Supporting Text

Transmission dynamics data

Three migration routes were studied: Tribune Channel (TR), Knight Inlet (KN), and Kingcome Inlet (KC). The TR sample sites are shown in Fig. 1 and the sample sites for KN and KC are shown in Fig 4. Juvenile salmon were collected at each location by beach seine and a random subsample was retained alive in buckets. In 2004, ≈50 juvenile pink and ≈50 juvenile chum salmon were randomly selected and assayed for sea lice. In 2005, 50-80 juvenile pink salmon were randomly selected and assayed for sea lice. Copepodid, chalimus, and motile stage lice were distinguished using an established nonlethal methodology (1). In 2005, chalimus lice were differentiated into chalimus I/II and chalimus III/IV stages. Salmon were subsequently released at the location of capture. A parallel lethal sampling program resolved the species distributions of motile lice (615 of 653 motile lice on 822 juvenile pink and chum salmon were *L. salmonis* in 2005 and 576 of 586 motile lice on 414 juvenile pink and chum salmon were *L. salmonis* in 2004).

Transmission dynamics model

The juvenile salmon migration routes are modelled as a one-dimensional infinite domain and the spatial stage-structured dynamics of lice infesting outmigrating juvenile salmon are given by the delay differential equations

$$\begin{aligned}\frac{dC}{dx} &= \frac{\beta}{v}\left[L(x) - L(x-\lambda_c)\right] \\ \frac{dH}{dx} &= \frac{s_c\beta}{v}\left[L(x-\lambda_c) - L(x-\lambda_h)\right] \\ \frac{dM}{dx} &= \frac{s_c s_h\beta}{v}\left[L(x-\lambda_h) - L(x-\lambda_m)\right]\end{aligned} \quad ,$$

which track the mean abundances of copepodid , chalimus, and motile lice, respectively. Salmon migrate at an average velocity v, lice attach to host fish at rate β, and s_c and s_h are the proportions of surviving copepodids and chalimi respectively. The λs are the cumulative distances salmon travel in the mean durations of successive developmental stages of lice (copepodids, chalimi, and motiles). These equations can also be written in their integral form

$$C(x) = \frac{\beta}{v} \int_{x-\lambda_c}^{x} L(u)du$$
$$H(x) = \frac{s_c\beta}{v} \int_{x-\lambda_h}^{x-\lambda_c} L(u)du$$
$$M(x) = \frac{s_c s_h\beta}{v} \int_{x-\lambda_m}^{x-\lambda_h} L(u)du,$$

to specify mean abundance of louse developmental stages at location x.

Advection-diffusion-decay equations were used to describe the dispersion of planktonic larval lice from a point-source (salmon farm). Planktonic lice must first pass through a noninfective naupliar stage before developing into infective copepodids. The spread of nauplii from farm salmon is

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} - \gamma\frac{\partial n}{\partial x} - (\mu_n + \theta_n)n,$$

with the conditions $\lim_{x\to\pm\infty} n(x)=0$. The diffusion coefficient D accounts for the combined effect of tides and winds and random movements of individuals, γ is the advection of larvae due to currents, and individuals die at a per capita rate μ_n and transform to copepodids at rate θ_n . We fixed γ = 1.56 km·day^{-1}, the average seaward advective flow for the Broughton Archipelago (2). We also fixed (μ_n +θ_n)= 4/5 days^{-1} according to experimental data of naupliar developmental and survival rates (3). The spatial steady-state solution yields a probability density function (PDF) for the distribution of nauplii around the source:

$$k_n(x) = c_n \begin{cases} e^{a_1 x}, & x \le y \\ e^{a_2 x}, & x > y \end{cases}, \quad a_2 < 0 < a_1,$$

where $a_{1,2} = [\gamma \pm (\gamma^2 + 4\mu_n D)^{0.5}] (2D)^{-1}$. The coefficient c_n ensures the PDF integrates to unity. Copepodids disperse according to the same advection-diffusion-decay equation as nauplii, with the exception that μ_c (the copepodid mortality rate) replaces ($\mu_n + \theta_n$). We fix $\mu_c = 1/5$ according to experimental data (3). The spread of copepodids from a point source is then

$$k_p(x) = c_p \begin{cases} e^{b_1 x}, & x \le z \\ e^{b_2 x}, & x > z \end{cases}, \quad b_2 < 0 < b_1,$$

where $b_{1,2} = [\gamma \pm (\gamma^2 + 4\mu_p D)^{0.5}] (2D)^{-1}$. The distribution of nauplii around $x = y$ forms a distributed source of copepodids, $k_n(x)$, and the PDF for the resulting distribution of copepodids around $x=y$ is given by the convolution

$$k(x) = \int_{-\infty}^{\infty} k_n(z)\, k_p(x - z)\, dz\,.$$

The distribution of copepodids produced by farms at x_i and copepodid numbers ϕ_i is then $L_1(x)=\Sigma_i \phi_i k(x-x_i)$.

The presence of gravid female lice parasitizing the juvenile salmon requires an accounting for successive generations of lice. Assuming that the spatial distribution of juvenile salmon is uniform, then M parasitic motile lice per juvenile salmon at location y will produce φ planktonic copepodids, and these copepodids will be distributed according to

$$L_2(x) = \varphi \int_{-\infty}^{\infty} M(y)\, k(x - y)\, dy\,.$$

The distribution of planktonic copepodids from natural sources is approximated by a uniform spatial distribution: $L_0(x)=\kappa$. The composite spatial distribution of planktonic copepodids from all three sources is simply their summation: $L=L_0+L_1+ L_2$.

Fitting the transmission dynamics model to field data

This model can also be formulated as a stochastic Poisson-binomial infection-survival process to create a probabilistic interface with data (4). Let $N_c(x)$, $N_h(x)$, and $N_m(x)$ be spatially explicit discrete random variables for the number of copepodid, chalimus, and motile lice on an individual juvenile salmon, respectively. If we assume infection events occur independently, then N_c is a variation on the Poisson process with a variable rate parameter (5), and spatially explicit mean, $C(x)$, given in the integral equations above. A count of k chalimus lice on an individual salmon could occur from any k of n attached copepodids surviving to the chalimus stage with probability s_c. It follows that

$$\begin{aligned} \mathrm{P}\{N_h = k\} &= \sum_{n=k}^{\infty}\left[\binom{n}{k}(s_c)^k(1-s_c)^{n-k}\left(\frac{[I_h(x)]^n}{n!}\mathrm{e}^{-\mathrm{I_h}(x)}\right)\right] \\ &= \frac{1}{k!}\left[s_c I_h(x)\right]^k \mathrm{e}^{-\mathrm{s_c} I_h(x)}, \end{aligned}$$

where I_h is the mean number of attached copepodids available for recruitment into the chalimus stage at location x:

$$I_h(x) = \frac{\beta}{\nu}\int_{x-\lambda_h}^{x-\lambda_c} L(u)\,du\,.$$

Thus, N_h is a Poisson random variable with mean, $H(x)$. In the same way, we define s_h as the probability a chalimus louse survives to the motile stage and arrive at a Poisson-distributed spatially explicit mean for motile stages, $M(x)$. This formulation allows us to assign probabilities to observations, write a likelihood function, compare models, estimate parameters, and so on.

We constrained the model by imposing independently estimated parameters for the advection, development, and mortality of planktonic larvae (described above). Further, for each spatiotemporal replicate (e.g., TR-I and TR-II are separate datasets as are TR-I and KN-I), pink and chum datasets shared four parameters (larval dispersion, louse demographic rates, and ratios of farm and ambient louse production rates), because pink and chum salmon data were collected simultaneously (there is no basis for a difference in these parameter values between host species). These common parameters were the diffusion coefficient of louse dispersion (D), ratios of source strengths (ϕ_i/κ; the subscript denotes the farm number), and the ratios of the mean durations of louse development stages (λ_h/λ_c and λ_h/λ_m). The host species-specific parameters were allowed to vary between host species. These parameters were louse survival (s_c, s_h), the mean distance salmon travel in the mean duration of the parasitic copepodid stage (λ_c), the ambient infection pressure ($\kappa\beta\cdot v^{-1}$), and if gravid females were present in the datasets, the average reinfection intensity that motile lice impose ($\varphi\beta\cdot v^{-1}$).

The likelihood function consisted of the product of probabilities of observed copepodid, chalimus, and motile counts on each fish of both species across all sample sites within a dataset. That is, if Θ is the set of parameters common to pinks and chums, and if Δ_i is the set of parameters specific to pinks ($i=p$) or chums ($i=c$), then the likelihood function is

$$\prod_{s}\prod_{i=p,c}\prod_{j=c,h,m}\prod_{k_{s,i}} P\left\{N_{i,j}=n_k \mid \Theta,\Delta_i\right\},$$

where s indexes the number of sample sites in a dataset, i indexes the host species (pink or chum), j indexes the developmental stage (copepodid, chalimus, motile), and $k_{s,i}$ indexes the number of fish of species i in sample s. The maximum-likelihood values of the six shared

and five to six species-specific parameters were estimated using the genetic algorithm toolbox in Matlab. Several optimizations were run on each dataset until a single optimum was consistently found.

We fit three different models to the data. The models consisted of only ambient-origin lice, farm-origin lice, and both. The model with only ambient-origin lice was nested within the model with both sources, permitting us to use a likelihood ratio test to test the null hypothesis that lice from farms do not infect wild salmon (all lice are ambient origin). Because not all the models were nested, we used Akiake Information criteria to select the best model from among the three posed. Across all datasets, the statistics show that farm salmon infected wild salmon with sea lice, and that the best model contained both farm and ambient sources of lice (Tables 1 and 2). Both the parameter estimates and the reconstructed spatial profiles of lice larvae indicate that farm salmon were the primary source of lice (Figs 2, 5, and 6; Table 3).

Mortality estimation

First, we show generally how parasite-induced host mortality appears as an unidentifiable parameter in the transmission dynamics model. Then we briefly describe the survival analysis and its coupling to the transmission dynamics model.

Recall that in the simplest host-parasite model (6) the rate of loss of parasites due to nonparasite related host mortality is $\sum_i \mu_N i \Phi_i N = \mu_N P$, where μN is the nonparasite-related host mortality rate, Φi is the probability a host has i parasites, N is the density of hosts, and

P is the density of parasites. Similarly, the loss of parasites due to parasite-induced host mortality is the summation $\sum_i [(\alpha i) i \Phi_i N]$, where the rate of host mortality induced by *i* parasites is αi (ref. 6). Building on this foundation, we consider a parasite that is divided into Ω substages, each of the same duration and indexed by *j*. The parasite infects a cohort of hosts, and we need not consider parasite reproduction. The model takes the form

$$\frac{dN}{dt} = -\mu_N N - \sum_j \alpha_j P_j$$

$$\frac{dP_j}{dt} = \chi_{j=1}\beta L N + \chi_{j>1}\theta P_{j-1} - (\mu_j + \chi_{j<\Omega}\theta + \mu_N)P_j - N \sum_{i_1 i_2 \ldots i_\Omega} i_j (\alpha_1 i_1 + \ldots + \alpha_\Omega i_\Omega)\Phi_{i_1 i_2 \ldots i_\Omega}$$

where P_j is the density of stage *j* parasites, *N* is the density of hosts, β is the infection rate, *L* is the density of parasite larvae, α_j is the rate of host mortality induced by parasite stage *j*, μ_j is the mortality rate of parasite stage *j*, θ is the transformation rate of parasites from one stage to the next, μN is the natural host mortality rate, and $\Phi_{i_1 i_2 \ldots i_\Omega}$ is the probability a host has exactly i_1 parasites of stage 1, i_2 parasites of stage 2, and so forth. (The characteristic function χ_A takes the value 1 when statement A is true and the value zero when A is false.) For a Poisson-distributed parasite (Fig. 9), this results in a linear equation for the dynamics of the mean abundance of the parasites $\tilde{P}_j = P_j / N$

$$\frac{d\tilde{P}_j}{dt} = \chi_{j=1}\beta L + (\chi_{j>1}\tilde{P}_{j-1} - \chi_{j<\Omega}\tilde{P}_j)\theta - (\mu_j + \alpha_j)\tilde{P}_j$$

When $\Omega \to \infty$, we arrive at the transmission dynamics model, which is a delay-differential equation that implicitly assumes zero variance in the duration copepodid, chalimus, and motile stages. This is consistent with other work (7) that found a long waiting time within a stage before parasites quickly developed into the next stage. Note that αj remains unidentifiable when fit to field data of parasite abundances, because it appears with the host mortality term μj in the above equation. Thus, lack of information on this parameter does

not affect the transmission dynamics results, and independent information is required to estimate the mortality impact of lice.

Analysis of 2004 mortality data

We analyzed a subset of data from Morton and Routledge (8), where juvenile salmon were sorted into copepodid and chalimus I/II stage infection categories, held in flow through ocean enclosures, and provisioned with fish feed (8). For robust model fitting, we required a broad range of infestation levels and stable physical conditions (e.g., temperature and salinity). For this purpose, we used data from the second replicate in Morton and Routledge (8). The first replicate did not capture the upper range in infestation levels. The third replicate had possible temperature stresses towards the end.

In the survival analysis, $Q(t)$ is the probability a host fish survives to time t. The probability density function of mortality events is

$$f(t) = \frac{d}{dt}\left[1 - Q(t)\right],$$

and, because the data were censored (the experiments ended before the fates of all fish were observed), the likelihood function is

$$\prod_i f(\tau_i) \prod_j Q(\tau_j)$$

where the τi are the mortality times for each dead fish, and the τj are the times each surviving fish was removed from its enclosure and released. The likelihood function includes all treatment levels and their replicates.

We considered two survival models that reflect possible changes in pathogenicity as lice progress through development and growth. Because the control treatments (no lice) experienced very low mortality (in four treatments with 60 fish each, there were 2, 2, 2, and 1 mortalities), we exclude natural mortality from the models.

The initial conditions in the 2004 observational studies consisted of copepodids and chalimus I/II stage lice, which are much smaller and probably less pathogenic than older and larger developmental stages. The first model assumes lice initially have no impact but increase in pathogenicity later in their developmental sequence. The second model assumes lice have an initial impact and then transition into a more pathogenic stage. We will describe the second model, of which the first model is a special case.

We approximate the change in pathogenicity by dividing the louse life cycle into two stages. The first pathogenic stage begins with chalimus lice, which induce mortality in their host at rate α_1 per parasite per unit time. The second stage of increased pathogenicity induces host mortality at rate α_2 per parasite per unit time. We leave the waiting time between these stages to be a free parameter, allowing us to identify where in the parasite's life cycle there is a marked change in pathogenicity. We also leave the variance in this waiting time to be a free parameter by dividing the first pathogenic stage into a series of n substages, each of equal pathogenicity and with exponentially distributed waiting times of equal duration. The waiting time therefore has a gamma distribution, ψ, with mean μ^{-1} and variance $(n\mu)^{-1}$ (refs. 9 and 10). The probability that a louse remains in the first pathogenic stage after t time units is

$$\xi(t) = 1 - \int_0^t \psi(\tau) d\tau .$$

Assuming the second stage persists over the time scale of the observational studies, the probability that a fish carrying H_0 young chalimus lice at time 0 is alive at time t is expressed by the survival function

$$Q(t) = \exp\left[- H_0 \int_0^t \Lambda(\tau) d\tau \right],$$

where $\Lambda(\tau) = \alpha_1 \xi(\tau) + p\alpha_2 [1 - \xi(\tau)]$ is a variable hazard rate determined by the progression of lice from one pathogenic stage to the next. Here p is the proportion of lice that survive to reach the second stage. There are four parameters (α_1, $p\alpha_2$, n, μ) to be estimated. The first model, where there is no initial mortality, occurs when α_1=0.

Connecting mortality estimates with survival of outmigrating juvenile salmon

To estimate the cumulative mortality of outmigrating juvenile salmon due to sea lice infestation, we coupled the survival model to the transmission dynamics model. In our model, time maps onto space by the mean migration velocity of juvenile salmon, $v = x \cdot t^{-1}$, such that any function describing the dynamics of salmon (or parasitic lice) in time $g(t)$ becomes a function of space $g(x)$ by using the chain rule $dg/dx = dg/dt \cdot dt/dx = v^{-1} \cdot dg/dt$. The model for the dynamics of lice through the pathogenic stages is

$$\begin{aligned}
\frac{dP_{1,1}}{dx} &= \frac{p_c \beta}{v} L(x - \lambda_h) - \frac{1}{v}(n\mu_1 + \alpha_1) P_{1,1} \\
\frac{dP_{1,2}}{dx} &= \frac{n\mu_1}{v} P_{1,1} - \frac{1}{v}(n\mu_1 + \alpha_1) P_{1,2} \\
&\vdots \\
\frac{dP_{1,n}}{dx} &= \frac{n\mu_1}{v} P_{1,n-1} - \frac{1}{v}(n\mu_1 + \alpha_1) P_{1,n} \\
\frac{dP_2}{dx} &= \frac{n\mu_1}{v} P_{1,n} - \frac{\sigma}{v} P_2
\end{aligned} \quad .$$

The first term in the first equation describes the influx of chalimus stage lice, similar to the transmission dynamics model. These lice then move through successive pathogenic substages, the number of which was estimated in the survival analysis. $1/\mu_1$ is the mean duration of the first pathogenic stage, which has variance $(n\mu)^{-1}$. Once arriving in the second pathogenic stage, lice die at rate σ, which represents the sum of natural parasite mortality and parasite-induced host mortality rates ($\sigma = \mu_2 + \alpha_2$), which were not separately identifiable. However, σ could be estimated directly from the transmission dynamics data as $\sigma = v \cdot (\lambda_m - \lambda_h)^{-1}$. The proportion of juvenile salmon at location x surviving sea lice infestation is then determined by

$$\frac{dN}{dx} = -\frac{1}{v}\left[\alpha_1 \sum_{i=1}^{n} P_{1,i}(x) + p\alpha_2 P_2(x)\right]N ,$$

where $N(x_0) = 1$ and x_0 is the landward extreme of the study area (Figs. 2, 5, and 6). There are four parameters (α_1, $p\alpha_2$, n, μ_1) that were estimated from the survival data and two parameters ($\beta p_c \cdot v^{-1}$, σ) estimated from the transmission dynamics data (Tables 3 and 4).

Analysis of 2005 mortality data

The survival analysis of 2005 data required a different formulation. The 2005 observational study collected random samples of juvenile salmon along a gradient of infection levels corresponding to their passage through a zone of salmon farms. We used the same flowthrough enclosures as in the 2004 study (8) but at a different location within the study area. The location was chosen for stable oceanographic conditions (temperature remained within 8-12°C and 28-32‰) and distance from salmon farms and suspected wild salmon migration routes to prevent new infections (only 1 copepodid and 22 chalimus I/II lice were

observed on 2,423 surviving salmon). Fish were transported from their source location in aerated buckets and transferred into the enclosures with a 15 × 15-cm dip net. An immediate decline in motile lice was observed after stocking fish in the observational vessels (motile lice are easily dislodged and can freely leave and swim in search of new hosts).

The initial conditions of sea lice infections for each treatment were distributed both within and across developmental stages. The initial abundances of lice were largely chalimus III/IV stage lice and motile lice (some of which were likely dislodged due to fish handling). The data were first sorted by the number of motile lice each fish carried at death or termination of the experiment. We assume there was relatively low mortality of motile lice over the time scale of the observational study (averaged 28 days), and that pathogenicity remained roughly constant from the chalimus stage III/IV lice onward. The probability of host survival to time t is

$$Q(t) = \exp\left[-\alpha m t\right],$$

where the constant hazard rate αm which is the number of motile lice each fish carried times the rate of mortality each motile louse imposes on its host (α). This model fit the data well, but it underestimated the mortality of heavily infected hosts (Fig. 8). The estimated value of α was 0.0229 (day·lice)$^{-1}$. If we equate α with α_1 estimated in the survival analysis of 2004 pink salmon we see that p=0.05, which is lower than the estimates of chalimus lice survival (s_h) in the analysis of lice transmission dynamics (Table 3).

The cumulative mortality of outmigrating juvenile salmon was simply the solution of

$$\frac{dN}{dx} = -\frac{\alpha m(x)}{v} N$$

where $m(x)$ is the spatial distribution of motile lice as estimated by the transmission dynamics model. This model, simpler than the survival model applied to 2004 data, yields mortality estimates similar to those predicted by the 2004 survival model (Figs. 2, 5, and 6).

1. Krkošek M, Morton A, Volpe JP (2005) *Trans Am Fish Soc* 134:711-716.
2. Brooks KM (2005) *Rev Fish Sci* 13:177-204.
3. Krkošek M, Lewis MA, Volpe JP, Morton A (2006) *Rev Fish Sci* 14:1-11.
4. Krkošek M, Lewis MA, Volpe JP (2005) *Proc R Soc Lond Ser B Biol Sci* 272:689-696.
5. Papoulis A (1963) *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York).
6. Anderson RM, May RM (1978) *J Anim Ecol* 47:219-247.
7. Stien A, Bjorn PA, Heuch PA, Elston DA (2005) *Mar Ecol Prog Ser* 290:263-275.
8. Morton A, Routledge R (2005) 11:146-152.
9. Lloyd AL (2001) *Proc R Soc Lond Ser B Biol Sci* 268:985-993.
10. Keeling MJ, Grenfell BT (2002) *Proc R Soc Lond Ser B Biol Sci* 269:335-343.