Contrasting genetic differentiation of urban and rural populations of two grassland lepidopterans across Europe

Sami M. Kivelä[1]*, Kyung Min Lee[2], George C. Adamidis[3], Simona Bonelli[4], Frederik Hendrickx[5], Peter Huemer[6], Tomáš Kadlec[7], Tuomas Kankaanpää[1], Guillaume Minard[8], Matthew E. Nielsen[1, 9], Elena Piano[4], Toomas Tammaru[10], Konstantina Zografou[11], Thomas Merckx[12]

[1] Ecology and Genetics Research Unit, University of Oulu, FI-90014 Oulu, Finland

[2] Zoology Unit, Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland

[3] Laboratory of Plant Physiology, Department of Biology, University of Patras, Patras, Greece

[4] Department of Life Sciences and Systems Biology, University of Turin, 10123, Turin, Italy

[5] Royal Belgian institute of natural Sciences, Brussels, Belgium

[6] Natural History Collection, Collections and Research Centre, Tyrolean Federal State Museums, Hall, Austria

[7] Department of Ecology, Czech University of Life Sciences Prague, 16500 Prague-Suchdol, Czech Republic

[8] Université Claude Bernard Lyon 1, Laboratoire d'Ecologie Microbienne, UMR CNRS 5557, UMR INRAE 1418, VetAgro Sup, 69622 Villeurbanne, France

[9] Current address: Faculty 2 Biology/Chemistry, University of Bremen, Bremen, Germany

[10] Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, EE-50409 Tartu, Estonia

[11]Department of Biological Applications and Technology, University of Ioannina, 45110 Ioannina, Greece

[12] WILD, Biology Department, Vrije Universiteit Brussel, 1050 Brussels, Belgium

* Corresponding author

Abstract

Urbanisation is transforming environments globally. The altered abiotic conditions and biotic interactions in urban habitats inflict divergent selection pressures on urban versus rural populations. Genetic drift may simultaneously be significant in typically small urban populations. A key question in urban evolution is whether urban genotypes have a single or multiple origins, but such assessments remain rare, precluding generalizations. We expect similar patterns in species that share ecological traits, but such replicates are missing. To address whether urban and rural populations are genetically differentiated and what is the origin of urban genotypes, we conducted an extensive population genomic ddRADseq analysis of two non-model grassland lepidopterans, *Coenonympha pamphilus* and *Chiasmia clathrata*, across Europe. Although the species share ecological traits, their genetic population structures were strikingly different: *C. pamphilus* had strong population differentiation, while this was almost absent in *C. clathrata*, which instead showed signs of recent strong population expansion. Although this signature of population expansion complicates inferences, it is nonetheless plausible that the urban populations of both species have multiple origins. This study highlights that urban evolution research needs to integrate phylogeography, because population history affects urban evolutionary dynamics.

Keywords: *Chiasmia clathrata*, *Coenonympha pamphilus*, ddRADseq, isolation-by-distance, isolation-by-environment, Lepidoptera, population admixture, urban evolution, urbanisation

Introduction

Urbanisation changes both the abiotic and biotic environment globally [1,2], and it is projected to extensively continue [2,3]. Consequently, wildlife is increasingly forced to face urban environments. Although urban settings filter out many species, other species do cope with urbanisation, and some even successfully exploit urban habitats [4-8]. Because both biotic interactions and abiotic conditions differ in urban environments from those of natural or rural habitats, it is likely that populations in urban habitats experience different selection pressures compared to those living in non-urban habitats. Accordingly, evidence of evolved adaptations to urban environments is accumulating (reviewed by [9-12]). Yet, urban populations are often small and isolated in green patches embedded within an urban matrix, where stochastic events such as habitat modification and destruction by human activities may be more frequent than in other contexts [13]. These conditions make genetic drift potentially an important evolutionary force too, resulting in neutral genetic differentiation between urban and rural populations [10,14]. Consequently, genetic differentiation of urban populations has been reported in some species [15-20] but not in all studied ones [19,21-24]. However, studies have typically focused on only a single or a few cities (but see [15,18,20,23]), which is why the large-scale dynamics of urban evolution remain insufficiently understood.

An important issue concerning evolution in urban environments is to identify whether urban genotypes have arisen only once and then spread across cities—which is a plausible mechanism in mobile species—or whether urban genotypes derive from independent

evolution in different cities [20,22,23,25]. An example of the former scenario is the recent urbanisation-associated range expansion of the butterfly *Pieris mannii* in Central Europe, driven by a single genetic lineage [26]. Industrial melanism in the moth *Biston betularia* in Britain has been traced to a single evolutionary origin too [27]. On the contrary, genetic evidence suggests multiple origins of urban Blackbird (*Turdus merula*) genotypes [15], although the observed timing and spatial spread of urban colonisations erroneously implied a scenario where urban genotypes have a single origin and have then spread from one city to another [28]. In Great tits (*Parus major*), haplotypes under selection in urban populations mainly differ across cities, suggesting that urban adaptations evolve independently and repeatedly in different cities, although selection consistently acts on the same behaviour-linked genes in different urban populations [20]. The owl *Athene cunicularia* shows a clear genetic signature of multiple urban colonisations [29] and repeated adaptation in neural function likely affecting cognition [25]. Finally, in the frog *Lithobates sylvaticus*, bumblebee *Bombus lapidarius*, and damselfly *Ischnura elegans*, urban populations are not genetically differentiated from nearby rural ones, although a few genetic markers associated with urbanisation were identified [21,22,24], rendering inferences on the dynamics of urban evolution impossible. Owing to this diversity of patterns, we cannot draw generalizations about the origin and spread of urban genotypes or the genetic differentiation between urban and rural populations. Although a diversity of patterns is expected across species [14], we would still expect ecologically similar species to show similar evolution in urban environments and, consequently, similar genetic population structures. Spatially extensive studies including several cities and replicated across species are necessary to address whether this prediction holds true.

Here, we conducted an extensive population genomic analysis across multiple European urban and rural populations of the butterfly *Coenonympha pamphilus* (L.) (Nymphalidae) and the moth *Chiasmia clathrata* (L.) (Geometridae) to find out the scale of genetic differentiation among urban and rural populations, and to address the origin of urban genotypes. Both species are small and sedentary, inhabit open meadow and grassland habitats in their (Western) Palearctic range and use very common and abundant larval host plants (grasses by *C. pamphilus*, leguminous plants by *C. clathrata*) [30-37]. The species have similar phenologies and produce one generation per year in the northern parts of their distribution and two or more at lower latitudes where the growing season is longer [32,35,36]. Because there is evidence of genetic population structure at a small spatial scale in *C. pamphilus* [38], we expected to find genetic differentiation also at a larger geographic scale and between urban and nearby rural populations. In *C. clathrata*, common garden experiments have demonstrated genetically-based urban adaptations in seasonal plasticity and heat tolerance [39-41] as well as local life-history adaptations along a subset of the climatic gradient included in the present study [42,43], suggesting genetic differences among populations. Besides testing for genetic population differentiation and the genetic signature of urban adaptation, we specifically aimed to reveal whether urban genotypes have a single or multiple evolutionary origin. A single origin of urban genotypes would manifest itself as shared genetic lineages across urban populations, with individuals from all urban populations belonging to the same genetic cluster, resulting in weaker isolation-by-distance among urban than among rural populations [15]. Alternatively, urban individuals clustering with geographically nearby rural ones would indicate multiple origins of urban genotypes. Moreover, if each urban population derives from a nearby rural

population, we expect similar isolation-by-distance across both urban and rural populations, and a threshold-type increase in isolation by the degree of urbanisation of the environment.

Results

We sampled *C. pamphilus* and *C. clathrata* from 22 and 28 sites across Europe, respectively. These sites consisted almost exclusively of pairwise urban and rural locations associated with the same city (Fig. 1; Table S1). We derived random single nucleotide polymorphisms (SNPs) from the genomes of the sampled individuals by using ddRAD [44]. This yielded 48459 polymorphic SNPs for the 326 *C. pamphilus* individuals that met our quality criteria for inclusion (see Methods), with 41.1% missing data. The corresponding numbers for *C. clathrata* were 17896 polymorphic SNPs for 325 individuals, and 37.8% missing data.

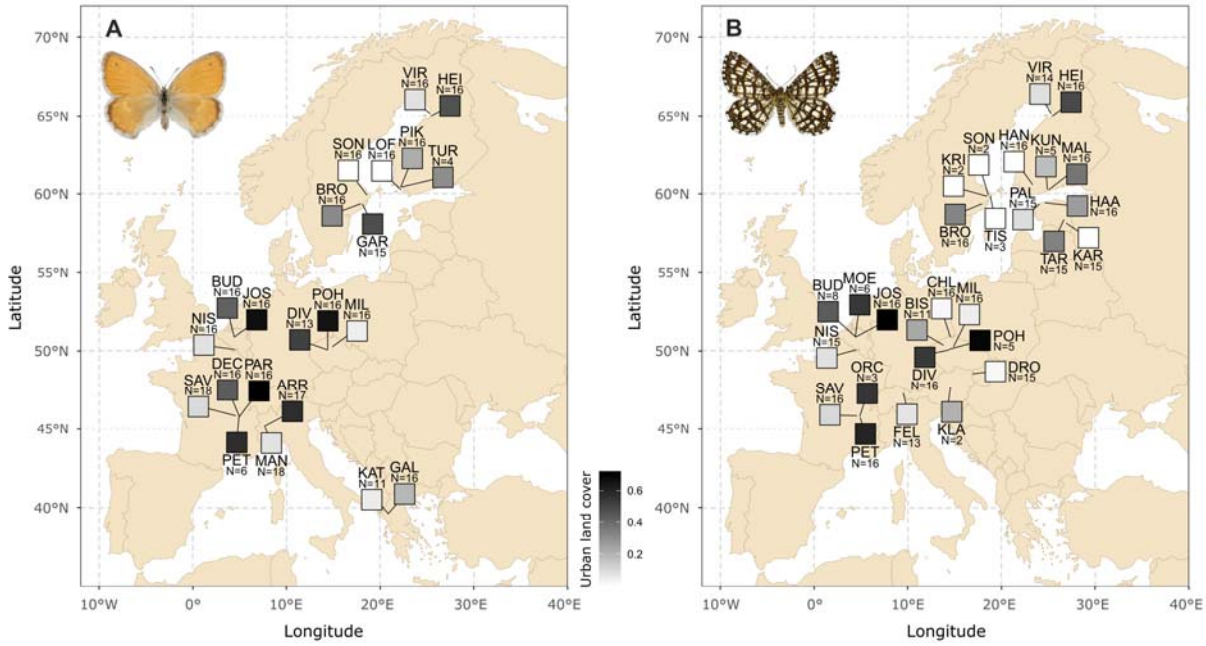

Figure 1. Maps showing the sampling sites for *Coenonympha pamphilus* (A) and *Chiasmia clathrata* (B). Squares indicate sampling sites, with lines connecting them to the exact sampling locations (the squares are spread out to improve readability). The colour of each square indicates the degree of urbanisation at the sampling sites (defined as the proportion of land covered by impervious structures within a 2500-m radius; darker colour means higher urbanisation). Population abbreviations and sample sizes (N) are given for each site. Pictures of the study species are shown at the top left corners of the maps (pictures sourced from the Finnish Biodiversity Information Facility [http://laji.fi] under the CC BY-SA 4.0 license, taken by Pekka Malinen, and accessed on 28 March 2024).

We investigated the genetic lineages present in the studied populations with a population admixture analysis using sparse non-negative matrix factorization algorithms as implemented in the R function ‘snmf’ [45]. This analysis identified nine genetic clusters as the best-supported number of clusters (*K*) in *C. pamphilus*, yet numbers from seven to 14 were almost equally well-supported (Fig. 2A). In this species, country-level genetic differentiation was evident, except that the Czech and Greek populations were assigned to the same cluster when $K \leq 10$ (Fig. 2A). This country-level genetic clustering was also recovered when analysing relationships between the individuals using a network analysis by SplitsTree [46](Fig. S1A). TreeMix [47], a tool designed for detecting historical population splits and gene flow among populations, yielded similar results, with the only inferred migration events having taken place from the French to the Czech lineages, and from the Italian to the French lineages (Fig. S2A). Within-country genetic clustering was found in Finland where the northern and southern populations formed separate clusters (Fig. 2A; Fig. S1A). Urban populations generally clustered with nearby rural populations, as expected under a scenario of multiple origin of urban genotypes. However, in Belgium, and to a lesser degree in Czech Republic, some urban populations showed an ancestry different from the associated rural populations.

In *C. clathrata*, the genetic clustering analysis best supported the presence of three genetic clusters (Fig. 2B). One of the clusters consisted of individuals of the urban Malmi population from Southern Finland, another one of all the other Finnish and Estonian populations, and the third of all the remaining populations (Fig. 2B). The assignment of individuals to these clusters was not as clear as in *C. pamphilus*, and the ancestry of many individuals was unclear. Lack of genetic population structure was also highlighted in a SplitsTree analysis (Fig. S1B). Consequently, drawing inferences concerning the origin of urban genotypes is difficult in this species. Nonetheless, a TreeMix analysis recovered a geographically meaningful clustering, although the degree of genetic differentiation was minimal in comparison to *C. pamphilus*, and several migration events between lineages were inferred (Fig. S2B).

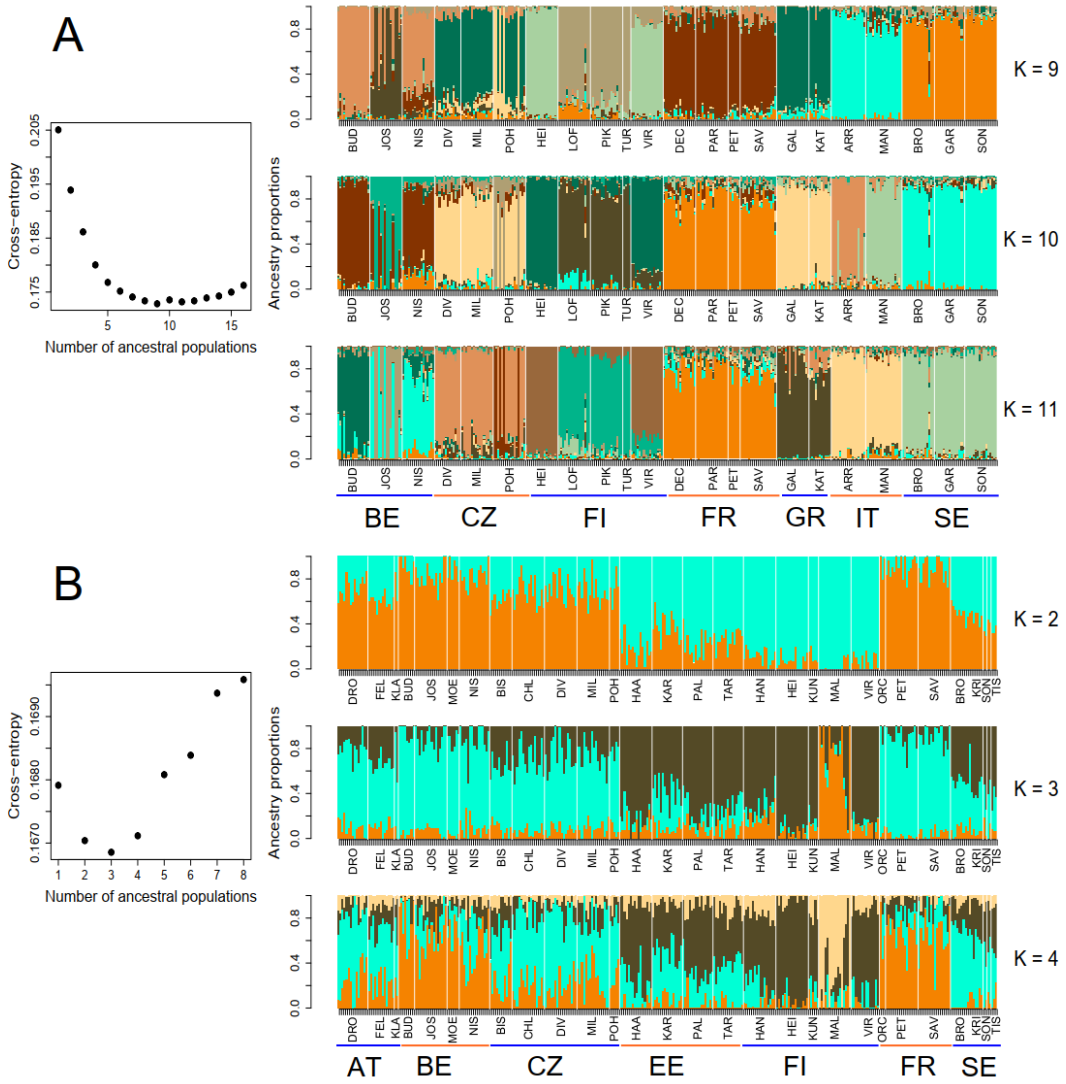

Figure 2. Results of population admixture analysis. Cross-entropy values (left column) and individual-specific ancestry coefficients (coloured vertical bars in the right column) derived from an analysis using sparse non-negative matrix factorization algorithms for *Coenonympha pamphilus* (A) and *Chiasmia clathrata* (B). The number of ancestral populations (*K*) that minimizes cross-entropy is the best-supported number of genetic clusters (nine in *C. pamphilus*; three in *C. clathrata*). The right column shows the ancestry coefficients for three well-supported numbers of ancestral populations (*K* = {9, 10, 11} for *C. pamphilus*, *K* = {2, 3, 4} for *C. clathrata*). The height of each colour within a vertical bar indicates the probability that the individual in question belongs to a particular genetic cluster. The order of individuals is kept the same across panels. White vertical lines separate populations, and population labels (see Fig. 1; Table S1) are given below the horizontal axis. Country abbreviations (see Table S1) are also given below the population labels in the bottom panels of both species, and the coloured horizontal lines above the country labels tell apart the populations from different countries.

To get more insight into population differentiation and the associated role of urbanisation and background climate, we ran a redundancy analysis (RDA) by using the R function 'rda' [48]. We did this for both species by setting urbanisation and growing season length as environmental covariates. In *C. pamphilus*, urbanisation and growing season length explained 8.08% of variation in the genotypic matrix (adjusted R^2), with both urbanisation ($F_{1,323}$ = 4.98, p = 0.001) and growing season length ($F_{1,323}$ = 25.6, p = 0.001) being significant predictors of genotypic variation. Populations were clearly separated in the space defined by the first two RDA axes, with urban populations situated close to nearby rural ones (Fig. 3A). This indicates a strong population structure, aligning with the results of the population admixture analysis.

In *C. clathrata*, urbanisation and growing season length explained only 3.90% of the genotypic variation in the RDA analysis (adjusted R^2). Nonetheless, both urbanisation ($F_{1,322}$ = 2.36, p = 0.001) and growing season length ($F_{1,322}$ = 12.8, p = 0.001) were again significant predictors of this genotypic variation. Populations were not separated in a space defined by the first two RDA axes but formed a continuum along the urbanisation and climatic gradients where geographically close populations were near each other (Fig. 3B). This aligns well with the results of the population admixture analysis, where clear genetic clusters could not be identified.

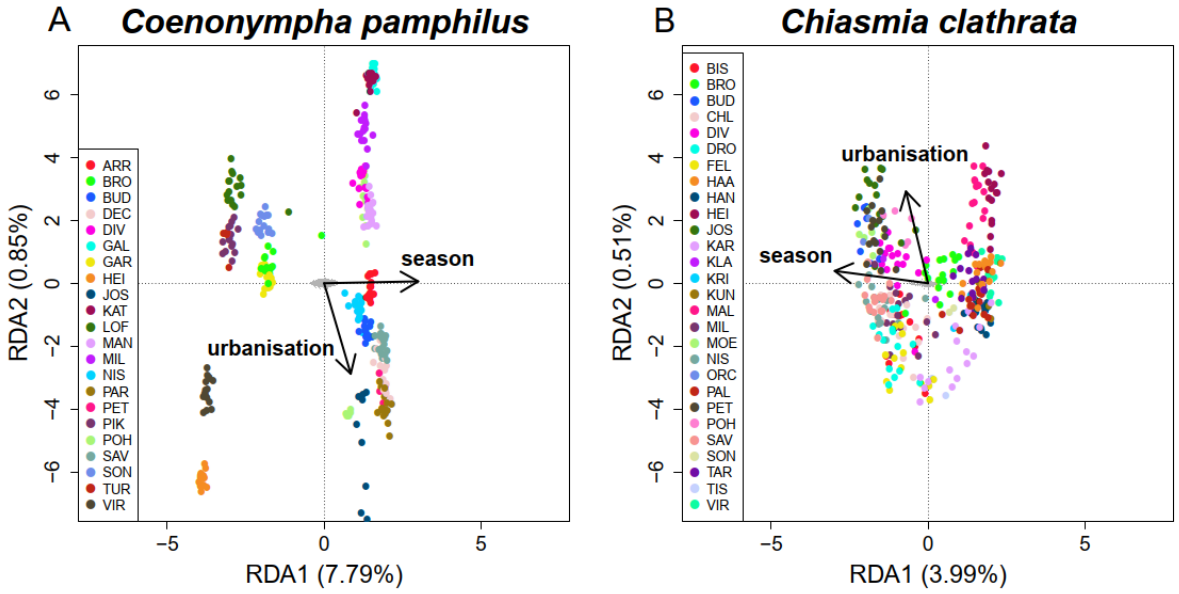

Figure 3. Visualization of results of the redundancy analysis (RDA) exploring the effects of urbanisation and climate (growing season length) on genotypic variation in *Coenonympha pamphilus* (A) and *Chiasmia clathrata* (B). Scaling is such that the lengths of the arrows—illustrating directions of increasing urbanisation and growing season length—show the effect sizes (arrow lengths are multiplied by three to improve readability), while the angles between the arrows indicate the correlation between the environmental predictors (almost no correlation here, because the angle is close to 90° in both species). The small grey points, densely concentrated in the middle, depict the locus scores, while the coloured large points indicate the individual scores, with colours corresponding to the population of origin.

Proportions of variance explained by the RDA axes are given in the axis labels. See Figure 1 and Table S1 for population abbreviations. Corresponding visualizations, scaled so that the distances among points depend on their similarity, are presented in Figure S3.

Because a part of the genetic variation was associated with variation in urbanisation and climate in the RDA analysis, we ran a more detailed investigation of associations between SNPs and either urbanisation or growing season length; genotype-environment association analysis with the latent factor mixed model approach by using the R function 'lfmm2' [45]. In *C. pamphilus*, the number of latent factors did not affect model goodness-of-fit much, but the model including 12 latent factors appeared to be best (see Fig. S4A, B) and we used it for drawing inferences concerning this species. That model found three candidate SNPs associated with variation in urbanisation with a 5% false discovery rate (FDR), none of which were significant at 1% FDR (Fig. 4A). In contrast, 4136 SNPs were associated with variation in growing season length at 5% FDR, 2547 of which were significant at 1% FDR (Fig. 4B). Hence, there is only a weak signature that few of the investigated SNPs are involved in urban adaptation, but there is a clear signal of climate adaptation across numerous loci.

In *C. clathrata*, the number of latent factors had a minimal effect on model goodness-of-fit too, yet the model including four latent factors seemed best (see Fig. S4C, D), so we used that model for drawing inferences. While there were no SNPs associated with variation in urbanisation, neither at 1% nor at 5% FDR (Fig. 4C), 73 SNPs were associated with variation in growing season length at 5% FDR, 53 of which remained significant at 1% FDR (Fig. 4D). *Chiasmia clathrata*, therefore, shows some genetic signature of climate adaptation, but no indication of urban adaptation. This is consistent with the known existence of local climate adaptations in this species but contradicts with the known presence of urban adaptations.

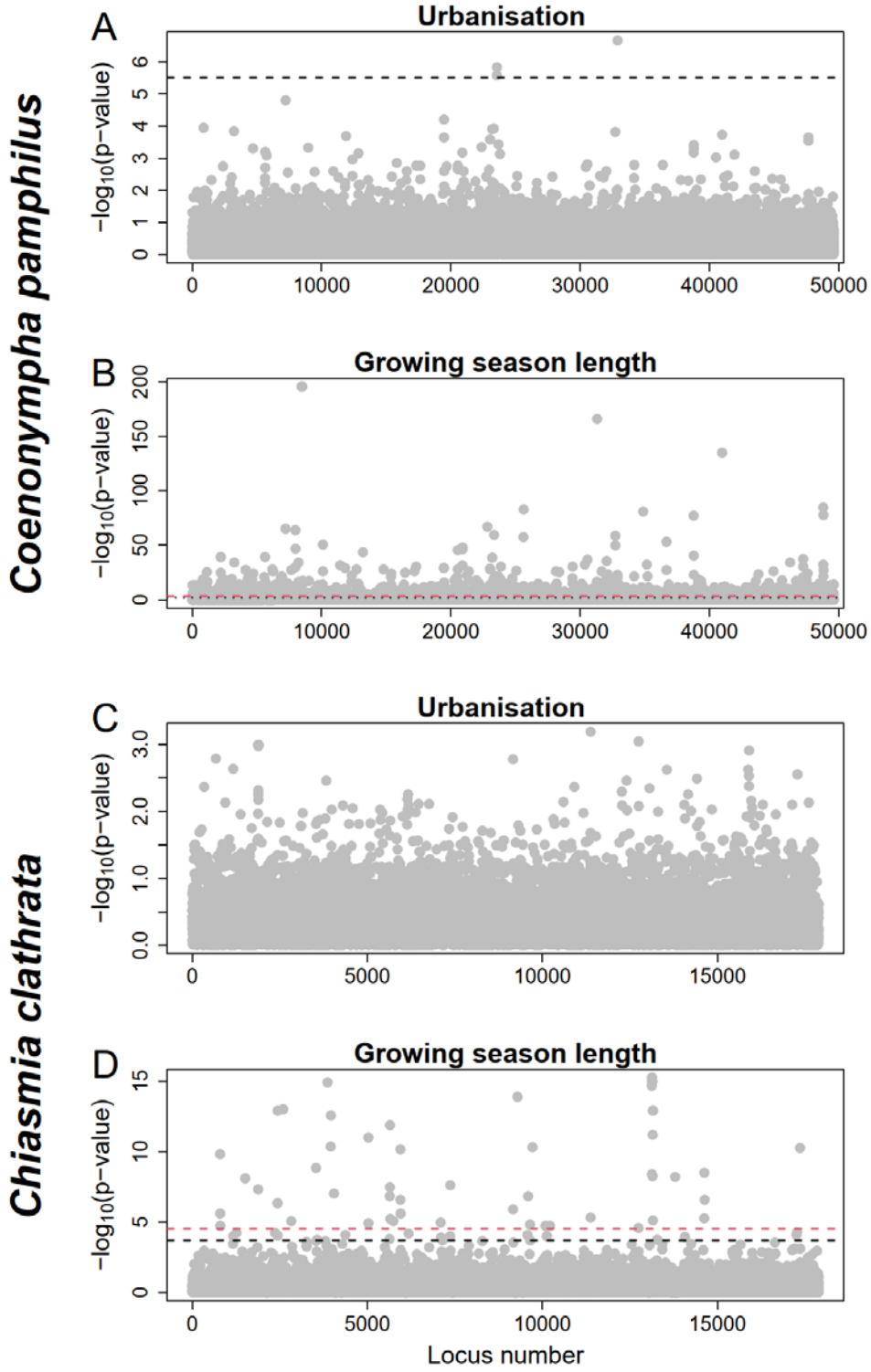

Figure 4. Manhattan plots showing the $\log_{10}$-transformed *p*-values for SNP-specific associations with urbanisation (A, C) and growing season length (B, D) in *Coenonympha pamphilus* (A, B) and *Chiasmia clathrata* (C, D). The red and black dashed lines indicate statistical significance thresholds adjusted for 1% and 5% false discovery rate, respectively (each threshold shown only when at least one SNP

exceeds the given threshold). Note that the scale of the y-axis varies greatly among panels A–D.

Finally, to test the predictions concerning single versus multiple origins of urban populations, we conducted isolation-by-distance (IBD) and isolation-by-(the degree of urbanisation of the) environment (IBE) analyses. We performed these analyses separately for rural and urban populations. We derived matrices of pairwise F_{ST} values (Tables S2 and S3), pairwise geographic distances, and pairwise differences in urbanisation between all populations for both species. In *C. pamphilus*, there was clear IBD both among rural and urban populations (Fig. 5A) as well as across all possible population comparisons (Mantel test [R function 'mantel'; [48]]: r=0.667, p=0.0002), but neither IBE in these population comparison categories (Fig. 5C) nor overall (Mantel test: r=-0.0313, p=0.62). There was also no threshold-type increase in genetic distance at any urbanisation difference (Fig. 5C). Hence, the patterns in *C. pamphilus* match most of the predictions for multiple origins of urban populations.

In *C. clathrata*, there was slight IBD among urban populations (Fig. 5B) and across all population comparisons (Mantel test: r=0.190, p=0.0036) but not when comparing only rural populations (Fig. 5B). In this species, there was also some IBE when comparing rural populations but neither among urban populations (Fig, 5D), nor across all population comparisons (Mantel test: r=-0.0320, p=0.59). These patterns contradict the predictions of both single and multiple origins of urban populations.

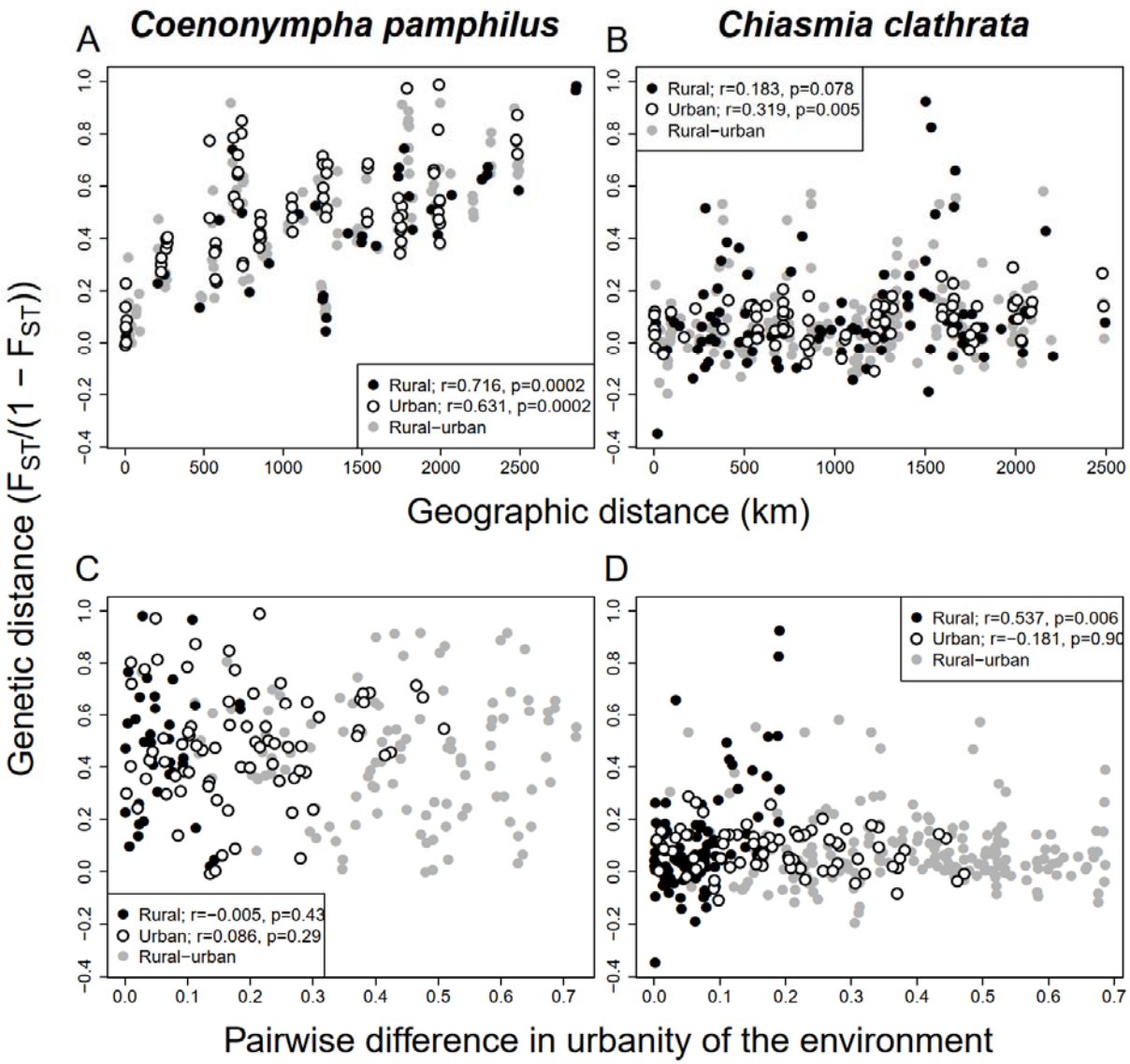

Figure 5. Isolation-by-distance (A, B) and isolation-by-environment (C, D) analyses for *Coenonympha pamphilus* (A, C) and *Chiasmia clathrata* (B, D). Genetic distances were based on transformation by Rousset [49], whereas difference in the urbanisation of the environment is the inter-population difference in proportion of land covered by impervious surfaces within 2500-m radii from the sampling locations. Black dots indicate rural-rural population comparisons, open circles urban-urban comparisons, and grey dots rural-urban comparisons. Pearson product-moment correlation coefficients, estimated based on Mantel tests with 5000 permutations, and associated *P*-values, are indicated in the legends for rural-rural and urban-urban comparisons. Note that the scale of the y-axis is kept constant across all panels to facilitate comparison between the species.

Discussion

We found contrasting genetic population structures in two grassland lepidopterans, although we expected similar patterns of population differentiation because of similarities in their body size, mobility, habitats, phenology, and host plant growth forms. Strong genetic population structure in *C. pamphilus* and almost no structure in *C. clathrata* is striking, especially when considering that common-garden experiments on *C. clathrata* have demonstrated heritable phenotypic differences between geographically close rural and urban populations [39-41] and geographically distant populations along a climatic gradient [42,43]. Genetic population structure in *C. pamphilus* is consistent with multiple origins of urban genotypes, whereas the lack of population structure in *C. clathrata* precludes inferences on the origins of urban genotypes. However, the results of the IBD and IBE analyses in *C. clathrata* are opposite to those predicted by a single origin of urban populations. Instead, the results imply a single origin of rural populations in *C. clathrata*, deriving from the strong and relatively recent population increase suggested by the star-like topology of the phylogeny of studied individuals [50] (Figs S1B and S5).

Our results concerning *C. pamphilus* generalize the small-scale population structure found by Wood and Pullin [38] to a continental scale. As the dispersal capacity of *C. pamphilus* appears limited [30], population differentiation may eventually evolve even between geographically nearby populations due to genetic drift. We found little evidence of selection in urban populations, as only three SNPs were associated with urbanisation, and only at a more tolerant 5% FDR threshold. This aligns with a ddRAD study on a damselfly that did not find strong evidence of specific SNPs being under selection in urban environments [24], yet a study on a bumblebee identified urban adaptation in some loci despite weak genetic population structure [21]. Whole-genome analyses revealed clear evidence of loci under selection in urban environments in two bird species, although genetic population structure was weak in one of them [20,25]. These examples suggest that methodological limitations of the ddRAD approach may partially explain why we did not find convincing evidence of urban adaptation in *C. pamphilus* despite clear evidence of local climatic adaptation.

Despite lacking identifiable SNPs associated with urbanisation, we know from common-garden experiments that adaptations to the urban environment have evolved in *C. clathrata* populations. There are five mutually non-exclusive explanations for this conundrum, and these explanations are also equally likely to explain the weak genetic signature of urban adaptation in *C. pamphilus*. (1) Urban adaptation may be polygenic, with different loci and alleles underlying the observed phenotypic evolution in different urban populations (see e.g., [20,51]), making the association analysis conservative. This is plausible because the known urban adaptations concern polygenic (quantitative) traits in *C. clathrata* and many other species (e.g., [20,51-54]). Consistent with this hypothesis, pairwise F_{ST} values between nearby urban-rural population pairs suggest slight differentiation associated with the urban environment in *C. pamphilus* (mean scaled F_{ST}=0.109 [95% confidence interval: 0.0226, 0.195], df=7, t=2.99, p=0.02), yet this is not the case in *C. clathrata* (mean scaled F_{ST}=-0.0100 [95% confidence interval: -0.0617, 0.0417], df=8, t=-0.448, p=0.67), which has very low population differentiation overall. (2) Low genetic variation between populations may have hindered the identification of adaptive markers. Accordingly, the number of polymorphic SNPs recovered in *C. clathrata* was only 37% of those recovered in *C. pamphilus*. (3) Key SNPs linked to urban adaptation may not have been captured in the sampled data due to the random genome

sampling inherent to the ddRADseq approach (see [44]). (4) Stringent bioinformatics quality control may have filtered out important low-quality SNPs, such as those with low read depth or high error rates. Finally, (5) small sample sizes of a few *C. clathrata* populations (e.g., KLA, ORC, KRI, SON, and TIS) and sampling imbalances between these and other populations may have constrained the detection of strong associations. Nevertheless, there were SNPs clearly associated with the known adaptations to local climate in *C. clathrata*, indicating that our approach at least had power to detect genotype-environment associations originating from selection that has continued longer and over a larger spatial scale than the selection imposed by the urban environment.

A plausible explanation for the lack of population structure in *C. clathrata* is a recent and strong population expansion that has left a strong signature across the whole genome. It remains uncertain whether such population expansion is linked to postglacial expansion from a glacial refugium, or with a more recent expansion associated with the spread of agriculture or extension of (agri)cultural landscapes. Alternatively, *C. clathrata* frequently disperses over long distances, even though it is considered as a sedentary species [31]. Long-distance dispersal could occur if adult moths are carried by strong winds, or if eggs or larvae are anthropogenically transported with leguminous fodder or mulching. We expect *C. pamphilus* to have expanded from a glacial refugium too, yet it remains also possible that it has been widespread in the grasslands during the ice age, similar to the butterfly *Polyommatus icarus* [55], and the current population structure reflects fragmentation of a once extensive and uniform population.

Both study species have been assumed poor dispersers [30,31]. Our findings not only make this assumption improbable for *C. clathrata,* but also suggest some long-distance gene flow in *C. pamphilus*, as individuals from the Lofsdal (LOF) population in southern Finland show some ancestry from the Swedish populations, which cannot be explained by gene flow through the continuum of populations around the Bothnian Bay because the northern (HEI, VIR) populations belonging to that continuum formed a separate genetic cluster (Fig. 2A). The TreeMix analysis also suggested migration from western to central Europe and over the Alps (Fig. S2A). Nevertheless, if long-distance dispersal is rare and, consequently, city-to-city dispersal very unlikely, urban genotypes likely have multiple origins in both species. The slight isolation by distance across urban *C. clathrata* populations (Fig. 5B) supports this reasoning; even though the rural *C. clathrata* populations appear to have originated from a single and strong population expansion, resulting in a very weak geographic population structure, this weak signature may be enhanced in urban populations if they have originated from nearby rural populations. We think that urban populations of both species have originated when parts of a formerly uniform population have been trapped in green spaces in the urban matrix once cities have grown, which is a similar process with similar genetic consequences as population subdivision due to habitat fragmentation in the butterfly *Zerynthia polyxena* [56]. Colonization of isolated habitat patches within urban environments would require crossing the urban matrix, which may be unlikely, or at least infrequent for the study species (cf. [30,31]). However, dispersal by human transport occurs in the spider *Latrodectus hesperus* [57] and some lepidopteran pests [58,59]. Anthropogenic dispersal along with fodder transportation remains a possibility for *C. clathrata* too, but we consider it unlikely for *C. pamphilus*. We also consider wind-assisted long-distance dispersal unlikely for *C. pamphilus*. Anyway, unlikely or unexpected possibilities need to be considered, as

exemplified by the plant *Lepidium virginicum* where genetic evidence suggests city-to-city colonization over long distances, contrary to *a priori* expectations [60].

Our study adds diversity to the reported patterns of genetic population structure in the urban evolutionary ecology context. The species studied here share several ecological traits, which would predict similar responses to urbanisation, and similar genetic population structures. However, *C. pamphilus* has a strong population structure, whereas *C. clathrata* shows almost none. These contrasting patterns seem to derive from different phylogeographies, with the genetic signature of population expansion being dominant in *C. clathrata*. It is not uncommon that the phylogeography differs between species sharing key functional traits [61,62]. Thus, caution is needed when generalizing results based on a few species. Because both the origin of urban populations as well as the evolution of genetic differentiation between urban and rural populations depend on phylogeography, the biogeographical and historical contexts need to be considered when studying urban evolution, independently of whether the focus is on a single or multiple urban populations. Therefore, continent-scale population genetic analyses should be integrated as a part of the urban evolutionary ecology research program. Currently, there is a paucity of such studies, but the existing ones suggest that urban adaptation occurs and is largely repeated [20,23].

Methods

Sampling

We sampled *C. pamphilus* with a hand net from 22 sites across Europe, the sites including both urban and rural environments (Fig. 1A; Table S1). *Chiasmia clathrata* was sampled from 28 sites across Europe, mainly with a hand net but a light trap was used in all Austrian sites, in two Czech sites (BIS, CHL; see Table S1 for an abbreviation key), and in two Estonian sites (KAR, TAR). Sampling sites included both urban and rural environments (Fig. 1B; Table S1).

DNA extraction and ddRADseq library preparation

Genomic DNA (gDNA) was extracted from leg and thorax tissues using the E.Z.N.A. Tissue DNA Kit (Omega Bio-tek), following the manufacturer's protocol. The gDNA concentration was measured with a Quant-iT PicoGreen dsDNA assay kit (Molecular Probes). The ddRADseq library preparation followed the protocol described by Lee et al. [63] with minor adjustments: PstI and MseI restriction enzymes (NEB) were used for digestion, and the size distribution was assessed via Bioanalyzer (Agilent Technologies). The final libraries were sequenced on the Illumina HiSeq X PE150 platform at Macrogen (Seoul, South Korea) and NovaSeq X Plus PE150 at Novogene (Cambridge, UK).

Bioinformatics of ddRADseq reads

Raw paired-end reads were demultiplexed using their unique index and adapter sequences, with no mismatches tolerated, with *ipyrad* v.0.9.92 [64]. The quality of the raw

demultiplexed reads was checked with FastQC (available at http://www.bionformatics.babraham.ac.uk/projects/fastq/). Poor-quality samples were removed after quality checks, with one sample excluded from the *C. pamphilus* dataset, leaving 326 samples, and two samples excluded from the *C. clathrata* dataset, leaving 325 samples. The demultiplexed paired reads were merged into single-end reads using PEAR v.0.9.8 [65] with default settings and then processed through the *ipyrad* pipeline for further analysis.

All *ipyrad* analyses were performed on the Puhti supercomputer at the Centre for Scientific computing Finland (CSC), using default settings with the following modifications: datatype was set to “ddrad”, assembly method to “denovo”, restriction overhang to “TGCAG,TAA”, maximum low-quality bases to 6, minimum statistical depth to 8, clustering threshold to 0.9, and the minimum number of samples required for a locus set to 120. Loci with excessive shared polymorphic sites (>50% of samples) were excluded to avoid clustering of paralogs. We generated two types of final matrices: one including all variable sites (SNP) and another containing one random SNP from each putatively unlinked locus (uSNP). These matrices were utilised for different analyses, depending on data volume and missing data, which could impact analysis speed or be constrained by dataset size.

Data analyses

First, we ran a frequentist population admixture analysis by using sparse non-negative matrix factorization algorithms as implemented in the function ‘snmf’ from the package ‘LEA’ [45] in R version 4.3.2 [66]. The genotypic matrix was used as an input for the derivation of individual-specific ancestry coefficients, and we repeated the analysis with the number of genetic clusters (*K*) spanning from one to 16 in *C. pamphilus* and from one to eight in *C. clathrata*. Higher values of *K* were included in *C. pamphilus* because initial investigations indicated strong population structure in this species.

We further investigated population structure by employing SplitsTree v.4.19.2 to generate an unrooted genetic network using the Neighbor-Net algorithm with uncorrected p-distances [46]. For admixture histories, we applied population tree inference model implemented in TreeMix v.1.13 [47]. The analysis, performed on the uSNP dataset, progressively added up to 8 migration events. Asymmetries in the covariance matrix of allele frequencies, relative to the ancestral population inferred from the ML trees, were used to determine the directionality of gene flow.

For a further investigation of the weak population structure in *C. clathrata*, we derived a phylogenomic maximum likelihood (ML) tree for this species. The tree was constructed by using the SNP dataset in IQ-TREE v.2.1.3 [67]. The integrated ModelFinder [68] within IQ-TREE tested 286 DNA substitution models, selecting ‘SYM+I+G4’ for the *C. clathrata* dataset. To reconstruct the phylogeny, ML analysis with ultrafast bootstrap approximation model (1,000 replicates) was applied [69]. The resulting tree was generated using FigTree v.1.4.2 [70] and further modified using CorelDRAW v24.

To get more insight into population differentiation and the role of urbanisation and climate in it, we ran a redundancy analysis (RDA) by using the R function ‘rda’ from the package ‘vegan’ [48] for both species and setting urbanisation and growing season length as

environmental covariates. We measured urbanisation as the proportion of land covered by artificial surfaces and constructions (water bodies excluded) within a 2500 m radius based on 10 m resolution land cover data from S2GLC Europe from 2017 (version v1.2; [71]). Growing season length was measured as the number of days during which mean daily temperature was above 5 °C, with the conditions that mean daily temperature was above the threshold for seven consecutive days since the beginning date of the season and below the threshold for seven consecutive days after the end date of the season. We used interpolated (30 arc second resolution) daily temperatures for 1990–2005 from CHELSA Europe (version v1.1; [72,73]) for the estimation of site-specific growing season lengths. Both urbanisation and growing season length were standardized to mean of zero and standard deviation of one. Before the RDA analysis, we imputed missing values in the genotypic matrices by using the function 'impute' from the R package 'LEA' [45] and setting the number of genetic clusters to the value that was best supported in the snmf-based admixture analysis (nine clusters in *C. pamphilus*, three in *C. clathrata*; see Results). We ran permutation tests for the effects of urbanisation and growing season length by using the 'anova.cca' function [48] with 1000 permutations.

Furthermore, we ran a genotype-environment association analyses with the latent factor mixed model approach to investigate whether any SNPs show an association with either urbanisation or background climate, measured as growing season length. We used the R function 'lfmm2' from the package 'LEA' [45] for fitting the latent factor mixed models. For both species, the response variables were the columns of the imputed genotypic matrix, and the environmental covariates included the above-mentioned urbanisation and growing season length metrics standardized to mean of zero and standard deviation of one. The number of latent factors was initially set to be equal to the number of genetic clusters used in missing value imputation, but we also fitted models with number of latent factors spanning from 7 to 15 in *C. pamphilus* and from 2 to 5 in *C. clathrata* to assess if genomic inflation factors measuring inflation of test statistics due to population structure, or the distribution of *p*-values (a model goodness-of-fit indicator), depended on the number of latent factors. We derived *p*-values separately for urbanisation and growing season length by using the 'lfmm2.test' function [45]. First, we adjusted the SNP-specific *p*-values of both covariates to population structure. We did this by calculating the genomic inflation factors (λ) as λ_i = median($\boldsymbol{z}_i^2$)/0.456, where $\boldsymbol{z}_i$ (i = {urbanisation, growing season length}) refers to a vector of *z*-scores of covariate *i*, and then deriving *p*-values for these approximately X_1^2-distributed scores (see [74]) from the X_1^2 distribution. Second, we made an adjustment for multiple testing by using the Benjamini-Hochberg procedure [75] so that we derived statistical significance thresholds (adjusted *p*-values) corresponding to 1% and 5% false discovery rates. For inferences, we chose the model where the number of latent factors resulted in λ_i values closest to unity and appropriate distributions of *p*-values in both species.

Finally, we conducted isolation by distance (IBD) and isolation by (urbanisation of the) environment (IBE) analyses. For this purpose, we derived matrices of pairwise F_{ST} values, pairwise geographic distances (in kilometres), and pairwise differences in urbanisation between all populations for both species. Pairwise F_{ST} values were calculated with Arlequin v.3.5.1.3 [76], employing 1,000 permutations. In the IBD analysis, we tested if the pairwise F_{ST} values between populations are positively correlated with pairwise geographic

distances between the populations. In the IBE analysis, we tested if the pairwise F_{ST} values between populations are positively correlated with pairwise differences in urbanisation or growing season length between the populations. By following Rousset [49], we scaled the proxies of genetic distance (F_{ST} values) as $F_{ST}/(1 - F_{ST})$, but we did not use a logarithmic transformation for geographic distances, although the geographic set up of the study is two-dimensional, because the relationship between genetic and geographic distances was best linearized by using untransformed distances. Isolation along the urbanisation gradient is one-dimensional, so we used the raw differences in urbanisation as the distance measures in the IBE analysis [49]. We tested whether genetic distances were correlated with geographic (IBD) or urbanisation (IBE) distances separately for rural-rural and urban-urban population pairs with a Mantel test (based on 5000 permutations) to test the predictions concerning origin of potential urban-specific genotypes by using the R function 'mantel' from the package 'vegan' [48]. We used an urbanisation value of 0.2 (proportion of land covered by impervious surfaces within a 2500-m radius) as a cutoff when classifying populations as rural (urbanisation <0.2) or urban (urbanisation ≥0.2) in this analysis, which correctly categorized the locations as rural or urban along with the subjective categorization done during fieldwork.

Data availability

The ddRAD raw demultiplexed reads are deposited at NCBI SRA under BioProject no. PRJNA1177370 and BioSample accession numbers SAMN44419406-SAMN44419730 for *C. clathrata* and SAMN44419731-SAMN44420056 for *C. pamphilus*. The bioinformatics output files and environmental data are available at Zenodo (DOI: 10.5281/zenodo.14220594).

Code availability

The R-scripts of the analyses are available at Zenodo (DOI: 10.5281/zenodo.14220594).

References

1. Grimm NB, Faeth SH, Golubiewski NE, Redman CL, Wu J, Bai X, Briggs JM (2008) Global change and the ecology of cities. Science 319: 756–760
2. Seto KC, Güneralp B, Hutyra LR (2012) Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. Proceedings of the National Academy of Sciences of the USA 109: 16083–16088
3. United Nations, Department of Economic and Social Affairs, Population Division (2019) World Urbanization Prospects: the 2018 Revision. United Nations, New York.
4. McKinney ML (2006) Urbanization as a major cause of biotic homogenization. Biological Conservation 127: 247–260
5. Merckx T, Souffreau C, Kaiser A, Baardsen LF, Backeljau T, Bonte D, Brans KI, Cours M, Dahirel M, Debortoli N, De Wolf K, Engelen JMT, Fontaneto D, Gianuca AT, Govaert L, Hendrickx F, Higuti J, Lens L, Martens K, Matheve H, Matthysen E,

Piano E, Sablon R, Schön I, Van Doninck K, De Meester L, Van Dyck H (2018) Body-size shifts in aquatic and terrestrial urban communities. Nature 558: 113–116

6. Callaghan CT, Major RE, Wilshire JH, Martin JM, Kingsford RT, Cornwell WK (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: 845–858
7. Fenoglio MS, Rossetti MR, Videla M (2019) Negative effects of urbanization on terrestrial arthropod communities: A meta-analysis. Global Ecology and Biogeography 29: 1412–1429
8. Callaghan CT, Bowler DE, Pereira HM (2021) Thermal flexibility and a generalist life history promote urban affinity in butterflies. Global Change Biology 27: 3532–3546
9. Alberti M, Marzluff J, Hunt VM (2017) Urban driven phenotypic changes: empirical observations and theoretical implications for eco-evolutionary feedback. Philosophical Transactions of the Royal Society B 372: 20160029
10. Johnson MTJ, Munshi-South J (2017) Evolution of life in urban environments. Science 358: eaam8327
11. Santangelo JS, Rivkin LR, Johnson MTC (2018) The evolution of city life. Proceedings of the Royal Society B 285: 20181529
12. Diamond SE, Prileson EG, Martin RA (2022) Adaptation to urban environments. Current Opinion in Insect Science 51: 100893
13. Parris KM (2016) Ecology of urban environments. Wiley Blackwell, UK.
14. Miles LS, Rivkin LR, Johnson MTJ, Munshi-South J, Verrelli BC (2019) Gene flow and genetic drift in urban environments. Molecular Ecology 28: 4138–4151
15. Evans KI, Gaston KJ, Frantz AC, Simeoni M, Sharp SP, McGowan A, Dawson DA, Walasz K, Partecke J, Burke T, Hatchwell BJ (2009) Independent colonization of multiple urban centres by a formerly forest specialist bird species. Proceedings of the Royal Society B 276: 2403–2410
16. Delaney KS, Riley SPD, Fisher RN (2010) A rapid, strong, and convergent genetic response to urban habitat fragmentation in four divergent and widespread vertebrates. PLoS ONE 5: e12767
17. Unfried TM, Hauser L, Marzluff JM (2013) Effects of urbanization on Song Sparrow (*Melospiza melodia*) population connectivity. Conservation Genetics 14: 41–53
18. Blumenfeld AJ, Eyer P-A, Helms AM, Buczkowski G, Vargo EL (2021) Consistent signatures of urban adaptation in a native, urban invader ant *Tapinoma sessile*. Molecular Ecology 31: 4832–4850
19. Markowski M, Minias P, Bańbura M, Glądalski M, Kaliński A, Skwarska J, Wawrzyniak J, Zieliński P, Bańbura J (2021) Genetic structure of urban and non-urban populations differs between two common parid species. Scientific Reports 11: 10428
20. Salmón P, Jacobs A, Ahrén D, Biard C, Dingemanse N, Dominoni DM, Helm B, Lundberg M, Senar JC, Sprau P, Visser ME, Isaksson C (2021) Continent-wide genomic signatures of adaptation to urbanisation in a songbird across Europe. Nature Communications 12: 2983
21. Theodorou P, Radzevičiūtė R, Kahnt B, Soro A, Grosse I, Paxton RJ (2018) Genome-wide single nucleotide polymorphism scan suggests adaptation to urbanization in an important pollinator, the red-tailed bumblebee (*Bombus lapidarius* L.). Proceedings of the Royal Society B 285: 20172806

22. Homola JJ, Loftin CS, Cammen KM, Helbing CC, Birol I, Schultz TF, Kinnison MT (2019) Replicated landscape genomics identifies evidence of local adaptation to urbanization in wood frogs. Journal of Heredity 110: 707–719
23. Santangelo JS, et al. (2022) Global urban environmental change drives adaptation in white clover. Science 375: 1275–1281
24. Babik W, Dudek K, Marszałek M, Palomar G, Antunes B, Sniegula S (2023) The genomic response to urbanization in the damselfly *Ischnura elegans*. Evolutionary Applications 16: 1805–1818
25. Mueller JC, Carrete M, Boerno S, Kuhl H, Tella JL, Kempenaers B (2020) Genes acting in synapses and neuron projections are early targets of selection during urban colonization. Molecular Ecology 29: 3403–3412
26. Blattner LA, Kulanek D, Ruffener S, Ziegler H, Wymann H-P, Wiemers M, Michalik P, Berner D (2024) Urbanization-associated range expansion genetically homogenizes a butterfly species. Current Biology 34: 4589–4595
27. van't Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, Hall N, Darby AC, Saccheri IJ (2016) The industrial melanism mutation in British peppered moths is a transposable element. Nature 534: 102–105
28. Evans KI, Hatchwell BJ, Parnell M, Gaston K (2010) A conceptual framework for the colonization of urban areas: the blackbird *Turdus merula* as a case study. Biological Reviews 85: 643–667
29. Mueller JC, Kuhl H, Boerno S, Tella JL, Carrete M, Kempenaers B (2018) Evolution of genomic variation in the burrowing owl in response to recent colonization of urban areas. Proceedings of the Royal Society B 285: 20180206
30. Dennis RLH, Shreeve TG (1997) Diversity of butterflies on British islands: ecological influences underlying the roles of area, isolation and the size of the faunal source. Biological Journal of the Linnean Society 60: 257–275
31. Kuussaari M, Saarinen M, Korpela E-L, Pöyry J, Hyvönen T (2014) Higher mobility of butterflies than moths connected to habitat suitability and body size in a release experiment. Ecology and Evolution 4: 3800–3811
32. Skinner B, Wilson D (1984) Colour identification guide to moths of the British Isles (Macrolepidoptera). Viking, London.
33. Tolman T, Lewington R (2002) Butterflies of Europe. Princeton: Princeton University Press.
34. van Swaay CAM (2003) Butterfly densities on line transects in the Netherlands from 1990 to 2001. Entomologische Berichten 63: 82–87
35. Haahtela T, Saarinen K, Ojalainen P, Aarnio H (2012) Suomen ja Euroopan päiväperhoset. Gummerus, Helsinki.
36. Mikkola K, Jalas I, Peltonen O (1989) Suomen perhoset, Mittarit 2, Recallmed/ Lepidopterological Society of Finland, Hangon Kirjapaino, Hanko.
37. Middleton-Welling J, Dapporto L, García-Barros E, Wiemers M, Nowicki P, Plazio E, Bonelli S, Zaccagno M, Šašić M, Liparova J, Schweiger O, Harpke A, Musche M, Settele J, Schmucki R, Shreeve T (2020) A new comprehensive trait database of European and Maghreb butterflies, Papilionoidea. Scientific Data 7: 351
38. Wood BC, Pullin AS (2002) Persistence of species in a fragmented urban landscape: the importance of dispersal ability and habitat availability for grassland butterflies. Biodiversity and Conservation 11: 1451–1468

39. Merckx T, Nielsen ME, Heliölä J, Kuussaari M, Pettersson LB, Pöyry J, Tiainen J, Gotthard K, Kivelä SM (2021) Urbanization extends flight phenology and leads to local adaptation of seasonal plasticity in Lepidoptera. Proceedings of the National Academy of Sciences of the USA 118: e2106006118
40. Merckx T, Nielsen M, Kankaanpää T, Kadlec T, Yazdanian M, Kivelä SM (2023) Dim light pollution prevents diapause induction in urban and rural moths. Journal of Applied Ecology 60: 1022–1031
41. Merckx T, Nielsen ME, Kankaanpää T, Kadlec T, Yazdanian M, Kivelä SM (2024) Continent-wide parallel urban evolution of increased heat tolerance in a common moth. Evolutionary Applications 17: e13636
42. Kivelä SM, Välimäki P, Carrasco D, Mäenpää MI, Oksanen J (2011) Latitudinal insect body size clines revisited: a critical evaluation of the saw-tooth model. Journal of Animal Ecology 80: 1184–1195
43. Välimäki P, Kivelä SM, Mäenpää MI (2013) Temperature- and density-dependence of diapause induction and its life history correlates in the geometrid moth *Chiasmia clathrata* (Lepidoptera: Geometridae). Evolutionary Ecology 27: 1217–1233
44. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest DARseq: An inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PLoS ONE 7: e37135
45. Frichot E, Francois O (2015) LEA: an R package for Landscape and Ecological Association studies. Methods in Ecology and Evolution 6: 925–929
46. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Molecular Biology and Evolution 23: 254–267
47. Pickrell JK, Pritchard JK (2012) Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genetics 8: 1002967
48. Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J (2024). vegan: Community Ecology Package. R package version 2.6-6.1, URL: https://CRAN.R-project.org/package=vegan
49. Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. Genetics 145: 1219–1228
50. Slatkin M, Hudson RR (1991) Pariwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129: 555–562
51. Miller JT, Clark BW, Reid NM, Karchner SI, Roach JL, Hahn ME, Nacci D, Whitehead A (2024) Independently evolved pollution resistance in four killifish populations is largely explained by few variants of large effect. Evolutionary Applications 17: e13648
52. Brans KI, De Meester L (2018) City life on fast lanes: Urbanization induces an evolutionary shift towards a faster lifestyle in the water flea *Daphnia*. Functional Ecology 32: 2225–2240
53. Campbell-Staton SC, Winchell KM, Rochette NC, Fredette J, Maayan I, Schweizer RM, Catchen J (2020) Parallel selection on thermal physiology facilitates repeated adaptation of city lizards to urban heat islands. Nature Ecology and Evolution 4: 652–658

54. Martin RA, Chick LD, Garvin ML, Diamond SE (2021) In a nutshell, a reciprocal transplant experiment reveals local adaptation and fitness trade-offs in response to urban evolution in an acorn-dwelling ant. Evolution 75: 876–887
55. Schmitt T, Giessl A, Seitz A (2003) Did *Polyommatus icarus* (Lepidoptera: Lycaenidae) have distinct glacial refugia in southern Europe? Evidence from population genetics. Biological Journal of the Linnean Society 80: 529–538
56. Piccini I, Pellegrino I, Bellone D, Cucco M, Di Pietro V, Mignone F, Nasuelli M, Bonelli S (2024) Combined demographic ecological and genetic tools reveal connectivity within a fragmented butterfly population. Global Ecology and Conservation 54: e03095
57. Miles LS, Johnson JC, Dyer RJ, Verrelli BC (2018) Urbanization as a facilitator of gene flow in a human health pest. Modelcular Ecology 27: 3219–3230
58. Ryne C, Bensch S (2008) Do anthropogenic transports facilitate stored-product pest moth dispersal? A molecular approach. Naturwissenschaften 95: 155–159
59. Bigsby KM, Tobin PC, Sills EO (2011) Anthropogenic drivers of gypsy moth spread. Biological Invasions 13: 2077–2090
60. Yakub M, Tiffin P (2017) Living in the city: urban environments shape the evolution of a native annual plant. Global Change Biology 23: 2082–2089
61. Michaux JR, Libois R, Filipucci M-G (2005) So close and so different: comparative phylogeography of two small mammal species, the Yellow-necked fieldmouse (*Apodemus flavicollis*) and the Woodmouse (*Apodemus sylvaticus*) in the Western Paelarctic region. Heredity 94: 52–63
62. Maliouchenko O, Palmé AE, Buonamici A, Vendramin GG, Lascoux M (2007) Comparative phylogeography and population structure of European *Betula* species, with particular focus on *B. pendula* and *B. pubescens*. Journal of Biogeography 34: 1601–1610
63. Lee KM, Kivelä SM, Ivanov V, Hausmann A, Kaila L, Wahlberg N, Mutanen M (2018) Information dropout patterns in restriction site associated DNA phylogenomics and a comparison with multilocus Sanger data in a species-rich moth genus. Systematic Biology 67: 925–939
64. Eaton DAR, Overcast I (2020) ipyrad: Interactive assembly and analysis of RADseq datasets. Bioinformatics 36: 2592–2594
65. Zhang J, Kobert K, Flouri T, Stamatakis A (2014) PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics 30: 614–620
66. R Core Team (2023) R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/
67. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R, Teeling E (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution 37: 1530–1534
68. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS (2017) ModelFinder: Fast model selection for accurate phylogenetic estimates. Nature Methods 14: 587–589
69. Minh BQ, Nguyen MAT, von Haeseler A (2013) Ultrafast Approximation for Phylogenetic Bootstrap. Molecular Biology and Evolution 30: 1188–1195

70. Rambaut A (2015) FigTree, v1.4.2: tree figure drawing tool. Molecular evolution, phylogenetics and epidemiology. Available from: http://tree.bio.ed.ac.uk/software/figtree/.
71. Malinowski R, Lewiński S, Rybicki M, Gromny E, Jenerowicz M, Krupiński M, Nowakowski A, Wojtkowski C, Krupiński M, Krätzschmar E, Schauer P (2020) Automated Production of a Land Cover/Use Map of Europe Based on Sentinel-2 Imagery. Remote Sensing 12: 3523
72. Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder P, Kessler M (2017) Climatologies at high resolution for the Earth land surface areas. Scientific Data. 4: 170122.
73. Karger DN, Dabaghchian B, Lange S, Thuiller W, Zimmermann NE, Graham CH (2020) High resolution climate data for Europe. EnviDat, http://dx.doi.org/doi:10.16904/envidat.150
74. Devlin B, Roeder K (1999) Genomic control for association studies. Biometrics 55: 997–1004
75. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society B 57: 289–300
76. Excoffier L, Lischer H (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10: 564–567

Acknowledgements

We are grateful to Ulrich Hiermann, Toni Mayr, Christin Wieser, and Manuel Vilgut for providing us with some *C. clathrata* specimens, to Anna Antinoja, Francesca Cochis, and Irene Piccini for helping us during fieldwork, to Laura Törmälä for carrying out ddRAD library preparation in the laboratory, and to Soile Alatalo and Hannele Parkkinen for assistance in DNA extraction. We also acknowledge the use of image sources from the Finnish Biodiversity Information Facility (FinBIF/laji.fi). The IT Center for Science (CSC) provided computational resources for bioinformatics and data storage services for the study.

Author contributions

S.M.K conceived the study. S.M.K., G.C.A., S.B., P.H., T. Kadlec, T. Kankaanpää, G.M., M.E.N., E.P., T.T., K.Z., and T.M. collected samples, with the sampling being coordinated by S.M.K. and T.M., K.M.L. did bioinformatics, and data were analysed by S.M.K., K.M.L., and F.H., with T. Kankaanpää extracting and handling land cover and climate data and helping in data visualization. The manuscript was initially drafted by S.M.K., K.M.L., F.H., and T.M. with input from all the other authors.

Funding

This work was Financed by the Academy of Finland (Grant Nos. 314833, 319898, and 345363 to S.M.K.) and the Kvantum Institute at the University of Oulu. Frederik Hendrickx and Thomas Merckx extend their gratitude to the Flemish Research Fund for their support of the Scientific Research Network EVENET (W001322N).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information file.