($p_{h,c}$), such that:

$$Npop_{h,c} = 10 + (400 - 10)\frac{\log p_{h,c} - \log(\min_{\forall i} p_{h,i})}{\log(\max_{\forall i} p_{h,i}) - \log(\min_{\forall i} p_{h,i})}$$

with $\min_{\forall i} p_{h,i}$ and $\max_{\forall i} p_{h,i}$ the minimum and maximum number of patches of type h across all cities.

Then, for each city and habitat type, we randomly sampled a number of patches equal to $Npop_{h,c}$ among the habitat patches of type h in the city c, before populating them with individuals. As we wanted the total population in a given city to reflect the total amount of habitat of that city, the total number of individuals $Nind_c$ in each city ranged from 500 to 10,000 and was proportional to the logarithm of the total habitat area in each city (using the same formula as above). Log-transformations normalized the distributions of population and individual numbers.

Each population (i.e., sampled habitat patch) was then populated with at least 10 individuals, in a way that reflected the area of its patch. We wanted to assign larger populations to large habitat patches while ensuring that the total number of individuals across populations was equal to $Nind_c$. Thus, the numbers of individuals in each population were randomly drawn following a multinomial distribution. Each sampled patch was associated with a probability of being assigned supplementary individuals (beyond the minimum effective of 10 individuals) that was proportional to its area (see Figure S2 for an histogram of patch areas). For each city, these probabilities summed to 1. Therefore, cities with many habitat patches had many populations, cities with large habitat areas had many individuals overall, and large habitat patches on average contained many individuals. On average, there were 23 individuals per population (median: 19), and the largest population included 229 individuals in a 1000 ha patch. Consequently, population sizes varied among patches and among cities, thereby affecting drift intensities. Yet, drift intensity was constant across dispersal scenarios in every patch, making it possible to directly attribute genetic response variations across scenarios to variations in dispersal patterns.

Dispersal and reproduction parameters

Dispersal between populations depended on cost-distances computed for the three cost scenarios. 10 % of the individuals of each population could disperse from their origin population to another at each generation, over a total of 250 generations. The dispersal probability from population i to population j decreased with cost-distance, such that: $p_{d_{ij}} = e^{-\alpha d_{ij}}$. We used the same α value as specified in the previous section.

The population size was constant over time and the sex-ratio initially equal to 1. After birth, individuals potentially dispersed (see above) and could then only mate and reproduce with individuals from the same population. Each female had 3 offspring, with a sex-ratio averaging 1, and supernumerary individuals, either juveniles or adults, died to keep the local population constant. Initial genotypes were randomly generated by drawing 1 in 20 alleles, twice for each of the 20 simulated loci and for every individual. Mutations could happen at a 0.0005 rate. We carried out the simulations using `PopGenReport` R package (Adamack and Gruber, 2014).

Genetic data analyses

We wanted to assess the influence of the dispersal scenarios on the genetic responses simulated in each type of habitat. Thus, at the end of the simulations, we assessed both (i) intra-population genetic diversity (local) and (ii) inter-population genetic differentiation (pairwise), and their variations according to the scenarios

and types of habitat (i.e., Forest, UGS, and their respective "interface" patches). Besides, to gain insights into gene flow influence on spatial genetic patterns in cities, we assessed the scaling properties of isolation by landscape resistance patterns, and whether populations formed genetic clusters coinciding with the spatial structure of habitat patches. Genetic analyses were performed using the `graph4lg` package (Savary et al., 2021b) in R (R Core Team, 2020).

Genetic diversity and differentiation

We estimated genetic diversity within each population as the mean number of alleles per locus when considering all individuals, i.e., allelic richness. We then averaged allelic richness for each habitat type ("Forest", "Forest Interface", "UGS", "UGS Interface"), each scenario and each city ($n = 4 \times 3 \times 325$). In parallel, we assessed pairwise genetic differentiation using the D_{PS} index (Bowcock et al., 1994), i.e., $1-$ proportion of shared alleles. We distinguished three types of population pairs, i.e., "Forest-Forest", "Forest.UGS" and "UGS.UGS", and averaged values for each scenario and city ($n = 3 \times 3 \times 325$).

Isolation by Landscape Resistance (IBLR)

Isolation by distance (IBD) patterns have been known for a long time for providing insights into the relative influence of drift and gene flow on genetic differentiation (Hutchison and Templeton, 1999; Slatkin, 1993). We wanted to assess to which degree differences in spatial patterns of dispersal due to cost scenarios could affect the spatial genetic structure. We therefore analyzed isolation by landscape resistance patterns (IBLR) and their scaling properties. For that purpose, we iteratively computed Mantel correlation coefficients between D_{PS} and cost-distances, while filtering population pairs according to increasing cost-distance thresholds until all population pairs became included. We identified the threshold at which this coefficient was maximal and called it Distance of Maximum Correlation (DMC), following Van Strien et al. (2015). To obtain comparable values across cities, we standardized the range of the DMC between 0 and 1 by dividing it by the maximum cost-distance between population pairs. Large values are supposed to indicate cases where IBLR leads to a continuous and linear relationship between D_{PS} and cost-distances at the scale of the whole study area (i.e., the equivalent of the case I IBD pattern *sensu* Hutchison and Templeton (1999), although we did not use the exact same framework as for classical Isolation By Distance analyses (Rousset, 1997)). On the contrary, values between 0 and 1 could indicate the presence of a plateau in the relationship (case IV IBD pattern *sensu* Hutchison and Templeton (1999))(Van Strien et al., 2015). Besides, the correlation coefficient value between D_{PS} and cost-distances at the DMC showed us to what extent genetic differentiation increased due to increases in cost-distance.

Module partitions

In each city and for each scenario ($n = 325 \times 3$), we modeled population genetic structure using a genetic graph. Each node represented a population, and the links were weighted by D_{PS} values. To identify genetic clustering patterns potentially emerging from dispersal limitations between sets of habitat patches, we identified modules in these graphs using the *fast greedy* modularity algorithm by Clauset et al. (2004). This algorithm identifies the partition of populations into modules maximizing a modularity index. This index takes genetic differentiation values into account and takes large values when populations from the same module are also genetically similar.

We wanted to determine (i) whether the spatial distribution of genetic modules reflected the dispersal constraints imposed by the different cost scenarios, and (ii) whether populations from different types of habitat tended to be assigned to different genetic clusters as well. We therefore compared the genetic modules with (i) modules computed in similar population graphs with links weighted by cost-distance values instead of D_{PS} values, and (ii) the classification of habitat patches into forests or UGS. In the former case, we set the number

of modules in each spatial cost-distance graph to be equal to the optimal number of modules identified in each corresponding genetic graph. Then, we compared these partitions using the Adjusted Rand Index (ARI, Hubert and Arabie (1985)), following the method described by Savary et al. (2022). This index takes its maximal value (1) when two nodes from the same module in one graph also belong to the same module in the other graph. It is equal to 0 when two partitions are comparable to random partitions. It takes its minimal value (-1) when partitions are totally distinct, i.e., when two nodes from the same module are in different modules in the other graph.

Genetic response modeling

We modeled the different genetic responses as a function of the dispersal cost scenarios (and habitat type or type of population pairs for the allelic richness and genetic differentiation, respectively) using mixed-effects models with random intercepts at the city level. This allowed us to take into account the lack of statistical independence between simulations made for three cost scenarios in the same city. In that way, we also accounted for the fact that all cities do not have the same habitat area, number of populations and individuals, which created overall differences in genetic structure, irrespective of cost scenarios.

Because all our genetic responses did not have the same range of values and distributions, we used mixed-effects models assuming various distributions and link functions (when generalized), as explained in the Results section and Supporting Information 3. Models were fitted with a Restricted Maximum Likelihood approach using the lme4 (Bates et al., 2007) and glmmTMB (Brooks et al., 2017) R packages. The adequation between the distribution of the residuals and the models' assumptions were checked using a simulation-based approach implemented in the DHARMa R package (Hartig, 2020). We only interpreted the models whose residuals matched these assumptions. Because we modeled simulated values for which we controlled the number of replicates, we did not interpret the p-values (which took the lowest value reported by the R software program in most cases)(White et al., 2014). Mean values per fixed effect level and their confidence intervals (95 %) were obtained using emmeans (Russel, V. L., 2021).

When both the cost scenarios and the habitat type (4 levels: Forest, Forest Interface, UGS, UGS Interface) or type of pairs (3 levels: Forest-Forest, Forest-UGS, UGS-UGS) were considered as fixed effects, we included an interaction between these two variables in the models. Indeed, variations in dispersal were not supposed to affect the genetic responses similarly according to the type of habitat considered.

Assessing results' sensitivity to city size and habitat amount variations among cities

We wanted to check whether the results were consistent across cities regardless of prominent differences in terms of city size and habitat cover. To this end, we analyzed the results by considering separate groups of cities, created from the quartiles of the city sizes (study area radii) and of the total amount of habitat (UGS and forest total area). Additionally, we assessed whether allelic richness contrasts between habitat types were similar when the connectivity of UGS ($EC_{UGS.UGS}$) was larger than that of forests ($EC_{Forest.Forest}$).

Results

Habitat connectivity

The 325 cities delineated in this study were covered at 20 % $\pm$ 1 % by artificial areas and had a radius ranging from 5 to 40 km, with a mean of 10.8 km and a median of 8.5 km. Their population averaged 300,000 inhabitants in 2018 (median: 140,000, maximum: 6,000,000). In these areas, the overall amount of reachable

habitat (EC) was mainly due to forest patches and much less to UGS, with $EC_{Forest.Forest}$ largely higher than $EC_{UGS.UGS}$, and than $EC_{Forest.UGS}$, although to a lesser extent (Figure 2A). This contrast was stronger when modeling connectivity according to dispersal cost scenarios 2 and 3 (Figures S3 and S4). However, in some urban areas where forest areas are very limited, the contribution of UGS to the amount of reachable habitat was the highest, as reflected by the proportional share of $EC_{UGS.UGS}$ (Figure 2B).

Genetic diversity

The simulated genetic diversity varied substantially both among cities and among cost scenarios and habitat types (Figure 3). We modeled the mean allelic richness across populations for each city and type of habitat occupied by the populations (Forest, UGS, Forest Interface, UGS Interface), as a function of habitat type and cost scenario using a linear mixed-effects model (LMM) with random intercept at the city level. The random effects explained 78 % of the overall variance (ICC: 0.783) due to large differences among cities. The residual distribution was satisfactory, as well as the model fit (conditional R^2 = 0.91, marginal R^2 = 0.56). Although population sizes, drift intensity, dispersal rates and dispersal distances were constant in a given city from one scenario to another, the overall allelic richness decreased substantially from scenario 1 to 3 in all types of populations when taking among-cities differences into account (Table 2), as expected from Figure 3. This resulted from stronger constraints on gene flow exerted by artificial areas, roads, and water bodies, which modified the respective intensities of drift and gene flow. The effect of the interaction between cost scenarios and habitat types was much lower than their main effects (χ^2 values from Wald test: main effect of cost scenario: χ^2=16037.6, main effect of habitat type: χ^2=5960.9, interaction: χ^2=890.1). In the cost scenario 1, allelic richness was in all habitats relatively higher than in other scenarios, although it took slightly lower values in UGS (Table 2). For a given type of habitat (Forest, UGS), the highest allelic richness was observed in Interface habitats. It strongly decreased from scenario 1 to 3 but the decrease depended on the type of patches in which the populations were located. In Forest and Forest Interface patches, allelic richness was halved from scenario 1 to 3 (Forest: 10.68 to 5.93, Forest Interface: 11.68 to 6.60). It was almost divided by three in UGS Interface (10.95 to 4.19), and almost by four in UGS (9.37 to 2.39) (see CI in Table 2).

Scenario	Habitat type	Estimate	Lower.CI	Upper.CI
Sc. 1	Forest	10.68	10.41	10.96
Sc. 2	Forest	9.18	8.90	9.46
Sc. 3	Forest	5.93	5.65	6.21
Sc. 1	UGS	9.37	9.09	9.65
Sc. 2	UGS	4.73	4.45	5.01
Sc. 3	UGS	2.39	2.12	2.67
Sc. 1	Forest Interface	11.68	11.40	11.95
Sc. 2	Forest Interface	10.05	9.77	10.33
Sc. 3	Forest Interface	6.60	6.32	6.87
Sc. 1	UGS Interface	10.95	10.67	11.23
Sc. 2	UGS Interface	8.13	7.85	8.41
Sc. 3	UGS Interface	4.19	3.91	4.47

Table 2. Results of the mixed-effects model of the simulated genetic diversity. Predicted values and confidence intervals of the mean allelic richness across populations at the city level as a function of dispersal cost scenario, habitat type and their interaction. "Forest Interface" patches correspond to the forest patches most connected to UGS according to the $F_{Forest \leftrightarrow UGS}$ metric, whereas "UGS Interface" patches correspond to the UGS patches most connected to forests according to the $F_{UGS \leftrightarrow Forest}$ metric.

Genetic differentiation

The mean genetic differentiation (D_{PS}) between populations was high overall (> 0.6) and increased from 0.62 to 0.72 and 0.85 in scenarios 1, 2 and 3, respectively (Figure 4). Genetic differentiation was lower between forest populations than between UGS populations and took intermediate values between forest and UGS populations. We first used a generalized linear mixed model (GLMM) assuming a beta distribution and using a logit link function to model the D_{PS} values as a function of the cost scenario and type of population pair (i.e., Forest-Forest, Forest-UGS, UGS-UGS). We also used a LMM assuming a normal distribution. The GLMM provided a slightly better fit than the LMM. However, its residuals were slightly over-dispersed, whereas the LMM residuals had a satisfactory distribution. Both models provided similar results and we here present the LMM results. The random effects (city-level random intercepts) explained 76 % of the overall variance (ICC: 0.759) due to large differences among cities. The model fit was good (conditional R^2 = 0.90, marginal R^2 = 0.58). There were large differences among both cost scenarios and types of population pairs, after accounting for among-cities differences (Table 3). In that case also, the effect of the interaction between cost scenarios and types of population pairs was much lower than their main effects (main effect of cost scenario: χ^2 =11824.2, main effect of population pairs type: χ^2 =3714.1, interaction: χ^2 =850.0). Overall, the genetic differentiation among populations was lower in scenario 1 than in scenario 2, and the latter lower than in scenario 3, in accordance with Figure 4 (Table 3). The genetic differentiation among Forest populations was overall lower than between Forest and UGS populations, itself lower than among UGS populations. The increase in genetic differentiation from scenarios 1 to 2 and 3 was slightly lower for pairs of populations located in forests (see mean values and their CI in Table 3).

Scenario	Population pair	Estimate	Lower.CI	Upper.CI
Sc. 1	Forest-Forest	0.599	0.589	0.609
Sc. 2	Forest-Forest	0.620	0.610	0.630
Sc. 3	Forest-Forest	0.764	0.754	0.774
Sc. 1	Forest-UGS	0.628	0.618	0.638
Sc. 2	Forest-UGS	0.752	0.742	0.762
Sc. 3	Forest-UGS	0.881	0.871	0.891
Sc. 1	UGS-UGS	0.644	0.634	0.654
Sc. 2	UGS-UGS	0.801	0.790	0.811
Sc. 3	UGS-UGS	0.902	0.892	0.912

Table 3. Results of the mixed-effects model of the simulated genetic differentiation. Predicted values and confidence intervals of the DPS among populations at the city level as a function of dispersal cost scenario, type of population pairs, and their interaction.

Distance of Maximum Correlation (DMC)

IBLR relationships were very different from one dispersal cost scenario to another. In scenario 1, the DMC was overall equal to the maximum cost-distance between populations (Figure 5A), suggesting that gene flow and drift jointly influenced genetic differentiation at the scale of the whole study area. The slope of the IBLR relationship, reflected by the Mantel correlation coefficient at the DMC (Figure 5B), was less steep in this scenario, indicating that the increase of genetic differentiation was somehow limited when cost-distances increased. In contrast, in scenario 2, the mean DMC was equal to 0.47 and the corresponding correlation coefficients were high. This implies that for a subset of populations separated by cost-distances lower than a given threshold, there was a strong linear relationship between genetic distances and cost-distances reflecting significant gene flow between neighbor populations. Finally, in scenario 3, the low DMC values (Figure 5A) suggest that the IBLR relationships were weak and that genetic drift had a much stronger influence than gene flow on genetic differentiation. The mixed-effects models did not have a satisfactory fit for the DMC.

In contrast, they confirmed that the differences in the correlation coefficients measured at the DMC across cost scenarios, visible in Figure 5B, were substantial even when taking among-cities differences into account. The values predicted by the LMM were equal to 0.55 (95% CI [0.54, 0.56]) in the scenario 1, 0.81 (95% CI [0.80, 0.83]) in the scenario 2, and 0.78 (95% CI [0.76, 0.79]) in the scenario 3 (see Supporting Information 3). However, given that the distribution of the DMC values in the latter scenario reflects an absence of IBLR relationship, the corresponding correlation coefficients should not be interpreted for the scenario 3.

Population graph modularity

The modules identified in genetic graphs best coincided with the modules from similar graphs with links weighted by cost-distances in scenario 2 (Figure 6A). Indeed, mean ARI values were equal to 0.12, 0.27 and 0.00 in scenarios 1, 2 and 3, respectively (corresponding median values: 0.02, 0.25, 0.00). Similarly, in scenario 2, genetic modules best reflected the distinction between forest and UGS patches (Figure 6B). This means that, in this scenario, it was more likely for two populations belonging to the same genetic cluster to be close when considering cost-distances, and to be located in the same habitat type. Thus, the genetic structure of populations and either their spatial structure when taking dispersal constraints into account, or their habitat type classification (UGS vs forests) matched in a stronger way in this intermediate cost scenario. This was not the case when gene flow was less constrained by unfavorable areas (i.e., scenario 1) or, on the contrary, highly constrained by these areas (i.e., scenario 3). We did not interpret the results of the mixed-effects models of ARI values, because the strong heteroscedasticity and atypical distribution of their values prevented us from obtaining satisfactory models.

Consistency of analysis results among cities having different sizes and habitat amount

The heterogeneity of city sizes, due to the varying radius of the disks containing 20 % of artificial areas, had negligible effects on the results. In the supplementary materials, we provide the same figures as included in this section plotted separately after splitting the urban areas in 4 quartiles based on their total area (Figures S5, S6, S7, S8, S9, and S10). Similarly, the results were highly consistent when considering cities including varying total habitat areas (Figures S11, S12, S13, and S14). Finally, in the few cities where the connectivity of UGS was higher than that of forests due to small forest areas, allelic richness contrasts were consistent with the ones observed for the whole set of cities in scenarios 2 and 3 (Table S3). However, in these cities, under the scenario 1, UGS and UGS interface patches tended to maintain higher diversity levels than forest patches.

Discussion

We simulated the genetic structure of urban tolerant forest species occupying forests and urban green spaces in 325 European cities, while varying species abilities to cross the least favorable areas. We thereby found that in urban contexts, variations in dispersal movement behaviors alone can shape highly variable genetic diversity contrasts between habitat types and isolation by landscape resistance patterns. The substantial variations in simulated genetic responses between forests and urban green spaces could be due to their connectivity differences, reflecting their respective extent, spatial configuration and location within the urban matrix. These results were obtained without making any assumption regarding the respective quality of these habitats in our simulations. Urban ecologists should thus bear in mind the strong influence that dispersal between urban habitats exhibiting different spatial distributions can have on genetic patterns when assessing the relative influence of dispersal, adaptation, resource availability or biotic interactions on species responses to urban environments. Our results also provide insights into connectivity modeling and biodiversity conservation in these contexts.

13

Variations in dispersal constraints can shape highly contrasted genetic patterns

In our simulations, as dispersal across the least favorable land cover types became more costly, genetic diversity tended to decrease in both forest and UGS, and genetic differentiation tended to increase between populations. Although overall genetic responses, such as the mean allelic richness, differed among cities, variations among cost scenarios for a given city were relatively consistent. Contrasts were mainly due to differences of total habitat areas and number of patches, which determined the number of populations and individuals, and consequently the intensity of drift in each city. In urban contexts, differences in effective population sizes among species are known for being an important driver of the variability of their genetic responses (Schmidt et al., 2020), and are frequently invoked as the main driver of genetic diversity (Miles et al., 2019). However, these population size differences do not explain the strongly contrasted genetic responses among cost scenarios we obtained for a given city because the number of individuals in each population stayed constant whatever the scenario. Thus, only gene flow variations can explain these contrasts.

The resistance of artificial areas, roads and water bodies to species dispersal varied according to the cost scenario, but the total cost-distance that a fixed proportion of individuals could cover at each generation did not vary. In other words, the scenarios essentially modified the spatial pattern of dispersal movements, but not the effective distance they could cover. Consequently, for a given number of dispersing individuals, dispersal paths crossing unfavorable land cover types became less likely to be followed from scenario 1 to 3 because the cost of paths crossing other land cover types (e.g., grasslands, agricultural areas) remained the same whatever the scenario. Our results thus reflect the potential genetic responses of several species having the same individual density within patches, the same absolute dispersal abilities, and dispersal rates, but different dispersal behaviors in urban environments. It is noteworthy that simulated variations in the spatial pattern of dispersal can reproduce the inter-specific variability of patterns of genetic diversity and differentiation commonly observed in empirical landscape genetic studies (as reviewed by Fusco et al. (2021) and Miles et al. (2019)). For instance, the very subtle differences in genetic diversity simulated under the scenario 1 recall the empirical results of Khimoun et al. (2020). In contrast, the sharper differences simulated under the scenario 3 are akin to the significant genetic contrasts between urban and non-urban settings observed by Delaney et al. (2010) in several species. Although our results do not provide an explanation for these specific empirical findings, they show that differences in dispersal patterns could be sufficient to generate similar genetic patterns.

When drift intensity is constant but the spatial pattern of gene flow changes, the relative influence of drift and gene flow on genetic differentiation is logically modified in a different way according to the respective location of populations. Variations in the spatial range at which the relationship between genetic differentiation and cost-distances was the strongest revealed how contrasted genetic differentiation patterns were across scenarios. Indeed, when gene flow was not strongly restricted across artificial areas, roads and water bodies, genetic differentiation only increased progressively as cost-distance increased, producing a continuous pattern of genetic variations at the scale of the entire urban areas. This pattern is similar to the case I IBD pattern, according to the typology by Hutchison and Templeton (1999), and suggests that stepping-stone dispersal movements can prevent strong genetic differentiation in species with good abilities to disperse within cities. On the contrary, this pattern was only apparent between subsets of populations in scenario 2, and genetic differentiation increased more strongly with cost-distances. This corresponds to the case IV IBD pattern, according to Hutchison and Templeton (1999), and indicates that considering species having intermediate abilities to move across cities, the continuous and progressive increase of genetic differentiation resulting from dispersal limitation is only observed at small scales, within well connected subsets of populations. Finally, when dispersal movements in unfavorable areas were highly constrained, the relationship between genetic differentiation and cost-distances flattened out because drift became the main driver of the genetic response, and not gene flow anymore, which somehow recalls the case III IBD pattern described by Hutchison and Tem-

14

pleton (1999). In sum, changes in dispersal cost scenarios led to contrasted genetic differentiation patterns because they modified the relative frequency of dispersal events over each path connecting two populations. In other words, they rewired dispersal networks. The consequences of these changes in dispersal spatial patterns reinforce previous calls for a better consideration of population network topology in landscape genetics (Savary et al., 2021a; Van Strien, 2017).

In landscape genetics, unbridgeable barriers to dispersal are commonly inferred by identifying landscape features separating spatially structured genetic clusters of populations (Manel et al., 2007; Safner et al., 2011). Interestingly, our results show that genetic modules best reflected the spatial structure of genetic differentiation patterns in the second cost scenario, and not in the third one, which exerted the strongest constraints on dispersal movements. We could have expected the opposite and this suggests that there is an analytical limit in our ability to detect barriers to dispersal when gene flow is so constrained that drift is the main driver of genetic differentiation.

Several studies have evidenced that other processes than gene flow drive different population dynamics and individual fitness in urban populations as compared with non-urban ones. For example, urban bird populations can feed on more diverse food items, sometimes at the expense of their quality (Sinkovics et al., 2021), and can exhibit shifted and/or more variable phenotypes (Thompson et al., 2021), due to plasticity or genetic adaptations. If our results do not deny the existence of these well-known processes, they nonetheless call for a better consideration of dispersal spatial patterns when inferring the respective influence of different drivers of population genetic structure in urban areas.

Influence of the spatial distribution of multiple urban habitat types

Forest populations tended to maintain a higher genetic diversity than UGS populations and to be less differentiated than pairs of UGS populations. Besides, genetic differentiation levels measured between these two types of habitats were intermediate, as compared with the high and low levels measured within UGS and forest patches, respectively. Similar genetic responses to urbanization have already been empirically observed in several species, from birds and reptiles (Delaney et al., 2010), to rodents (DeMarco et al., 2021; Gortat et al., 2015) and larger mammals (Stillfried et al., 2017; Wandeler et al., 2003). In our simulations, they mainly stem from the fact that the contribution of forest areas to the overall amount of reachable habitat was much larger than that of UGS in most cities. Besides, UGS patches are usually smaller and also harder to reach due to their location within the urban fabric. This explains why, except in scenario 1, even in cities where UGS contributed more to the amount of reachable habitat than forests, similar genetic contrasts were observed. These differences in terms of area and connectivity between UGS and surrounding natural areas provide a likely explanation to previous empirical observations in urban landscape genetics, as habitat amount and connectivity are often mentioned as key drivers of urban biodiversity (Beninde et al., 2015).

The stronger relative isolation of UGS was also apparent in the genetic clustering pattern. In the second cost scenario, we observed that forest and UGS populations tended to form separate genetic clusters. This sub-structuration of genetic differentiation patterns within the urban core areas had also been empirically evidenced in striped field mouse populations in Warsaw (Gortat et al., 2015) or in wild boar populations in Berlin (Stillfried et al., 2017).

However, we also observed that the presence of habitat patches well connected to patches of another habitat type seemed to locally buffer these differences among habitat types by promoting gene flow at the interface between forest and UGS. Therefore, the connectivity analyzes and genetic simulations together suggest that peri-urban and less anthropized areas can be important sources of biodiversity in cities when they are connected to intra-urban habitat patches, in accordance with previous simulations (Snep et al., 2006) and

15

empirical observations (Stillfried et al., 2017). The corollary of this is the potential sink role of UGS, as previously raised by Lepczyk et al. (2017) and Verrelli et al. (2022). This could have a negative influence on the long-term persistence and genetic adaptation of wild populations both within and outside cities, and remains to be investigated.

Implications for biodiversity management in urban areas

Making urban planning policies compatible with the conservation of biodiversity is crucial. These policies are commonly based on biodiversity surveys and on the conservation of so-called "green infrastructures", including several types of natural areas, UGS and the corridors connecting them. Our results stress several points that should deserve more attention in that context. First, genetic diversity and connectivity differences between forest and UGS were substantial regardless of the spatial extent of the city under consideration. In most cases, the largest habitat areas and biodiversity levels are to be found in natural areas surrounding city centers and not in UGS. This should encourage planners to consider large areas including the most biodiverse and favorable places to wildlife, often located at the periphery of cities.

Our connectivity analyzes and genetic simulations in 325 European cities also suggest that urban planners should identify habitat interface areas and consider them as "gateways" through which species can move from less to more anthropized habitats, as suggested by Gortat et al. (2017). We could also expect these areas to play a crucial role for maintaining species diversity within cities as long as gene flow and drift effects affect single species genetic structure in a comparable way as colonization and extinction processes affect species diversity (Vellend and Geber, 2005).

Finally, our results confirm that species which can hardly move across artificial areas and roads will not maintain high levels of genetic diversity within cities. This can explain why some species are very rarely observed in urban areas, but this could also mean that urban populations of some species are already engaged in a local extinction vortex. As such, considering the long-term effect of urbanization on genetic structure and its potential consequences for population persistence is key for biodiversity management (Sarrazin and Lecomte, 2016), and particularly in urban areas where these processes can be fast (Szulkin et al., 2020).

Limitations and perspectives

Our analyzes focused only on urban tolerant forest species occupying forest and UGS. These species only represent a small proportion of urban biodiversity. Reproducing these analyses by considering either specialist species using another type of habitat or more generalist species could help obtain a broader picture of biodiversity dynamics in urban settings. Besides, most forest patches were located at the periphery of cities, whereas UGS were more central, and this peculiar spatial distribution largely affected our results. Assessing how the spatial distribution of other types of urban and/or peri-urban habitats affects genetic patterns would also be needed.

For comparison purposes, we here assumed that UGS mapped by the Urban Atlas database were suitable for urban tolerant forest species. However, habitat patches may not all be suitable for these species and many other patches could fit their needs within cities (e.g., private wooded backyards in residential areas). This limitation should encourage the use of fine-grained remote sensing data to map urban habitats with more accuracy. Yet, acquiring such maps for a number of cities providing sufficient statistical power in a standardized way remains a challenge.

Although the genetic simulations we carried out can help prioritize habitat patches in urban planning context, they only reflect the potential genetic responses of a single species. These simulations could be advanta-

geously completed by empirical surveys for assessing whether they closely reflect actual ecological processes. Our simulation framework could also be implemented in areas where genetic data are already available for testing whether dispersal limitations can explain the empirically observed genetic structure. Besides, adding to the simulation local habitat features, and how they locally influence several species based on their niche optimum, breadth, and competitive interactions (alike in the meta-community simulation framework of Thompson et al. (2020)) could provide insights into species diversity patterns in urban areas. This could help determine whether connectivity restoration measures are always positive for biodiversity conservation in urban areas; which may not always be guaranteed when invasive species also benefit from them.

Finally, although substantial differences among urban areas in terms of reachable amounts of habitat (i.e., connectivity) probably explain most of the variability in genetic responses across cities, we did not investigate the structural causes of these differences. Indeed, the interplay of urban form with the spatial configuration of forest and UGS likely determines habitat connectivity patterns and species genetic structure. Further research is needed for understanding these relationships and providing broad guidance on urban planning at a time when increased urbanization and biodiversity conservation too often seem incompatible.

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Data, script, code, and supporting information availability

Data and codes are available online: https://doi.org/10.5281/zenodo.8180746. Supporting information is also available online: https://doi.org/10.5281/zenodo.7926345

References

Adamack, A. T. and Gruber, B. (2014). PopGenReport: simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution*, 5(4):384–387.

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

Aronson, M. F., La Sorte, F. A., Nilon, C. H., Katti, M., Goddard, M. A., Lepczyk, C. A., Warren, P. S., Williams, N. S., Cilliers, S., Clarkson, B., et al. (2014). A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. *Proceedings of the Royal Society B*, 281(1780):20133330.

Balbi, M., Petit, E. J., Croci, S., Nabucet, J., Georges, R., Madec, L., and Ernoult, A. (2019). Ecological relevance of least cost path analysis: An easy implementation method for landscape urban planning. *Journal of Environmental Management*, 244:61–68.

Balkenhol, N., Cushman, S., Storfer, A., and Waits, L. (2016). *Landscape genetics: concepts, methods, applications*. John Wiley & Sons.

Bates, D., Sarkar, D., Bates, M. D., and Matrix, L. (2007). The lme4 package. *R package version*, 2(1):74.

Beninde, J., Veith, M., and Hochkirch, A. (2015). Biodiversity in cities needs space: a meta-analysis of factors determining intra-urban biodiversity variation. *Ecology letters*, 18(6):581–592.

Blair, R. B. (1996). Land use and avian species diversity along an urban gradient. *Ecological applications*, 6(2):506–519.

Bowcock, A. M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J. R., and Cavalli-Sforza, L. L. (1994). High resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, 368(6470):455–457.

Brooks, M. E., Kristensen, K., Van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Machler, M., and Bolker, B. M. (2017). glmmtmb balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal*, 9(2):378–400.

Clauset, A., Newman, M. E., and Moore, C. (2004). Finding community structure in very large networks. *Physical review E*, 70(6):1–6.

Delaney, K. S., Riley, S. P., and Fisher, R. N. (2010). A rapid, strong, and convergent genetic response to urban habitat fragmentation in four divergent and widespread vertebrates. *PLoS ONE*, 5(9):e12767.

DeMarco, C., Cooper, D. S., Torres, E., Muchlinski, A., and Aguilar, A. (2021). Effects of urbanization on population genetic structure of western gray squirrels. *Conservation Genetics*, 22(1):67–81.

Díaz, S. M., Settele, J., Brondízio, E., Ngo, H., Guèze, M., Agard, J., Arneth, A., Balvanera, P., Brauman, K., Butchart, S., et al. (2019). The global assessment report on biodiversity and ecosystem services: Summary for policy makers. Technical report, Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services.

Fall, A., Fortin, M.-J., Manseau, M., and O'Brien, D. (2007). Spatial graphs: principles and applications for habitat connectivity. *Ecosystems*, 10(3):448–461.

Fanelli, R. E., Martin, P. R., Robinson, O. J., and Bonier, F. (2022). Estimates of species-level tolerance of urban habitat in north american birds.

Fidino, M., Gallo, T., Lehrer, E. W., Murray, M. H., Kay, C., Sander, H. A., MacDougall, B., Salsbury, C. M., Ryan, T. J., Angstmann, J. L., et al. (2020). Landscape-scale differences among cities alter common species' responses to urbanization. *Ecological Applications*, 2253.

Foltête, J.-C., Vuidel, G., Savary, P., Clauzel, C., Sahraoui, Y., Girardet, X., and Bourgeois, M. (2021). Graphab: an application for modeling and managing ecological habitat networks. *Software Impacts*, 8:100065.

Frankham, R., Ballou, J. D., and Briscoe, D. A. (2004). *A primer of conservation genetics*. Cambridge University Press.

Fusco, N. A., Carlen, E. J., and Munshi-South, J. (2021). Urban landscape genetics: Are biologists keeping up with the pace of urbanization? *Current Landscape Ecology Reports*, pages 1–11.

Gortat, T., Rutkowski, R., Gryczynska, A., Kozakiewicz, A., and Kozakiewicz, M. (2017). The spatial genetic structure of the yellow-necked mouse in an urban environment–a recent invader vs. a closely related permanent inhabitant. *Urban Ecosystems*, 20(3):581–594.

Gortat, T., Rutkowski, R., Gryczyńska, A., Pieniążek, A., Kozakiewicz, A., and Kozakiewicz, M. (2015). Anthropopressure gradients and the population genetic structure of apodemus agrarius. *Conservation Genetics*, 16(3):649–659.

Hartig, F. (2020). Dharma: residual diagnostics for hierarchical (multi-level/mixed) regression models. *R package version 0.3*, 3(5).

Hubert, L. and Arabie, P. (1985). Comparing partitions. *Journal of classification*, 2(1):193–218.

Hutchison, D. W. and Templeton, A. R. (1999). Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, 53(6):1898–1914.

Khimoun, A., Doums, C., Molet, M., Kaufmann, B., Peronnet, R., Eyer, P., and Mona, S. (2020). Urbanization without isolation: the absence of genetic structure among cities and forests in the tiny acorn ant temnothorax nylanderi. *Biology letters*, 16(1):20190741.

Lemoy, R. and Caruso, G. (2020). Evidence for the homothetic scaling of urban forms. *Environment and Planning B: Urban Analytics and City Science*, 47(5):870–888.

Lepczyk, C. A., Aronson, M. F., Evans, K. L., Goddard, M. A., Lerman, S. B., and MacIvor, J. S. (2017). Biodiversity in the city: fundamental questions for understanding the ecology of urban green spaces for biodiversity conservation. *BioScience*, 67(9):799–807.

Lourenço, A., Álvarez, D., Wang, I. J., and Velo-Antón, G. (2017). Trapped within the city: Integrating demography, time since isolation and population-specific traits to assess the genetic effects of urbanization. *Molecular Ecology*, 26(6):1498–1514.

Manel, S., Berthoud, F., Bellemain, E., Gaudeul, M., Luikart, G., Swenson, J., Waits, L., Taberlet, P., and Consortium, I. (2007). A new individual-based spatial approach for identifying genetic discontinuities in natural populations. *Molecular Ecology*, 16(10):2031–2043.

McDonald, R. I., Mansur, A. V., Ascensão, F., Crossman, K., Elmqvist, T., Gonzalez, A., Güneralp, B., Haase, D., Hamann, M., Hillel, O., et al. (2020). Research gaps in knowledge of the impact of urban growth on biodiversity. *Nature Sustainability*, 3(1):16–24.

Miles, L. S., Dyer, R. J., and Verrelli, B. C. (2018). Urban hubs of connectivity: Contrasting patterns of gene flow within and among cities in the western black widow spider. *Proceedings of the Royal Society B*, 285(1884):20181224.

Miles, L. S., Rivkin, L. R., Johnson, M. T., Munshi-South, J., and Verrelli, B. C. (2019). Gene flow and genetic drift in urban environments. *Molecular Ecology*, 28(18):4138–4151.

Moreno-Monroy, A. I., Schiavina, M., and Veneri, P. (2021). Metropolitan areas in the world. delineation and population trends. *Journal of Urban Economics*, 125:103242.

Munshi-South, J. and Richardson, J. (2020). Landscape genetic approaches to understanding movement and gene flow in cities. In Szulkin, M., Munshi-South, J., and Charmantier, A., editors, *Urban evolutionary biology*, pages 54–73. Oxford University Press, USA.

Ooms, J. (2014). The jsonlite package: A practical and consistent mapping between json data and r objects. *arXiv preprint arXiv:1403.2805*.

Piano, E., Souffreau, C., Merckx, T., Baardsen, L. F., Backeljau, T., Bonte, D., Brans, K. I., Cours, M., Dahirel, M., Debortoli, N., et al. (2020). Urbanization drives cross-taxon declines in abundance and diversity at multiple spatial scales. *Global Change Biology*, 26(3):1196–1211. [716] [717] [718]

Pulliam, H. R. (1988). Sources, sinks, and population regulation. *The American Naturalist*, 132(5):652–661. [719]

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. [720] [721]

Rochat, E., Manel, S., Deschamps-Cottin, M., Widmer, I., and Joost, S. (2017). Persistence of butterfly populations in fragmented habitats along urban density gradients: motility helps. *Heredity*, 119(5):328–338. [722] [723]

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, 145(4):1219–1228. [724] [725]

Russel, V. L. (2021). *emmeans: Estimated marginal means, aka least-squares means*. CRAN R package version 1.5.4. [726] [727]

Safner, T., Miller, M. P., McRae, B. H., Fortin, M.-J., and Manel, S. (2011). Comparison of Bayesian clustering and edge detection methods for inferring boundaries in landscape genetics. *International Journal of Molecular Sciences*, 12(2):865–889. [728] [729] [730]

Sahraoui, Y., Foltête, J.-C., and Clauzel, C. (2017). A multi-species approach for assessing the impact of land-cover changes on landscape connectivity. *Landscape Ecology*, 32(9):1819–1835. [731] [732]

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., et al. (2022). Global urban environmental change drives adaptation in white clover. *Science*, 375(6586):1275–1281. [733] [734] [735]

Sarrazin, F. and Lecomte, J. (2016). Evolution in the anthropocene. *Science*, 351(6276):922–923. [736]

Saura, S., Estreguil, C., Mouton, C., and Rodríguez-Freire, M. (2011). Network analysis to assess landscape connectivity trends: application to European forests (1990–2000). *Ecological Indicators*, 11(2):407–416. [737] [738]

Savary, P., Clauzel, C., Foltête, J.-C., Vuidel, G., Girardet, X., Bourgeois, M., Martin, F.-M., Ropars, L., and Garnier, S. (Preprint). Multiple habitat graphs: how connectivity brings forth landscape ecological processes. *Submitted to Ecological Applications*. [739] [740] [741]

Savary, P., Foltête, J.-C., Moal, H., and Garnier, S. (2022). Combining landscape and genetic graphs to address key issues in landscape genetics. *Landscape Ecology*, X(X):1–17. [742] [743]

Savary, P., Foltête, J.-C., Moal, H., Vuidel, G., and Garnier, S. (2021a). Analysing landscape effects on dispersal networks and gene flow with genetic graphs. *Molecular Ecology Resources*, 21(4):1167–1185. [744] [745]

Savary, P., Foltête, J.-C., Moal, H., Vuidel, G., and Garnier, S. (2021b). graph4lg: a package for constructing and analysing graphs for landscape genetics in R. *Methods in Ecology and Evolution*, 12(3):539–547. [746] [747]

Schmidt, C., Domaratzki, M., Kinnunen, R., Bowman, J., and Garroway, C. (2020). Continent-wide effects of urbanization on bird and mammal genetic diversity. *Proceedings of the Royal Society B*, 287(1920):20192497. [748] [749]

Schmidt, C. and Garroway, C. J. (2021). The population genetics of urban and rural amphibians in North America. *Molecular Ecology*. [750] [751]
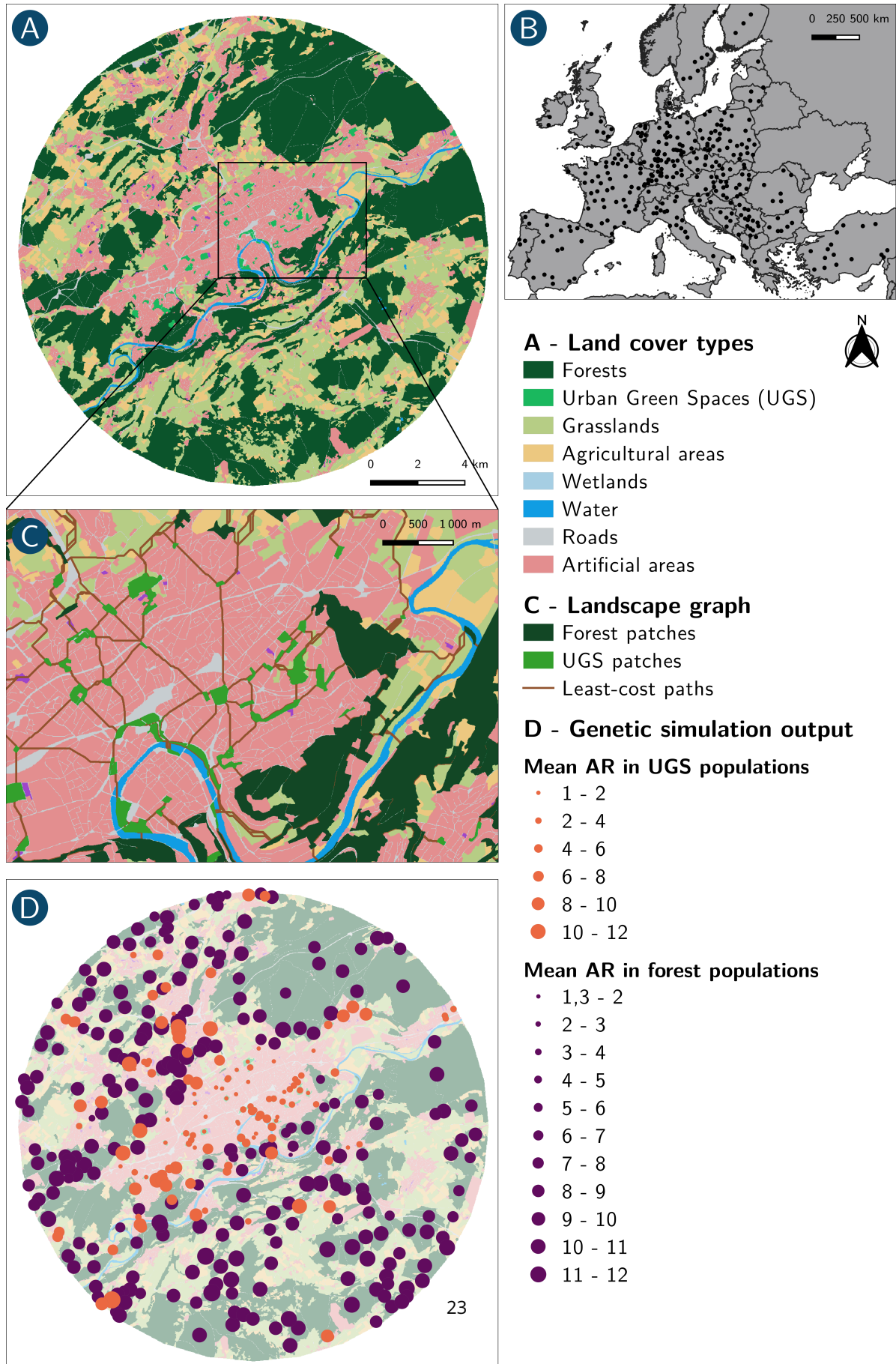
Schoville, S. D., Widmer, I., Deschamps-Cottin, M., and Manel, S. (2013). Morphological clines and weak drift along an urbanization gradient in the butterfly, *Pieris rapae*. *PLoS ONE*, 8(12):e83095. [752] [753]

Seto, K. C., Güneralp, B., and Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences*, 109(40):16083–16088.

Sinkovics, C., Seress, G., Pipoly, I., Vincze, E., and Liker, A. (2021). Great tits feed their nestlings with more but smaller prey items and fewer caterpillars in cities than in forests. *Scientific reports*, 11(1):1–13.

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, 47(1):264–279.

Snep, R. P., Opdam, P., Baveco, J., Wallis De Vries, M., Timmermans, W., Kwak, R., and Kuypers, V. (2006). How peri-urban areas can strengthen animal populations within cities: A modeling approach. *Biological Conservation*, 127(3):345–355.

Stillfried, M., Fickel, J., Börner, K., Wittstatt, U., Heddergott, M., Ortmann, S., Kramer-Schadt, S., and Frantz, A. C. (2017). Do cities represent sources, sinks or isolated islands for urban wild boar population structure? *Journal of Applied Ecology*, 54(1):272–281.

Szulkin, M., Munshi-South, J., and Charmantier, A. (2020). *Urban evolutionary biology*. Oxford University Press, USA.

Tannier, C., Bourgeois, M., Houot, H., and Foltête, J.-C. (2016). Impact of urban developments on the functional connectivity of forested habitats: a joint contribution of advanced urban models and landscape graphs. *Land Use Policy*, 52:76–91.

Thompson, M., Capilla-Lasheras, P., Dominoni, D., Réale, D., and Charmantier, A. (2021). Phenotypic variation in urban environments: mechanisms and implications. *Trends in Ecology & Evolution*, 37(2):171–182.

Thompson, P. L., Guzman, L. M., De Meester, L., Horváth, Z., Ptacnik, R., Vanschoenwinkel, B., Viana, D. S., and Chase, J. M. (2020). A process-based metacommunity framework linking local and regional scale community ecology. *Ecology Letters*, 23(9):1314–1329.

Tournant, P., Afonso, E., Roué, S., Giraudoux, P., and Foltête, J.-C. (2013). Evaluating the effect of habitat connectivity on the distribution of lesser horseshoe bat maternity roosts using landscape graphs. *Biological Conservation*, 164:39–49.

Urban, D. and Keitt, T. (2001). Landscape connectivity: a graph-theoretic perspective. *Ecology*, 82(5):1205–1218.

Van Strien, M. J. (2017). Consequences of population topology for studying gene flow using link-based landscape genetic methods. *Ecology and Evolution*, 7(14):5070–5081.

Van Strien, M. J., Holderegger, R., and Van Heck, H. J. (2015). Isolation-by-distance in landscapes: considerations for landscape genetics. *Heredity*, 114(1):27–37.

Vellend, M. and Geber, M. A. (2005). Connections between species diversity and genetic diversity. *Ecology Letters*, 8(7):767–781.

Verrelli, B. C., Alberti, M., Des Roches, S., Harris, N. C., Hendry, A. P., Johnson, M. T., Savage, A. M., Charmantier, A., Gotanda, K. M., Govaert, L., et al. (2022). A global horizon scan for urban evolutionary ecology. *Trends in Ecology & Evolution*.

Wandeler, P., Funk, S. M., Largiader, C., Gloor, S., and Breitenmoser, U. (2003). The city-fox phenomenon: Genetic consequences of a recent colonization of urban habitat. *Molecular Ecology*, 12(3):647–656.

White, J. W., Rassweiler, A., Samhouri, J. F., Stier, A. C., and White, C. (2014). Ecologists should not use statistical 793
significance tests to interpret simulation model results. *Oikos*, 123(4):385–388. 794

Zeller, K. A., McGarigal, K., and Whiteley, A. R. (2012). Estimating landscape resistance to movement: a review. 795
*Landscape Ecology*, 27(6):777–797. 796

Figure 1. (A) Land cover map of a city under consideration (city of Besançon, East of France), represented by a disc including 20% artificial areas. (B) The 325 European cities considered in our analyzes. (C) Example of least-cost paths (brown lines) connecting forest patches (dark green) and UGS patches (light green). (D) Example of genetic simulation output. The purple and orange dot sizes represent the simulated allelic richness in forest and UGS populations, respectively, according to the cost scenario 2.

A - Land cover types

- Forests
- Urban Green Spaces (UGS)
- Grasslands
- Agricultural areas
- Wetlands
- Water
- Roads
- Artificial areas

C - Landscape graph

- Forest patches
- UGS patches
- Least-cost paths

D - Genetic simulation output

Mean AR in UGS populations

- 1 - 2
- 2 - 4
- 4 - 6
- 6 - 8
- 8 - 10
- 10 - 12

Mean AR in forest populations

- 1,3 - 2
- 2 - 3
- 3 - 4
- 4 - 5
- 5 - 6
- 6 - 7
- 7 - 8
- 8 - 9
- 9 - 10
- 10 - 11
- 11 - 12

23

Figure 2. Absolute and relative values of the different EC components computed in the 325 cities according to cost scenario 1. (A) $EC_{Forest.Forest}$ (green), $EC_{Foret.UGS}$ (purple) and $EC_{UGS.UGS}$ (orange) divided by the total study area, for each city. (B) Respective contributions of $EC_{Forest.Forest}$ (green), $EC_{Forest.UGS}$ (purple) and $EC_{UGS.UGS}$ (orange) to the connectivity of the habitat network. The total connectivity value of the network is the sum of the three EC components, which is slightly different from the global EC value because of square root properties.
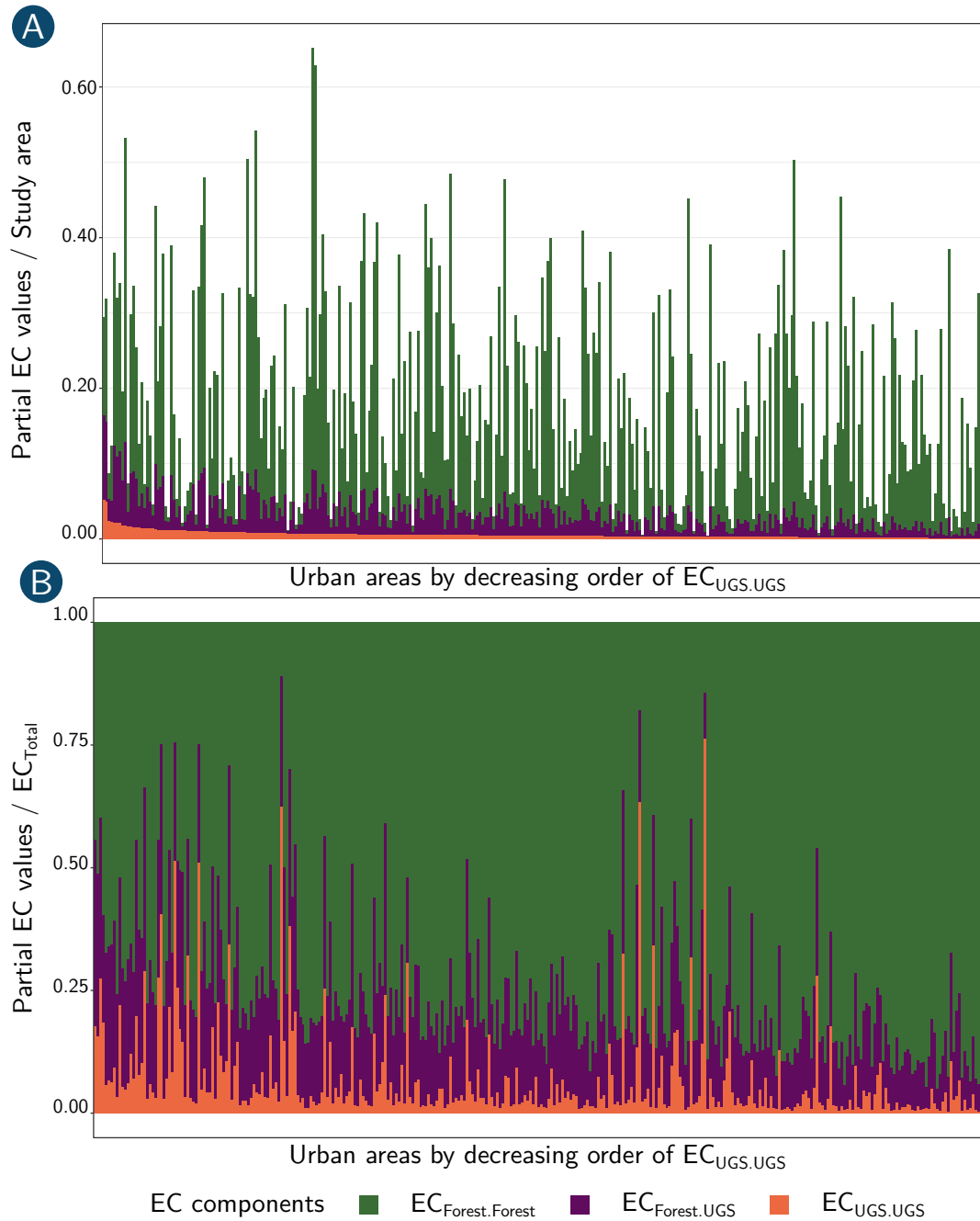
Figure 3. Distribution of the mean allelic richness of populations located in "Forest", "Forest Interface", "UGS" and "UGS Interface" patches in the 325 cities for the three dispersal cost scenarios. "Forest Interface" patches correspond to the forest patches most connected to UGS according to the $F_{Forest \rightarrow UGS}$ metric, whereas "UGS Interface" patches correspond to the UGS patches most connected to forests according to the $F_{UGS \rightarrow Forest}$ metric. $n = 325$ values per box
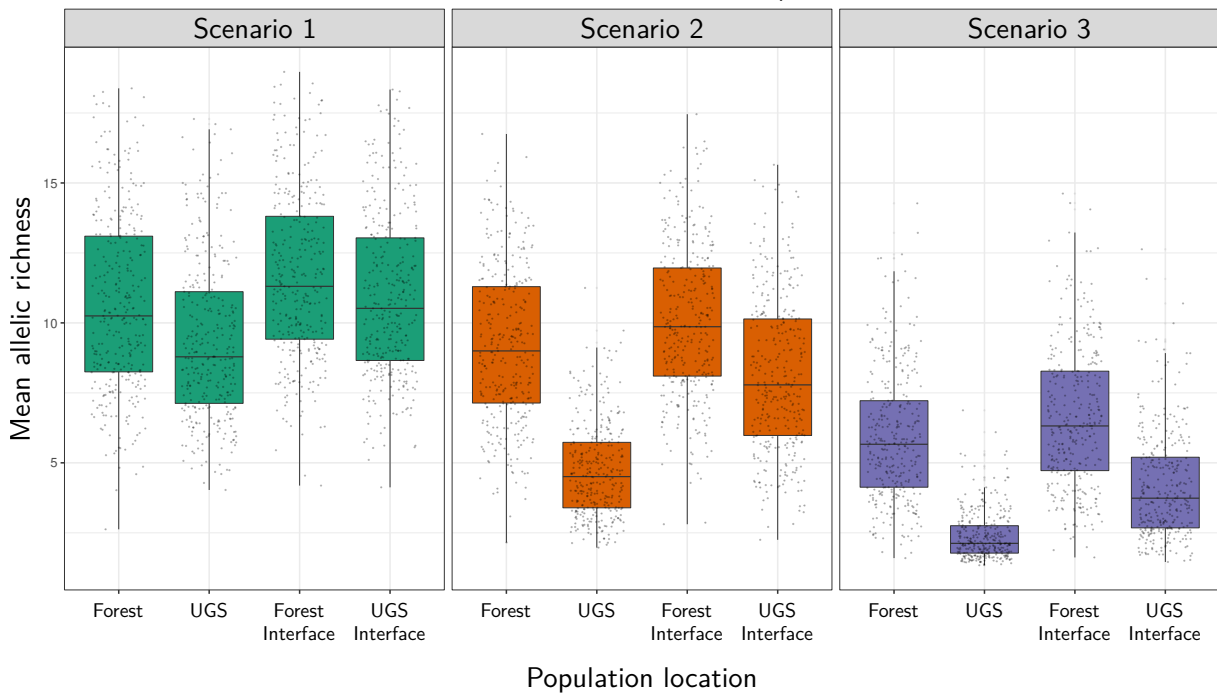
Figure 4. Distribution of the mean genetic differentiation (D_{PS}) computed between forest patches (Forest.Forest), forest and UGS patches (Forest.UGS) and UGS patches (UGS.UGS) in the 325 cities for the three dispersal cost scenarios. $n = 325$ values per box
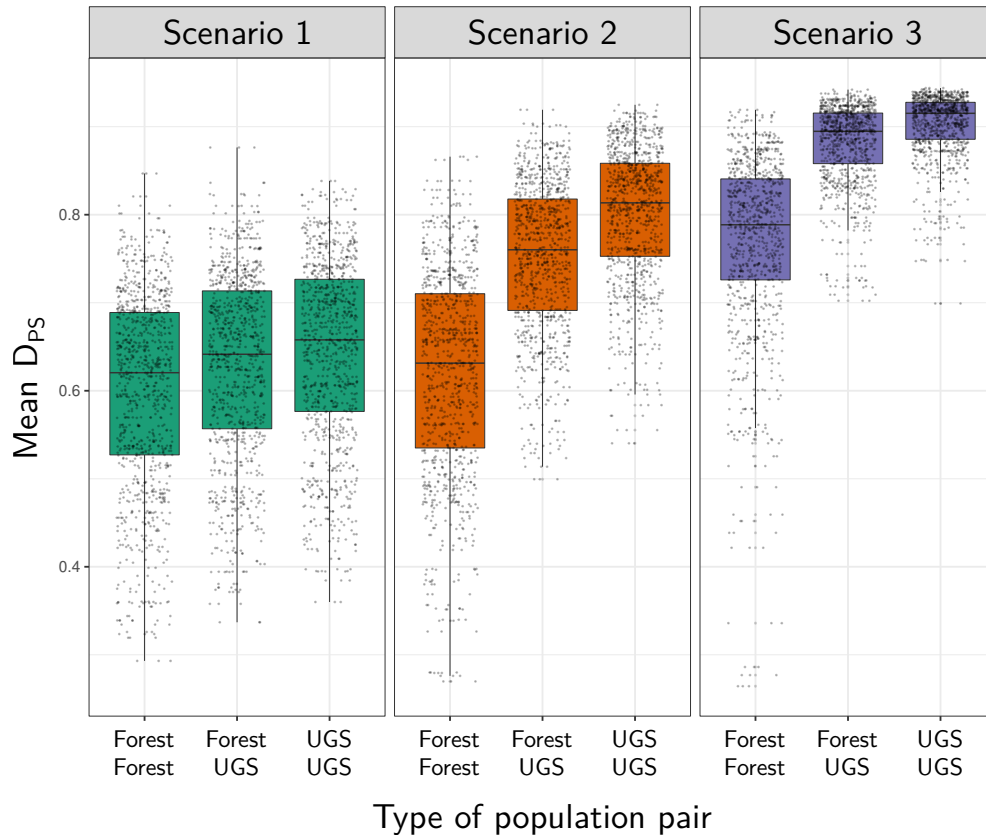
Figure 5. (A) Distance of Maximum Correlation (DMC), computed as the threshold distance used for selecting the subset of population pairs giving the maximum Mantel correlation coefficient between genetic distances (D_{PS}) and cost-distances, in the 325 cities and for the three dispersal cost scenarios. The DMC is divided by the maximum cost-distances between populations in the corresponding city and cost scenario and therefore ranges from 0 to 1. (B) Mantel correlation coefficients measured at the DMC. $n = 325$ values per box
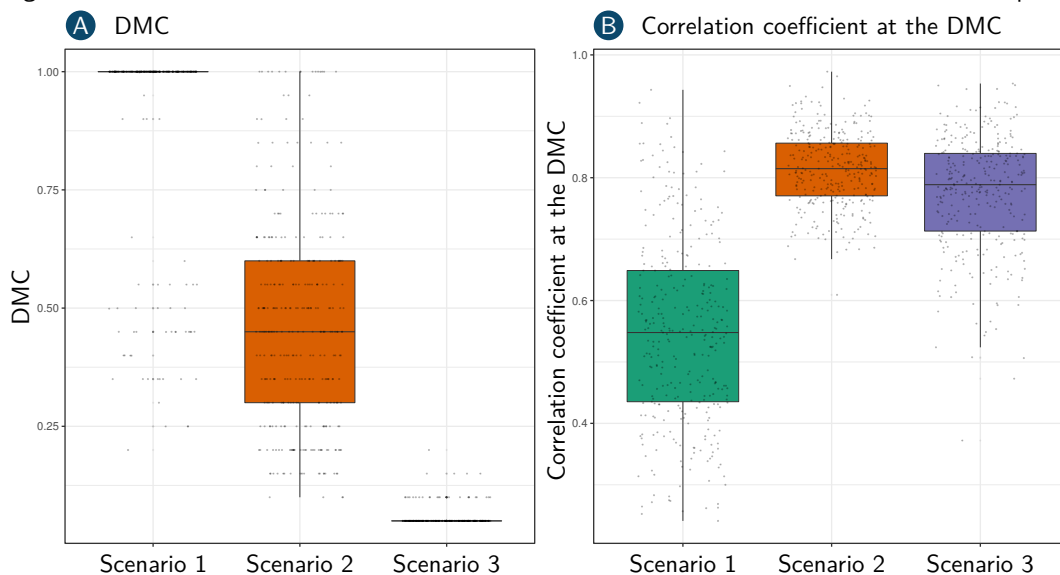
Figure 6. Adjusted Rand Index (ARI) comparing the partitions into modules of the genetic graph with links weighted by genetic distances (D_{PS}) with (A) module partitions obtained from similar graphs with links weighted by cost-distances or with (B) the classifications of populations into forest or UGS populations according to the type of patch they occupy. An ARI value is computed for each city and each cost scenario, such that $n = 325$ values per box