# HOKKAIDO UNIVERSITY

Title	Socially induced reproductive synchrony in a salmonid: an approximate Bayesian computation approach
Author(s)	Koizumi, Itsuro; Shimatani, Ichiro K.
Citation	Behavioral ecology, 27(5), 1386-1396 https://doi.org/10.1093/beheco/arw056
Issue Date	2016
Doc URL	http://hdl.handle.net/2115/66985
Rights	This is a pre-copyedited, author-produced version of an article accepted for publication in Behavioral Ecology following peer review. The version of record Behavioral Ecology, Volume 27, Issue 5, 1 January 2016, Pages 1386-1396 is available online at: https://doi.org/10.1093/beheco/arw056.
Type	article (author version)
Additional Information	There are other files related to this item in HUSCAP. Check the above URL.
File Information	SupplementaryMaterials.pdf

Instructions for use

Supplementary Fig. S1.

Location of the 30 small tributaries surveyed in the Sorachi River, Hokkaido, Japan. Each tributary (filled circle) is too small to see in the regional map. A zoomed map is shown in the right panel. An example of a typical spawning nest (i.e., redd) made by Dolly Varden is also shown in the bottom. Detailed information on the tributaries is described in Supplementary Table S1 and previous studies (e.g. Koizumi et al. 2006, 2008, 2012).
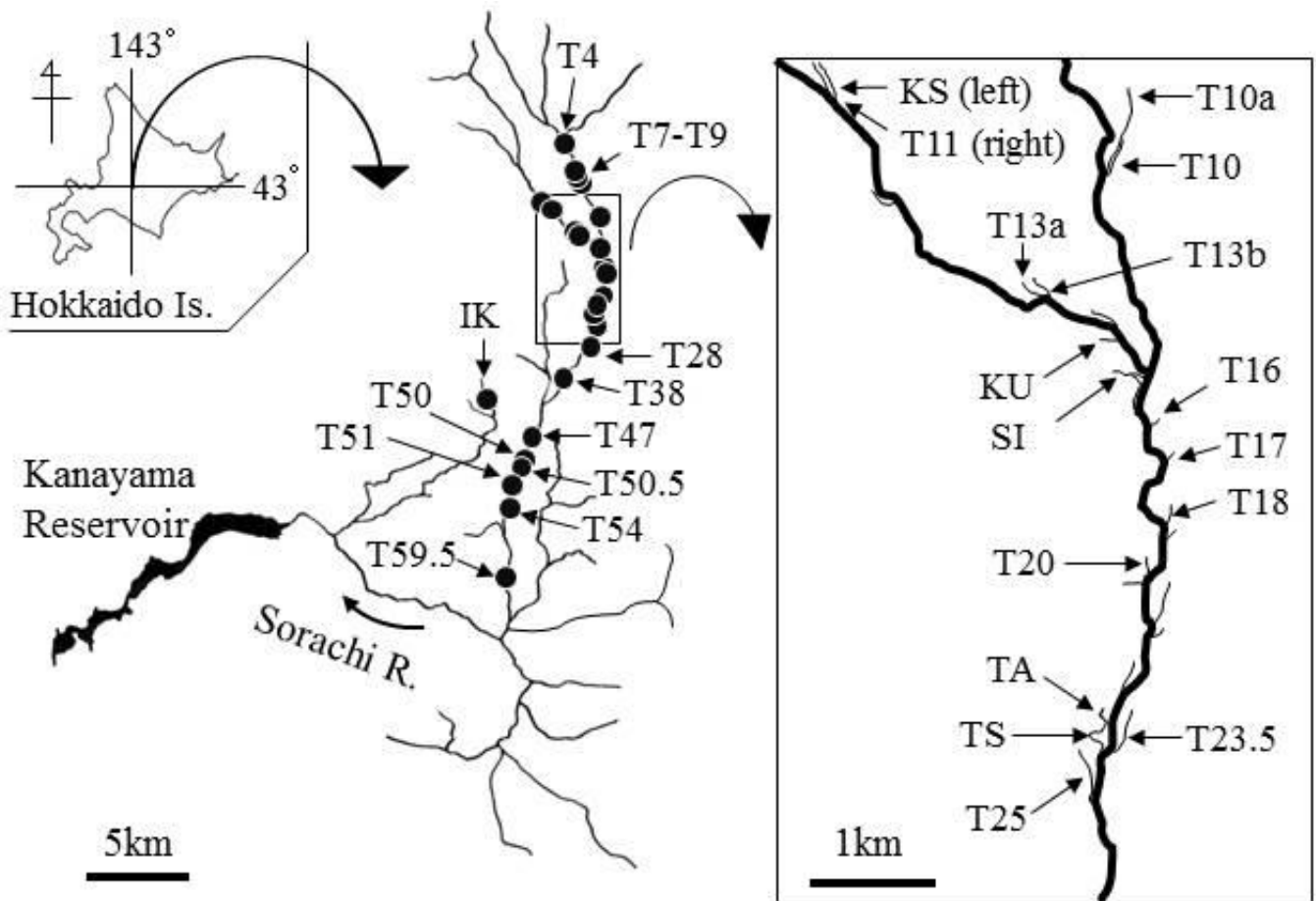
Supplementary Table S1

Basic information on the 30 tributaries surveyed in the Sorachi River, Japan.

| Tributary	Main stem distance (km)[a]	Elevation (m)	Total length (m)	Mean width (m)	Mean depth (cm)	Substrate score[b]
T4	0.0	697	131	1.1 ± 0.3	7.5 ± 3.3	1.8 ± 1.0
T7	1.2	673	682	1.8 ± 0.9	12.3 ± 10.9	4.3 ± 1.7
T7.5	1.3	671	306	1.9 ± 0.3	12.0 ± 4.0	4.3 ± 2.1
T8	1.4	667	261	2.1 ± 1.1	5.5 ± 4.2	4.0 ± 2.4
T9	2.0	658	334	1.8 ± 0.5	9.2 ± 5.6	3.2 ± 2.1
T10a	4.1	619	1110	1.3 ± 0.5	8.7 ± 6.0	3.1 ± 1.5
T10	4.2	617	399	1.6 ± 0.5	12.6 ± 11.4	4.9 ± 1.1
KS	9.2	649	328	1.3 ± 0.5	11.4 ± 8.1	2.9 ± 1.6
T11	9.2	647	423	1.8 ± 0.5	7.1 ± 3.7	2.3 ± 1.2
T13	6.8	610	198	2.3 ± 0.8	10.4 ± 7.7	3.7 ± 1.8
T13b	6.8	610	8[c]	na	na	na
KU	6.3	599	283	3.2 ± 1.7	5.5 ± 4.7	2.6 ± 1.1
SI	5.8	591	297	1.9 ± 0.8	3.7 ± 2.4	3.2 ± 1.2
T16	6.2	586	140	0.8 ± 0.3	8.5 ± 4.8	1.7 ± 1.6
T17	6.5	584	284	0.7 ± 0.2	6.8 ± 4.0	2.1 ± 1.2
T18	7.0	578	211	1.1 ± 0.3	4.2 ± 2.9	2.2 ± 0.4
T20	7.5	572	377	1.0 ± 0.3	9.6 ± 3.9	3.3 ± 1.1
T23.5	9.0	556	324	1.1 ± 0.4	9.4 ± 3.8	2.7 ± 1.4
TA	8.9	561	146	1.6 ± 1.0	5.5 ± 3.0	2.9 ± 1.5
TS	8.9	556	388	2.2 ± 0.4	9.3 ± 4.8	2.6 ± 1.5
T25	9.6	552	794	1.1 ± 0.3	8.8 ± 3.6	3.3 ± 1.5
T28	10.6	553	541	1.4 ± 0.5	9.3 ± 1.8	3.6 ± 1.6
T38	13.3	513	430	2.5 ± 1.2	19.2 ± 8.1	3.1 ± 1.8
T47	15.7	488	771	2.2 ± 1.2	9.6 ± 4.5	3.8 ± 1.5
T50	17.2	477	682	2.0 ± 0.9	13.4 ± 9.1	3.9 ± 1.2
T50.5	17.4	475	128	1.2 ± 0.6	8.0 ± 3.6	3.4 ± 1.4
T51	18.2	472	323[c]	1.4 ± 0.4	6.6 ± 2.3	4.0 ± 0
T54	19.8	451	106[c]	1.6 ± 0.1	10.0 ± 4.6	5.3 ± 1.5
T59.5	23.3	425	260[c]	1.0 ± 0.3	8.3 ± 5.0	2.5 ± 0.5
IK	35.1	585	302[c]	2.1 ± 0.9	4.9 ± 4.1	3.4 ± 2.1

[a] Watercourse distance from the uppermost tributary T4.

[b] 1, silt (< 0.063 mm); 2, sand (0.063-2 mm); 3, gravel (2-16 mm); 4, pebble (16-64 mm); 5, cobble (64-256 mm); 6, boulder (> 256 mm); 7, bedrock.

[c] Length for the reach where redd count survey was conducted.

Supplementary Fig. S2

Number of redds observed in the 30 local populations of stream-dwelling Dolly Varden in the Sorachi River, Hokkaido, Japan. The x-axis is the same scale for all panels (from 1 September to 23 December 2007).
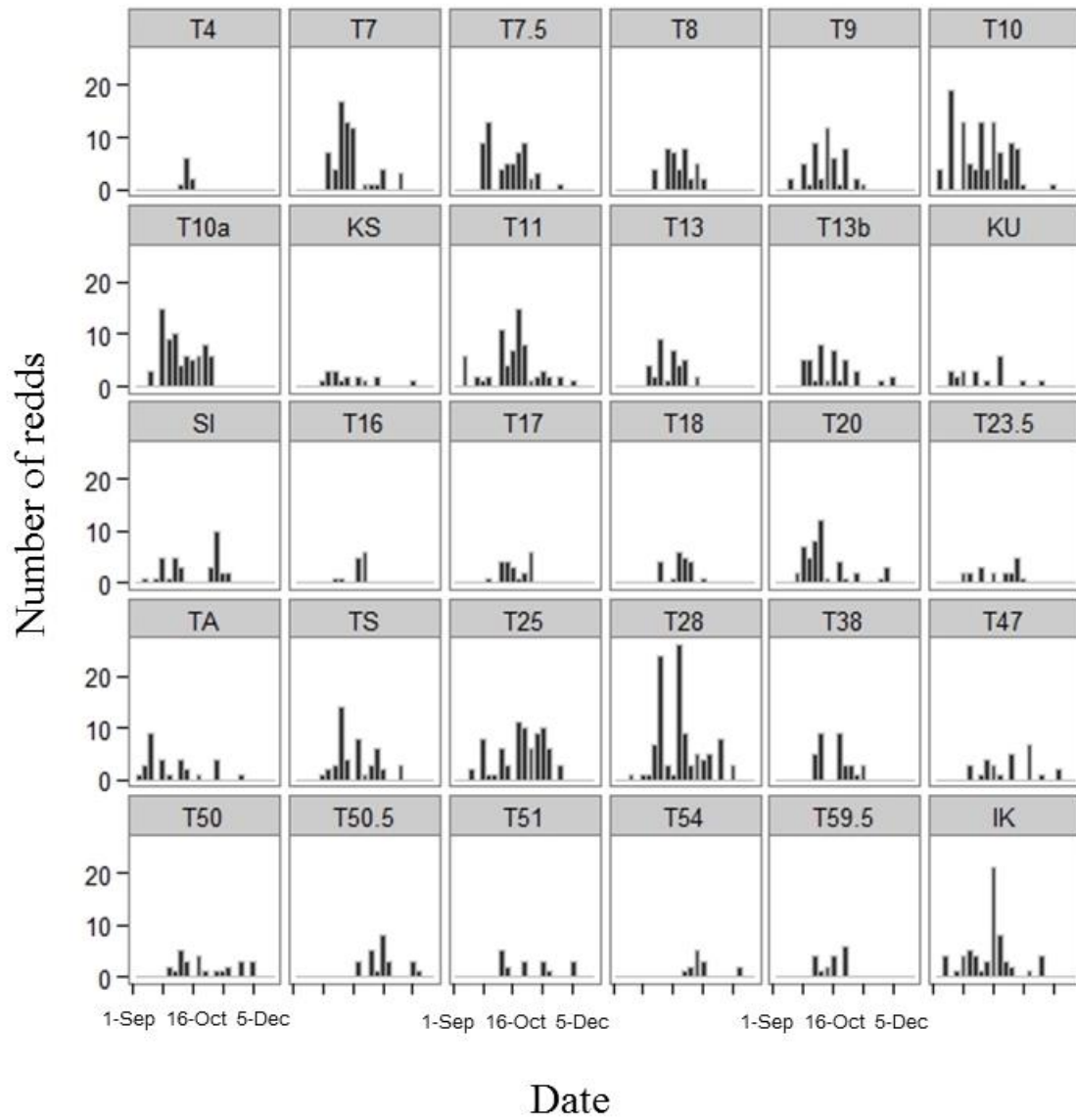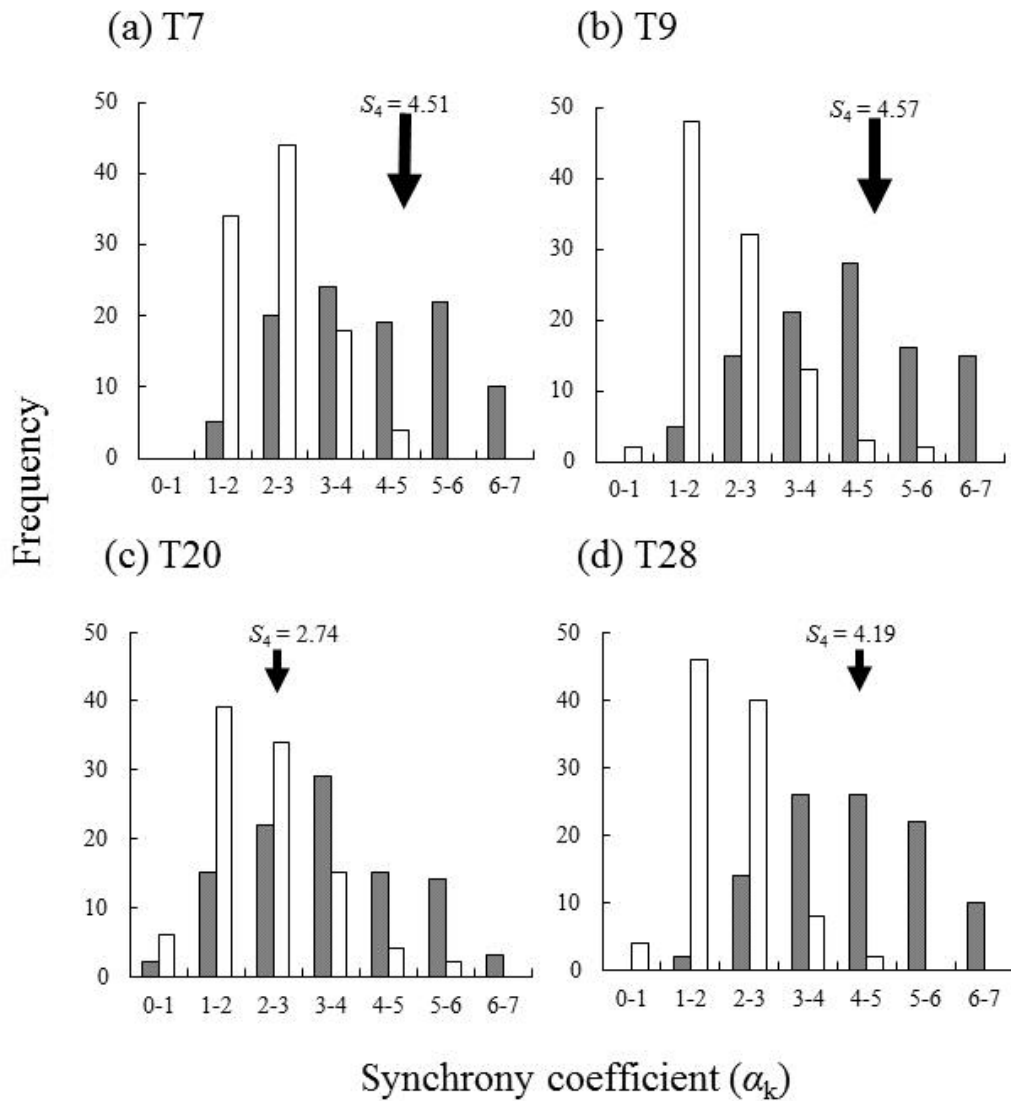
Supplementary Table S2

Values of summary statistics, estimates of the parameters in the null model, and AKB-estimates of the synchrony model for the 30 local populations of Dolly Varden.

Tributary	T4	T7	T7.5	T8	T9	T10a	T10	KS	T11	T13	T13b	KU	SI	T16	T17
Summary statistics															
Number of redds S_1	9	63	58	40	49	108	72	16	67	34	39	20	33	13	21
Mean spawning date S_2	10/12	10/15	10/13	10/22	10/12	10/5	10/8	10/15	10/17	10/15	10/15	10/8	10/18	10/21	10/19
SD over spawning date S_3	3.40	14.98	15.63	11.28	13.47	21.75	13.63	18.95	19.28	10.35	18.82	20.69	20.98	7.20	9.49
Auto-covariance S_4	-1.96	-7.08	-1.58	-2.48	-5.66	-1.47	-0.92	-0.28	-0.91	-2.41	-2.22	-0.90	-1.70	0.45	-0.09
Null model															
$S_1/1.8$	5	35	32	22	27	60	40	9	37	19	22	11	18	7	12
Mean spawning date μ_k	10/11	10/13	10/11	10/20	10/10	10/4	10/6	10/13	10/17	10/13	10/12	10/6	10/23	10/19	10/17
SD σ_k	3.26	14.34	15.02	11.09	13.25	25.42	13.58	18.29	22.66	10.13	20.23	20.09	29.24	7.06	9.15
AKB-estimates															
Number of females N_k	5	34	32	22	27	60	40	9	37	19	22	11	18	7	12
Mean spawning date μ_k	10/11	10/15	10/12	10/22	10/12	10/3	10/6	10/13	10/15	10/14	10/14	10/7	10/17	10/19	10/17
SD σ_k	4.06	14.78	15.68	11.30	13.41	22.26	13.84	19.02	19.06	10.45	19.15	20.72	20.96	7.30	9.62
Synchonicity coefficient a_k	3.05	4.51	2.26	2.69	4.57	1.97	1.59	2.37	0.91	3.02	4.81	4.36	3.93	1.47	1.81

Tributary	T18	T20	T23.5	TA	TS	T25	T28	T38	T47	T50	T50.5	T51	T54	T59.5	IK
Summary statistics															
Number of redds S_1	21	46	19	30	47	76	101	33	27	26	24	17	13	17	63
Mean spawning date S_2	10/22	10/11	10/21	10/1	10/20	10/26	10/24	10/19	10/31	10/29	11/13	10/31	11/12	10/17	10/13
SD over spawning date S_3	9.59	19.95	17.39	22.26	16.55	18.40	18.45	13.29	21.99	23.09	13.51	21.82	14.11	10.29	21.12
Auto-covariance S_4	-0.71	-1.41	-1.07	-0.32	-1.77	-1.67	-9.58	-1.29	-0.75	-0.34	-1.70	-0.51	-0.30	-1.30	-0.85
Null model															
$S_1/1.8$	12	26	11	17	26	42	56	18	15	14	13	9	7	9	35
Mean spawning date μ_k	10/20	10/14	10/19	9/29	10/18	10/23	10/21	10/17	10/28	10/26	11/11	10/28	11/9	10/15	10/10
SD σ_k	9.58	27.57	17.48	22.08	16.03	18.20	17.88	12.80	21.47	22.12	12.83	21.89	13.21	9.66	21.79
AKB-estimates															
Number of females N_k	12	26	11	17	26	42	56	18	15	15	13	10	7	9	35
Mean spawning date μ_k	10/21	10/9	10/20	9/29	10/19	10/24	10/24	10/18	10/29	10/27	11/12	10/30	11/10	10/16	10/10
SD σ_k	9.76	20.10	17.47	22.53	16.42	18.23	18.43	13.25	22.02	22.75	13.30	21.76	13.96	10.28	21.40
Synchonicity coefficient a_k	2.30	2.74	3.82	2.14	2.36	1.76	4.19	2.41	2.68	1.51	3.22	3.00	2.91	3.41	1.42

Supplementary Fig. S3

The distributions of the synchrony coefficients estimated by the AKB-algorithm from redd count data simulated by the null model (open bars) and synchrony model (filled bars) in populations T7, T9, T20 and T28. The arrows indicate the values of the synchrony coefficients estimated by the observed data.

Supplementary Table S3.

Results of Poisson and ZIP regressions for the number of redds and an environmental variable (i.e., precipitation or temperature) in each population.

Population	Poisson regression			ZIP regression (count model)		
	Intercept	Precipitation	Temperature	Intercept	Precipitation	Temperature
T4	-1.528	-0.005	0.045	4.192	0.136 **	-0.369
T7	1.981	-0.017 *	-0.410	1.213	-0.037 **	0.066
T7.5	-1.137	-0.029 *	0.152 *	0.565	-0.024	0.082
T8	3.354	-0.020 *	-0.157 *	1.368	-0.045 **	0.047
T9	-0.510	-0.019 *	0.098	-0.070	0.023	0.004
T10a	1.220	0.007	0.007	2.173	0.003	-0.031
T10	0.779	-0.001	0.026	1.618	-0.010	0.021
KS	-0.930	0.002	0.025	3.011	0.018	-0.211
T11	1.321	-0.001	-0.016	1.576	-0.001	-0.010
T13	1.702	-0.062	-0.622	1.472	0.012	-0.030
T13b	1.375	0.007	-0.052	2.173	0.006	-0.074
KU	-2.308	0.005	0.126 *	-0.102	-0.011	0.055
SI	-0.506	0.004	0.050	2.632	-0.019	-0.089
T16	0.140	-0.021	-0.016	-8.306	-0.149 *	0.697 *
T17	1.031	-0.018	-0.045	3.488	0.108 **	-0.261 *
T18	-0.235	-0.031	0.041	-3.764	-0.017	0.295
T20	-2.672	0.013 *	0.188 **	-0.697	0.016 *	0.089
T23.5	-0.297	-0.008	0.018	-5.379	-0.024	0.380 **
TA	1.601	0.009	-0.101	2.519	0.003	-0.125 *
TS	0.255	-0.007	0.039	-2.233	-0.017	0.225 *
T25	1.274	-0.015 *	0.013	1.208	-0.021 *	0.045
T28	0.111	-0.011 *	0.100 **	0.778	-0.009 *	0.068 *
T38	-1.370	-0.004	0.114 *	-1.188	-0.005	-0.070
T47	1.303	-0.006	-0.062	2.271	0.003	-0.091
T50	0.002	-0.019	0.026	-0.949	0.012	0.072
T50.5	1.124	0.007	-0.076	-0.890	0.044 *	0.069
T51	-2.820	-0.011	0.160 *	-4.550	0.048	0.235 *
T54	3.798	-0.052	-0.234 *	2.916	-0.019	-0.135
T59.5	-0.995	-0.055 *	0.080	-5.686	-0.032	0.398 *
IK	1.871	-0.009	-0.041	5.093	-0.014	-0.211 *
(mean)	0.298	-0.012	-0.017	0.215	-0.001	0.038

* $P < 0.05$, ** $P < 0.00167$ (after Bonferroni correction, 0.05/30)

Supplementary Fig. S4.

Cross-correlation coefficients (i.e., synchrony) of redd counts between each pair of the 30 tributary populations. The cross-correlation was plotted against geographic distance (km, in watercourse distance).
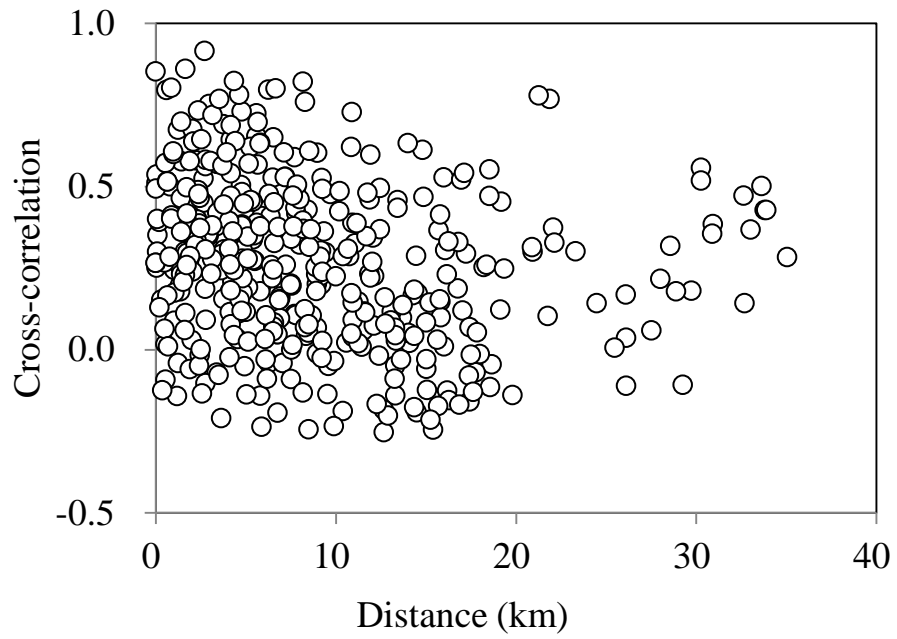
Supplementary Appendix A

Redd count procedure for Dolly Varden in the upper Sorachi River

The redds can be discriminated in the field with varying degrees of certainty depending on the fish species, fish size, stream environment (velocity, substrate, cover, etc.), and the observer's experience (Dunham et al. 2001; Gallagher et al. 2007). Because Dolly Varden in the Sorachi River are small (< 200 mm) and construct relatively small redds (30-70 cm in length), they can sometimes be difficult to identify. In order to minimize the potential for observational error, all redd counts were done by the same person (I. Koizumi) who had been studying this population for 10 years by the time of the survey.

The procedure for counting redds involved first selecting 30 suitable tributaries as mentioned in the main text. Overhead cover (mostly shrubs) was removed by the start of redd counts to make observation easier. Suitable spawning areas in the tributaries are generally limited and can roughly be recognized. Second, because redds can disappear for long periods of time, we conducted redd counts every 3-5 days for most of the breeding period. Disappearance mainly occurs due to to water currents that make typical pot-tailspill shapes of redds ambiguous (Gallagher et al. 2007). Also, animal footprints (by sika deer and brown bear) occasionally disrupted the redd's shape. Fortunately, there were no heavy rains during the study period to make redds disappear or make field observation difficult. Third, to prevent double counts, each redd was labeled with pink tape indicating the date and other pertinent remarks. Fourth, four levels of certainty criteria (C1-C4) were recorded for each redd based on past experience digging and confirming nests (I. Koizumi, personal observation). A redd was given a C1 classification if its appearance was unambiguous. C1 is, therefore, the highest level of certainty. A C2 classification, also highly certain, is given if the shape of the fish was clear. C3 is only potentially a redd. Such potential redds tend to be old and have less typical appearances, often together with obstacles, such as rocks, logs, or rapids, that can generate water currents that produce redd-like hole-mount shapes. C4 is probably a non-redd. We conservatively used only redds having a high reliability (C1-C2) of being identified correctly. When we observed spawning behavior (i.e., construction of redds, often accompanied with multiple males), we recorded them as redds on the same day.

Nonetheless, we do acknowledge that there might still be observation errors. First, redds of very small females (< 120 mm) can be difficult to identify. We might have missed such redds, but this omission should be negligible because small females are rare (Koizumi et al. 2006). Second, from 20 November roads were closed due to the accumulation of snow. Since we then had to travel by foot (or snowmobile), the frequency of redd counts in each tributary was reduced to every 8-10 days. Because the probability of redd disappearance is expected to increase with increasing time intervals, we might have missed some redds. However, we believe that such biases, if any, would not be significant because spawning activities were largely finished by that time (see Results) and because the probability of redd disappearance would have been reduced as water discharge became more stable with the end of rain (changed to snow).

References

Dunham J, Rieman B, Davis K. 2001. Sources and magnitude of sampling error in redd counts for bull trout. North Am J Fish Manag. 21:343–352.

Gallagher S., Hahn P. K., Johnson D. H. 2007. Redd counts. In: John-son D. H., Shrier B. M., O'Neil J. S., Knutzen J. A., Augerot X., O'Neil T. A., Pearsons T. N., editors. Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations. American Fisheries Society, Bethesda, Maryland. p. 197–234.

Supplementary Appendix B

The AKB-algorithm and the prior distributions for the synchrony model

Here, we describe the general procedure of the AKB algorithm, after the prior distributions used in this study. Mathematical details are described in Nakagome et al. (2013a) and a brief summary can be seen in Osada et al. (2013) and Nakagome et al. (2013b).

Prior distributions

We used the following prior distributions for the four parameters:

$N_k \sim$ Unif $(S_1^k/1.8 - 2, S_1^k/1.8 + 2)$ (S1)

$\mu_k \sim$ Unif $(S_2^k - 10, S_2^k + 10)$ (S2)

$\sigma_k \sim$ Unif $(S_3^k - 3, S_3^k + 3)$ (S3)

$a_k \sim$ Unif $(0, 8)$ (S4)

The first prior (i.e., number of females in population k) is because, on average, one female produced 1.8 redds (equation 2), and S_1^k should be close to 1.8 $\times N_k$. The next two (i.e., mean and SD of spawning date) are because each of these parameters should be close to the values of each summary statistic. For the synchrony coefficient, a_k, female reproductive periods in salmonids are heritable and restricted to a short time (Hendry et al. 1999), and in cases of artificial breeding, fertilization success quickly drops 1-2 weeks after ovulation (Craik and Harvey 1984; Brooks et al. 1997). Thus, we assumed that females were able to adjust spawning timing within a week or so, and we used equation (S4) and (3). Note that the synchrony affects not only S_4^k but also S_1^k and S_2 as hastening and aggregating spawning dates and that the AKB algorithm reflects all the summary statistics simultaneously..

Simulating redd count data

Let random samples from the prior distributions be $(N^{(j)}, \mu^{(j)}, \sigma^{(j)}, a^{(j)})$ $(j = 1, …, n = 2000$; here, we omit the index for population k for simplicity). Using these parameters, and equations (1)-(4) we simulated spawning dates and produced redd count data, O_t^j, for each population separately.

Equations for AKB estimates

We then calculated the summary statistics and let their values be $\bar{s}^j = (S_1^j, S_2^j, S_3^j, S_4^j)$.

For real data, denote as $\bar{s} = (S_1, S_2, S_3, S_4)$.

In the AKB algorithm, the estimator of the posterior mean (hereafter we called it *AKB estimate*) is given by, in the case of parameter μ_k,

$$\hat{\mu}_k = \sum_{j=1}^{n} w_j \mu_k^{(j)} \qquad (S5)$$

(for the other parameters, replace μ with N, σ, or a) in which w_j is the weight for the *j*-th sample and given by the following

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \left(\begin{pmatrix} k(\bar{s}^1,\bar{s}^1) & k(\bar{s}^1,\bar{s}^2) & \cdots & k(\bar{s}^1,\bar{s}^n) \\ k(\bar{s}^2,\bar{s}^1) & k(\bar{s}^2,\bar{s}^{21}) & \cdots & k(\bar{s}^2,\bar{s}^n) \\ \vdots & \vdots & \ddots & \vdots \\ k(\bar{s}^n,\bar{s}^1) & k(\bar{s}^n,\bar{s}^2) & \cdots & k(\bar{s}^n,\bar{s}^n) \end{pmatrix} + n\varepsilon I_n \right)^{-1} \begin{pmatrix} k(\bar{s}^1,\bar{s}) \\ k(\bar{s}^2,\bar{s}) \\ \vdots \\ k(\bar{s}^n,\bar{s}) \end{pmatrix} \qquad (S6)$$

Here, $k(x, y)$ is the Gaussian kernel;

$$k(\bar{s}^i,\bar{s}^j) = \exp(-\sum_{l=1}^{4}(\bar{s}_l^i - \bar{s}_l^j)^2 / 2\sigma_0^2), \qquad (S7)$$

in which $\bar{s}^j = (\bar{s}_1^j,...,\bar{s}_4^j)$ consists of standardized S_l^j by subtracting the mean and dividing by the standard deviation over *n* replicates. The same standardization was done for the real data and denote them as $\bar{s}$. I_n is the $n \times n$ identity matrix. σ_0 is the band width of the kernel function. ε is an arbitrary fixed constant for stabilizing the matrix inversion (the Tykhonov-type regularization. Equation S6 is a modification of a commonly seen statistical technique called the ridge regression). In general, if a small ε is used, weights (w_j) sensitively depend on simulated datasets, resulting in different estimates, whereas a large ε stabilizes estimates but may not capture properties of simulated datasets.

In the AKB algorithm, determining σ_0 and ε are arbitrary. Following the previous studies (Nakagome et al. 2013ab; Osada et al. 2013), σ_0 was fixed as

$$\sigma_0^2 = median(\sum_{l=1}^{4}(\bar{s}_l^i - \bar{s}_l^j)^2) \quad (i, j = 1, ..., n, i \neq j). \qquad (S8)$$

We tested $\varepsilon = 1, 10^{-1}, ..., 10^{-4}$ for various real and artificial datasets and determined to use $\varepsilon = 0.01$ because this value was the smallest value that visually stabilized estimates. (Nakagome et al. (2013) proposed the cross-validation for determining ε, although this

requires huge computation but generally involves few differences from when

determined as described above.)

The AKB algorithm was conducted by Mathcad ver. 14 (PTC Inc.). We carried out $n =$ 2000 simulations. The AKB algorithm does not necessarily require so many simulations, and partly because of the simplicity of our model, $n \geq 1000$ gave similar estimates, suggesting convergence as proved by Fukumizu et al. (2013).

References

Brooks S, Tyler CR, Sumpter JP. 1997. Egg quality in fish: what makes a good egg? Rev Fish Biol Fish. 7:387–416.

Craik JCA, Harvey SM. 1984. Egg quality in rainbow trout: the relation between egg viability, selected aspects of egg composition and time of stripping. Aquaculture. 40:115–134.

Fukumizu K, Song L, Gretton A. 2013. Kernel Bayes' rule: Bayesian inference with positive definite kernels. J Mach Learn Res. 14:3753-3783.

Hendry AP, Berg OK, Quinn TP. 1999. Condition dependence and adaptation-by-time: breeding date, life history, and energy allocation within a population of salmon. Oikos. 85:499–514

Nakagome S, Fukumizu K, Mano S. 2013a. Kernel approximate Bayesian computation in population genetic inferences. Stat Appl Genet Mol Biol. 12:667-678.

Nakagome S, Mano S, Hasegawa M. 2013b. Abcestral polymorphisms and sex-biased migration shaped the demographic history of brown bears and polar bears. PLoS ONE. 8:e78813.

Osada, N, Nakagome S, Mano S, kameoka Y, Takahashi I, Terao K. 2013. Finding the factors of reduced genetic diversity on X chromosomes of *Mavava fascicularis*: male-driven evolution, demography, and natural selection. Genetics. 195:1027-1035.

Supplementary Appendix C

Comparison of conventional ABC and AKB methods

We have also conducted a conventional ABC method, the rejection method.

We simulated 100,000 datasets using the same prior distributions as above and accepted

parameter sets if all the following was satisfied.

$$\left|S_1 - S_1^{(j)}\right| \leq 2, \ \left|S_2 - S_2^{(j)}\right| \leq 2, \ \left|S_3 - S_3^{(j)}\right| \leq 1, \ \left|S_4 - S_4^{(j)}\right| \leq 1$$

Table S2 compares the posterior means estimated by the conventional ABC-method

with those by the AKB-algorithm for three local populations. The means are close for

all populations.

Table S2. The posterior means estimated by the ABC rejection method and those by the

AKB-algorithm.

	T7	T9	T28
ABC rejection	4.78	4.58	4.40
AKB	4.51	4.57	4.19