SI APPENDIX

Table S1. Explanatory variables used in this study. Descriptors in bold were retained for statistical analyses after collinearity analyses. Data were downloaded from the Climatic Research Unit (CRU) at the British Atmospheric Data Centre (http://badc.nerc.ac.uk/), NASA Earth Observations (http://neo.sci.gsfc.nasa.gov/), NASA Socioeconomic Data and Applications Center (http://sedac.ciesin.columbia.edu/), World Soil Information (http://www.isric.org/), National Agricultural Statistics Service of the United States Department of Agriculture (http://www.nass.usda.gov/), the IUCN Red List (http://www.iucnredlist.org/), and Natural Earth (http://www.naturalearthdata.com/).

Table S2. Bacterial indicator taxa identified. For each indicator taxa, the number of samples where it was present is indicated.

Figure S1. Map of sampling locations **(A)** and information on the number of samples per state **(B)**. Points indicate sampling locations. Color intensity indicates overlapping points.

Figure S2. Comparison between normalization and rarefaction pre-processing for bacterial communities for **(A)** richness ($r = 0.84$, $P < 0.001$) and **(B)** community similarity as NMDS axis ($r = 0.95$, $P < 0.001$).

Figure S3. Sample-based species accumulation curves (mean and 95% confidence interval).

Figure S4. Proportions of sequences representing bacterial taxa across samples at the phylum level (top) and within the Proteobacteria phylum (bottom). Samples are ordered along the bacterial NMDS ordination axis.

Figure S5. Proportions of sequences representing fungal taxa across samples at the phylum level (top) and within the Dothideomycetes class (bottom). Samples are ordered along the fungal NMDS ordination axis.

Figure S6. Mean relative abundances (where present) and occupancy relationships for **(A)** bacteria and **(B)** fungi.

Figure S7. Mantel correlogram. Solid points indicate significant ($P < 0.05$ using progressive correction of multiple tests) correlations, while empty points indicate non-significant ($P > 0.05$) correlations.

Figure S8. Geographical distribution maps for some bacterial sea indicator taxa: **(A)** *Prochlorococus*, **(B)** *Synechococcus*, **(C)** Pelagibacteraceae and **(D)** *Pseudoalteromonas*. Points represent sampling locations.

Figure S9. Effects of urbanization on richness ($P > 0.05$, Mann-Whitney test) and community similarity ($R^2 < 0.005$, PERMANOVA) for **(A)** bacterial and **(B)** fungal communities.

Class	Descriptors
Climate	**Mean annual temperature (MAT)**, temperature variability, **mean annual precipitation (MAP)**, precipitation variability, cloud fraction, days of frost during a year, net radiation.
Soil	**pH**, moisture, **dust deposition**, clay content.
Plant productivity	**Net primary productivity (NPP)**, normalized difference vegetation index (NDVI).
Organisms	**Human population density**, **cattle density**, chickens density, **hogs density**, **vascular plant diversity**, mammal diversity, bird diversity.
Geographic isolation	**Elevation**, **distance to the coast**.
House	**Construction age**, house space, number of bedrooms, **number of people**, **carpeting**, **number of house plants**, **number of dogs**, **number of cats**, **self-reported presence of mold**.

Table S1

Source environment	Bacterial indicator taxa
Plants	Chloroplasts (n=1186)
Insects	*Wolbachia* (n=459), *Buchnera* (n=429), *Rickettsiella* (n=737), *Cardinium* (n=152), *Rickettsia* (n=489), *Rhabdochlamydia* (n=153), *Fritschea* (n=30), *Regiella* (n=21), *Hamiltonella* (n=48), *Blochmannia* (n=18), *Portiera* (n=109), *Tremblaya* (n=19), Endomicrobia (n=20), Entomoplasmatales (n=217), Bartonellaceae (n=386), *Blattabacterium* (n=48), *Sulcia* (n=59), *Thorsellia* (n=1), *Baumannia* (n=2)
Sea	Pelagibacteraceae (n=11), *Prochlorococcus* (n=8), *Synechococcus* (n=78), *Polaribacter* (n=29), *Pseudoalteromonas* (n=166), *Marinobacter* (n=141), *Nitrosopumilus* (n=9)
Skin	*Propionibacterium* (n=176), *Staphylococcus* (n=1121), *Corynebacterium* (n=1144), *Streptococcus* (n=1057), *Rothia* (n=356), *Micrococcus* (n=706), *Anaerococcus* (n=507), *Brevibacterium* (n=1026)
Soil	Solibacteraceae (n=990), Chloracidobacteria (n=1153), Koribacteraceae (n=790), Acidobacteria-iii1.15 (n=1156), Acidobacteriaceae (n=1132), *Rhizobium* (n=873), *Bradyrhizobium* (n=1061), *Mesorhizobium* (n=1072), *Rhodoplanes* (n=1092), *Chitinophaga* (n=581), *Solirubrobacter* (n=728), *Opitutus* (n=689)
Stool	*Bacteroides* (n=889), *Faecalibacterium* (n=763), *Lachnospira* (n=176), *Oscillospira* (n=811), *Roseburia* (n=610), *Coprococcus* (n=845), *Ruminococcus* (n=762), *Parabacteroides* (n=459), *Phascolarctobacterium* (n=535), *Sutterella* (n=88), *Blautia* (n=836)

Table S2

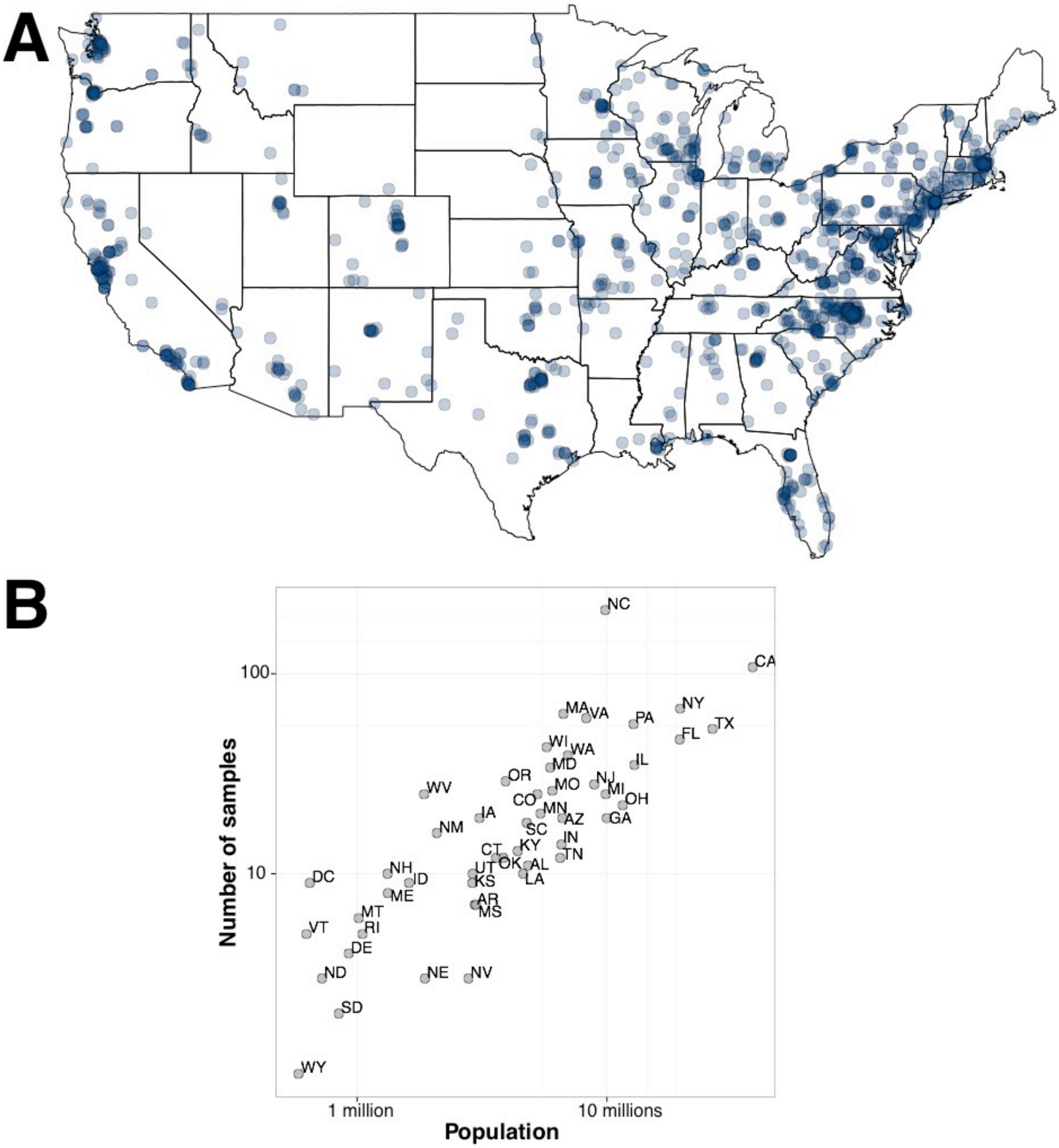
A
B
NC
CA
100
MA
VA
NY
PA
TX
FL
WI
WA
MD
IL
OR
NJ
MO
MI
WV
CO
OH
MN
AZ
IA
GA
NM
SC
KY
IN
CT
OK
AL
TN
UT
NH
ID
KS
LA
DC
10
ME
AR
MS
MT
VT
RI
DE
ND
NE
NV
SD
WY
Number of samples
1 million
10 millions
Population

Fig. S1

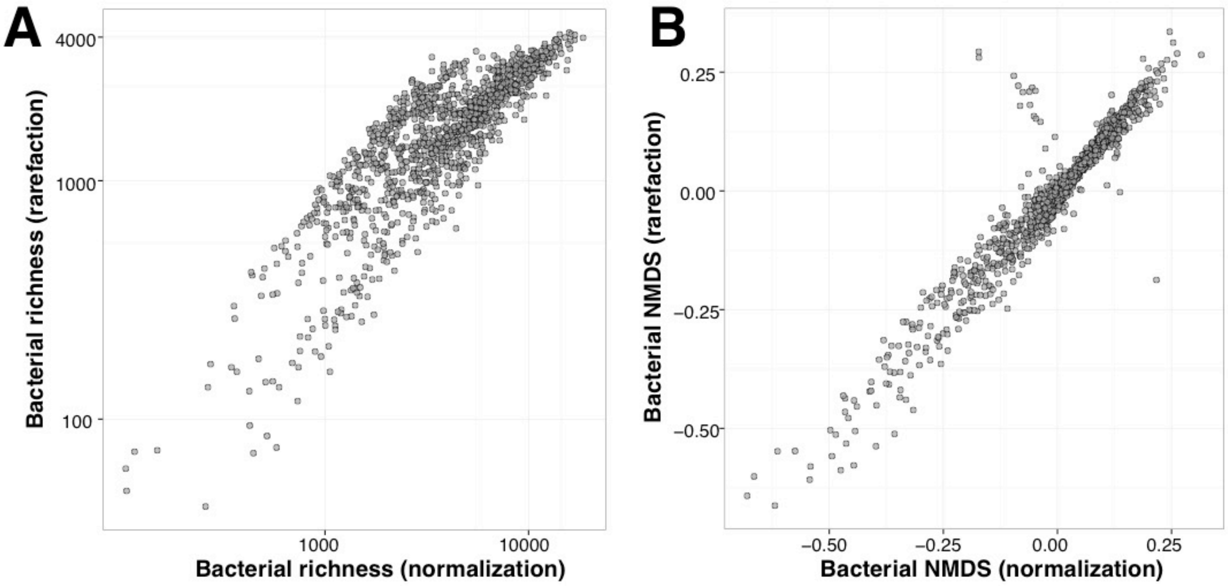
A
B
Bacterial richness (rarefaction)
Bacterial richness (normalization)
4000
1000
100
1000
10000
Bacterial NMDS (rarefaction)
Bacterial NMDS (normalization)
0.25
0.00
−0.25
−0.50
−0.50
−0.25
0.00
0.25

Fig. S2

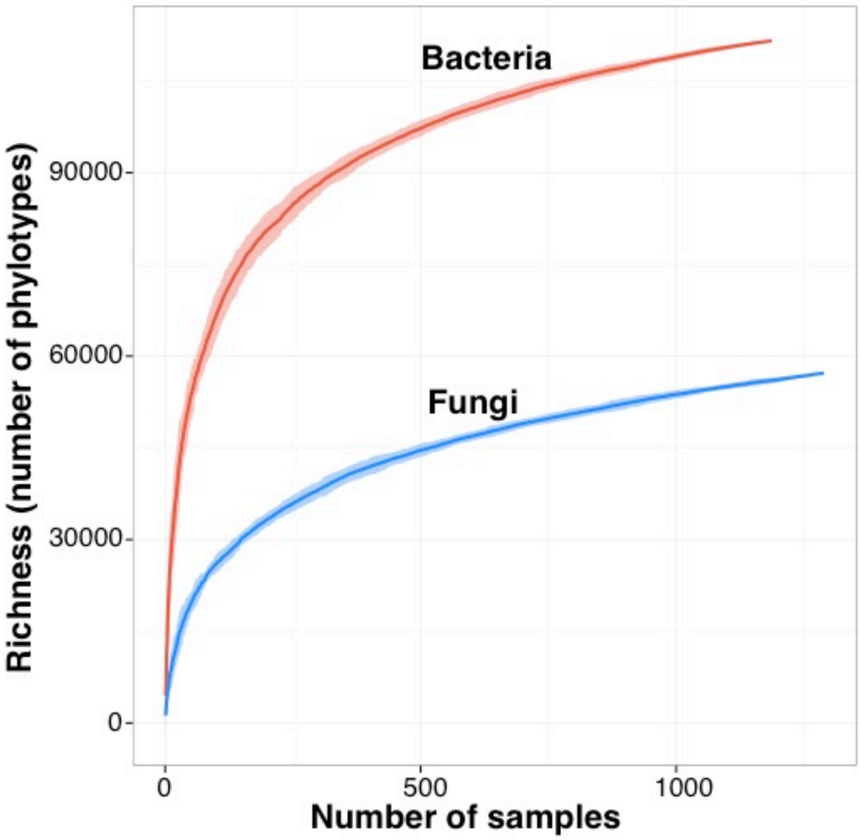
Bacteria
Fungi
Richness (number of phylotypes)
0
30000
60000
90000
0
500
1000
Number of samples

Fig. S3

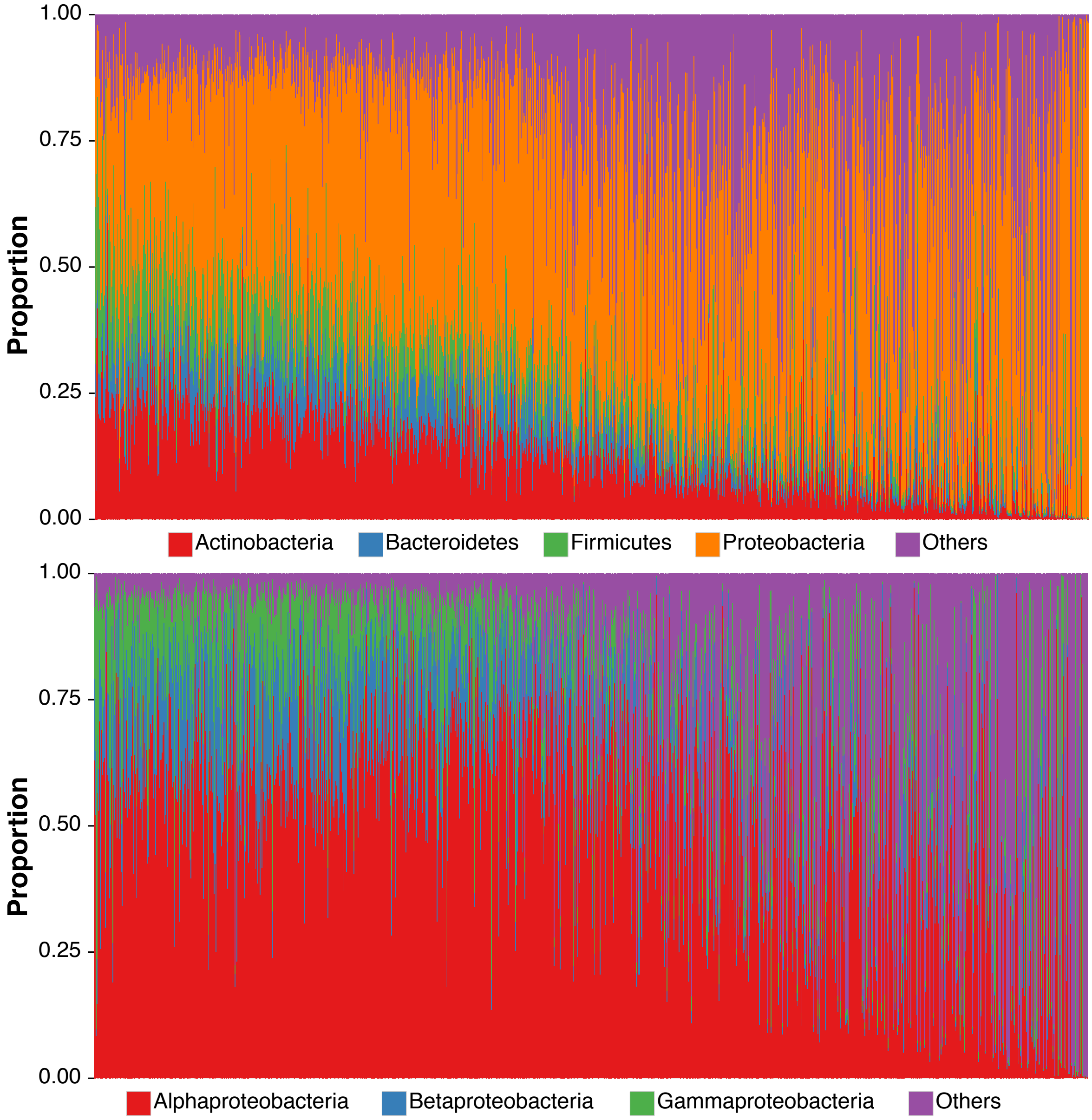
1.00
0.75
0.50
0.25
0.00
Proportion
Actinobacteria
Bacteroidetes
Firmicutes
Proteobacteria
Others
1.00
0.75
0.50
0.25
0.00
Proportion
Alphaproteobacteria
Betaproteobacteria
Gammaproteobacteria
Others

Fig. S4

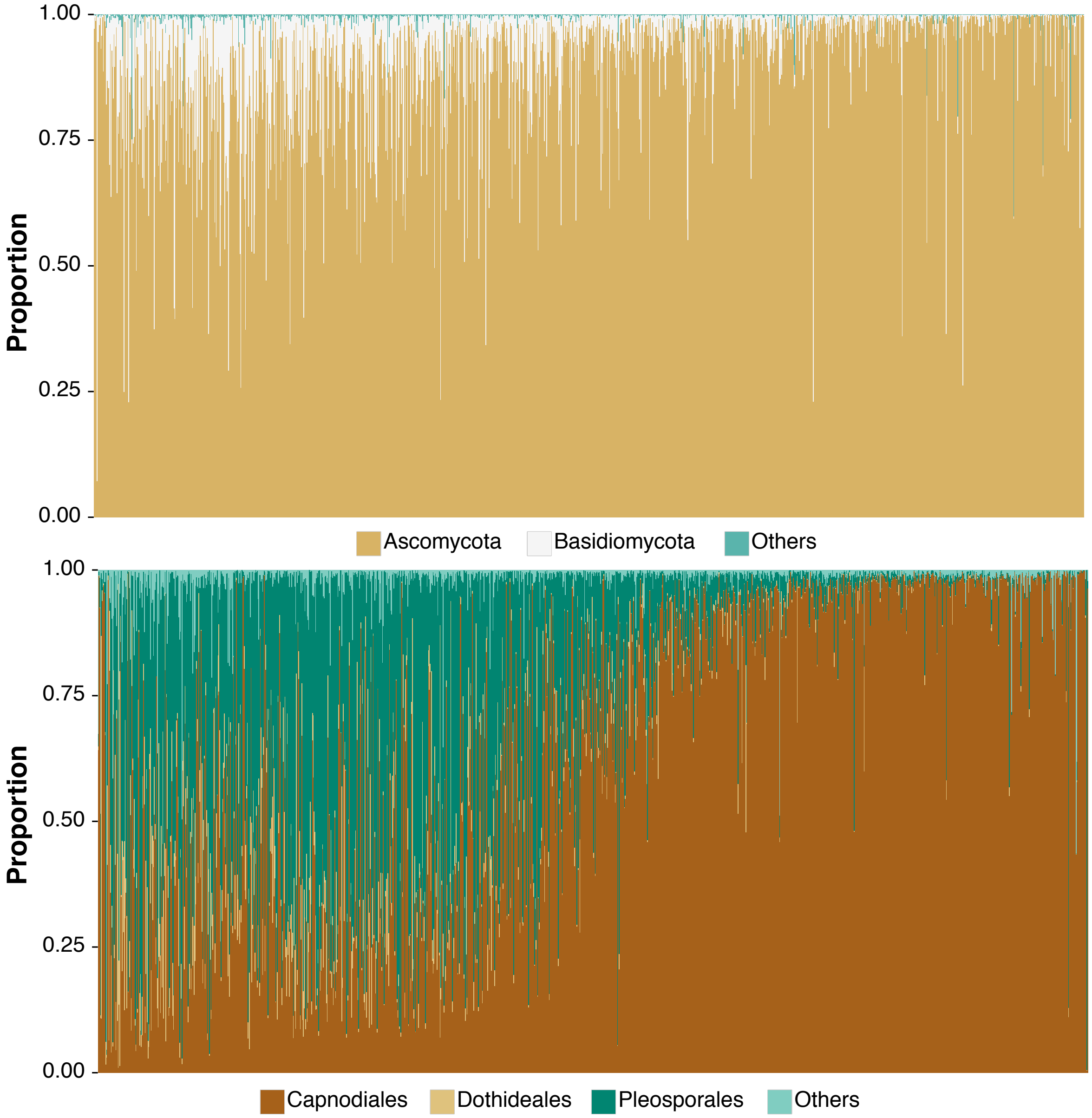
Proportion
1.00
0.75
0.50
0.25
0.00
Ascomycota
Basidiomycota
Others
Proportion
1.00
0.75
0.50
0.25
0.00
Capnodiales
Dothideales
Pleosporales
Others

Fig. S5

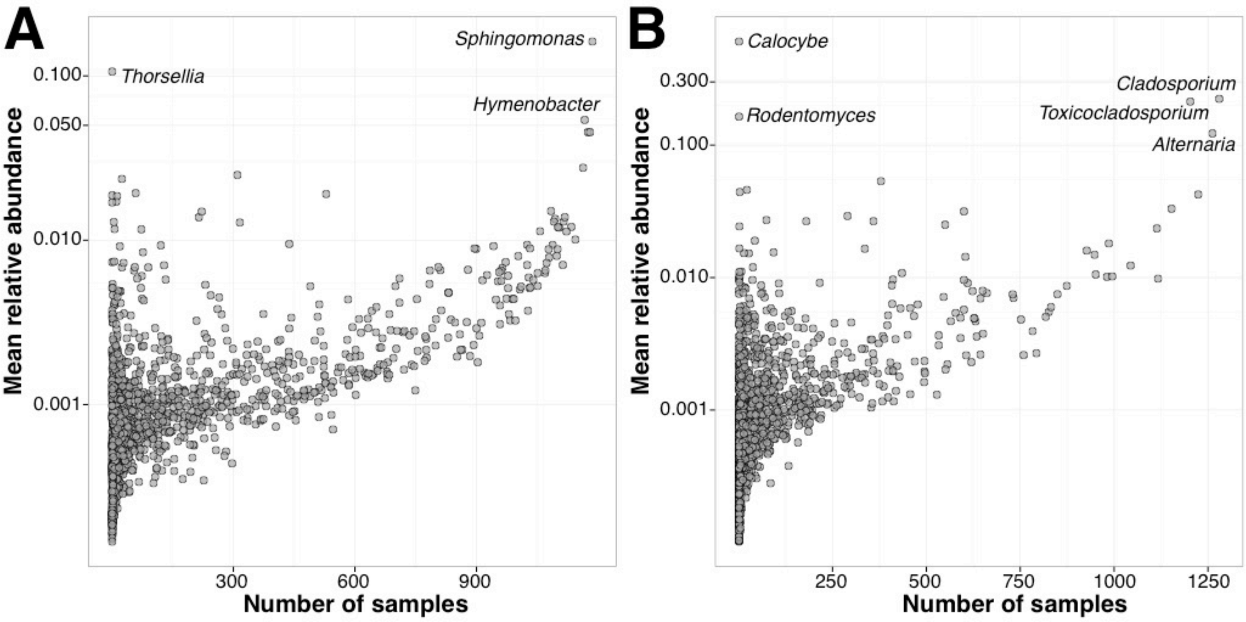
A
Sphingomonas
Thorsellia
Hymenobacter
0.100
0.050
0.010
0.001
Mean relative abundance
300
600
900
Number of samples
B
Calocybe
Rodentomyces
Cladosporium
Toxicocladosporium
Alternaria
0.300
0.100
0.010
0.001
Mean relative abundance
250
500
750
1000
1250
Number of samples

Fig. S6

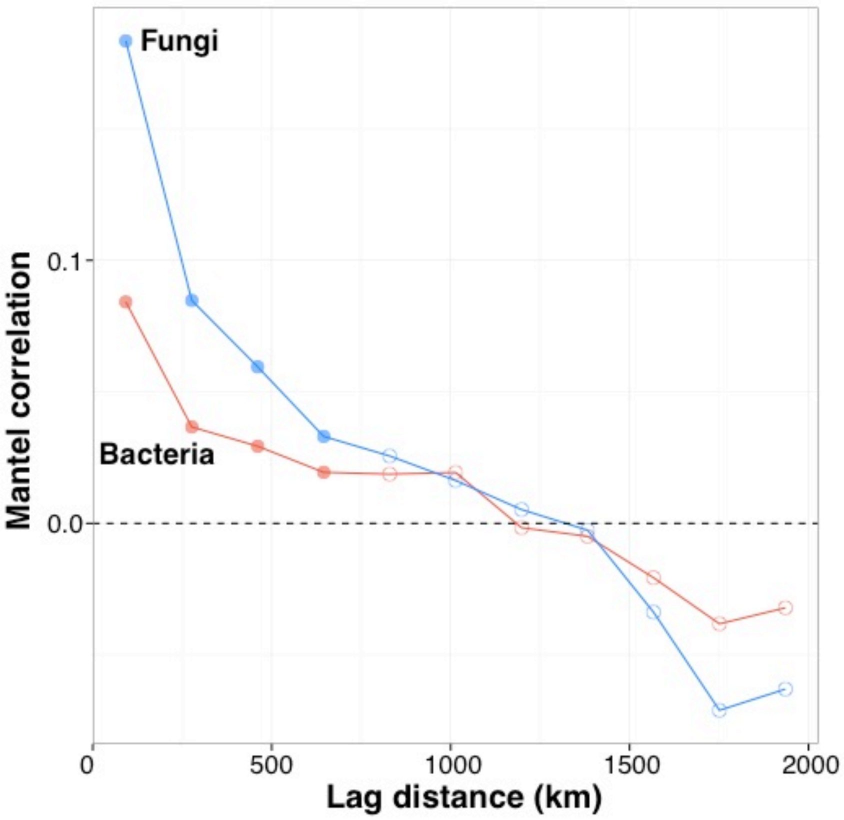
Fungi
Bacteria
Mantel correlation
0.1
0.0
0
500
1000
1500
2000
Lag distance (km)

Fig. S7


A
Proportion of
Prochlorococcus
0.00020
0.00015
0.00010
0.00005
B
Proportion of
Synechococcus
6e−04
4e−04
2e−04
C
Proportion of
Pelagibacteraceae
7.5e−05
5.0e−05
2.5e−05
D
Proportion of
Pseudoalteromonas
0.004
0.003
0.002
0.001

Fig. S8

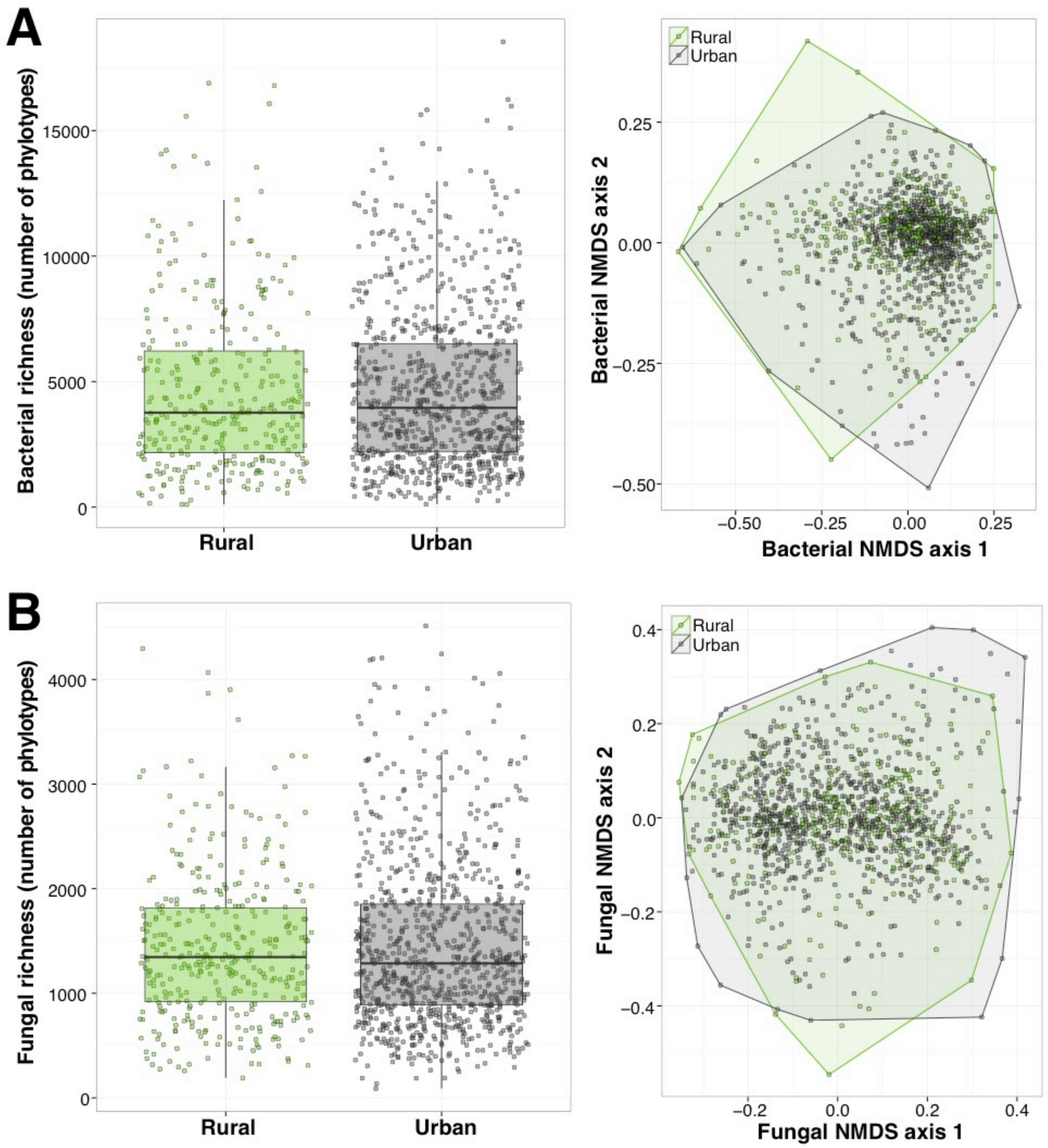
A
Bacterial richness (number of phylotypes)
0
5000
10000
15000
Rural
Urban
Rural
Urban
Bacterial NMDS axis 2
0.25
0.00
−0.25
−0.50
−0.50
−0.25
0.00
0.25
Bacterial NMDS axis 1
B
Fungal richness (number of phylotypes)
0
1000
2000
3000
4000
Rural
Urban
Rural
Urban
Fungal NMDS axis 2
0.4
0.2
0.0
−0.2
−0.4
−0.2
0.0
0.2
0.4
Fungal NMDS axis 1

Fig. S9