**Supplementary methods**

*Data*

We compiled a global data set of species abundances per unit area or unit time spent observing for bird assemblages in 27 regions from Europe, North America, Australia, Africa and South America (Appendix S2). Information came from published studies and our own surveys in Newcastle, Australia (Sol *et al.* 2012) and Valdivia, Chile (Sol *et al.* 2014). The data included 185,908 bird detections belonging to 1,219 species from urbanised environments and the little urbanised surrounding environments, 47 of which were exotic in at least one region.

For each region, information was available for one or more spatially-defined bird assemblages belonging to urbanised environments and the little urbanised, surrounding environments (n = 161 assemblages). To define the degree of urbanisation, we differentiated (1) urban habitats ('highly urbanised' environments, hereafter), where buildings were densely packed and parks were small or absent, and (2) suburbs ('moderately urbanised' environments), which were residential areas with single-family houses, commonly with backyards (Marzluff et al. 2001). The surrounding habitats ('little urbanised' environments) primarily included woodland (27.6%), grassland (21.3%) and rural (34.0%) habitats but also habitat mosaics and gardens. Of the 161 bird assemblages, 39 were in highly urbanised environments, 36 in moderately urbanised environments and 86 in the little urbanised, surrounding environments. While the survey method varied across studies, the same method was used for each habitat within a particular region, making data comparable using appropriate models (see details below).

The comparison among habitats raises issues about detectability and sampling effort (Lahoz-Monfort *et al.* 2014). Although all habitats within each region were intensively surveyed, we further tackled this limitation in four ways. First, we tested whether certain methodological decisions could bias the results (see details below), including the survey method (line count, point count or cell records), whether the study controlled or not for detectability, and whether the sampling effort was equal or not across habitats. Second, rarefaction of communities was used to estimate phylogenetic diversity with similar sampling effort (see details below). Third, we estimated phylogenetic diversity loss taking into account the relative abundance of the species in the habitat, which

reduces the influence of rare species which are more likely to remain undetected (see below). Finally, we considered in the models geographic and human-related factors that could influence variation in PD across habitats and regions, including season in which the survey was carried out (breeding, non-breeding, all-year), altitude (in m), distance to Equator (in Km) and degree of urbanisation of the study region. To estimate the later factor, we obtained information for each city on the following variables: i) urbanisation period (year of foundation of the city minus last year of the study), ii) urban area (in $km^2$) and iii) human population density (habitants per $km^2$). As the latter two variables were highly correlated, we only used urban areas in the analyses.

*Phylogenetic hypotheses*

Our analyses were based on the complete, dated phylogeny of birds proposed by Jetz et al. (2012). This phylogeny includes over 3,300 species without genetic information, which are partially constrained within their clade based on a combination of consensus trees and taxonomic information. In our case, the influence of species without genetic information was considered negligible as genetic information was available for 91% of the 1,219 species.

We dealt with phylogenetic uncertainty in two ways. First, we used a maximum clade credibility phylogeny (CCP). The CCP was built with the TREEANNOTATOR software in BEAST (Drummond *et al.* 2012), based on 200 phylogenies from the pseudoposterior distribution of 10,000 trees available in www.birdtree.org. Second, we repeated the most important analyses with 50 phylogenies randomly chosen from the pseudoposterior distribution of 10,000 trees. The results were highly congruent and in the main text we present the results with the CCP.

*Loss of phylogenetic richness in urbanised environments*

We used the phylogenies to estimate Faith's phylogenetic diversity (PD, hereafter) for each bird assemblage. PD was measured as the sum of the lengths of all those branches that are members of the corresponding minimum spanning path (Faith 1992). Rarefaction was also used to account for differences in sampling effort across habitats. Expected number of species in each assemblage was drawn at random from the regional pool, using the minimum abundance among all assemblages within a region (Cardoso *et al.* 2015). To further assess the influence of rare species on phylogenetic diversity measures (Fine & Kembel 2011), we also used a recently proposed phylogenetic

diversity metric (PDw hereafter) based on Hill numbers that is sensitive to both species abundance and phylogenetic distances (Chao *et al.* 2010). We set the *q* parameter to 1, meaning that a species weight in the Hill number is proportional to its abundance.

To assess whether the loss of phylogenetic diversity merely results from species loss, we standardized PD and PDw by subtracting from the observed values the mean values expected under a null model in which taxa labels were shuffled across the tips of the regional phylogenetic tree 999 times. This null model generates randomized phylogenetic relationships among taxa from the regional pool while maintaining the observed phylogenetic tree, species abundance distribution, and community structure. The effect size was then divided by the standard deviation of the randomized values (hereafter phylogenetic diversity index, PDI and PDIw). The more negative PDI and PDIw, the more phylogenetically clustered species are locally compared to the regional species pool. We further explored the influence of species loss on PD decline by means of Piecewise structural equation modelling (Lefcheck 2015).

All the above metrics were estimated for each assemblage with the *R*-packages "Picante" (Kembel *et al.* 2010) and "BAT" (Cardoso *et al.* 2015) as well as our own scripts. Differences across urbanised and little urbanised environments were then evaluated by means of linear mixed models, using the Bayesian approximation implemented in the R-package 'MCMCglmm' (Hadfield 2010). As the response variables were continuous and exhibited symmetric "bell curve" shapes, we used models with a Gaussian structure of errors. Country and region within country were included as random effects. By allowing the intercept of the models to vary across regions and countries, phylogenetic diversity was compared within regions instead of across regions. As the defined random structure did not always account for spatial autocorrelation (see Appendix S3), in some models we included geographic longitude and/or distance to Equator as fixed effects. The survey method, sampling effort and whether detectability was or was not taken into account were also included as fixed effects, along with season, altitude, urbanisation period and urban area (see also Sol et al. 2014). To check for model convergence, we ran the model twice with different starting values, sampling 1000 iterations from a total of 1100000. When the model did not adequately converge, we sequentially increased the total number of iterations and number of iterations passed before samples were stored.

*Beta diversity and mitigation of PD loss by the addition of native urban exploiters*

To calculate dissimilarity in phylogenetic composition among urbanised and the surrounding little urbanised environments, we estimated UniFrac beta diversity indices among urbanised and little urbanised environments. The UniFrac beta diversity index, a derivation of the Jaccard dissimilarity index, quantifies the proportion of shared branch length between pairs of assemblages (Leprieur *et al.* 2012). We compared highly-urbanised vs moderately urbanised, highly-urbanised vs little urbanised and moderately urbanised vs little urbanised by means of the MCMCglmm approach outlined above. Then, we further explored how differences in phylogenetic composition arise by separating 'True' turnover of lineages (turnover, hereafter) from phylogenetic diversity gradients (nestedness), following Leprieur et al. (2012). A predominance of turnover over nestedness would imply that beta diversity does not merely reflect a reduction of diversity but a different combination of phylogenetic lineages.

As we detected substantial turnover, we estimated the fraction of the total phylogenetic diversity (PD) coming from exploiters (see definitions below), exotics and non-exploiters. To this purpose, we estimated the evolutionary distinctiveness (ED) of the three groups based on subtrees only containing the species of the assemblage. ED was calculated as the weighted sum of the edge lengths from the root to a tip of the tree, with the weights being 1/number of species that share that edge (Isaac *et al.* 2007). Following Sol et al. (2014), we operationally defined an exploiter as a native species that is more abundant in the urbanised environment than expected by their abundance in the little urbanised surrounding environments (see Blair 2001 for a more general definition). To assess this, we constructed 999 null communities representing the abundance of species expected by random. We considered a species as an urban exploiter if the observed abundance in the highly urbanised habitat was equal to or higher than the 95th percentile of its random abundances. As some species were abundant in highly urbanised environment despite exhibiting lower densities than in the little urbanised surrounding environments, we also used a second, less restrictive classification in which we also considered as exploiter those species with relative abundances in highly urbanised environment higher than 5%. We note that the use of a higher threshold does not alter the conclusions.

*Evolutionary distinctiveness and urbanisation*

To investigate whether evolutionary distinctive species are more sensitive to urbanisation, we again calculated the weighted sum of the edge lengths from the root to a tip of the tree (Isaac et al. 2007). In this case, however, ED was estimated based on the entire phylogeny of 9,993 avian species (a global measure referred to as EDg hereafter), and the mean EDg of each community was estimated as the average value for all species present. Differences across habitats in EDg were evaluated with the Gaussian MCMCglmm approach outlined in previous sections. As we found a consistent decline in EDg with the intensity of urbanisation (Morelli *et al.* 2016), we conducted additional analyses to further interpret it. On one hand, we investigated the unique contribution of each species to evolutionary distinctiveness, we also estimated three additional metrics: (1) evolutionary uniqueness (EU), measured as the length of the terminal branch of the species, (2) the average distance of each species with all other species of the bird assemblage (AD), and (3) the mean pairwise distance (MPD) of all species of the assemblage. The mean of each metric across species was then used to characterize the assemblages. As AD and MPD were non-significantly associated with urbanisation in all analyses ($P > 0.35$ in all cases), they are not reported in the results. On the other hand, we asked whether vulnerable species fall within species poor, phylogenetically isolated clades by comparing EDg between urban exploiters and avoiders. Whether or not a species was an exploiter was modelled as a function of EDg with a binomial MCMCglmm. Because being an exploiter does not only depend on the features of the species itself but also on how it interacts with other species from the region, we conducted the analyses at the assemblage level and included region and country as random factors in a nested structure. The species identity and the phylogeny were also included as random factors.

To further investigate changes in the phylogenetic composition of communities, we assessed whether species vulnerability to urbanisation was phylogenetically clustered. Using again the distinction between exploiter and avoider as a way to assess differences in the vulnerability of species to urbanisation, we investigated phylogenetic clustering in two ways. First, we used the *D* statistic (Fritz & Purvis 2010) as a measure of phylogenetic signal in binary traits, as implemented in the R-package Caper (Orme *et al.* 2013). For this analysis, we considered a species as exploiter if it was identified as exploiter in at least one of the study regions. We illustrated the results with a

phylogenetic reconstruction based on stochastic character mapping (Revell 2011). Second, we tested whether some avian families had more or less exploiters than expected by chance by means of randomization tests, as described in Frishkoff *et al.* (2014). For each family of Jetz et al. (2012) taxonomy, we estimated the number of exploiters expected by chance by randomly sampling species without replacement from all exploiter and non-exploiter native species included in the present study (n = 1172 species) and counting the number of exploiters in the sample. The total number of species sampled per family corresponded to the number of species present in our dataset for that particular family. We repeated this 10,000 times to create a null distribution of random expectations. We considered a family to contain more exploiters than expected by chance if the observed number of exploiters was equal to or higher than the 95th percentile of the random number and fewer exploiters than expected by chance if the observed value was equal to or lower than the 5th percentile. We used four ways to identify exploiters: i) if it was categorised as exploiter at least in one of the study regions based on metric 1; ii) 1) if it was categorised as exploiter at least in one of the study regions based on metric 2; iii) if it was identified as exploiter in more than 50% of the studied regions based on metric 1; and vi) if it was identified as exploiter in more than 50% of the studied regions based on metric 2. Only the 20 families with 15 or more species were used in the analyses.