RESEARCH PAPER

The compositional similarity of urban forests among the world's cities is scale dependent

Jun Yang[1,2]*, Frank A. La Sorte[3], Petr Pyšek[4,5,6], Pengbo Yan[7], David Nowak[8] and Joe McBride[9]

[1]*Ministry of Education Key Laboratory for Earth System Modeling, Center for Earth System Science, Tsinghua University, 100084 Beijing, China,* [2]*Joint Center for Global Change Studies (JCGCS), Beijing 100875, China,* [3]*Cornell Laboratory of Ornithology, Cornell University, Ithaca, NY 14850, USA,* [4]*Institute of Botany, Department of Invasion Ecology, The Czech Academy of Sciences, CZ-252 43 Průhonice, Czech Republic,* [5]*Department of Ecology, Faculty of Science, Charles University in Prague, CZ-128 44 Prague, Czech Republic,* [6]*Centre for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, 7602 Matieland, South Africa,* [7]*College of Forestry, Beijing Forestry University, 100083 Beijing, China,* [8]*Northern Research Station, c/o SUNY ESF, Syracuse, NY 13210, USA,* [9]*Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720, USA*

*Correspondence: Jun Yang, Center for Earth System Science, Tsinghua University, S1013 Meng Man Wai Science Building, Beijing 100084, China.
E-mail: larix001@gmail.com

ABSTRACT

Aim We examined species composition of urban forests from local to global scales using occurrence and abundance information to determine how compositional similarity is defined across spatial scales. We predicted that urban forests have become more homogeneous world-wide, which should result in minimal scale dependence that is more pronounced for non-native species, especially when considering abundance information.

Location Thirty-eight cities world-wide.

Methods We estimated compositional dissimilarities of urban forests, including both spontaneous and cultivated trees, from local to global spatial scales using six dissimilarity metrics. We used redundancy analysis to determine how climate, geographic distance and anthropogenic factors are related to compositional dissimilarity among cities. These analyses were implemented for all species combined and for native and non-native species separately.

Results The 38 cities contained a median of 77 tree species, with a greater percentage of these classified as native (median = 58%). The similarity of urban forests was scale dependent, declining as the spatial scale increased – an outcome that did not differ when considering native and non-native species separately. Climate, geographic distance and city age were the main factors describing variation in tree species composition among cities. The addition of abundance information resulted in lower dissimilarity across spatial scales.

Main conclusions Compositional similarity of urban forests is a scale-dependent phenomenon that is not affected by the presence or absence of non-native species, suggesting a limited role for biotic interchange in promoting homogenization. However, compositional similarity across spatial scales increased uniformly with the addition of abundance information, suggesting that patterns of abundance may have greater biological relevance when homogenization trends among urban forests are considered.

Keywords

Abundance, biotic homogenization, composition dissimilarity, occurrence, trees, urban areas.

INTRODUCTION

Increasing genetic, taxonomic and functional similarity among plant and animal communities has been observed in different regions of the world. McKinney & Lockwood (1999) defined this phenomenon as biotic homogenization. Biotic homogeni-zation can be considered as an indicator of the impact of anthropogenic activities on natural communities (McKinney & Lockwood, 1999), and is regarded as one of the most prominent forms of biotic impoverishment occurring world-wide, with important ecological and evolutionary consequences (Olden *et al.*, 2004).

Urbanization is a major anthropogenic disturbance that can lead to biotic homogenization (McKinney & Lockwood, 1999; McKinney, 2006). Homogenization of urban floras has been reported within cities (Trentanovi *et al.*, 2013) and among cities at regional (Kühn & Klotz, 2006), continental (La Sorte *et al.*, 2008; Lososová *et al.*, 2012a,b; Ricotta *et al.*, 2012) and intercontinental spatial scales (La Sorte *et al.*, 2007, 2014; Aronson *et al.*, 2014). These studies suggest that the compositions of urban floras are becoming more similar across spatial scales. Environmental filtering and human preference are considered to be two major drivers underlying the homogenization of urban floras (McKinney, 2006; Williams *et al.*, 2009). Similar environmental conditions among different regions favour species with particular traits (Duncan *et al.*, 2011; Williams *et al.*, 2015). Intentional introduction and transportation of species by humans help spread non-native species and can lead to increased similarity or differentiation of urban floras depending on mediating factors such as the spatial distances between urban areas (McKinney, 2008), planting intensity (Loeb, 2012), environmental conditions and residence times of non-native plants (Lososová *et al.*, 2012a). Other variables that mediate the effects of the two driving forces of homogenization include climate, human population size, the age of the city and the spatial extent of the city (Qian & Ricklefs, 2006; Winter *et al.*, 2010; La Sorte *et al.*, 2014).

Nevertheless, there are still questions as to whether the homogenization trend of urban floras is a ubiquitous global phenomenon. Findings from recent studies examining urban forests do not support this generalization. For example, no homogenization effect was detected in the species compositions of urban forests among 27 US cities occurring in different biomes (Ramage *et al.*, 2013). This finding contradicts the conclusions of another study that found that urban forests in eight US cities displayed less regional variation in species composition than natural forests (Nock *et al.*, 2013). Unlike other studies that only examined spontaneous floras, Ramage *et al.* (2013) and Nock *et al.* (2013) included both spontaneous and cultivated species in their comparisons. These studies raised the question as to whether the factors that contribute to the homogenization of spontaneous urban floras also have an impact on urban forests when both spontaneous and cultivated tree species are considered. The homogenization effect at the intercontinental scale is also divergent in plant assemblages based on species places of origin. La Sorte *et al.* (2014) found that invasive species and European archaeophytes had similar compositions, but native species were unique among the 110 studied cities. It is clear that more studies are needed to determine how homogenization patterns among the world's urban floras are defined across spatial scales and the role of native and non-native species in defining these patterns. Multi-scale studies at intercontinental and global spatial extents that include regions other than Europe and North America are especially needed (Castro & Jaksic, 2008; La Sorte *et al.*, 2014).

In addition to conducting more studies, a better understanding of how biotic homogenization may be occurring among urban floras can be obtained by examining different kinds of biological data. Currently, the majority of studies use species occurrence data compiled from various sources including field surveys and herbarium records. The varied spatial and temporal grains of these data sources may contribute to potential biases in how biotic homogenization is estimated (Olden & Rooney, 2006). Furthermore, there are different opinions on whether the use of occurrence data alone affects the characterization of biotic changes. Some researchers suggest that the lack of information on species abundance does not change the basic results in most cases (Cassey *et al.*, 2008), while others suggest that a reliance on species occurrence data might constrain our understanding of the mechanisms contributing to homogenization (Olden & Poff, 2003; McKinney & La Sorte, 2007). The inclusion of abundance data could provide a better understanding of the influence of non-native species on homogenization (La Sorte & McKinney, 2007). McKinney (2008) therefore recommended that studies of taxonomic homogenization should include species abundance data wherever possible.

In this study we ask the following questions: (1) is the composition of species among urban forests (including both spontaneous and cultivated tree species) similar across spatial scales, and (2) do these patterns differ when we examine native and non-native species separately? We compared the species composition of urban forests in 38 cities world-wide. Our objectives were as follows: (1) to test for the presence of globally homogeneous urban forests; (2) to examine the role of native and non-native species in these patterns; (3) to examine the influence of climatic, geographic and anthropogenic factors on the dissimilarity of species composition among urban forests; and (4) to contrast the patterns derived from abundance and occurrence data.

METHODS

Data collection

We compiled abundance data of urban tree species for 38 cities (30 in North America, five in China, two in Europe and one in South America) (Fig. 1). To minimize sampling biases, we used tree abundance data that were collected using the same sampling protocol in all cities. All data were collected from 1996 to 2012 using the same field protocol developed by Nowak *et al.* (2005). The protocol recommended that investigators randomly distribute 150 to 200 0.1-acre (*c.* 405 m^2) circular plots in each city. Nowak *et al.* (2008) showed that a sample size of 200 would give population estimates of urban forests within an average relative standard error of 12% for 14 studied US cities, including New York and Philadelphia. All spontaneously occurring and cultivated tree species encountered in the sample plots were identified and counted. Details on the data sources, sampling time and sample size for each city are included in Appendix S1 in the Supporting Information.

We first removed records from the data that included unknown species and those only recorded at the genus level. The median percentage of removed trees to the total number of trees in each city was 4%. We then converted the counts of trees in each city to relative abundances, with the total value summing to
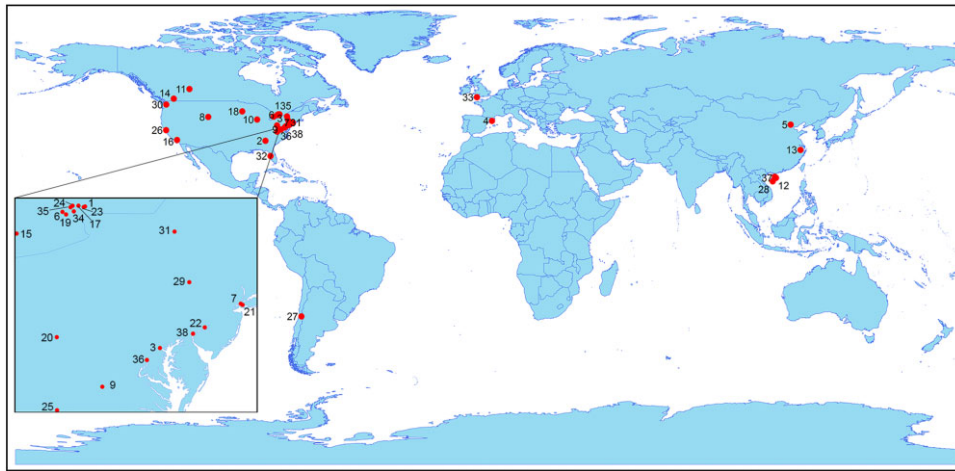
Figure 1 Geographical distribution of the 38 studied cities. The inset shows details of cities in the north-eastern USA and south-western Ontario, Canada. Numbers are: 1, Ajax; 2, Atlanta; 3, Baltimore; 4, Barcelona; 5, Beijing; 6, Brampton; 7, Brooklyn; 8, Casper; 9, Charlottesville; 10, Chicago; 11, Edmonton; 12, Haikou; 13, Hangzhou; 14, Kelowna; 15, London; 16, Los Angeles; 17, Markham; 18, Minneapolis; 19, Missauga; 20, Morgantown; 21, New York; 22, Philadelphia; 23, Pickering; 24, Richmond Hill; 25, Roanoke; 26, San Francisco; 27, Santiago; 28, Sanya; 29, Scranton; 30, Seattle; 31, Syracuse; 32, Tampa; 33, Torbay; 34, Toronto; 35, Vaughan; 36, Washington, DC, 37, Wenchang; 38, Wilmington.

one. We merged all varieties and subspecies identified in each city into the species rank and standardized taxonomic nomenclature using the Plant List (http://www.theplantlist.org, accessed July 2014). Species in North American cities were classified as either native or non-native by referring to the Plants Database (http://plants.usda.gov, accessed July 2014). For the five cities in China, we used the Flora of China (http://www.efloras.org, accessed July 2014) to classify species as native or non-native. For Santiago (Chile) and Barcelona (Spain) we adopted the classifications provided by the investigators. We classified the native status for tree species in Torbay (UK) by referring to PLANTATT (Hill *et al.*, 2004).

Compositional analysis

We compared the floristic dissimilarity among cities using multiple-site dissimilarity indices. First, we calculated total beta diversity (BD_{Total}), which is an unbiased estimator of the total variance of the community data (Legendre & De Cáceres, 2013). BD_{Total} can be further partitioned into two components: (1) species contributions (SCBD), which shows degree of variation of individual species across the study area, and (2) local contributions (LCBD), which are comparative indicators of the ecological uniqueness of the sites. In this study, we calculated the values of BD_{Total} and LCBD using abundance-based Sørensen and Jaccard metrics. Then the abundance data for trees were transformed to presence/absence. BD_{Total} and LCBD were again calculated using occurrence-based Sørensen and Jaccard metrics. For BD_{Total} calculated using the Sørensen and Jaccard related metrics, values had a range from zero to 0.5 with larger values indicating higher dissimilarity. To facilitate the comparison of results among different methods, we converted BD_{Total} to relative BD_{Total} (BD_{Rel}) by multiplying the values of BD_{Total} by two

(Legendre & De Cáceres, 2013). The significance of LCBD was tested by using a permutation procedure with 999 runs.

We also used the presence/absence data to calculate the multiple-site versions of the Sørensen ($M\beta_{SOR}$) and Jaccard ($M\beta_{JAC}$) dissimilarity indices (Baselga, 2013). Both indices have values between zero and one with values approaching one indicating higher dissimilarity. We adopted six kinds of multiple-site dissimilarity indices in this study, for two reasons: (1) to facilitate the comparison of abundance-based and occurrence-based methods, and (2) to exclude the possibility that the trend detected by a particular dissimilarity index was an artefact of the particular method.

We applied the analysis to four geographic classifications of cities: (1) intercontinental, which considered all 38 cities combined; (2) continental, the 30 North American cities; (3) regional, the 24 cities in the Eastern Temperate Forests (ETF) ecoregion of North America (CEC, 1997); and (4) local, the nine cities in south-western Ontario. The nine cities in south-western Ontario could be treated as independent sample units because they are municipalities at the same administrative level and the urban forests in these cities are managed independently. Two classifications of tree species were considered in the analysis: native and non-native. We used permutation tests to estimate the likelihood of observed differences in compositional dissimilarity among cities occurring by chance alone between native and non-native species. We randomly shuffled native and non-native classifications among tree species in each city independently. After each permutation procedure, the dissimilarity indices were recalculated for native and non-native species. The difference between these two estimates was then calculated. The permutation procedure was implemented 1000 times and a one-sided *P*-value was estimated following the method in La Sorte *et al.* (2014).

To examine the dissimilarity between any two urban forests, we calculated the paired abundance-based and paired occurrence-based Sørensen indices for all 38 cities. We used Ward's hierarchical clustering method to group urban forests in the 38 cities into clusters based on their dissimilarity of species compositions. We also conducted a paired *t*-test to assess the significance of the difference in observed values between the two sets of indices.

To examine the relationship between abundance and occupancy, we conducted a Spearman rank correlation test between the mean relative abundance and occupancy for tree species in different geographic classifications. We first applied a square root transformation to all abundance estimates. Then we calculated the mean relative abundance and standard deviation for each tree species. The number of cities where the species occurred was summed to provide an estimate of occupancy.

Analysis of influencing factors

To identify factors that contributed to the observed variation in species composition of different urban forests, we conducted a forward-selection redundancy analysis (RDA) including climate, geographic distance, human population size, the length of the time that a city has been established (city age) and the land areas of the cities. We determined city age by subtracting the year the city was officially founded from 2014. Except for climatic variables, all of these variables were log-transformed before analysis to improve their distributional properties. We also included sampling effort, defined as the ratio of sampled areas to the total area of the city, to determine whether sampling effort contributed to the observed difference in community compositions. The RDA analysis was applied to all geographic classifications, excluding the nine cities in south-western Ontario because of small sample size.

We obtained six bioclimatic variables from WorldClim (Hijmans *et al.*, 2005): mean annual temperature, mean coldest monthly temperature, temperature seasonality, mean annual precipitation, precipitation of driest month and seasonality of precipitation. The data were interpolated from weather records between 1950 and 2000 and saved in a gridded format with a spatial resolution of 30 arcsec (*c.* 1 km at the equator). We used the boundaries of the 38 cities obtained from the Global Rural–

Urban Mapping Project (GRUMP, v.1) to clip out the corresponding grid cells for the six climate variables (CIESIN *et al.*, 2011). The values contained in these grid cells for each city were then averaged. Following Qian & Ricklefs (2006), we conducted a principal component analysis (PCA) to reduce the collinearity between the bioclimatic variables.

To estimate geographic distances, we calculated the average geodesic distances from each city to all the remaining cities. We obtained estimates of city area and human population size for US cities from the US Census Bureau (http://Factfinder.census.gov, accessed May 2014). For the remaining cities, we retrieved data from each city's official website and from reports produced by investigators. The founding years of the cities were collected from various sources, mainly the cities' official websites.

All data analysis in this study was implemented using R version 3.1.0 (R Development Core Team, 2014). The function Beta.div developed by Legendre & De Cáceres (2013) was used to calculate BD_{total} and LCBD. The R package betapart developed by Baselga (2013) was used to calculate $M\beta_{SOR}$ and $M\beta_{JAC}$. The geodesic distances between cities were calculated using the R package fields (Fields Development Team, 2006). The forward-selection RDA analysis was implemented using the R package packfor developed by Dray *et al.* (2009). The PCA analysis was conducted using the base function of R.

RESULTS

Species richness

The 38 cities contained a total of 748 tree species. The median number of tree species was 77. The median percentage of native tree species was 58%, while the median percentage of non-native tree species was 42% (Fig. 2). On average the 38 cities had a statistically significant proportion of native tree species that was greater than that of non-native tree species (Welch two-sample *t*-test, $t = 3.383$, d.f. = 76, $P < 0.01$).

Compositional dissimilarity

Values of the six multiple-site dissimilarity indices all showed that dissimilarities of species compositions among urban forests
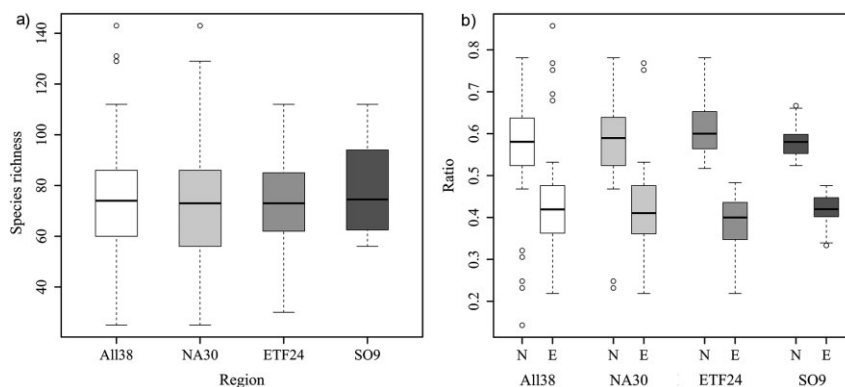

Figure 2 Summary of tree species in the 38 studied cities. (a) Species richness, and (b) proportions of native (N) and non-native (E) tree species of urban forests in all 38 cities (All38), the 30 North American cities (NA30), the 24 cities in the Eastern Temperate Forest ecoregion (ETF24) and the nine cities in south-western Ontario (SO9).

Table 1 Multiple-site dissimilarity indices for four geographic classifications of cities: all 38 cities (All38), 30 North American cities (NA30), 24 cities in the Eastern Temperate Forest ecoregion (ETF24) and nine cities in south-western Ontario (SO9)

Geographaic classification	Species status	Abundance based		Occurrence based			
		BD_{rel} Absor	BD_{rel} Abjac	BD_{rel} Sor	BD_{rel} Jac	$M\beta_{SOR}$	$M\beta_{JAC}$
All38	All	0.551	0.647	0.743	0.837	0.950	0.974
	Native	0.608	0.674	0.762	0.843	0.951	0.973
	Non-native	0.623	0.720	0.785	0.868	0.956	0.978
NA30	All	0.367	0.486	0.641	0.769	0.925	0.961
	Native	0.375	0.481	0.622	0.750	0.916	0.956
	Non-native	0.466	0.648	0.670	0.814	0.937	0.967
ETF24	All	0.191	0.303	0.549	0.700	0.884	0.939
	Native	0.182	0.290	0.514	0.670	0.876	0.934
	Non-native	0.335	0.473	0.634	0.767	0.908	0.951
SO9	All	0.019	0.055	0.408	0.578	0.693	0.818
	Native	0.009	0.028	0.363	0.530	0.660	0.796
	Non-native	0.084	0.179	0.472	0.638	0.731	0.845

Absor, abundance-based Sørensen; Abjac, abundance-based Jaccard; Sor, Sørensen; Jac, Jaccard.

increased with increasing spatial scale (Table 1). As we expanded the spatial extent of our study area from the 24 cities in the ETF ecoregion to the 30 cities in North America, the multiple-site dissimilarity indices estimated using abundance data almost doubled. The six multiple-site dissimilarity indices were highly correlated, with a mean Pearson correlation coefficient of 0.95. However, values of occurrence-based indices were constantly higher than abundance-based indices for all spatial scales. In all geographic classifications, dissimilarities of non-native tree assemblages did not differ significantly from those of native tree assemblages (*P*-values of 0.104–1.0).

Values of LCBD indicated that native tree assemblages in the eight cities outside North America were unique compared with those in North America (Fig. 3a). Nevertheless, when comparing non-native tree assemblages, the two European cities did not differ significantly from cities in the eastern part of North America (Fig. 3b). The species compositions of urban forests in Los Angeles, San Francisco and Tampa were significantly different from the other North American cities. As the geographic extents decreased, the number of urban forests with significant LCBD values also decreased (Fig. 3c–f). For example, of the 24 cities located inside the ETF ecoregion, only Tampa had unique compositions of native and non-native tree species (Fig. 3e, f).

The means of paired abundance-based and occurrence-based Sørensen dissimilarity indices for all 38 cities were 0.678 and 0.854, respectively. The difference between the two sets of paired dissimilarity indices was significant (*t* = −24.514, d.f. = 702, *P* < 0.001). The clustering results based on the dissimilarity indices both indicated that there were four distinctive groups of cities. Three groups largely followed geographic boundaries: three cities in south China, cities in the north-eastern USA, and a group of mainly Canadian cities. The fourth group, a mixture of cities from Europe, China and the USA (Fig. 4), had assemblages of native and non-native tree species that were significantly different from the other cities (Fig. 3a, b).

When plotting the transformed relative abundance and occupancy of the tree species in all 38 cities together (Fig. 5a), an L-shaped pattern was obtained, with most species concentrated at the low end of the occupancy distribution where relative abundances had the greatest variability. Several species such as arborvitae (*Thuja occidentalis* L.), boxelder (*Acer negundo* L.), honey locust (*Gleditsia triacanthos* L.) and red maple (*Acer rubrum* L.) were widespread (in more than 24 cities) but their mean relative abundances were less than 5% in cities where they occurred. The value of correlation coefficient revealed a weak correlation among the two variables (ρ = 0.289, *P* < 0.0001). Variability in estimates of relative abundance was explained marginally by occupancy (Fig. 5b; adjusted R^2 = 0.05, *P* < 0.0001). The correlations became stronger as the spatial extent decreased to 30 North American cities (ρ = 0.285, *P* < 0.0001; adjusted R^2 = 0.128, *P* < 0.0001), 24 cities in the ETF region (ρ = 0.494, *P* < 0.0001; adjusted R^2 = 0.290, *P* < 0.0001) and nine cities in south-western Ontario (ρ = 0.658, *P* < 0.0001; adjusted R^2 = 0.324, *P* < 0.0001).

Influencing factors

The first two principal components of the PCA, accounting for 90% of total variance (Appendix S2), were retained for further analysis. PC1 represented a north–south gradient of temperature and precipitation. PC2 mainly represented an east–west gradient of precipitation.

When assessing the species composition of native tree assemblages in all 38 cities, forward selection RDA produced a model containing four significant predictors at α = 0.05: PC1, PC2, city age and the average geodesic distance (Fig. 6a). For the species composition of non-native tree assemblages, only the climate variables were significant predictors (Fig. 6b). Climate variables and geodesic distances contributed significantly to predicting variations among native tree assemblages in all cities. Climate and city age contributed significantly to predicting the variation
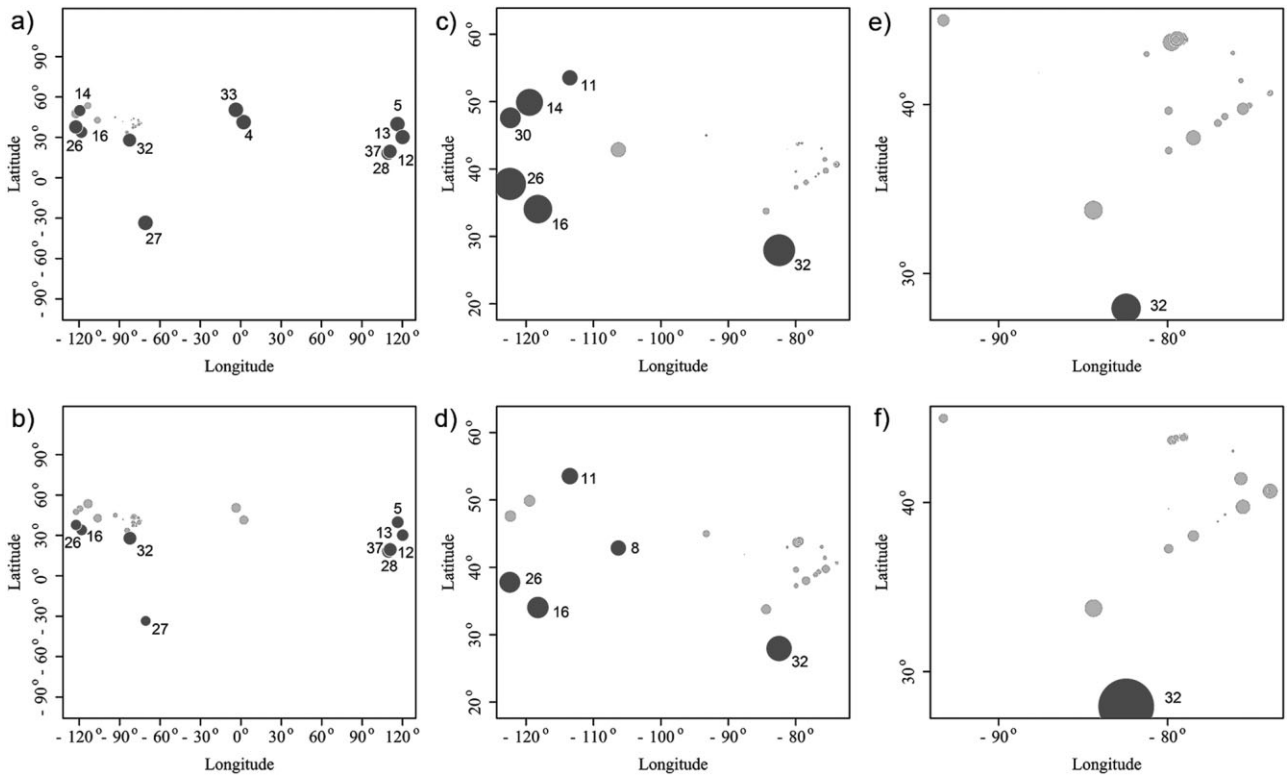
Figure 3 (a) The local contributions to beta diversity (LCBD) of the native tree assemblages in all 38 cities, (b) the LCBD of the non-native tree assemblages in all 38 cities, (c) the LCBD of native tree assemblages in the 30 North American cities, (d) the LCBD of non-native tree assemblages in the 30 North American cities, (e) the LCBD of native tree assemblages in the 24 cities in the Eastern Temperate Forest ecoregion, and (f) the LCBD of non-native tree assemblages in the 24 cities in the Eastern Temperate Forest ecoregion, all plotted against latitude and longitude. The LCBD values were calculated using the abundance-based Sørensen method. The size of the circles is proportional to the LCBD values. Black represents significant LCBD indices at the 0.05 significance level. Numbers stand for the names of cities, given in the caption of Fig. 1.
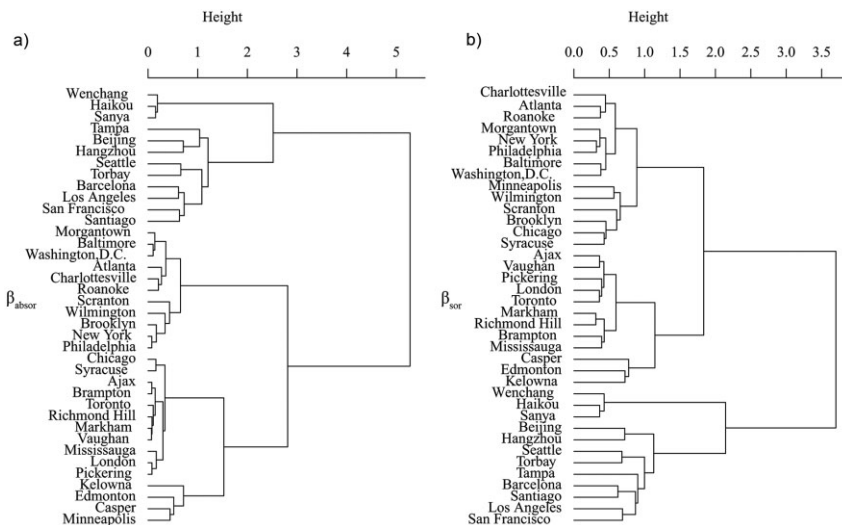

Figure 4 Dendrograms of urban forests based on their paired dissimilarity indices: (a) the dendrogram from the abundance-based Sørensen dissimilarity indices, and (b) the dendrogram from the occurrence-based Sørensen dissimilarity indices.

among non-native tree assemblages in the 30 North American cities (Fig. 6c, d). For the 24 cities in the ETF, climate again contributed significantly to predicting the variation in species composition for both native and non-native tree assemblages (Fig. 6e, f).

DISCUSSION

Our results show that the compositional dissimilarity of urban trees among cities varies based on the spatial scale considered. Dissimilarity is highest at the intercontinental scale and

Figure 5 The mean (a) and standard deviation (b) in relative abundance as a function of occupancy for all urban tree species in the 38 studied cities. A square root transformation was applied to the relative abundance. The mean and standard deviation of the relative abundance for each species were calculated over cities where the species occurred. Occupancy was estimated as the number of cities where each species occurred. The standard deviation was fitted with a simple linear regression.
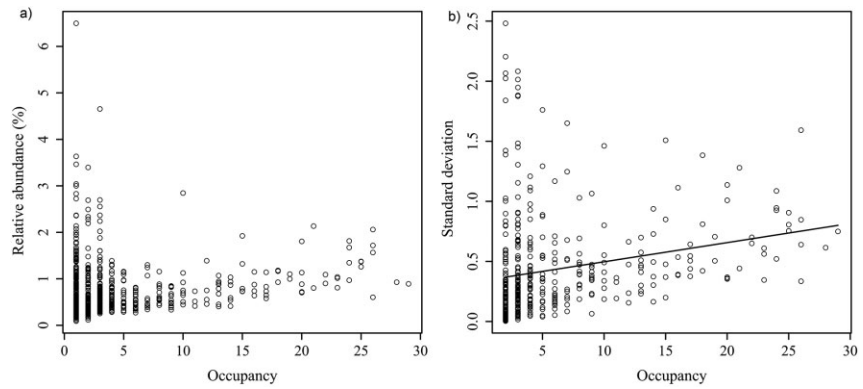
Figure 6 Correlation triplots based on a redundancy analysis (RDA) depicting the relationship between the selected geographic, climatic and anthropogenic variables and the variation of species composition among different tree assemblages. (a) Native tree assemblages in all 38 cities. (b) Non-native tree assemblages in all 38 cities. (c) Native tree assemblages in the 30 North American cities. (d) Non-native tree assemblages in the 30 North American cities. (e) Native tree assemblages in the 24 cities in the Eastern Temperate Forest ecoregion. (f) Non-native tree assemblages in the 24 cities in Eastern Temperate Forest ecoregion. Avedis, average geodesic distances; Etime, length of time that a city has been established. Values of adjusted R^2 are shown in the figure.
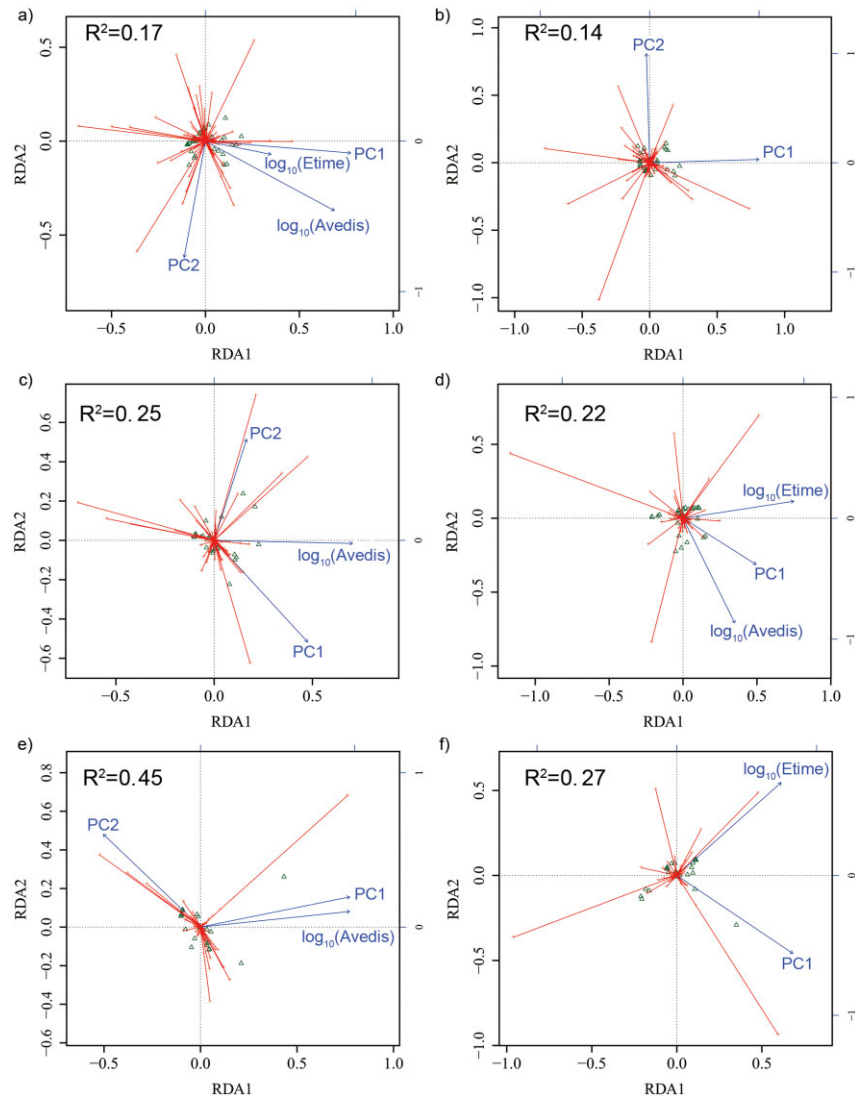
decreases at the regional and local scales. This finding supports Olden & Poff's (2003) conclusion that biotic homogenization is a scale-dependent phenomenon. However, contrary to other studies that have examined compositional similarity among plant assemblages (Mokany *et al.*, 2013), our results indicate that similarity among urban forests is greater at smaller spatial scales, suggesting a unique scale dependence for these particular urban systems. We also found evidence that these scale-dependent patterns did not differ between native and non-native species. This finding replicates those found with eight urban floras located in the ETF ecoregion of the USA (La Sorte & McKinney, 2006) but not those found for urban floras in

Europe (La Sorte *et al.*, 2008), where non-native species were associated with more homogeneous species compositions across scales.

Our results offer explanations for the differing conclusions reached by Nock *et al.* (2013) and Ramage *et al.* (2013) on biotic homogenization of urban forests. The eight cities studied by Nock *et al.* (2013) overlapped with the 24 ETF cities examined in our study. We found that species composition of urban forests in the eight cities did not differ significantly. In fact, the compositional dissimilarity among all 24 urban forests in the ETF ecoregion was low except for Tampa, which is located at the boundary between the ETF ecoregion and the Tropical Humid Forests ecoregion. Therefore, our results supported Nock *et al.*'s (2013) conclusion that the urban forests were becoming similar in this region. The six additional cities all differed significantly from those 24 in the ETF ecoregion, in species composition of either native or non-native tree assemblages. A closer look showed that those six cities belonged to five different ecoregions: Great Plains, Temperate Sierras, Mediterranean California, Marine West Coast Forests and North American Deserts. The 27 cities studied by Ramage *et al.* (2013) belonged to nine biome classes. Based on our results, we would expect the dissimilarity among the 27 urban forests to be even larger. We corroborated their conclusion that the compositions of urban tree species in different biomes are still significantly different.

One pattern captured by our study was that the species compositions of non-native tree assemblages in the two European cities were not significantly different from those of non-native tree assemblages in the ETF ecoregion, except for Tampa. The two European cities and cities in the EFT ecoregion shared high proportions of non-native tree species that originated from Asia or Eurasia. About 43% and 36% of non-native species in Barcelona and Torbay, respectively, originated in Asia or Eurasia. For the 23 cities in the ETF ecoregion, excluding Tampa, 70% of non-native tree species had Asian or Eurasian origins. La Sorte *et al.* (2014) found that European archaeophytes increased homogenization of urban floras between European and North American cities. Here we suspect that tree species originating from Asia or Eurasia played a similar role for urban forests at an intercontinental scale. This possibility needs to be verified in future studies examining a larger number and distribution of cities.

We found that climate and geographic distances made a significant contribution to the variation of species compositions among native tree assemblages across all spatial scales. This supports the view that the distribution of native species was constrained primarily by ecological factors (Qian & Ricklefs, 2006; Winter *et al.*, 2010; La Sorte *et al.*, 2014). For non-native tree species, climate was still a significant factor across all spatial scales. Qian & Ricklefs (2006) found that there was no strong climatic signal in the distribution of non-native species, because non-native species occupied habitats defined more closely by human activities. Conversely, Winter *et al.* (2010), Lososová *et al.* (2012b) and La Sorte *et al.* (2014) identified climate as a significant factor influencing compositional dissimilarity among non-native plant assemblages. Winter *et al.* (2010) sug-

gested that successful establishment and spread of non-native species were determined by their climatic suitability. It is believed that climate acts as a long-term filter, removing non-native tree species that are poorly adapted to local climates (Loeb, 2011). As a consequence, decisions on planting non-native trees in human-made habitats are often guided by an understanding of their climatic suitability (e.g. USDA plant hardiness zone maps).

The effects of other variables on the variation of species composition were not consistent at different spatial scales and reflected the complicated processes involved in shaping the composition of urban forests. City age only had a marginally significant effect for native tree assemblages at the intercontinental scale ($P = 0.049$) and was not significant at the other two spatial scales. The lack of a significant effect of city age on the compositional dissimilarity of native tree assemblages in North America and the ETF ecoregion is likely a reflection of the fact that North American cities retain regionally distinct remnants of native floras (La Sorte *et al.*, 2014). In contrast, the significant effect of city age on non-native tree assemblages may reflect the importance of the length of time that an area has been subject to human disturbance and the importation of non-native plants by humans (Pyšek *et al.*, 2009; Lososová *et al.*, 2012a; Kowarik *et al.*, 2013). The marginally significant relationship at the intercontinental scale may have been caused by the five Chinese cities; the longer establishment times for these cities were often associated with more significant habitat modifications through intensive change in land use (Lin & Ho, 2003). The inconsistent effect of geographic distance on the dissimilarity of non-native species assemblages, the lack of influence by variables considered as significant in other studies (e.g. population size) and the low explained variation suggest that the compositional dissimilarity among urban forests is affected by additional factors not considered here. For example, composition of different land use types (Kendal *et al.*, 2014) and commercial availability of species (Pincetl *et al.*, 2013) have the potential to affect species diversities of urban forests.

The high correlations among different types of multiple-site dissimilarity indices indicated that abundance- and occurrence-based information responded to the change of dissimilarities at different spatial scales monotonically. However, the dissimilarity indices estimated from abundance data were consistently lower than those estimated from occurrence data. For example we would expect urban forests from the nine cities in south-western Ontario to have highly similar species compositions because they are subjected to similar natural and anthropogenic factors. The values of the multiple-site dissimilarity indices based on occurrence data were still high (e.g. $M\beta_{JAC} = 0.818$), while those estimated using abundance-based methods were closer to our expectations (e.g. BD_{Rel}, abundance-based Jaccard = 0.055). Patterns revealed by paired dissimilarity indices using abundance data were also a better reflection of reality than those using occurrence data. For example, because Brooklyn is part of New York City, we would expect the compositional dissimilarity between the two urban forests to be low and to be caused mainly by sampling effects. In the dendrogram of clusters, the two

locations were clustered together on the basis of abundance-based Sørensen dissimilarity indices, but these locations separated in clusters formed on occurrence-based Sørensen dissimilarity indices. Thus, as has been identified in other studies (McKinney & La Sorte, 2007), the addition of abundance information resulted in a stronger pattern of homogenization.

The addition of abundance information can potentially provide better descriptions of compositional similarity among urban forests and the biological relevance of the observed patterns. First, our analysis of abundance and occupancy showed that the abundance of species was not well correlated with occupancy from regional to intercontinental scales. Non-native species with low abundance and low occupancy were prevalent in our data. This reflects the fact that urban forests often contain a large number of non-native species planted in small quantities (Sjöman et al., 2012). This situation could have produced biases in our occurrence-based results because highly abundant and rare species were treated equally (Chao et al., 2006). It also partly explained why we did not find evidence that non-native species were associated with more homogeneous species compositions across scales (La Sorte et al., 2008).

The abundance information on tree species also helps to explain the observed scale-dependent pattern of compositional dissimilarity. At the local scale, because species were from similar species pools (e.g. natural vegetation, nurseries) and were subject to similar human management, there was less variation in the levels of abundance of species, leading to a higher level of similarity across cities. As the spatial extent increased, tree species were taken from a more diverse collection of species pools, resulting in a greater range of abundance. This is in agreement with an early finding that species introduced from nearby sources have a more homogenizing effect than species from distant sources (McKinney, 2005). Although mainly through human activities, some landscape tree species achieved wide occupancy among cities but their influences on homogenization were limited due to low local abundance.

Finally, abundance-based Sørensen and Jaccard indices could correct for an under-sampling bias in sample data (Chao et al., 2006; Legendre & De Cáceres, 2013). Under-sampling can be a common issue when sampling urban floras due to the high heterogeneity of urban landscapes (Zhao et al., 2010).

Several limitations of the current study should be kept in mind when interpreting the results. In our investigation we strove to remove the influence of spatial and temporal variation by using a uniform sampling approach which resulted in a total sample size that was smaller than those used in other studies (Aronson et al., 2014; La Sorte et al., 2014). Geographically, the data were dominated by North American urban forests – the inclusion of more data from other regions of the world would help address questions concerning the global ubiquity of biotic homogenization. Another issue is that our city-wide surveys included both spontaneous trees and cultivated trees. Spontaneous trees are affected more by the filtering effect of urban environments, while human preferences have a dominant effect in defining the composition of cultivated trees. Trees that occur spontaneously do not generally form large stands in urban areas,

except in places that receive little human attention, such as wastelands (Kunick, 1990). Loeb (2011) found that many of the native tree species in Philadelphia and New York City that were reported to have grown spontaneously could have originated from horticultural plantings. Indeed, when doing ad hoc surveys it is virtually impossible to separate trees that were originally planted in a given site from those that are the result of spontaneous dispersal. Because of these factors, the inclusion of both spontaneous and cultivated trees in the study was necessary, especially when abundance data were used. The inclusion of cultivated trees makes it more difficult to explain the variation of compositional dissimilarity of urban trees in different regions. Nevertheless, our results showed that the scale-dependent pattern was strong enough not to be hidden by data pooled across the pathway of introduced trees.

In summary, our results showed that compositional similarity of urban forests is a scale-dependent phenomenon that is not affected by the addition of non-native species, and that the addition of abundance information resulted in a uniform decline in compositional dissimilarity across spatial scales. Because urban forests represent a component of urban floras that is strongly influenced by the selection pressure of human activities, we argue that the global transportation and introduction of plant species by humans has not led to globally homogeneous urban floras. However, as globalization increases together with the rate of biotic exchange we expect that the biotic homogenization of urban floras across the globe will increase, especially through changes in patterns of abundance, suggesting the need for additional studies focusing on developing richer spatial and especially temporal perspectives.

REFERENCES

Aronson, M.F., La Sorte, F.A., Nilon, C.H. 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: Biological Sciences, **281**, 20133330.

Baselga, A. (2013) Multiple site dissimilarity quantifies compositional heterogeneity among several sites, while average pairwise dissimilarity may be misleading. Ecography, **36**, 124–128.

Cassey, P., Lockwood, J.L., Olden, J.D. & Blackburn, T.M. (2008) The varying role of population abundance in structuring indices of biotic homogenization. *Journal of Biogeography*, **35**, 884–892.

Castro, S.A. & Jaksic, F.M. (2008) How general are global trends in biotic homogenization? Floristic tracking in Chile, South America. *Global Ecology and Biogeography*, **17**, 524–531.

CEC. (1997) *Ecological regions of North America: toward a common perspective*. Commission for Environmental Cooperation, Quebec, Canada.

Chao, A., Chazdon, R.L., Colwell, R.K. & Shen, T.J. (2006) Abundance-based similarity indices and their estimation when there are unseen species in samples. *Biometrics*, **62**, 361–371.

CIESIN, Center for International Earth Science Information Network, C.U., International Food Policy Research Institute & the World Bank and Centro Internacional de Agricultura Tropical. (2011) *Global Rural–Urban Mapping Project, version 1 (GRUMP v1): urban extents grid*. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY.

Dray, S., Legendre, P. & Blanchet, F.G. (2009) *Packfor: forward selection with permutation* (Canoco p.46). R package version 0.0-7lr58. The R Project for Statistical Computing, Vienna, Austria. Available at: http://r-forge.r-project.org/R/?group_id=195 (accessed 21 October 2014).

Duncan, R.P., Clemants, S.E., Corlett, R.T., Hahs, A.K., McCarthy, M.A., McDonnell, M.J., Schwartz, M.W., Thompson, K., Vesk, P.A. & Williams, N.S. (2011) Plant traits and extinction in urban areas: a meta-analysis of 11 cities. *Global Ecology and Biogeography*, **20**, 509–519.

Fields Development Team. (2006). *Fields: tools for spatial data*. National Center for Atmospheric Research, Boulder, CO. Available at: http://www.cgd.ucar.edu/Software/Fields (accessed 21 October 2014).

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hill, M.O., Preston, C.D. & Roy, D.D. (2004) *PLANTATT. Attributes of British and Irish plants: status, size, life history, geography and habitats*. CEH Publication Sales, Wallingford.

Kendal, D., Dobbs, C. & Lohr, V. (2014) Global patterns of diversity in the urban forest: is there evidence to support the 10/20/30 rule? *Urban Forestry and Urban Greening*, **13**, 411–417.

Kowarik, I., Von Der Lippe, M. & Cierjacks, A. (2013) Prevalence of alien versus native species of woody plants in Berlin differs between habitats and at different scales. *Preslia*, **85**, 113–132.

Kunick, W. (1990) Spontaneous woody vegetation in cities. *Urban ecology* (ed. by H. Sukopp, S. Hejny and I. Kowarik), pp. 167–174. SPB Academic Publishing, Hague.

Kühn, I. & Klotz, S. (2006) Urbanization and homogenization – comparing the floras of urban and rural areas in Germany. *Biological Conservation*, **127**, 292–300.

La Sorte, F.A. & McKinney, M.L. (2006) Compositional similarity and the distribution of geographical ranges size for assemblages of native and non-native species in urban floras. *Diversity and Distributions*, **12**, 679–688.

La Sorte, F.A. & McKinney, M.L. (2007) Compositional changes over space and time along an occurrence–abundance continuum: anthropogenic homogenization of the North American avifauna. *Journal of Biogeography*, **34**, 2159–2167.

La Sorte, F.A., McKinney, M.L. & Pyšek, P. (2007) Compositional similarity among urban floras within and across continents: biogeographical consequences of human-mediated biotic interchange. *Global Change Biology*, **13**, 913–921.

La Sorte, F.A., McKinney, M.L., Pyšek, P., Klotz, S., Rapson, G.L., Celesti-Grapow, L. & Thompson, K. (2008) Distance decay of similarity among European urban floras: the impact of anthropogenic activities on β diversity. *Global Ecology and Biogeography*, **17**, 363–371.

La Sorte, F.A., Aronson, M.F.J., Williams, N.S.G., Celesti-Grapow, L., Cilliers, S., Clarkson, B.D., Dolan, R.W., Hipp, A., Klotz, S., Kühn, I., Pyšek, P., Siebert, S. & Winter, M. (2014) Beta diversity of urban floras among European and non-European cities. *Global Ecology and Biogeography*, **23**, 769–779.

Legendre, P. & De Cáceres, M. (2013) Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters*, **16**, 951–963.

Lin, G. & Ho, S.P. (2003) China's land resources and land-use change: insights from the 1996 land survey. *Land Use Policy*, **20**, 87–107.

Loeb, R.E. (2011) What and where are old growth urban forests? *Old growth urban forests* (ed. by R.E. Loeb), pp. 1–14. Springer, New York.

Loeb, R.E. (2012) Arboricultural introductions and long-term changes for invasive woody plants in remnant urban forests. *Forests*, **3**, 745–763.

Lososová, Z., Chytrý, M., Tichý, L., Danihelka, J., Fajmon, K., Hájek, O., Kintrová, K., Láníková, D., Otýpková, Z. & Řehořek, V. (2012a) Biotic homogenization of Central European urban floras depends on residence time of alien species and habitat types. *Biological Conservation*, **145**, 179–184.

Lososová, Z., Chytrý, M., Tichý, L., Danihelka, J., Fajmon, K., Hájek, O., Kintrová, K., Kühn, I., Láníková, D., Otýpková, Z. & Řehořek, V. (2012b) Native and alien floras in urban habitats: a comparison across 32 cities of central Europe. *Global Ecology and Biogeography*, **21**, 545–555.

McKinney, M.L. (2005) Species introduced from nearby sources have a more homogenizing effect than species from distant sources: evidence from plants and fishes in the USA. *Diversity and Distributions*, **11**, 367–374.

McKinney, M.L. (2006) Urbanization as a major cause of biotic homogenization. *Biological Conservation*, **127**, 247–260.

McKinney, M. (2008) Do humans homogenize or differentiate biotas? It depends. *Journal of Biogeography*, **35**, 1960–1961.

McKinney, M.L. & La Sorte, F.A. (2007) Invasiveness and homogenization: synergism of wide dispersal and high local abundance. *Global Ecology and Biogeography*, **16**, 394–400.

McKinney, M.L. & Lockwood, J.L. (1999) Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends in Ecology and Evolution*, **14**, 450–453.

Mokany, K., Jones, M.M. & Harwood, T.D. (2013) Scaling pairwise β-diversity and α-diversity with area. *Journal of Biogeography*, **40**, 2299–2309.

Nock, C.A., Paquette, A., Follett, M., Nowak, D.J. & Messier, C. (2013) Effects of urbanization on tree species functional diversity in eastern North America. *Ecosystems*, **16**, 1487–1497.

Nowak, D.J., Crane, D.E., Stevens, J.C. & Hoehn, R.E. (2005) *The urban forest effects (UFORE) model: field data collection manual.* US Department of Agriculture, Forest Service, Northeastern Research Station, Newtown Square, PA.

Nowak, D.J., Walton, J.T., Stevens, J.C., Crane, D.E. & Hoehn, R.E. (2008) Effect of plot and sample size on timing and precision of urban forest assessments. *Arboricculture and Urban Forestry*, **34**, 386–390.

Olden, J.D. & Poff, N.L. (2003) Toward a mechanistic understanding and prediction of biotic homogenization. *The American Naturalist*, **162**, 442–460.

Olden, J.D. & Rooney, T.P. (2006) On defining and quantifying biotic homogenization. *Global Ecology and Biogeography*, **15**, 113–120.

Olden, J.D., Poff, N.L., Douglas, M.R., Douglas, M.E. & Fausch, K.D. (2004) Ecological and evolutionary consequences of biotic homogenization. *Trends in Ecology and Evolution*, **19**, 18–24.

Pincetl, S., Prabhu, S., Gillespie, T., Jenerette, G. & Pataki, D. (2013) The evolution of tree nursery offerings in Los Angeles County over the last 110 years. *Landscape and Urban Planning*, **118**, 10–17.

Pyšek, P., Křivánek, M. & Jarošík, V. (2009) Planting intensity, residence time, and species traits determine invasion success of alien woody species. *Ecology*, **90**, 2734–2744.

Qian, H. & Ricklefs, R.E. (2006) The role of exotic species in homogenizing the North American flora. *Ecology Letters*, **9**, 1293–1298.

Ramage, B.S., Roman, L.A. & Dukes, J.S. (2013) Relationships between urban tree communities and the biomes in which they reside. *Applied Vegetation Science*, **16**, 8–20.

Ricotta, C., La Sorte, F.A., Pyšek, P., Rapson, G.L., Celesti-Grapow, L. & Thompson, K. (2012) Phylogenetic beta diversity of native and alien species in European urban floras. *Global Ecology and Biogeography*, **21**, 751–759.

R Development Core Team (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/ (accessed 21 June 2014)

Sjöman, H., Östberg, J. & Bühler, O. (2012) Diversity and distribution of the urban tree population in ten major Nordic cities. *Urban Forestry and Urban Greening*, **11**, 31–39.

Trentanovi, G., von der Lippe, M., Sitzia, T., Ziechmann, U., Kowarik, I. & Cierjacks, A. (2013) Biotic homogenization at the community scale: disentangling the roles of urbanization and plant invasion. *Diversity and Distributions*, **19**, 738–748.

Williams, N.S., Schwartz, M.W., Vesk, P.A., McCarthy, M.A., Hahs, A.K., Clemants, S.E., Corlett, R.T., Duncan, R.P., Norton, B.A., Thompson, K. & McDonnell, M.J. (2009) A conceptual framework for predicting the effects of urban environments on floras. *Journal of Ecology*, **97**, 4–9.

Williams, N.S., Hahs, A.K. & Vesk, P.A. (2015) Urbanisation, plant traits and the composition of urban floras. *Perspectives in Plant Ecology, Evolution and Systematics*, **17**, 78–86.

Winter, M., Kühn, I., La Sorte, F.A., Schweiger, O., Nentwig, W. & Klotz, S. (2010) The role of non-native plants and vertebrates in defining patterns of compositional dissimilarity within and across continents. *Global Ecology and Biogeography*, **19**, 332–342.

Zhao, J., Ouyang, Z., Xu, W., Zheng, H. & Meng, X. (2010) Sampling adequacy estimation for plant species composition by accumulation curves. A case study of urban vegetation in Beijing, China. *Landscape and Urban Planning*, **95**, 113–121.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Appendix S1 Information on 38 studied cities.
Appendix S2 Result of the principal component analysis.

BIOSKETCH

Our research team is interested in studying the patterns of urban biota and the underlying ecological and social processes that drive the dynamic changes in these patterns at a global scale. Our goal is to gain a better understanding of the impact of urbanization on the biodiversity and ecosystem services that are the foundation for the sustainable development of our society (http://www.yanglab.org).

Editor: Karl Evans