
THE PHENOTYPIC SIGNATURE OF HOST-PARASITE COEVOLUTION IN CONTINUOUS SPACE

A PREPRINT

Bob Week
Integrative Biology
Michigan State University
East Lansing, MI 48824
weekrobe@msu.edu

Gideon Bradburd
Integrative Biology
Michigan State University
East Lansing, MI 48824
bradburd@msu.edu

February 1, 2021

Abstract

Here we identify the phenotypic signature of host-parasite coevolution in continuous space.

Keywords blah · blee · bloo · these are optional and can be removed

1 Introduction

Space is hypothesized to play a fundamental role in the coevolutionary process. In particular, the combined effects of local co-adaptation and spatial movement are thought to produce complex spatial patterns of matched and mismatched traits. However, theoretical studies investigating the spatial signal of coevolution have for the most part restricted themselves to understanding patterns of spatial correlation using spatially implicit models. Hence, the fine-grained details that can be produced in spatially explicit settings have yet to be understood. Here we close this gap in the context of host-parasite coevolution using continuous-space models inspired by classical quantitative genetic theory. In particular, we study the autocorrelations of mean traits within species and the cross-correlation of mean traits between species when trait dynamics are driven by (1) random genetic drift, gene-flow and abiotic stabilizing selection, (2) drift, gene-flow, abiotic stabilizing selection and unilateral adaptation of the host species, (3) drift, gene-flow, stabilizing selection and unilateral adaptation of the parasite species and (4) drift, gene-flow, stabilizing selection and host-parasite coevolution. In each case we assuming abiotic stabilizing selection is spatially homogeneous. By comparing patterns of spatial autocorrelation and cross-correlation in these four scenarios we are able to elucidate the signature of coevolution.

2 The Model

2.1 Fitness

The host is denoted H and the parasite P . Individual trait values of hosts and parasites are respectively denoted by $z_H, z_P \in \mathbb{R}$. Assuming a trait-matching model, where host fitness is minimized and parasite fitness is maximized when $z_H = z_P$, we have the following two fitness functions:

$$w_H(z_H, z_P) \propto \exp\left(-\frac{A_H}{2}(\theta_H - z_H)^2 + \frac{B_H}{2}(z_P - z_H)^2\right), \quad (1a)$$

$$w_P(z_P, z_H) \propto \exp\left(-\frac{A_P}{2}(\theta_P - z_P)^2 - \frac{B_P}{2}(z_H - z_P)^2\right), \quad (1b)$$

where $A_H, A_P > 0$ capture the strengths of abiotic stabilizing selection, $\theta_H, \theta_P \in \mathbb{R}$ are the abiotic optimal phenotypes and $B_H, B_P > 0$ capture the strengths of biotic selection experienced by each species.

2.2 Non-spatial Dynamics

Using a standard approach, we use the above fitness functions to derive the following non-spatial coevolutionary model:

$$\frac{d\bar{z}_H}{dt} = G_H A_H (\theta_H - \bar{z}_H) - G_H B_H (\bar{z}_P - \bar{z}_H), \quad (2a)$$

$$\frac{d\bar{z}_P}{dt} = G_P A_P (\theta_P - \bar{z}_P) + G_P B_P (\bar{z}_H - \bar{z}_P), \quad (2b)$$

where $G_H, G_P > 0$ denote the additive genetic variances of each species. In this linear set of equations, when abiotic stabilizing selection is absent ($A_H = A_P = 0$), the parasite mean trait \bar{z}_P evolves to match the host mean trait \bar{z}_H . Simultaneously \bar{z}_H evolves away from \bar{z}_P . In particular, when $\bar{z}_H > \bar{z}_P$, the host mean trait will evolve upwards. Similarly, when $\bar{z}_H < \bar{z}_P$, the host mean trait will evolve downwards. This and related models have been thoroughly studied by Sergey Gavrillets and others. For the case when $A_H = A_P = 0$ two possible outcomes have been demonstrated. First, when $G_P B_P > G_H B_H$ so that the parasite evolves faster than the host, the system asymptotically evolves to the equilibrium $\bar{z}_H = \bar{z}_P$. Second, when $G_P B_P < G_H B_H$ so that the host evolves faster than the parasite, there is no stable equilibrium. This implies the host evolves to escape the interaction. In the case when $A_H, A_P \neq 0$ limit cycles can emerge.

2.3 Space

In this case mean traits become functions of spatial variables (eg., $\bar{z}_H(x), \bar{z}_P(x)$ for $x \in \mathbb{R}$ and $\bar{z}_H(x_1, x_2), \bar{z}_P(x_1, x_2)$ for $(x_1, x_2) \in \mathbb{R}^2$). As a first step, we assume all model parameters are spatially homogeneous. If we assume interactions occur locally (individuals interact only when they ‘collide’) and abundance densities are spatially homogeneous, then we can obtain the following continuous space model:

$$\frac{\partial \bar{z}_H}{\partial t} = G_H A_H (\theta_H - \bar{z}_H) - G_H B_H (\bar{z}_P - \bar{z}_H) + \frac{D_H}{2} \Delta \bar{z}_H, \quad (3a)$$

$$\frac{\partial \bar{z}_P}{\partial t} = G_P A_P (\theta_P - \bar{z}_P) + G_P B_P (\bar{z}_H - \bar{z}_P) + \frac{D_P}{2} \Delta \bar{z}_P, \quad (3b)$$

where $D_H, D_P > 0$ are dispersal rates for each species and Δ is the diffusion operator representing spatial movement. In particular, $\Delta = \frac{\partial^2}{\partial x^2}$ in one-dimensional space and $\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}$ in two-dimensional space.

2.4 Random Genetic Drift

To incorporate the effects of random genetic drift, we should be able to justify the following SPDE model:

$$\frac{\partial \bar{z}_H}{\partial t} = G_H A_H (\theta_H - \bar{z}_H) - G_H B_H (\bar{z}_P - \bar{z}_H) + \frac{D_H}{2} \Delta \bar{z}_H + \sqrt{\frac{G_H}{N_H}} \dot{W}_H, \quad (4a)$$

$$\frac{\partial \bar{z}_P}{\partial t} = G_P A_P (\theta_P - \bar{z}_P) + G_P B_P (\bar{z}_H - \bar{z}_P) + \frac{D_P}{2} \Delta \bar{z}_P + \sqrt{\frac{G_P}{N_P}} \dot{W}_P, \quad (4b)$$

where N_P, N_H are the local abundance densities (here assumed to be constant in space and time) and \dot{W}_H, \dot{W}_P are space-time white noise processes. This SPDE model may possess some major benefits over more traditional continuous-space models of population genetics. In particular, it is known to have solutions for all spatial dimensions. The effect of abiotic stabilizing selection may also imply function-valued solutions.

3 Methods

To study the continuous-space signature of coevolution, we can start with an analytical approach. However, if the corresponding calculations prove to be analytically intractable or time costly, we can employ simulation-based methods to describe the shape of the autocorrelations and cross-correlation in relevant regions of parameters space. The analytical approach would begin by identifying stable attractors of the deterministic system using classical PDE theory. Following this, we may possibly apply theory of SPDE to describe stationary distributions associated with each attractor. This may lend analytical solutions for the expectation and variance of intraspecific autocorrelations and the interspecific cross-correlation, which are defined in Appendix A below.

3.1 A Note on Biological Diversity vs Statistical Uncertainty

An important distinction to make when analyzing spatial patterns of coevolving species is the difference between statistical uncertainty and biological variation across space. In particular, we can ignore the stochastic effects of random genetic drift and focus on patterns produced by deterministic processes through studying the PDE (3). In this case, we might study quantities such as the global mean trait values (ie., mean trait values averaged across space) and the spatial variance of mean trait values. These quantities are purely functions of existing biological diversity and do not take into account uncertainties inherent in the process of evolution (eg., random genetic drift). Once drift is accounted for, the global mean and spatial variance of local mean trait values themselves become random variables.

To understand why spatial expectations and variances become random variables in the presence of drift, we can think of each possible instance of the white-noise processes \dot{W}_H, \dot{W}_P . In probability theory these instances are called sample paths and are typically denoted by the variable ω . The set of all ω is denoted Ω and is referred to as the probability space. Hence, we would refer to $\dot{W}_H(\omega), \dot{W}_P(\omega)$ as particular sample paths of the processes \dot{W}_H, \dot{W}_P . This implies that a solution to the stochastic system (4) is really a set of solutions indexed by $\omega \in \Omega$. For each ω we can compute spatial averages, variances and covariance of mean traits. Hence, the probabilistic expectation and variance of these spatial statistics are found by integrating across Ω .

This heavily theoretical discussion is important for understanding the difference between spatial autocorrelation and the covariance function of a spatial stochastic process. In particular, the spatial autocorrelation we are interested in (denoted $\rho_f(r)$ for a spatial function f) corresponds to deterministic patterns. Hence, when accounting for random genetic drift, ρ_f will be a random function. To compute the expectation, variance and covariance of this function, we can use properties of the associated spatial stochastic process (the expectation and covariance function, denoted $m(x)$ or $m(x_1, x_2)$ and $C(r)$ respectively, in particular).

Finally, for the sake of completion it is worthwhile to note that sampling error adds a second layer of stochasticity.

A Definitions of Spatial Autocorrelation and Cross-correlation

A.1 Autocorrelation

In one-dimensional continuous space, the autocorrelation of a function f is defined by

$$\rho_f(x) = \int_{-\infty}^{+\infty} f(y)f(y-x)dy. \quad (5)$$

Sometimes the autocorrelation $\rho_f(x)$ is written as $(f \star f)(x)$. To extend this definition to two-dimensional space, we can calculate a bivariate autocorrelation function of f ,

$$\rho_f(x_1, x_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(y_1, y_2)f(y_1 - x_1, y_2 - x_2)dy_1dy_2, \quad (6)$$

and then average this quantity across distances $r = \sqrt{x_1^2 + x_2^2}$:

$$\bar{\rho}_f(r) = \frac{1}{2\pi r} \int_0^{2\pi} \rho_f(r \cos \theta, r \sin \theta) d\theta. \quad (7)$$

A.2 Cross-correlation

The cross-correlation of two functions f, g is defined by

$$\rho_{fg}(x) = \int_{-\infty}^{+\infty} f(y)g(y-x)dy. \quad (8)$$

Sometimes this is written as $f \star g$. Just as with the bivariate autocorrelation above, we can extend this definition to two-dimensional space. In particular, the bivariate cross-correlation of two functions f, g can be written

$$\rho_{fg}(x_1, x_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(y_1, y_2)g(y_1 - x_1, y_2 - x_2)dy_1dy_2. \quad (9)$$

We can also reduce this quantity to a function of distance using the average

$$\bar{\rho}_{fg}(r) = \frac{1}{2\pi r} \int_0^{2\pi} \rho_{fg}(r \cos \theta, r \sin \theta) d\theta. \quad (10)$$

B A Stable Equilibrium of PDE (3)

To find an equilibrium of the deterministic PDE (3) we set the time derivatives of \bar{z}_H and \bar{z}_P equal to zero. This leads to following systems of ODE's

$$0 = G_H A_H(\theta_H - \bar{z}_H) - G_H B_H(\bar{z}_P - \bar{z}_H) + \frac{D_H}{2} \Delta \bar{z}_H, \quad (11a)$$

$$0 = G_P A_P(\theta_P - \bar{z}_P) + G_P B_P(\bar{z}_H - \bar{z}_P) + \frac{D_P}{2} \Delta \bar{z}_P. \quad (11b)$$

Solving this system for the spatial functions $\bar{z}_H(x_1, x_2)$ and $\bar{z}_P(x_1, x_2)$ then returns equilibrium solutions to PDE (3). However, we can start simple by considering spatially homogeneous solutions. In these cases the spatial derivatives will return zero (in particular, $\Delta \bar{z}_H = \Delta \bar{z}_P = 0$). Finding these solutions amounts to solve the following system of algebraic equations

$$0 = G_H A_H(\theta_H - \bar{z}_H) - G_H B_H(\bar{z}_P - \bar{z}_H), \quad (12a)$$

$$0 = G_P A_P (\theta_P - \bar{z}_P) + G_P B_P (\bar{z}_H - \bar{z}_P). \quad (12b)$$

This linear system is uniquely solved by

$$\bar{z}_P(x_1, x_2) \equiv \frac{(A_H - B_H)A_P \theta_P + B_P A_H \theta_H}{(A_H - B_H)A_P + B_P A_H}, \quad (13a)$$

$$\bar{z}_H(x_1, x_2) \equiv \frac{(A_P + B_P)A_H \theta_H - B_H A_P \theta_P}{(A_P + B_P)A_H - B_H A_P}, \quad (13b)$$

where the symbol \equiv is used in place of $=$ to emphasize these functions are constant across space.

To understand whether this equilibrium is stable we perform stability analysis on the non-spatial model. Notice that system (2) can be rewritten in matrix notation as $\vec{z} = M\vec{z} + \vec{b}$ where

$$\vec{z} = \begin{pmatrix} \bar{z}_H \\ \bar{z}_P \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} G_H A_H \theta_H \\ G_P A_P \theta_P \end{pmatrix}, \quad (14a)$$

$$M = \begin{pmatrix} G_H(B_H - A_H) & -G_H B_H \\ G_P B_P & -G_P(A_P + B_P) \end{pmatrix}. \quad (14b)$$

It turns that if the real parts of the eigenvalues of M are negative then equilibrium (13) is stable. This condition holds in general when $B_H < A_H + (A_P + B_P)G_P/G_H$ and $B_H < A_H(1 + B_P/A_P)$. A sufficient condition with clear biological intuition is $B_H < A_H$, which just means the strength of biotic selection on the host is less than the strength of abiotic stabilizing selection. In this case abiotic stabilizing selection prevents the host from escaping via evolution.

C Stationary Solutions of SPDE (4)

To focus on stationary solutions of the stochastic system (4) we set the time derivatives to zero to obtain the following system

$$G_H A_H (\theta_H - \bar{z}_H) - G_H B_H (\bar{z}_P - \bar{z}_H) + \frac{D_H}{2} \Delta \bar{z}_H = \sqrt{\frac{G_H}{N_H}} \dot{W}_H, \quad (15a)$$

$$G_P A_P (\theta_P - \bar{z}_P) + G_P B_P (\bar{z}_H - \bar{z}_P) + \frac{D_P}{2} \Delta \bar{z}_P = \sqrt{\frac{G_P}{N_P}} \dot{W}_P. \quad (15b)$$

If we further restrict our focus to the single species case we recover a SPDE of the form

$$b^2(a - u) + \Delta u = \sigma \dot{W}. \quad (16)$$

In the case of two spatial dimensions, it is known that equation (16) is satisfied by a Gaussian field with a Whittle covariance function (Sigrist, Künsch, and Stahel 2014). Denoting K_1 the modified Bessel function of order 1, the Whittle covariance function is given by $C(r) = (b\sigma^2 r)K_1(br)$. Hence, we can postulate that solutions of the stationary system (15) will have Whittle cross-covariance matrices.

References

Sigrist, Fabio, Hans R. Künsch, and Werner A. Stahel. 2014. “Stochastic Partial Differential Equation Based Modelling of Large Space-Time Data Sets.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 77 (1): 3–33. <https://doi.org/10.1111/rssb.12061>.