Using multi-response models to investigate pathogen coinfections across scales: Insights from emerging diseases of amphibians

William E. Stutz1 | Andrew R. Blaustein2 | Cheryl J. Briggs3 | Jason T. Hoverman4 | Jason R. Rohr5 | Pieter T. J. Johnson1

1Department of Ecology and Evolutionary Biology, University of Colorado at Boulder, Boulder, CO, USA
2Integrative Biology, Oregon State University, Corvallis, OR, USA
3Ecology, Evolution and Marine Biology, University of California, Santa Barbara, Santa Barbara, CA, USA
4Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN, USA
5Department of Integrative Biology, University of South Florida, Tampa, FL, USA

Correspondence
Pieter Johnson
Email: pieter.johnson@colorado.edu

Funding information
National Institutes of Health, Grant/Award Number: R01GM109499; National Science Foundation, Grant/Award Number: DEB-1149308; David and Lucile Packard Foundation

Handling Editor: Steven Kembel

Abstract

1. Associations among parasites affect many aspects of host–parasite dynamics, but a lack of analytical tools has limited investigations of parasite correlations in observational data that are often nested across spatial and biological scales.

2. Here we illustrate how hierarchical, multiresponse modelling can characterize parasite associations by allowing for hierarchical structuring, offering estimates of uncertainty and incorporating correlational model structures. After introducing the general approach, we apply this framework to investigate coinfections among four amphibian parasites (the trematodes *Ribeiroia ondatrae* and *Echinostoma* spp., the chytrid fungus *Batrachochytrium dendrobatidis* and ranaviruses) and among >2,000 individual hosts, 90 study sites and five amphibian host species.

3. Ninety-two percent of sites and 80% of hosts supported two or more pathogen species. Our results revealed strong correlations between parasite pairs that varied by scale (from among hosts to among sites) and classification (microparasite versus macroparasite), but were broadly consistent across taxonomically diverse host species. At the host-scale, infection by the trematode *R. ondatrae* correlated positively with the microparasites, *B. dendrobatidis* and ranavirus, which were themselves positively associated. However, infection by a second trematode (*Echinostoma* spp.) correlated negatively with *B. dendrobatidis* and ranavirus, both at the host- and site-level scales, highlighting the importance of differential relationships between micro- and macroparasites.

4. Given the extensive number of coinfesting symbiont combinations inherent to natural systems, particularly across multiple host species, multiresponse modelling of cross-sectional field data offers a valuable tool to identify a tractable number of hypothesized interactions for experimental testing while accounting for uncertainty and potential sources of co-exposure. For amphibians specifically, the high frequency of co-occurrence and coinfection among these pathogens—each of which is known to impair host fitness or survival—highlights the urgency of understanding parasite associations for conservation and disease management.
INTRODUCTION

Many, if not all, organisms become infected by multiple parasite species during their lives (Petney & Andrews, 1998). Concurrent infections can alter fitness consequences for the host either additively or through interactions between parasites (e.g., Cox, 2001). Parasite interactions—which can be positive or negative—occur through a variety of mechanisms, including indirect interactions mediated by the host immune system as well as direct pathways, such as competition for host resources (Graham, 2002; Mideo, 2009). Within-host interactions have recently been implicated as influential in a number of human and wildlife diseases, including HIV and malaria in humans (Korenromp et al., 2005), bovine tuberculosis in African buffalo (Ezenwa & Jolles, 2015), colony collapse disorder in honeybees (Ratnieks & Carreck, 2010) and emerging infections in coral reefs (Bromenshenk et al., 2010). Coinfection-mediated changes in infection or pathology within hosts can also affect the resulting transmission rates between hosts or among populations (Ezenwa & Jolles, 2015; Hellard, Fouchet, Vavr, & Pontier, 2015; Johnson, De Roode, & Fenton, 2015). Given the potential consequences of coinfection for pathology, mortality and transmission, understanding the strength and extent of parasite associations in natural populations has become a major goal for disease ecologists (Hellard et al., 2015; Jolles, Ezenwa, Etienne, Turner, & Olff, 2008; Lello, Boag, Fenton, Stevenson, & Hudson, 2004; Telfer et al., 2010; Tompkins, Dunn, Smith, & Telfer, 2011).

Many factors influence parasite exposure and susceptibility both among host individuals (e.g., age, sex, diet, genetics) and across host populations or species (e.g., environmental, spatial or epidemiological factors) (Hawley & Altizer, 2011; Hellard et al., 2015; Wilson et al., 2002). In some cases, parasite correlations observed at one scale (e.g., infections within individual hosts) may be absent or reversed at other scales (e.g., average infection load among sites), emphasizing the importance of explicitly considering scale (the inappropriate application of correlations at one scale to another scale is sometimes referred to as the "ecological fallacy") (Joseph, Stutz, & Johnson, 2016). Thus, even in the absence of interspecific parasite interactions, spatial or temporal non-independence of the factors influencing infection can generate correlations in parasite occurrence or abundance. For instance while infections with two plant viruses, barley- and cereal yellow dwarf virus (B/CYDV), often correlate positively within Oregon meadows (Seabloom, Borer, Mitchell, & Power, 2010), whether such patterns reflect parasite interaction (e.g., facilitation) or instead stem from shared vectors or other environmental drivers is challenging to determine from surveys alone. This obstacle is particularly relevant to cross-sectional studies in multisymbiont and multihost systems, where the number of potential parasite associations is large and not easily tested experimentally. Thus, rigorous analytical methods that can help to characterize covariance in parasite abundance across different scales (e.g., among individuals, populations, communities, regions) and among multiple host species are becoming increasingly important (Hellard et al., 2015). Effective tools to identify pathogen correlations from cross-sectional field data—while characterizing sources of uncertainty and scale-specific covariates that can drive variation in parasite exposure—are essential for understanding the degree to which particular hosts (e.g., superspreaders), species (e.g., reservoirs) or geographical locations (e.g., hotspots) are disproportionately influential to the transmission of multiple parasites concurrently (e.g., Paull et al., 2012; Streicker, Fenton, & Pedersen, 2013).

To address these challenges, we advance hierarchical, multiresponse models as a flexible analytical framework for investigating multiscale parasite associations in natural populations (Cressie, Calder, Clark, Hoef, & Wikle, 2009; Downs & Dochtermann, 2014; Hadfield, 2010). This approach offers three core advantages over standard regression-type approaches: (1) direct modelling of associations and correlations, (2) incorporation of multiscale covariance (e.g., individual, sites and host species) within the same statistical models and (3) estimation of uncertainty while accounting for multiscale covariance. Although multiresponse approaches build on commonly used methods for analysing non-normal data, such as generalized linear mixed models (Schielzeth & Nakagawa, 2013), they account for uncertainty in all included variables (rather than simply the "response" variable) and emphasize correlational structures (rather than imposing a covariate-response approach). To illustrate and apply the outlined multiresponse model framework, we examined multiscale (individual host, community, host species) associations in four important amphibian parasites—the trematodes Riberoia ondatrae and Echinostoma spp., the chytrid fungus Batrachochytrium dendrobatidis and the viral pathogen ranavirus—across >2,000 individual hosts in >90 wetland communities (Figure 1). These parasites have been linked to declines or severe pathologies in many amphibian species and populations (Green, Converse, & Schrader, 2002; Johnson & McKenzie, 2009; Johnson et al., 2002; Vredenburg, Knapp, Tunstall, & Briggs, 2010), but little is known about the importance of concomitant infections in natural host communities. After describing how hierarchical, multiresponse models can be applied to study associations between parasites species generally, we apply it to the amphibian–parasite system as a case study to: (1) estimate correlations in parasite abundance at the individual, site and host-species levels, (2) determine how the strengths and directions of associations varied by scale and (3) illustrate how inclusion of environmental or individual-level covariates can provide insight into the factors driving observed associations. By helping to identify potential correlations among parasites and across scales, this approach is intended to be complementary to experimental investigations and predictive modelling, which collectively can offer insights...
FIGURE 1  Map of study system that illustrates the multilevel data structure. (a) Study sites are divided into four distinct metacommunities (shapes) across three counties in California. Each metacommunity is composed of many individual wetlands (b), each containing an assemblage of between one and five host species (c). Within each site, we collected multiple individuals (frogs, toads or newts) for each host species present and determined the abundance of three different parasites for each individual (d). Our framework provides an approach to assessing associations between parasite species within each scale (site, species, individual). B. dendrobatidis and ranavirus pictures are courtesy of CSIRO into the importance or frequency of parasite associations within host communities.

2 | MATERIALS AND METHODS

2.1 | Study system

We sampled 93 wetlands within a c. 10,000 m² area covering three counties in central California (Figure 1). Sampled wetlands are grouped within neighbouring regional parks or reserves, creating multiple clusters (i.e. metacommunities) of sites within the larger study area (Hoverman, Mihaljevic, Richgels, Kerby, & Johnson, 2012). Although up to seven species of amphibians can be found in the study area, five species comprise the majority of non-threatened amphibian hosts (Johnson, Preston, Hoverman, & Richgels, 2013): Pseudacris regilla (Pacific chorus frog), Anaxyrus boreas (western toad), Lithobates catesbeianus (American bullfrog), Taricha torosa (California newt) and Taricha granulosa (rough-skinned newt). These amphibians are infected by trematode parasites such as Ribeiroia ondatrae and Echinostoma spp., which enter amphibians through the skin (R. ondatrae) or cloaca (Echinostoma) and encyst in the hindlimb region and kidneys, respectively. Ribeiroia ondatrae causes severe malformations and both parasites can reduce survival to metamorphosis in many host species (Johnson & McKenzie, 2009; Johnson et al., 2012). Ranaviruses are intracellular viral pathogens transmitted between amphibians through direct contact with infected animals, necrophagy or exposure to contaminated water or substrates (Gray, Miller, & Hoverman, 2009), in some cases causing a severe external and internal pathologies (Densmore & Green, 2007; Miller, Gray, & Storfer, 2011). Batrachochytrium dendrobatidis is a water-borne fungal pathogen with both extra- and intracellular stages (Densmore & Green, 2007). Its emergence has been linked to massive die-offs and severe population declines across multiple continents (Rohr & Raffel, 2010; Kilpatrick, Briggs, & Daszak, 2010; Vredenburg et al., 2010). Based on experimental infections, host species vary in their susceptibility to—and pathology resulting from—all four parasites (Blaustein et al., 2005; Hoverman, Gray, Haislip, & Miller, 2011; Johnson & Buller, 2011; Johnson et al., 2012; Vredenburg et al., 2010).

2.2 | Data collection

We sampled 2,156 amphibian hosts during the summer of 2013 (859 P. regilla, 275 A. boreas, 172 L. catesbeianus, 681 T. torosa, 169 T. granulosa). Within sites, sample sizes for most species reached at least 10 individuals when present (maximum 30 individuals). Upon capture, hosts were swabbed for B. dendrobatidis using a standardized protocol (Briggs, Knapp, & Vredenburg, 2010) and then examined for R. ondatrae and Echinostoma infection. Pieces of liver and kidney tissue were removed during necropsy, frozen and screened for ranavirus. Ribeiroia ondatrae is discernible from other larval trematodes by the presence of oesophageal diverticula, whereas Echinostoma spp. were identified through characteristic collar spines (Johnson & McKenzie, 2009). We use “Echinostoma” here inclusively to refer to echinostomes (e.g. E. trivolvis, E. revolutum, Echinoparyphrium spp.) because their morphological similarity often precludes species-level identification (Johnson & McKenzie, 2009). The total number of metacercariae per host was used as our measure of parasite abundance. Snout-vent lengths (SVL) of each host were mean-centred and scaled to unit variance within host species prior to inclusion in analyses.

Batrachochytrium dendrobatidis DNA was extracted and quantified using standardized protocols (Boyle, Boyle, Olsen, Morgan, & Hyatt, 2004; Hyatt et al., 2007) and swabs were run in triplicate with an internal positive control to test for DNA inhibition. Samples were declared “positive” only if they amplified at least twice across runs. Total B. dendrobatidis zoospore-equivalents (ZE) per swab were averaged across swabs and rounded to the nearest whole number. Ranaviral DNA was extracted
from liver and kidney tissue samples and infection was determined using standard quantitative PCR protocols (Forson & Storfer, 2006). Viral load (i.e. viral copies ng\(^{-1}\) of DNA) was estimated based on a known standard curve. We used a synthetic double-stranded DNA standard by synthesizing a 250bp fragment of the major capsid protein (MCP) gene (gBlocks Gene Fragments; Integrated DNA Technologies), which is conserved among ranaviruses. Viral equivalents (VE) were rounded to the nearest whole number.

### 2.3 | Data analysis

Multiresponse models offer an opportunity to determine whether parasite abundances are correlated among sampled units, such as individuals, sites and host species. These models differ from generalized linear mixed models (GLMMs) in two key ways: (1) two (or more) response variables are modelled simultaneously and (2) correlations between response variables at each level of the model can be estimated directly through correlated random effects. As we show below, these correlation parameters are included directly in the models, allowing for the simultaneous and joint estimation of correlations at each level alongside other model parameters. Here we fit one model for each pair of parasites in our dataset (six total models), with each model containing four levels of random effects: individuals (i.e. amphibian hosts), sites (i.e. ponds), host species and metacommunities (i.e. collections of ponds within a contiguous geographical area, Figure 1). Specifically, we modelled the observed parasite counts \(y\) for individual hosts \(i = 1,\ldots, I\) at site \(j = 1,\ldots, J\) within the metacommunity \(l = 1,\ldots, L\). Each individual host also was a member of a host species \(k = 1,\ldots, K\), which could be present at any site. These models took the following general form:

\[
\begin{align*}
\eta_i & \sim \text{Poisson}(\lambda_i) \\
\eta_i & = \exp(\alpha_k + \eta_{iA} + \eta_{iB} + \delta_{i,j} + \gamma_{i} + \epsilon_i) \\
\gamma_j & = \exp(\sigma_{\text{SVL}}) \\
\lambda_k & = \exp(\sigma_{\text{SVL}} + \delta_{l})
\end{align*}
\]

where the subscripts \(A\) and \(B\) indicate identical parameters associated with the two parasites. The response variables \(y\) are parasite counts, which are distributed as Poisson random variables with rate parameter (and expected value) \(\lambda\). Overdispersed-Poisson distributions were chosen for computational tractability within the MCMCGLM package framework (see below), although in practice other probability distributions (negative binomial or overdispersed lognormal) may work equally well for parasite count data.

As seen in Equation 1, each \(\lambda\) is a function of six, distinct terms which together account for the model-estimated abundance of each parasite within each individual host:

1. The \(\alpha\) terms are global intercepts for each parasite
2. The \(\eta\) terms are the site-level random effects on parasite counts for each site \(j\)
3. The \(\theta\) terms are host species-level random effects for each of the five host species, \(k\), and indicate how abundant each parasite is within each host
4. The \(\gamma\) terms are metacommunity-level random effects for each of the metacommunities, \(l\), and indicate how abundant each parasite was within each metacommunity
5. The \(\delta\) terms are the host species-specific random effects of snout-vent length (SVL) for each host species, \(k\), which account for the variation in parasite counts associated with individual host size
6. The \(\epsilon\) terms are individual-level random effects, which allow us to estimate correlations in parasite abundance at the individual level and account for overdispersion in parasite counts

Three of the six terms above (sites \(\eta\), host-species \(\theta\) and individuals \(\epsilon\)) were modelled as multivariate normal random variables (using \(\eta\) as an example):

\[
\begin{pmatrix}
\eta_{iA} \\
\eta_{iB}
\end{pmatrix}
\sim N\left(0, \begin{pmatrix}
\sigma^2_{\eta_A} & \rho_{\eta_A\eta_B} \sigma_{\eta_A}\sigma_{\eta_B} \\
\rho_{\eta_B\eta_A} \sigma_{\eta_A}\sigma_{\eta_B} & \sigma^2_{\eta_B}
\end{pmatrix}\right)
\]

In the example above, the \(j\)-total site-level random effects for each parasite have a mean of zero and a standard deviation parameter \(\sigma_{\eta_j}\).

Most importantly, this formulation also includes a \(\rho_{\eta_j}\) parameter, which represents the correlation between the site-level random effects for the two parasites. The 18 total \(\rho\) parameters (one each at the individual-, site- and host-species level for each pair of parasite species) are the parameters we are most interested in evaluating to assess parasite associations across scales. Positive and negative values of \(\rho\) indicate the degree to which abundances are correlated among individuals, sites and host species, respectively, when controlling for the other effects in the model.

The other three terms in Equation 1 (the intercepts \(\alpha\), metacommunity effects \(\gamma\) and SVL effects \(\delta\)) were estimated as univariate normal random variables. Both the intercept and metacommunity effects were modelled with mean equal to 0 and standard deviations \(\sigma\). For simplicity, we chose not to include correlations among metacommunities, but in practice there is no barrier to doing so. Lastly, the \(\delta\) were modelled with a mean equal to \(\bar{\text{SVL}}\) (e.g. the average effect of SVL across host species) and standard deviation \(\sigma_{\text{SVL}}\).

Between any two pairs of parasite species, this model formulation estimates a single correlation parameter \(\rho_{\eta_j}\) at the individual-level regardless of host species. In other words, we assume a priori that there is no variation in the individual-level correlations for different host species. We relaxed this assumption in a set of six additional multiresponse models, where the individual-level variance/covariance matrix (equivalent to Equation 2 for \(\epsilon\)) was subdivided into five separate covariance matrices, one for each host species. We refer to these modified models as species-specific multiresponse models. Model parameters were estimated by drawing 1,000 samples from their joint posterior distributions using the Markov Chain Monte Carlo (MCMC) algorithm implemented the MCMCGLM package (Hadfield, 2010) in R (R Core Team, 2013) (see prior distributions and MCMC chain specifications in Appendix S1). Although the outlined framework does not depend on Bayesian model fitting, current limitations of existing likelihood-based implementations for modelling correlations and associations emphasize the additional utility of Bayesian procedures for assessing parasite correlations across scales (see Joseph et al., 2016).
3 | RESULTS

Overall, we detected each specific amphibian pathogen at 57.7% (B. dendrobatidis), 65.9% (ranaviruses), 85.7% (Echinostoma) and 63.7% (R. ondatrae) of the 93 sampled sites. With respect to co-occurrence, 7.9% of sites supported only one of the parasites, 17% supported two, 31.8% supported three and 43.2% had all four detected. No sites were free of infections, although 312 individual hosts lacked any of the included parasites. Among the 2,152 sampled hosts, Echinostoma was the most commonly encountered parasite overall, which occurred in 45.6% of hosts, followed by R. ondatrae (38.8%), ranavirus (32.7%) and B. dendrobatidis (17.1%). The full dataset and computer code necessary to reproduce all of our results in r are deposited at Data Dryad. Posterior distributions for all estimated parameters (e.g. site and species effects) and species-mean-abundance estimates based on single response models are reported in the supplementary material (Appendix S1). Below we focus specifically on correlation estimates between parasites at different hierarchical levels.

3.1 | Correlations among individual hosts

Infections by different parasite species showed a mix of positive and negative correlations across both host species and parasite classifications (micro- vs. macroparasites). We found positive correlations between the abundance of R. ondatrae and both Echinostoma \( p = 0.31 \), 95% CI = [0.22, 0.40], Figure 2) and ranavirus \( p = 0.13 \), [-0.01, 0.27], Figure 2). These associations were also consistent in direction across different host species. The abundances of Batrachochytrium dendrobatidis and ranavirus also correlated positively within P. regilla \( p = 0.23 \), [0.01, 0.44], Figure 3), despite showing a weak overall correlation when host species were pooled (Figure 2). Intriguingly, despite the generally positive associations between R. ondatrae and Echinostoma and between R. ondatrae and ranavirus, the abundance of Echinostoma correlated negatively with B. dendrobatidis \( p = -0.18 \) [-0.29, -0.06]) and ranavirus \( p = -0.13 \) [-0.28, 0.01], Figure 2). Both negative correlations were consistent across host species, although the host species with the best-supported correlations varied (Figure 3).

3.2 | Correlations among sites and species

At the site (pond) level, parasite associations tended to mirror those observed in individual hosts with a few important exceptions. Consistent with individual host-level patterns, we detected positive correlations between the abundances of R. ondatrae and Echinostoma \( p = 0.32 \) [0.04, 0.56], Figure 4) and negative correlations between Echinostoma and B. dendrobatidis among sites \( p = -0.22 \) [-0.50, 0.09], Figure 4). However, in contrast to observations at the individual-level, R. ondatrae and ranavirus abundances correlated negatively at the site-level \( p = -0.35 \) [-0.64, -0.03], Figure 4), indicating that sites with greater R. ondatrae abundance had lower ranaviral abundance. Additionally, despite finding no individual-level correlation between R. ondatrae and B. dendrobatidis, these infections correlated positively at the site-level \( p = 0.28 \) [-0.11, 0.57], Figure 4). Among host species, we found no evidence for correlations in parasite mean abundance (Figure S1), although our power to detect such correlations with only five host species was low.

4 | DISCUSSION

Most empirical work on parasite coinfections to date has focused either on assessing the consequences of polyparasitism for individual host susceptibility and pathology (Ezenwa & Jolles, 2015; Munson et al., 2008; Pedersen & Antonovics, 2013; Romancic et al., 2011), or on evaluating the degree to which parasite species correlate within host populations (Jolles et al., 2008; Lello et al., 2004; Telfer et al., 2010). Relatively little research has explored how parasite interactions or associations at one scale extend to influence—or are influenced by—patterns at other scales, despite theoretical and empirical evidence suggesting such cross-scale effects could be important (Hellard et al., 2015; Kuris & Lafferty, 1994; Tompkins et al., 2011). This line of inquiry is important for at least two reasons. In some cases, field-based correlations between parasites may help identify candidate interactions between symbionts for subsequent experimental investigation. However, given that even strong correlations may not reflect true parasite interactions (e.g. Fenton, Knowles, Petchey, & Pedersen, 2014), an equally relevant goal is to characterize the relative importance of specific hosts, species or sites in affecting host exposure to multiple parasites concurrently.

Here we illustrate a framework for investigating parasite associations and coinfections in multiscale datasets that (1) explicitly models correlations between parasite species, (2) uses a hierarchical structure to explore effects across multiple scales and (3) captures uncertainty in the parameters of interest. Our results highlight four primary advantages of the framework for investigating associations among parasites in multiscale data. Most importantly, multiresponse models do not imply a direction of causality for parasite associations; instead, correlations between parasites are modelled directly through the use of multivariate normal random effects, making these models particularly applicable to cross-sectional data. Second, by utilizing both parasites as response variables, researchers can apply the appropriate probability distributions (e.g. overdispersed Poisson, binomial, negative binomial, gamma or mixture models) for each parasite—rather than just the (often arbitrarily) assigned “response” species. The use of appropriate distributions also removes the necessity for data transformations that can sometimes impede biological interpretation (e.g. log + 1, arcsin, square-root). Third, a hierarchical modelling framework allows for correlations and associations to be estimated over an unlimited number of nested and crossed scales, provided sufficient data are available. Including multiple scales within the same statistical model controls for associations at one level when calculating associations at another (Joseph et al., 2016). Finally, hierarchical modelling also allows one to avoid subsetting the data based on arbitrary sample size cutoffs: parameters for sites or species can be estimated even when sample sizes are low by “shrinking”
estimates towards the mean using partial pooling (Gelman, Hill, & Yajima, 2012).

By applying this framework to four pathogens of amphibians, including micro- and macroparasitic infections, we identified multiple correlations between infections that extended among sites, individuals and host species. Most sampled wetlands and hosts supported more than one pathogen species, emphasizing the importance of investigating parasite associations. Among individual hosts, correlations varied in both magnitude and direction as a function of the specific pair of coinfesting parasites. Many of the estimated coefficients were large enough to suggest that parasite interactions could be important in determining infection outcomes, despite the relative

![Figure 2](image-url)
rarity of work on coinfection in amphibian hosts (e.g. Hoverman et al., 2012; Johnson & Buller, 2011; Romansic et al., 2011). Although the outcome of parasite interactions is often believed to depend on the types of co-infecting parasites, with facilitation being more likely when a micro- and macroparasite co-occur than when both parasites are of the same type, associations detected in this

FIGURE 3 Individual-level correlations of parasite abundance for each host species. Correlations for individual host species were obtained by fitting multiresponse models with separate individual-level variance-covariance matrices (rather than a single matrix) for all host species. Each column indicates results from a single fitted model. Circles show random draws of 1,000 point estimates (per column) of individual effects ($\varepsilon_i$) from the posterior distribution. The number of points in each panel is proportional to the number of individuals sampled for each host species. A solid red line indicates cases where zero fell outside the 95% credible interval for the correlation, whereas dashed red lines indicate where zero fell outside the 90% (but not 95%) interval. All remaining correlations are indicated by dotted lines in order to visualize the direction of the mean correlation. Blank panels indicate cases where $\varepsilon_i$ was not estimated due to just a single $T. \text{granulosa}$ individual infected with $Echinostoma$
methods in ecology and evolution

STUTZ et al.

Study did not follow a simple classification paradigm. For instance, despite their ecological similarities, the trematodes *R. ondatrae* and *Echinostoma* exhibited contrasting correlations in abundance with the two micro-parasites: *Ribeiroia ondatrae* and ranavirus loads correlated positively within hosts, whereas *Echinostoma* infections in amphibian kidneys correlated negatively with both ranavirus and *B. dendrobatidis*. The two micro-parasites also correlated positively with one another. Thus, while the positive correlation between *R. ondatrae* and ranavirus is consistent with T-helper cell induced facilitation between macro- and micro-parasites (Graham, 2002), the negative relationship between *Echinostoma* and ranavirus suggests an alternative mechanism could be operating, such as direct

**FIGURE 4** Site-level correlations of parasite abundance. Results from multiresponse models for pairs of parasites are given at the intersection of rows and columns. Histograms below the diagonal show the sampled posterior distributions of the site-level correlations ($\rho_{ij}$). The histograms with shaded blue and green areas indicate correlations with positive and negative posterior means respectively. The percentage given indicates the largest mean-centred credible interval that did not include $\rho_{ij}$ equal to zero. Scatterplots above the diagonal show posterior mean estimates (circles) of the site-effects $\eta_j$ for each parasite. Horizontal and vertical grey bars show the 95% credible intervals for the $\eta_j$'s. A solid black line indicates cases where zero fell outside the 95% credible interval for the correlation, whereas dashed lines indicate where zero fell outside the 90% (but not 95%) interval. A lack of line indicates that the 90% credible interval included zero.
competition for host tissue space (e.g. within the kidneys) or indirect pathways such as immunological regulation or priming (Taylor, van der Werf, & Maizels, 2012). Positive correlations in infection patterns could also be related to host behaviour: infections that cause reductions in host activity or avoidance behaviours can enhance exposure to additional parasites (see Johnson and Hoverman 2014; Preston, Boland, Hoverman, & Johnson, 2014).

Perhaps most importantly, the applied analysis offered an opportunity to assess how patterns of correlated infection shifted from individual hosts to the among-site scale. In some cases, the form and magnitude of the correlation persisted between pairs of parasites as the analysis moved across scales; in others, it changed or even reversed directions. For instance while the positive association between R. ondatrae and Echinostoma spp. persisted at both scales, the correlation between ranavirus and the two trematode species either changed direction (with R. ondatrae) or was present at the individual but not site level (with Echinostoma). If the positive link between ranavirus and R. ondatrae is caused by immune-mediated facilitation, then the negative correlation in infection among sites could be the result of increased mortality of coinfected individuals, as has also been suggested for interactions between bovine tuberculosis and nematode worms in African buffalo (Ezenwa & Jolles, 2015). Such a result would be consistent with the deleterious effects of both ranavirus and R. ondatrae on larval amphibian survival (Hoverman et al., 2011; Johnson et al., 2012). However, such site-level relationships may also reflect strong environmental or biotic differences among sites that outweigh parasite interactions. The strong positive correlation between the trematodes R. ondatrae and Echinostoma, for instance likely reflects covariation in host exposure consistent with their shared use of rams horn snails (Helisoma spp.) as intermediate hosts (Joseph et al., 2016), particularly given that previous experimental research has shown these parasites interact antagonistically within hosts (Johnson & Buller, 2011). Similar forms of scale-dependent associations have also been reported when using alternative forms of hierarchical modelling approaches (e.g. Bayesian model-fitting; Joseph et al., 2016) and for other trematode parasite systems (e.g. Kuris & Lafferty, 1994). Collectively, these observations set the stage for collection of additional covariate data to help account for co-exposure as well as targeted, follow-up experiments to test hypothesized links between parasite pairs.

Finally, these results showed that, in many cases, the estimated individual-level correlations were highly similar among host species. Because most previous studies have focused on a single host species, whether host species respond differentially to coinfection remains unclear. Our findings indicated that taxonomically diverse amphibian hosts, including representatives of four families and two orders, had similar patterns of parasite association and coinfection within and among wetlands. This consistency suggests that drivers (both environmental and interaction-induced) of correlated infections may not be limited to particular species, and thus the impact of concomitant infection on emerging infectious diseases could influence the entire host community. Additionally, similar responses to infection may facilitate further experimental investigation, because easier-to-use host species may act as representatives of other hosts that cannot be collected from the wild. Thus far, the complexities of incorporating data on multiple host species, multiple parasites and across replicate communities have limited opportunities to explore such questions.

From a conservation standpoint, pathogen co-occurrence and coinfection in amphibians has the potential to influence patterns of individual pathology and, by extension, host population persistence. Each of the parasites included here has been shown to cause host damage or elevated mortality in amphibian hosts. For the trematodes, these effects are intensity-dependent, such that the risk of osmoregulatory disruption (Echinostoma) or limb malformations and mortality (R. ondatrae) depends on the infection load (Holland et al. 2007; Johnson et al., 2012). Although both ranavirus and B. dendrobatis are capable of replicating on or within amphibian hosts, evidence also highlights the importance of infection intensity in driving pathology (e.g. Briggs et al., 2010; Wuerthner, Hua, & Hoverman, 2017), emphasizing the value of analytical approaches such those described here that test for correlations in load among parasites. One of the remarkable observations from this dataset is how widespread and prevalent coinfections were among amphibian hosts: all sites sampled supported at least one infection, and 75% supported two or more pathogens—despite variation in sample sizes. Similarly, a 2010 survey in the same region indicated that each infection occurred in 45% or more of sampled wetlands, with at least two of the four co-occurring in 68% of sites and all four pathogens evident in 13% (Hoverman et al., 2012). Collectively, these observations—as alongside laboratory experiments that illustrate additive or synergistic effects of coinfection in amphibians (Romansic et al., 2011; Johnson & Buller, 2011; Johnson et al., 2013; Wuerthner et al. 2017)—emphasize the widespread nature of pathogen co-occurrence and the importance of sampling for infection assemblages in ongoing monitoring or conservation efforts.

The role of parasite interactions in the spread of disease in wildlife remains an open question in disease ecology (Ezenwa & Jolles, 2015; Ratnieks & Carreck, 2010; Tompkins et al., 2011). Evidence from experimental (Ezenwa & Jolles, 2015; Graham, 2008; Pedersen & Antonovics, 2013) and observational (Jolles et al., 2008; Lello et al., 2004; Telfer et al., 2010) studies has confirmed that parasite interactions can affect the risk of infection, the severity of disease pathology and transmission to other hosts. However, clear examples of how associations between parasites extend beyond individual hosts to affect processes across space and host species remain rare, despite their importance for understanding disease management in complex ecological communities. Large-scale, observational studies will become increasingly important for identifying potential linkages among parasite species in multihost, multiparasite communities. The framework presented here offers an enhanced opportunity to fit statistical models to such large, multiscale and multihost cross-sectional datasets, allowing for direct comparisons of the direction and magnitude of correlations across levels (site and individual) or within different host species. We emphasize, however, that the multiresponse approach outlined here is
ACKNOWLEDGEMENTS

We thank Dana Calhoun, Melina Allahverdian, Kelly DeRolf, Jackie Gregory, Emily Hannon, Jeremy Henderson, Megan Housman, Aaron Klingborg, Bryan LaFonte, Keegan McCaffrey, Travis McDevitt-Galles, Mary Toothman and Vanessa Wuerthner for assistance with data collection and sample processing and Max Joseph for helpful discussions of data analysis. This research was supported by funds from the National Institutes of Health, Ecology and Evolution of Infectious Diseases Program (R01GM109499), the National Science Foundation (DEB-1149308) and the David and Lucile Packard Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This work utilized the Janus supercomputer, which is supported by the National Science Foundation (award number CNS-0821794) and the University of Colorado.

AUTHORS’ CONTRIBUTIONS

A.B., C.B., J.H., J.R., and P.J. planned the research; C.B., J.H. and P.J. collected the data; W.S. devised the approach, performed the analysis and wrote the draft of the manuscript; all authors contributed to revisions and approved final submission.

DATA ACCESSIBILITY

The data, metadata and R script associated with this article are available through Dryad Digital Repository https://doi.org/10.5061/dryad.14nr4 (Stutz et al., 2017).

ORCID

Pieter T. J. Johnson http://orcid.org/0000-0002-7997-5390

REFERENCES


**SUPPORTING INFORMATION**

Additional Supporting Information may be found online in the supporting information tab for this article.

---

Appendix S1.

Description of priors and MCMC chain parameters for multiresponse and occurrence models

In applying our framework to other systems, we emphasize that care must be taken with prior specifications for covariances in Bayesian hierarchical models (Barnard et al. 2000; Gelman & Hill 2006). Standard inverse-wishart prior for variance/covariance matrices are easy-to-apply, conjugate priors for Gaussian latent variables, and are the default priors used in MCMCglmm. Generally speaking, if most of the likelihood distribution for variance parameters falls away from variance = 0, then these inverse-wishart parameters will be largely uninformative for the variances (especially when the number of groups is large). Non-informative (i.e. flat) improper inverse-Wishart priors can also be specified in MCMCglmm. An alternative, also implemented in MCMCglmm, are zero-centered half-Cauchy priors, which are vaguely informative (in that they reduce the prior support for extremely large variances), but can handle likelihoods close to zero and may work better when the number of groups is small (Gelman 2006).

However, inverse-wishart priors do not separate the variance and correlation components of the covariance matrix, and difficulties in model fitting can occasionally arise when using them to fit covariance models of the type illustrated here (Barnard et al. 2000). In these cases, the variance priors may need to be strengthened considerably for models to be fit. This occurred for us in some of the species-specific, individual-level variance-covariance matrices when sample sizes and variances were low for particular host-species/parasite-species combinations. The posterior distributions for these host/parasite combinations should be considered somewhat provisional until more data are available.

We note that a number of alternative prior specifications for variance-covariance matrices
have been proposed (Barnard et al. 2000; Gelman 2006; Huang & Wand 2013), but their application is limited to programming languages such as BUGS (Lunn et al. 2009), Stan (Stan Development Team 2014), or non MCMC approaches like R-INLAA (Rue et al. 2009). These languages allow more freedom for prior specification and model parametrization than *MCMCglmm*, at the cost of less initial user-friendliness for most practicing ecologists. Consequently, using an R package like *MCMCglmm* may make the most sense for beginning Bayesian analysis. Information on fitting priors in MCMCglmm is given by Hadfield (2010), though we recommend also reading the notes available for the MCMCglmm package at the CRAN website (https://www.cran.r-project.org/) as well as Gelman (2006) for further reference and clarification.

*MCMC chain parameters*

The length of all MCMC chains and the thinning intervals were heuristically optimized to reduce autocorrelations below 0.1 for all estimated parameters. Chains were subsequently run long enough to sample 1000 times from the joint posterior. Burn-in periods (i.e. the number of iterations prior to sampling from the joint posterior) were set to be one-half of the entire MCMC chain after optimizing the thinning interval.

*Single response models for estimating and comparing mean abundances of parasites (within hosts) and host-species effects.*

One additional goal of our study that was unrelated to coinfection was to show how hierarchical, generalized linear models can be used to better understand observed differences in the rates of infection among host species. Species may differ in their infection levels either
because they tend to occur in sites or metacommunities that vary in parasite abundance due to external (e.g. environmental) factors or because of factors relating specifically to their species identity (exposure, susceptibility, immune response). The latter is best thought of as the effect of host-species *sensu stricto* on infection, while the former encompasses all the factors (both internal and external) that determine the level of infection within each host species (on average).

We used hierarchical, single-response models to estimate: (1) the mean abundance of each parasite within each host species and (2) species-specific effects on the abundance of each parasite. The former measure indicates which species tended to be more infected overall, while the latter estimates the contribution of species-identity to parasite abundance after accounting for variation across individuals, sites, and metacommunities.

Our single response models were standard hierarchical, GLMMs (Gelman & Hill 2006).

These models took the form:

\[
\begin{align*}
    y_i & \sim \text{Poisson} \left[ \exp \left( \alpha + \eta_j + \theta_k + \gamma_l + \beta_j \text{SVL}_{i[j]} + \epsilon_i \right) \right] \\
    \eta_j & \sim N(0, \sigma_{\eta}^2) \\
    \theta_k & \sim N(0, \sigma_{\theta}^2) \\
    \gamma_l & \sim N(0, \sigma_{\gamma}^2) \\
    \beta_k & \sim N(0, \sigma_{\beta}^2) \\
    \epsilon_i & \sim N(0, \sigma_{\epsilon}^2)
\end{align*}
\]

where \(y_i\) is the estimated parasite count. The \(\eta_j\), \(\theta_k\), and \(\gamma_l\) terms indicate site-, species-, and metacommunity-level intercept coefficients (i.e. effects). These are modeled as normal random variables with mean = 0 and variances \(\sigma_{\eta}^2\), \(\sigma_{\theta}^2\), and \(\sigma_{\gamma}^2\) respectively. The \(\beta_k\) term indicates the \(k\) species-specific slope coefficients associated with individual SVL. \(\beta_k\) is also modeled as a normal random variable with mean = 0 and variance \(\sigma_{\beta}^2\). These models are identical in structure to equation 1 in the main text, except they are fit to only a single parasite, and thus only variances (rather than full variance-covariance matrices) are estimated for the random effects.
To fit a given model, we excluded individuals from sites where the parasite in question was not observed so that we were modeling the predicted abundance only at sites where we were 100% confident that exposure was possible (this also reduces zero-inflation associated with the lack of parasite presence). To estimate species' mean abundances and 95% prediction intervals, we first simulated 1000 abundances for each individual from the joint posteriors of the fitted models and calculated the mean abundance within each host species in each simulation. The host species effects ($\theta_k$) and 95% posterior intervals we pulled directly from fitted joint posteriors for each model.

Estimated parasite mean abundance estimates across host species varied by two orders of magnitude for both *R. ondatrae* and ranavirus and over four orders of magnitude for *B. dendrobatidis* and *Echinostoma* (Fig. S1 and Fig. S2). The 95% posterior intervals for *B. dendrobatidis* abundance were much wider than for the other parasites, reflecting greater uncertainty in *B. dendrobatidis* abundance estimates generally. A few host species stood out in their respective mean abundances. *B. dendrobatidis* abundances were estimated to be one to three orders of magnitude higher in both *B. boreas* and *P. regilla* than in the other three species. Similarly, the estimated *R. ondatrae* abundance was about an order of magnitude lower in *T. granulosa* than in other species, while the estimated Ranaviral abundance was somewhat lower in *L. catesbeianus* than in other species. Lastly, both *Taricha* species were substantially less infected overall with *Echinostoma* than the other three (Anuran) species.

The species effects indicated that there was more uniformity among species in their proclivity to infection than was indicated by the mean abundance estimates (Fig. S1 vs. Fig. S2). The *T. granulosa* effect on *R. ondatrae* abundance was no longer lower than the other species' effects (except perhaps *T. torosa*). This suggests that the substantially lower abundance of *R.
ondatrae in *T. granulosa* was likely due to the newts' primarily occurring in sites where *R.*
ondatrae abundance was lower than average over all. Similarly, the *L. catesbeianus* effect on
Ranaviral abundance was only marginally lower than the other species' effects. On the other
hand, *B. boreas* and *P. regilla* effects' on *B. dendrobatidis* abundance were roughly equal, but
both remained substantially higher than the other species' effects. Additionally, both *Taricha*
species effects remained lower than the other species' effects on *Echinostoma*. All together, these
results suggest that our host species are roughly equally prone to infection with all four parasites,
with the exception of *B. boreas* and *P. regilla* infection with *B. dendrobatidis* and *Taricha* sp.
infection with *Echinostoma*.

**References**

standard deviations and correlations, with application to shrinkage. *Statistica Sinica*, 10,
1281–1312.

on article by Browne and Draper). *Bayesian Analysis*, 1, 515–534.


Huang, A. & Wand, M.P. (2013). Simple marginally noninformative prior distributions for
covariance matrices. *Bayesian Analysis*, 8, 439–452.
