Title

Spatial Ecology of Parasites: Integrating Models and Frameworks

Permalink

https://escholarship.org/uc/item/65w6b5dg

Author

Fornberg, Johanna

Publication Date

2023

UNIVERSITY OF CALIFORNIA

Santa Barbara

Spatial Ecology of Parasites: Integrating Models and Frameworks

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Ecology, Evolution, and Marine Biology

by

Johanna Lee Fornberg

Committee in charge:

Professor Stephen R. Proulx, Chair

Professor Cheryl J. Briggs

Professor Deron E. Burkepile

Dr. Kevin D. Lafferty

December 2023

The dissertation of Johanna Lee Fornberg is approved.

Cheryl J. Briggs

Deron E. Burkepile

Kevin D. Lafferty

Stephen R. Proulx, Committee Chair

December 2023

DEDICATION

To my parents Connie and Doug, my sister Zoe, my family and friends,
and my love Sebastian. You make me brave and strong, and I love you endlessly.

VITA OF JOHANNA LEE FORNBERG

December 2023

EDUCATION

University of California, Santa Barbara (2017-2023)

Ph.D. Ecology, Evolution, and Marine Biology
Committee chair: Dr. Stephen Proulx
Committee members: Dr. Kevin Lafferty, Dr. Cherie Briggs, Dr. Deron Burkepile
Thesis: “Spatial Ecology of Parasites: Integrating Models and Frameworks”

University of Michigan (2015-2017)

M.S. Natural Resources and Environment
Thesis: “Evolution of Host-Parasite Interactions in Fragmented Landscapes”

University of South Florida (2012-2015)

B.S. Environmental Science and Policy Honors *Summa cum laude*
Thesis: “Island Biodiversity Conservation and Recovery in Context of a Growing Need for Sustainable Agricultural Development”

PUBLICATIONS

Abonyi, A. *et al.* The Chytrid Insurance Hypothesis: integrating parasitic chytrids into a biodiversity-ecosystem functioning framework at the phytoplankton-zooplankton interface. *In review* (2023).

Semegen, S. L. & Fornberg, J. L. What need for speed? Lizards from islands missing predators sprint slower. *In review* (2023).

Sokolow, S. H. *et al.* Ecological and socioeconomic factors associated with the human burden of environmentally mediated pathogens: a global analysis. *The Lancet Planetary Health* **6**, e870–e879 (2022).

Fornberg, J. L. & Semegen, S. L. Biogeographic patterns of blood parasitism in the Aegean Wall Lizard across the cycladic islands. *Frontiers of Biogeography* **13**, (2021).

PROFESSIONAL EMPLOYMENT

Data Scientist – Clean Technology
Greenpeace (2023-Present)

- Developing custom data tools and machine learning analyses for tracking financial investments, cryptocurrency mining, political power, and fossil fuels emissions in cryptocurrency
- Managing, cleaning, analyzing data from public, disparate sources including APIs, web scraping, and public records
- Creating GIS campaign mapping and visualization tools and dashboards for stakeholder and public use

Consultant
Impact Consulting, Santa Barbara, CA (2021)

- Strategy and marketing consulting for sustainable coffee company to expand B2B operations and enter B2C sales growth internationally
- Developed strategy for marketing, fundraising, investment opportunities, and financial planning to increase profits and enter new (international) markets

Project Manager
Community Environmental Council, Santa Barbara, CA (2018-2020)

- Led the Community Kitchen Mapping Project for the Santa Barbara County Food Action Network (SBCFAN) which assessed the county's active community food spaces and future needs for improving food access
- Led several focus groups and roundtables with stakeholders to determine future goals for SBCFAN
- Contributed to CalRecycle programming at local primary schools to track food waste reduction work

Research Assistant
Ecology Center, Ann Arbor, MI (2016-2017)

- Identified chemical toxins in several hundred samples of consumer goods across several industries, especially BPA's and other plasticizers in food packaging using IR spectroscopy
- Lab assessment of lichen and moss as bioindicators of air pollutants and heavy metals across space
- Identification and analysis of heavy metals and other toxins in consumer goods using XRF techniques
- Results reporting for scientific publication and public use

Research Assistant II
University of Michigan, Ann Arbor, MI (2016)

- Literature review for book which examines the lives and experiences of contemporary women and women of the past throughout the world through an evolutionary lens

Assistant Manager
Aramark, Tampa, FL (2012-2015)

- Executed coordination of operations as well as team development, leadership, delegation, and training for >100 employees
- Leader in sustainable initiatives through project development and advocacy of opportunities for improve sustainability in operations
- Responsible for ensuring company execution standards in food production, safety, customer service, and labor management

Conservation Intern
Jatun Sacha Foundation, Galápagos Islands (2014)

- Study of invasive species issues and threats to biodiversity, especially in agriculture

- Development and execution of effective conservation and restoration projects and improvements to current projects

Research Intern
Azafady, Madagascar (2013)

- Researched biodiversity, including population and behavior studies, especially on lemur populations, herpetology, and local flora
- Implemented conservation of local habitats via habitat restoration and sustainable development initiatives in local communities
- Educated local populations on environmental processes, conservation, and sustainable living and trained adults in English language

TEACHING AND MENTORING

- EEMB 112/L – Invertebrate Zoology* (2018-2019)
- EEMB 111/211/L – Parasitology* (2018-2019)
- EEMB 111/211/L – Parasitology *head TA* (2020)
- EEMB 116/L – Invertebrate Zoology (2018-2019)
- EEMB 116/L – Invertebrate Zoology *head TA* (2020)
- EEMB 159 – Tropical Ecology* (2018-2022)
- EEMB 142A – Aquatic Communities* (2020-2022)
- EEMB 148 – River Ecology* (2021-2022)
- EEMB 179/279 – Ecological Modeling (2021)
- INT 95A – The Modern Research University (2018-2021)
- * Indicates direct mentorship of undergraduate teaching assistants while working as an instructor

Research Mentor
University of California, Santa Barbara, Santa Barbara, CA (2019-2020)

- Supervision of undergraduate student research on insect trophic (food web) interactions in Cedar Creek Reserve, Minnesota (student research is literature-based data collection)
- Ensuring student data collection is comprehensive and attuned to cover the complexity of insect dynamics

Research Mentor
University of Michigan, Ann Arbor, MI (2016-2017)

- Direct oversight of five undergraduate students during field research in remote, international field sites
- Ensured data gathering and organization was maintained at highest quality
- Trained and directed students in field techniques, especially working with reptiles and birds, and collection of blood samples in field
- Administered training and supervision of laboratory work, specifically with microscopy, parasitological blood sample preparation and staining, hematological assessments, and cryostat use

PRESENTATIONS

Midwest Fish and Wildlife Conference (2017): “Evolution of host-parasite interactions in fragmented landscapes”

AWARDS

- EEMB Departmental Grant (2021)
- EEMB Departmental Grant (2017)
- Rackham Graduate Student Research Grant (2016)
- University of Michigan Academic Leadership Fellowship (2015 – 2016)
- USF Undergraduate Scholar (2015)
- University of South Florida Director’s Award (annually) (2012 – 2015)
- USF Honors Scholar Award (annually) (2012 – 2015)
- Florida Bright Futures Academic Scholars Award (annually) (2012 – 2015)
- USF Honors Travel Scholarship Award (2013, 2014)

SELECTED STATISTICAL AND QUANTITATIVE SKILLS

- R, Python, SQL, GIS, Git/GitHub, Julia, Mathematica, Tableau, Microsoft Office
- Advanced knowledge of Dynamical Models, Bayesian and Frequentist Statistics, Predictive Models, Bioinformatics, Machine Learning, Visualization, Data Management, Data Wrangling, Data Mining
- Certified Data Science Professional (IBM, 2023)
- Certified in SQL/MySQL (Udemy, 2023)
- Certified in Animal Welfare and Behaviour (University of Edinburgh, 2016)
- Extensive knowledge of field techniques for wildlife and plants, especially relating to mammals, reptiles, invertebrates, and parasites
- Experienced in microscopy, cryostat, microtome, XRF, spectroscopy techniques

ABSTRACT

Spatial Ecology of Parasites: Integrating Models and Frameworks

by

Johanna Lee Fornberg

Parasites are increasingly recognized as essential components in ecological communities. Hosts and parasites experience intimate, long-lasting associations that impact where they are found in space and how they interact in the larger ecological community. The spatial ecology of parasites is therefore an interesting and complex area of research that requires careful study. In this dissertation I explore three areas of parasite spatial ecology building on some of the latest statistical and mathematical tools in ecology. Specifically, I concentrate on a guild of trematodes in estuaries along the Pacific west coast. I begin by quantifying how parasite distributions vary across space, with special consideration of how to do so when parasite data is limited or affected by error. In the second chapter, I explore how to statistically predict trematode abundance across several estuaries and what these findings suggest about the mechanisms driving parasite abundance. Finally, I define a mathematical model describing host-parasite transmission dynamics in systems that demonstrate an encounter dilution effect as a result of infective stage depletion in the environment. These three works not only improve what is known about the focal trematode guild, but they also improve understanding of how statistical and mathematical models can be used to gain valuable, accurate insight about parasites more generally.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Overview

When asking fundamental questions about species communities, ecologists are almost always thinking about spatial relationships of ecological phenomena. Even research questions without explicit consideration for space are concerned with ecological processes that are both a cause and effect of spatial patterns. This emphasizes one of ecology's major questions: where do we find species, and why?

Communities are associations of species that are defined by geographic space and ecological interactions and may also be bounded by ecological time (Vellend 2016). The importance of parasites and other infectious organisms has been accepted as an important facet of the definition of a community (Dobson and Hudson 1986, Dobson *et al.* 2008). Parasites and their hosts form durable communities, with complex interactions that may differ from free-living counterparts alone(Holmes and Price 1986, Combes 2001). Central to both community and disease ecology theory are paradigms and models of species occurrence and abundance across space, as well as mechanisms that drive patterns of species distributions and interactions. Data collection on host-parasite communities can be challenging since parasite data often requires destruction of the host for sampling and can be cryptic, making parasites more susceptible to sampling error (Poulin 2004). From a practical perspective, this lack of clarity about spatial relationships makes valuable insight difficult to achieve. The application of most parasite ecology work is to provide insight about species distributions and interactions in order to make predictions about future change scenarios, such as climate change and habitat modification, and to make decisions about how to mitigate, adapt, or restore ecological systems in light of changing conditions. Creating better tools and models to describe and predict how hosts and parasites vary across space and how they interact is essential to understand biodiversity and ecological complexity.

Parasites and their hosts alike may exhibit predictive variation or patterns across space, and often these patterns give insight to the underlying mechanisms that drive spatial and ecological relationships between hosts and parasites (MacArthur 1972, Levin 1992). Work on this has generated significant momentum toward understanding the complexity of how species occur and interact in space (MacArthur and Pianka 1966, Paine 1966, Chave 2013). However, there are still unanswered questions about how to accurately describe and predict parasites in space, especially when data may be limited or imperfect (Poulin 2004, Lafferty *et al.* 2008). In applied ecology, exploring mechanisms that drive spatial relationships has been essential for informing predictions about how species distributions and interactions may affect other aspects of ecological communities, ecosystems, and how to anticipate and prepare for responses to changing conditions (MacArthur 1972, Levin 1992, Lafferty *et al.* 2008, May *et al.* 2008). Improving tools for predictions and inference on parasites can enhance how ecologists forecast changing parasite transmission or host-parasite dynamics in future change scenarios (Dunn *et al.* 2009, Carlson *et al.* 2017, Mellado and Zamora 2020). Finding the most accurate and flexible tools for description and prediction is thus an essential area of research.

Parasite ecology has demonstrated that parasites can be essential to driving dynamics in a community (Loreau *et al.* 2007, Lafferty *et al.* 2008, Michalska-Smith *et al.* 2018). Host-parasite assemblages offer a tractable community that can be more easily bounded than assemblages of free-living species, in part because parasites are inherently dependent on hosts for survival and thus create a biologically meaningful definition of what species are in the community. The intimacy of host-parasite interactions also tends to make parasite distributions highly nested within the range of hosts (Poulin 2003, Guégan *et al.* 2005, Moss *et al.* 2020). Parasites may exist as free-living stages in the environment during periods of dispersal between hosts, and directly experience the external environmental conditions, but most often parasites are subjected to conditions within hosts and only indirectly to the external environment.

Identifying the community level of interest is the initial decision for studying the spatial ecology of parasites. Parasite infracommunities (parasites within a single host individual) and component communities (parasites within a host population or species) are two common definitions of a parasite community (Holmes and Price 1986). Infracommunities and component communities can each display different spatial patterns that also differ from those of hosts alone (Kuris *et al.* 1980, Stella *et al.* 2017, White *et al.* 2018). Moreover, parasites with complex life cycles may have spatial distributions that are driven by multiple host species at different life stages, as well as environmental or spatial conditions (Huspeni and Lafferty 2004, Hechinger and Lafferty 2005, Lafferty and Dunham 2005). Research on these complex drivers of parasite spatial ecology is vital for understanding dynamics of hosts and parasites but may also provide novel insight for other organisms that engage in intimate, durable relationships, such as mutualists and symbionts.

In this dissertation, I will investigate the spatial ecology of host-parasite community structure in estuaries in California (USA) and Baja California (México). Specifically, I will work with the component community of a guild (sensu Root 1967) of digenean trematode species in the first intermediate host snail (*Cerithideopsis californica*) of their life cycle. This trematode guild is the most notable parasite infecting *C. californica* by prevalence and biomass, and the guild also maintains a strikingly high biomass compared to other parasitic and free-living groups in the estuaries explored in this dissertation (Kuris 1990, Kuris *et al.* 2008, Torchin *et al.* 2015). In my first chapter, I propose analyzing data from a series of estuaries to explore the spatial distribution of trematode species as a metacommunity. Here, I focus on both describing species distributions as well as quantifying and correcting sampling error and exploring how that error may affect patterns in observational data. In the second chapter, I assess the relative significance of environmental, species trait, and spatial covariates for predicting the abundance of trematodes in estuary sites with special consideration of the role of spatial autocorrelation. In the third chapter, I develop a transmission model that examines the relationship between host density and parasitism and test the model using field and experimental data. This model especially focused on encounter dilution in host-parasite interactions and the contribution of infectious stage depletion to this process. Throughout this dissertation, I use and reference extensive work and data collected by many researchers on this parasite guild from several decades of dedicated research.

Spatial Patterns of Salt Marsh Trematodes: Detection Error and the Classification of Parasite Distributions

2.0 Abstract

The intricate life cycles of parasites, especially those that use multiple hosts, create spatial distributions that may be difficult to understand. Metacommunity theory has provided the elements of metacommunity structure (EMS) as a framework to untangle the complexity of species distributions by classifying communities with several common metrics from community ecology. Although this framework is a simple way to characterize ecological communities, it is not able to account for possible errors in species-occurrence data. Species-occupancy models, which account for detection error probability in species communities, can be used to improve the performance of this framework. We apply detection error-corrected EMS to trematode parasite data from Carpinteria Salt Marsh, where a diverse trematode guild exhibits hierarchical structures and competitive interactions within hosts, forming metacommunities. By combining EMS and occupancy models, the research provides deeper insights into how parasites are distributed which can guide inference about the ecological factors shaping these patterns. The results highlight the significance of detection errors in understanding spatial patterns in ecological systems and contribute to the broader understanding of species distribution and metacommunity dynamics in parasites.

2.1 Introduction

Despite being some of the most widespread and diverse organisms on the planet (Windsor 1998, Dobson *et al.* 2008, Lafferty *et al.* 2008), parasites are understudied. Complex life cycles, in which parasites are dependent on multiple hosts throughout their life cycle, are not uncommon and can lead to complex distributions of parasites in space (Anderson 1993, Combes 2001). Successfully finding and infecting several suitable hosts, in sequence, in the course of different life stages is an essential challenge for parasites with complex life cycles. Parasites experience environmental and biological processes that establish whether or not a host is accessible and can be infected (Combes 2001). Free-living hosts, in turn, experience their own environmental and biological pressures that create "structure", or discernable patterns in their own distributions. Unlike their free-living counterparts, parasites exist in a nest of filtering processes that lead to complex, heterogenous spatial distributions. Parasites that are able to use a wider diversity of hosts at one or multiple life stages may be distributed in ways that could alternatively reflect host dynamics or may reflect multiple drivers of occurrence at different scales. The complexity of where and how parasites are distributed can be immense but may lead to deeper understanding of host-parasite interactions (Guégan *et al.* 2005, Pullan *et al.* 2012).

Metacommunity theory is a sub-discipline of ecology that has attempted to breakdown complexity of large, disconnected systems that are linked by dispersal and other interactions over space and time (Leibold and Chase 2017). Metacommunity ecology conceptually echoes much of the progress in spatial parasitology to integrate biogeography, ecology, and evolution into the study of parasites and parasitism (Guégan *et al.* 2005, Mihaljevic 2012). Recent work in both fields has sought general principles to characterize how organisms distribute in space, and what processes lead to observed patterns (Guégan *et al.* 2005, Leibold and Chase 2017). One approach for detecting and describing spatial patterns of species has been elements of metacommunity structure (EMS), which uses pattern-based classification of species spatial distributions which can be used to infer processes that may generate those patterns. The EMS workflow assesses occurrence data on species in a metacommunity based on the value of three metrics commonly used in community ecology and biogeography (Leibold and Mikkelson 2002, Presley *et al.* 2010), see Box 1 for details. The resulting pattern is a categorization of these metrics and is interpreted to suggest possible processes or gradients that may inform a distribution pattern. Patterns detected with EMS also provide a way of describing species distributions with more consistent terminology than when communities are simply compared to null models of "randomness".

While EMS provides a procedure for detecting patterns in species data, on its own it is limited by the quality of data. Work on detection error in ecology and biogeography has demonstrated that correcting for imperfect detection of species greatly improves the accuracy of metacommunity analyses, including EMS classification and species distribution models (Lahoz-Monfort *et al.* 2013, Iknayan *et al.* 2014, Mihaljevic *et al.* 2015, Mihaljevic *et al.* 2018). EMS classification, for example, is subject to misclassification in the form of type II errors due to incomplete or inaccurate occurrence data (Mihaljevic *et al.* 2015, 2018). Species occupancy models are meant to test the probability of a species occurring at a site, while considering that the likelihood of detecting that species is imperfect (MacKenzie *et al.* 2002, Royle and Dorazio 2009). By incorporating detection probabilities into occupancy (or co-occupancy) forecasting, species occupancy modelling has been an important tool for

Box 1: Classification framework of Elements of Metacommunity Structure

EMS classifies spatial distributions of species into described patterns by assessing species occurrence data (Figure 1). Occurrence (presence-absence) data is ordinated by reciprocal averaging (*i.e.*, correspondence analysis, Gauch 1982) to generate scores that represent a gradient of similarity of species composition. The ordinated matrix is then used to calculate three metrics: coherence, species turnover, and boundary clumping (Leibold and Mikkelson 2002, Presley *et al.* 2009, 2010). Calculations of community metrics are done successively beginning with coherence.

Coherence is measured by the number of embedded absences in the ordinated site-by-species matrix in a range of sites that make up the metacommunity. Embedded absences are species absences at sites they are expected to occur based on the ordination of the empirical matrix (Leibold and Mikkelson 2002, Presley *et al.* 2010). Conceptually, many embedded absences in an ordinated matrix suggests that species are not responding to the same structuring conditions in the metacommunity. Positive coherence is defined by the presence of significantly fewer embedded absences in the empirical matrix, negative coherence is defined as significantly more embedded absences in the empirical matrix when compared to simulated null matrices. Negative coherence is classified as a checkerboard pattern and is expected to arise when strong competition for resources lead to mutual exclusion of species pairs within their range (sensu Diamond 1975). When a community does not demonstrate positive or negative coherence, it is classified as “random”. Non-interactive generalist communities may generate this random pattern, with structure based on processes such as dispersal limitation. A random pattern in EMS is an indication that species distributions are independent and not determined by the same gradient(s) identified in reciprocal averaging (*i.e.*, it is not synonymous with true randomness).

Turnover and boundary clumping are assessed after the community matrix is ordinated and coherence is calculated, using the same ordinated matrix. Community structure might also be structured by niche differences, with species responding to a shared latent gradient. Niche differences can differ in breadth (*i.e.*, generalists and specialists along the same resource), which can create nestedness along resource or environmental gradients when species turnover is low. Species turnover is calculated after coherence, once the number of embedded absences in species distributions has been minimized. Species turnover is calculated as the number of replacements of a species by another between sites, where positive turnover indicates more species replacements are observed compared to a null model. When turnover is negative, communities are considered nested. Nested species distributions occur when some species ranges are subsets of ranges of other species; often this leads to rare species occurring only at high diversity sites, and common species at low and high diversity sites. The ecological processes creating nestedness can be wide-ranging, such as geographical isolation leading to nested subsets of species in islands metacommunities (Patterson and Atmar 1986) or mutualist networks that display coexistence of more species by reducing the negative effect of interspecific competition among pollinators and seed dispersers (Bastolla *et al.*, 2009). When the species turnover is positive, and species are replaced along their coherent range, the trend of replacement of each species can determine the distribution pattern (Presley *et al.* 2010).

Finally, boundary clumping is used to measure dispersion of species and is calculated using Morisita’s index (I) (Morisita 1971). When turnover is positive, species may differ due to differences in niche centroids (*i.e.*, specialists on different resources). Boundary clumping is used to distinguish between these distributions that have high turnover but are not nested. Species are considered hyperdispersed if I is significantly less than one and clumped when I is significantly greater than one. Hyperdispersion is suggested to occur when species exhibit strong interspecific competition that lead to trade-offs in competitive ability (sensu Tilman 1982) creating an evenly-spaced EMS pattern. Clumped distributions are suggested to occur when species inhabit a patchy environment leading to aggregated species loss (*i.e.*, species loss occurs in groups of species, Clements 1916) and this pattern is called Clementsian. Species may instead experience different, individual responses along an environmental gradient leading to I being non-significantly different from one (Gleason 1926) which is classified as a Gleasonian pattern (Figure 1).

studies that are sensitive to false negatives (absence) or false positives (presence) in data (MacKenzie *et al.* 2002, Guillera-Arroita 2017, Devarajan *et al.* 2020). By incorporating species occupancy models to correct for false negatives in species data, Mihaljevic *et al.* (2015) demonstrated that misclassification can be reduced by upwards of 20% in some datasets, even when detection errors are low (*e.g.*, ~ 10%). Combining these two analyses (termed detection error-corrected EMS, or DECEMS) can be a more robust way to describe how species are distributed and make inferences about the factors that may drive these patterns (Mihaljevic *et al.* 2015, 2018). Occupancy models rely on replicated data collection of species presence (or abundance) in order to jointly find the detection and occupancy probability for one or multiple species (Dorazio and Royle 2005, Dorazio *et al.* 2006) and can incorporate additional covariates that describe, for example, biotic interactions or environmental conditions (Kéry and Royle 2016, Rota *et al.* 2016).

Species classification using the EMS framework has been conducted on a wide range of organisms in a variety of different ecosystems to help describe and understand free-living metacommunities (Cottenie 2005, Presley and Willig 2010, Dallas and Drake 2014, Heino *et al.* 2015, Gascón *et al.* 2016). Parasites and other symbiotic organisms have been studied using the same framework to understand how such organisms are organized in space (Dallas and Presley 2014, Mihaljevic *et al.* 2018, Costa-Neto *et al.* 2019, Cardoso *et al.* 2020, Álvarez-Mendizábal *et al.* 2021). One study on flatworms in freshwater snails suggested parasites exhibit nestedness and suggested this distribution was predictable under a mass effects or species sorting process (Leibold *et al.* 2004, Richgels *et al.* 2013). Other parasite taxa have been linked to other EMS patterns, including nestedness (Mihaljevic *et al.* 2018, Álvarez-Mendizábal *et al.* 2021), checkerboard (Costa-Neto 2019), “Clementsian” (Dallas and Presley 2014, Álvarez-Mendizábal *et al.* 2021), and “random” (Dallas and Presley 2014, Cardoso *et al.* 2018). Many studies, however, do not explicitly account for detection error in

EMS assessments. Aside from correcting for detection errors, occupancy models have the additional benefit of considering the effect of covariates (Mihaljevic *et al.* 2015, Doser *et al.* 2022) which can provide added species-level responses to the environmental or other factors may be related to the standalone EMS pattern that describes empirical data.

Estuaries comprise distinct communities linked by dispersal across large distances. Past work in the tidal estuary Carpinteria Salt Marsh (CSM) along the west coast of California (USA) has identified a community of digenean trematodes parasites (Platyhelminthes) that are infective to a diverse range of hosts that occupy CSM and use the marine snail *Cerithideopsis californica* as the first intermediate host. In CSM, 19 identified species across 8 families maintain complex life cycles that involve two or more hosts (Hechinger 2019). Parasites form infracommunities within host individuals through shared dispersal mediated by biotic interactions (Holmes and Price 1986, Lafferty *et al.* 2008). Parasite communities have additional hierarchical structures among host individuals within a species, and among host species within a host community (the parasite component community) (Bush *et al.* 1997). As such, parasites form metacommunities within host populations and host metacommunities (Guégan *et al.* 2005, Moss *et al.* 2020). This guild of trematodes forms spatial patterns that can be described by metacommunity theory.

Community structure could change depending on the scale at which parasites are observed. For instance, interactions between parasites are typically stronger at fine spatial scales (*e.g.*, within a snail), whereas evolutionary and biographic associations are typically stronger at large spatial scales (across estuaries). We focus on the component community (*i.e.*, the community of parasite species within the population of a single host at a single site) of this system for two reasons. Component communities are considered to be long-lived associations, making them appropriate indicators of long-term processes that may affect the host and parasites over time (Poulin 2004, 2007). Therefore, we expect inference from patterns of distribution within the component community will provide valuable inference about what processes shape that group of organisms. An important caveat is that statistical inference from component community data is also more sensitive to rare species and, consequently, detection errors, as rare species are less likely to be observed in community sampling (Poulin 2004). The component community also represents a diverse group of parasites that occur very commonly in the *C. californica* host and have well-documented ecological roles and competitive interactions. Parasite-parasite competition, where parasite species directly compete with other parasite species, occurs implicitly as parasites seek access to susceptible hosts and directly within hosts for dominant access to host tissue (Kuris 1990, Sousa 1993, Kuris and Lafferty 1994). The latter form of parasite-parasite competition has an established dominance hierarchy within which parasites predictably compete when species encounter each other within an individual first intermediate host (Kuris 1990, Kuris and Lafferty 1994, Garcia-Vedrenne *et al.* 2016). Notably, trematodes also experience competition-colonization tradeoffs, as less competitive species are stronger colonizers (Mordecai *et al.* 2016). Processes at the infracommunity structure the component community, but to a limited extent (Lafferty *et al.* 1994).

Trematode species may not occur together due to uniformly low prevalence and limited host dispersal, in which case communities would be neutrally coherent. For instance, if infection is a relatively rare event, and prevalence varies little among trematode species, and host movement does not homogenize the trematode community, then communities would be sparse and scattered, with no clear patterns. Such a pattern is called a random community.

Alternatively, trematodes may not occur together primarily due to interspecific competition for shared resources, *i.e.*, competition for hosts (sensu Diamond 1975). Competitive exclusion could lead to negative coherence (a “checkerboard” pattern) in the EMS framework (Box 1). At the infracommunity level, many fewer double infections occur than expected by random chance (Kuris 1990). This competition effect can scale up to the component community so that subordinate species are less prevalent in the component community than they would be if competition did not occur (Lafferty *et al.* 1994). This could lead to spatial patterns if infection rates were patchy in space, or if infection rates were high and snail populations differed strongly in mean age, so that some populations were comprised of young snails with only subordinate trematode species and other populations comprised of old snails infected mostly with dominant trematode species. Such a pattern is called a checkerboard community (Tilman 1982).

If communities are neither random, nor a checkerboard, then some positive associations among species will arise (positive coherence). There are four types of such coherent communities, depending on turnover and boundary clumping. An exposure gradient could lead to some sites having a full complement of species, whereas other sites only have a few numerically dominant species. Assuming trematode species respond to local conditions, variation in species occurrence may be primarily due to different niche breadths or abundances. This may result in communities that are positively coherent and have less turnover than expected. For instance, trematode infections are more likely in large snails. Because snails do not lose infections, larger (older) snails are more likely to have become infected over time (Lafferty 1993), including with relatively rare trematode species. Populations comprised of younger snails seem more likely to have a subset of the trematode community. Another shared gradient may be an abiotic gradient such as the amount of organic material in sediment, which may be a proxy for host conditions or resources such as bird feces or diatoms. Such mechanisms lead to communities defined as nested subsets (all sites have the same common species, and some sites have rare species). Nested patterns may have negative clumping (hyperdispersed species loss), neutral clumping (random species loss), or positive clumping (clumped species loss), respectively.

Alternatively, trematodes may not occur together primarily due to their having different niche centroids, in which case communities may have positive coherence, and have more turnover than expected rather than less, which results in few widely distributed species, but where associations form due to environmental gradients (Gleason 1926), or a patchy environment (Clements 1916). Although the signature of different niche centroids is similar for measures of boundary clumping (negative, neutral, or positive clumping, respectively), they would result in different EMS pattern classification (evenly-spaced, Gleasonian, and Clementsian, respectively) (Figure 1). If the environment has a strong, continuous gradient, and species have different responses to that gradient, then the community will transition gradually in space. For instance, estuaries often create a salinity gradient from marine to freshwater. If trematodes (or their hosts) were to have narrow salinity tolerances, one could imagine a gradient in the trematode community from the mouth to upstream. Such a community would be called evenly spaced. On the other hand, a gradient may be related to shared intermediate or final hosts. For instance, some trematode species use crabs as second intermediate hosts whereas others use fish and others use molluscs. Thus, snail populations visited by crab-eating hosts (such as raccoons or night herons) might have different trematode communities than those visited by piscivores (such as herons and egrets) than

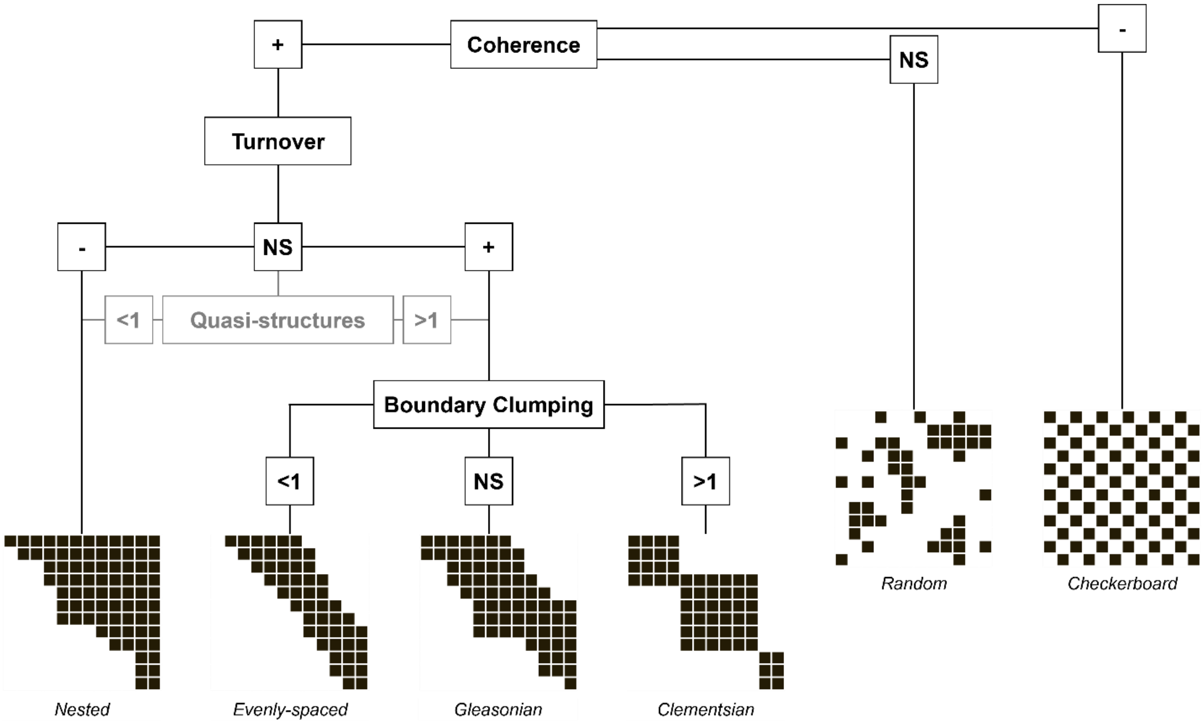


Figure 1: Decision tree for EMS patterns based on the outcome of coherence, turnover, and boundary clumping metrics. The classification of EMS pattern begins with interpretation of coherence and follows a decision tree based on the significance and sign of calculated community metrics when compared to a null model (Leibold and Mikkelson 2002). Quasi-structures may exist when turnover is non-significant but are conceptually interpreted in the same way as the traditional EMS patterns (Presley *et al*. 2010).

those visited by molluscivores (some shorebirds). If the host distributions fell along a gradient, the community would be called Gleasonian (Gleason 1926). If the host patches were non-overlapping, the community would be called Clementsian (Clements 1916).

However, turnover might not be due to gradients or niches, but due to sampling error. In such cases, all sites would have prevalent species, but some rare species would be less likely to be observed, especially where sampling effort was low. This would inflate beta diversity so that nested communities might appear to be Clementsian, Gleasonian or evenly spaced. It is also possible that sampling error would inflate boundary clumping, especially to the extent that sampling effort varies among sites.

By incorporating multi-species occupancy models to assessment of metacommunity distribution patterns with EMS, we were able to account for detection error in observational data. Using the posterior samples (*i.e.*, inferred community occurrence matrices) from a multi-species occupancy model led to different EMS pattern classification. Although observational data was dominantly classified as Gleasonian, inferred communities were dominantly nested (see Box 1 for details on these patterns). We also used simulated communities to determine if the random removal of species from nested communities inherently generated a reclassification to a Gleasonian community. The majority of simulated communities did not yield pattern reclassification to a particular EMS pattern with random species removal, suggesting using a species occupancy model reduces turnover due to error, providing more accurate insight about species occurrence in observed communities.

2.2 Methods

Our work is based on a combination of EMS pattern classification, as developed in metacommunity literature (see Leibold and Mikkelson 2002, and Presley *et al.* 2009, 2010), with species occupancy models that will correct error detection and allow for closer exploration of species-specific covariates and how they relate to identified patterns. This workflow was developed in Mihaljevic *et al.* (2015) as detection error-corrected EMS (DECEMS) and we apply this technique to a novel system on trematodes in the component community of the first intermediate host, *C. californica*.

2.2.1 Study System

CSM is a coastal estuary consisting of sub-tidal channels, mudflats, and upland habitats. Digenean trematodes (Platyhelminthes) are prevalent in CSM and infective to a wide variety of hosts at several life stages. Here, we use observational data on trematodes that use the marine gastropod *C. californica* (Haldeman) as the first intermediate host. Viable second intermediate hosts include fish, gastropods, bivalves, and other invertebrates, though some species do not use a second intermediate host (Huspeni and Lafferty 2004 and references therein). Transmission typically occurs in the environment as infective stages are released and encounter a susceptible host (Box 2). The final host is often birds, including migratory waterfowl, or mammals, and transmission may occur in the environment or trophically. In this study we focus specifically on trematodes that were identified in *C. californica* infections (Hechinger *et al.* 2017). 19 species of trematodes were identified across 8 families of Digenea in 5882 infected snail hosts out of the 11643 snails that were collected in the field (Appendix Table 1). In total, there were 5882 single infections, 386 double infections, and 13 triple infections. We counted each identified trematode species in an infected *C. californica* host as a single infection (*i.e.*, double infections were considered two singles and triple infections were considered three singles) per Lafferty *et al.* (1994), for a total of 6281 infections. Data on the occurrence of trematodes in *C. californica* hosts were completed between 2012-2014 on 27 sample dates across each season (wet and dry) of each year. We considered the wet season to be from November to April, and the dry season from May to October. During this period, 5 sample communities were created from the wet season of 2012 to the wet season of 2014 by pooling sampling records from sampling events in the wet or dry season of a particular year. 58 snails of representative sizes were collected at each of 10 sites in two habitat types of the salt marsh (tidal channels and mudflats), and sites were not consistently sampled on every sample date (Appendix Table 2), with infected snails ranging from 5 to 40 mm and the average infected snail measuring 28 mm. Snails were collected and dissected, and trematodes were identified to species level (see Hechinger *et al.*

Box 2: Parasitology of Digenean Trematodes in Focal Estuaries

The focal host of this study, *C. californica*, is found in coastal estuaries from North America (California, USA) to as far south as Piura (Peru) in South America (Keen 1971, Miura *et al*. 2010). Of the parasites infective to this organism, the most notable (by prevalence and size) is the guild of digenean trematode species (Kuris 1990, Kuris *et al*. 2008, Torchin *et al*. 2015). In the focal estuary, Carpinteria Salt Marsh (CSM), this guild of trematode parasites includes at least 19 identified species that all use a complex life cycle (Appendix Figure 1). The complex life cycle includes several life stages (eggs, miracidia, cercaria, redia/sporocysts, metacercaria, adults) and is transmitted by one or several processes to each host, generally environmental or trophic pathways. Parasites infect multiple invertebrate and vertebrate hosts throughout their life cycle, the first being an invertebrate, typically the California horn snail, which comes in contact with larval, free swimming parasite miracidia in water and substrate. Upon infection, castrate both sexes of host by altering hormones in the host reproductive system or displacing or consuming host reproductive tissue (Adema and Loker 1997). Within the snail host, redia or sporocyst stages may experience interference competition as species compete for access to the host. Parasite life stages reproduce and cercaria are shed from the host (who remains infected for life) into the environment. Some species use a second intermediate host, often a fish, and others do not, and are infective to final hosts, which include many bird species and some mammals (Huspeni and Lafferty 2004, Hechinger 2019). Cercaria may proceed directly to penetrate the final host. In some species, cercaria seek out and encyst on a second intermediate host, where they become metacercaria. Cercaria may also encyst as metacercaria in the environment. Metacercaria (in the second intermediate host or the environment) infect the final host when they are ingested. Adult trematodes develop and sexually reproduce in the final host, and eggs are shed into the environment in hosts' feces (Appendix Figure 1).

2017 for details on collection and dissection, and Hechinger 2019 for identification details).

Occupancy models were fit with covariates that described environmental conditions (*i.e.*, abiotic conditions and host conditions) as well as spatial context. Occupancy was fit to six environmental covariates predicted to relate to the true occupancy of trematode species: the percent of organic matter in soil, the percent of pickleweed (*Sarcocornia virginica*) coverage, density of birds, diversity (Shannon index) of birds, the average body size of *C. californica* hosts, and the habitat type, and each of these covariates were collected at and represent each site (Appendix Table 3). The season and year of trematode sampling and the average body size of snail hosts were used to model detection of trematodes in the occupancy model. Environmental covariates were collected at the same sites used for sampling hosts and parasites, as described in Kuris *et al.* (2008). Spatial coordinates were used to create spatially explicit models and describe the location of each sampling site within the salt marsh (Kuris *et al.* 2008).

2.2.2 Elements of Metacommunity Structure

Metacommunity data was assessed using the EMS pattern classification framework (Leibold and Mikkelson 2002) in the package metacom (Dallas 2014) in R (version 4.1.3) (R Core Team 2022). Patterns of species distributions are defined by metrics calculated from replicated presence-absence data of species occurrence across sites in a metacommunity: coherence, species turnover, and boundary clumping (Leibold and Mikkelson 2002). These metrics are calculated using a site-by-species dataset that is ordinated by reciprocal averaging, which orders interactions based on similarity of species ranges and species composition in sites, regardless of abundance (Gauch 1982, Dallas 2014).

Both the observational data and posterior samples from occupancy models (described below) were used as input for EMS classification. For both observed and inferred communities (*i.e.*, error-corrected data) we conducted EMS on the entire community of trematodes for each season (wet, dry) over a three-year period. Pattern identification in EMS uses the sign (positive/negative) and the significance of three community-level metrics to assign a pattern type (or "structure") to a community matrix (Leibold and Mikkelson 2002). Calculations of community metrics are done successively beginning with coherence. Coherence is measured by the number of embedded absences in the ordinated site-by-species matrix in a range of sites that make up the metacommunity. Embedded absences are species absences at sites they are expected to occur based on the ordination of the observational matrix (Mihaljevic *et al.* 2015). Significance is assessed by comparing the observed to simulated number of embedded absences in 1000 ordinated null matrices using a z-test. Null matrices were simulated in the metacom package using the "r1" (fixed row-proportional column) method (Wright *et al.* 1997). This null method maintains site (row) frequencies, making the null randomization robust against Type 1 errors (false positives) (Gotelli 2000) and treats species (columns) as probabilities determined by columns' marginal frequencies. This null randomization method assumes that sites differ in suitability (Gotelli 2000).

Species turnover is calculated once the number of embedded absences in species distributions has been minimized (*i.e.*, when they have been made completely coherent). Species turnover is calculated as the number of replacements of a species by another between sites, following the generalization of beta diversity in Eq. 25 of Jost (2007, Harrison *et al.* 1992). Significance of turnover is determined with a z-test by comparing observed turnover to the distribution of turnover values in 1000 ordinated null simulations.

Boundary clumping is calculated as Morisita's index (*I*) measures dispersion of species among sites (Morisita 1971) and uses a chi-square test for significance to determine if the index is significantly different from 1. Based on the outcome of three community metrics, the metacommunity of species is interpreted as one of several patterns of distribution (Figure 1, Box 1). Once EMS identifies a pattern, the user can make interpretations about species distributions based on how the data was ordinated and what environmental/other gradient was considered to be impacting the community.

2.2.3 Occupancy Modeling

Multi-species occupancy modeling was used to correct for potential detection error in the observational data used in EMS analyses, as well as to estimate species-specific covariate effects. We used the spOccupancy package (Doser *et al.* 2022) in R (version 4.1.3) for occupancy models, which uses Bayesian inference and parameter estimation and Markov chain Monte Carlo (MCMC) sampling. Using observational data, we conducted multi-species spatial and non-spatial occupancy models. While our observational data has uneven sampling across dates, species occupancy models are accepted to be robust to this variation (Doser *et al.* 2022). Both spatial and non-spatial models used the same species and covariate data (Appendix Table 3).

The observational dataset of trematodes, $y_{i,j,k}$ describes the detection or non-detection of species i at site j for each replicate sampling event k for $i = 1,\ldots,N$. $y_{i,j,k}$ was used to estimate the true presence or absence of species, $z_{i,j}$, by a Bernoulli process (per Dorazio and Royle 2005) such that

$$z_{i,j} \sim Bernoulli(\psi_{i,j})$$

where $\psi_{i,j}$ is the probability of occurrence of species i at site j (MacKenzie *et al.* 2002, Tyre *et al.* 2003), and is defined by a logit link as

$$logit(\psi_{i,j}) \;=\; x_{i,j}^{T}\,\beta_i$$

where β_i represents species-specific regression coefficients for the effect of covariates $x_{i,j}$. The superscript ^{T} indicates transposition of column vector $x_{i,j}$.

$\psi_{i,j}$ is estimated by obtaining $k = 1,\ldots,K_{ij}$ samples (sampling replicates) for each species i at each site j. The observed observational detection or non-detection was described by a Bernoulli process and dependent on the true occupancy, $z_{i,j}$, and the probability, $p_{i,j,k}$, of species i being detected at site j in replicate k, such that

$$y_{i,j,k} \sim Bernoulli(p_{i,j,k} z_{i,j})$$

where the probability $p_{i,j,k}$ is modeled with a logit link function as

$$logit(p_{i,j,k}) \;=\; v_{i,j,k}^{T}\,\alpha_i$$

where α_i represents species-specific detection regression coefficients from the site and detection covariates $v_{i,j,k}$ (Appendix Table 3). Both β_i and α_i coefficients are treated as random effects which increases precision for each species (Zipkin *et al.* 2009, Doser *et al.* 2022).

In spatial models, geographical coordinates are described by s_j for each site j in $j=1,\ldots,J$ are included in the estimation of $\psi_{i,j}$ such that $\psi_i(s_j)$ becomes

$$logit(\psi_i(s_j)) = x_i(s_j)^T \beta_i + w_i(s_j)$$

with $w_i(s_j)$ being a realization from a zero-mean spatial Gaussian process (Doser *et al.* 2022), which assumes that

$$w_i(s_j) \sim Normal(0, \Sigma_i(s, s', \theta_i))$$

where $\Sigma(s,s',\theta_i)$ is a covariance matrix and is a function of the distance between any pair of site coordinates, s and s', and a set of parameters, θ_i, that determines the spatial process using a spatial correlation function (Doser *et al.* 2022). We used exponential spatial correlation function in these models based on preliminary comparisons with other correlation function (spherical, Gaussian, and Matérn) using widely applicable information criterion (WAIC, Watanabe 2010) on the full community dataset. Exponential correlation function is described as $\theta_i = \{\sigma^2, \phi\}$ with σ^2 representing the spatial variance parameter and ϕ a spatial decay parameter. Exponential correlation is commonly assumed in ecology and suggests exponential decay of similarity or covariance of species with distance (Legendre and Fortin 1989, Bjørnstad *et al.* 1999).

Occupancy models were implemented with Pólya-Gamma latent variables (Polson *et al.* 2013, Doser *et al.* 2022) for both the occurrence and detection model components in both spatial and non-spatial models. The model used weakly informative priors, with occurrence and detection means described by multivariate normal distributions, and occurrence and detection variance described by independent inverse-Gamma distributions (Doser *et al.* 2022

Appendices). For the spatial component of the model, we used default priors which were an inverse-Gamma prior for spatial variance (σ^2) and uniform priors for spatial decay (ϕ) which follow standard recommendations for point data (Banerjee *et al.* 2015).

Models ran with twelve MCMC chains with 10,000 samples per chain, with an initial burn-in of 1,000 and a thinning rate of 1, for a total of 108,000 posterior samples. We assessed convergence by confirming the potential scale reduction factor (Rhat) was < 1.1 (Gelman 1996) and that effective sample size was > 1000 (Bürkner 2017) before using posterior samples for EMS classification. We also evaluated importance of each covariate using the density intervals of the posterior and considered covariates to be "significant" to the distribution of trematodes when the 95% highest density interval does not overlap with zero (Mihaljevic *et al.* 2018). Finally, we ran a posterior predictive check to assess how well the occupancy model represented patterns in the observed trematode data (Rubin 1984).

2.2.4 Simulation of Species Communities

We simulated 1000 site-by-species matrices from a binomial distribution that met the same conditions of the inferred communities created with occupancy modelling, *i.e.*, a nested EMS pattern, in order to evaluate if using an occupancy model led to different insight than randomly adjusting species presence-absence. Simulations included 10 sites and 20 species, similar to what we analyzed with observed trematode data (10 sites and 19 trematode species). By incorporating a species occupancy model (described above) we corrected for detection error and thus added species occurrences that were potentially missed in the observed community data. Doing so adds richness to each site. We were worried that some community measures might be affected by variation in richness rather than by species associations. So, to compare our approach with a random adjustment of species occurrence, we randomly removed species without replacement and then reintroduced EMS analyses on the resulting simulated matrices. In each trial, we randomly removed between 10 and 80% of species and tracked if the resulting frequency of EMS patterns was changed in each trial.

2.3 Results

2.3.1 EMS Observed Communities

EMS classification of observational data showed that the full trematode community primarily followed a Gleasonian pattern in 3 of 5 replicate surveys, or a quasi-Gleasonian for 2 of 5 replicate surveys (Figure 2). Turnover, the number of times one species replaced another between sites, ranged from 40 to 183 (Table 1). One survey followed a quasi-Clementsian pattern. A (quasi-) Gleasonian pattern is defined by positive coherence, positive turnover, and positive boundary clumping (Figure 1, Table 1) and suggests a gradual spatial gradient of trematode niches.

2.3.2 Occupancy Model

Of the six covariates used to predict the true occupancy of trematodes, only pickleweed was a strong predictor in the occupancy model based on the 95% credible interval in the posterior (Figure 3). Snail densities tend to be lower at sites where pickleweed is present because snails avoid shade. Trematode detection was best predicted by the size of snail hosts (older hosts are more likely to be infected), and the spatial covariance of sampling was also a strong predictor of occupancy (Figure 3). Variance of covariates (for occupancy and detection parts of the model) were high compared to means, suggesting species may be responding differently to most occupancy and detection covariates. No covariates were important predictors of occupancy or detection for individual species that were not also

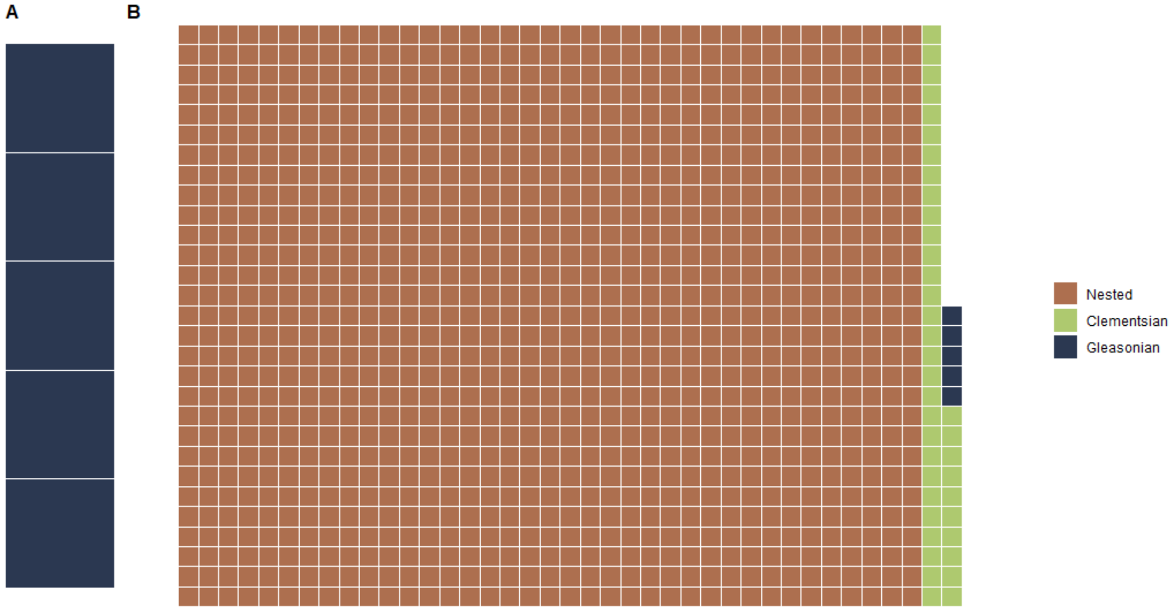


Figure 2: Frequency plot of EMS patterns between observed and inferred communities. Each square in the plot represents a single community matrix and the color of each square represents its EMS pattern. Panel A shows samples from the observed trematode data (N=5) and panel B shows inferred samples (N=2000). We include quasi-structures with traditional structures because we interpret them as ecologically synonymous (for example, all quasi-Gleasonian samples are labelled Gleasonian).

Community Type	**Embedded Absences**	**Turnover**	**Morisita's *I* index**
Observed	19*	95	1.4
Inferred	2*	6*	5.33*

Table 1: Results of EMS classification of metacommunity structure using observed community matrices and 2000 randomly sampled inferred community matrices following multi-species spatial occupancy models. Each metric represents the median value of calculations for each community (Mihaljevic *et al.* 2018). Significance compared to null models is indicated with * for $p < 0.05$.

important at the community level. The posterior predictive check suggests that the fit occupancy model adequately represented variation in species across space as most Bayesian p-values were close to 0.5 (Gelman 2013; see Appendix Table 4 for details).

We also assessed sampling error based on the change in species richness (the number of species) at each site between observational data and inferred communities from the occupancy model. Sampling error tended to underestimate the species present at a site. The

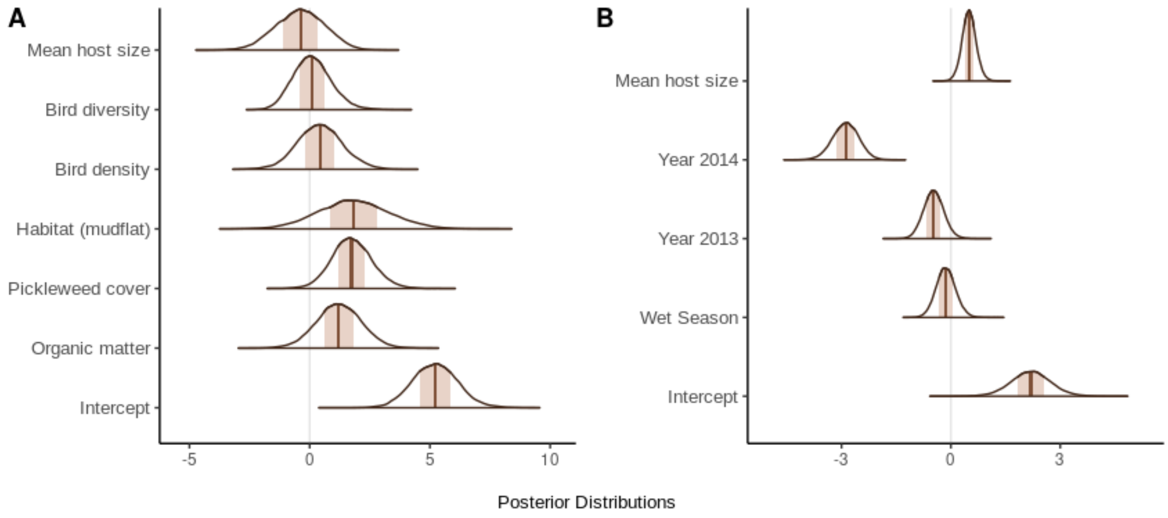


Figure 3: Density plots of the posterior distributions of occupancy and detection model covariates, showing the median (dark line) with the 90% uncertainty interval (outer interval, dark line) and the 50% interval in the inner shaded region. Panel A shows the posterior distribution for samples describing true occupancy state of the trematode community, and the intercept represents channel habitat type. Panel B shows the posterior distribution for samples describing detection probability of trematodes, and the intercept represents the dry season in the year 2012.

mean richness at a site in the observed community data was 17.8 (range: 16 to 19 species) whereas the mean estimated richness from the occupancy model output was 19 species. Species varied in the extent to which they were likely to be missed by sampling, with less prevalent species being less likely to be sampled despite being present.

2.3.3 EMS Inferred Communities

The occupancy model generated 108,000 posterior samples (site-by-species matrices) of trematode occupancy, and 2,000 of these were randomly chosen for EMS classification. Turnover ranged from 0 to 55 (Table 1). We found that EMS classification generated different patterns to that in observational data. Inferred community data were identified primarily with nested patterns (~69%) followed by quasi-nested patterns (~26%) in the trematode community (Figure 2). A (quasi-) nested pattern is defined by positive coherence, negative turnover, and positive boundary clumping (Figure 1, Table 1).

2.3.4 Simulation of Species Communities

When species were randomly removed, there was no change in the classification of EMS patterns compared to the original, simulated nested communities. With progressively more species removal, EMS patterns began to deviate from nested, but not in a particular direction. Overall, reclassification assigned either a Gleasonian or Clementsian pattern, but occasionally also a random or evenly-spaced pattern (Figure 4).

2.4 Discussion

EMS has been a useful tool for exploring and describing species distribution. The identified patterns in the EMS framework are meant to guide inference about what processes or conditions may be important for species occurrence or abundance patterns, but its susceptibility to imperfections in observational data can lead to misinterpretation (Mihaljevic *et al.* 2018). By incorporating species occupancy models as a means for detection error estimation and correction, we've been able to correct for possible detection error in the observational data on trematodes in *C. californica* hosts. This technique led to significant reclassification of the trematode community and the detection error-corrected data showed nested patterns in inferred community matrices, which were not found in observed communities.

Our initial assessment of EMS patterns in observational data suggested that trematodes were best described with a Gleasonian or quasi-Gleasonian pattern. The Gleasonian pattern is defined by positive, significant coherence, positive, significant turnover, and nonsignificant boundary clumping (Table 1). Quasi-structures, as proposed in

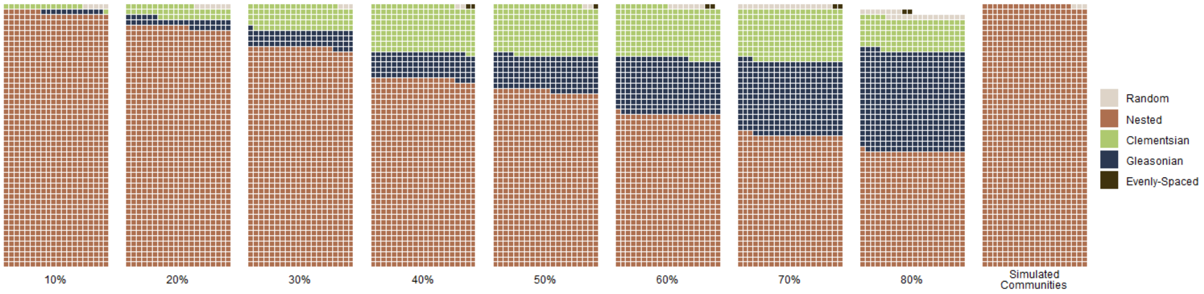


Figure 4: Frequency plot of EMS patterns in simulated communities. Community matrices were purposefully created to represent a nested pattern, and species were removed at 10-80% from the original simulated matrices.

Presley *et al.* (2010) are extensions of the original patterns described by Leibold and Mikkelson (2002). These patterns differ in that they are used to classify communities that exhibit non-significant species turnover. Quasi-structures are meant to improve insight from communities that may exhibit weaker "structuring" processes (such as environmental gradients, competition, *etc.*) than communities with significant turnover between sites (Presley *et al.* 2010). The flexibility of quasi-structures is valuable for the sake of better understanding a metacommunity and have been widely adopted in EMS literature. In our observational data we found very frequent support for quasi-Gleasonian in cases with non-significant, positive turnover, and agree that these can be interpreted ecologically in a similar way. In short, the observed communities showed patterns consistent with a gradual but inconsistent gradient in niches where sites closer in attributes had more similar communities.

Applying a multi-species occupancy model to the observational data allowed us to detect and correct for potential detection errors. Errors led to an underestimate of richness by 6%. Correcting these false negatives led to a noticeable change in the distribution patterns in trematodes based on classification with EMS. To eliminate the possibility that the reclassification of the detection error-corrected community was due only to underestimating richness, we evaluated trials of species removals in simulated communities. The original simulations had a nested EMS pattern, and removing between 10 and 80% of species and reclassifying them in the EMS analysis did not shift the majority of communities to a Gleasonian pattern. Although simulated communities were progressively more likely to be reclassified as more species were removed, reclassification at any level of species removal favored two of the most similar EMS patterns, Gleasonian and Clementsian (Figure 1), but did not strongly favor one of the two. At and above 50% removal, we do see an inclusion of more EMS patterns (Figure 4), however because the number of species occurrences becomes very small, we argue that any statistical interpretation is not advisable. For the sampling error in our data (6% missing species), the apparent reclassification of species patterns with inferred communities supports that classification changes were not likely related to removing species at random, but they are related to undersampling rare species, which inflates turnover.

Using inferred communities, we found that detection error-corrected data predominantly corresponded to nested patterns. Unlike observational samples, almost no inferred communities registered as having a Gleasonian pattern when reclassifying detection error-corrected data. Nested patterns in this framework are quite similar to the Gleasonian pattern we found in observational data, with the main difference being that turnover in nested patterns is negative rather than positive (Figure 1). This specifically indicates a change in the calculation of turnover from positive to negative, suggesting that species exhibit fewer replacements than expected in null models from season-to-season when detection error is accounted for. In nested metacommunities, species are distributed within larger ranges of another (or several) species, rather than being replaced intermittently over space. While EMS or DECEMS has not been used widely for trematode communities, one other study on freshwater snails found that trematodes also demonstrated nested distributions among pond sites (Richgels *et al.* 2013).

Coherence in the trematode community suggests that trematodes respond to a similar, latent environmental gradient. Occupancy models supported that the pickleweed cover at sites where hosts were sampled was correlated with parasite distributions (Figure 3). It is well-documented that *C. californica* densities (infected and uninfected) are negatively correlated with pickleweed (Lorda and Lafferty 2012). In empirical studies in the same system, it’s been shown that trematodes of this guild vary in density and abundance over space, and areas with fewer snails may also have more infective stages and transmission per snail (see however Buck *et al.* 2017).

EMS classification assumes that there are no errors in species detection, which for more inconspicuous organisms like parasites, can be more difficult to ensure. Error in species detection can lead to over- or underestimation of occupancy probability (Dorazio *et al.* 2011, Devarajan *et al.* 2020). Digenean trematodes in this system may be especially sensitive to detection error in instances where infection dynamics are in-progress, for example during interference competition among competing trematodes (Sousa 1983, Kuris 1990). Given the ubiquity of sampling error in community studies, the reclassification of EMS patterns using inferred communities suggests that other community classifications are likely sensitive to detection error and over-estimated turnover.

Past work has focused on how competition and environmental heterogeneity affect trematode interactions within snails. Trematode recruitment tends to be positively correlated in space (Lafferty *et al.* 1994), perhaps due to bird aggregations in the estuary (Hechinger and Lafferty 2005). Despite this positive correlation among trematodes within a sample of snails, few double infections are observed (which may create a checkerboard pattern within snails). Although competition reduces the observed prevalence of subordinate species below that expected if competition did not occur, it is not enough to create a checkerboard pattern for the trematode community at the snail population level.

Other critiques of the EMS framework have detailed similar methodological limitations of null models and ordering of species matrices for, *e.g.*, user-defined environmental gradients and the impact of these decisions on calculations of EMS metrics, particularly coherence (Ulrich and Gotelli 2013, Schmera *et al.* 2018, Presley *et al.* 2019). Although occupancy models don't explicitly correct these limitations, they do add an additional perspective to the potential influence of covariates on the distribution of species aside from their value at correcting detection error. We used a multi-species occupancy model structure in this study that incorporates environmental, spatial, and other covariates that may be important for species occurrence and distribution. Our simulated communities illustrate that using a multi-species occupancy model does not simply add species at random to account for sampling error, but it does so based on statistical distributions of species in the conditions described by covariates. While DECEMS provides a useful improvement of EMS on raw data, it does not shift the framework from a pattern-based technique for understanding species communities. Classifying species distributions does not provide a mechanistic understanding of how species create or maintain detected patterns. But DECEMS does provide an avenue for guiding further study and eliminating mechanisms that are unlikely to be occurring.

Our results indicate that trematode species are best described by a nested distribution pattern, and that empirical parasite data is sensitive to detection error. We speculate that nested distribution patterns observed in some inferred communities are a signature of the trematode community having core and satellite species. Little work has been done describing trematode distributions in a metacommunity framework, and classification provides a comparable pattern that could be emulated in other studies on parasites. This work also provides evidence for DECEMS as a way to account for detection error in data in order to avoid misclassification that may occur when using EMS alone, while considering that satellite species may be more sensitive to detection error, thereby inflating turnover. While the EMS framework alone as a pattern detection scheme may be limited in its inference capacities for some organisms, it is a simple mechanistic addition to other statistical models like occupancy models that can generate at least a comparative framework between systems and observation scales (Ulrich and Gotelli 2013, Dallas and Drake 2014). Additionally, EMS uses three well-supported metrics for assigning patterns, which may be useful for additional understanding outside of the EMS framework (Presley 2020). Moreover, EMS often presents these structures as discrete, but actual communities often lie along a continuum and pattern classification should be treated as one of a series of tools ecologists use to understand community assembly or structure (Thompson *et al.* 2020). Future studies may incorporate multiple hosts or additional covariates for a richer description of the metacommunity experienced by trematodes throughout their life cycle.

Spatial Autocorrelation in Species Distribution Models of Trematode Parasites

3.0 Abstract

Species distribution models are a powerful tool for making inferences about species presence and abundance in metacommunities. Using a metacommunity of trematode parasites infecting gastropod hosts, we explore what factors best predict trematodes at local sites. Prior work has demonstrated the significance of parasite traits, host conditions, and the physical environment in trematode spatial distributions. We highlight the role of spatial autocorrelation in predicting trematodes relative to other ecological factors. In doing so we shed light on the often-understated role of spatial autocorrelation in parasite communities and the value of considering explicit spatial effects in ecological analyses. In this system, spatial autocorrelation seems to be driven by differences in the trematode community among the three estuaries sampled.

3.1 Introduction

Early theoretical research on metacommunities has suggested that complex metacommunities often have measurable dynamical similarities (Leibold *et al.* 2004, Shoemaker and Melbourne 2016). While parasite metacommunities are increasingly appreciated in empirical and theoretical studies, there is a lot of room for insight from parasites, and particularly parasites in less-charismatic hosts. Parasites of invertebrates, while easier to sample and host to more macroparasites than many vertebrates, are neglected in many community studies (Wilson *et al.* 2015). Parasites exist in complex communities that reflect intimate interactions with hosts, the environment, and each other. From free-living species, we know that community assembly is an outcome of many spatiotemporal processes (HilleRisLambers *et al.* 2012). The assembly and distribution of local ecological communities can be described by processes that "filter" the establishment or persistence of species at different spatiotemporal scales (Cornell and Lawton 1992, Zobel 1997) meaning a subset of species are found in realized local communities (Combes 2001, Sattler et al, 2011). These filtering processes can be environmental, corresponding to abiotic conditions, or biotic interactions, or their interactions (Wisz *et al.* 2013, Kraft *et al.* 2015). Parasites experience these processes in a hierarchal way, because they are influenced by both their own responses to environmental conditions, as well as the responses of their hosts to (sometimes different) processes.

There is undeniable evidence of the importance of ecological processes in determining where parasites may be found and at what abundance. Consequentially, we expect that these processes will also serve as the best statistical predictors when we make predictions about parasite occurrence, or abundance, in a host community. But there is a rich discourse in geography and spatial ecology that indicates the spatial autocorrelation is also important for understanding what species we find locally (Fortin and Dale 2005). Spatial autocorrelation (SAC) is the phenomenon of observations (*e.g.*, environmental conditions or species occurrence patterns) closer together being more similar than observations that are farther apart (Legendre and Fortin 1989, Legendre 1993).

SAC has been a source of insight for many ecological analyses to help understand the spatial structures found in nature (Fortin and Dale 2005). But SAC also adds a complex, even problematic, component to any analysis because it challenges the independence of samples in ecological data (Legendre 1993). If samples of species or environmental characteristics are autocorrelated, there can be statistical bias in how we understand the impact of different covariates in predicting species occurrence or abundance (Fortin and Dale 2005). For example, Dormann showed that regression models had higher type I error due to SAC and can affect the distribution of residuals in model results (2007), but other work has shown that SAC doesn't necessarily create bias (Diniz-Filho *et al.* 2003). An empirical analysis on plant communities along an altitudinal gradient found that incorporating SAC even reversed some of the relationships in model results (Kühn 2007).

A guild of digenean trematodes in coastal estuaries of the west coast of North America have been richly studied over several decades. Not only does this guild exceed the biomass of other parasite functional groups in the estuaries they inhabit, but they also exceed the biomass of top predators (Kuris *et al.* 2008), suggesting an immense potential for affecting community and ecosystem dynamics. These trematodes are infective to a wide range of hosts at different life stages (Huspeni and Lafferty 2004, Hechinger 2019) and, at some stages, have direct, negative fitness impacts on hosts. First intermediate hosts of both sexes are castrated by infection by trematodes in this guild, and some second intermediate fish hosts are subject to observable behavioral modification that makes them more susceptible prey to birds that serve as final host (Lafferty and Morris 1996). Interference competition among trematodes in hosts can be fierce and may lead to fine-scale distribution patterns that reflect dominance among competing species (Kuris 1990, Sousa 1993, Kuris and Lafferty 1994) and the competition-colonization trade-off seen among trematode species (Mordecai *et al.* 2016). There is evidence for competitive exclusion by competitively dominant species, particularly in larger snails that have been exposed to parasite stages for longer. However, in snails that are younger, there may be higher species richness of trematodes (Kuris 1990, Kuris and Lafferty 1994, Mordecai *et al.* 2016). Parasites may also vary based on the transmission strategy to the first intermediate host, which in this system may occur by an egg (which is consumed by the host) or by a miracidium larva (which actively seeks out a host). Field experiments have shown that egg transmission often leads to higher per capita risk of infection to hosts and higher prevalence (Buck *et al.* 2017), and these species-specific strategies may also help predict trematode abundance. Adding to these parasite-parasite interactions are additional impacts from interactions with host species in the trematode life cycle, the physical environment, and the effect of spatial autocorrelation (see, for instance, Hechinger and Lafferty 2005, Hechinger *et al.* 2008, Mordecai *et al.* 2016). However, we still lack a consensus about the best way to predict trematode occurrence and abundance within and across host metacommunities.

The diverse metacommunity of digenean trematodes (Platyhelminthes) in their invertebrate, first intermediate host *Cerithideopsis californica* (Haldeman) exists in a series of estuaries in the North American west coast that are along the Pacific bird migration corridor. Trematodes often occur at high prevalence and density in the estuaries in which they are found, and eggs and larval stages can be transported within and between estuaries by birds, which are the most frequent final host (Kuris and Lafferty 1994, Kuris *et al.* 2008). There are currently 19 identified species of trematodes in this system (Hechinger 2019). This guild of trematodes has a complex life cycle, and while the first intermediate host is the marine gastropod *C. californica*, the second intermediate and final host can vary widely including fish and invertebrates, and birds and mammals, respectively (Huspeni and Lafferty 2004, Hechinger 2019).

Adding SAC to statistical analyses, including SDMs, is increasingly common in ecology (Diniz-Filho *et al.* 2003, 2012, Fortin and Dale 2005). In the well-studied estuarine metacommunity we focus on in this study, SAC has not been incorporated into predictions about trematode presence or abundance. We hypothesize that the occurrence and abundance of the guild of trematodes infecting *C. californica* hosts is determined by a combination of environmental, spatial, and species traits. We attempt to determine which factors are most influential on trematode occurrence and abundance with explicit consideration of the component community in the system as described at local sites across three estuaries. Using SDMs fitted with Bayesian inference we identified which parameter(s) best explained variation in species occurrence and abundance. We consider covariates that described environmental conditions and species traits in SDMs that either did or did not include SAC. Broadly, we found that while environmental characteristics were strong predictors of trematodes in sample sites, the importance of those predictors was diluted where SAC was explicitly added to the SDM.

3.2 Methods

We explored the variation in trematode occurrence and abundance using species distribution models. Multi-species SDMs extend the framework of single-species distribution modelling to consider the responses of multiple species in a community to their environment and to biological interactions to improve inference about species that are interacting or coexisting in a community (Zimmermann *et al.* 2010, Guisan and Rahbek 2011). All analyses were performed in R (version 4.1.3) (R Core Team 2022). This study uses the brms R package to build and fit SDMs, using Bayesian inference which offers flexibility for handling non-independence of hierarchal and nested relationships in ecological community data (Bürkner 2017).

3.2.1 Study Sites and Observational Data

We used published observational data on trematodes identified in the first intermediate host, *C. californica*, collected from three tidal estuaries on the Pacific west coast of North America (Figure 5). Carpinteria Salt Marsh (CSM—Carpinteria, California, USA) is a coastal estuary and natural reserve with 0.93 km^2 of upland, wetland, and sub-tidal habitat. Estero de Punta Banda (EPB—Ensenada, Baja California, México) is a 16.4 km^2 coastal estuary and Ramsar site. Bahía Falsa (BSQ—San Quintín, Baja California, México) is a coastal estuary situated in San Quintín Bay. All three estuaries are situated along the pacific migration corridor and are used by migratory birds. Within these three estuaries, trematodes were sampled once in January 2012 at a series of 54 sites that were each roughly 25 m^2 among the three estuaries (Figure 5). At each site, *C. californica* were collected and brought to a lab for dissection and trematode identification (if individual hosts were infected) (Appendix Table 5; see Kuris *et al.* 2008 for details on sample collection and processing). All identified infections were not distinguished with respect to whether they came from the same or different individual snails, meaning rare cases of double or triple infections in a snail were treated as two or three single infections (per Lafferty *et al.* 1994).

Several studies produced field data that served as either trait or environmental covariates in analyses (Appendix Table 6). Trait covariates included the transmission strategy of each trematode species (*i.e.*, whether the egg or miracidia larval stage was infective to the

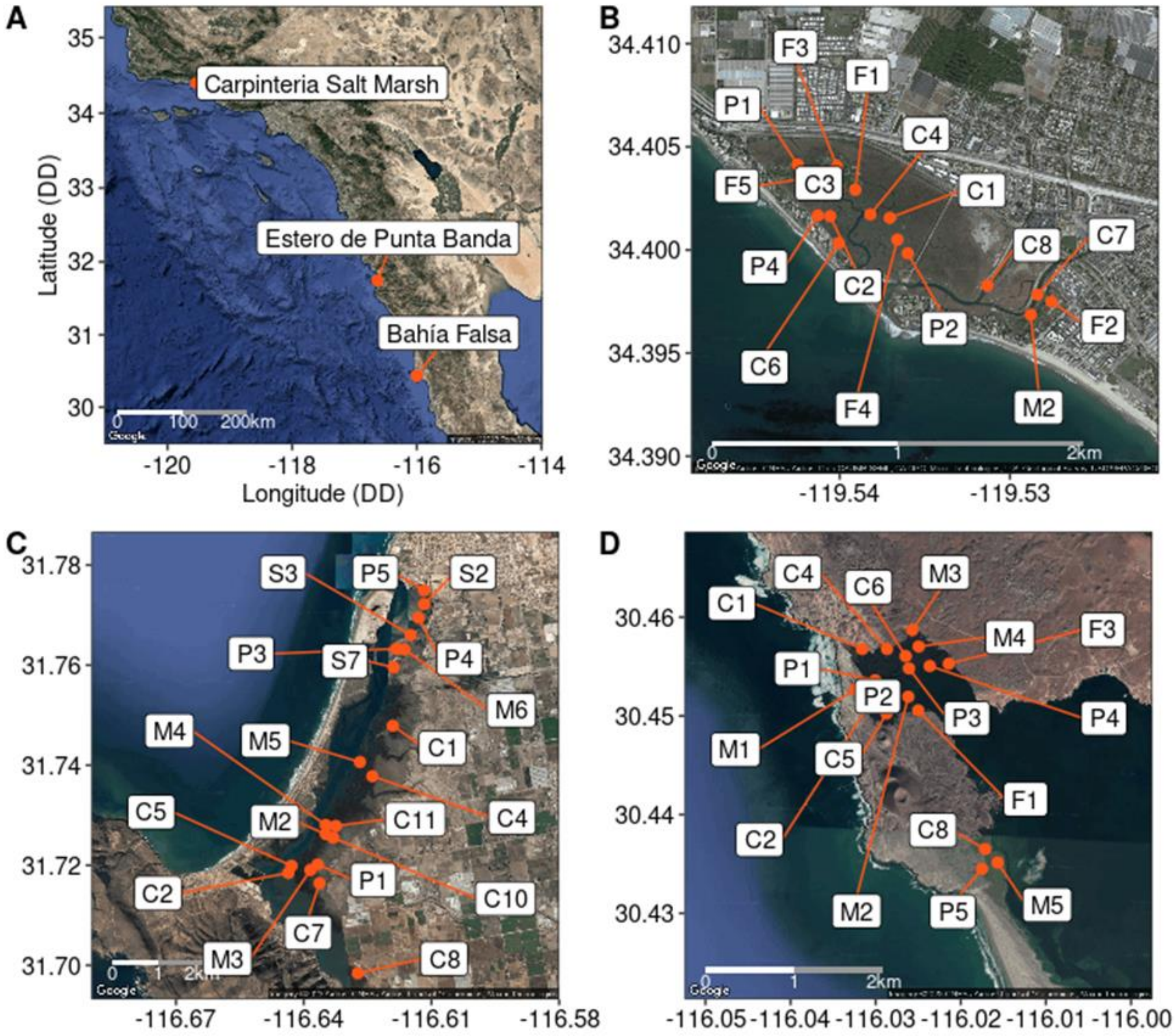


Figure 5: Map of estuaries and sites described in this study. Panel A shows the three focal estuaries in California and Baja California. Panel B, C, and D show site locations in Carpinteria Salt Marsh, Estero de Punta Banda, and Bahía Falsa, respectively. Site names in Panel B-D indicate habitat type of each site, where "C" (*e.g.*, C1) represents subtidal channels, "P" represents pans, "M" represents marshes, and "F" represents mudflats. Latitude and longitude axes are shown in decimal degrees.

first intermediate host) and the rank of each species in the dominance hierarchy observed in this guild of trematodes (Appendix Figure 2). Transmission strategy data was collected from Buck *et al.* (2017) and dominance rank was collected from Hechinger (2019). We considered environmental covariates that described both the abiotic and biotic (host) conditions for trematodes. These were collected at sample sites and included the average biomass of *C.*

californica at each site, the average size of infected *C. californica* individuals, the diversity of birds at each site (based on the Shannon index, H), the density of birds at each site, the percent cover of pickleweed (*Sarcocornia virginica*), the percent organic matter in sediment, and the habitat type. Details of how these field measurements were collected are available in Kuris *et al.* (2008).

3.2.2 Species Distribution Modeling

We fit Bayesian species distribution models using the brms R package (Bürkner 2017) to evaluate the influence of environmental conditions on species occurrence, as well as the contribution of parasite traits on occurrence, using abundance data on trematodes in *C. californica* hosts collected in the three focal estuaries. All models were stacked SDMs, which consider the spatial distribution of each trematode species in the community without explicitly considering species co-occurrence (Guisan and Thuiller 2005, Zurell *et al.* 2020). Trematode occurrence was described as the number (abundance) of infections in sampled hosts at each sample site by trematode species. Models used a combination of environmental, trait, and spatial covariates hypothesized to correlate with the presence and abundance of trematodes, as well as a random effect representing the trematode species. In total, two models were created with or without a spatial autocorrelation term. In each model iteration, environmental and trait covariates were modeled as fixed effects. All continuous covariates were scaled before use in SDMs, and we confirmed covariates were not correlated with Spearman's rank correlation coefficient (Spearman 1904). The model that contained explicit spatial information used a spatial conditional autoregressive (CAR) term to model spatial autocorrelation between sites by calculating the Euclidean distance between the latitude and longitude of sites. This term accounts for SAC in the trematode abundance data while controlling for all other covariates, including any of their inherent SAC. Each model was fit with a zero-inflated Poisson distribution, which accounted for both presence and abundance of the observational data. We also tried models with lognormal-Poisson, Poisson, and normal distributions but excluded these options due to poor convergence and fit.

We used weakly informative priors for environmental and trait covariates on a normal distribution ($\mu = 0$, $\sigma = 1$) and represented species abundance on a zero-inflated Poisson distribution for all models. Models sampled the posterior from 4 Markov Chain Monte Carlo (MCMC) chains and each chain ran for 5000 iterations with 400 samples as burn-in. This resulted in 5000 posterior samples per chain, or 20400 posterior samples per model.

We confirmed convergence of each model if the potential scale reduction factor (Rhat) was < 1.1 (Gelman 1996) and effective sample size was > 1000 (Bürkner 2017). We used posterior predictive checks to assess the distribution of the observed and posterior samples, and Bayesian R^2 (Gelman *et al.* 2019) to explore model performance. Model comparison was performed using model WAIC (Watanabe 2010, Devine *et al.* 2023). We interpreted the importance of model covariates from the best performing models using posterior distributions by identifying cases that had a 95% credible interval that did not include zero.

3.2.3 Non-Metric Multidimensional Scaling

To further explore the relationship between spatial autocorrelation and differences between estuaries, we used non-metric multidimensional scaling (NMDS) to see if there was clustering of sites in their respective estuaries, and if clustering reflected the geographic location of each estuary. CSM is the northern-most estuary and NMDS is a nonparametric, rank-order ordination technique to visualize patterns of trematode abundance at each site in low dimensional space. We imposed two dimensions using the metaMDS function in the vegan package in R (Oksanen *et al.* 2022). Because sampling sites were clustered within three broadly separated estuaries, an estuary effect will necessarily result in spatial autocorrelation. Thus, with three estuaries, it may be difficult to separate the effects of differences among estuaries and the effect of distance per se. If communities are primarily affected by spatial distance, then we would expect the northernmost site (CSM) to be most different from the southernmost site (BSQ), with the middle site (EPB) intermediary.

3.2.4 Autocorrelation of Environmental Data

While we included a term to test for SAC as a predictor of trematodes, we also wanted to determine if environmental data themselves were spatially autocorrelated. We used Moran's *I* (Moran 1950) to test for SAC in the environmental covariates (Appendix Table 6) using the inverse Euclidean distance matrix of sample sites to consider distance between all sites, as well as a binary matrix that indicated whether sites were in the same estuary or not. We used the R package ape (Paradis and Schliep 2019) to calculate Moran's *I* using the method described in Gittleman and Kot (1990).

3.3 Results

We predicted the abundance of 18 species of trematodes in the 54 sampled sites at three focal estuaries (CSM, EPB, and BSQ). In the SDM defined by environmental and trait data only, we found that most environmental covariates were important predictors of trematode occurrence and abundance. Trematodes were positively associated with the mean snail biomass within a sample site, the density of birds in a site, the amount of pickleweed cover, and were observed more in the mudflat and tidal channel habitat types. Trematodes were observed less in the marsh habitats, and were negatively correlated with the mean host size, bird diversity in a site, and the amount of organic material in sediment. The dominance rank and the transmission type were not significant predictors (Figure 6).

When we accounted for spatial autocorrelation in the SDM using the Euclidean distance between sampling sites, we found that the credible interval for the environmental and trait covariates were not significant in predicting trematodes. Instead, spatial autocorrelation between sites was significantly, positively associated with trematodes (Figure 6, Appendix Table 7). The WAIC for the SDM without spatial information was 10904 and was 9929.8 for the SDM with spatial autocorrelation. The Bayesian R^2 for the non-spatial SDM was 0.136 and was 0.247 for the spatial SDM. Both models also had appropriate posterior predictive check distributions, as indicated by the similar distribution of the observed data and the posterior samples (Appendix Figure 3).

The NMDS had noticeable clustering of sites in their respective estuaries (Figure 7). However, contrary to the expectation of a spatial effect, there was no clustering that reflected the geographic location of estuaries along the Pacific west coast, for example with clusters ordered by latitude (Figure 5). Each cluster of estuaries was approximately equidistant to each other (Figure 7).

Some of the environmental covariates we tested did have significant SAC based on our calculations of Moran's *I* (Appendix Table 8). We found that the amount of organic matter in sediment, the pickleweed cover, and the mean snail size at a site were positively spatially autocorrelated ($p<0.05$) when we considered the Euclidean distance between all 54 sites. When we considered whether sites were in the same estuary or not, we found that organic matter, pickleweed cover, and mean snail size were negatively autocorrelated ($p<0.05$).

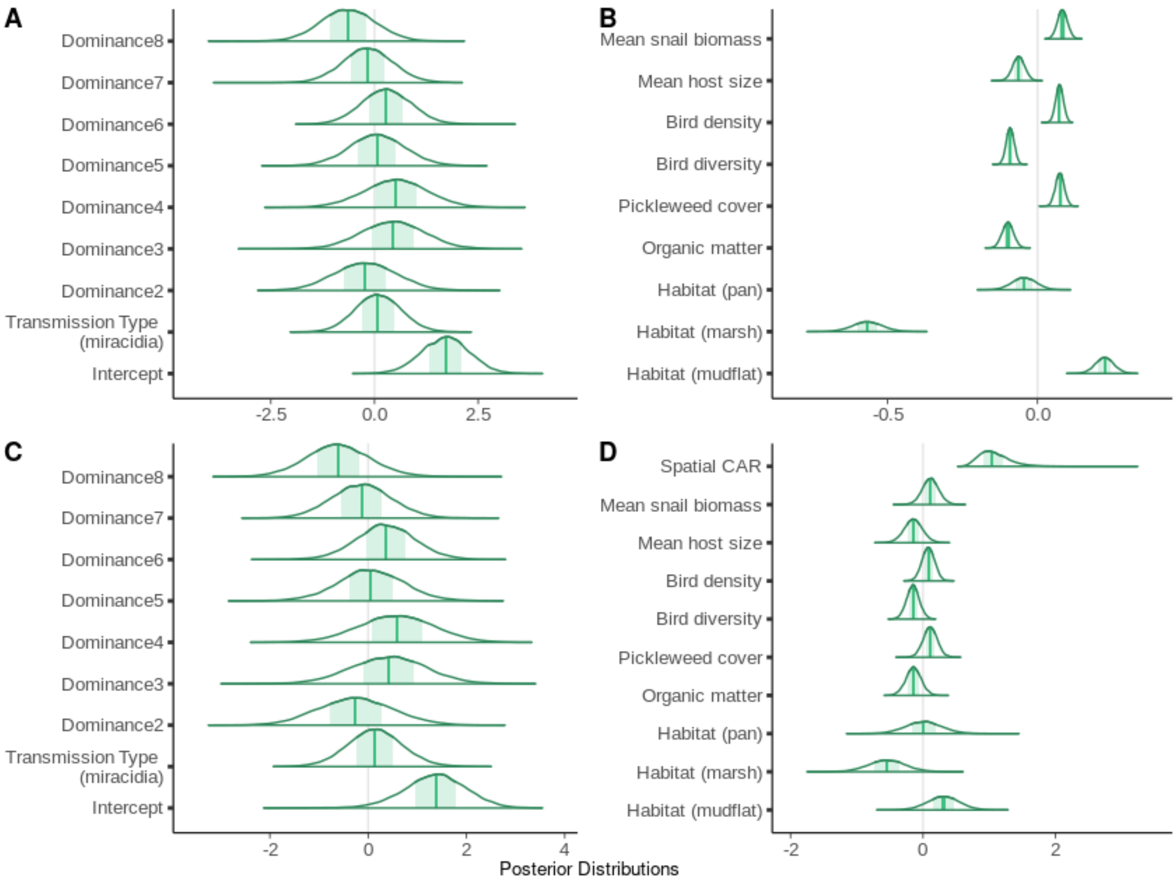


Figure 6: Density plots of posterior distributions of model covariates, with the median (dark line) with the 90% uncertainty interval (outer interval, dark line) and the 50% interval in the inner shaded region. Panel A and B show the posterior distribution for the non-spatial model, and Panel C and D show the posterior distribution for the spatial model. Panel A and C show distributions for the intercept (which is the trematode with dominance rank 1, egg transmission, and the tidal channel habitat type) and for trait covariates. Panel B and D show the environmental covariates, and Panel D also shows the spatial autocorrelation term.

3.4 Discussion

Species distribution models fit with Bayesian inference are an increasingly popular method in ecology for exploring the impact of different parameters and conditions on species presence and abundance. In fact, SDMs are now one of the most common and flexible tools for investigating the processes that drive species community assembly and patterns (Guisan and Thuiller 2005, Elith and Leathwick 2009). We present two SDMs that describe trematode communities across local sites that highlight the impact of environmental, trait, and spatial

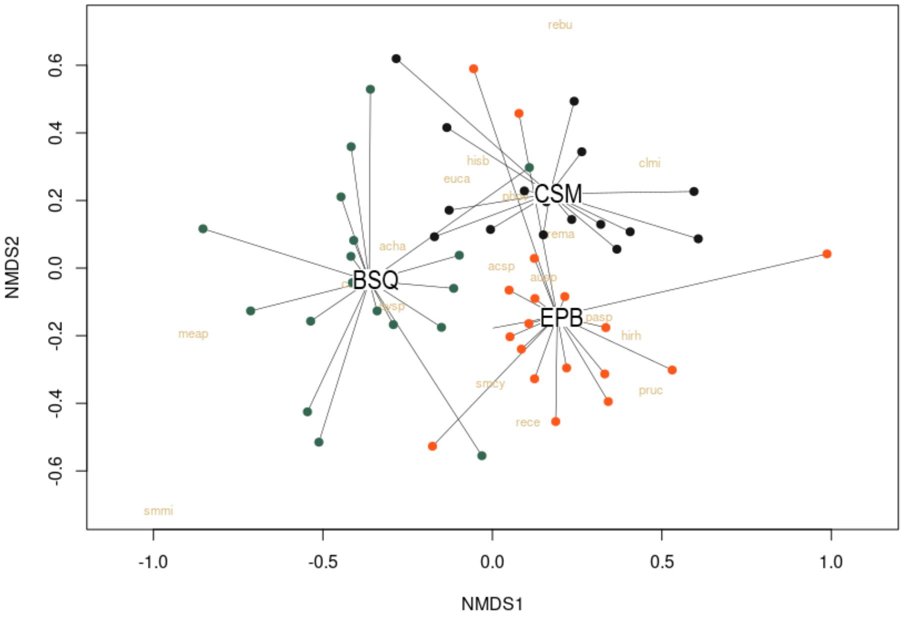


Figure 7: NMDS plot of trematode abundance at each of 54 sites. Species abbreviations (in grey) reflect those in Appendix Table 1. Each site is named by the abbreviated estuary name (BSQ, CSM, and EPB) and the site name as in Figure 5. The horizontal axis indicates the primary ordination axis of the NMDS. Site names are colored according to the estuary in which they are located.

information on presence and abundance from field data. While the environmental and trait covariates explained some variation in trematode species, spatial autocorrelation among sites was the driving predictor of trematodes. Since we used a Bayesian approach, we were able to make inference about this system from probability distributions of the posterior which allowed us to better assess the relative importance of SAC in our models.

The first model we described did not explicitly account for SAC and used only environmental and trait data to make inference about the trematode metacommunity. Environmental characteristics ended up being strong predictors of trematodes across sites in the metacommunity (Figure 6). All host characteristics for both snails and birds (the final host) were significant for trematodes when SAC was not considered in the model. We found that the mean snail biomass within a site (analogous to snail density) and bird density within a site were both positively associated with trematodes. This is not a surprising finding, as we generally expect that more snails should mean more trematodes, and similarly that more final hosts, which shed trematode eggs and larvae that are infectious to snails, would also mean more trematodes (Huspeni and Lafferty 2004, Hechinger and Lafferty 2005, Buck *et al.* 2017). These patterns have been observed in other trematode communities as well. For example, Smith found that trematodes were more prevalent in first intermediate gastropods in areas with higher bird density in mangroves (2001).

Contrary to our expectations, trematodes were negatively associated with the mean size of snail hosts as well as the diversity of birds (based on Shannon index) (Figure 6). Observational data has widely supported that snail body size is a close analog to snail age, and that older snails should harbor trematodes more often being that they've had more time to be exposed and infected (Kuris 1990, Sousa 1990, Lafferty *et al.* 1994). This is a trend often seen within sites, however in our model we compare mean snail body size between sites. While these results may seem counterintuitive, we expect that this is a finding conflated by factors like snail density or snail growth rates. For example, snails may grow larger and live longer in sites with less overall trematode abundance, leading to a negative correlation between mean snail size and trematode counts (Sousa 1983, Lafferty 1993).

A higher prevalence of trematodes in the mudflat and tidal channel habitats than in the marsh habitat is typical in field observations for these trematodes and corresponds with where bird final hosts are more common (Kuris and Lafferty 1994). The positive association between trematodes with pickleweed cover and negative association with the amount of organic material in sediment (Figure 6) is less obvious. *C. californica* density tends to be negatively associated with pickleweed cover because they avoid shade, and positively associated with organic material that reflects resource availability (Lorda and Lafferty 2012).

This might arise because per-capita infection rates can decline with snail density due to the encounter dilution effect (Buck *et al.* 2017).

Trematodes were spatially autocorrelated indicating that trematodes (presence or abundance) at a given site are similar between (*i.e.*, conditional on) the trematodes at neighboring sites (Besag 1974). Indeed, we found that some environmental characteristics were positively autocorrelated as a function of Euclidean distance between sites and that sites between estuaries were negatively autocorrelated (Appendix Table 8). This suggests that drivers of trematode communities also have spatial autocorrelation and that sites within an estuary are more similar to one another than sites across estuaries and that sites within an estuary are more similar if they are closer together. Our SDM, however, indicates that spatial autocorrelation exists in the trematode community regardless of other predictors we considered, including their inherent SAC (or lack thereof). This suggests that spatial distance is a strong predictor of trematode abundance among sites. Additionally, SAC of trematode abundance may also indicate that trematodes differ between estuaries more than within estuaries. Because data were collected between three estuaries that are many kilometers apart, and therefore clustered, differences between estuaries may be a strong driver of this outcome (Figure 6).

An NMDS of trematode abundance across the 54 sites suggested that SAC may be driven by differences between estuaries (a large observation scale effect) (Figure 7). However, we expected that a strong signature of an estuary effect would reflect the geographic locations of each estuary, meaning BSQ and EPB would be clustered more closely together than each would be to CSM, and that the clusters would be ordered by latitude (Figure 7). Instead, clusters of each estuary were approximately equidistant to each other. We interpret this to mean that the SAC in trematode abundance is likely explainable by an estuary effect rather than solely by physical distance between sites. Future work is needed to more closely explore this in more detail and quantify between- versus within-estuary effects as it relates to SAC of this trematode guild.

These findings tell us about the spatial structure of trematode presence and abundance, but also inform what we can understand about the observation scale of the data. All of the data in this study reflects 25 m^2 sample sites (N=54) that are randomly located across three estuaries (Figure 7). These estuaries are largely connected by dispersal of bird final hosts which also shed trematode eggs and larvae. At the spatial scale of the site, we find that there is a strong, positive autocorrelation of trematode occurrence and abundance. Other systems that have quantitatively compared SAC at different observation scales have found that broader scales (such as an entire ecosystem or region) often demonstrate low or negative SAC relationships within ecological communities (Collingham *et al.* 2000, Chevalier *et al.* 2021, König *et al.* 2021). Indeed, a review of parasite spatial patterns indicates that parasites readily exhibit small spatial autocorrelation in field data scale, even when entire the study area is <0.01 km^2 (Albery *et al.* 2022). Studies on free-living communities have found that deterministic processes, such as the physical environment, better predict species associations and distribution at broader observation scales (Chase and Myers 2011, Chase 2014, König *et al.* 2021). One explanation for this scale effect on SAC is that endogenous processes, that is processes that occur in the interaction between (or within) organisms such as competition or dispersal, are more apparent on communities at smaller spatial resolutions (Dormann 2007, Dormann *et al.* 2007) where the factors that drive communities are also spatially correlated.

We found that trematode traits were not strong predictors of trematode abundance. Experimental field work has shown that trematodes in this guild exhibit different infection risks in *C. californica* hosts depending on whether transmission occurs by egg or larva (miracidium) stages, a species-specific trait (Buck *et al.* 2017). While trematode infection risk may vary by transmission type, this may not describe trematode occurrence and abundance across the metacommunity at the scale we've used here. Similarly, there is clear support for variation in spatial distribution of dominant and submissive species based on the competitive dominance hierarchy in this system (Appendix Figure 2; Mordecai *et al.* 2016). This hierarchy creates a competition-colonization tradeoff, where more dominant species generally disperse less widely than more submissive trematode species. Mordecai *et al.* showed that while this tradeoff contributes to trematode coexist, additional mechanisms likely explain trematode abundance in this metacommunity, which may include SAC or other sources of spatial heterogeneity (2016).

In parasites, the range at which SAC is evident and the strength of SAC can also be affected by host characteristics. For example, Albery *et al.* found that the signature of SAC in statistical models varied by parasite taxonomic group and can be informed by both abiotic and host characteristics (2019). Work on human helminths in Brazil found that while all parasites had strong SAC, *Schistosome mansoni* was spatially correlated with water sources used by human hosts while human hookworm was less influenced by environmental characteristics and exhibited SAC at much finer resolution (Brooker *et al.* 2006). We did not explore the SAC of snail or bird hosts in this system, however this could also be a mechanism that contributes to SAC in trematodes. For instance, Koenig found that in California (USA) migratory birds exhibited stronger SAC than resident bird species, but that the strength of SAC was mediated by season, demographic processes, and the physical environment, and varied by observation scale (1998). Therefore, further exploration of SAC in parasite systems at different spatial scales in tandem with other ecological processes will be essential to deepening our understanding of the causes of SAC in parasites.

We present an example of how SAC can affect the model predictions of parasite metacommunities using two Bayesian SDMs. Clearly SAC has a strong role for understanding trematodes and probably reflects the smaller scales at which biological and environmental drivers occur. This work also shows that the scale of sample sites in these data show a positive SAC pattern among trematodes. At broader observation scales, SAC may be weaker or even negative, and other covariates such as the physical environment or traits may be more important predictors. SAC is often considered a nuisance in ecological analyses, but SAC does not necessarily create bias in statistical analyses (Diniz-Filho *et al.* 2003). Research that doesn't explicitly account for SAC are violates an important statistical assumption about the number of independent samples available in data, which can lead to statistical significance (as we found in our SDM without SAC) when there is actually not enough statistical power to detect an effect. While we don't a suggest that past work on this build of trematodes is invalidated, our work does emphasize that SAC is an essential and interesting phenomenon that can improve what we understand about a metacommunity and add detail to aspects like what information we can gain at different spatial scales (Diniz-Filho *et al.* 2003, Chevalier *et al.* 2021). This work adds additional detail and context to the long research history of this system and how SAC can improve predictions about parasite metacommunities more generally.

A Theoretical Model and Case Study Analysis on Encounter Dilution and Infectious Stage Depletion in Host-Parasite Transmission

4.0 Abstract

We investigate a theoretical model that describes the occurrence and conditions of the encounter dilution effect in disease ecology. Encounter dilution in host-parasite systems is defined by a negative density dependence between host density and per capita parasite transmission, and here we especially focus on how this effect arises due to infectious stage depletion in environmentally transmitted parasites. We develop a theoretical model that can be adapted to a variety of host-parasite communities, and we produce a case study with a well-studied digenean trematode community to test this model. Our model results produced evidence for encounter dilution that matches past field and experimental findings, demonstrating that higher host densities result in lower per capita infection risk but higher parasite recruitment. This model provides a general mathematical framework for considering encounter dilution in host-parasite systems with limited, environmentally transmitted infective stages, offering insights into the dynamics of these systems.

4.1 Introduction

A key challenge in disease ecology is to understand the transmission dynamics of parasites with complex, multi-host life cycles (Buhnerkempe *et al.* 2015). Parasites that use a single host species are coupled in that the density of parasite stages that are infective to the host are directly dependent on host density and parasitism prevalence. For example, parasites that are directly transmitted tend to be more prevalent and have higher infection intensities when there are more abundant hosts, but indirectly transmitted parasites may not demonstrate the same pattern depending on transmission mode, host and parasite mobility, and other factors (Mooring and Hart 1992, Côté and Poulin 1995). Environmental transmission by parasites occurs when hosts become infected when they come in contact with a parasite that is free-living in the environment, is driven by the opportunities for hosts and infective parasite stages (such as parasite eggs or larvae) to interact, and these interactions often do not occur homogenously in a landscape. The outcome of transmission (for instance, whether parasitism is positively or negatively related to host density) for parasites with complex life cycles is therefore more nuanced (Dobson 1990).

Hosts may naturally occur at low or high densities, and basic theory generally predicts that per capita contact rate between hosts and infective stages increases with host density (Anderson and May 1979, May and Anderson 1979, Dobson 1990). Many diseases contradict this pattern, and instead show a negative relationship between host density and per capita contact rate as well as per capita infection risk (Mooring and Hart 1992, Côté and Poulin 1995, Rifkin *et al.* 2012). This is often observed in cases where parasites are limited in the environment, such as when parasite stages are not produced locally. For instance, parasites with complex life cycles may be dispersed long distances from one life stage to another if hosts travel long distances. This limitation of parasites effectively divides the number of individual parasites among the available hosts, known as infective stage depletion (Turner and Pitcher 1986, Mooring and Hart 1992, Côté and Poulin 1995). Infection stage depletion decouples the production of parasites by one host from the density of parasites that are available to infect the subsequent host in a life cycle. Evidence for the negative relationship between parasite prevalence and host density has been called the encounter dilution effect, and infective stage depletion is one mechanism that may lead to this effect (Mooring and Hart 1992). The encounter dilution effect has been best described empirically, mostly in systems with social animals infected by mobile parasites, however the relationships between host density and parasitism in field tests vary idiosyncratically and have yet to be generalized in epidemiological theory (see Rifkin *et al.* 2012 for a review). Thus, there is opportunity for further exploration to understand not only the conditions where encounter dilution occurs, but how to describe such an effect theoretically.

Spatial variation in host density makes the outcome of transmission more complex than in simple contact rate models (McCallum *et al.* 2001). In particular, spatial variation in host density can create transmission heterogeneity as encounter likelihood (*i.e.*, contact rate) with limited parasite stages changes (Mooring and Hart 1992, Côté and Poulin 1995); the possible significance of this heterogeneity for host-parasite dynamics could be large but is not well-known, especially when infective parasite stages are mobile (May and Anderson 1984, Dobson 2004, Rifkin *et al.* 2012, Roberts *et al.* 2013). For example, Buck *et al.* (2017) demonstrated in a system of intermediate hosts and environmentally transmitted trematode parasites that a non-linearity exists between host density and per capita infection risk when infectious parasite stages are limited in the environment. When host density is high, individual host risk is decreased because parasite stages are depleted among hosts. Limited, mobile parasite stages create a “safety in numbers” effect where per capita risk of infection decreases with higher host density, at least at short time scales or fine spatial scales (Buck *et al.* 2017). When host density varies naturally across space, complex transmission patterns may not be accurately represented in generic host-parasite models that do not represent host density patterns in a landscape. If host density mediates the intensity of an encounter dilution effect, then accounting for variation in host density is essential to capturing transmission. The presence of an encounter dilution effect may be lost if certain mathematical assumptions (such as homogenous mixing) are used to described host-parasite dynamics.

Considering host density in a theoretical model can also make it easier to assess transmission patterns at different observation scales, especially when density does not change linearly with observation scale. Most classic models use a mean-field approach, where transmission at a broad observation scale, such as an entire ecosystem, is based on averaged densities of both hosts and parasites. At finer observation scales, there may be observed variation in the density of hosts and/or parasites. There is a rich literature exploring the causes of parasite aggregation in host-parasite communities that challenges the idea that parasite infections appear the same in all hosts, however variation in host density can also create interesting infection patterns. Spatial heterogeneity in host density may lead to hotspots of transmission, even if parasite stages are not similarly heterogenous. If hosts occur at high densities, then both per capita host risk and parasite recruitment may increase (Paull *et al.* 2012, Streicker *et al.* 2013). Alternatively, if there is negative density dependence between hosts and infective stages, we may see the opposite occur (Mooring and Hart 1992, Côté and Poulin 1995, Rifkin *et al.* 2012, Buck *et al.* 2017). Especially when parasite stages are actively seeking hosts, the effects of encountering dense or sparse host patches can lead to different transmission consequences at a community-level.

To approach a general theoretical model for systems that exhibit an encounter dilution effect, we formulate a model that accounts for hosts and infective parasite stages. We then use the model to explore a host-parasite community of digenean trematodes and their intermediate host studied by Buck *et al.* (2017). This system was chosen as a case study for this model not only because it has been experimentally tested for encounter dilution in the parasitism of the host, the California horn snail, but also because the parasite guild and their hosts have been studied for several decades. Since this is a well-studied host-parasite relationship, we believe it is an appropriate system to speculate about how the observed ecology relates to theoretical considerations of the model produced herein.

We produce a series of model results illustrating how encounter dilution can be observed in a system of environmentally transmitted macroparasites. This is the first published model that describes the encounter dilution effect in parasite transmission mathematically. The model is described by compartmental ordinary differential equations that describe the change in susceptible and infected host density (*i.e.*, an SI model) with vital rates and transmission, and accounts for host density that may be adjusted to a range of densities. We make two important changes to other SI models, such as those used for microparasites: 1) explicitly accounting for the supply and depletion of infective parasite stages in the environment (May and Anderson 1979), and 2) incorporating host density as a determinant of how intense density dependence (that is, the shape of the transmission function) between hosts and parasitism is in the system. The proposed model suggests that encounter dilution can be mathematically described by the depletion of infective stages as a function of host density by explicitly modeling the parasite stage in the ODE system. We explore transmission dynamics and parameterize the general model using experimental data and surveys on digenean trematodes and their intermediate host from Carpinteria Salt Marsh (CSM) as a case study. Our model gives similar results to those found in field and experimental studies in the same system, and we also explore sensitivity of the model results to a range of parameter values.

4.2 Methods

We used a series of model runs to explore the transmission dynamics of infectious parasite stages and a host in conditions with different host carrying capacities. These models reflect a parasite that is mobile in the environment and hosts the become infected by environmental transmission when they encounter parasite stages. Models accounted for dynamics of susceptible hosts and infected hosts, and explicitly considered parasite searching stages with infection and vital rates. Field and experimental data were used to inform model parameters. Models were evaluated with a sensitivity analysis to consider robustness of the results.

4.2.1 Estuary Case Study

We used a well-studied estuary community of hosts and parasites to test the general model described below. Carpinteria Salt Marsh (CSM) is an estuary in Southern California, USA and is habitat for a range of wetland species, including a digenean trematode community (Platyhelminthes) consisting of approximately 19 species. This trematode guild is present within a diverse assemblage of hosts at different life stages, but predominantly infects the California horn snail, *Cerithideopsis californica*, as the first intermediate host (Kuris and Lafferty 1994, Huspeni and Lafferty 2004). *C. californica*, can occur at high densities in

CSM and become infected by environmental transmission when they encounter an egg or larval miracidia of a trematode species (Sousa 1993, Kuris and Lafferty 1994). Buck *et al.* (2017) showed that miracidia, which are released by the final host and seek *C. californica* snails, may more easily find dense patches of snails since miracidia rely on snail mucus as a cue and are mobile. *C. californica* shed cercaria larvae once infected which are not infective to other snails. Most trematodes infective to *C. californica* use a second intermediate host, which may be a fish or an invertebrate (including gastropods, crustaceans, bivalves) including *C. californica*. Birds and mammals are the typical final host (see Huspeni and Lafferty 2004, Hechinger 2019, and references therein).

From this estuary, we especially focus on the findings of Buck *et al.* which experimentally tested for encounter dilution in *C. californica* snails (2017). In a field experiment, uninfected snails were caged at sites in the estuary and exposed to natural (unmanipulated) infection processes over the course of four months. Cages on 30-cm diameter contained 1, 5, 25, or 100 snails, and after the period of exposure were collected and dissected. Aside from this field experiment, they also collected additional information on observed infections in field surveys as well as estimates of infectious stage supply by surveying final bird hosts. From these data, Buck *et al.* estimated prevalence, per capita infection risk, and parasite recruitment at different host densities (see Buck *et al.* 2017 for additional details).

4.2.2 Model Descriptions

We used a nonlinear, deterministic compartmental system of ordinary differential equations to describe the infection dynamics of a single host and an infectious parasite life stage that builds on earlier microparasite models (*e.g.*, May and Anderson 1979, Anderson *et* *al.* 1986, Dobson and Hudson 1992, McCallum *et al.* 2001). Namely, we explicitly model not only compartments for hosts, but also infective parasite stages. Explicitly including infective stages allows us to account for both supply and depletion of parasite stages. In other models, it is generally assumed that the amount of total parasite stages is directly proportional to the density of infected hosts and that the parasite does not have a complex (multi-host) life cycle. Here, we assume infectious stages are produced independently of density of hosts in part because the parasite uses a complex life cycle. The resulting system is such that:

$$S'(t) = \alpha S(1 - \frac{S+I}{K}) - \beta M S \gamma - \mu_S S$$

$$I'(t) = \beta M S \gamma - \mu_I I$$

$$M'(t) = \zeta - \beta M(S+I) - \mu_M M$$

where the host may be either susceptible (*S*) or infected (*I*) by infective parasite stages (*M*). Because we parameterized this model using *C. californica* hosts and digenean trematodes, we assumed that recovery was not possible in the model, meaning infected individuals remain infected for life and there is no recovered class of hosts. In this system, infective stages are both parasite eggs and larvae (miracidia) which are shed in feces of the final host. Vital rates assume a different mortality rate (μ) for uninfected (μ_S) and infected (μ_I) individuals, and a birth rate (α) with a logistic growth function based on total host density ($S + I$). While there is evidence of higher mortality in infected hosts, there is no evidence that parasite load affects parasite-induced mortality, so we do not explicitly track parasite load in hosts in this model, making it more typical of microparasite models. The *K* parameter acts to determine how much density dependence is observed in the system between hosts and parasites and denotes the host carrying capacity. Consistent with the symptoms of trematode parasitism in *C.*

californica, only susceptible (uninfected) individuals reproduce as infected hosts in this community are parasitically castrated (Sousa 1983, Kuris 1990). Hosts transition from the susceptible class to the infected class with a functional infection term that assumes infective parasite stages are depleted based on the density of hosts in the perceived community. This term is defined by the effective contact rate (β) as well as the current densities of both hosts classes and of parasite stages. Infection is also affected by a parameter (γ) that is the inverse of the number of stages a host must encounter to be infected (Table 2).

Free-living infective parasite stages have a constant rate of background mortality (μ_M), are assumed to be homogenous in the landscape, and are maintained by a constant input (ζ). This latter assumption indicates a small spatial scale model whereby parasites are widely distributed (in this case by infected birds that fly from estuary to estuary) and infections within the model do not thereafter lead to future infections within the observed host population (in this case, snails that move just a few meters). Thus, the model reflects the small-scale observations of the experimental system it attempts to explain. We do not manipulate parasite density to observe the effect of host density on transmission outcomes. When a host becomes infected, we also remove infective stages from the system using the same infection term to ensure parasites do not infect multiple hosts. This is notably different from familiar SI models where infected hosts can spread infections to multiple susceptible hosts. Infected hosts can continue to encounter parasite stages in the environment. This means there may be loss of a parasite from the environment, but not in the production of new infected hosts.

Parameter	**Definition**	**Value**	**Dimension**	**Source**
α	intrinsic birth rate of hosts	0.05	% hosts per month	Sousa 1983
μ_S	natural mortality of uninfected (susceptible) hosts	0.012	% susceptible hosts per month	Lafferty 1993
μ_I	natural mortality of infected hosts	0.021	% infected hosts per month	Lafferty 1993
K	carrying capacity of hosts	varies	host density	Buck *et al.* 2017 and references therein
β	contact rate of hosts by parasite stages	0.2	(month*host density)$^{-1}$	Buck *et al.* 2017
γ	likelihood of a parasite stage encountering a viable host	0.0001	% hosts*month^{-1}	NA
ζ	shedding input of parasite stages from final hosts	1000	parasite density per month	NA
μ_M	natural mortality of parasite stages	0.8	% parasite stages per month	NA

Table 2: Model parameters used in the ODE system describing host-parasite dynamics. The chosen parameters values for the model were either drawn from field and experimental data on the CSM community of trematodes and *C. californica* or were inferred when data or estimates were unavailable.

We analytically derived the stable equilibria for this ODE system such that:

$$S^* = -\frac{K(\beta M\gamma + \mu_S)}{\alpha} - I + K$$

where α and $K \neq 0$, and

$$I^* = \frac{\beta MS\gamma}{\mu_I}$$

where $\mu_I \neq 0$, and

$$M^* = \frac{\zeta}{\beta(S+I) + \mu_M}$$

where $\beta(S+I) + \mu_M \neq 0$.

4.2.3 Model Specifications

Model parameters were chosen based on field and experimental data that was collected in several studies on *C. californica* hosts in Carpinteria Salt Marsh when available (Table 2). We assumed a logistic birth rate term for hosts, and we inferred a birth rate per unit time from the average lifespan of California horn snails (Sousa 1983) as it was not directly estimated in field studies. The mortality rate of hosts was collected from field work by Lafferty (1993) for both infected and uninfected hosts. The effective contact rate (β) was inferred from experimental data in Buck *et al.* (2017) on the proportion infected hosts at given host densities, as was the input (ζ) of parasite stages per month. We took an educated guess for values for remaining parameters that have not been previously estimated in this system (Table 2).

We performed 20 time series iterations to reach a stable equilibrium at a range of host densities that included the naturally observed densities of the host *C. californica*. Initial densities were K susceptible hosts and zero infected hosts, which include the range of natural densities observed in CSM (Buck *et al.* 2017). The range of host densities (10 to 100 hosts at 5 host increments) were chosen to explore the variance in transmission outcomes that occurred at densities that would be observed naturally and beyond, to consider how intensely density may affect an encounter dilution effect. Initial values for infective parasite stages were set to 1000 individual larval stages. All analyses were performed in R (version 4.1.3) (R Core Team 2022). We used the deSolve package with the lsoda ODE solver (Soetaert *et al.* 2010) for solving the SI model which is appropriate for both stiff and non-stiff systems of first-order ordinary differential equations (see details in Petzold 1983, Soetaert *et al.* 2010).

At the end of each model run, we also calculated prevalence as the proportion of infected hosts in the final host population.

4.2.4 Sensitivity Analysis

We used Latin hypercube sampling (LHS) to assess the sensitivity of the model to our parameter choices (Blower and Dowlatabadi 1994, Helton and Davis 2003). We drew 50 values without replacement from a uniform distribution using with the maximinLHS function in the lhs package (Carnell 2022) in R (version 4.1.3) which optimizes sampling by maximizing the minimum distance between points. The parameter values drawn for the sensitivity analyses were within 0.5 to 1.5 times the parameter values in Table 2 so we could comprehensively explore parameter sensitivity. Each set of parameter values was used in sensitivity trials using the same model equations and structure (Eq. 1).

Following sensitivity trials, we used partial rank correlation coefficients to assess importance of each parameter on the model results (Iooss and Lemaître 2015). In line with experimental findings in Buck *et al.*, we compared two factors between experimental and simulated trials: prevalence, which reflects the per capita infection risk to hosts, and the density of infected snails (I) at the end of each time series, which reflects the change in parasite recruitment (*i.e.*, the number of successful infections by parasites). Partial rank correlation coefficients (PRCC) were used to measure the sensitivity of the response variable of interest to each model parameter when the effect of other parameters is removed. PRCC calculate pairwise correlation between two variables (here, a parameter and a model output variable) while controlling for the effect of others. We used the Spearman method in the R package ppcor to calculate PRCC describe the monotonic relationship between parameters and model results (Kim 2015).

4.3 Results

Our model predicted that when transmission rate is dependent on parasite stage density by the process of infectious stage depletion, infection prevalence is negatively correlated with host density (Figure 8). Model outputs showed that final parasite prevalence ranged from approximately 6% to 33% of hosts infected by parasite stages in the time series at equilibrium. Field experiments that helped parameterize the model and tested for encounter dilution found infection prevalence ranged from 16.4 – 27.5% (Buck *et al.* 2017). Simulated infection risk did not fluctuate over time but was asymptotic within the time series and reached a stable equilibrium. In the same simulations, we also found that the density of infected hosts (I) increased with greater host carrying capacity (Figure 8).

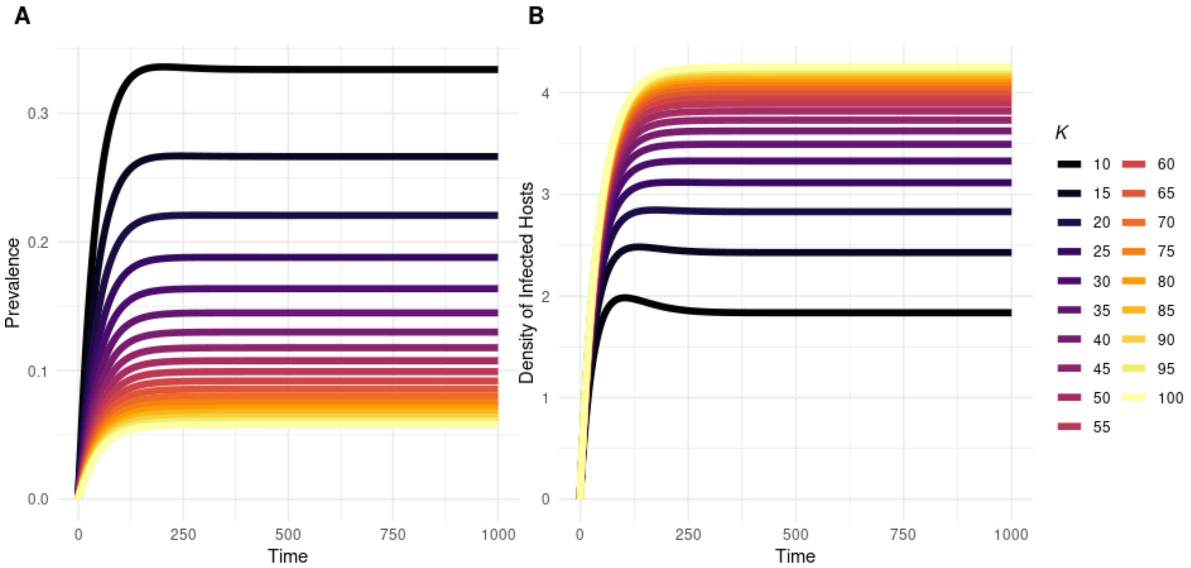
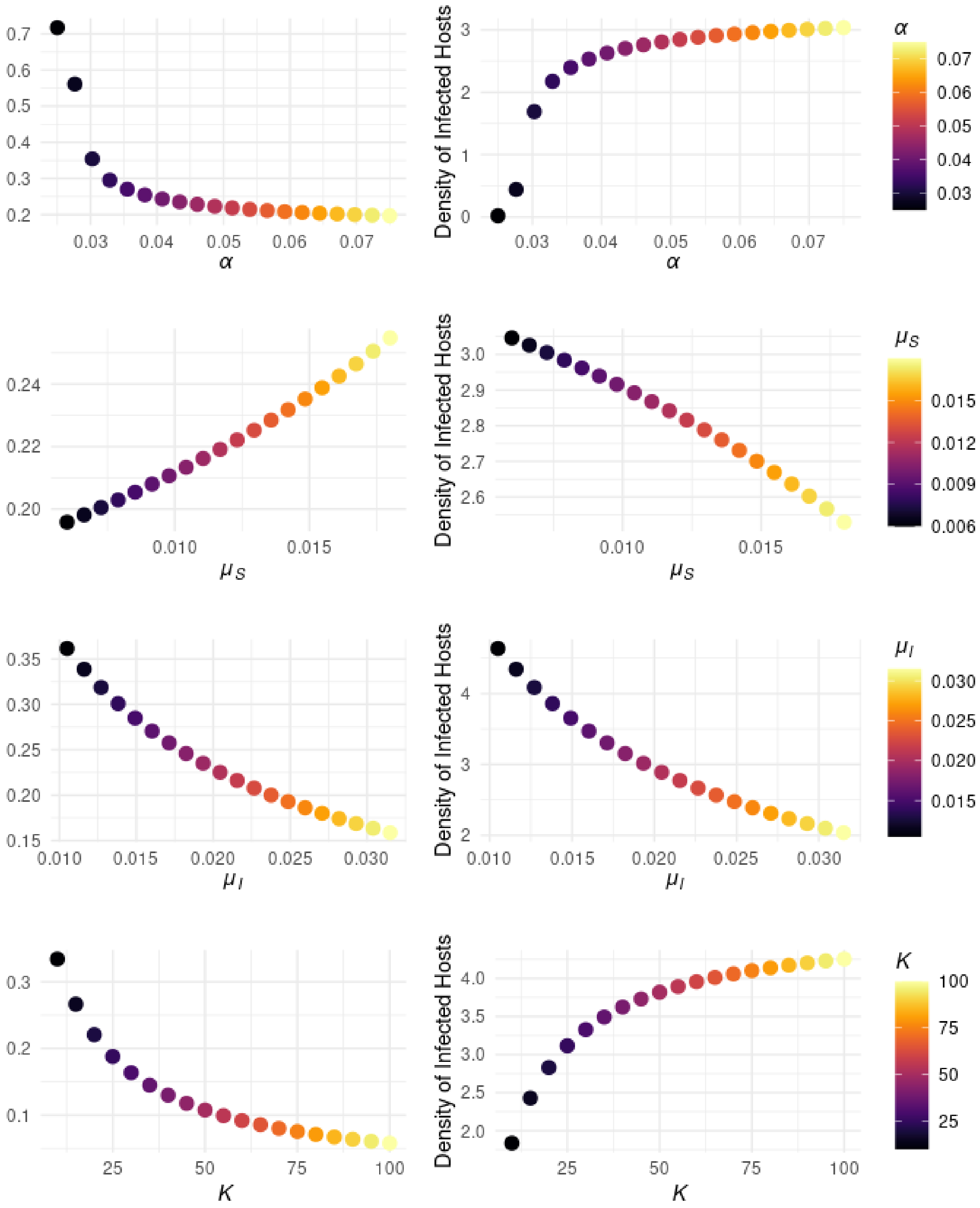


Figure 8: Time series of transmission of *C. californica* hosts by infective larval trematode stages using the model framework in Equation (1) as defined by the host carrying capacity, *K*. Panel A shows the change in infection prevalence over the course of each simulation, and Panel B shows the change in the density of infected hosts (I). Infection prevalence reflects per capita infection risk of hosts, and the density of infected hosts reflects parasite recruitment in the host-parasite community.

4.3.1 Sensitivity Analysis

We focused sensitivity analyses on the impact of model parameters on prevalence and infected host density at the end of each model run following LHS. Overall, both infection terms and vital rates were important for influencing infection prevalence and density of infected host according to PRCC (Figure 9, Table 3). Specifically, prevalence was negatively correlated with infected host mortality (PRCC: -0.64, $p<0.01$) but was positively correlated with β (PRCC: +0.38, $p<0.05$), γ (PRCC: +0.66, $p<0.01$), and ζ (PRCC: +0.37, $p<0.05$). Susceptible host density was positively correlated with host birth rate (PRCC: +0.41, $p<0.05$) and host carrying capacity (PRCC: +0.6, $p<0.01$) and negatively correlated with susceptible host mortality (PRCC: -0.45, $p<0.01$). The density of infected hosts was negatively correlated with infected host mortality (PRCC: -0.45, $p<0.01$) and positively correlated with γ (PRCC: +0.48, $p<0.01$) and ζ (PRCC: +0.55, $p<0.01$). Parasite stage density was sensitive to the parasite input constant (ζ) (PRCC: +0.43, $p<0.05$), β (PRCC: -0.79, $p<0.01$), and parasite mortality (PRCC: -0.37, $p<0.05$).


0.7
0.6
0.5
0.4
0.3
0.2
0.03
0.04
0.05
0.06
0.07
α
Density of Infected Hosts
3
2
1
0
0.03
0.04
0.05
0.06
0.07
α
α
0.07
0.06
0.05
0.04
0.03
0.24
0.22
0.20
0.010
0.015
μ_S
Density of Infected Hosts
3.0
2.9
2.8
2.7
2.6
0.010
0.015
μ_S
μ_S
0.015
0.012
0.009
0.006
0.35
0.30
0.25
0.20
0.15
0.010
0.015
0.020
0.025
0.030
μ_I
Density of Infected Hosts
4
3
2
0.010
0.015
0.020
0.025
0.030
μ_I
μ_I
0.030
0.025
0.020
0.015
0.3
0.2
0.1
25
50
75
100
K
Density of Infected Hosts
4.0
3.5
3.0
2.5
2.0
25
50
75
100
K
K
100
75
50
25

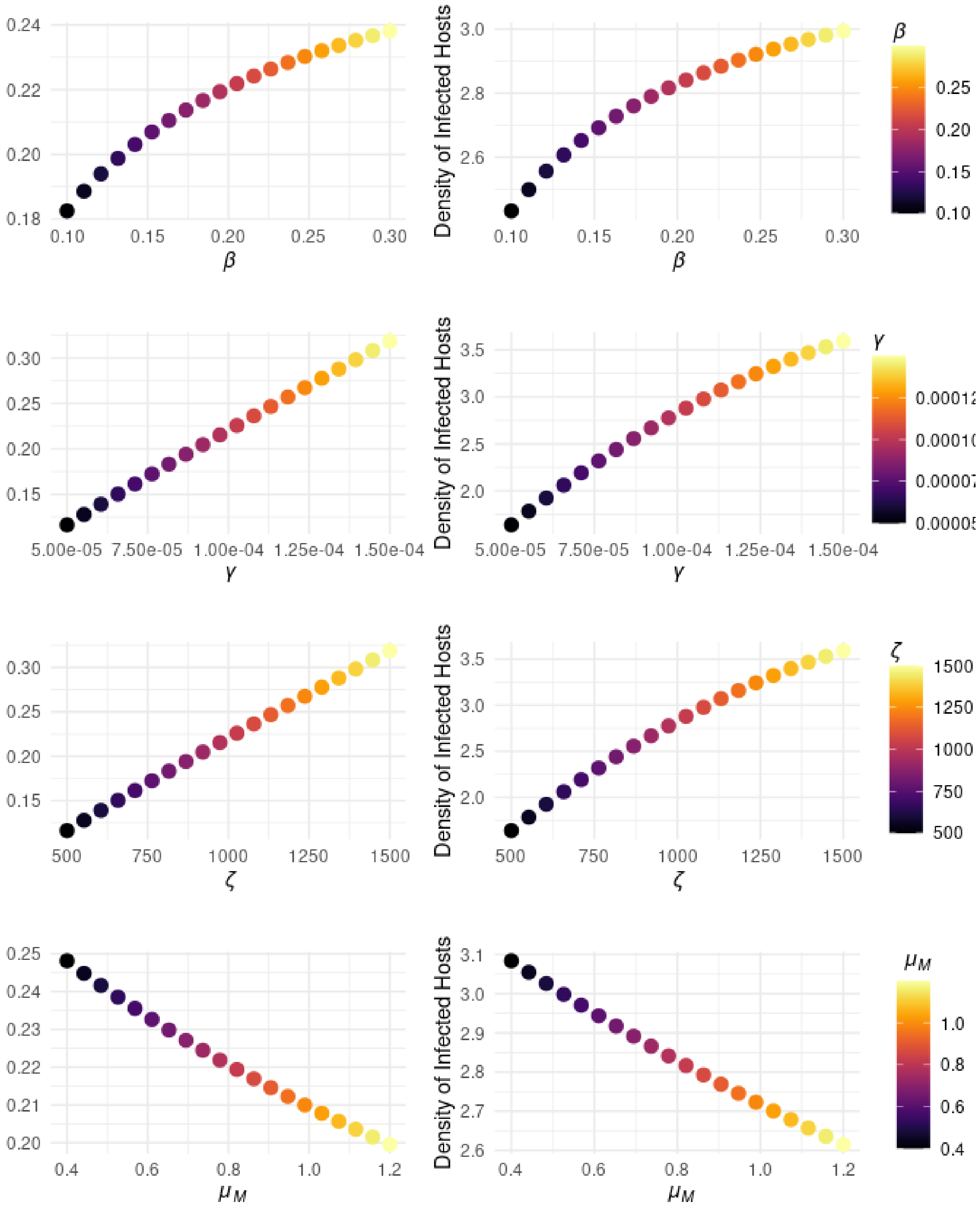


Figure 9: We extended each parameter value (Table 2) to range from 0.5 to 1.5 times in magnitude of the value used in our initial model runs when exploring sensitivity of the model outputs to model parameters. *K* was set to 20 when exploring ranges of other parameter values in simulations. Here we show the resulting prevalence and density of infected hosts for the equilibrium, calculated numerically.

	S	***I***	***M***	***Prevalence***
Correlation Coefficient				
α	+ 0.41*	+ 0.01	- 0.1	- 0.15
μ_S	- 0.45*	+ 0.11	+ 0.1	- 0.08
μ_I	- 0.18	- 0.45*	+ 0.12	- 0.64*
K	+ 0.6*	- 0.04	- 0.19	- 0.13
β	- 0.14	- 0.1	- 0.79*	+ 0.38*
γ	+ 0.17	+ 0.48*	- 0.07	+ 0.66*
ζ	+ 0.19	+ 0.55*	+ 0.43*	+ 0.37*
μ_M	- 0.07	- 0.21	- 0.37*	- 0.03
p-value				
α	0.01	0.97	0.53	0.36
μ_S	3.9e-3	0.5	0.52	0.64
μ_I	0.26	3.9e-3	0.44	1.1e-5
K	3.7e-5	0.82	0.25	0.43
β	0.4	0.55	1.2e-9	0.02
γ	0.29	1.9e-3	0.68	4.0e-6
ζ	0.23	2.7e-4	0.01	0.02
μ_M	0.65	0.2	0.02	0.84

Table 3: Sensitivity of prevalence, infected host density, and effective contact rate (β) to changes in model parameters based on 50 Latin hypercube samples (see *Methods*). Partial rank correlation coefficients were used to quantify sensitivity, including the correlation coefficient and p-value. Significant relationships ($p \leq 0.05$) are indicated with an asterisk (*).

4.4 Discussion

Models in disease ecology have yet to find a general framework for considering encounter dilution in mathematical representations of natural systems. While the number of field studies exploring this pattern is growing, we intend to present a theoretical technique for identifying and describing potential encounter dilution effects in macroparasites. We accomplish this is by making several manipulations of traditional disease models created for microparasites (Anderson and May 1979, May and Anderson 1979) to give flexibility for infective stage depletion which may occur between free-living parasite stages and a host population. A key modification is to match models to the small spatial scales at which encounter dilution is observed. By parameterizing this model with field and experimental data from a well-studied host-parasite community, we were able to 1) echo experimental findings that suggest an encounter dilution effect, 2) use this system as a case study to explore a more comprehensive set of conditions than may be possible in field experiments to describe this phenomenon, and 3) assess model sensitivity.

Model results using the CSM trematode community confirmed experimental work that found evidence for an encounter dilution effect. Model results indicated that parasitism prevalence was higher in trials with lower *K* (Figure 8). Additionally, all model outputs showed that model results leveled off at a stable equilibrium. We set a contact rate term in the model (Table 2) but do not follow the assumption used in simpler SI models that parasite stages are coupled with host density and instead are depleted as host density increases. Instead, we incorporate a density effect by assuming the observed transmission will be dependent on the density of hosts and parasite stages, as well as the likelihood of them encountering each other (McCallum et al. 2001). Additionally, we account for infective stage depletion of free-living trematode parasite stages by removing stages from the system when they encounter either susceptible or infected hosts. There is evidence from field and lab studies that hosts in high densities experience reduced per capita infection risk from trematode parasite stages as a result of this process (Ewers 1964, Johnson *et al.* 2012, Rohr *et al.* 2015, Buck and Lutterschmidt 2017, Buck *et al.* 2017). We expect that infective stage depletion is a more general phenomenon that occurs in macroparasites that produce free-living stages and are candidates for exhibiting an encounter dilution effect. Free-living, environmentally transmitted parasite stages have been suggested as a strategy benefiting both hosts and parasites in cases where host density is high, including in social animals: hosts experience a lower per capita risk of infection, and parasites enjoy a higher likelihood of successfully finding a host (Mooring and Hart 1992, Côté and Poulin 1995, Rifkin *et al.* 2012, Patterson and Ruckstuhl 2013). This pattern is echoed in our findings using data from the CSM system as a spatial encounter dilution effect.

Depletion of infective stages is essential to the encounter dilution effect and is not only a result of the limited supply of parasites shed from a host into the environment. Parasites such as the trematodes in CSM are also limited because transmission (*i.e.*, infection of the intermediate host) is spatially decoupled from infective stage production (*i.e.*, infective stages that are infective to an intermediate host) when the parasite uses a mobile final host (Buck and Lutterschmidt 2017, Buck *et al.* 2017). This can be the case of trematodes and other parasites that exhibit complex life cycles. For example, the focal trematode guild in CSM uses birds and mammals as final hosts, and other trematodes with complex life cycles also use vertebrates (including humans) as final hosts (Roberts *et al.* 2013). Vector-transmitted parasites and macroparasites have also been shown to experience parasite stage depletion and would likely show similar evidence of encounter dilution in this theoretical model (Ostfeld *et al.* 1996, Fauchald *et al.* 2007, Samsing *et al.* 2014). Our results suggest that we can account for infective stage depletion in these communities by incorporating a host and parasite density transmission term and removing parasite stages based on contact rate (β).

We also found that the density of infected hosts is positively related to host carrying capacity, K, in model time series as well as in sensitivity analyses (Figure 8, 9). This follows previous theoretical and empirical explorations about how encounter dilution and infective stage depletion can benefit parasites as there are more successful infections in hosts, suggesting improved parasite recruitment (Mooring and Hart 1992, Côté and Poulin 1995, Rifkin *et al.* 2012, Patterson and Ruckstuhl 2013). While hosts experience lower per capita infection risk when at high densities, parasites are more likely to find and infect a susceptible host.

The increased success of parasites at high host density should lead to an increase in the production of infective stages. This feedback is not considered in our model, and it is important to emphasize that an encounter dilution effect is a small-scale phenomenon that might reverse at larger spatial scales. That is, overall increases in host density would be expected to eventually increase the infective stage input term (which we keep constant). How that affects parasite dynamics is a difficult problem in need of more sophisticated models.

Infected host density and prevalence were both sensitive to a number of model parameters (Figure 9). Prevalence and infected host density naturally declined with higher mortality in infected hosts and increased with more parasite stages regardless of whether K was high or low. Prevalence and infected host density were also positively associated with γ. The γ term reflects the likelihood of infectious stages to successfully find and infect a susceptible host and was positively correlated with prevalence. Conceptually, many systems such as the CSM community explored here will have a generous inflow of infective stages from final hosts, but only a fraction of these will find physical contact with and successfully infect an encountered, susceptible intermediate host. As our sensitivity analysis shows, when the likelihood of individual infectious stages encountering a host increases, parasite prevalence and infected host density increases (Figure 9, Table 3). Interestingly, prevalence and infected host density were both sensitive to contact rate (β) but in opposite directions. Infected hosts were negatively correlated with higher contact rate, while prevalence increased with contact rate. This is an indication of infective stage depletion leading to an encounter dilution effect. Despite hosts and parasites having an increased rate of transmission and thus a higher prevalence of parasitism, the density of infected host individual still declines as parasites become depleted and reduce the per capita risk to susceptible hosts. This is further supported by the negative relationship between parasite stage density and contact rate in our sensitivity analysis (Table 3).

In the CSM system, Buck *et al.* found that the effect of host density on contact rate was more pronounced for egg transmission over larval transmission in experiment treatments (2017). In this guild of parasites, both eggs and mobile larva (miracidia) can be infectious to the first intermediate host, depending on trematode species (Sousa 1983, Kuris 1990). Buck *et al.* suspected that because mobile miracidia could actively search for hosts, the effect of host density on parasite infection was less intense at broader observation scales, despite having shorter viable period than an egg (2017). We speculate that the impact of mobility in parasite stages can be represented as a variance in the γ term, where more active stages have a higher likelihood of encountering a host. This may also extend to the experimental difference in infection between eggs and miracidia seen in different trematode species in the CSM system. While mobile miracidia larva seek out snail hosts, trematodes that are egg transmitted rely on being consumed by snails. The likelihood of an egg being consumed by a snail may be greater than the likelihood of an active larva successfully finding a host and infecting it. This relationship between parasite stage motility and transmission also aligns with experimental and theoretical expectations that transmission saturates in systems that exhibit encounter dilution as host density increases to very high densities (McCallum *et al.* 2001, Buck *et al.* 2017).

Our mathematical model provides a general modeling approach for considering the encounter dilution effect sometimes observed in environmentally transmitted macroparasites. Infective stage depletion caused by spatial decoupling of transmission and infectious stage supply is an essential component to consider for generalizing the dynamics of encounter dilution in this type of host-parasite system. By using a community of digenean trematodes infecting an invertebrate snail host as a case study, we demonstrated how encounter dilution can be represented with simple adjustments to more traditional macroparasite models. Our case study from CSM mirrors experimental findings that encounter dilution effects exist in this system result when host density is variable in space (Buck *et al.* 2017). Generally, we would expect to observe an encounter dilution effect in any host-parasite system that also had limited, environmentally transmitted infective stages. In the CSM observed here, limitation of stages came as decoupling transmission from stage production. Transient dynamics like seasonal activity of hosts or movement patterns may also generate depletion of parasite stages and thus encounter dilution patterns over space-time and are a noteworthy direction for future work (Woolhouse 1996, Atkinson *et al.* 2013, Tao *et al.* 2021). We were also able to explore a wider distribution of conditions in this case study by using a theoretical approach. In doing so, we found that the observed encounter dilution was notably sensitive to certain infection terms and vital rates. We expect that this general model could be adapted to other systems.

References

Adema, C. M. For Loker, E. S. Specificity and immunobiology of larval digenean-snail associations. in *Advances in trematode biology* (eds. Fried, B. & Graczyk, T. K.) 229–263 (CRC Press, 1997).

Albery, G. F., Sweeny, A. R., Becker, D. J. & Bansal, S. Fine-scale spatial patterns of wildlife disease are common and understudied. *Functional Ecology* **36**, 214–225 (2022).

Álvarez-Mendizábal, P. *et al.* Metacommunity structure reveals that temperature affects the landscape compositional patterns of avian malaria and related haemosporidian parasites across elevations. *Acta Oecologica* **113**, 103789 (2021).

Anderson, R. M. Epidemiology. in *Modern Parasitology* (ed. Cox, F. E. G.) 75–116 (Blackwell Publishing Ltd., 1993). doi:10.1002/9781444313963.ch4.

Anderson, R. M. & May, R. M. The invasion, persistence and spread of infectious diseases within animal and plant communities. *Phil. Trans. R. Soc. Lond. B* **314**, 533–570 (1986).

Anderson, R. M. & May, R. M. Population biology of infectious diseases: Part I. *Nature* **280**, 361–367 (1979).

Atkinson, J.-A. M. *et al.* Synthesising 30 Years of Mathematical Modelling of Echinococcus Transmission. *PLOS Neglected Tropical Diseases* **7**, e2386 (2013).

Banerjee, S., Carlin, B. P. & Gelfand, A. E. *Hierarchical modeling and analysis for spatial data*. (CRC Press, Taylor & Francis Group, 2015).

Bastolla, U. *et al.* The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature* **458**, 1018–1020 (2009).

Besag, J. Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society: Series B (Methodological)* **36**, 192–225 (1974).

Bjørnstad, O. N., Stenseth, N. Chr. & Saitoh, T. Synchrony and scaling in dynamics of voles and mice in northern japan. *Ecology* **80**, 622–637 (1999).

Blower, S. M. & Dowlatabadi, H. Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: An HIV Model, as an Example. *International Statistical Review / Revue Internationale de Statistique* **62**, 229 (1994).

Brooker, S. *et al.* Contrasting patterns in the small-scale heterogeneity of human helminth infections in urban and rural environments in Brazil. *International Journal for Parasitology* **36**, 1143–1151 (2006).

Buck, J. C. *et al.* Host density increases parasite recruitment but decreases host risk in a snail–trematode system. *Ecology* **98**, 2029–2038 (2017).

Buck, J. C. & Lutterschmidt, W. I. Parasite abundance decreases with host density: evidence of the encounter-dilution effect for a parasite with a complex life cycle. *Hydrobiologia* **784**, 201–210 (2017).

Buhnerkempe, M. G. *et al.* Eight challenges in modelling disease ecology in multi-host, multi-agent systems. *Epidemics* **10**, 26–30 (2015).

Bürkner, P.-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *J. Stat. Soft.* **80**, (2017).

Bush, A. O., Lafferty, K. D., Lotz, J. M. & Shostak, A. W. Parasitology Meets Ecology on Its Own Terms: Margolis et al. Revisited. *The Journal of Parasitology* **83**, 575 (1997).

Cardoso, T. S. *et al.* Helminth metacommunity of small mammals in a Brazilian reserve: the contribution of environmental variables, host attributes and spatial variables in parasite species abundance. *Community Ecology* **21**, 159–170 (2020).

Cardoso, T. dos S. *et al.* Helminth metacommunity structure of wild rodents in a preserved area of the Atlantic Forest, Southeast Brazil. *Rev. Bras. Parasitol. Vet.* **27**, 495–504 (2018).

Carlson, C. J. *et al.* Parasite biodiversity faces extinction and redistribution in a changing climate. *Science Advances* **3**, (2017).

Carnell, R. lhs: Latin Hypercube Samples. (2022).

Chase, J. M. Spatial scale resolves the niche versus neutral theory debate. *J Veg Sci* **25**, 319–322 (2014).

Chase, J. M. & Myers, J. A. Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 2351–2363 (2011).

Chave, J. The problem of pattern and scale in ecology: What have we learned in 20 years? *Ecology Letters* **16**, 4–16 (2013).

Chevalier, M. *et al.* Low spatial autocorrelation in mountain biodiversity data and model residuals. *Ecosphere* **12**, e03403 (2021).

Clements, F. E. *Plant Succession An Analysis of the Development of Vegetation*. (The Carnegie Institution of Washington, 1916).

Collingham, Y. C., Wadsworth, R. A., Huntley, B. & Hulme, P. E. Predicting the spatial distribution of non-indigenous riparian weeds: issues of spatial scale and extent. *Journal of Applied Ecology* **37**, 13–27 (2000).

Combes, C. *Parasitism: The Ecology and Evolution of Intimate Interactions*. (The University of Chicago Press, 2001).

Cornell, H. V. & Lawton, J. H. Species Interactions, Local and Regional Processes, and Limits to the Richness of Ecological Communities: A Theoretical Perspective. *The Journal of Animal Ecology* **61**, 1 (1992).

Costa-Neto, S. F., Cardoso, T. S., Boullosa, R. G., Maldonado, A. & Gentile, R. Metacommunity structure of the helminths of the black-eared opossum *Didelphis aurita* in peri-urban, sylvatic and rural environments in south-eastern Brazil. *J. Helminthol.* **93**, 720–731 (2019).

Côté, I. M. & Poulin, R. Parasitism and group size in social animals: a meta-analysis. *Behavioral Ecology* **6**, 159–165 (1995).

Cottenie, K. Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters* **8**, 1175–1182 (2005).

Dallas, T. metacom: an R package for the analysis of metacommunity structure. *Ecography* **37**, 402–405 (2014).

Dallas, T. & Drake, J. M. Relative importance of environmental, geographic, and spatial variables on zooplankton metacommunities. *Ecosphere* **5**, art104 (2014).

Dallas, T. & Presley, S. J. Relative importance of host environment , transmission potential and host phylogeny to the structure of parasite metacommunities. *Oikos* **123**, 866–874 (2014).

Devarajan, K., Morelli, T. L. & Tenan, S. Multi-species occupancy models: review, roadmap, and recommendations. *Ecography* **43**, 1612–1624 (2020).

Devine, S., Falk, C. F. & Fujimoto, K. A. Comparing the Accuracy of Three Predictive Information Criteria for Bayesian Linear Multilevel Model Selection. Preprint at https://doi.org/10.31234/osf.io/p2n8a (2023).

Diamond, J. M. Assembly of species communities. in *Ecology and Evolution of Communities* (eds. Cody, M. L. & Diamond, J. M.) 342–444 (Harvard University Press, 1975).

Diniz-Filho, J. A. F. *et al.* Spatial autocorrelation analysis allows disentangling the balance between neutral and niche processes in metacommunities. 201–210 (2012) doi:10.1111/j.1600-0706.2011.19563.x.

Diniz-Filho, J. A. F., Bini, L. M. & Hawkins, B. A. Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography* **12**, 53–64 (2003).

Dobson, A. P. Models for multi-species parasite-host communities. in *Parasite Communities: Patterns and Processes* (eds. Esch, G. W., Bush, A. O. & Aho, J. M.) 261–288 (Springer Netherlands, 1990). doi: 10.1007/978-94-009-0837-6_10.

Dobson, A. P. & Hudson, P. J. Parasites, disease and the structure of ecological communities. *Trends in Ecology & Evolution* **1**, 11–15 (1986).

Dobson, A. Population Dynamics of Pathogens with Multiple Host Species. *The American Naturalist* **164**, S64–S78 (2004).

Dobson, A. P. & Hudson, P. J. Regulation and Stability of a Free-Living Host-Parasite System: Trichostrongylus tenuis in Red Grouse. II. Population Models. *The Journal of Animal Ecology* **61**, 487 (1992).

Dobson, A., Lafferty, K. D., Kuris, A. M., Hechinger, R. F. & Jetz, W. Homage to Linnaeus: How many parasites? How many hosts? *Proc. Natl. Acad. Sci. U.S.A.* **105**, 11482–11489 (2008).

Dorazio, R. M. & Royle, J. A. Estimating Size and Composition of Biological Communities by Modeling the Occurrence of Species. *Journal of the American Statistical Association* **100**, 389–398 (2005).

Dorazio, R. M., Gotelli, N. J. & Ellison, A. M. Modern Methods of Estimating Biodiversity from Presence-Absence Surveys. in *Biodiversity Loss in a Changing Planet* (eds. Grillo, O. & Venora, G.) (InTech, 2011). doi: 10.5772/23881.

Dorazio, R. M., Royle, J. A., Söderström, B. & Glimskär, A. Estimating species richness and accumulation by modeling species occurrence and detectability. *Ecology* **87**, 842–854 (2006).

Dormann, C. F. *et al.* Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* **30**, 609–628 (2007).

Dormann, C. F. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography* **16**, 129–138 (2007).

Doser, J. W., Finley, A. O., Kéry, M. & Zipkin, E. F. spOccupancy: An R package for single-species, multi-species, and integrated spatial occupancy models. *Methods Ecol Evol* **13**, 1670–1678 (2022).

Dunn, R. R., Harris, N. C., Colwell, R. K., Koh, L. P. & Sodhi, N. S. The sixth mass coextinction: are most endangered species parasites and mutualists? *Proceedings. Biological sciences / The Royal Society* **276**, 3037–45 (2009).

Elith, J. & Leathwick, J. R. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annu. Rev. Ecol. Evol. Syst.* **40**, 677–697 (2009).

Ewers, W. H. The influence of the density of snails on the incidence of larval trematodes. *Parasitology* **54**, 579–583 (1964).

Fauchald, P. *et al.* Escaping parasitism in the selfish herd: age, size and density-dependent warble fly infestation in reindeer. *Oikos* **116**, 491–499 (2007).

Fortin, M.-J. & Dale, M. R. T. *Spatial analysis: a guide for ecologists*. (Cambridge University Press, 2005).

Garcia-Vedrenne, A. E. *et al.* Social Organization in Parasitic Flatworms—Four Additional Echinostomoid Trematodes Have a Soldier Caste and One Does Not. *Journal of Parasitology* **102**, 11 (2016).

Gascón, S. *et al.* Environmental filtering determines metacommunity structure in wetland microcrustaceans. *Oecologia* **181**, 193–205 (2016).

Gauch, H. G. *Multivariate analysis in community ecology*. (Cambridge University Press, 1982).

Gelman, A. Inference and monitoring convergence. in *Markov chain Monte Carlo in practice* (Chapman & Hall, 1996).

Gelman, A. Two simple examples for understanding posterior p-values whose distributions are far from uniform. *Electronic Journal of Statistics* **7**, 2595–2602 (2013).

Gelman, A., Goodrich, B., Gabry, J. & Vehtari, A. R-squared for Bayesian Regression Models. *The American Statistician* **73**, 307–309 (2019).

Gittleman, J. L. & Kot, M. Adaptation: Statistics and a Null Model for Estimating Phylogenetic Effects. *Systematic Zoology* **39**, 227 (1990).

Gleason, H. A. The Individualistic Concept of the Plant Association. *Bulletin of the Torrey Botanical Club* **53**, 7–26 (1926).

Gotelli, N. J. Null model analysis of species co-occurrence patterns. **81**, 16 (2000).

Guégan, J.-F., Morand, S. & Poulin, R. Are there general laws in parasite community ecology? The emergence of spatial parasitology and epidemiology. in *Parasitism and Ecosystems* (eds. Thomas, F., Renaud, F. & Guegan, J.-F.) 22–42 (Oxford University Press, 2005). doi:10.1093/acprof:oso/9780198529873.003.0003.

Guisan, A. & Rahbek, C. SESAM - a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages: Predicting spatio-temporal patterns of species assemblages. *Journal of Biogeography* **38**, 1433–1444 (2011).

Guisan, A. & Thuiller, W. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* **8**, 993–1009 (2005).

Harrison, S., Ross, S. J. & Lawton, J. H. Beta Diversity on Geographic Gradients in Britain. *The Journal of Animal Ecology* **61**, 151 (1992).

Hechinger, R. F., Lafferty, K. D. & Kuris, A. M. Diversity increases biomass production for trematode parasites in snails. *Proc. R. Soc. B.* **275**, 2707–2714 (2008).

Hechinger, R. F. Guide to the trematodes (Platyhelminthes) that infect the California horn snail (*Cerithideopsis californica*: Potamididae: Gastropoda) as first intermediate host. *Zootaxa* **4711**, (2019).

Hechinger, R. F. & Lafferty, K. D. Host diversity begets parasite diversity: Bird final hosts and trematodes in snail intermediate hosts. *Proceedings of the Royal Society B: Biological Sciences* **272**, 1059–1066 (2005).

Hechinger, R. F. *et al.* Monthly densities of the snail *Cerithideopsis* (*Cerithidea*) *californica* at Carpinteria Salt Marsh, California USA, February 2012 to January 2014. (2017) doi:10.5066/F71J9879.

Heino, J., Soininen, J., Alahuhta, J., Lappalainen, J. & Virtanen, R. A comparative analysis of metacommunity types in the freshwater realm. *Ecol Evol* **5**, 1525–1537 (2015).

Helton, J. C. & Davis, F. J. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety* **81**, 23–69 (2003).

HilleRisLambers, J., Adler, P. B., Harpole, W. S., Levine, J. M. & Mayfield, M. M. Rethinking Community Assembly through the Lens of Coexistence Theory. *Annual Review of Ecology, Evolution, and Systematics* **43**, 227–248 (2012).

Holmes, J. C. & Price, P. W. Communities of Parasites. in *Community Ecology: Patterns and Processes* (eds. Anderson, D. J. & Kikkawa, J.) 187–213 (Blackwell Scientific Publications, 1986).

Huspeni, T. C. & Lafferty, K. D. Using larval trematodes that parasitize snails to evaluate a saltmarsh restoration project. *Ecological Applications* **14**, 795–804 (2004).

Iknayan, K. J., Tingley, M. W., Furnas, B. J. & Beissinger, S. R. Detecting diversity: emerging methods to estimate species diversity. *Trends in Ecology & Evolution* **29**, 97–106 (2014).

Iooss, B. & Lemaître, P. A Review on Global Sensitivity Analysis Methods. in *Uncertainty Management in Simulation-Optimization of Complex Systems* (eds. Dellino, G. & Meloni, C.) vol. 59 101–122 (Springer US, 2015).

Johnson, P. T. J. *et al.* Species diversity reduces parasite infection through cross-generational effects on host abundance. *Ecology* **93**, 56–64 (2012).

Jost, L. Partitioning diversity into independent alpha and beta components. *Ecology* **88**, 2427–2439 (2007).

Keen, A. M. *Sea shells of tropical west America: marine mollusks from Baja California to Peru*. (Stanford University Press, 1971).

Kéry, M. & Royle, J. A. *Applied hierarchical modeling in ecology: analysis of distribution, abundance and species richness in R and BUGS*. (Elsevier/AP, Academic Press is an imprint of Elsevier, 2016).

Kim, S. ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients. *CSAM* **22**, 665–674 (2015).

Koenig, W. D. Spatial Autocorrelation in California Land Birds. *Conservation Biology* **12**, 612–620 (1998).

König, C. *et al.* Scale dependency of joint species distribution models challenges interpretation of biotic interactions. *Journal of Biogeography* **48**, 1541–1551 (2021).

Kraft, N. J. B. *et al.* Community assembly, coexistence and the environmental filtering metaphor. *Funct Ecol* **29**, 592–599 (2015).

Kühn, I. Incorporating spatial autocorrelation may invert observed patterns. *Diversity and Distributions* **13**, 66–69 (2007).

Kuris, A. Guild structure of larval trematodes in molluscan hosts prevalence, dominance and significance of competition. in *Parasite Communities: Patterns and Processes* (ed. Esch, G. W.) 69–100 (Springer Netherlands, 1990). doi:10.1007/978-94-009-0837-6.

Kuris, A. M., Blaustein, A. R. & Alio, J. J. Hosts as Islands. *The American Naturalist* **116**, 570–586 (1980).

Kuris, A. M. *et al.* Ecosystem energetic implications of parasite and free-living biomass in three estuaries. *Nature* **454**, 515–518 (2008).

Kuris, A. M. & Lafferty, K. D. Community Structure: Larval Trematodes in Snail Hosts. *Annual Review of Ecology and Systematics* **2**, 189–217 (1994).

Lafferty, K. D., Sammond, D. T. & Kuris, A. M. Analysis of Larval Trematode Communities. *Ecology* **75**, 2275 (1994).

Lafferty, K. D. Effects of parasitic castration on growth, reproduction and population dynamics of the marine snail *Cerithidea californica*. **96**, 229–237 (1993).

Lafferty, K. D. *et al.* Parasites in food webs: The ultimate missing links. *Ecology Letters* **11**, 533–546 (2008).

Lafferty, K. D. *et al.* Parasites in food webs: the ultimate missing links. *Ecology Letters* **11**, 533–546 (2008).

Lafferty, K. D. & Dunham, E. J. Trematodes in Snails near Raccoon Latrines Suggest a Final Host Role for this Mammal in California Salt Marshes. *Journal of Parasitology* **91**, 474–476 (2005).

Lafferty, K. D. & Morris, A. K. Altered behavior of parasitized killifish increases susceptibility to predation by bird final hosts. *Ecology* **77**, 1390–1397 (1996).

Lahoz-Monfort, J. J., Guillera-Arroita, G. & Wintle, B. A. Imperfect detection impacts the performance of species distribution models. *Global Ecology and Biogeography* **23**, 504–515 (2014).

Legendre, P. Spatial Autocorrelation: Trouble or New Paradigm? *Ecology* **74**, 1659–1673 (1993).

Legendre, P. & Fortin, M. J. Spatial pattern and ecological analysis. *Vegetatio* **80**, 107–138 (1989).

Leibold, M. A. & Chase, J. M. *Metacommunity Ecology*. (Princeton University Press, 2017).

Leibold, M. A. *et al.* The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters* **7**, 601–613 (2004).

Leibold, M. A. & Mikkelson, G. M. Coherence, species turnover, and boundary clumping: elements of meta-community structure. *Oikos* **97**, 237–250 (2002).

Levin, S. A. The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. *Ecology* **73**, 1943–1967 (1992).

Lorda, J. & Lafferty, K. D. Shading decreases the abundance of the herbivorous California horn snail, *Cerithidea californica*. *Journal of Experimental Marine Biology and Ecology* **432–433**, 148–155 (2012).

Loreau, M., Roy, J. & Tilman, D. Linking ecosystem and parasite ecology. in *Parasitism and Ecosystems* (eds. Thomas, F., Renaud, F. & Guegan, J.-F.) 13–21 (Oxford University Press, 2005). doi:10.1093/acprof:oso/9780198529873.003.0002.

MacArthur, R. H. *Geographical Ecology*. (Harper & Row Publishers, 1972).

MacArthur, R. H. & Pianka, E. R. On Optimal Use of a Patchy Environment. 603–609 (1966) doi:10.2307/2459298.

MacKenzie, D. I. *et al.* Estimating site occupancy rates when detection probabilities are less than one. *Ecology* **83**, 2248–2255 (2002).

May, R. M. & Anderson, R. M. Spatial heterogeneity and the design of immunization programs. *Mathematical Biosciences* **72**, 83–111 (1984).

May, R. M., Levin, S. A. & Sugihara, G. Ecology for bankers. *Nature* **451**, 893–894 (2008).

McCallum, H. How should pathogen transmission be modelled? *Trends in Ecology & Evolution* **16**, 295–300 (2001).

Mellado, A. & Zamora, R. Ecological consequences of parasite host shifts under changing environments: More than a change of partner. *Journal of Ecology* **108**, 788–796 (2020).

Michalska-Smith, M. J., Sander, E. L., Pascual, M. & Allesina, S. Understanding the role of parasites in food webs using the group model. *Journal of Animal Ecology* **87**, 790–800 (2018).

Mihaljevic, J. R. Linking metacommunity theory and symbiont evolutionary ecology. *Trends in Ecology and Evolution* **27**, 323–329 (2012).

Mihaljevic, J. R., Hoye, B. J. & Johnson, P. T. J. Parasite metacommunities: Evaluating the roles of host community composition and environmental gradients in structuring symbiont communities within amphibians. *J Anim Ecol* **87**, 354–368 (2018).

Mihaljevic, J. R., Joseph, M. B. & Johnson, P. T. J. Using multispecies occupancy models to improve the characterization and understanding of metacommunity structure. *Ecology* **96**, 1783–1792 (2015).

Miura, O., Torchin, M. E. & Bermingham, E. Molecular phylogenetics reveals differential divergence of coastal snails separated by the Isthmus of Panama. *Molecular Phylogenetics and Evolution* **56**, 40–48 (2010).

Mooring, M. S. & Hart, B. L. Animal grouping for protection from parasites: Selfish herd and encounter-dilution effects. *Behaviour* **123**, 173–193 (1992).

Moran, P. A. P. Notes on Continuous Stochastic Phenomena. *Biometrika* **37**, 17 (1950).

Mordecai, E. A., Jaramillo, A. G., Ashford, J. E., Hechinger, R. F. & Lafferty, K. D. The role of competition - colonization tradeoffs and spatial heterogeneity in promoting trematode coexistence. *Ecology* **97**, 1484–1496 (2016).

Morisita, M. Composition of the *Iδ* -index. *Population Ecology* **13**, 1–27 (1971).

Moss, W., McDevitt-Galles, T., Calhoun, D. & Johnson, P. T. J. Tracking the assembly of nested parasite communities: using β-diversity to understand variation in parasite richness and composition over time and scale. *Journal of Animal Ecology* (2020) doi:10.1111/1365-2656.13204.

Oksanen, J. *et al.* vegan: Community Ecology Package. (2020).

Ostfeld, R. S., Miller, M. C. & Hazler, K. R. Causes and Consequences of Tick (*Ixodes scapularis*) Burdens on White-Footed Mice (*Peromyscus leucopus*). *Journal of Mammalogy* **77**, 266–273 (1996).

Paine, R. T. Food Web Complexity and Species Diversity. 65–75 (1966) doi:10.2307/2459379.

Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).

Patterson, B. D. & Atmar, W. Nested subsets and the structure of insular mammalian faunas and archipelagos. *Biological Journal of the Linnean Society* **28**, 65–82 (1986).

Patterson, J. E. H. & Ruckstuhl, K. E. Parasite infection and host group size: a meta-analytical review. *Parasitology* **140**, 803–813 (2013).

Paull, S. H. *et al.* From superspreaders to disease hotspots: linking transmission across hosts and space. *Frontiers in Ecology and the Environment* **10**, 75–82 (2012).

Petzold, L. Automatic Selection of Methods for Solving Stiff and Nonstiff Systems of Ordinary Differential Equations. *SIAM J. Sci. and Stat. Comput.* **4**, 136–148 (1983).

Polson, N. G., Scott, J. G. & Windle, J. Bayesian Inference for Logistic Models Using Pólya–Gamma Latent Variables. *Journal of the American Statistical Association* **108**, 1339–1349 (2013).

Poulin, R. Are there general laws in parasite ecology? *Parasitology* **134**, 763–776 (2007).

Poulin, R. The decay of similarity with geographical distance in parasite communities of vertebrate hosts. *Journal of Biogeography* **30**, 1609–1615 (2003).
139.

Poulin, R. Macroecological patterns of species richness in parasite assemblages. *Basic and Applied Ecology* **5**, 423–434 (2004).

Presley, S. J. On the detection of metacommunity structure. *Community Ecology* 1–4 (2020) doi:10.1007/s42974-020-00004-9.

Presley, S. J., Higgins, C. L. & Willig, M. R. A comprehensive framework for the evaluation of metacommunity structure. *Oikos* **119**, 908–917 (2019).

Presley, S. J. & Willig, M. R. Bat metacommunity structure on Caribbean islands and the role. 185–199 (2010) doi:10.1111/j.1466-8238.2009.00505.x.

Presley, S. J., Higgins, C. L., López-González, C. & Stevens, R. D. Elements of metacommunity structure of Paraguayan bats: Multiple gradients require analysis of multiple ordination axes. *Oecologia* **160**, 781–793 (2009).

Presley, S. J., Higgins, C. L. & Willig, M. R. A comprehensive framework for the evaluation of metacommunity structure. *Oikos* **119**, 908–917 (2010).

Pullan, R. L., Sturrock, H. J. W., Soares Magalhães, R. J., Clements, A. C. A. & Brooker, S. J. Spatial parasite ecology and epidemiology: a review of methods and applications. *Parasitology* **139**, 1870–1887 (2012).

R Core Team. R: A language and environment for statistical computing. (2022).

Richgels, K. L. D., Hoverman, J. T. & Johnson, P. T. J. Evaluating the role of regional and local processes in structuring a larval trematode metacommunity of *Helisoma trivolvis*. *Ecography* **36**, 854–863 (2013).

Rifkin, J. L., Nunn, C. L., Garamszegi, L. Z., Steppan, A. E. S. J. & Bronstein, E. J. L. Do Animals Living in Larger Groups Experience Greater Parasitism? A Meta-Analysis. *The American Naturalist* **180**, 70–82 (2012).

Roberts, L. S., Janovy, J., Nadler, S. & Roberts, L. S. *Foundations of parasitology*. (McGraw Hill, 2013).

Rohr, J. R. *et al.* Predator diversity, intraguild predation, and indirect effects drive parasite transmission. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3008–3013 (2015).

Rota, C. T. *et al.* A multispecies occupancy model for two or more interacting species. *Methods Ecol Evol* **7**, 1164–1173 (2016).

Royle, J. A. & Dorazio, R. M. *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities*. (Academic Press, 2009).

Rubin, D. B. Bayesianly Justifiable and Relevant Frequency Calculations for the Applied Statistician. *The Annals of Statistics* **12**, 1151–1172 (1984).

Samsing, F., Oppedal, F., Johansson, D., Bui, S. & Dempster, T. High host densities dilute sea lice *Lepeophtheirus salmonis* loads on individual Atlantic salmon, but do not reduce lice infection success. *Aquacult. Environ. Interact.* **6**, 81–89 (2014).

Sattler, T., Obrist, M. K., Duelli, P. & Moretti, M. Urban arthropod communities: Added value or just a blend of surrounding biodiversity? *Landscape and Urban Planning* **103**, 347–361 (2011).

Schmera, D., Podani, J., Botta-Dukát, Z. & Erős, T. On the reliability of the Elements of Metacommunity Structure framework for separating idealized metacommunity patterns. *Ecological Indicators* **85**, 853–860 (2018).

Shoemaker, L. G. & Melbourne, B. A. Linking metacommunity paradigms to spatial coexistence mechanisms. *Ecology* **97**, 2436–2446 (2016).

Smith, N. F. Spatial heterogeneity in recruitment of larval trematodes to snail intermediate hosts. *Oecologia* **127**, 115–122 (2001).

Soetaert, K., Petzoldt, T. & Setzer, R. W. Solving Differential Equations in R: Package deSolve. *J. Stat. Soft.* **33**, (2010).

Sousa, W. P. Host life history and the effect of parasitic castration on growth: A field study of *Cerithidea californica* Haldeman (Gastropoda: Prosobranchia) and its trematode parasites. *Journal of Experimental Marine Biology and Ecology* **73**, 273–296 (1983).

Sousa, W. P. Spatial scale and the processes structuring a guild of larval trematode parasites. in *Parasite Communities: Patterns and Processes* (eds. Esch, G. W., Bush, A. O. & Aho, J. M.) 41–67 (Springer Netherlands, 1990). doi:10.1007/978-94-009-0837-6_3.

Sousa, W. P. Interspecific Antagonism and Species Coexistence in a Diverse Guild of Larval Trematode Parasites. *Ecological Monographs* **63**, 104–128 (1993).

Spearman, C. 'General Intelligence,' Objectively Determined and Measured. *The American Journal of Psychology* **15**, 201 (1904).

Stella, M., Andreazzi, C. S., Selakovic, S., Goudarzi, A. & Antonioni, A. Parasite spreading in spatial ecological multiplex networks. *Journal of Complex Networks* **5**, 486–511 (2017).

Streicker, D. G., Fenton, A. & Pedersen, A. B. Differential sources of host species heterogeneity influence the transmission and control of multihost parasites. *Ecology Letters* **16**, 975–984 (2013).

Tao, Y., Hite, J. L., Lafferty, K. D., Earn, D. J. D. & Bharti, N. Transient disease dynamics across ecological scales. *Theor Ecol* **14**, 625–640 (2021).

Thompson, P. L. *et al.* A process-based metacommunity framework linking local and regional scale community ecology. *Ecol Lett* **23**, 1314–1329 (2020).

Tilman, D. *Resource Competition and Community Structure*. (Princeton University Press, 1982).

Torchin, M. E., Miura, O. & Hechinger, R. F. Parasite species richness and intensity of interspecific interactions increase with latitude in two wide-ranging hosts. *Ecology* **96**, 3033–3042 (2015).

Turner, G. F. & Pitcher, T. J. Attack Abatement: A Model for Group Protection by Combined Avoidance and Dilution. *The American Naturalist* **128**, 228–240 (1986).

Tyre, A. J. *et al.* Improving precision and reducing bias in biological surveys: estimating false-negative error rates. *Ecological Applications* **13**, 1790–1801 (2003).

Ulrich, W. & Gotelli, N. J. Pattern detection in null model analysis. *Oikos* **122**, 2–18 (2013).

Vellend, M. *The Theory of Ecological Communities (MPB-57)*. (Princeton University Press, 2016).

Watanabe, S. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *Journal of Machine Learning Research* **11**, 3571–3594 (2010).

White, L. A., Forester, J. D. & Craft, M. E. Dynamic, spatial models of parasite transmission in wildlife: Their structure, applications and remaining challenges. *Journal of Animal Ecology* **87**, 559–580 (2018).

Wilson, H., Miles, A. F., Daane, K. M. & Altieri, M. A. Landscape Diversity and Crop Vigor Influence Biological Control of the Western Grape Leafhopper (*E. elegantula* Osborn) in Vineyards. *PLoS ONE* **10**, e0141752 (2015).

Windsor, D. A. Most of the species on Earth are parasites. *Int J Parasitol* **28**, 1939–1941 (1998).

Wisz, M. S. *et al.* The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews* **88**, 15–30 (2013).

Woolhouse, M. E. Mathematical models of transmission dynamics and control of schistosomiasis. *Am J Trop Med Hyg* **55**, 144–148 (1996).

Wright, D. H., Patterson, B. D., Mikkelson, G. M., Cutler, A. & Atmar, W. A comparative analysis of nested subset patterns of species composition. *Oecologia* **113**, 1–20 (1997).

Zimmermann, N. E., Edwards, T. C., Graham, C. H., Pearman, P. B. & Svenning, J.-C. New trends in species distribution modelling. *Ecography* **33**, 985–989 (2010).

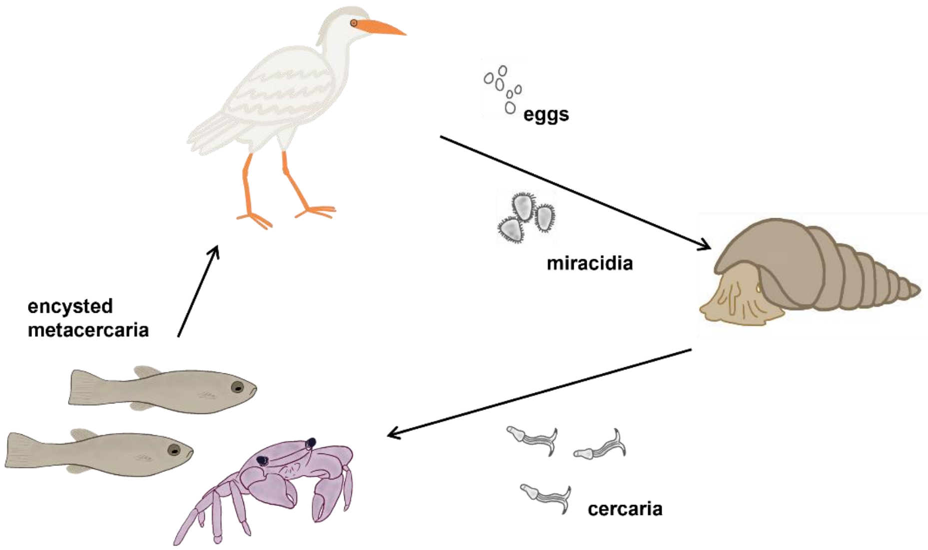
Zipkin, E. F., DeWan, A. & Andrew Royle, J. Impacts of forest fragmentation on species richness: a hierarchical approach to community modelling. *Journal of Applied Ecology* **46**, 815–822 (2009).

Zobel, M. The relative of species pools in determining plant species richness: an alternative explanation of species coexistence? *Trends in Ecology & Evolution* **12**, 266–269 (1997).

Zurell, D. *et al.* Testing species assemblage predictions from stacked and joint species distribution models. *Journal of Biogeography* **47**, 101–113 (2020).

Appendices

APPENDIX CHAPTER 1



Appendix Figure 1: A simplified life cycle of the trematode guild described in Carpinteria Salt Marsh. The life cycle begins (and ends) in the definitive host, which are often birds, where adult trematodes release trematode eggs into the environment through host feces. Miracidia and seek out the first intermediate host (such as *C. californica*) and cercaria are shed into the environment where they may encyst on the second intermediate host (such as a fish or crab) or on substrate in the environment. The final host may encounter metacercaria environmentally or trophically.

Scientific Name	**Family**	**Superfamily**	**Species Abbreviation**
Acanthotrema hancocki	Heterophyidae	Opisthorchioidea	acha
Acanthoparyphium spinulosum	Himasthlidae	Echinostomatoidea	acsp
Austrobilharzia sp.	Schistosomatidae	Schistosomatoidea	ausp
Catatropis johnstoni	Notocotylidae	Pronocephaloidea	cajo
Cloacitrema michiganensis	Philophthalmidae	Echinostomatoidea	clmi
Euhaplorchis californiensis	Heterophyidae	Opisthorchioidea	euca
Himasthla rhigedana	Himasthlidae	Echinostomatoidea	hirh
Himasthla sp. B	Himasthlidae	Echinostomatoidea	hisb
Mesostephanus appendiculatus	Cyathocotylidae	Diplostomoidea	meap
Parorchis sp.	Philophthalmidae	Echinostomatoidea	pasp
Phocitremoides ovale	Heterophyidae	Opisthorchioidea	phov
Probolocoryphe uca	Microphallidae	Microphalloidea	pruc
Pygidiopsoides spindalis	Heterophyidae	Opisthorchioidea	pysp
Renicola buchanani	Renicolidae	Microphalloidea	rebu
Renicola cerithidicola	Renicolidae	Microphalloidea	rece
Renicola sp. "martini"	Renicolidae	Microphalloidea	rema
Renicola sp. "polychaetophila"	Renicolidae	Microphalloidea	repo
Small cyathocotylid	Cyathocotylidae	Diplostomoidea	smcy
Small microphallid	Microphallidae	Microphalloidea	smmi

Appendix Table 1: Trematode species described in this study across 5 replicated field surveys of the trematode community. Species abbreviations were used for trematode species and were consistent with other studies conducted on this guild of trematodes (see details in Hechinger 2019).

Season	Collection Date	C2	C3	C6	C7	C8	F1	F2	F3	F4	F5
Wet	2012-02-14	33	16	44	0	0	0	0	0	0	26
Wet	2012-02-15	0	0	0	14	16	26	36	17	43	0
Wet	2012-04-09	49	33	0	32	26	31	36	16	28	35
Dry	2012-05-08	42	29	0	18	48	31	32	25	35	18
Dry	2012-06-08	40	23	0	35	28	0	40	22	50	5
Dry	2012-06-13	0	0	0	0	0	50	0	0	0	15
Dry	2012-07-09	51	26	0	41	36	23	38	37	36	26
Dry	2012-08-16	38	22	0	36	21	28	44	27	55	17
Dry	2012-09-14	45	19	0	38	36	21	46	22	43	28
Dry	2012-10-16	26	34	0	42	36	27	37	35	47	19
Wet	2012-11-09	47	33	0	34	30	27	42	29	39	32
Wet	2012-12-04	56	38	0	42	20	33	39	25	57	21
Wet	2013-01-08	58	27	0	43	37	22	40	20	47	44
Wet	2013-02-05	23	32	0	34	27	29	35	16	45	24
Wet	2013-03-05	52	28	0	37	20	26	24	11	44	15
Wet	2013-04-03	0	31	0	24	26	0	31	0	0	0
Wet	2013-04-04	34	0	0	0	0	36	0	21	35	26
Dry	2013-06-14	43	27	0	17	18	32	33	20	40	9
Dry	2013-07-16	31	37	0	13	20	19	38	17	56	35
Dry	2013-08-12	44	24	0	17	31	26	33	16	46	29
Dry	2013-09-17	55	26	0	15	24	31	34	18	28	25
Dry	2013-10-15	42	12	0	14	29	37	38	5	46	34
Wet	2013-11-12	54	32	0	17	36	22	29	6	55	40
Wet	2013-12-12	46	7	0	0	0	24	0	35	46	22
Wet	2013-12-13	0	0	0	29	25	0	19	0	0	0
Wet	2014-01-13	54	29	0	0	0	48	0	32	50	45
Wet	2014-01-14	0	0	0	31	22	0	28	0	0	0

Appendix Table 2: Number of snails (*i.e.*, *C. californica* individuals) out of the 58 snails collected on a given sample date at each of the 10 sites in CSM that were infected with trematodes. Sites were not consistently sampled on every survey date, but all but one site was sampled consistently every season. On each sample date, 58 snail hosts were collected and dissected, and trematodes were identified to species (see Kuris *et al.* 2008 for details).

Variable	**Description**	**Unit**	**Use**
Species occurrence	Presence/absence of each trematode species in an infected *C. californica* host, collected at each site on multiple sample dates via dissection (Appendix Table 1)	Binary	Dependent variable
Mean host size	Size (body length) of the infected host, including soft and hard parts, collected during dissections of hosts that generated species occurrence data	mm	Occupancy covariate, Detection covariate
Bird diversity	Shannon index of bird species identified at sample sites	index	Occupancy covariate
Bird density	The density of birds at a sample site, based on biomass per unit area collected at each site	g/m^2	Occupancy covariate
Habitat	Habitat type of the site sampled (may be either sub-tidal channel or mudflat)	character	Occupancy covariate
Pickleweed cover	The percent of space in a sample site that was occupied by *Sarcocornia virginica* vegetation, calculated as the average of three vegetation surveys collected at each site	%	Occupancy covariate
Organic matter	The percent of sediment mass that is organic material, based on samples collected once at each site	%	Occupancy covariate
Season	The season (wet, dry) of the sample date, based on the quarter of the year	character	Detection covariate
Year	Year of the sample date	Year	Detection covariate
Coordinates	Spatial coordinates (latitude/longitude) of each site sampled	Spatial coordinates	Spatial covariate

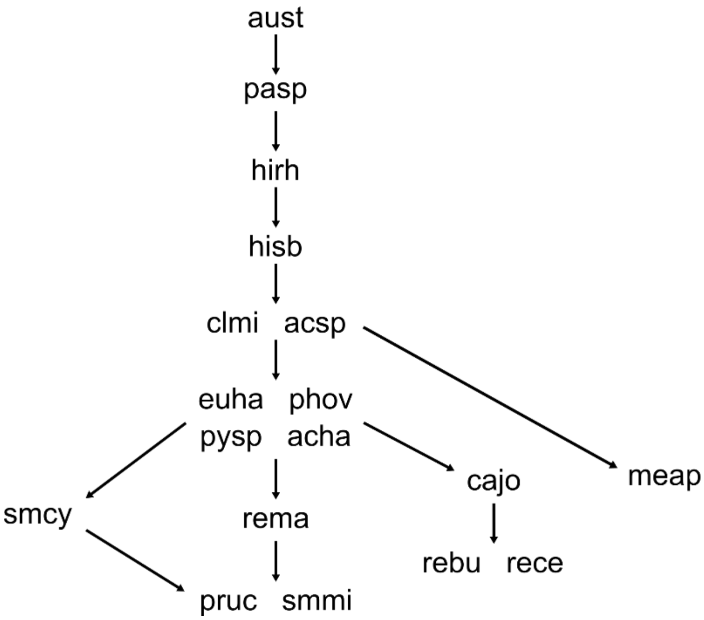
Appendix Table 3: Covariates in each multi-species occupancy model described environmental and spatial conditions of *C. californica* hosts and their trematode parasites.

Bayesian p-values

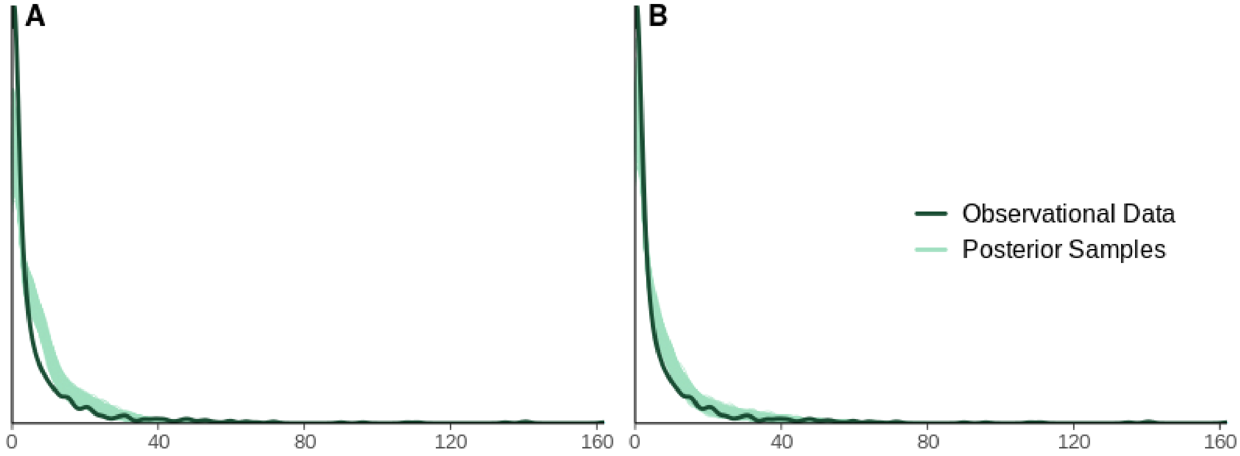
Species	Freeman-Tukey	Chi-Squared
acha	0.44	0.44
acsp	0.33	0.24
ausp	0.63	0.48
cajo	0.46	0.45
clmi	0.38	0.35
euca	0.45	0.45
hirh	0.62	0.61
hisb	0.53	0.52
meap	0.19	0.28
pasp	0.78	0.77
phov	0.14	0.14
pruc	0.64	0.62
pysp	0.55	0.55
rebu	0.39	0.38
rece	0.45	0.63
rema	0.38	0.33
repo	0.18	0.20
smcy	0.46	0.46
smmi	0.35	0.32
Community	0.44	0.43

Appendix Table 4: Bayesian p-values reported for each species described in this DECEMS analysis, generated in posterior predictive checks of the species occupancy model. See Appendix Table 1 for species abbreviations. We concluded that our model predictions are similar to observed data since both the community-level and most species-specific Bayesian p-values are close to 0.5 and none are below 0.1 or above 0.9 (*i.e.*, p-values do not suggest a strong bias for high or low predicted values). Below we report the posterior predictive check Bayesian p-values using both a Freeman-Tukey and a chi-squared fit statistic.

APPENDIX CHAPTER 2



Appendix Figure 2: Dominance hierarchy of trematode species (add detail from Hechinger 2019). The top species (“aust”) is considered the most dominant species, meaning it will outcompete other species that are shown below it within a snail host. The most dominant position is referred to as rank 1 for shorthand (*e.g.*, Figure 6). Trematode species that are grouped together (*e.g.*, “clmi” and “acsp”) are considered competitive equals. Species abbreviations are listed in Appendix Table 1 with additional taxonomic detail.



Appendix Figure 3: Posterior predictive plots of the trematode community data (dark line) and the predicted (posterior) samples (light lines). Panel A shows the model without a spatial term, and Panel B shows the model with the spatial autocorrelation term.

Scientific Name	**Family**	**Superfamily**	**Species Abbreviation**
Acanthotrema hancocki	Heterophyidae	Opisthorchioidea	acha
Acanthoparyphium spinulosum	Himasthlidae	Echinostomatoidea	acsp
Austrobilharzia sp.	Schistosomatidae	Schistosomatoidea	ausp
Catatropis johnstoni	Notocotylidae	Pronocephaloidea	cajo
Cloacitrema michiganensis	Philophthalmidae	Echinostomatoidea	clmi
Euhaplorchis californiensis	Heterophyidae	Opisthorchioidea	euca
Himasthla rhigedana	Himasthlidae	Echinostomatoidea	hirh
Himasthla sp. B	Himasthlidae	Echinostomatoidea	hisb
Mesostephanus appendiculatus	Cyathocotylidae	Diplostomoidea	meap
Parorchis sp.	Philophthalmidae	Echinostomatoidea	pasp
Phocitremoides ovale	Heterophyidae	Opisthorchioidea	phov
Probolocoryphe uca	Microphallidae	Microphalloidea	pruc
Pygidiopsoides spindalis	Heterophyidae	Opisthorchioidea	pysp
Renicola buchanani	Renicolidae	Microphalloidea	rebu
Renicola cerithidicola	Renicolidae	Microphalloidea	rece
Renicola sp. "martini"	Renicolidae	Microphalloidea	rema
Small cyathocotylid	Cyathocotylidae	Diplostomoidea	smcy
Small microphallid	Microphallidae	Microphalloidea	smmi

Appendix Table 5: 18 digenean trematodes were described in the empirical data that informed species distribution models. We used species abbreviations in analyses that were consistent with other studies conducted on this guild of trematodes (see details in Hechinger 2019).

Variable	**Description**	**Unit**	**Source**
Trematode data	Trematodes were identified in infected *C. californica* snails at local sampling sites; trematode abundance was calculated as the number of infections by a trematode species at a site	NA	Kuris *et al.* 2008
Dominance rank	The rank of a trematode species in the dominance hierarchy described for this system (Appendix Figure 2) with 1 indicating the most dominant species (*e.g.* Figure 6)	NA	Hechinger 2019
Transmission type	Whether the trematode species is transmitted to the first intermediate host as an egg or a miracidium larva	NA	Buck *et al.* 2017
Mean snail biomass	The mean biomass density of all (uninfected and infected) snail hosts in a sample site	g/m^2	Kuris *et al.* 2008
Mean host size	Mean (infected) snail host body mass within a sample site	grams	Kuris *et al.* 2008
Bird density	The density of bird biomass observed within a sample site	g/m^2	Kuris *et al.* 2008
Bird diversity	The Shannon index (*H*) of bird diversity observed in each sample site, calculated from bird observations using the diversity function from the R package 'vegan' (Oksanen *et al.* 2022)	NA	Kuris *et al.* 2008
Pickleweed cover	The percent cover of pickleweed (*Sarcocornia virginica*) within a sample site, based on three transects	%	Kuris *et al.* 2008
Organic matter	The percent of sediment from sediment cores that was organic matter; natural log-transformed in all models and averaged by replicate for the estuary- and habitat-scale	%	Kuris *et al.* 2008
Habitat	The habitat type (marsh, mudflats, pans, sub-tidal channel) occurring at a sample site; covariate not used in estuary-scales analyses, and used as a random effect rather than fixed effect at the habitat-scale	NA	Kuris *et al.* 2008
Euclidean distance	The latitude and longitude for local sites that were sampled in CSM, EPB, and BSQ were used for calculated the Euclidean distance	x-y coordinates; distance matrix	Kuris *et al.* 2008

	between each pair of sites, and informed the spatial conditional autoregressive term		

Appendix Table 6: List and description of environmental and trait covariates in SDMs. See *Methods* for details on data sources and collection.

	Non-spatial model		**Spatial model**	
Predictors	***Estimate***	***CI***	***Estimate***	***CI***
Intercept	1.71	0.57 - 2.83	1.35	0.046 - 2.55
Habitat (flat)	0.22	0.16 - 0.28	0.31	-0.17 - 0.78
Habitat (marsh)	-0.57	-0.67 - -0.47	-0.55	-1.10 - 0.005
Habitat (pan)	-0.05	-0.12 - 0.033	0.01	-0.59 - 0.55
Organic matter	-0.10	-0.14 - -0.062	-0.14	-0.38 - 0.089
Pickleweed cover	0.08	0.046 - 0.10	0.11	-0.1 - 0.32
Bird diversity	-0.09	-0.12 - -0.066	-0.15	-0.34 - 0.039
Bird density	0.07	0.048 - 0.099	0.08	-0.11 - 0.27
Mean host size	-0.06	-0.10 - -0.023	-0.15	-0.42 - 0.13
Mean snail biomass	0.08	0.051 - 0.12	0.11	-0.13 - 0.37
Spatial CAR term	NA	NA	0.51	0.03 - 0.98
Transmission (miracidia)	0.09	-0.98 - 1.20	0.14	-0.96 - 1.26
Dominance 2	-0.21	-1.69 - 1.27	-0.25	-1.72 - 1.27
Dominance 3	0.44	-1.06 - 1.93	0.41	-1.10 - 1.89
Dominance 4	0.51	-1.01 - 1.98	0.57	-0.95 - 2.05
Dominance 5	0.06	-1.26 - 1.40	0.06	-1.28 - 1.40
Dominance 6	0.29	-0.89 - 1.47	0.36	-0.85 - 1.55
Dominance 7	-0.16	-1.37 - 1.05	-0.13	-1.33 - 1.07
Dominance 8	-0.63	-1.91 - 0.66	-0.61	-1.88 - 0.69

Appendix Table 7: Each species distribution model describing trematode abundance was run with and without spatial autocorrelation informed by the Euclidean distance between sample site locations. This produced two total model versions. Here we provide the covariate estimate (model coefficients) and 95% credible intervals. Note that the intercept of both models accounts for the tidal channel habitat type, species with dominance rank 1, and species that transmit using eggs (Appendix Table 2).

	Observed *I*	**Expected *I***	**Std Dev**	**p-value**
Euclidean Distance				
Organic matter	0.222	-0.019	0.058	2.87e-5*
Pickleweed cover	0.762	-0.019	0.060	0*
Bird diversity	0.019	-0.019	0.059	0.51
Bird density	-0.008	-0.019	0.058	0.85
Mean host size	0.479	-0.019	0.059	0*
Mean snail biomass	0.090	-0.019	0.058	0.061
Binary for Estuary Connectivity				
Organic matter	-0.140	-0.019	0.018	9.37e-12*
Pickleweed cover	-0.370	-0.019	0.018	1.42e-80*
Bird diversity	-0.006	-0.019	0.018	0.49
Bird density	-0.011	-0.019	0.018	0.65
Mean host size	-0.242	-0.019	0.018	1.55e-34*
Mean snail biomass	-0.027	-0.019	0.018	0.66

Appendix Table 8: We tested for spatial autocorrelation using Moran's *I* for all environmental covariates, which were measured at each sample site. We calculated Moran's *I* using both Euclidean distance between all sites, and using a binary that indicated whether pairs of sites were in the same estuary or not. We presented the standard deviation and p-value (p-values < 0.05 are indicated by an asterisk) of each test as well as the observed (the computed Moran's *I*) and expected value under a null hypothesis.