Ecography

ECOG-02946
Giam, X. and Olden, J. D. 2017. Drivers and interrelationships among multiple dimensions of rarity for freshwater fishes. – Ecography doi: 10.1111/ecog.02946

Supplementary material

Supplementary material

Drivers and interrelationships among multiple dimensions of rarity for freshwater fishes

Xingli Giam & Julian D. Olden

Appendix 1. Suitability of different local abundance metrics

Appendix 2. Sampling intensity among regions

Appendix 3. Phylogenetic tree and metrics of rarity for 292 species

Appendix 4. Model rankings for all response variables

Appendix 5. Final phylogenetic logistic regression (PLR) models for use of lacustrine habitats and use of large rivers

Appendix 6. Shipley's d-separation test for validity of final path models

Appendix 1. Suitability of different local abundance metrics

Raw (or log-transformed) relative abundance

Neither raw nor log-transformed relative abundance is appropriate for our analyses, which aims to estimate a composite measure of local abundance for every species in our dataset. The relative abundance of the *i*-th species in a given site (RA_i),

$$RA_i = A_i / \sum A_i \qquad (1),$$

is sensitive to the species richness and abundance at the site (Fig. A1a, b). This is because the minimum value (= 1 / $\sum A_i$) decreases with total abundance $\sum A_i$, which becomes clear when $log(RA_i)$ is plotted against $log(\sum A_i)$ (Fig. A1c). As species richness increases, the mean RA_i at a given site decreases (Fig. A1c,d) because the total relative abundance remains constant at 1. As a result of these relationships, raw or log-transformed relative abundance is conflated with across-site variation in site richness and abundance, and consequently complicate our analyses of rarity drivers and interrelationships.

Standardized log relative abundance

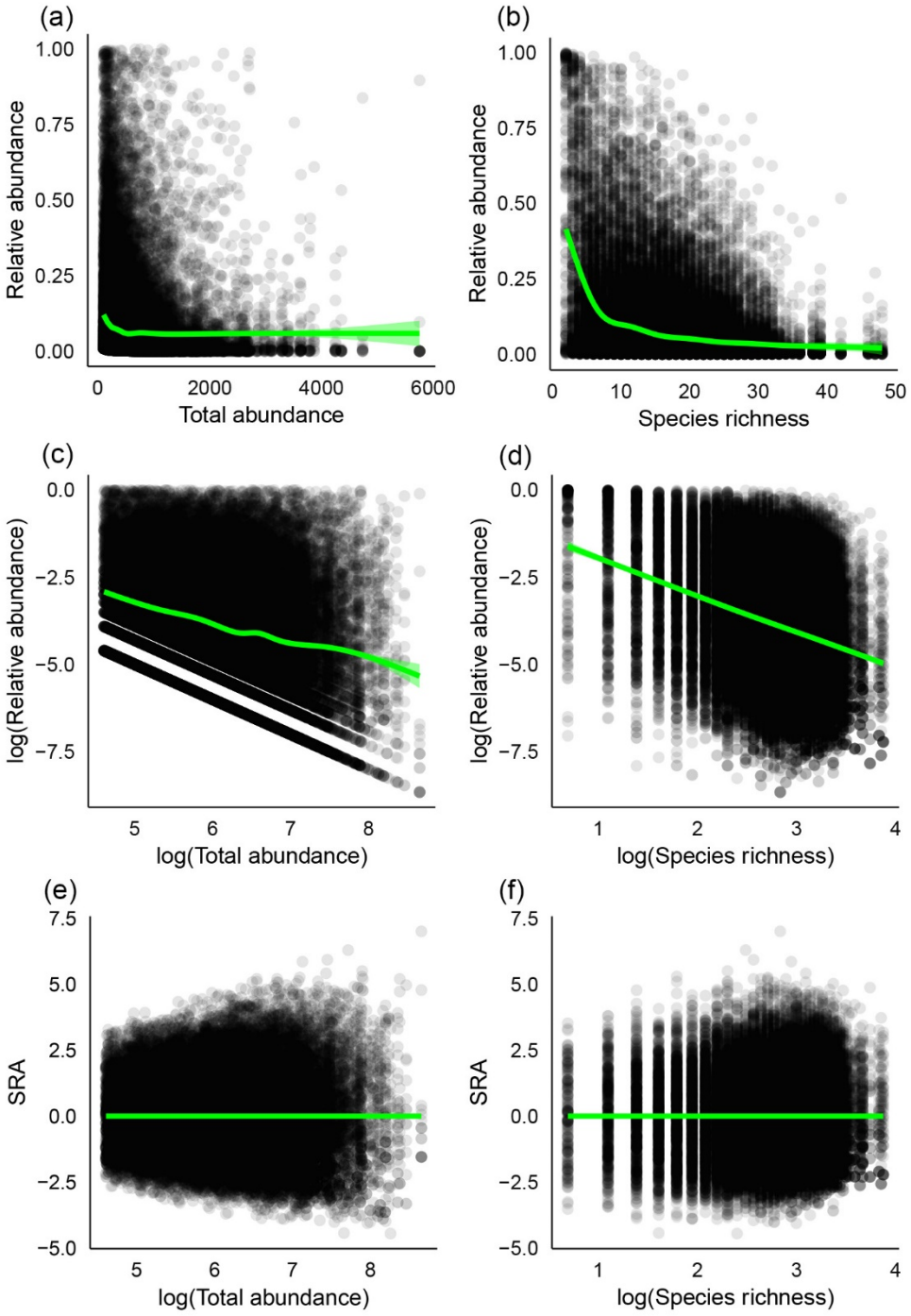
We needed a measure of local abundance that is independent from site richness and abundance. Such a measure would ensure comparability of local abundance values among sites, and allow us to aggregate the abundance value of each species across sites in which it occurs into a single composite measure for our analyses of rarity drivers and interrelationships. To do this, we centered the $log(RA_i)$ on its mean value at each given site. The resulting quantity, standardized log relative abundance (SRA_i) for the *i*-th of *S* species is expressed as,

$$SRA_i = log(RA_i) - {\sum_{i=1}^{S} log(RA_i)}/{S} \qquad (2).$$

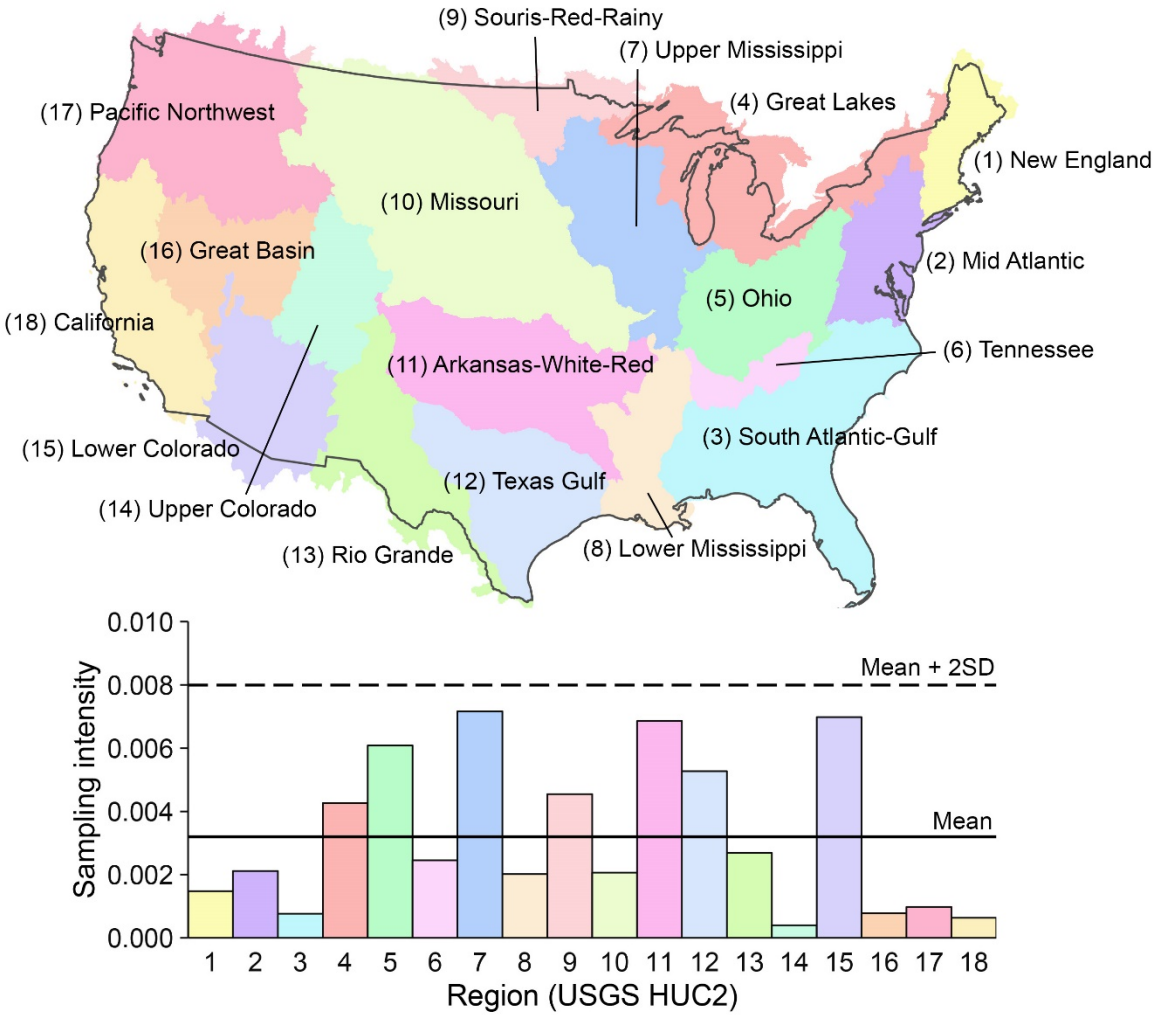
34 The centering procedure fixes the mean SRA_i is 0 at every site, thus removing the artefactual
35 effects of site richness and abundance on relative abundance (Fig. A1e, f). We therefore use SRA
36 as an aggregable measure of local abundance in our analyses.

37

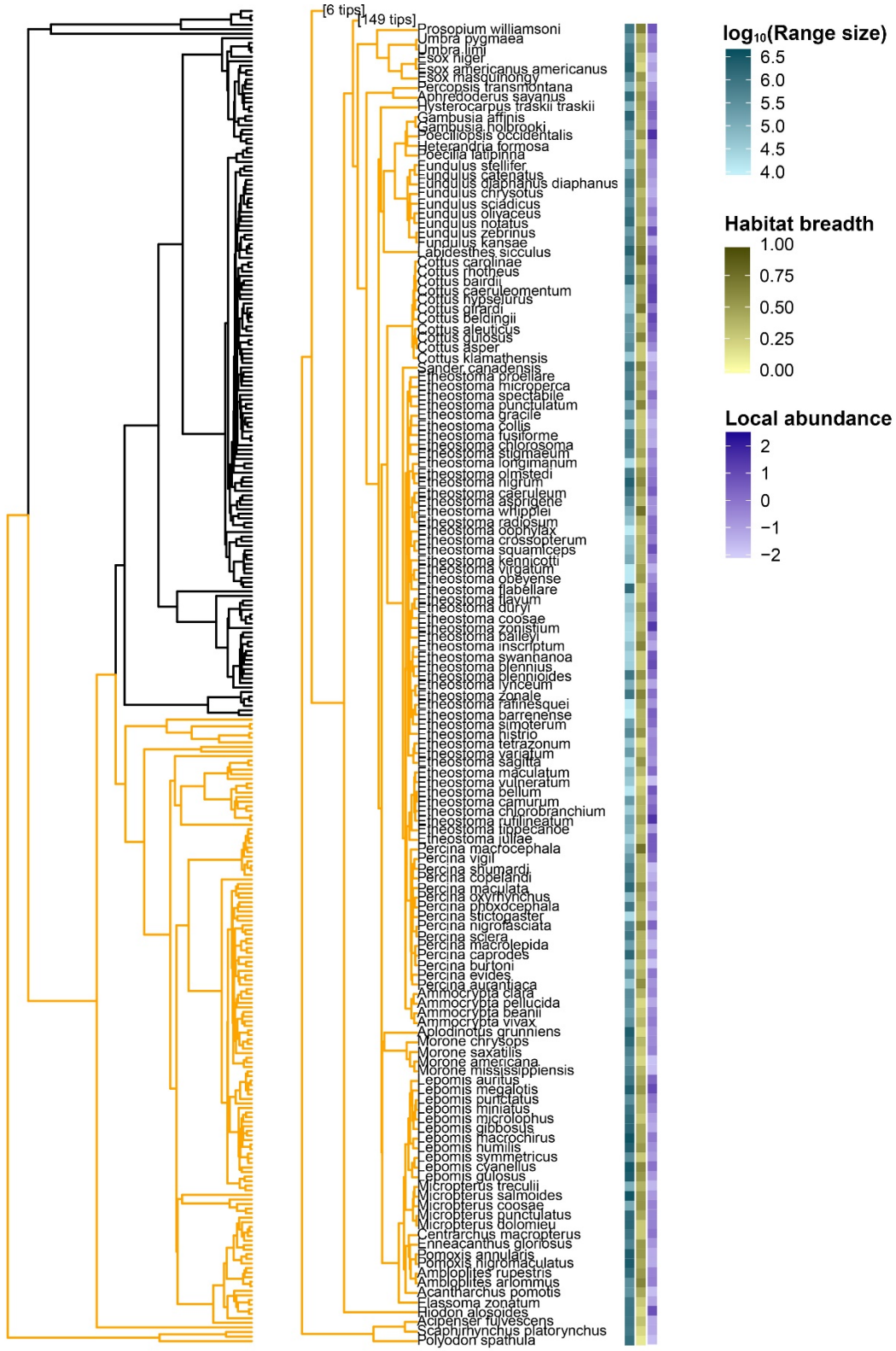
Figure A1. Relationships between relative abundance (a, b), log-transformed relative abundance (c, d), and standardized log relative abundance (e, f) and site total abundance and richness. Green lines are smoothers based on generalized additive models, which characterize the general relationship between each pair of variables.

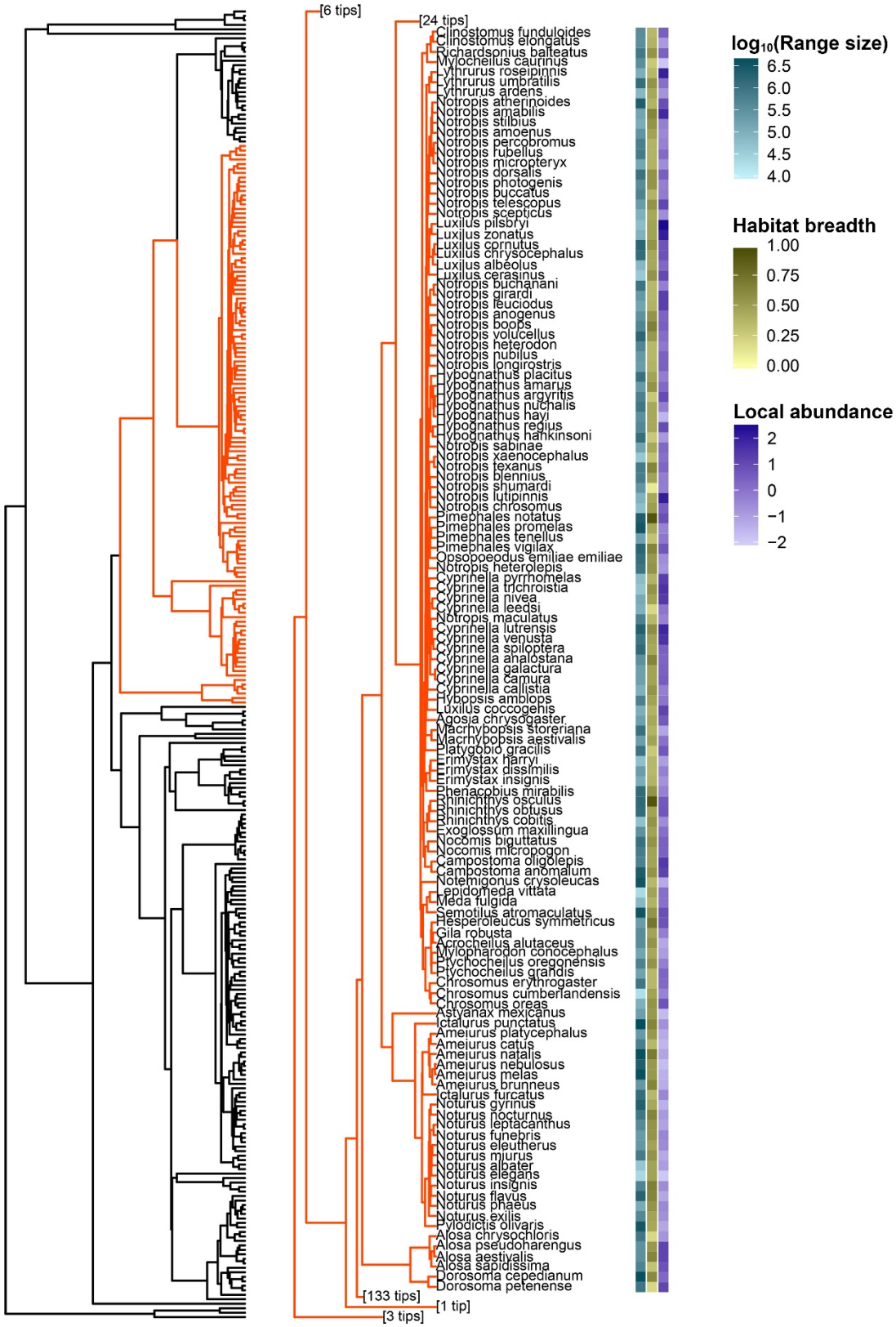


Appendix 2. Sampling intensity among regions. Sampling intensity is measured as number of sampling sites divided by the number of perennial stream reaches sampled within each USGS HUC2 region. Sampling intensity of all regions are within 2 standard deviations of the mean.



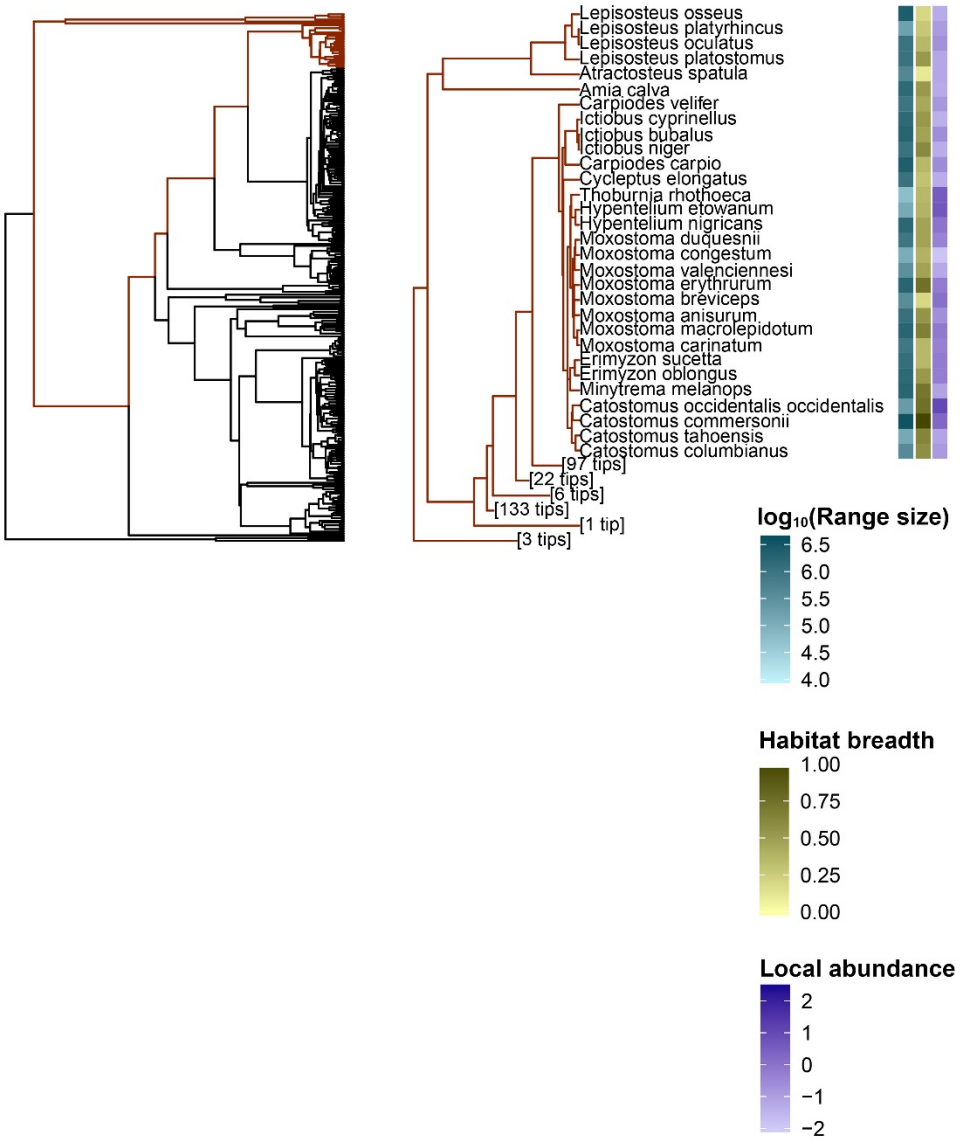
Appendix 3. Phylogenetic tree and metrics of rarity for 292 species. The different colors (yellow, orange, brown) indicate the position of the sub-trees within the overall phylogenetic tree.




[6 tips]
[24 tips]
Clinostomus funduloides
Clinostomus elongatus
Richardsonius balteatus
Mylocheilus caurinus
Lythrurus roseipinnis
Lythrurus umbratilis
Lythrurus ardens
Notropis atherinoides
Notropis amabilis
Notropis stilbius
Notropis amoenus
Notropis percobromus
Notropis rubellus
Notropis micropteryx
Notropis dorsalis
Notropis photogenis
Notropis buccatus
Notropis telescopus
Notropis scepticus
Luxilus pilsbryi
Luxilus zonatus
Luxilus cornutus
Luxilus chrysocephalus
Luxilus albeolus
Luxilus cerasinus
Notropis buchanani
Notropis girardi
Notropis leuciodus
Notropis anogenus
Notropis boops
Notropis volucellus
Notropis heterodon
Notropis nubilus
Notropis longirostris
Hybognathus placitus
Hybognathus amarus
Hybognathus argyritis
Hybognathus nuchalis
Hybognathus hayi
Hybognathus regius
Hybognathus hankinsoni
Notropis sabinae
Notropis xaenocephalus
Notropis texanus
Notropis blennius
Notropis shumardi
Notropis lutipinnis
Notropis chrosomus
Pimephales notatus
Pimephales promelas
Pimephales tenellus
Pimephales vigilax
Opsopoeodus emiliae emiliae
Notropis heterolepis
Cyprinella pyrrhomelas
Cyprinella trichroistia
Cyprinella nivea
Cyprinella leedsi
Notropis maculatus
Cyprinella lutrensis
Cyprinella venusta
Cyprinella spiloptera
Cyprinella analostana
Cyprinella galactura
Cyprinella camura
Cyprinella callistia
Hybopsis amblops
Luxilus coccogenis
Agosia chrysogaster
Macrhybopsis storeriana
Macrhybopsis aestivalis
Platygobio gracilis
Erimystax harryi
Erimystax dissimilis
Erimystax insignis
Phenacobius mirabilis
Rhinichthys osculus
Rhinichthys obtusus
Rhinichthys cobitis
Exoglossum maxillingua
Nocomis biguttatus
Nocomis micropogon
Campostoma oligolepis
Campostoma anomalum
Notemigonus crysoleucas
Lepidomeda vittata
Meda fulgida
Semotilus atromaculatus
Hesperoleucus symmetricus
Gila robusta
Acrocheilus alutaceus
Mylopharodon conocephalus
Ptychocheilus oregonensis
Ptychocheilus grandis
Chrosomus erythrogaster
Chrosomus cumberlandensis
Chrosomus oreas
Astyanax mexicanus
Ictalurus punctatus
Ameiurus platycephalus
Ameiurus catus
Ameiurus natalis
Ameiurus nebulosus
Ameiurus melas
Ameiurus brunneus
Ictalurus furcatus
Noturus gyrinus
Noturus nocturnus
Noturus leptacanthus
Noturus funebris
Noturus eleutherus
Noturus miurus
Noturus albater
Noturus elegans
Noturus insignis
Noturus flavus
Noturus phaeus
Noturus exilis
Pylodictis olivaris
Alosa chrysochloris
Alosa pseudoharengus
Alosa aestivalis
Alosa sapidissima
Dorosoma cepedianum
Dorosoma petenense
[133 tips]
[1 tip]
[3 tips]
log10(Range size)
6.5
6.0
5.5
5.0
4.5
4.0
Habitat breadth
1.00
0.75
0.50
0.25
0.00
Local abundance
2
1
0
−1
−2

54

55


Lepisosteus osseus
Lepisosteus platyrhincus
Lepisosteus oculatus
Lepisosteus platostomus
Atractosteus spatula
Amia calva
Carpiodes velifer
Ictiobus cyprinellus
Ictiobus bubalus
Ictiobus niger
Carpiodes carpio
Cycleptus elongatus
Thoburnia rhothoeca
Hypentelium etowanum
Hypentelium nigricans
Moxostoma duquesnii
Moxostoma congestum
Moxostoma valenciennesi
Moxostoma erythrurum
Moxostoma breviceps
Moxostoma anisurum
Moxostoma macrolepidotum
Moxostoma carinatum
Erimyzon sucetta
Erimyzon oblongus
Minytrema melanops
Catostomus occidentalis occidentalis
Catostomus commersonii
Catostomus tahoensis
Catostomus columbianus
[97 tips]
[22 tips]
[6 tips]
[133 tips]
[1 tip]
[3 tips]
log10(Range size)
6.5
6.0
5.5
5.0
4.5
4.0
Habitat breadth
1.00
0.75
0.50
0.25
0.00
Local abundance
2
1
0
−1
−2

56

57

58 **Appendix 4.** Model rankings for all response variables. Acronyms for the predictor variables are
59 BS: body size, EGG: egg size, HB: habitat breadth, LAC: use of lacustrine habitats, LG: use of
60 large rivers, PISC: piscivory, PC: parental care, WEST: fishes belonging to western USA.

61

62 **Table A1.** Top 10 models for local abundance among fishes occurring in all streams. AIC_c:
63 sample-size corrected Akaike Information criterion; w: Akaike weight (the probability that the
64 model is the best out of the candidate model set); λ: measure of phylogenetic correlation among
65 model residuals; R^2: Cox-Snell R^2.

66

Rank	Model coefficients	AIC_c	w	λ	R^2
1	–0.23LG – 0.31LAC + 1.14HB – 0.75	668.75	0.120	0.67	0.09
2	0.14PC – 0.22LG – 0.32LAC + 1.12HB – 0.78	669.87	0.068	0.66	0.09
3	0.06EGG – 0.23LG – 0.31LAC + 1.14HB – 0.91	670.24	0.057	0.68	0.09
4	–0.13PISC – 0.21LG – 0.31LAC + 1.13HB – 0.70	670.29	0.055	0.66	0.09
5	–0.21LG – 0.30LAC + 1.13HB – 0.05BS – 0.41	670.39	0.053	0.65	0.09
6	0.09EGG – 0.19LG – 0.28LAC + 1.13HB – 0.09BS – 0.45	671.32	0.033	0.66	0.09
7	–0.14PISC + 0.14PC – 0.20LG – 0.32LAC + 1.11HB – 0.73	671.36	0.032	0.64	0.09
8	– 0.36LAC + 1.11HB – 0.90	671.46	0.031	0.69	0.07
9	0.14PC – 0.20LG – 0.31LAC + 1.12HB – 0.06BS – 0.43	671.49	0.030	0.64	0.09
10	–0.16PISC + 0.07EGG – 0.20LG – 0.30LAC + 1.12HB – 0.88	671.58	0.029	0.67	0.09

67

68

Table A2. Top 10 models for habitat breadth among fishes occurring in all streams. AIC_c: sample-size corrected Akaike Information criterion; w: Akaike weight (the probability that the model is the best out of the candidate model set); λ: measure of phylogenetic correlation among model residuals; R^2: Cox-Snell R^2.

Rank	Model coefficients	AIC_c	w	λ	R^2
1	$0.03LAC + 0.35BS - 0.03BS^2 - 0.53$	–342.51	0.173	0.17	0.05
2	$0.01EGG + 0.04LAC + 0.37BS - 0.03BS^2 - 0.58$	–341.52	0.105	0.13	0.06
3	$0.02PC + 0.03LAC + 0.34BS - 0.03BS^2 - 0.50$	–341.40	0.099	0.15	0.05
4	$0.35BS - 0.03BS^2 - 0.52$	–341.29	0.094	0.15	0.04
5	$-0.02PISC + 0.03LAC + 0.35BS + 0.03BS^2 - 0.53$	–340.76	0.072	0.15	0.05
6	$0.02PC + 0.34BS - 0.03BS^2 - 0.49$	–340.76	0.072	0.13	0.05
7	$0.01EGG + 0.02PC + 0.03LAC + 0.36BS + 0.03BS^2 - 0.55$	–339.95	0.048	0.13	0.06
8	$-0.03PISC + 0.02PC + 0.03LAC + 0.34BS - 0.03BS^2 - 0.51$	–339.87	0.046	0.13	0.06
9	$-0.02PISC + 0.01EGG + 0.04LAC + 0.37BS - 0.03BS^2 - 0.59$	–339.77	0.044	0.11	0.06
10	$0.01EGG + 0.36BS - 0.03BS^2 - 0.57$	–339.77	0.044	0.12	0.05

Table A3. Top 10 models for geographic range size among fishes occurring in all streams. AIC_c: sample-size corrected Akaike Information criterion; w: Akaike weight (the probability that the model is the best out of the candidate model set); λ: measure of phylogenetic correlation among model residuals; R^2: Cox-Snell R^2.

Rank	Model coefficients	AIC_c	w	λ	R^2
1	–0.14EGG + 0.11PC + 0.17LG + 0.42LAC – 0.43WEST + 0.61HB + 0.20BS + 4.26	389.89	0.141	0.07	0.38
2	–0.12EGG + 0.17LG + 0.44LAC – 0.44WEST + 0.63HB + 0.19BS + 4.28	390.38	0.110	0.11	0.36
3	0.09PISC – 0.14EGG + 0.10PC + 0.17LG + 0.42LAC – 0.43WEST + 0.62HB + 0.18BS + 4.34	391.40	0.066	0.07	0.38
4	–0.13EGG + 0.10PC + 0.18LG + 0.43LAC – 0.44WEST + 0.58HB + 0.48BS – 0.03BS^2 + 3.53	391.42	0.065	0.06	0.38
5	0.11PISC – 0.12EGG + 0.16LG + 0.43LAC – 0.44WEST + 0.64HB + 0.16BS + 4.39	391.42	0.065	0.10	0.36
6	–0.11EGG + 0.17LG + 0.45LAC – 0.45WEST + 0.59HB + 0.55BS – 0.03BS^2 + 3.34	391.48	0.063	0.09	0.37
7	–0.14EGG + 0.11PC + 0.18LG + 0.42LAC – 0.44WEST + 0.61HB + 0.01ABUN + 0.20BS + 4.26	392.02	0.048	0.07	0.38
8	0.12PISC – 0.11EGG + 0.17LG + 0.45LAC – 0.44WEST + 0.60HB + 0.56BS – 0.04BS^2 + 3.36	392.29	0.042	0.08	0.37
9	–0.12EGG + 0.17LG + 0.44LAC – 0.44WEST + 0.62HB + 0.01ABUN + 0.19BS + 4.28	392.49	0.038	0.11	0.36
10	0.10PISC – 0.13EGG + 0.09PC + 0.18LG + 0.43LAC – 0.43WEST + 0.59HB + 0.49BS – 0.03BS^2 + 3.53	392.78	0.033	0.07	0.38

Table A4. Top 10 models for local abundance among fishes occurring in small streams (1st–3rd order). AIC_c: sample-size corrected Akaike Information criterion; w: Akaike weight (the probability that the model is the best out of the candidate model set); λ: measure of phylogenetic correlation among model residuals; R^2: Cox-Snell R^2.

Rank	Model coefficients	AIC_c	w	λ	R^2
1	0.22EGG + 0.26PC – 0.32LG – 0.29LAC – 0.21BS + 0.51HB + 0.25	499.56	0.065	0.68	0.18
2	0.23EGG + 0.26PC – 0.30LG – 0.29LAC – 0.21BS + 0.39	499.87	0.056	0.69	0.17
3	0.24EGG + 0.24PC – 0.29LG – 0.28LAC + 0.82BS – 0.10BS^2 – 2.29	500.22	0.047	0.68	0.17
4	0.26EGG – 0.33LG – 0.26LAC + 0.50HB – 0.22BS + 0.27	500.33	0.044	0.70	0.16
5	0.23EGG + 0.25PC – 0.30LG – 0.28LAC + 0.46HB + 0.67BS – 0.08BS^2 – 2.01	500.46	0.042	0.68	0.18
6	0.27EGG – 0.32LG – 0.26LAC – 0.21BS + 0.41	500.54	0.040	0.71	0.15
7	0.28EGG – 0.30LG – 0.25LAC + 0.91BS – 0.10BS^2 – 2.51	500.57	0.039	0.70	0.16
8	0.27EGG – 0.32LG – 0.25LAC + 0.44HB + 0.76BS – 0.09BS^2 – 2.24	500.91	0.033	0.69	0.17
9	–0.12PISC + 0.22EGG + 0.26PC – 0.31LG – 0.30LAC + 0.50HB – 0.18BS + 0.12	501.49	0.025	0.67	0.18
10	–0.14PISC + 0.23EGG + 0.26PC – 0.30LG – 0.29LAC – 0.18BS + 0.24	501.71	0.022	0.68	0.17

Table A5. Top 10 models for habitat breadth among fishes occurring in small (1st–3rd order) streams. AIC_c: sample-size corrected Akaike Information criterion; w: Akaike weight (the probability that the model is the best out of the candidate model set); λ: measure of phylogenetic correlation among model residuals; R^2: Cox-Snell R^2.

Rank	Model coefficients	AIC_c	w	λ	R^2
1	$0.33BS - 0.03BS^2 - 0.53$	–203.85	0.091	0.17	0.02
2	$0.03EGG + 0.35BS - 0.03BS^2 - 0.60$	–203.75	0.087	0.14	0.03
3	$-0.05PISC + 0.03EGG + 0.33BS - 0.03BS^2 - 0.56$	–203.01	0.060	0.09	0.04
4	$-0.05PISC + 0.31BS - 0.03BS^2 - 0.49$	–202.84	0.055	0.12	0.03
5	0.37	–202.81	0.054	0.20	0.00
6	$0.02EGG + 0.32$	–202.44	0.045	0.16	0.01
7	$0.03EGG + 0.02LAC + 0.35BS - 0.03BS^2 - 0.61$	–202.01	0.036	0.15	0.04
8	$0.01LAC + 0.33BS - 0.03BS^2 - 0.54$	–201.89	0.034	0.18	0.02
9	$0.03EGG - 0.01PC + 0.35BS - 0.03BS^2 - 0.63$	–201.79	0.033	0.13	0.04
10	$-0.00003PC + 0.33BS - 0.03BS^2 - 0.53$	–201.73	0.032	0.17	0.02

Table A6. Top 10 models for geographic range size among fishes occurring in small (1st–3rd order) streams. AIC_c: sample-size corrected Akaike Information criterion; w: Akaike weight (the probability that the model is the best out of the candidate model set); λ: measure of phylogenetic correlation among model residuals; R^2: Cox-Snell R^2.

Rank	Model coefficients	AIC_c	w	λ	R^2
1	–0.11EGG + 0.18LG + 0.44LAC – 0.43WEST + 0.40HB + 0.21BS + 4.39	300.02	0.087	0.13	0.36
2	–0.13EGG + 0.09PC + 0.19LG + 0.42LAC – 0.42WEST + 0.40HB + 0.21BS + 4.39	300.89	0.056	0.09	0.37
3	–0.10EGG + 0.20LG + 0.43LAC – 0.39WEST + 0.21BS + 4.52	301.32	0.045	0.13	0.35
4	0.21LG + 0.46LAC – 0.47WEST + 0.37HB + 0.17BS + 4.38	301.47	0.042	0.17	0.34
5	0.11PISC – 0.11EGG + 0.18LG + 0.44LAC – 0.42WEST + 0.42HB + 0.18BS + 4.51	301.47	0.042	0.11	0.37
6	–0.11EGG + 0.19LG + 0.44LAC – 0.43WEST + 0.39HB + 0.34BS – 0.01BS^2 + 4.02	302.13	0.030	0.12	0.36
7	–0.11EGG + 0.18LG + 0.43LAC – 0.43WEST + 0.40HB – 0.01ABUN + 0.21BS + 4.39	302.18	0.029	0.13	0.36
8	–0.12EGG + 0.09PC + 0.21LG + 0.42LAC – 0.38WEST + 0.21BS + 4.51	302.18	0.029	0.10	0.36
9	0.22LG + 0.46LAC – 0.43WEST + 0.17BS + 4.50	302.33	0.027	0.16	0.33
10	0.09PISC – 0.13EGG + 0.08PC + 0.19LG + 0.42LAC – 0.42WEST + 0.41HB + 0.19BS + 4.48	302.59	0.024	0.09	0.38

Table A7. Top 10 models for local abundance among fishes occurring in medium-to-large (≥4th order) streams. AIC_c: sample-size corrected Akaike Information criterion; w: Akaike weight (the probability that the model is the best out of the candidate model set); λ: measure of phylogenetic correlation among model residuals; R^2: Cox-Snell R^2.

Rank	Model coefficients	AIC_c	w	λ	R^2
1	–0.20LG – 0.27LAC + 1.38HB – 0.81	610.09	0.111	0.60	0.08
2	–0.31LAC + 1.37HB – 0.94	610.87	0.075	0.61	0.07
3	–0.15PISC – 0.17LG – 0.30LAC + 1.36HB – 0.75	611.56	0.053	0.58	0.09
4	–0.21PISC – 0.30LAC + 1.35HB –0.84	611.66	0.051	0.58	0.08
5	0.05EGG – 0.19LG – 0.27LAC + 1.38HB – 0.93	611.93	0.044	0.61	0.08
6	–0.20LG – 0.28LAC + 1.38HB + 0.02BS – 0.90	612.18	0.039	0.60	0.08
7	0.02PC – 0.19LG – 0.27LAC + 1.37HB –0.81	612.19	0.039	0.60	0.08
8	0.05EGG – 0.30LAC + 1.38HB – 1.08	612.61	0.032	0.62	0.07
9	0.05PC – 0.31LAC + 1.36HB – 0.95	612.86	0.028	0.61	0.07
10	–0.30LAC + 1.37HB – 0.03BS – 0.78	612.88	0.028	0.60	0.07

Table A8. Top 10 models for habitat breadth among fishes occurring in medium-to-large (≥4th order) streams. AIC_c: sample-size corrected Akaike Information criterion; w: Akaike weight (the probability that the model is the best out of the candidate model set); λ: measure of phylogenetic correlation among model residuals; R^2: Cox-Snell R^2.

Rank	Model coefficients	AIC_c	w	λ	R^2
1	$0.03LAC + 0.33BS - 0.03BS^2 - 0.47$	–287.50	0.113	0.15	0.05
2	$0.03PC + 0.33BS - 0.03BS^2 - 0.46$	–287.06	0.091	0.12	0.05
3	$0.35BS - 0.03BS^2 - 0.51$	–286.98	0.088	0.13	0.04
4	$0.03PC + 0.03LAC + 0.32BS - 0.03BS^2 - 0.43$	–286.88	0.084	0.14	0.06
5	$0.02EGG + 0.03LAC + 0.36BS - 0.03BS^2 - 0.55$	–286.67	0.075	0.11	0.06
6	$-0.04PISC + 0.04PC + 0.33BS - 0.03BS^2 - 0.48$	–286.32	0.063	0.07	0.06
7	$-0.03PISC + 0.03LAC + 0.34BS - 0.03BS^2 - 0.49$	–286.00	0.054	0.12	0.05
8	$-0.04PISC + 0.03PC + 0.03LAC + 0.32BS - 0.03BS^2 - 0.45$	–285.86	0.050	0.09	0.06
9	$0.01EGG + 0.38BS - 0.03BS^2 - 0.59$	–285.79	0.048	0.09	0.05
10	$-0.03PISC + 0.35BS - 0.03BS^2 - 0.53$	–285.60	0.044	0.10	0.05

Table A9. Top 10 models for geographic range size among fishes occurring in medium-to-large (≥4th order) streams. AIC_c: sample-size corrected Akaike Information criterion; w: Akaike weight (the probability that the model is the best out of the candidate model set); λ: measure of phylogenetic correlation among model residuals; R^2: Cox-Snell R^2.

Rank	Model coefficients	AIC_c	w	λ	R^2
1	–0.15EGG + 0.15PC + 0.13LG + 0.38LAC – 0.45WEST + 0.65HB + 0.18BS + 4.42	314.36	0.187	0.06	0.38
2	0.07PISC – 0.15EGG + 0.14PC + 0.13LG + 0.38LAC – 0.45WEST + 0.66HB + 0.17BS + 4.49	316.15	0.076	0.06	0.38
3	–0.16EGG + 0.15PC + 0.39LAC – 0.46WEST + 0.65HB + 0.21BS + 4.33	316.16	0.076	0.07	0.36
4	–0.14EGG + 0.14PC + 0.14LG + 0.38LAC – 0.46WEST + 0.63HB + 0.38BS – 0.02BS^2 + 3.91	316.26	0.072	0.05	0.38
5	–0.15EGG + 0.15PC + 0.13LG + 0.37LAC – 0.45WEST + 0.65HB – 0.01ABUN + 0.18BS + 4.42	316.52	0.063	0.06	0.38
6	–0.12EGG + 0.12LG + 0.41LAC – 0.47WEST + 0.65HB + 0.18BS + 4.47	317.08	0.048	0.09	0.35
7	0.08PISC – 0.16EGG + 0.13PC + 0.39LAC – 0.45WEST + 0.67HB + 0.19BS + 4.42	317.74	0.034	0.07	0.36
8	0.07PISC – 0.14EGG + 0.13PC + 0.13LG + 0.38LAC – 0.45WEST + 0.63HB + 0.38BS – 0.02BS^2 + 3.91	318.00	0.030	0.06	0.38
9	0.11PISC – 0.12EGG + 0.12LG + 0.41LAC – 0.46WEST + 0.69HB + 0.14BS + 4.58	318.13	0.028	0.09	0.35
10	–0.16EGG + 0.14PC + 0.39LAC – 0.46WEST + 0.64HB + 0.34BS – 0.01BS^2 + 3.98	318.21	0.027	0.06	0.37

Appendix 5. Final phylogenetic logistic regression (PLR) models for use of lacustrine habitats and use of large rivers

Table A10. Backward selection to obtain final PLR model for the use of lacustrine habitats among fishes occurring in all streams.

Step	Model coefficients	*p*-values			
		BS	PISC	EGG	PC
1	1.32BS – 0.86PISC – 0.52EGG + 1.08PC – 6.33	<0.0001	0.12	0.03	0.003
2 (Final)	1.17BS –0.61EGG + 1.02PC – 5.94 (R^2 = 0.22, α = 0.106)	<0.0001		0.007	0.001

Table A11. Backward selection to obtain final PLR model for the use of large rivers among fishes occurring in all streams.

Step	Model coefficients	*p*-values			
		BS	PISC	EGG	PC
1	1.58BS – 0.21PISC – 0.50EGG – 0.54PC – 7.82	<0.0001	0.72	0.06	0.20
2	1.65BS – 0.50EGG – 0.56PC – 8.15	<0.0001		0.06	0.18
3 (Final)	1.66BS – 0.64EGG – 8.19 (R^2 = 0.25, α = 0.033)	<0.0001		0.01	

133 **Table A12.** Backward selection to obtain final PLR model for the use of lacustrine habitats
134 among fishes occurring in small (1st–3rd order) streams.

Step	Model coefficients	*p*-values			
		BS	PISC	EGG	PC
1	1.50BS – 1.16PISC – 0.86EGG + 1.31PC – 6.33	<0.0001	0.11	0.005	0.002
2 (Final)	1.18BS – 0.68EGG + 0.82PC – 5.67 (R^2 = 0.24, α = 0.088)	<0.0001		0.01	0.02

135
136
137 **Table A13.** Backward selection to obtain final PLR model for the use of large rivers among
138 fishes occurring in small (1st–3rd order) streams.

Step	Model coefficients	*p*-values			
		BS	PISC	EGG	PC
1	1.66BS – 0.10PISC – 0.45EGG – 0.39PC – 8.47	<0.0001	0.88	0.13	0.37
2	1.68BS –0.48EGG – 0.42PC – 8.54	<0.0001		0.11	0.33
3 (Final)	1.67BS – 0.54EGG – 8.58 (R^2 = 0.28, α = 0.051)	<0.0001		0.07	

139

Table A14. Backward selection to obtain final PLR model for the use of lacustrine habitats among fishes occurring in medium-to-large (≥4th order) streams.

Step	Model coefficients	*p*-values			
		BS	PISC	EGG	PC
1	1.17BS – 0.17PISC – 0.45EGG + 0.91PC – 6.24	<0.0001	0.73	0.05	0.006
2 (Final)	1.16BS – 0.54EGG + 0.89PC – 5.98 (R^2 = 0.26, α = 0.083)	<0.0001		0.02	0.01

Table A15. Backward selection to obtain final PLR model for the use of large rivers among fishes occurring in medium-to-large (≥4th order) streams.

Step	Model coefficients	*p*-values			
		BS	PISC	EGG	PC
1	1.54BS +0.04PISC – 0.47EGG – 0.42PC – 7.60	<0.0001	0.96	0.12	0.17
2	1.59BS –0.49EGG – 0.40PC – 7.83	<0.0001		0.07	0.34
3 (Final)	1.61BS – 0.62EGG – 7.88 (R^2 = 0.23, α = 0.033)	<0.0001		0.02	

Appendix 6. Shipley's d-separation test for validity of final path models.

Shipley's d-separation (or d-sep) test evaluates the validity of a given path model by examining whether patterns of conditional independence implicit in the path model are actually observed in the empirical data (Shipley 2016). In a path model, variables that are not directly connected by an arrow are independent conditional upon (i.e., after controlling for) their direct causes. For example, in the final model for all fishes (Fig. 4), piscivory is not directly connected to local abundance, therefore, these two variables are expected to be independent after controlling for the effects of all of their direct causes (i.e., use of lacustrine habitats, use of large rivers, and habitat breadth). This conditional independence statement or claim can be written as (piscivory, local abundance) | {use of lacustrine habitats, use of large rivers, habitat breadth}, which can be read as piscivory and local abundance should be independent after controlling for use of lacustrine habitats, use of large rivers, and habitat breadth.

For a given path model, we need to identify all conditional independence claims and calculate the *p*-value associated with each of these claims using the statistical test appropriate for the response and predictor variables. These probabilities are then combined into a *C* statistic, which is tested against a χ^2 distribution with $2k$ degrees of freedom where k is the number of conditional independence claims, to determine if the path model should be rejected. In our analysis, we used an alpha value of 0.05 (i.e., path models with $p < 0.05$ would be rejected).

We performed the d-sep test on the path models describing mechanisms driving rarity among all fishes (Table A16), fishes in small streams (Table A17), and fishes in medium-to-large streams (Table A18). Based on their *p*-values ($p = 0.81$, 0.19, 0.94, respectively), we concluded that all three models were valid.

Table A16. Evaluation of conditional independence claims implicit in the final path model for fishes in all streams (Fig. 4a). Abbreviations for predictors are: range size (RANGE), habitat breadth (HB), local abundance (ABUN), use of large rivers (LG), use of lacustrine habitats (LAC), native to western US (WEST), body size (BS), egg size (EGG), parental care (PC), and piscivory (PISC). Ø represents a null set of control variables (i.e., no control variables are required).

Conditional independence claim	*p*-value
1. (EGG, WEST) | {Ø}	0.3867
2. (HB, EGG) | {BS, BS^2, LAC}	0.2722
3. (ABUN, EGG) | {LG, HB, LAC}	0.4362
4. (BS, WEST) | {Ø}	0.5360
5. (ABUN, BS) | {LG, HB, LAC}	0.4928
6. (PC, WEST) | {Ø}	0.9904
7. (LG, PC) | {EGG, BS}	0.1796
8. (HB, PC) | {BS, BS^2, LAC}	0.3213
9. (ABUN, PC) | {LG, HB, LAC}	0.3257
10. (LG, WEST) | {BS, EGG}	0.4475
11. (HB, LG) | {BS, BS^2, LAC, EGG}	0.3402
12. (LAC, WEST) | {BS, EGG, PC}	0.1147
13. (HB, WEST) | {BS, BS^2, LAC}	0.2787
14. (ABUN, WEST) | {LG, HB, LAC}	0.9288
15. (RANGE, ABUN) | {WEST, LG, BS, PC, EGG, HB, LAC}	0.8751
16. (PISC, WEST) | {Ø}	0.5390
17. (LG, PISC) | {EGG, BS}	0.9150
18. (RANGE, PISC) | {WEST, LG, BS, PC, EGG, HB, LAC}	0.4277
19. (HB, PISC) | {BS, BS^2, LAC}	0.5472
20. (ABUN, PISC) | {LG, HB, LAC}	0.4554
21. (LAC, PISC) | {EGG, BS, PC}	0.5153
22. (PC, BS) | {Ø}	0.9492
23. (PISC, PC) | {Ø}	0.3066

$C = -2\sum_{j=1}^{k} ln(p_j)$	37.5442
Overall p-value ($\chi^2_{df=46}$)	0.8083

176

177

178 **Table A17.** Evaluation of conditional independence claims implicit in the final path model for
179 fishes in small (1st–3rd order) streams (Fig. 4b). Abbreviations follow those in Table A1.

Conditional independence claim	*p*-value
1. (EGG, WEST) | {Ø}	0.1364
2. (HB, EGG) | {BS + BS^2}	0.1475
3. (BS, WEST) | {Ø}	0.6012
4. (PC, WEST) | {Ø}	0.9915
5. (LG, PC) | {EGG, BS}	0.3320
6. (RANGE, PC) | {WEST, LG, BS, EGG, HB, LAC}	0.2384
7. (HB, PC) | {BS, BS^2}	0.9991
8. (LG, WEST) | {BS, EGG}	0.4509
9. (HB, LG) | {BS, BS^2, EGG}	0.0650
10. (LAC, WEST) | {BS, EGG, PC}	0.0098
11. (HB, LAC) | {BS, BS^2, PC, EGG}	0.4846
12. (HB, WEST) | {BS, BS^2}	0.0171
13. (ABUN, WEST) | {LG, EGG, PC, BS, LAC, HB}	0.9546
14. (RANGE, ABUN) | {WEST, LG, BS, PC, EGG, HB, LAC}	0.7706
15. (PISC, WEST) | {Ø}	0.6473
16. (LG, PISC) | {EGG, BS}	0.8869
17. (LAC, PISC) | {EGG, BS, PC}	0.1101
18. (RANGE, PISC) | {WEST, EGG, LG, BS, LAC, HB}	0.3875
19. (HB, PISC) | {BS, BS^2}	0.2467
20. (ABUN, PISC) | {LG, EGG, BS, LAC, PC, HB}	0.6064
21. (PC, BS) | {Ø}	0.9966
22. (PISC, PC) | {Ø}	0.7992
$C = -2\sum_{j=1}^{k} ln(p_j)$	52.1006
Overall *p*-value ($\chi^2_{df=44}$)	0.1879

180
181

Table A18. Evaluation of conditional independence claims implicit in the final path model for fishes in medium-to-large (≥4th order) streams (Fig. 4b). Abbreviations follow those in Table A1.

Conditional independence claim	*p*-value
1. (EGG, WEST) | {Ø}	0.2413
2. (HB, EGG) | {BS, BS^2, LAC}	0.2374
3. (ABUN, EGG) | {LG, HB, LAC}	0.5983
4. (BS, WEST) | {Ø}	0.8305
5. (ABUN, BS) | {LG, HB, LAC}	0.8605
6. (PC, WEST) | {Ø}	0.9997
7. (LG, PC) | {EGG, BS}	0.3398
8. (ABUN, PC) | {LG, HB, LAC}	0.8899
9. (LG, WEST) | {BS, EGG}	0.8067
10. (HB, LG) | {BS, BS^2, LAC, EGG}	0.5571
11. (LAC, WEST) | {BS, EGG, PC}	0.0688
12. (HB, PC) | {BS, BS^2, LAC, EGG}	0.3516
13. (HB, WEST) | {BS, BS^2, LAC}	0.3255
14. (ABUN, WEST) | {LG, HB, LAC}	0.9930
15. (RANGE, ABUN) | {WEST, LG, BS, PC, EGG, HB, LAC}	0.8795
16. (PISC, WEST) | {Ø}	0.2305
17. (LG, PISC) | {EGG, BS}	0.8301
18. (RANGE, PISC) | {WEST, LG, BS, PC, EGG, HB, LAC}	0.5336
19. (HB, PISC) | {BS, BS^2, LAC}	0.4087
20. (ABUN, PISC) | {LG, HB, LAC}	0.4188
21. (LAC, PISC) | {EGG, BS, PC}	0.3570
22. (PC, BS) | {Ø}	0.9991
23. (PISC, PC) | {Ø}	0.8592
$C = -2\sum_{j=1}^{k} ln(p_j)$	31.8313
Overall *p*-value ($\chi^2_{df=46}$)	0.9443

References

Shipley, B. 2016. *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference with R, 2nd Edition*. - Cambridge University Press.