Appendix S4

Exploring geographic patterns in evolutionary richness and divergence

Biodiversity is known to exhibit global spatial patterns. To assess whether and how these patterns affect the main findings of our study, we proceeded as follows. We used the package Ape (Paradis et al. 2004) to formally assess whether we really needed to deal with spatial autocorrelation by estimating Moran's I for PD, PDraref, PDw, ED and EU, considering all the habitats together.

METRIC	OBSERVED	EXPECTED	SD	P
PD	0.597	-0.006	0.057	<0.0001
PD_{RAREF}	0.540	-0.006	0.057	<0.0001
PD_{W}	0.578	-0.006	0.057	<0.0001
ED	0.763	-0.006	0.057	<0.0001
EU	0.696	-0.006	0.057	<0.0001

As P < 0.05 is all cases, we rejected the null hypothesis that there was zero spatial autocorrelation present in the variables.

We next used a mixed model approach to investigate the influence of spatial autocorrelation in the results, as implemented in the nlme package. We started with two simple random structures, region alone and region nested within country. We included habitat and all confounding factors as fixed factors (see main text for details).

PD

	MODEL	DF	AIC	BIC	LOGLIK	TEST	L.RATIO	P-VALUE
M.NULL	1	13	206.1898	246.2480	-90.09488			
M0	2	13	103.3661	143.4244	-38.68306			
M0.1	3	14	105.3661	148.5058	-38.68306	2 VS 3	5.070945E-08	0.9998

PD_{W}

	MODEL	DF	AIC	BIC	LOGLIK	TEST	L.RATIO	P-VALUE
M.NULL	1	13	83.50337	123.56163	-28.75168			
M0	2	13	-15.95504	24.10322	20.97752			
M0.1	3	14	-17.41807	25.72159	22.70904	2 VS 3	3.463031	0.0628

PD_{RAREF}

	MODEL	DF	AIC	BIC	LOGLIK	TEST	L.RATIO	P-VALUE
M.NULL	1	13	-29.39822	10.66004	27.69911			
M0	2	13	-108.03184	-67.97359	67.01592			
M0.1	3	14	-110.34480	-67.20514	69.17240	2 VS 3	4.312958	0.0378

ED

	MODEL	DF	AIC	BIC	LOGLIK	TEST	L.RATIO	P-VALUE
M.NULL	1	13	-73.24944	-33.19118	49.62472			
M0	2	13	-166.11347	-126.05521	96.05674			
M0.1	3	14	-171.53786	-128.39820	99.76893	2 VS 3	7.424389	0.0064

EU

	MODEL	DF	AIC	BIC	LOGLIK	TEST	L.RATIO	P-VALUE
M.NULL	1	13	22.49886	62.55711	1.75057			
M0	2	13	-37.65338	2.40488	31.82669			
M0.1	3	14	-40.80886	2.33080	34.40443	2 VS 3	5.155486	0.0232

The best model was in all cases the one including region as random factor. Next, we fitted a number of spatial structures into this model and again used AICc to assess the fit. The spatial structures analysed were the quadratic (m1), exponential (m2), spherical (m3) and linear (m4). We also included x and y directly in the model to account for the possibility that spatial autocorrelation depended on more than just distance (m5).

We then compared all models with m0.

PD

	MODEL	DF	AIC	BIC	LOGLIK	TEST	L.RATIO	P-VALUE
M0	1	13	103.3661	143.4244	-38.68306			
M1	2	14	105.3661	148.5058	-38.68306	1 VS 2	0.000	0.9999
M2	3	14	105.3661	148.5058	-38.68306			
M3	4	14	160.8002	203.9398	-66.40008			
M4	5	14	163.7525	206.8922	-67.87626			
M5	6	16	105.6235	154.9260	-36.81176	5 VS 6	62.129	<.0001

PD_{RAREF}

	MODEL	DF	AIC	BIC	LOGLIK	TEST	L.RATIO	P-VALUE
M0	1	13	-108.03184	-67.97359	67.01592			
M1	2	14	-106.03184	-62.89218	67.01592	1 VS 2	0.00000	0.9999
M2	3	14	-106.03184	-62.89218	67.01592			
M3	4	14	-106.03184	-62.89218	67.01592			
M4	5	14	-70.05527	-26.91561	49.02763			
M5	6	16	-111.59519	-62.29272	71.79759	5 VS 6	45.53992	<.0001

PD_W

	MODEL	DF	AIC	BIC	LOGLIK	TEST	L.RATIO	P-VALUE
M0	1	13	-15.95504	24.10322	20.977520			
M1	2	14	-13.95504	29.18462	20.977520	1 VS 2	0.00000	0.9999
M2	3	14	-13.95504	29.18462	20.977520			
M3	4	14	-13.95504	29.18462	20.977520			
M4	5	14	44.95515	88.09481	-8.477577			
M5	6	16	-18.35252	30.94995	25.176261	5 VS 6	67.30768	<.0001

ED

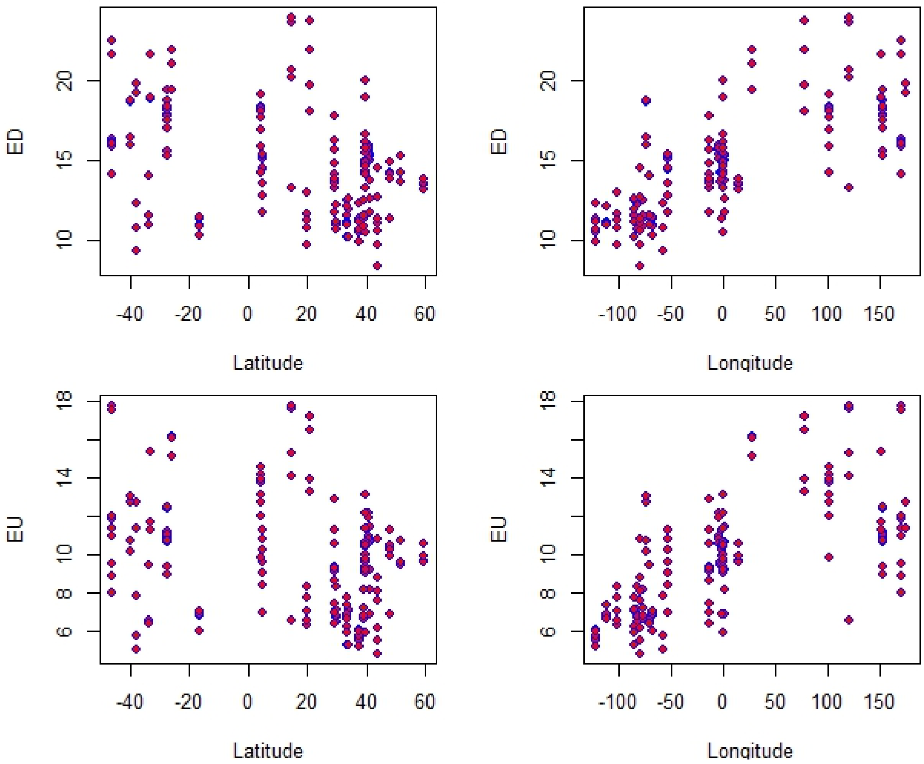
	MODEL	DF	AIC	BIC	LOGLIK	TEST	L.RATIO	P-VALUE
M0	1	14	-165.4334	-122.29373	96.71669			
M1	2	15	-164.3783	-118.15723	97.18915	1 VS 2	0.94491	0.331
M2	3	15	-164.0449	-117.82383	97.02245			
M3	4	15	-164.1868	-117.96571	97.09339			
M4	5	15	-129.8477	-83.62659	79.92383			
M5	6	17	-195.3640	-142.98009	114.68198	5 VS 6	69.51630	<.0001

EU

MODEL	DF	AIC	BIC	LOGLIK	TEST	L.RATIO	P-VALUE

M0	1	14	-39.41041	3.72925	33.70521			
M1	2	15	-39.62925	6.59182	34.81462	1 VS 2	2.21883	0.1363
M2	3	15	-40.51142	5.70965	35.25571			
M3	4	15	-18.77756	27.44351	24.38878			
M4	5	15	-13.46969	32.75137	21.73485			
M5	6	17	-64.38136	-11.99748	49.19068	5 VS 6	54.91166	<.0001

While including spatial autocorrelation did not improve the fit of the models for phylogenetic richness, autocorrelation seemed to be an issue for the metrics associated with phylogenetic divergence. Specifically, longitude seemed to affect the degree of ED and EU. We therefore included longitude in the formal analyses. Latitude also had an effect, but it seemed unimodal. In this case we included the distance to Equador as fixed factor in the models.



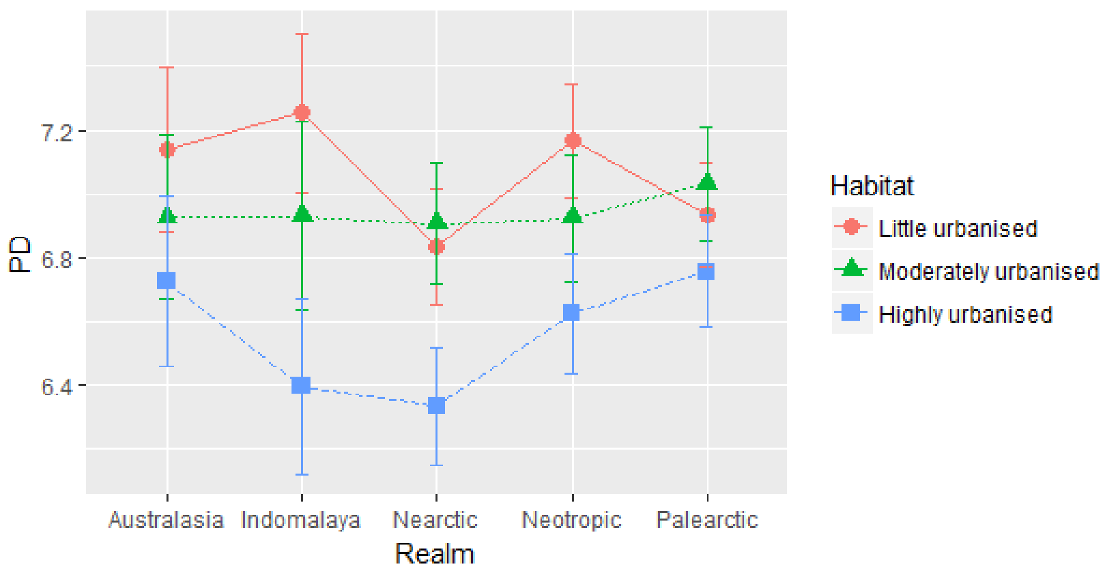
Estimating the Moran's I for the residuals indicated that including both spatial covariates to model ED and EU eliminates the effect of spatial autocorrelation (P = 0.38 and 0.64, respectively).

The existence of geographical variation in phylogenetic richness was further investigated by means of mixed models. We started by examining the interaction between habitat and biogeographic realms to assess whether differences in PD between habitats varied geographically. Interactions were explored with the package phia (Helios de Rosario-Martinez 2015). Region was included as random factor so as the compare PD within each urbanisation gradient.

```
Analysis of Deviance Table (Type II Wald chisquare tests)

 Response: log(PD)
                 Chisq Df Pr(>Chisq)
 habitat           71.2048  1  < 2.2e-16 ***
 bioregion          2.2011  4   0.698835
 habitat:bioregion 13.4176  4   0.009406 **
 ---
 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
 P-value adjustment method: holm
```

```
Pr(>Chisq)
 Little VS Highly urbanised : Australasia-Indomalaya   0.419238
 Little VS Highly urbanised :   Australasia-Nearctic   1.000000
 Little VS Highly urbanised :  Australasia-Neotropic   1.000000
 Little VS Highly urbanised : Australasia-Palearctic   0.761696
 Little VS Highly urbanised :     Indomalaya-Nearctic   0.419238
 Little VS Highly urbanised :    Indomalaya-Neotropic   0.419238
 Little VS Highly urbanised :   Indomalaya-Palearctic   0.004833
**
 Little VS Highly urbanised :      Nearctic-Neotropic   1.000000
 Little VS Highly urbanised :     Nearctic-Palearctic   0.252018
 Little VS Highly urbanised :    Neotropic-Palearctic   0.272896
 ---
 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



For PD, the difference in PD across highly and little urbanised environments changes across realms. In particular, the decline appeared to be lower in the Palearctic. Note that the Afrotropics were ignored in the figure because data was available for only one region.

To further investigate geographic effects, we estimated the decline in PD as PD_{loss} measured as log(PDLittle urbanised environs)-log(PDHighly urbanised environs) for all the regions. We log-transformed the variables to avoid confounding differences with mean values, and focused on the comparison between highly-little urbanised environments. PD_{loss} was then used as response variables in a mixed model to investigate factors explaining regional differences in decline. Here we considered as fixed factors variables like species richness in the little urbanised environment, urbanisation period, urban area, geographical coordinate of longitude and distance to Equator. The variable biogeographic realms was included as random factor. As shown

	Estimate	*Std. Error*	*Adjusted SE*	*z value*	*Pr(>|z|)*
(Intercept)	1.084	0.617	0.640	1.695	0.090
log(years of urbanization)	-0.104	0.093	0.096	1.088	0.277
detectability: used	0.056	0.140	0.145	0.385	0.700
log(urbanised surface)	-0.004	0.022	0.024	0.162	0.871
longitude	0.000	0.001	0.001	0.287	0.774
balanced surveys	0.013	0.072	0.076	0.166	0.868
altitude	0.000	0.000	0.000	0.127	0.899
survey method	0.022	0.126	0.133	0.165	0.869
log(species richness)	0.005	0.053	0.057	0.083	0.934
Distance to Equator	0.000	0.002	0.003	0.024	0.981
season: winter	-0.002	0.031	0.033	0.054	0.957
season: all year	0.001	0.055	0.059	0.017	0.987

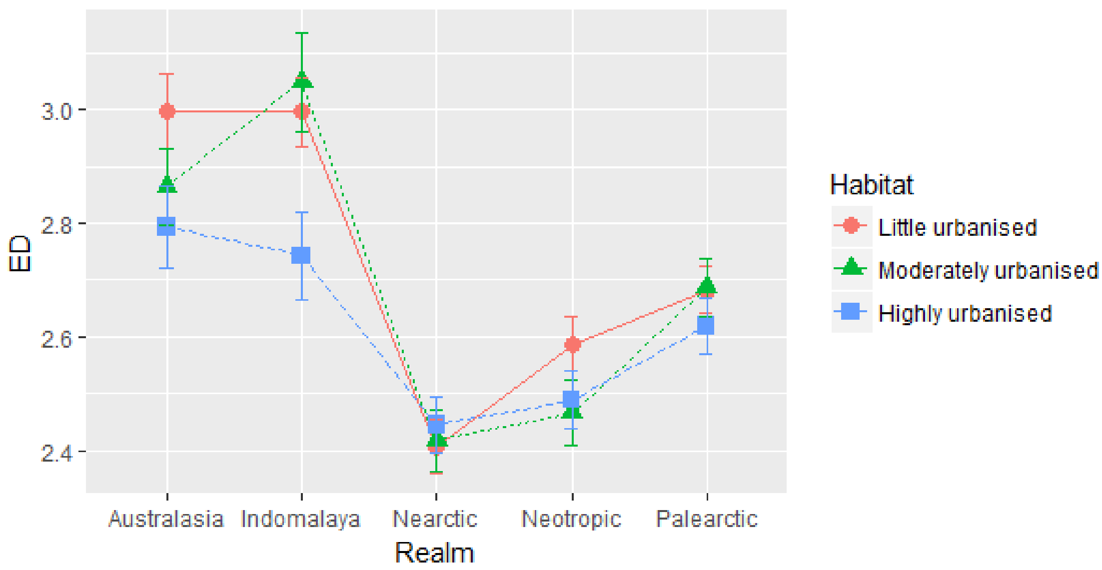
in the next table, none of the studied factors explains variation in PD_{loss}.

For ED, the visual inspection of the differences across highly and little urbanised environments also suggests some variation among biomes. Indeed, the loss in ED with urbanisation was lower in the Nearctic region than the Australasian and Indomalayan regions.

```
Analysis of Deviance Table (Type II Wald chisquare tests)

Response: log(ED)
               Chisq Df Pr(>Chisq)
habitat           14.105  1  0.0001729 ***
bioregion      70.218  4  2.042e-14 ***
habitat:bioregion 19.493  4  0.0006285 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
P-value adjustment method: holm
                                                      Pr(>Chisq)
Little VS Highly urbanised : Australasia-Indomalaya   1.000000
Little VS Highly urbanised :   Australasia-Nearctic   0.009027 **
Little VS Highly urbanised :  Australasia-Neotropic   0.415250
Little VS Highly urbanised : Australasia-Palearctic   0.306004
```

```
Little VS Highly urbanised :    Indomalaya-Nearctic   0.001885 **
Little VS Highly urbanised :   Indomalaya-Neotropic   0.306004
Little VS Highly urbanised :  Indomalaya-Palearctic   0.126409
Little VS Highly urbanised :      Nearctic-Neotropic   0.212501
Little VS Highly urbanised :     Nearctic-Palearctic   0.325384
Little VS Highly urbanised :   Neotropic-Palearctic   1.000000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



From the previous figure, it seemed that the most dramatic loss in ED occurred in regions with higher mean ED. This was further confirmed with mixed models. No other factor was significant in the model.

	Estimate	*Std. Error*	*Adjusted SE*	*z value*	*Pr(>|z|)*
(Intercept)	-0.339	0.160	0.168	2.019	0.043
Mean ED	0.030	0.010	0.010	3.033	0.002
season: winter	-0.161	0.088	0.091	1.767	0.077
season: all year	0.091	0.116	0.124	0.733	0.464
detectability: yes	0.013	0.038	0.039	0.326	0.745
survey method	0.019	0.060	0.062	0.300	0.764
log(years of urbanization)	-0.001	0.008	0.009	0.136	0.892
log(urbanised surface)	-0.001	0.006	0.006	0.119	0.905
altitude	0.000	0.000	0.000	0.124	0.901
distance to Equator	0.000	0.001	0.001	0.112	0.911
longitude	0.000	0.000	0.000	0.202	0.840
balanced surveys	-0.001	0.016	0.017	0.045	0.964
log(species richness)	0.000	0.014	0.015	0.005	0.996

None of the other factors was associated with ED_{loss}. Variation in ED_{loss} between moderately-little urbanised environments were not explained by any of the studied factors, $p > 0.2$ in all cases.

References

Helios De Rosario-Martinez (2015). phia: Post-Hoc Interaction Analysis. R package version 0.2-1. https://CRAN.R-project.org/package=phia

Paradis E., Claude J. & Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20: 289-290.

Pinheiro J, Bates D, DebRoy S, Sarkar D and R Core Team (2016). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-128, http://CRAN.R-project.org/package=nlme.