SI Appendix

An integrated field, laboratory and theoretical study of PKD spread in a Swiss prealpine river

Luca Carraro, Enrico Bertuzzo, Lorenzo Mari, Inês Fontes, Hanna Hartikainen, Nicole Strepparava, Heike Schmidt-Posthaus, Thomas Wahli, Jukka Jokela, Marino Gatto, Andrea Rinaldo

Supporting materials and methods

Study area and hydro-geomorphology

The river Wigger is highly anthropized, with scarce restoration areas, regular channel widths and slopes engendered by a system of check dams along the main courses which does not hinder fish movement.

The catchment's morphology was obtained from a 25-m resolution digital terrain model. The river network was extracted using the Taudem routine in a GIS software [1]. Flow directions were determined by following steepest descent paths (D8 algorithm) [1]. Pixels belonging to the channeled portion of the landscape were those whose drained area was greater than or equal to 0.5 km^2 whereas details on slope- of curvature-dependent area thresholds prove immaterial at the scale of interest [2]. The extracted fluvial network was compared with the vectorial map provided by the Swiss Federal Office for the Environment (FOEN) [3], and the overall match proved satisfactory. In order to subdivide the catchment into sub-units where reasonably constant local morphological conditions apply, i.e. nodes of the metacommunity scheme, stream reaches longer than 5 km were split into equally long stretches. A subcatchment was defined by the direct contributing area drained by a single stream reach. Overall, the river network was divided into 166 subcatchments hierarchically arranged according to the network connectivity. The geological characterization of the catchment was obtained by a vectorized geological map of Switzerland provided by the Swiss Federal Office of Topography (Swisstopo) [4]. The map was processed on a GIS software and the initial 16 geological classes were grouped into five main classes, as displayed in Fig. 1e

of the main text. Note that the 'superficial water' class only refers to ponds or small lakes, while all pixels covering the river network proper were conventionally classified as 'alluvial rocks'.

Owing to the relatively small extension of the catchment, it is reasonable to assume that the daily discharge at a station is proportional to the total contributing area therein [2]. This hypothesis was tested for the two discharge series available, namely measured by FOEN in Zofingen (corresponding to site #3 of Fig. 1a of the main text - drained area A = 366 km^2) and Nebikon (site #7 - A = 105 km^2). Comparing observed and modelled daily discharges in Zofingen for the period 2010-2015 (where the latter is obtained by multiplying the daily Nebikon discharge by $366/105 \approx 3.5$), we achieved a satisfactory fitting (Nash-Sutcliffe coefficient NS = 0.89).

Average river width for each stretch was estimated from Swisstopo aerial images and checked against at-a-station scaling relationships linking landscape-forming discharges to total contributing area A, which yield for consistency the scaling of widths and depths of the cross section to A [5, 6]. Cross-sections are assumed to be rectangular and endowed with large width-to-depth ratios. Given the assumptions and the scales of interest, it is customary to hypothesize the maintenance of uniform flow conditions for individual river stretches. As such, a Manning's roughness coefficient n = 0.033 m$^{-1/3}$s was assumed for the whole river network; according to the U.S. Federal Highway Authority [7], this value corresponds to a fine gravel bedrock, with low vegetation and negligible cross-sectional variation, surface irregularities obstructions and sinuosity. Such features were qualitatively observed during *in-situ* campaigns. Through a systematic use of Manning's equation of uniform flow resistance, time series of water depths for all stretches were computed and used by the metacommunity model.

Field data collection

eDNA collection. Stream water samples were collected in 15 different locations along the river network. For each site, 21 500-mL samples were taken at approximately bi-weekly (or monthly during December, January and February) intervals (except site #5, which was abandoned after 12 samples).

Pre-sterilized plastic bottles (with 10% bleach followed by UV-B treatment) were used to collect water from the river by submerging the bottle with a gloved hand. The samples were transported to the laboratory on ice and filtered within the same day on to 5-cm diameter, 0.45-μm pore size individually packaged sterile membrane filters (Merck Millipore). A vacuum pump with a borosilicate glass filtration setup was used and sterilized in 10% bleach between each sample. Negative controls were created by filtering MilliQ water through a sterile filter at the start and end of the filtration session, as well as once during the filtration (after sample 7). Filter papers were placed in 2-mL bead beating tubes (obtained from the kit described below) and frozen at 80 ºC until extraction. Filter papers were cut with sterilized scissors to break them up and eDNA was extracted from all filter papers, including controls, using a PowerSoil®

Target DNA	Primers and probe	Sequences (5′ - 3′)	Length [bp]	GC [%]	Calc. T_m [°C]	Calc. T_a [°C]	Assay T_m [°C]	Optimal Conc.
Fredericella sultana	Fs_16S_F1q	CATTGAGCTTC-GGGAATGTT	20	45.0	54.4	49.4	60	300
	Fs_16S_R1q	ATGAAACCTCG-TCCCTTGTG	20	50.0	56.3	49.4	60	900
	Fs_Probe_16S_1	Cy5-GGGGTCAG-GTTGCTAAGC-CATGA-BHQ-2	23	56.5	62.9	-	60	200
Internal Positive Control	MIMf	GTATTCCTGGTTT-GTAGGTTGAGC	25	48.0	59.0	52.8	60	50
	MIMr	ATGTGACTGGAC-TTCCGTATCG	22	50.0	57.8	52.8	60	900
	IPC_probe	Cy3-CGACGGCC-AGTGAATTGTA-ATACGA-BHQ-2	25	48.0	59.9	-	60	250

Table S1: Oligonucleotide primers and probe sequences for the qPCR assay. Calculated melting (T_m) and annealing (T_a) temperatures are at 0.5 μM for primers and at 0.25 μM for probes. Reporter and quencher dyes for each probe are presented in the sequences: Cy®3 = cyanine 3; Cy®5 = cyanine 5; BHQ-1 = Black Hole Quencher®-1; and BHQ-2 = Black Hole Quencher®-2. The optimal primer and probe concentrations (Optimal conc) in multiplex reactions are reported in nM.

DNA kit (MO BIO Laboratories) in a dedicated clean laboratory (free of PCR products). The kit includes a bead beating step and a separate inhibitor removal step. The eDNA was eluted in 60 μL of Solution C6 and subsequently preserved at -20 °C. eDNA samples were only removed from the -20-°C freezer for screens and remained at room temperature for a maximum of 2 h.

qPCR assay. Specific primers for bryozoan detection (Fs_16S_F1q and Fs_16S_R1q) were designed based on inspection of 16S mitochondrial sequences of all major clades of phylactolaemates. The primers were designed to amplify fredericellid phylactolaemates, with 100% identity in primer and probe sequence with the most common host in Europe, *F. sultana*. The size of the *F. sultana* PCR fragment was 71 bp. A custom internal positive control (IPC) template was spiked into all reactions and amplified in multiplex with the *F. sultana* probe assay. The inclusion of the IPC allowed the detection of possible PCR inhibition. The IPC primers and probe are reported in Table S1.

The 71-bp fragment of *F. sultana* was amplified from genomic DNA samples derived field collected colonies. PCR reaction composition and cycling conditions were the same as those used to test primer pairs. The amplification was performed in 50-μL reactions and 5 reactions were pooled prior to gel purification using the QIAquick Gel Extraction Kit (QIAGEN). The concentration of the pooled purified products and the IPC template stock solution (100 mM) were measured using a Qubit® 2.0 Fluorometer double-stranded DNA (dsDNA) high-sensitivity Assay (in ng/μL; Invitrogen) and adjusted to 1 nM. A 1:10 serial dilution of the standardized 1-nM solution of each of the two fragments was performed and pooled. Seven standards of the resulting serial dilution were included in each run. A master standard curve for each target obtained from a single multiplex experiment (*F. sultana* and IPC) was applied to all

subsequent experiments using the first standard point as a calibrator (i.e. $8.33 \cdot 10^{-12}$ mol/L per PCR reaction).

PCR amplification and target quantification were performed using a LightCycler® 480 II (Roche Diagnostics) with color compensation. The detection format, analysis mode and color compensation data were the same across all experiments. The optimal forward primer concentrations are reported in Table S1. The amplification was performed in a final volume of 10 μL containing 7.5 μL of master mix with probes and 2.5 μL of template DNA. Each PCR reaction contained 1x LightCycler® 480 Probes Master (Roche Diagnostics) and 1x optimized primer-probe mix. Two master mixes were prepared per experiment: one without IPC to generate the standard curve and test the non-template control (NTC); and a second one with IPC for the quantification of the sample and IPC calibrator reactions, which contained 1 μL of $1.00 \cdot 10^{-13}$ mol/L of IPC. The reaction volumes of both master mixes were made up to 10 μL with PCR-grade water (Sigma-Aldrich). PCR amplification was performed in triplicates with an initial DNA polymerase activation/denaturation step of 95 °C for 10 minutes (4.8 °C/s ramp rate) followed by 45 cycles of denaturation at 95 °C for 10 s (4.8 °C/s ramp rate) and annealing/extension at 60 °C for 1 minute (2.5 °C/s ramp rate), followed by a cooling step of 40 °C for 10 s (2.0 °C/s ramp rate). PCR amplification took a total of 1.5 h. Fluorescence was acquired once per cycle at the end of the annealing/extension phase. A QIAgility pipetting robot (QIAGEN) was used to prepare and dispense the master mix and template DNA into 384-well white qPCR microplates (TreffLab) in a maximum of 2 h to avoid evaporation. All samples used in this study were kept at -20 °C and defrosted in a fridge at 4 °C for 30 minutes before mixing and pipetting into qPCR plates. At the end of each experiment, an absolute quantification of samples was performed using the second derivative maximum method using a high confidence algorithm on the LightCycler® software.

The Limit of Detection (LOD, see Table S2) of the assay was 37.097 for *F. sultana*, defined as the highest C_q mean observed for a truly positive sample with all triplicates fluorescing. Limit of Quantification (LOQ) was set at C_q value where target concentration no longer exhibited linear relationship with C_q readings. The LOQ of the assay was 34.977 for *F. sultana*. Reactions were considered to be positive when C_q value was not greater than 34.98 for *F. sultana*, corresponding to 0.78 template DNA copies. NTCs either did not generate fluorescence, this was higher than the LOQ or it was not present in all triplicates. Experimental samples were considered positive if there was fluorescence in all three triplicates and the C_q mean was not greater than the LOQ.

Fish sampling and prevalence assessment. Electrofishing was performed by FORNAT AG (Forschung für Naturschutz und Naturnutzung, Zurich) with an EFKO FEG8000 machine (Power: 8 kW, number of machine 130708). Two runs of electrofishing per date and stretch were performed. For density assessment purpose, caught fish were kept in tanks supplied with oxygen and fish were set back into the river after

Target		C_q	Calculated conc. [mol/L]	Q [mol]	MW [ng]	Copy number
Fredericella sultana	LOQ	34.977 (0.191)	$8.68 \cdot 10^{-19}$ $(1.30 \cdot 10^{-19})$	$8.68 \cdot 10^{-24}$ $(1.30 \cdot 10^{-24})$	$4.07 \cdot 10^{-10}$ $(6.07 \cdot 10^{-11})$	5.229 (0.781)
	LOD	37.097 (0.967)	$4.02 \cdot 10^{-20}$ $(5.74 \cdot 10^{-20})$	$4.02 \cdot 10^{-25}$ $(5.74 \cdot 10^{-25})$	$1.88 \cdot 10^{-11}$ $(2.69 \cdot 10^{-11})$	0.242 (0.346)
IPC	LOQ/LOD	34.930 (0.645)	$9.12 \cdot 10^{-20}$ $(3.80 \cdot 10^{-20})$	$9.12 \cdot 10^{-25}$ $(3.80 \cdot 10^{-25})$	$5.84 \cdot 10^{-11}$ $(2.43 \cdot 10^{-11})$	0.549 (0.229)

Table S2: Limits of quantification (LOQ) and detection (LOD) for each qPCR assay, given as C_q values, calculated molar concentration, molecular weight (MW) and copy number. Results are given as mean, standard deviation is included between parentheses. IPC: internal positive control. The LOQ and LOD for the IPC assay are the same.

the second fishing run. Prior to release fish length was measured. As for PKD prevalence assessment, the trout were euthanized using 3-aminobenzoic acid ethyl ester (MS 222®, Argent Chemical Laboratories) and length was recorded. The kidney was removed and preserved in RNAlater (Sigma Aldrich, Switzerland) for qPCR based assessment of *T. bryosalmonae* infection presence. Fish kidneys were weighted and homogenized in a 1.5-mL tube with a 2-mm diameter steel bead (QIAGEN, Switzerland) with a tissue lyser (QIAGEN, Switzerland) at a shaking frequency of 30 shakes per second for 3 min. DNA was extracted using the Blood & Tissue DNA extraction kit (QIAGEN, Switzerland). Prevalence was assessed by qPCR using the *T. bryosalmonae* specific TaqMan method according to [8]. All reactions were carried out in duplicates. As positive control, a linearized plasmid containing the amplified fragment [9] was used. Non-target control (water) within the qPCR never showed any amplification while internal control was always amplified showing no qPCR inhibition.

Table S3 reports prevalence data; Fig. S1 displays results from the brown trout density assessment. All details on water temperature measurements and model are reported in the section *Water temperatures*.

Site	Age-class	2014		2015		2016	
		Aug 14	Sep 18	Jun 10	Oct 2	Jul 7	Sep 21
#4	YOY	34/34	20/20	4/5	1/1	25/25	25/25
	Adults	-	6/6	-	5/5	-	5/5
#8	YOY	25/25	30/30	26/30	25/25	25/25	25/25
	Adults	-	4/4	-	5/5	-	5/5
#16	YOY	22/25	26/26	5/6	0/25*	9/25	23/25
	Adults	-	5/5	-	5/5	-	5/5

Table S3: Prevalence data. The first number counts the amount of PKD-positive individuals, the second refers to the total number of individuals tested.
* This sampling was conducted in a different location (see main text).

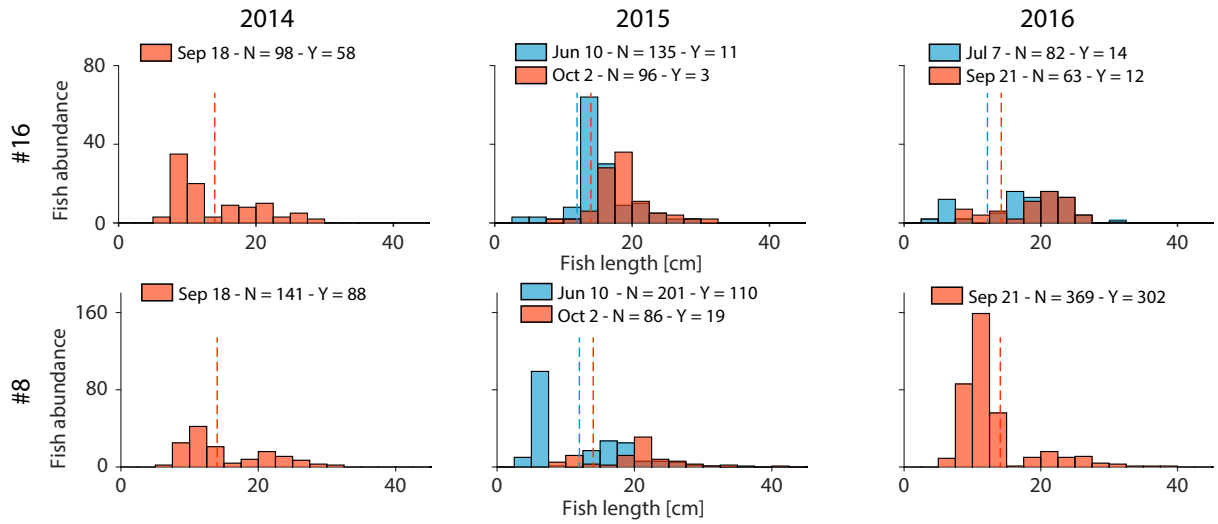

Figure S1: Brown trout population sampling data. Blue indicates early summer sampling, red stands for late summer sampling, maroon colour results from overlay of blue and red bars. N: total fish abundance: Y: YOY abundance. Dashed lines highlight the threshold lengths used to distinguish between YOY and adults: 12 cm in early summer and 14 cm in late summer. Population sampling in site #4 (not shown) was only performed in late 2016 summer.

Bryozoan habitat suitability

The whole river network is divided into a suitable number of stretches N_s, some of which include the sampling sites. Let A_i^L be the direct drainage area of stretch i, namely the subcatchment pertaining to stretch i only, and A_i the upstream contributing area to stretch i, that is the catchment surface upstream of i. A_i obviously includes A_i^L. Let c_i be the average eDNA concentration that would be attained in stretch i in absence of water flow. We assume that c_i is proportional to the density of bryozoan biomass pertaining to site i, and that c_i does not change throughout the season. Moreover, we hypothesize that the eDNA produced from bryozoans in stretch i does not decay until it reaches the exit cross-section of stretch i. We term instead C_i the eDNA concentration measured in stretch i, i.e. when waterflow is considered.

If one assumes that there are no bryozoans upstream of stretch i, which is at hydraulic steady state, it follows that the flow of incoming eDNA (from the sole source, namely stretch i) equals the flow of outgoing eDNA through the exit cross-section of stretch i. Since the contribution to discharge from subcatchment i is proportional to A_i^L, while the outgoing flow is proportional to A_i (as already stated in Section *Study area and hydro-geomorphology*, see *Materials and Methods* in the main text), one can state that $C_i A_i = c_i A_i^L$. Notice that, because A_i^L can be much smaller than A_i (especially for downstream stretches), the measured concentration at site i can be small even though the bryozoan biomass in the subcatchment may be quite large. By considering all eDNA contributions upstream of the sampling site i, and weighting them by their distance L (according to the decay term $\exp(-L/\lambda_B)$), it is possible to derive an expression (Eq. (1)

of the main text) for the measured eDNA concentration C_i.

Extension	Variable	Abbreviation
Local	Elevation	*LocElv*
	Annual mean water temperature	*LocMwt*
	Slope	*LocSlp*
Upstream	Mean Elevation	*UpElv*
	Mean Slope	*UpSlp*
	Contributing Area	*UpCAr*
	Percentage of molasses	*GeoMls*
	Percentage of alluvial rocks	*GeoAll*
	Percentage of moraines	*GeoMrn*
	Percentage of peat	*GeoPea*
	Percentage of superficial water	*GeoWat*

Table S4: List of covariates used for the model of bryozoan habitat suitability. *UpSlp* was computed as the mean slope of all upstream stretches weighted by their respective lengths. Underlined covariates were finally used in the model after multicollinearity test.

Starting from the initial set of eleven covariates of Table S4, we checked for possible multicollinearity effects by computing the variance inflation factors (VIFs) and discarding the predictor with the highest VIF [10]. We repeated this procedure until all remaining covariates had VIFs all lower than 10. The discarded covariates were, in the order: *GeoMls*; *UpSlp*; *GeoPea*; *LocElv*. VIF values for the set of residual explanatory variables (*LocMwt*; *LocSlp*; *UpElv*; *UpCAr*; *GeoAll*; *GeoMrn*; *GeoWat*) are: 5.8; 4.9; 2.4; 5.4; 7.8; 2.3; 1.9. None of the correlation coefficients R between any of these seven variables were above the rule-of-thumb threshold $|R| = 0.8$ indicating strong correlation (Fig. S2). We tested the performance of all models (total number: 127) generated by any subset of the residual explanatory variables. Leave-One-Out-Cross-Validation (LOOCV) [11] was used to evaluate the performance of each model. For a given subset of covariates, a site k at a time was removed from the dataset; a calibration was then performed by maximizing the Nash-Sutcliffe index NS_k

$$NS_k = 1 - \frac{\sum_{j=1, j\neq k}^{N_s} \left(C_j^m - C_j\right)^2}{\sum_{j=1, j\neq k}^{N_s} \left(C_j^m - \overline{C_k^m}\right)^2},$$

where C_j^m is the mean eDNA concentration measured at site j and $\overline{C_k^m} = \frac{1}{N_s - 1} \sum_{j\neq k} C_j^m$. Once calibrated, the overall performance of the model was calculated as

$$s = 1 - \frac{\sum_{k=1}^{N_s} (C_k^m - C_k)^2}{\sum_{k=1}^{N_s} \left(C_k^m - \frac{1}{N_s} \sum_{j=1}^{N_s} C_j^m\right)^2}.$$

We retained all models achieving $s > 0$ (i.e. 16 models). These models were subsequently re-calibrated against the whole set of sites by maximizing the Nash-Sutcliffe index. All calibrations were performed via

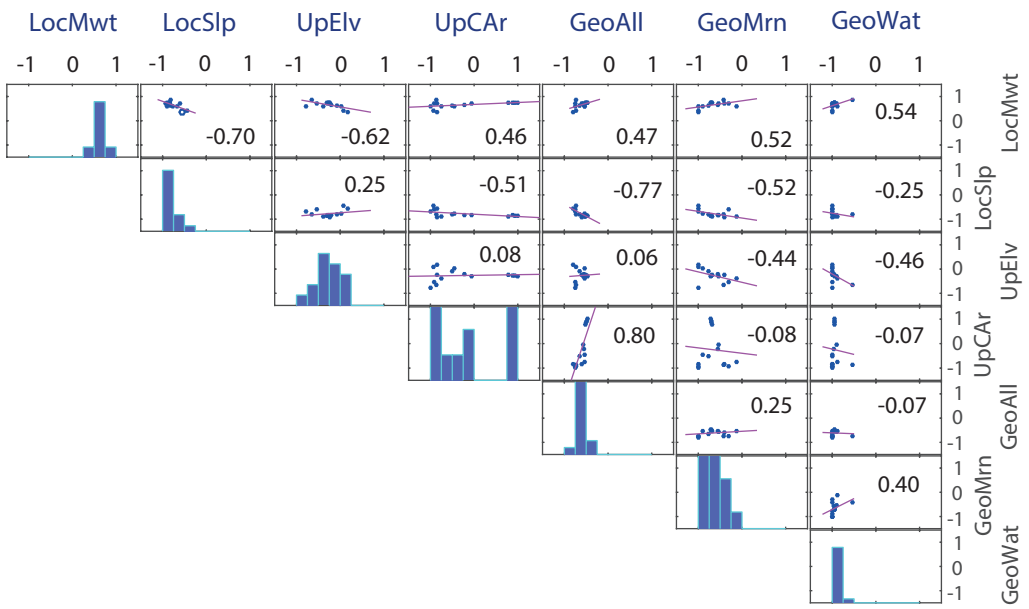

Figure S2: Plot of the correlation matrix of the seven chosen predictors for the model of bryozoan suitability.

a particle swarm optimization algorithm [12].

Epidemiological model

The system of equations for the time-hybrid epidemiological model is a modified version of [13]. The main modification pertains the inclusion of the YOY compartment. A list of state variables is reported in Table S5; Table S6 displays all model parameters and their reference values.

Intra-seasonal model. The set of ordinary differential equations describing the intra-season disease dynamics in stretch i reads:

$$\frac{dB_{S,i}}{d\tau} = r_i \left[1 - \frac{\rho_i}{V_i}(B_{S,i} + B_{C,i} + B_{O,i})\right] B_{S,i} - \beta_B \frac{Z_{F,i}}{V_i} B_{S,i}; \tag{S1a}$$

$$\frac{dB_{C,i}}{d\tau} = r_i \left[1 - \frac{\rho_i}{V_i}(B_{S,i} + B_{C,i} + B_{O,i})\right] B_{C,i} + \beta_B \frac{Z_{F,i}}{V_i} B_{S,i} - d_{CO,i} B_{C,i} + d_{OC,i} B_{O,i}; \tag{S1b}$$

$$\frac{dB_{O,i}}{d\tau} = r_{O,i} \left[1 - \frac{\rho_i}{V_i}(B_{S,i} + B_{C,i} + B_{O,i})\right] B_{O,i} + d_{CO,i} B_{C,i} - d_{OC,i} B_{O,i}; \tag{S1c}$$

$$\frac{dS_{S,i}}{d\tau} = f_B B_{S,i} + \phi f_B B_{C,i}; \tag{S1d}$$

$$\frac{dS_{I,i}}{d\tau} = (1 - \phi) f_B B_{C,i}; \tag{S1e}$$

$$\frac{dY_{S,i}}{d\tau} = -\mu_F Y_{S,i} + \zeta Y_{C,i} - \beta_F \frac{Z_{B,i}}{V_i} Y_{S,i} + \sum_{k=1}^{N} d_{ki} l_k Y_{S,k} - \sum_{k=1}^{N} d_{ik} l_i Y_{S,i}; \tag{S1f}$$

Symbol	Variable
B_S	Biomass of susceptible bryozoans
B_C	Biomass of covertly infected bryozoans
B_O	Biomass of overtly infected bryozoans
S_S	Non-infected statoblast abundance
S_I	Infected statoblast abundance
Y_S	Abundance of susceptible YOY
Y_E	Abundance of exposed YOY
Y_I	Abundance of acutely infected YOY
Y_C	Abundance of carrier YOY
F_S	Abundance of susceptible adult fish
F_E	Abundance of exposed adult fish
F_I	Abundance of acutely infected adult fish
F_C	Abundance of carrier adult fish
Z_B	Abundance of spores released by bryozoans
Z_F	Abundance of spores released by fish

Table S5: List of variables for the epidemiological model. All state variables are dimensionless.

$$\frac{dY_{E,i}}{d\tau} = \beta_F \frac{Z_{B,i}}{V_i} Y_{S,i} - (\mu_F + h_i) Y_{E,i} + \sum_{k=1}^{N} d_{ki} l_k Y_{E,k} - \sum_{k=1}^{N} d_{ik} l_i Y_{E,i}; \tag{S1g}$$

$$\frac{dY_{I,i}}{d\tau} = (1-\epsilon) h_i Y_{E,i} - (\mu_F + a_i + \gamma) Y_{I,i} + \sum_{k=1}^{N} d_{ki} l_k Y_{I,k} - \sum_{k=1}^{N} d_{ij} l_i Y_{I,i}; \tag{S1h}$$

$$\frac{dY_{C,i}}{d\tau} = \epsilon h_i Y_{E,i} + \gamma Y_{I,i} - (\mu_F + \zeta) Y_{C,i} + \sum_{k=1}^{N} d_{ki} l_k Y_{C,k} - \sum_{k=1}^{N} d_{ij} l_i Y_{C,i}; \tag{S1i}$$

$$\frac{dF_{S,i}}{d\tau} = -\mu_F F_{S,i} + \zeta F_{C,i} - \beta_F \frac{Z_{B,i}}{V_i} F_{S,i} + \sum_{k=1}^{N} d_{ki} l_k F_{S,k} - \sum_{k=1}^{N} d_{ik} l_i F_{S,i}; \tag{S1j}$$

$$\frac{dF_{E,i}}{d\tau} = \beta_F \frac{Z_{B,i}}{V_i} F_{S,i} - (\mu_F + h_i) F_{E,i} + \sum_{k=1}^{N} d_{ki} l_k F_{E,k} - \sum_{k=1}^{N} d_{ik} l_i F_{E,i}; \tag{S1k}$$

$$\frac{dF_{I,i}}{d\tau} = (1-\epsilon) h_i F_{E,i} - (\mu_F + a_i + \gamma) F_{I,i} + \sum_{k=1}^{N} d_{ki} l_k F_{I,k} - \sum_{k=1}^{N} d_{ij} l_i F_{I,i}; \tag{S1l}$$

$$\frac{dF_{C,i}}{d\tau} = \epsilon h_i F_{E,i} + \gamma F_{I,i} - (\mu_F + \zeta) F_{C,i} + \sum_{k=1}^{N} d_{ki} l_k F_{C,k} - \sum_{k=1}^{N} d_{ij} l_i F_{C,i}; \tag{S1m}$$

$$\frac{dZ_{B,i}}{d\tau} = \pi_B B_{O,i} - \mu_Z Z_{B,i} + \sum_{k=1}^{N} w_{ki} \frac{Q_k}{V_k} Z_{B,k} - \sum_{k=1}^{N} w_{ik} \frac{Q_i}{V_i} Z_{B,i}; \tag{S1n}$$

$$\frac{dZ_{F,i}}{d\tau} = \pi_F (Y_{I,i} + F_{I,i}) + \kappa \pi_F (Y_{C,i} + F_{C,i}) - \mu_Z Z_{F,i} + \sum_{k=1}^{N} w_{ki} \frac{Q_k}{V_j} Z_{F,j} - \sum_{k=1}^{N} w_{ik} \frac{Q_i}{V_i} Z_{F,i}. \tag{S1o}$$

For the sake of economy in the model formulation, we introduce two new state variables $Z_F^* = \beta_B Z_F$, $Z_B^* = \beta_F Z_B$ which we term *equivalent spores*. They represent the abundance of spores needed to infect a unit bryozoan biomass (or a single susceptible fish) per unit time and unit water volume. This assumption allows us to discard the exposure rates β_B and β_F with the introduction of two synthetic contamination rates $\pi_B^* = \beta_F \pi_B$ and $\pi_F^* = \beta_B \pi_F$ (see Table S6). Synthetic contamination rates π_B^* and π_F^* are then

Type	Model	Parameter	Symbol	Value	Unit
Constant	Bryo	Baseline growth rate of B_S, B_C	r_i	$0.004 \cdot T_i$	d^{-1}
		Baseline growth rate of B_O	$r_{O,i}$	$0.002 \cdot T_i$	d^{-1}
		Rate of covert-to-overt transition	$d_{OC,i}$	$0.00032 \cdot T_i^2$	d^{-1}
		Rate of covert-to-overt transition	$d_{CO,i}$	$0.00032 \cdot T_i^2 + -0.016 \cdot T_i + 0.2$	d^{-1}
		Recovery rate	ψ	0.01	d^{-1}
		Overwintering probability (B_S, B_C)	σ_B	0.1	-
		Overwintering probability (B_O)	σ_O	0.05	-
		Biomass generated by one statoblast	ν	0.035	-
		Statoblast production rate	f_B	$0.005 \cdot T_i + +0.00025 \cdot t$	d^{-1}
		Fraction of B_C generating S_S	ϕ	0.7	-
	Fish	Rate of complete recovery	ζ	0.001	d^{-1}
		Relative cont. rate operated by Y_C, F_C	κ	0.2	-
		Natural mortality rate	μ_F	$5 \cdot 10^{-4}$	d
		Spore mortality rate	μ_Z	1.33	d
		Overwintering probability	σ_F	0.9	-
		Baseline reproduction rate	η	1	-
		Relative overwintering probability (Y_E, F_E)	p_E	0.57	-
		Relative overwintering probability (Y_I, F_I)	p_I	0.55	-
		Characteristic value of fish density	f_c	0.7	m^{-3}
Spatially heterogeneous	Bryo	Inverse of carrying capacity	ρ_i	$\overline{V_i} \max_i(c_i)/c_i$	m^3
	Fish	Strength of density dependence	ξ_i	$\max_i(\overline{D_i})/(f_c\overline{D_i})$	m^3
		Mobility rate	l_i	from Eq. (S4)	d^{-1}
Calibrated	Bryo	Rate of contamination operated by B_O	π_B^*	$\beta_F \pi_B$	$\mathrm{m}^3\mathrm{d}^{-2}$
	Fish	Rate of contamination operated by Y_I, F_I	π_F^*	$\beta_B \pi_F$	$\mathrm{m}^3\mathrm{d}^{-2}$
		PKD-caused mortality rate	a_i	$a_{15}/225 \cdot T_i^2$	d^{-1}
		Rate of disease development	h_i	$h_{15}/225 \cdot T_i^2$	d^{-1}
		Rate of recovery from acute infection	γ		d^{-1}
		Fraction of acute infections	ϵ		-
		Average fish mobility rate	l_{avg}		d^{-1}

Table S6: Reference set of parameters for the epidemiological model. Subscript i identifies site-dependent parameters. Water temperature T_i and time τ are in Celsius degrees and ordinal days, respectively. The reader is referred to [14] for a list of references.

calibrated (see *Details on model simulations and calibration*).

Inter-seasonal model. The following difference equation system relates the state of the model variables at the end of a season (y) with that at the beginning of the following season ($y+1$).

$$B_{S,i}(y+1) = \sigma_B B_S(y) + \nu S_S(y); \tag{S2a}$$

$$B_{C,i}(y+1) = \sigma_B B_C(y) + \sigma_O B_O(y) + \nu S_I(y); \tag{S2b}$$

$$B_{O,i}(y+1) = 0; \tag{S2c}$$

$$S_{S,i}(y+1) = 0; \tag{S2d}$$

$$S_{I,i}(y+1) = 0; \tag{S2e}$$

$$Y_{S,i}(y+1) = \sum_{j \neq i} W_{ij} \eta \mathcal{F}_j(y) \exp\left(-\frac{\xi_j}{\overline{V_j}} \mathcal{F}_j(y)\right); \tag{S2f}$$

$$Y_{E,i}(y+1) = 0; \tag{S2g}$$

$$Y_{I,i}(y+1) = 0; \tag{S2h}$$

$$Y_{C,i}(y+1) = 0; \tag{S2i}$$

$$F_{S,i}(y+1) = \sigma_F(Y_{S,i}(y) + F_{S,i}(y)); \tag{S2j}$$

$$F_{E,i}(y+1) = 0; \tag{S2k}$$

$$F_{I,i}(y+1) = 0; \tag{S2l}$$

$$F_{C,i}(y+1) = \sigma_F \left\{ p_E \left[Y_{E,i}(y) + F_{E,i}(y) \right] + p_I \left[Y_{I,i}(y) + F_{I,i}(y) \right] + Y_{C,i}(y) + F_{C,i}(y) \right\}; \tag{S2m}$$

$$Z_{B,i}(y+1) = 0; \tag{S2n}$$

$$Z_{F,i}(y+1) = 0. \tag{S2o}$$

where $\mathcal{F}_j(y) = Y_{S,j}(y) + F_{S,j}(y) + p_E \left[Y_{E,j}(y) + F_{E,j}(y) \right] + p_I \left[Y_{I,j}(y) + F_{I,j}(y) \right] + Y_{C,j}(y) + F_{C,j}(y)$, while W_{ji} is the fraction of newborns generated by adults living in i that hatch in j ($j \in \mathcal{U}_i$, where $\mathcal{U}_i$ contains all stretches upstream of i and i itself) and is calculated via a gravity model [15]:

$$W_{ji} = \frac{W_j^A \, e^{-L_{ji}/\lambda_F}}{\sum_{j \in \mathcal{U}_i} W_j^A \, e^{-L_{ji}/\lambda_F}} \tag{S3}$$

where W_j^A is a dimensionless score for spawning suitability and λ_F the shape factor of the exponential kernel representing the deterrence factor. An exponential function was also used to express spawning suitability: $W_j^A = e^{-A_j/A_F}$, implying that eggs tend to be deposited in small, low-velocity stretches (as the upstream contributing area is a proxy of stretch cross-section and mean water velocity [5, 6, 2]). The normalization parameters $A_F = 100$ km^2 and $\lambda_F = 2000$ m chosen such that, with the reference value

l_{avg} = 0.02 d^{-1}, the distribution of YOY at the end of the season is sufficiently close to the equilibrium distribution. Note that the value of λ_F is congruent with observed values of typical migration distances covered by brown trout [16, 17].

Fish mobility. Movement rules for fish are assumed as in [13]. For the sake of completeness, details are here briefly reported.

The maximum flux of fish exiting from stretch i is $l_i F_i$ (see Eqs. (S1f)-(S1m)), where F_i is the local fish abundance regardless of the epidemiological class and age, while l_i is a mobility rate. The flux of fish from stretch i to stretch j reads $d_{ij} l_i F_i$, where $\mathbf{D}$ is a mobility matrix whose entries d_{ij} are positive when either w_{ij} or w_{ji} are equal to unity, and null otherwise. Furthermore, one has $\sum_{j=1}^{N_s} d_{ij} \leq 1$. d_{ij} expresses the relative probability that a fish exiting from stretch i chooses stretch j out of all stretches connected to i, either downstream or upstream. The strict inequality $\sum_{j=1}^{N_s} d_{ij} < 1$ holds when stretch i represents either a headwater or the outlet, where, assuming the river network as a closed system, only one direction for fish movement is allowed. The diffusion matrix henceforth used is derived thanks to the following assumptions: $\sum_{j=1}^{N_s} d_{ij} = 0.5$ if stretch i is a headwater or the outlet; for other stretches, fish have equal probability to move downstream or upstream; the probability to choose a given upstream stretch is proportional to its cross-sectional area.

The values of l_i such that a distribution of fish abundances F_i is an equilibrium state of the diffusion process are obtained by solving the following linear system:

$$d_{ij} l_i F_i = d_{ji} l_j F_j \qquad \forall i \leq N_s,\, j \leq N_s \tag{S4}$$

Only $N_s - 1$ of the above equations are not trivial identities: in fact, $d_{ij} \neq 0$ only if stretches i and j are directly connected, and every stretch has one downstream connection, with the exception of the outlet stretch. The system has ∞^1 solutions (i.e. if a set of l_i is a solution, then also $k l_i$, $k \in \mathcal{R}^+$ is a solution); a set of mobility rates for all stretches is thus defined by specifying its mean value l_{avg}.

Details on model simulations and calibration

In order to allow the system to lose memory of the initial conditions which are hardly determinable, the epidemiological model was run for 20 years (1996-2016) with real discharge data from the FOEN Zofingen station. Water temperatures before July 2014 were derived from sinusoidal interpolations of the regressed temperature time series at a local level (see *Details on model simulations and calibration*). At the beginning of the simulations, we assumed that half of the bryozoan biomass at each site is covertly infected, while all fish are uninfected. Initial sizes of local YOY and adult fish populations were set to their dynamical disease-free trajectory value (as in [13]). After around 10 years, spatial patterns of prevalence

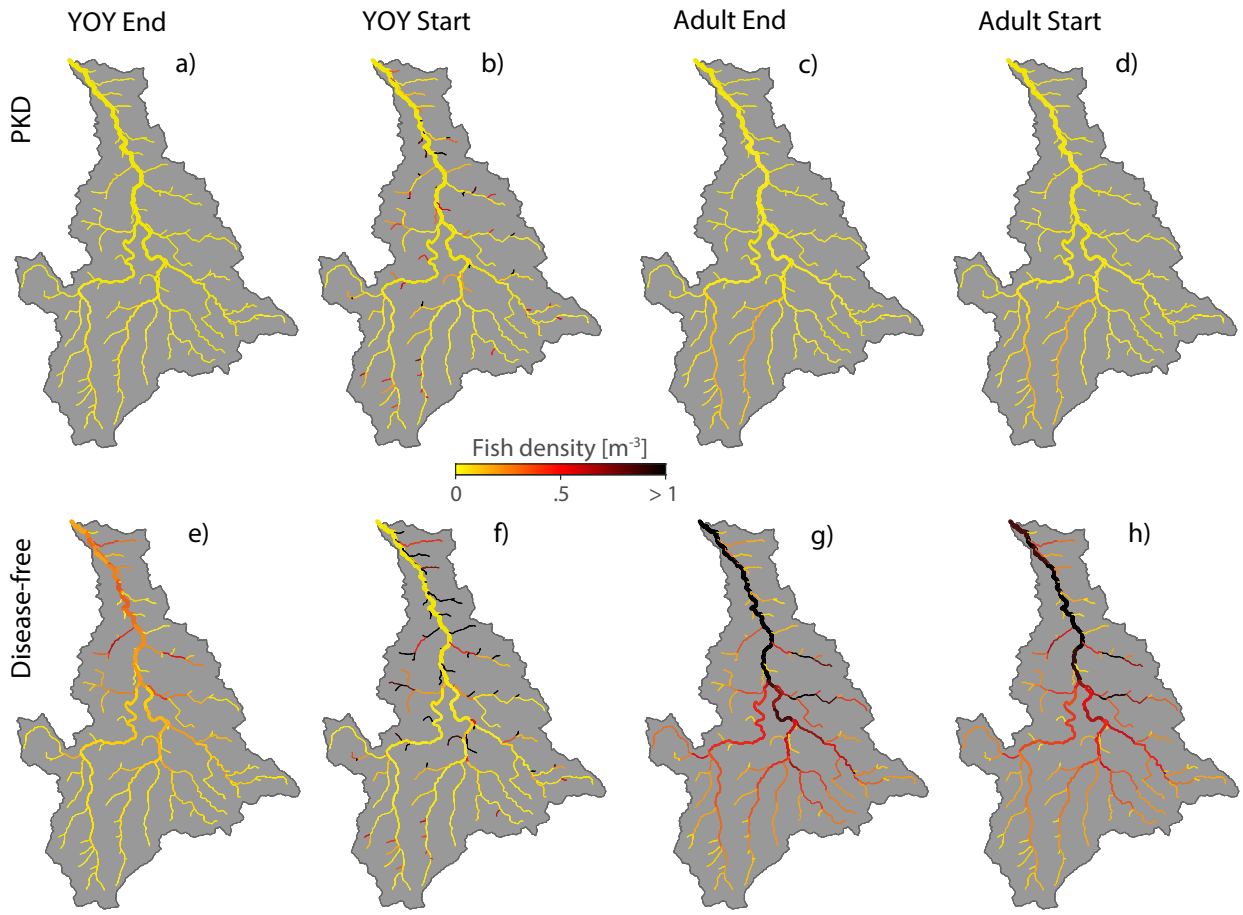

Figure S3: Maps of modelled fish density for both age-classes at the start or end of the 2016 season. Top row: simulation when PKD is present. Bottom row: simulation for the disease-free case. Calibrated parameters are set to their best fit value (see Fig. S5); other parameters are taken from Table S6.

and population size reach an annual cycle with seasonal variations only due to hydrothermal variability.

Based on a previous sensitivity analysis [14], seven epidemiological parameters were chosen for calibration: π_B^* (rate of contamination operated by bryozoans), π_F^* (rate of contamination operated by fish), l_{avg} (average fish mobility rate), a_i (PKD-induced mortality rate), h_i (rate of disease development), γ (rate of recovery from acute infection), ϵ (probability of not developing an acute infection). As both a_i and h_i depend on temperature, the parabolic relationships of [14] were used. The calibrated parameters were then the values of a_i and h_i at 15 ^{o}C. Other parameters were set to their reference values (Table S6). Calibration was performed via a Metropolis-Hastings algorithm [18]. As both prevalence and decline are constrained between 0 and 1, we hypothesized that errors are distributed according to a normal distribution truncated between 0 and 1, with standard deviation equal to 0.1.

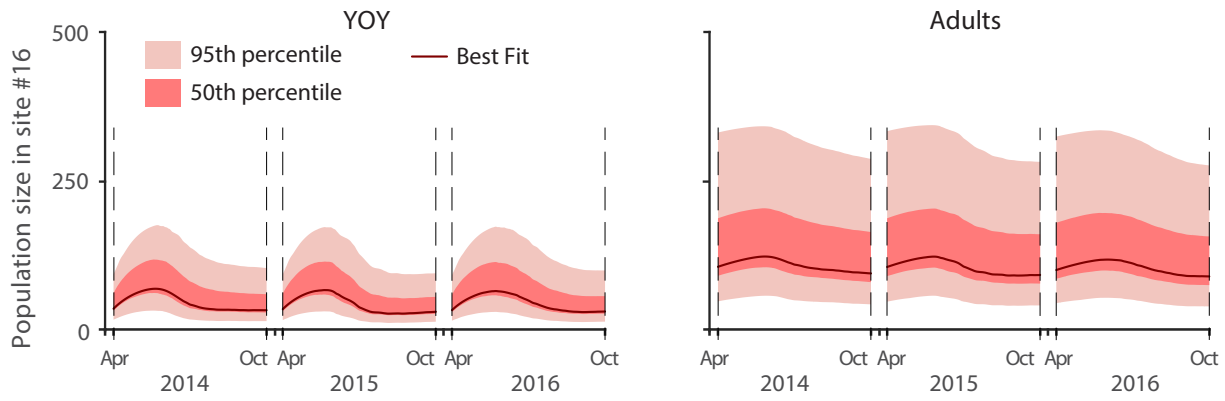

Figure S4: Time evolution of modelled fish population in site #16.

Results

Bryozoan habitat suitability

Confirming the sensitivity of the eDNA based detection method, multiple manual field searches for bryozoan colonies on all of the 15 eDNA sampling sites were able to locate populations only in the three sites (#9, #13, #15) corresponding to the highest eDNA concentrations and the most frequent detections (see Fig. 2d). Similarly, sites known to be bryozoan negative through exhaustive field searches (e.g. #5) were consistently negative via eDNA.

Epidemiological model

Modelled spatial distributions of fish densities in the absence of PKD (Fig. S3; bottom row) expectedly follow the imposed patterns at equilibrium (see *Fish mobility*), according to which fish density is proportional to mean stretch depth that, in turn, scales with the contributing area. This is particularly evident for adults (Fig. S3g, h), where seasonal variations of the spatial density distributions appear negligible. Conversely, at the beginning of the warm season YOY are mainly concentrated in small, peripheral reaches, whose spawning suitability is higher, but they later move towards the main river course (Fig. S3e, f), as confirmed by the bump in the YOY population in site #16 (see Fig. S4a). When PKD is present, the equilibrium distribution is perturbed by the enhanced disease severity at the sites characterized by large contributing area [13], thus resulting in a more spatially homogeneous distribution of adult and YOY fish density (Fig. S3a, c, d), whereas the initial YOY density remains higher in small headwaters (Fig. S3b).

Overall, PKD engenders a remarkable decline in the fish population abundance with respect to the disease-free case (92.5% for YOY and 95.7% for adults at the end of the warm season for the best fit simulation shown in Fig. S3). The bottom row of Fig. S3 represents a possible scenario of PKD complete

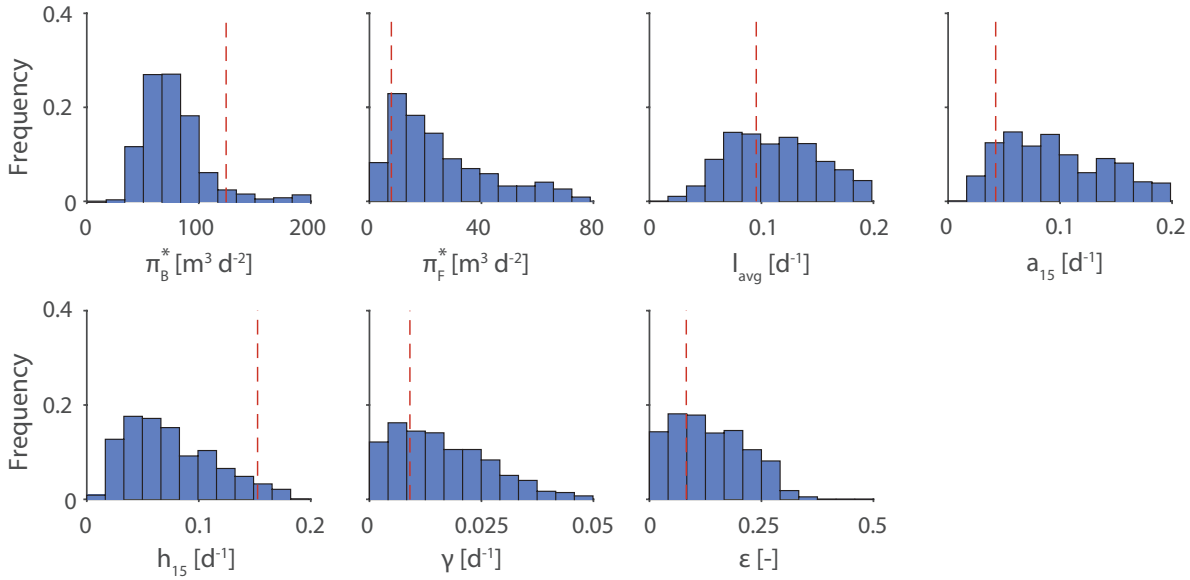

Figure S5: Posterior distributions of the calibrated parameters. The total length of the Markov Chain is 1250 steps. Dashed lines represent best-fit values.

eradication according to the present model and knowledge, although other density dependent effects regulating the fish population not included here might play a role in reducing fish abundance. Best fit and posterior distributions of the calibrated parameters are reported in Fig. S5.

Water temperatures

Water temperatures were measured at 15-minute intervals in 11 sites (see Fig. 1a). The sampling started on June 28, 2014, except for site #17, which was installed on September 27, 2015. Data are available until October 4, 2016. Daily mean temperatures were then extrapolated to all ungauged subcatchments by means of a linear regression against five morphological covariates: local elevation (*LocElv*), local slope (*LocSlp*), upstream contributing area (*UpCAr*), upstream elevation (*UpElv*), upstream slope (*UpSlp*). All covariates were normalized in the range [-1; 1], where bounds correspond to the lowest/highest values of the covariates among all subcatchments. The daily mean water temperature $T_i(t)$ in reach i at day τ is then expressed as:

$$\begin{aligned} T_i(\tau) = T_b(\tau) + \alpha_{LE}(\tau)LocElv(i) + \alpha_{LS}(\tau)LocSlp(i) + \\ + \alpha_{UC}(\tau)UpCAr(i) + \alpha_{UE}(\tau)UpElv(i) + \alpha_{US}(\tau)UpSlp(i) \end{aligned} \tag{S5}$$

where the baseline temperature $T_b(\tau)$ (i.e. the water temperature at day t of a reach whose covariates are at a null level) and the vector of coefficients $\boldsymbol{\alpha}(\tau) = [\alpha_{LE}(\tau); \alpha_{LS}(\tau); \alpha_{UC}(\tau); \alpha_{UE}(\tau); \alpha_{US}(\tau)]$ were cali-

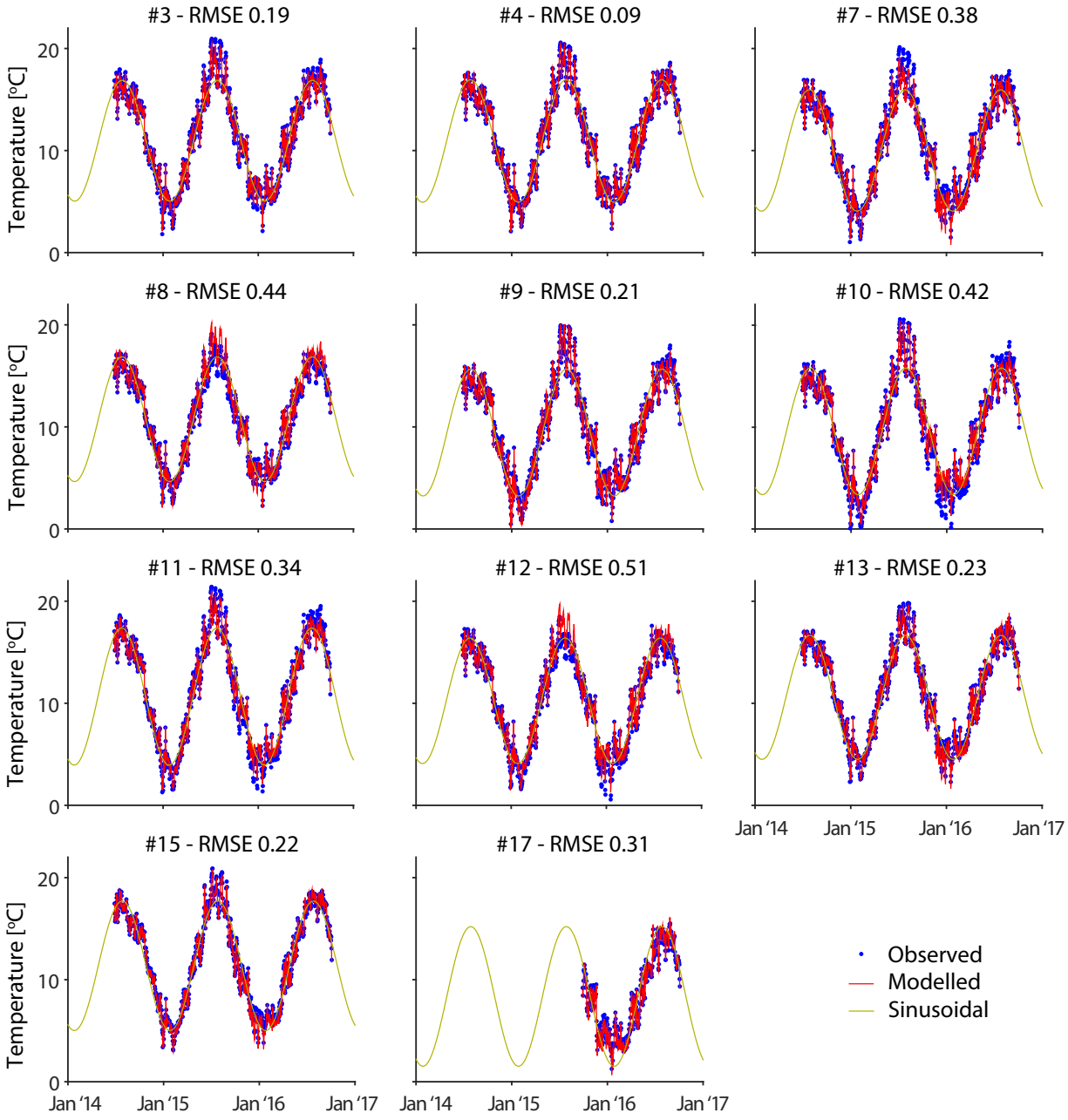

Figure S6: Modelled and observed mean daily water temperatures. For each station, the root mean square error achieved by the linear regression model is displayed. Sinusoidal fitted temperatures are also presented.

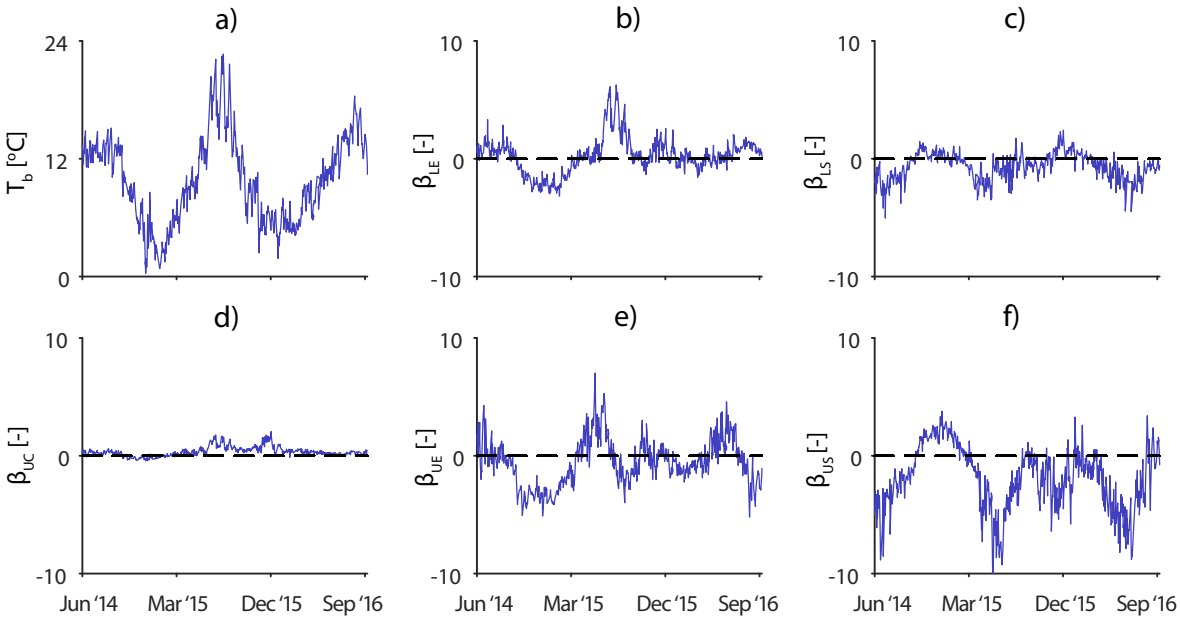

Figure S7: Time series of the calibrated values of the linear regression parameters.

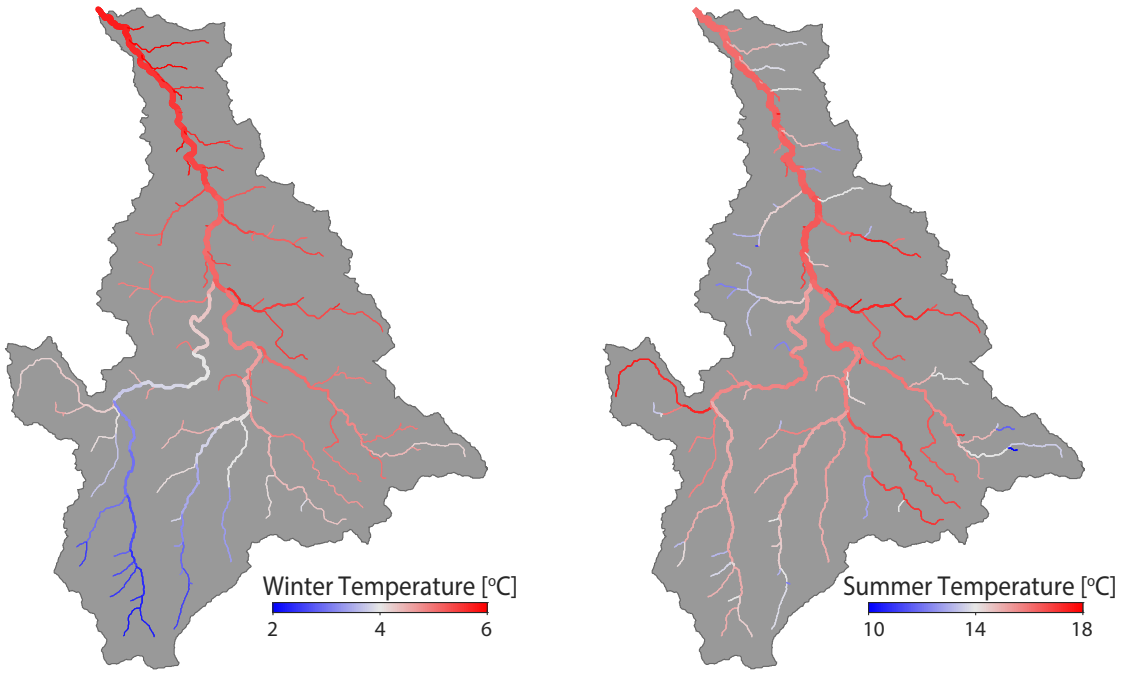

Figure S8: Map of mean water temperatures in winter (mean of winter 2015 and 2016) and summer (mean of summer 2014 and 2015) obtained by linear regression.

brated separately for each day. Modelled temperatures were extended beyond the observation period via sinusoidal functions $T_{s,i}$ fitted for each subcatchment based on linear regression outputs:

$$T_{s,i}(\tau) = T_{m,i} - T_{a,i} \cos\left(2\pi \frac{\tau - \tau_{min,i}}{365}\right)$$

where $T_{m,i}$ is the annual mean temperature in stretch i; $T_{a,i}$ the mid-amplitude of the sinusoidal signal; $\tau_{min,i}$ the day of minimum temperature (expressed in ordinal days). These three coefficients were estimated via a least square technique. A comparison between observed and modelled water temperatures is presented in Fig. S6, while Fig. S7 shows the evolution of the calibrated parameters over time.

Expectedly, the baseline temperature follows a sinusoidal trend with yearly period (Fig. S7a). During warm periods, local and upstream elevations have a positive effect on water temperature (Fig. S7b, e), while local and upstream slope are negatively correlated with high temperatures (Fig. S7c, f). During winter, such trends are reversed. The effect of contributing area is generally positive (Fig. S7d) but its magnitude is modest.

Overall, the linear regression provides a straightforward and reliable tool for the spatial extrapolation of water temperature data at a catchment scale. The model performs well in reproducing the observed data (root mean square errors for all sites are generally below 0.5 ^{o}C) and the temperature patterns obtained for intermediate and downstream reaches appear reasonable. A shortcoming has been observed in the prediction of excessively low summer temperatures in short, upstream reaches (see Fig. S8), due to the fact that no sampling sites are located close to the headwaters. However, this drawback has little impact in the outcomes of the PKD epidemiological model, since the reaches affected are those whose fish density is lower. On the other hand, we acknowledge that there might be some repercussions in the determination of bryozoan suitability maps, as the mean water temperature during the warm season (*LocMwt*) is used as a covariate for the prediction of local bryozoan density. The prediction of spatio-temporal water temperature patterns could be improved by means of a deterministic model (see [19]), but we deemed this approach to be out of the scope of this work.

List of symbols

Table S7 features a list of all symbols used in the main text and the SI Appendix. This table does not include state variables and parameters of the epidemiological model (presented in Tables S5 and S6, respectively).

Symbol	Description	Dimension
A_i	Upstream drainage area (of stretch i)	L^2
A_i^L	Watershed area directly drained by stretch i	L^2
A_R	Shape factor for spawning suitability W^A	L^2
C_i	*F. sultana* eDNA concentration measurable in stretch i	NL^{-3}
c_i	*F. sultana* eDNA concentration in an unconnected stretch i	NL^{-3}
c_0	Scale factor for exponential link function used for c_i	NL^{-3}
C_i^m	Mean measured *F. sultana* eDNA concentration at stretch i	NL^{-3}
$\overline{C_i^m}$	Mean of C_i^m at all sampling sites except i	NL^{-3}
$\overline{D_i}$	Mean water depth in stretch i	L
d_{ij}	Entry of the mobility matrix $\mathbf{D}$	-
l_i	Fish mobility rate at stretch i	T^{-1}
L_{ij}	Downstream length from stretch i to j	L
n	Manning's roughness coefficient	$TL^{-1/3}$
N_s	Total number of river stretches	-
NS_k	Nash-Sutcliffe index used for calibration in LOOCV	-
p_{ij}	Entry of the connectivity matrix $\mathbf{P}$	-
$Q_i(\tau)$	Discharge at time τ in stretch i	L^3T^{-1}
s	Performance index used to rank models of bryozoan habitat suitability	-
$T_i(\tau)$	Temperature in stretch i at day τ obtained by Eq. (S5)	Θ
$T_{a,i}$	Mid-amplitude of the sinusoidal signal $T_{s,i}$	Θ
$T_b(\tau)$	Baseline temperature at day τ	Θ
$T_{m,i}$	Annual mean temperature in stretch i	Θ
$T_{s,i}(\tau)$	Sinusoidal fitted temperature in stretch i at day τ	Θ
$V_i(\tau)$	Water volume at time τ in stretch i	L^3
$\overline{V_i}$	Mean water volume in stretch i	L^3
w_{ij}	Entry of the adjacency matrix $\mathbf{W}$	-
W_i^A	Spawning suitability score	-
W_{ji}	Fraction of newborn fish born in stretch j generated by fish dwelling in stretch i at the end of the previous season	-
$\mathbf{X}_i$	Vector of normalized covariates evaluated at stretch i	-
$\boldsymbol{\alpha}(\tau)$	Vector of parameters for the temperature model at day τ	-
$\boldsymbol{\beta}$	Vector of parameters for the bryozoan habitat suitability model	-
λ_B	Characteristic decay length for *F. sultana* eDNA	L
λ_F	Characteristic distance covered by *S. trutta* for spawning	L
τ	Time in days	T
$\tau_{min,i}$	Ordinal day of annual minimum temperature in stretch i	T

Table S7: List of other symbols.

References

[1] D. G. Tarboton. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33(2):309–319, 1997.

[2] I. Rodriguez-Iturbe and A. Rinaldo. *Fractal River Basins. Chance and self-organization*. Cambridge University Press, New York, US, 2001.

[3] Federal Office of Topography Swisstopo. Vector200 Hydrography. Retrieved from `https://www.geocat.ch/geonetwork/srv/eng/md.viewer#/full_view/f22843ce-ee17-43b3-adc8-ad7015043984`, 2016. Accessed: 2017-06-27.

[4] Federal Office of Topography Swisstopo. Geological map of Switzerland 1:500000. Retrieved from `https://www.geocat.ch/geonetwork/srv/eng/md.viewer#/full_view/ca917a71-dcc9-44b6-8804-823c694be516`, 2005. Accessed: 2017-06-27.

[5] L.B. Leopold and T Maddock. The hydraulic geometry of stream channels and some physiographic implications. Report Geological Survey Professional Paper 252, Washington DC, US, 1953.

[6] L.B. Leopold, M.G. Wolman, and J.P. Miller. *Fluvial Processes in Geomorphology*. Dover Publications Inc., New York, US, 1964.

[7] G. J. Arcement Jr and V. R. Schneider. Guide for selecting Manning's roughness coefficients for natural channels and flood plains. *US Geological Survey Water-Supply Paper*, 2339, 1989.

[8] K. Bettge, T. Wahli, H. Segner, and H. Schmidt-Posthaus. Proliferative kidney disease in rainbow trout: time- and temperature-related renal pathology and parasite distribution. *Diseases of Aquatic Organisms*, 83(1):67–76, 2009.

[9] K. Bettge, H. Segner, R. Burki, H. Schmidt-Posthaus, and T. Wahli. Proliferative kidney disease (PKD) of rainbow trout: Temperature- and time-related changes of *Tetracapsuloides bryosalmonae* DNA in the kidney. *Parasitology*, 136(6):615–625, 2009.

[10] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson. *Multivariate data analysis*. Pearson Education Limited, UK, 7th edition, 2013.

[11] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., New York, NY, USA, 2001.

[12] J. Kennedy and R. Eberhart. Particle swarm optimization. In *IEEE International Conference on Neural Networks - Conference Proceedings*, volume 4, pages 1942–1948, 1995.

[13] L. Carraro, L. Mari, M. Gatto, A. Rinaldo, and E. Bertuzzo. Spread of proliferative kidney disease in fish along stream networks: A spatial metacommunity framework. *Freshwater Biology*, 2017. Article in Press.

[14] L. Carraro, L. Mari, H. Hartikainen, N. Strepparava, T. Wahli, J. Jokela, M. Gatto, A. Rinaldo, and E. Bertuzzo. An epidemiological model for proliferative kidney disease in salmonid populations. *Parasites and Vectors*, 9(1), 2016.

[15] S. Erlander and N. F. Stewart. *The gravity model in transportation analysis: theory and extensions*, volume 3. Vsp, The Netherlands, 1990.

[16] A. Rustadbakken, J. H. L'Abee-Lund, J. V. Arnekleiv, and M. Kraabøl. Reproductive migration of brown trout in a small Norwegian river studied by telemetry. *Journal of Fish Biology*, 64(1):2–15, 2004.

[17] C. Gosset, J. Rives, and J. Labonne. Effect of habitat fragmentation on spawning migration of brown trout (*Salmo trutta* l.). *Ecology of Freshwater Fish*, 15(3):247–254, 2006.

[18] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. Cited By :5362.

[19] D. Caissie. The thermal regime of rivers: A review. *Freshwater Biology*, 51(8):1389–1406, 2006.