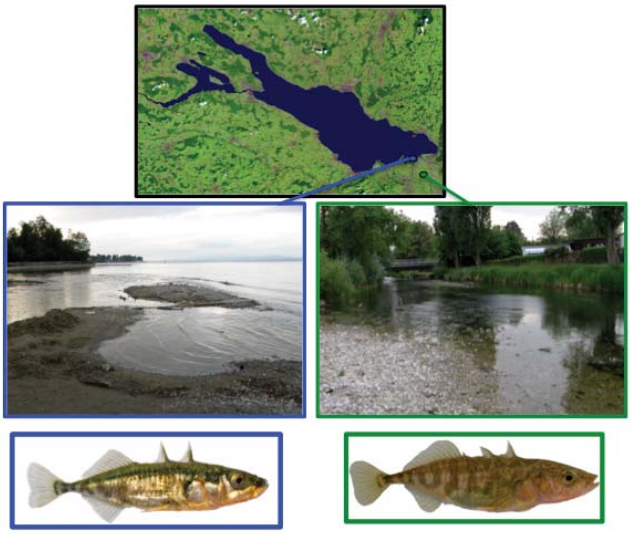
Supplemental Figures (S1-S3):

Figure S1.

(1)



(2)

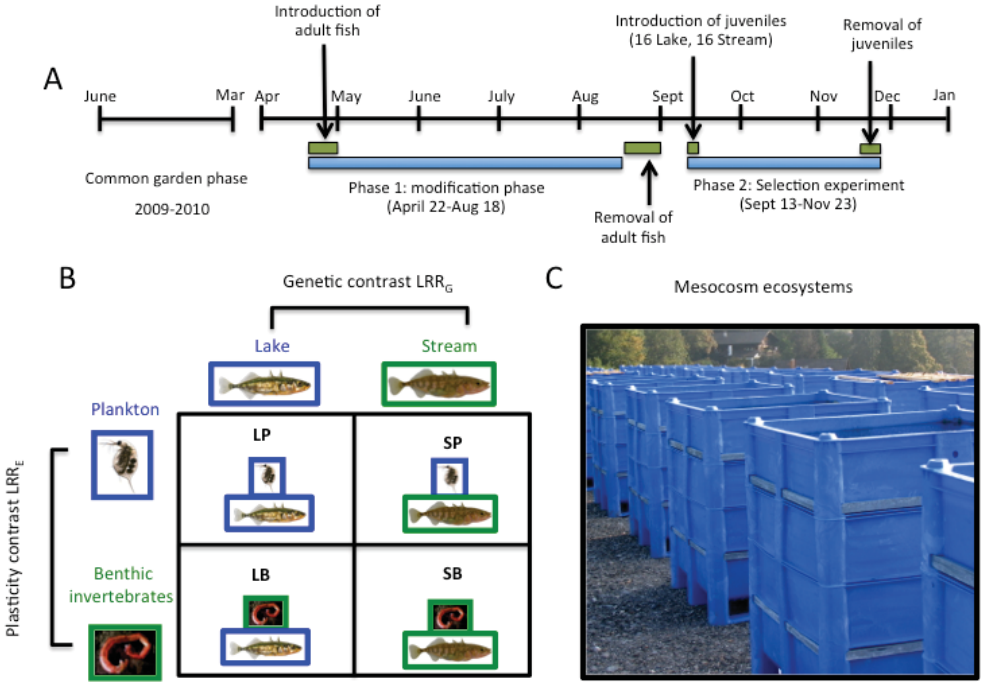


Figure S2.

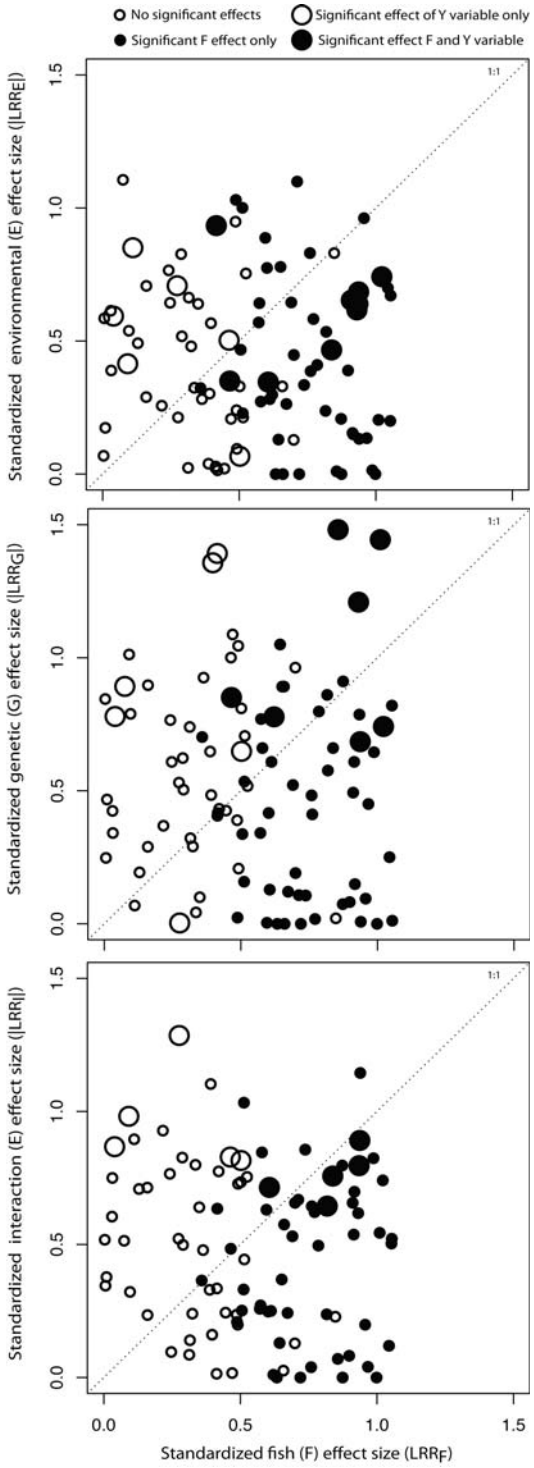

No significant effects
Significant effect of Y variable only
Significant F effect only
Significant effect F and Y variable
1:1
Standardized environmental (E) effect size (|LRR$_E$|)
Standardized genetic (G) effect size (|LRR$_G$|)
Standardized interaction (E) effect size (|LRR$_I$|)
0.0
0.5
1.0
1.5
Standardized fish (F) effect size (LRR$_F$)

Figure S3.

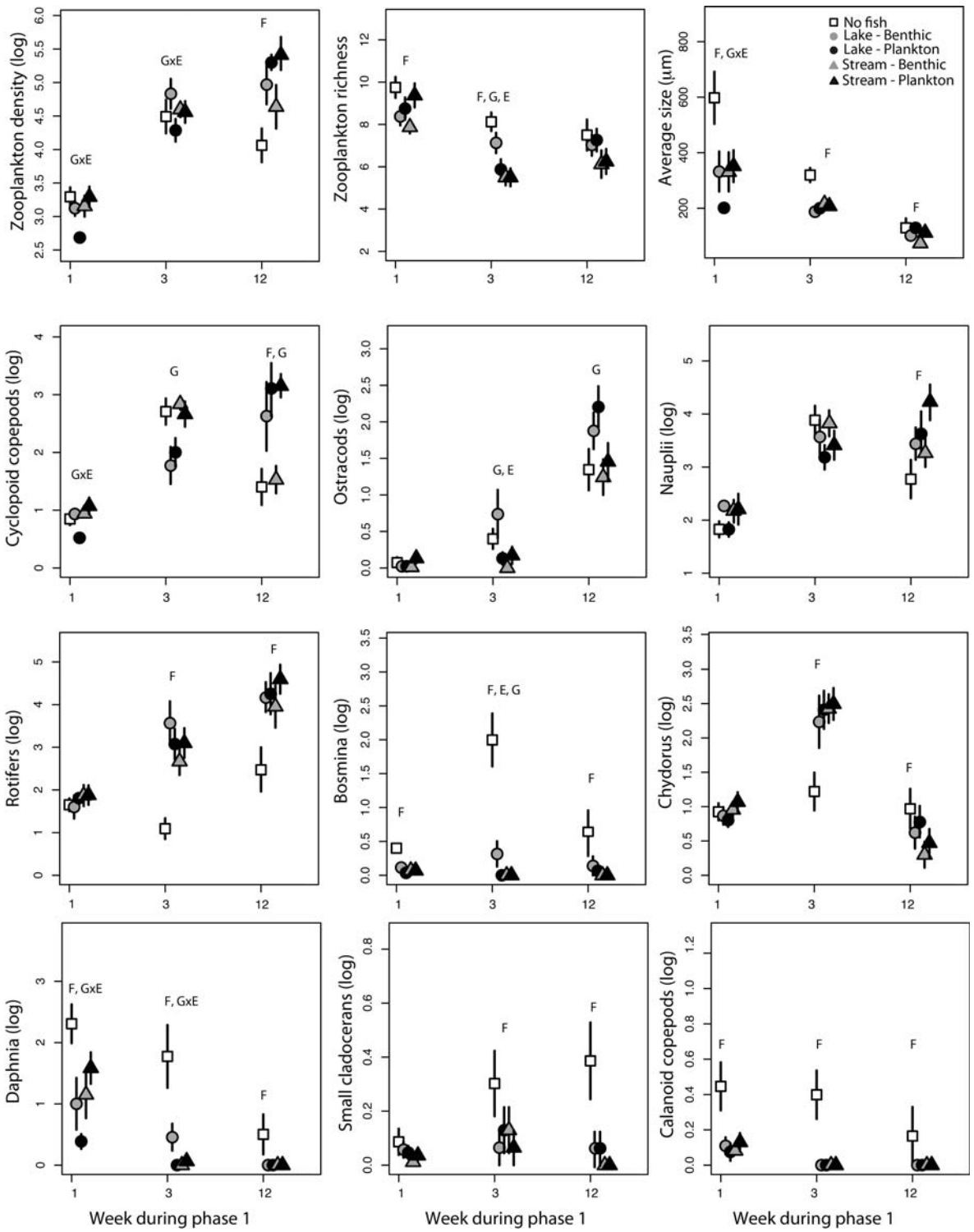

Zooplankton density (log)
Zooplankton richness
Average size (µm)
No fish
Lake - Benthic
Lake - Plankton
Stream - Benthic
Stream - Plankton
Cyclopoid copepods (log)
Ostracods (log)
Nauplii (log)
Rotifers (log)
Bosmina (log)
Chydorus (log)
Daphnia (log)
Small cladocerans (log)
Calanoid copepods (log)
Week during phase 1
Week during phase 1
Week during phase 1

Figure S1, related to Figure 1. (Panel 1) Photographs from the two sampling locations of the adult fish used in the mesocosm experiment. Lake fish (Blue border) were caught from the southeastern shores of Lake Constance (47°29'08.29"N 09°32'37.85"E), and stream fish were caught several kilometers upstream at Aubach, Oberriet (47°19'36.80"N 09°34'04.13"E). **(**Panel 2**)** Experimental design to complement results from Figure 1, showing (A) the time course of the experiment, (B) the design of the common garden phase where sticklebacks were reared in the lab prior to being used in Phase 1. (C) Shows the array of mesocosms.

Figure S2, related to Figure 2. Comparison of the absolute value of ecosystem effect contrasts of (on the x-axes) fish (LRR_F) with those of (on the y-axes) rearing environment (LRR_E), genotype (LRR_G), and genotype by environment interaction (LRR_{GxE}). All ecosystem metrics from Table S1 are included. Filled circles indicate a significant fish (F) effect, and large circles indicate a significant effect of the contrast on the y-axis. Significance testing was done as indicated in (Table S1).

Figure S3, related to Figure 3. Change in the density (#/L) of different zooplankton species over the three sampling weeks (Weeks 1, 3, and 12 correspond to May 9, May 25, and July 29, respectively) during phase 1. Significant effects of fish (F), environment (E), genetic (G), or genetic by environment interactions (GxE) are labeled accordingly (Table S1). For clarity, E and G effects are not labeled when GxE effects are significant. All error bars are +/- one standard error.

Table S1, related to Figure 1, 2, S2, and S3. Summary of effect sizes for every ecosystem metric measured in the experiment. Log response ratios (i.e. effects sizes) were calculated as described in the text, and significance was determined either by non-parametric tests (ORM: ordinal regression models) or by generalized linear mixed effects models (glme). Uncorrected p-values are reported, and those <0.05 are in bold. All p-values <0.007 are deemed significant when controlling for false discovery rates using the Benjamini & Hochberg correction [S1].

See accompanying excel table

Table S2, related to Table S1. Results of multivariate analysis for the last sampling date of phase 1. Contrasts are as defined in the main text: F is fish presence/absence, G is genetic background of fish, E is Rearing environment, and GxE is the interaction of GxE. Values with $p<0.10$ are highlighted in bold. The Multivariate Effect Size (MVES) was calculated as the Euclidean distance between centroids from a non-metric dimensional scaling analysis (NMDS).

Parameter	Contrasts	F-value (RDA)	p-value	MVES (NMDS)
Macrozooplankton	G	$F_{1,21}=4.26$	**<0.01**	0.21
	E	$F_{1,21}=1.23$	0.32	0.11
	GxE	$F_{1,21}=0.51$	0.62	0.07
	F	$F_{1,37}=2.548$	**0.08**	0.19
Benthos	G	$F_{1,21}=1.19$	0.27	0.09
	E	$F_{1,21}=0.78$	0.61	0.09
	GxE	$F_{1,21}=0.68$	0.7	0.06
	F	$F_{1,37}=3.09$	**<0.001**	0.3
Phytoplankton	G	$F_{1,21}=1.31$	0.26	0.2
	E	$F_{1,21}=0.81$	0.49	0.14
	GxE	$F_{1,21}=1.68$	0.16	0.22
	F	$F_{1,37}=6.87$	**<0.001**	0.45
Microzooplankton	G	$F_{1,21}=1.25$	0.31	0.15
	E	$F_{1,21}=0.63$	0.6	0.1
	GxE	$F_{1,21}=0.74$	0.54	0.11
	F	$F_{1,37}=28.22$	**<0.001**	0.8
Macrophytes	G	$F_{1,21}=0.65$	0.62	0.17
	E	$F_{1,21}=1.95$	0.13	0.23
	GxE	$F_{1,21}=1.5$	0.23	0.2
	F	$F_{1,37}=2.13$	0.1	0.24
Richness	G	$F_{1,21}=0.9$	0.48	0.18
	E	$F_{1,21}=0.42$	0.82	0.14
	GxE	$F_{1,21}=0.65$	0.63	0.21
	F	$F_{1,37}=31.48$	**<0.001**	2.06
Biomass	G	$F_{1,21}=1.11$	0.35	0.21
	E	$F_{1,21}=2.47$	**0.03**	0.17
	GxE	$F_{1,21}=0.46$	0.84	0.4

	F	$F_{1,37}$=1.94	1.94	0.6
Ecosystem Functions	G	$F_{1,21}$=1.39	0.27	0.77
	E	$F_{1,21}$=1.21	0.3	0.66
	GxE	$F_{1,21}$=0.32	0.77	0.4
	F	$F_{1,37}$=1.98	0.13	0.96
Physical/Chemical	G	$F_{1,21}$=0.1	0.99	0.07
	E	$F_{1,21}$=0.31	0.92	0.2
	GxE	$F_{1,21}$=2.16	**0.06**	0.18
	F	$F_{1,37}$=3.43	**0.004**	1.15
Abundance	G	$F_{1,21}$=0.18	0.97	0.05
	E	$F_{1,21}$=1.37	0.24	0.67
	GxE	$F_{1,21}$=1.06	0.39	0.46
	F	$F_{1,37}$=5.59	**<0.001**	1.52

Table S3, related to Figure 4. Statistical analysis of adult fish total biomass and growth over phase 1. Analysis was done using linear mixed effects models, with block as a random factor.

Factor	Total Biomass (Phase 1 start)	Total Biomass (Phase 1 end)	Growth (Length)	Growth (Weight)
Genetic	$F_{1,21}$= 1.61 (0.22)	$F_{1,21}$= 1.30 (0.27)	$F_{1,21}$= 0.02 (0.87)	$F_{1,21}$= 0.19 (0.66)
Environment (Rearing)	$F_{1,21}$= 0.48 (0.49)	$F_{1,21}$= 0.21 (0.65)	$F_{1,21}$= 1.62 (0.22)	$F_{1,21}$= 2.75 (0.22)
Interaction (GxE)	$F_{1,21}$= 0.36 (0.56)	$F_{1,21}$= 0.16 (0.70)	$F_{1,21}$= 1.51 (0.23)	$F_{1,21}$= 0.40 (0.22)

Table S4, related to Figure 4. Statistical analysis of juvenile survival during phase 2 as shown in Figure 4A. Factors codes are as follows: Fish (F), Genetic (G), Rearing Environment (E), Juvenile Ecotype (JE). In all cases, we used the function *lmer* in R with Tank nested within Block (both as random factors), and did the analysis of variance with a Type II Wald Chisquare test.

Response Variable	Factor	χ (df)	p- value
Overall Juvenile Survival (Fish effects)	F	0.47 (1)	0.49
	JE	1.07 (1)	0.30
	FxJE	0.17 (1)	0.68
Overall Juvenile Survival (Treatment effects)	G	5.83 (1)	**0.016**
	E	0.087 (1)	0.768
	JE	0.44 (1)	0.508
	GxE	0.17 (1)	0.680
	GxJE	0.005 (1)	0.941
	ExJE	0.05 (1)	0.825
	GxExJE	0.005 (1)	0.942
Juvenile Lake Survival	G	3.44 (1)	0.063
	E	0.138 (1)	0.71
	GxE	0.138 (1)	0.71
Juvenile Stream Survival	G	4.46 (1)	**0.035**
	E	0.01 (1)	0.92
	GxE	0.09 (1)	0.76

Table S5, related to Figure 4. Statistical analysis of growth accrual, and growth accrual differential during phase 2. Factors codes are as follows: Fish (F), Genetic (G), Rearing Environment (E). An initial length (or length difference) was only included in the model if it lowered the overall model AIC value.

Response Variable	Factor	χ (df)	p- value
Growth Accrual (Figure 4B)			
Entire population (diamonds)	F	9.24 (1)	**0.002**
	Initial length	2.26 (1)	0.13
Stream juveniles (closed circles)	F	6.49 (1)	**0.01**
Lake juveniles (open circles)	F	9.43 (1)	**0.002**
	Initial length	2.26 (1)	**0.01**
Growth Accrual Differential (Figure 4C)	G	0.86 (1)	0.35
	E	2.68 (1)	0.10
	GxE	4.20 (1)	**0.04**

Supplemental Experimental Procedures:

Design and setup of the mesocosm experiment

Our experiment used a single lake-stream ecotype pair (from the Constance region) that has evolved in the past 150 years since the introduction of stickleback into the drainage system [S2]. Lake and stream ecotypes of stickleback have evolved independently in multiple lineages in coastal areas of the northern hemisphere [S3-5] and exhibit a broad range of adaptive phenotypic changes associated with lake and stream habitats [S6]. However, as our design did not include replicate lake-stream pairs, our experimental results cannot be safely extrapolated beyond the studied pair from Lake Constance [S7].

In the spring of 2010, we collected wild lake (L) and stream (S) stickleback from the Constance region (Figure S1), made within-population crosses in the lab (LxL and SxS), and reared the progeny to adulthood on either benthic or pelagic food sources. The following spring, these adult lab-reared fish were used in an outdoor experiment with 40 mesocosm ecosystems (1136 L, ~1.3m max depth). In each mesocosm, we added a mixture of gravel, sand, and soft sediment from a local lake (Lake Luzern), supplemented each mesocosm with nutrients, and added invertebrates and benthic substrate originating from three lakes and their tributary streams.

Phase 1 – The adult fish we used in this phase were individuals from the same lab populations that have been used in two previous studies [S7]. Briefly, in May of 2010, we collected the Lake fish ecotype from Lake Constance at Altenrhein 47°29'08.29"N 09°32'37.85"E and Stream fish ecotype from a small stream at Aubach, Oberriet 47°19'36.80"N 09°34'04.13"E. The two sampled populations are connected by streams and canals but separated by a minimal distance of

30 km from each other. Pairs of males and females were allowed to spawn in aquaria (60x30x40 cm), and then the adults were removed, and full-sib juveniles were reared on a common food sources for 5 weeks, namely brine shrimp (*Atermia sp*.) and small nematodes (*Panagrellus sp*.). After this time, each full-sib family was split into two groups and reared for 11 months on a "pelagic" diet, or a "benthic" diet, so as to establish contrasting rearing environments for both the lake and stream ecotypes. The pelagic diet consisted of live zooplankton (mainly *Daphnia* sp., *Diaptomus* and Cyclopoid copepods) collected multiple times per week from Lake Luzern, using a 170µm zooplankton net. The benthic diet consisted of frozen bloodworms (*Chironomidae spp*. larvae). Rearing lab bred sticklebacks on contrasting food sources from these populations has previously shown that lake fish grow faster than stream fish when reared on plankton [S2]. In addition, both linear morphometrics and body shape show some degree of plasticity, for both lake and stream ecotypes [S7]. Because we used individuals from the same rearing populations as two previous studies [S2,S7], we are confident that our plasticity manipulation created fish with different morphology, body shape, and potentially feeding behaviour. The previous morphological analysis on these fish [S7] is particularly relevant for understanding the phenotypic differentiation of the adults at the start of phase 1. In short, [S7] found that wild lake fish were larger, had shallower and more elongated bodies, deeper heads with longer jaws, and longer gill rakers (same number of rakers) than wild stream fish. Multivariate effect sizes for linear morphology (Mahalanobis distance (MD)=1.167) and shape (MD=1.66) were both significant. When fish were reared on a common food source (i.e. zooplankton or chironomids), lake fish had wider gapes (although there was an interaction between ecotype and rearing environment), more terminal mouths, dorso-caudally shifted maxilla, and shorter heads. In this case, multivariate effect sizes for linear morphology

(Mahalanobis distance (MD)=1.31) and shape (MD=2.201) were also significant. Finally, when comparing fish reared under different food conditions, those raised on chironomids had shorter heads with smaller eyes, deeper bodies, larger orbit, increased suspensorium, and a more downturned snout. Here as well, multivariate effect sizes for both linear morphology (Mahalanobis distance (MD)=1.63) and shape (MD=1.81) were significant. Overall, there was clear phenotypic differentiation between lake and stream fish, both in the wild and after rearing them on a common food source [S7]. This is broadly consistent with other previous studies that have found significant morphological and shape variation between this lake-stream pair [S8-10], two of which show that the amount of divergence in the Constance pair is small relative to other lake-stream pairs in Canada [S10], and other Swiss and Irish pairs [S9].

Phase 2 – The breeding was done in a similar way as in Phase 1, except that the juveniles were all kept on a constant diet (a mix of both plankton, and chironomids) throughout their entire rearing period. We used clutches from 10-12 independent lake and stream families, with several half-sib clutches (i.e. offspring had the same father). Families were raised separately (i.e. full and half sibs were in the same tank) until they were used in the mesocosm experiment. Each mesocosm received individuals from multiple families, and we did not track variation in survival and growth among individuals of different families. Hence, family was not used as a covariate in our analyses.

Design - The first phase of our mesocosm experiment (April 22-Aug 18, 2011) consisted of a factorial manipulation of stickleback genetic background (2 levels: stream [S] and lake ecotypes [L]) and rearing environment (2 levels: benthic [B] and pelagic [P]) and was set up in randomized complete block design, with 40 experimental mesocosms in 8 blocks of 5

mesocosms, where four were intended to hold either only males or only females. The blocks were set up spatially within the grid of mesocosms, and some sampling was done temporally by block. Sex determination of the adults was done following well established methods [S11]. We refer to the treatment combinations as LP, LB, SP, and SB (Figure S1), and a no fish control mesocosm (NF). This design allowed us to measure interactive effects of the genetic background and rearing diet of the lake and stream ecotypes on these aquatic ecosystems (Figure S1), and compare them to the ecosystem effects of fish presence. All of our analyses include block as a random factor.

Phase 1 set up - In April of 2011, we established experimental ecosystems in 1136 L mesocosms (dimensions: 1.5x1.5x1.3m) (Figure S1). We added approximately 40L of a mixture of sand, gravel, and mud to the bottom of each mesocosm to obtain a substrate depth of about 5 cm. We filled tanks with water from Lake Luzern and fertilized them with 2.46g $NaNO_3$ and 0.18 g NaH_2PO_4. After 1 week, we added to each mesocosm zooplankton collected from Lake Luzern (mesh size 65 μm), and lake sediment and benthic substrate (~5 L/tank) collected from multiple lakes (Lake Luzern, Lake Rot, and Lake Zug) and from multiple stream tributaries. We then waited for 2 weeks to allow the ecosystem to develop. This procedure ensured that each mesocosm received substrate, propagules, and a high diversity of living organisms from both benthic and pelagic habitats. Unfortunately the addition of sediment from these particular lakes unintentionally introduced two individual fish larvae into two of the tanks (see below). From April 22-May 4 2011 we added 7 adults sticklebacks to each of the 32 mesocosms, proceeding in 1-2 blocks per day. We chose fish from a narrow range of lengths and randomly assigned them to the tanks such that neither starting fish density (N=7 in all tanks), nor average length (mean

length=44.8mm, SD=4.2), nor biomass (mean weight=0.71g/fish, SD=0.22: mean biomass per tank: 4.95 g, SD=0.68) were significantly different among treatments (linear mixed effect models: all p-values >0.1 for length and biomass). Over the course of phase 1, mean survival rate of the adults was very high (Mean survival rate=90%, SD=14%: 17 tanks with 7/7 fish survived, 10 tanks with 6/7, 3 with 5/7, and 1 with 4/7, and 1 with 3/7) and did not differ among treatments (Friedman rank sum test: χ^2=3.16, df=3, p=0.37). The adults increased in both length (average increase=2.8 mm, SD=2.5) and weight (average increase=0.39 g, SD=0.18), and this growth did not differ among treatments, and so the biomass of adults at the end of phase 1 (mean biomass per tank: 6.92 g, SD=1.87) was not different among treatments (Table S3).

Phase 2 set up - The second phase of our experiment (Sept 13 - Nov 23, 2011) aimed to test for an effect of ecosystem modification in the mesocosms on the survival and growth of juvenile stickleback. At the end of phase 1, we removed all the adults by a combination of trapping and dip netting. This procedure took approximately 3 weeks, because we had to ensure that all tanks had no remaining adults and we did not want to disturb the ecosystem too much by our fishing activity. To start phase 2 we decided to only use a subset of 20 of mesocosms, as opposed to all 40. This was because of the following three reasons. First, we found a loach (*Barbatula barbatula*: 6.5cm, 2.1g) in one mesocosm (Mesocosm 17, Block 4) and a burbot (*Lota lota*: 5.8 cm, 2.4g) in another (Mesocosm 27, Block 6). These benthic fish presumably came into the experiment as larvae in the sediment (probably <1 cm) and avoided our detection throughout phase 1, probably by burrowing and foraging at the sediment interface. No additional individuals of either species were found in any of tanks at the end of phase 2 despite a thorough search of the sediment when the mesocosms were drained and emptied. Second, despite trying to keep the

sexes of the adult sticklebacks separate during phase 1 (4 blocks of females, and 4 blocks of males) we had two bouts of reproduction in two separate tanks (both with males). At the end of phase 1, we discovered 25 juveniles (< 1 week old) in each of tanks 25 (Block 6) and 35 (Block 8). Third, despite the high adult stickleback survival rate across the entire experiment the tanks with the lowest survival (5 tanks with a <6/7 survival rate) were spread across four blocks (Blocks =2, 4, 6, and 8). For these three reasons we decided to only continue with the four blocks (Blocks 1, 3, 5, and 7) for phase 2 of the experiment (all female tanks). Hence we compared ecosystem metrics between phases (Figure 3) as well as the survival and growth analyses (Figure 4) for only 20 tanks (4 blocks of 5 tanks).

Analysis of Ecosystem metrics (phase 1 and 2)

Light extinction – Transmission of photosynthetic available radiation (PAR: 400-700nm mmols/s m^2) was measured on cloudless sunny days using a 4π light sensor (Licor LI-193), and recorded on a LI-1000 Data logger (Licor Biosciences) at 6 different depths in each mesocosm: at 1, 10, 20, 30, 40, and 50 cm. A light extinction parameter (k) was calculated as the slope of the relationship between depth and ln(PAR).

Filtration- To sample water for analysis of algal biomass (chlorophyll-a), nutrients, and dissolved organic carbon, we filled a 1 L glass bottle with water from a central location in every mesocosm from a depth of about 30 cm. Samples were kept in the dark until filtration (within 3 hours of collection), whereupon we filtered between 500-1000 mL of water through 47mm ashed GF/Fs (6 hours at 450°C).

Nutrients - The filtrate was used to measure dissolved nutrients (NO_3^-, and PO_4^-). Nitrate (NO_3-) was measured with a Metrohm 761 Compact IC (Metrohm Schweiz AG Zofingen) with a detection detection limit of 0.25 mg/L (±0.1). Phosphate (PO_4-) was measured using a spectrophotometer Cary 50 (Varian) with a detection limit of 1.0 mg/L (±0.5). DOC was measured using total organic carbon analyzer Shimadzu TOC-V (Shimadzu Coporation 2011), with a detection limit of 0.5 mg/L (±0.5).

Algal biomass – Chlorophyll-a (μg/L) was extracted from the filters by immersing them in 8 mL of 90% ethanol, putting the vials in a 78°C water bath (10 min), and sonicating them (10 min). The extraction solution was then stored at 4°C until analysis on the subsequent day. Just prior to analysis, the solution was filtered with a FP 030/30 Rotband filter (to remove any filter particulates from the extraction procedure), and absorption was measured at 665nm using a Hitachi U 2000 spectrophotometer.

Periphyton biomass - The growth of periphyton (mg/cm^2) was measured by quantifying the chlorophyll-a concentration from plastic strips that were placed in the center of every mesocom tank. We used a commercial 33m x 50mm PVC yellow flagging tape as substrate for the benthic algae to grow on. The tape was anchored to the mesocosm's crossbeam and submerged (with weights) to about 50cm below the water surface. At sampling, we cut a strip of tape (Surface: 15 cm^2) and extracted cholorophyll from it as described above.

Bacteria abundance and richness - Sterile glass bottles were used to collect between ~150 mL of water from mesocosm tanks. To measure bacterial abundance, we used flow cytometry on raw water that was fixed in a final concentration of 0.01% paraformaldehyde and 0.1% glutaraldehyde. It was stored for 4-8 months at 4 °C until analysis. Bacterial cells were then stained with 10 μL mL−1 SYBR Green (Invitrogen) and analyzed for flow cytometry [S12].

To quantify bacteria richness, bacteria was filtered through a 0.2 μm Supor® membrane filter (Pall Scientific). Filters were immediately placed in liquid nitrogen and kept in a -80 °C freezer until DNA extraction. Prior to DNA extraction, sterile forceps and scissors were used to cut up the filters into small pieces. Mechanical lysis was used to extract DNA from filters by a bead-beating step with 2 glass beads (2 mm) for 35 seconds at 5 ms^{-1} on a FastPrep®-24 instrument in 1.2 ml STE buffer (10 mM Tris hydrochloride [pH 8], 1 mM EDTA, 100 mM NaCI). After lysis, we added 50 μL of 20 % sodium dodecyl sulfate (SDS), and then boiled the extract for 2 min. We transferred the liquid extract to new tubes with 400 μL phenol (pH 8) and 400 μL 24:1 chloroform isoamyl alcohol (CIA), followed by a vortexing step for 15 s. The extract was centrifuged for 5 min at 13200 rpm, followed by the repetition of the phenol/chloroform step. 100 μL volumes of 3M sodium acetate (pH 5.2) and 600 μL of isopropanol was added with the extract prior to precipitation on ice for approximately 2 hours. After precipitation, the samples were centrifuged at 13200 rpm for 30 min at 4°C. an aspirator was used to remove the supernatant, the pellet was washed with 70% ethanol and re-suspended in TE buffer. Quant-iT Picogreen (Invitrogen, Oregon, USA) was used to quantify DNA concentration following manufacturer's protocols.

PCR reactions in volumes of 25 μL with 5ng of template DNA were made with the bacterial specific primers: 1406F fluorescently labeled with 6FAM on the 5' end and 23Sr. These were used to target the intergenic spacer region between 16S and 23S rRNA for ARISA analysis. A mixture of 1 mL aliquot of PCR product, 9 mL highly deionized (HiDi) formamide, and 0.5 mL Liz1200 size standard (Applied Biosystems) was denatured on a PCR thermocycler for 3 min at 95 °C. Denaturing capillary electrophoresis of each fragment was run on a 3130XL Capillary Genetic Analyzer (Applied Biosystems) with a 50 cm capillary using POP-7 polymer. Southern size-calling method was used on ARISA fragments between 200 bp and 1250 bp, and a background cut-off level of 50 fluorescence units was employed. A binning window size of 2 bp was used to bin AIRSA peaks with the automatic and interactive binning R scripts [S13]. We exported the relative fluorescence intensity of binned peaks data for further analysis, and excluded any peaks present in blank extractions and amplifications. We operationally defined as the number of peaks detected in ARISA profiles as bacterial richness (R).

Ecosystem metabolism - Oxygen concentration (mg/L), temperature (°C), and light (PAR measured on a 2π sensor, mmols/s m^2) were measured simultaneously for time periods of at least 24 hrs (up to 96 hours) in order to calculate gross primary productivity (GPP), net primary productivity (NPP) and respiration (R) for every mesocosm ecosystem. We used temperature sensors (Hobo Tidbit) to record temperature on 15-minute intervals, and an OXY-10 mini (PreSens) equipped with 10 m long fiber optic dipping probes (PreSens) to make oxygen measurements on compatible intervals. Because we only had 10 channels per on the OXY-10 mini, we spilt the 40 mesocosms into 4 sampling blocks of ten tanks, where each sampling block consisted of two of our 8 experimental blocks (each with 5 treatment combinations). From the

simultaneous measurements of water temperature, incident radiation (PAR), and dissolved oxygen (DO) over a minimum of 24 hours, we calculated gross primary production (GPP), net primary production (NPP), and ecosystem respiration (R). We calculated NPP as $[DO]_{t1-t0}$, R as $[DO]_{t1-t2,}$ and GPP as NPP+R, where t_0=sunrise on first day, t_1=sunset on the first day, and t_2=sunrise on the following day [S14].

Sedimentation – Sedimentation rate (mg/day) was measured using a 50 mL falcon tube that was weighted with a stainless steel nut, and suspended (from a wooden cross beam) in the middle of each mesocosmat a depth of 50 cm. The tubes were deployed for 3 weeks near the end of phase 1, and upon removal from the mesocosm we removed the sediment content of the tube and dried (at 50°C for 24 hrs) and weighed (analytical balance WBA-620) it on an aluminum weigh boat.

Zooplankton sampling - We sampled 9.5L of water using a tennis ball sampler (2 m long plexiglass tube, with a 5cm width) from 6 different locations in the mesocosm(all from the open water column), filtered it through a 100 μm Nitex mesh, suspended the zooplankton in approximately 35 ml tap water and preserved it in a 50ml Falcon tube at -20°C. The zooplankton community was identified to the lowest taxonomic level possible, counted and measured using a Leica DMI 6000B inverted microscope. The body length of either all or at a minimum 30 individuals per taxon was measured using the Leica microscope system at 1x to 100x magnification and the software LSA 3.7. All individuals were used for the calculation of abundance and species richness for each sampling date, but in order to compare similar taxonomic groupings among time periods, we aggregated organisms into the following taxonomic groupings: Calanoid copepods, *Bosmina*, *Daphnia*, Cyclopoids, Nauplii, Ostracods,

Small Cladocerans (*Alona*, *Eurycercus*, *Pleuroxus*), and Rotifers (*Asplanchna*, *Keratella*, *Trichocerca*, *Astramoeba*). Count data (i.e. number of individuals per sample of 9.5 L) was analyzed in order to compare with species from other communities in a common statistical analysis for the last sampling date of phase 1. In addition, zooplankton densities and biomass were calculated and subsequently analyzed using generalized linear mixed effect models (Figure 1B, Figure 2, 3). Macrozooplankton biomass was calculated from published length-weight regressions [S15].

Macrozoobenthos sampling – At the end of phase 1, we used a hand net (mesh size of 0.2 mm) to collect ~1520 cm^3 of sediment from the top 0.5 cm of the benthic substrate in each tank, and stored the samples at -20°C. To quantify macrozoobenthos, each sample was thawed and individually washed through a stack of stainless steel mesh filters with a size range of 0.25-6.3 mm. All individuals retained on the sieve were then identified to the taxonomic level shown in Table S1. We measured the size of individuals in taxonomic grouping with >30 individuals across all the tanks, and used published length-weight regressions to estimate biomass [S16].

Visual surveys –At the end of each phase, visual surveys of *Dreissena* mussels were counted from the upper half of all 4 sides of the mesocosm walls (2 m^2). We used the upper half because this was the region where variation in water clarity had no effect on our ability to see *Dreissena*, and because of how the mesocosms were constructed there was a clear demarcation at this level. A single observer counted all tanks within an 8-hour period (30-60 min/block).

Statistical analysis of ecosystem variables – We used generalized linear mixed effects (glme) models with a randomized block design to analyze all ecosystem metrics with a continuous data type, including abundance, density, richness, biomass, chemical, and physical parameters (Table S1). We made no mathematical corrections of p-values for multiple comparisons, however we indicate in Table S1 which p-values remain significant after controlling for false discovery rates using the Benjamini & Hochberg correction. All metrics based on count data were analyzed with ordered probit ordinal regression [S17], which assumes the response variable is ordinal. Such a semi-parametric analysis has several advances over other related approaches, for example, unlike many other rank related analyses of factorial experiments ordinal regression models allow for robust testing of interaction terms in a complete randomized block design [S17]. In addition, it provides a useful way to compare across multiple response variables that would otherwise need very different distributional assumptions for standard parametric analysis (e.g. glme models).

In order to compare the magnitude of effects across multiple metrics we quantified log response ratios (LRR) as:

$LRR_G = (\ln Y_{SP} + \ln Y_{SB}) - (\ln Y_{LP} + \ln Y_{LB})$

$LRR_E = (\ln Y_{SB} + \ln Y_{LB}) - (\ln Y_{SP} + \ln Y_{LP})$

$LRR_{GxE} = (\ln Y_{SB} + \ln Y_{LP}) - (\ln Y_{SP} + \ln Y_{LB})$

$LRR_F = \ln (Y_{NF}) - (\ln Y_{LB} + \ln Y_{LP} + \ln Y_{SB} + \ln Y_{SP})/4$

where Y is the mean ecosystem metric among mesocosms, after correcting for block means. The significance of F, G, E, and GxE effects were tested with either parametric (generalized linear mixed effect models) or non-parametric tests (ordinal regression).

For the last sampling date in phase 1, we performed a multivariate analysis (Redundancy analysis: RDA) of multiple groups of ecosystem metrics and communities, as listed in Table S1. For each community and grouping of ecosystem metrics, we performed an RDA on data that was either hellinger transformed (Macrozooplankton, Microzooplankton, Zoobenthos, Phytoplankton, and Macorphyte communities), log-transformed (Richness, Abundance), or standardized (Biomass, Ecosystem Function, Physical/Chemical). In all cases, significance testing of the fish effects and treatment effects were done via permutation (N=1000) within each block [S18]. To estimate multivariate effect sizes of each grouping, we performed non-metric multidimensional scaling (on transformed data) and calculated Euclidean distances between treatment centroids (using both NMDS axes) (Table S2).

Quantifying survival and growth rate of juvenile populations (phase 2)

Juvenile rearing and selection - Phase 2 began (Sept 13, 2011) with the addition to each mesocosm of 32 juvenile stickleback (16 stream, and 16 lake), which had been reared (June – September 2011) on a common food source, consisting of brine shrimp, nematodes for 6 weeks, and both plankton and chironomids for 10 weeks. These juveniles were reared from matings in the lab between adult Lake and Stream ecotypes that were collected in May 2011 from Lake Constance at Altenrhein (47°29'08.29"N 09°32'37.85"E) and Steinach (47°30'10.85"N 09°26'49.82"E), and in a small stream at Aubach, Oberriet (47°19'36.80"N 09°34'04.13"E). All the lake and stream juveniles were reared from about 25 successful matings in aquaria (60x30x40 cm), wherein we provided a petridish with sand, mud, and textile fibers. After mating, the eggs were removed from the nest, put in separate aquaria equipped with air pumps.

We reared the juveniles in the lab from June to September 2011. In order to minimize family effects we reared juveniles of different clutches in 25 different aquariums (11 Lake and 14 Stream families), and when selecting juveniles for phase 2, we randomly selected 16 juveniles from 8 different Lake families and 16 juveniles from 6 different Stream families, ensuring to get 1 to 4 individuals per family per mesocosm. Overall we used 1024 juveniles, 512 of each ecotype.

Juvenile staining - In order to distinguish lake and stream juveniles at the end of phase 2, we stained either the lake or stream population intended for each mesocosmwith a Calcein dye ($C_{30}H_{26}N_2O_{13}$, CAS number 1461-15-0, Sigma # C0875), which fluoresces under UV light with excitation and emission wavelengths of 495/515 nm, respectively. We used a pH buffered solution of 125 mg/L of Calcein (Sigma, St. Louis, MO) and incubation time of 2 hours [S19]. We randomized whether the lake or stream fish were stained in each mesocosm.

After the staining of fish with calcein, every juvenile was photographed in a cuvette with a scale bar and a color reference. Standard length was measured using the photographs and the software ImageJ [S20]. Fish were transported to mesocosms in small plastic fish tanks, and acclimated to the new environment by slowly mixing water from the mesocosm with water from the lab tanks.

Juvenile collection and identification - At the end of phase 2, we collected juveniles by trapping and hand netting, and then froze them in individual tubes at -20°C. Each individual was observed under UV light using a Leica DMI6000 B inverted microscope. Every individual was screened for florescence at 6 different bony structures of the body. To minimize errors arising from weak

auto-florescence of some body parts, we inspected florescence of (i) bony scales on the head, (ii) pectoral spine, (iii) pectoral girdle, (iv) dorsal spine, (v) vertebrae, and (vi) teeth. A single observer did all the screening and the determination of whether an individual was stained or not. Standard length of each individual was measured with callipers prior to dissection.

Statistical analysis of juvenile growth –

At the start of phase 1, stream juveniles were slightly larger than lake juveniles (Lake juveniles: mean=19.2, SD=3.5, N=320; Stream juveniles: mean=20.6, SD=3.2, N=320), but were a similar length by the end of phase 2 (Lake juveniles: mean=25.8, SD=3.9, N=255; Stream juveniles: mean=26.0, SD=3.6, N=241). The entire juvenile population (lake and stream individuals) experienced positive growth (*g*) over the course of phase 2 in all mesocosms (mean growth accrual: *g*=5.14 mm, SD=1.5), suggesting that mesocosm conditions were amenable to test for variation in performance over the time period of the experiment. Consistent with previous laboratory experiments using the same population pair [S2], we found that lake fish grew consistently faster (higher GA) than stream fish (Figure 4B, Table S5). In addition, average growth accrual of the entire population was significantly higher in tanks that previously did not have adult fish (Figure 4B, Fish effect in Table S5). This provides additional support for a transgenerational effect of fish on juvenile growth, but provides no direct evidence of a feedback, because the effect is mediated by fish density rather than phenotype. In order to test for an eco-evolutionary feedback (prediction 4), we focused our analyses on the survival and growth accrual of co-occurring juveniles. We focused on growth accrual rather than final body length because there was a slight difference in mean body length of the two ecotypes at the start of phase 2 (Lake juveniles: mean=19.2, SD=3.5, N=320; Stream juveniles: mean=20.6, SD=3.2, N=320, $t=5.37$, $p<0.001$). We used linear mixed effects models (*lmer* package in R, Type II Wald

Chisquare test) with Tank nested within Block (both as random factors), to analyse the variation in juvenile length and growth accrual over the course of phase 2 (N=20 tanks). In all models, either an initial length (for growth accrual) or an initial length difference (for growth accrual differential) was included as a random covariate in the analysis, but it was only retained if the model had a lower AIC than the model without the covariate (Table S5).

Supplemental References

[S1] Benjamini Y, Hochberg Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B-Methodological 57:289–300.

[S2] Lucek, K., Sivasundar, A., and Seehausen, O. (2012). Evidence of adaptive evolutionary divergence during biological invasion. PLoS ONE *7*, e49377.

[S3] Hendry AP, Taylor EB, McPhail JD. (2002). Adaptive divergence and the balance between selection and gene flow: Lake and stream stickleback in the misty system. Evolution 56:1199–216.

[S4] McKinnon JS, Rundle HD. (2002). Speciation in nature: the threespine stickleback model systems. Trends in Ecology & Evolution 17: 480–8.

[S5] Berner D, Adams DC, Grandchamp AC, Hendry AP. (2008). Natural selection drives patterns of lake-stream divergence in stickleback foraging morphology. J Evol Biol 21:1653–65.

[S6] Kaeuffer R, Peichel CL, Bolnick DI, Hendry AP. (2012). Parallel and Nonparallel Aspects of Ecological, Phenotypic, and Genetic Divergence Across Replicate Population Pairs of Lake and Stream Stickleback. Evolution 66:402–18.

[S7] Lucek, K., Sivasundar, A., and Seehausen, O. (2014). Disentangling the role of phenotypic plasticity and genetic divergence in contemporary ecotype formation during a biological invasion. Evolution 68, 2619–2632.

[S8] Roy, D., Lucek, K., Bühler, E., and Seehausen, O. (2010). Correlating shape variation with feeding performance to test for adaptive divergence in recently invading stickleback populations from Swiss peri-alpine environments. In Morphometrics for Nonmorphometricians (Springer Berlin Heidelberg), pp. 233–257.

[S9] Lucek, K., Sivasundar, A., Roy, D., and Seehausen, O. (2013). Repeated and predictable patterns of ecotypic differentiation during a biological invasion: lake–stream divergence in parapatric Swiss stickleback. J. Evol. Biol. *26*, 2691–2709.

[S10] Berner, D., Roesti, M., Hendry, A. P., and Salzburger, W. (2010). Constraints on speciation suggested by comparing lake-stream stickleback divergence across two continents. Mol. Ecol. *19*, 4963–4978.

[S11] Peichel, C. L., Ross, J. A., Matson, C. K., Dickson, M., Grimwood, J., Schmutz, J., Myers, R. M., Mori, S., Schluter, D., and Kingsley, D. M. (2004). The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. Curr. Biol. *14*, 1416–1424.

[S12] Prest, E. I., Hammes, F., Kötzsch, S., van Loosdrecht, M. C. M., and Vrouwenvelder, J.

S. (2013). Monitoring microbiological changes in drinking water systems using a fast and reproducible flow cytometric method. Water Research *47*, 7131–7142.

[S13] Ramette, A. (2009). Quantitative community fingerprinting methods for estimating the abundance of operational taxonomic units in natural microbial communities. Appl. Environ. Microbiol. *75*, 2495–2505.

[S14] Harmon, L. J., Matthews, B., Roches, Des, S., Chase, J. M., Shurin, J. B., and Schluter, D. (2009). Evolutionary diversification in stickleback affects ecosystem functioning. Nature *458*, 1167–1170.

[S15] Culver, D. A., Boucherle, M. M., Bean, D. J., and Fletcher, J. W. (1985). Biomass of freshwater crustacean zooplankton from length–weight regressions. Can. J. Fish. Aquat. Sci. *42*, 1380–1390.

[S16] Greig, H. (2008). Community assembly and food web interactions across pond permanence gradients. PhD Thesis. 1–142.

[S17] Harrell FE. Harrell, F. E. (2001). Regression modeling strategies (Springer Science & Business Media).

[S18] Pantel, J. H., Duvivier, C., and Meester, L. D. (2015). Rapid local adaptation mediates zooplankton community assembly in experimental mesocosms. Ecol. Lett. *18*, 992–1000.

[S19] Wilson, C. A., Beckman, D. W., and Dean, J. M. (1987). Calcein as a fluorescent marker of otoliths of larval and juvenile fish. Transactions of the American Fisheries Society *116*, 668–670.

[S20] Abramoff, M. D., Magalhaes, P. J., and Sunanda, J. R. (2004). Image processing with ImageJ. Biophotonics International, 1–7.