Associations between patterns of human intestinal schistosomiasis and snail and mammal species richness in Uganda: can we detect a decoy effect?

Anna-Sofie Stensgaard\textsuperscript{a*}, Thomas K. Kristensen\textsuperscript{b,c}, Aslak Jørgensen\textsuperscript{d}, Narcis B. Kabatereine\textsuperscript{e} and Carsten Rahbek\textsuperscript{a}

\textsuperscript{a}Center for Macroecology, Evolution and Climate, The Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen Ø, Denmark. www.macroecology.ku.dk
\textsuperscript{b}Section for parasitology and Aquatic Diseases, Department of Veterinary Disease Biology, University of Copenhagen, Dyr lægevej 100, DK-1871 Frederiksberg C, Denmark.
\textsuperscript{c}School of Biological & Conservation Sciences, Faculty of Science and Agriculture, University of Kwa zulu-Natal, Durban, South Africa.
\textsuperscript{d}Laboratory of Molecular Systematics, Sølvgade 83, The Natural History Museum of Denmark, University of Copenhagen, DK-1307 Copenhagen K, Denmark
\textsuperscript{e}Vector Control Division, Ministry of Health, P.O. Box 1661, Kampala, Uganda.

* Corresponding author, asstensgaard@snm.ku.dk, http://macroecology.ku.dk/people/

Abstract. In recent years, ecological research has suggested several mechanisms by which biodiversity might affect the risk of acquiring infectious diseases (i.e., the decoy, dilution or amplification effects), but the topic remains controversial. While many experimental studies suggest a negative relationship between biodiversity and disease, this relationship is inherently complex, and might be negative, positive or neutral depending on the geographical scale and ecological context. Here, applying a macroecological approach, we look for associations between diversity and disease by comparing the distribution of human schistosomiasis and biogeographical patterns of freshwater snail and mammal species richness in Uganda. We found that the association between estimated snail richness and human infection was best described by a negative correlation in non-spatial bi- and multivariate logistic mixed effect models. However, this association lost significance after the inclusion of a spatial component in a full geostatistical model, highlighting the importance of accounting for spatial correlation to obtain more precise parameter estimates. Furthermore, we found no significant relationships between mammal richness and schistosomiasis risk. We discuss the limitations of the data and methods used to test the decoy hypothesis for schistosomiasis, and highlight key future research directions that can facilitate more powerful tests of the decoy effect in snail-borne infections, at geographical scales that are relevant for public health and conservation.

Keywords. Biodiversity, Decoy effect, Dilution effect, Disease ecology, Health, Macroecology, Parasites, Schistosomiasis, Snails.
refractory or resistant to infection, resulting in ‘wasted’ transmission events (LoGuidice et al. 2003). However, the idea has since been expanded to encompass other, non-vector-borne, pathogens such as parasites with free-living stages (Johnson et al, 2009). A recent meta-analysis concludes that dilution effects occur commonly in nature and may modulate human disease risk (Civitello et al. 2015). In contrast, an earlier meta-analysis (Salkeld et al. 2013) suggests an idiosyncratic effect of biodiversity on human parasites, with a bias towards publishing only significant, negative associations between biodiversity and disease risk. As such, the generality of the dilution effect remains fiercely debated (Randolph and Dobson 2012, Salkeld et al. 2013, Wood and Lafferty 2013, Wood et al. 2014): Critics argue that the direction of such a relationship is just as likely to be neutral or positive (referred to as an amplification effect) and of varying magnitude, depending on the disease system in question, the ecological context and the geographical scale of study (Wood et al. 2014). In general, due to a shortage of vital empirical data, the understanding of the extent to which patterns of biodiversity influence different disease systems across spatial scales is still very limited (Hough 2014, Johnson et al. 2015).

While contemporary discussions of the dilution effect mainly have focused on vector-borne micro-parasites (Keessing et al, 2006, Johnson et al. 2015) historically, a “decoy effect”, describing similar mechanisms by which community diversity can reduce disease risk, has been discussed in the macro-parasite literature in the context of snail-borne infections (i.e., Chernin 1968, Upatham and Sturrock 1973, Laracuente et al. 1979, Christensen 1980, Moné and Combes 1986, Combes and Moné 1987, Johnson et al. 2009, Johnson and Thieltges 2010). The main mechanism behind the decoy effect is a reduction in disease risk, caused by attempts of the free living infectious parasite stages to infect poor or unsuitable hosts, instead of reaching suitable hosts (Upatham 1972, Upatham and Sturrock 1973, Laracuente et al. 1979, Yousif et al. 1998, Johnson et al. 2009). The decoy effect has mainly been demonstrated experimentally and several interrelated underlying mechanisms have been suggested (summarized in Johnson et al. 2010): (i) an encounter or transmission reduction, where a richer community of non-host snails acts as a “sponge” intercepting snail infective stages (Laracuente et al. 1979), eventually leading to lower or delayed output of the human infective stage, the cercaria larva (Frandsen and Christensen 1977, Johnson et al. 2009), and (ii) a reduction of susceptible hosts by adding competing or pre-dating non-host snail species to the system (Madsen 1990, Stryker et al. 1991, Mkoji et al. 1992, Pointier and Giboda 1999, Keessing et al. 2006). In the following we will refer to these mechanisms collectively as the decoy effect, as this is the most commonly used term in the snail-borne disease literature.

Lab-based experimental studies have confirmed the existence of a decoy effect for a number of snail-borne trematodes, but few studies consider the interplay between such an effect and abiotic factors, and whether a decoy effect is detectable at the large geographic scales relevant to conservation and public health. Furthermore, most studies have focused on the role of invertebrates as decoys for the miracidia larvae, and only a few studies have demonstrated similar effects of adding less susceptible vertebrate species to divert the free-living cercarial stages (but see Johnson et al. 2008). Thus, the potential for other mammals to dilute or amplify the transmission of infective stages to the definitive human hosts remains largely unknown.

Here, we investigate whether the patterns predicted by the decoy effect are detectable at a country-wide scale in Uganda, using human intestinal schistosomiasis (causative agent Schistosoma mansoni), one of the most prevalent parasitic infections of medical importance in the tropics and sub-tropics (Steinmann et al. 2006), as our disease model system. Specifically, we hypothesize the following scenarios: If a more diverse freshwater snail community exerts a strong decoy effect, this would yield a pattern in which areas with higher freshwater snail species richness have lower human parasite prevalence (i.e., a negative relationship). Likewise, we also hypothesize that a more
diverse mammal community may dilute the infection success of the parasite cercaria (the human infective stage), such that areas of higher mammal richness also have lower human parasite prevalence (i.e., a negative relationship); alternatively, more mammal species may facilitate transmission by providing more transmission pathways for *S. mansoni*, thus amplifying the disease (i.e., a positive relationship). Humans are clearly the most important definitive host for the maintenance of the *S. mansoni* life cycle, but other mammalian groups have also been demonstrated to be naturally infected (Hanelt et al. 2010, Standley et al. 2012). However, the effect of mammal diversity on *S. mansoni* infection risk has never been formally tested (Modena et al. 2008).

We apply a macroecological approach to search for patterns in accordance with the expectation of the above outlined hypotheses. We focus on species richness as our measure of diversity, but recognize that other community structural aspects, e.g., species evenness and species trait variation are likely to play equally important roles (Ostfeld and Keesing 2000, Keesing et al. 2006, Johnson et al. 2008). We use data on human schistosomiasis infection prevalence as a proxy for disease risk, and seek to account for confounding individual level and environmental level risk factors often found to be associated with schistosomiasis in a multivariate modelling framework.

**Materials and Methods**

**Data**

Schistosomiasis prevalence data

Parasitological data were available from a national epidemiological survey of *S. mansoni* in Ugandan school-children undertaken by the Ministry of Health in the period 1998-2002. Details on survey layout and design are described elsewhere (Kabatereine et al. 2004). In brief, individual level infection data from 9347 children from 153 inland schools covering all districts of Uganda (Fig. 1) were available to the current study. This dataset was previously applied to map the distribution of intestinal schistosomiasis in Uganda (Stensgaard et al. 2005), and the full dataset is publicly available in the open access global neglected tropical disease database$^1$ (Hurlimann et al. 2011).

Species richness data

In macroecological analysis the unit of investigation often consists of gridded layers of species richness at varying resolution. Mammal species richness data were extracted at a $1^\circ \times 1^\circ$ grid resolution from the extensive African database (Brooks et al. 2001), held at the Natural History Museum of Denmark, University of Copenhagen. These data currently represent the most complete dataset available on the distribution of sub-Saharan vertebrates, compiled from published as well as unpublished survey data sources.

However, no similar data are currently available for mollusks. Thus, to construct a comparable species richness layer for freshwater snails, a grid of $1^\circ \times 1^\circ$ resolution matching that of the mammal richness grid was overlaid on Uganda (using ArcGIS v.10.3 (ESRI, Redlands, USA)), and all

---

$^1$ http://www.gntd.org
known records of freshwater snails were aggregated within this grid (summing the number of snail species observed in each grid cell). The majority of the snail data came from a country-wide malacological survey undertaken in 2000-2003, at 94 sampling sites across Uganda (see Jørgensen 2003 and Stensgaard et al. 2006 for further details). These data were supplemented with records from the mollusk specimen collections held at Natural History Museum, University of Copenhagen (45 localities), and an additional 2 localities from literature (Brown 1994). To minimize effects of under-sampling, the ‘true’ number of species in each grid cell was estimated using EstimateS v.7.00, (R.K. Colwell)\(^2\) using the non-parametric species richness estimator Chao2 (Chao 1987, Colwell and Coddington 1994).

**Environmental data**

Climate and other environmental factors, such as temperature and water availability are known to highly influence the epidemiological patterns of schistosomiasis (Brooker 2007). To account for these effects, a number of climatic and environmental proxies such as land surface temperature (LST), rainfall, normalized difference vegetation index (NDVI) and altitude were included in the analysis (full list in Table S1; details of sources and resolution of the these data can be found in Stensgaard et al. 2005).

**Statistical analysis**

**Multilevel, non-spatial analysis**

To test the hypotheses outlined in the introduction the data were analysed using progressively more complex and highly parameterised logistic regression models, with the infection status of each child as the binary response variable. Given the hierarchical structure of the data, logistic mixed effect models were applied to identify and account for as many potential factors associated with \(S. \text{mansoni}\) infection as possible. Clustering of individuals in schools and of schools in grids was accounted for by including location-specific exchangeable random effects at school- and/or grid-level, modelled with a mean zero normal distribution and unknown variance \(\sigma^2\). Individual age and gender were included as individual-level explanatory variables. Potential environmental confounders were included as school-level explanatory variables, whereas species richness measures were included as grid-level explanatory variables.

Before fitting the multivariate models, bivariate mixed effects logistic regression models were first used to examine the relationship between the outcome variable and each potential covariate in isolation (see Table S1, Supplementary Materials). This was done to avoid the problems associated with multicollinearity among similar variables. From each “theme” of highly collinear significant environmental covariates (Table S1), one was chosen based on the fit according to the Akaike Information Criterion (AIC) (Akaike 1973). The bivariate analysis was also used to examine for potential non-linear relationships between the covariates and parasitaemia outcome. The bi-variate analysis was performed using the `melogit` command in STATA/SE v.10, StataCorp LP, College Station, USA. All multivariate analyses were performed in a Bayesian framework using OPENBUGS v.3.3.1 (Imperial College & Medical Research Council, London, UK) (Lunn et al. 2009), where model parameters were estimated using Markov chain Monte Carlo (MCMC) simulation (Gelfand and Smith 1990).

**Bayesian geostatistical (spatial) models**

Bayesian geostatistical formulations of the above models were fitted to take into account the spatial correlation often found among parasitaemia survey locations (Diggle et al. 1998). Specifically, a generalised linear geostatistical model in a Bayesian framework that account explicitly for spatial autocorrelation and uncertainty in the input data and model parameters was formulated. Spatially correlated random effects were introduced at every location (school latitude-longitude coordinates or grid centroid coordinates), generally assuming that the spatial correlation decreases as distances between locations increases. The Bayes-
ian geostatistical approach was chosen because the effect of covariates and spatial heterogeneity, or clustering, can be modelled simultaneously, to ensure more accurate parameter and significance estimates when modelling spatially correlated data (Diggle et al. 1998). Model fit was implemented in OPENBUGS v.3.1.1 via Markov Chain Monte Carlo methods which allow flexibility when fitting complex models and avoid the computational problems and asymptotic inference encountered in likelihood-based fitting (Diggle et al. 1998).

To assess the direction and strength of any associations between infection and species richness measures, three models were compared: the first, model A, was a non-spatial mixed effect logistic regression model, including all demographic, environmental and species richness co-variates, along with two exchangeable random effects to take the clustering at the school and grid-level into account. Model B was a spatial model, similar to model A, except spatial correlation between schools or grid cells was introduced on the school and grid specific random effect. Finally, model C (identical to model B, but without the species richness co-variates) was constructed to evaluate the contribution of species richness to overall model fit.

The best fitting model (i.e., spatial versus non-spatial model, and models with/without species richness covariates), was identified based on the deviance information criterion (DIC), which is a Bayesian model comparison criterion (Spiegelhalter et al. 2002) A smaller value of the DIC indicates a better model fit to the data. Details on the Bayesian geostatistical model formulation are available in Supplementary Materials (Appendix S2).

Figure 2. Gridded species richness maps for Uganda at 1° resolution. (A) Observed freshwater snail richness grid based on snail sampling surveys conducted in 2000-2003, supplemented with historical records. Each grid cell is based on 2–36 sampling locations, except northern grids which had to be excluded due to under-sampling. (B) Estimated freshwater snail richness (Chao2) and (C) mammal species richness.
Results

The gridded snail species richness measures, i.e., observed numbers and Chao2 estimated richness, can be seen in Figure 2, along with the gridded mammal species richness. In total 39 freshwater snail species were identified from Uganda. The true diversity is likely to be even higher. Due to unresolved taxonomic issues some species could not be discriminated more finely than the level of genus or species complex (details provided in Appendix S1, Supplementary Materials).

Bi-variate analysis

The initial bivariate regression revealed significant negative associations between infection status and both measures of snail species richness (Table S1). The strongest correlation was found with observed snail species richness ($OR = 0.88$, $p < 0.001$), although the estimated richness as measured by Chao2 ($OR = 0.91$, $p < 0.001$) showed a better fit to the data as measured by AIC. Hereafter, we focus on Chao2 as the measure of snail richness.

Multivariate modelling

In the multivariate non-spatial regression model the negative association with snail richness remained significant after adjusting for child age, gender and environmental factors (Model A, Table 1).

However, when a school-level spatial random effect was included, only individual level factors remained significantly correlated with schistosome infection prevalence (model B, Table 1). There was a strong spatial correlation between schools up to a distance of 114.8 km ($95\% \text{ CI} =$...
50.3–475.3) (model B, Table 1), which was not accounted for in the non-spatial model (model A, Table 1). Compared to the non-spatial model A, the spatial model B also showed a substantially better fit (DIC value improved from 3885 to 2509), justifying the inclusion of a spatial random effect at this level. Model B also had a substantially lower DIC than model C (2509 vs 3225, Table 1), indicating that including species richness co-variates did improve model fit considerably, even though these co-variates were not significant in model B. Note that the inclusion of a grid-level random effect in the multivariate models was not justified, as its variance approached zero and it caused an increase in DIC (see Table S2 and S3). Furthermore, the spatial range at which autocorrelation became less than 5%, was only 63.4 km (which is below the resolution of the gridded richness data). Adding an extra random effect also prolonged the MCMC chain convergence time, and as there were no substantial changes to parameter estimates and significance levels, we choose to present the results for the more parsimonious models (without the grid-level random effects). However, the results from the models with both random effects can be seen in Table S2 and Table S3.

Discussion

To our knowledge, this is the first time the decoy effect (A.K.A. the dilution or amplification effect) has been investigated empirically for schistosomiasis at a country-wide scale, providing valuable input to this emerging field of study. Using non-spatial mixed effect logistic regression models, we first observed a geographical co-variation between schistosomiasis prevalence patterns and measures of freshwater snail richness in Uganda that was best described by a negative correlation, in accordance with the expectations of the decoy effect. However, this association became non-significant after incorporating a spatial component in the model (model B, Table 1). The inclusion of a spatial component appears to be justified (i.e., lower model DIC in the spatial model, and evidence of strong spatial autocorrelation between survey sites), which demonstrates the importance of accounting for spatial correlation among surveys locations, in order to avoid over-estimation of the significance of the regression coefficients (Cressie 1993).

Despite the lack of a significant association in the final model, it must be pointed out that the association with snail richness was always negative (never positive), and that the inclusion of species richness as a covariate in the infection risk model substantially improved model fit. Other reasons for a lack of a strong negative association could be an off-setting effect caused by increased host snail reproductive output: Because trematode infection often causes snail castration (Gerard and Theron 1997), a low snail infection rate in more species diverse communities could result in such an apparent facilitating effect (see Johnson et al. 2009).

With regards to mammal richness, no significant associations were detected with schistosomiasis at any stage of the analysis. Nor did we at any time observe any significant positive associations between infection risk and mammal or snail species richness, and thus cannot find support for an amplification effect of higher species richness (i.e., a positive associations between species richness and parasitaemia).

Several issues pertaining to the data and study design are of course likely to have influenced the results and the ability of the applied models to detect country-scale signals of decoy and/or amplification effects and needs addressing. Here we will discuss the major limitations of the current study, and suggest possible ways to improve survey designs and methods used to investigate disease-diversity relationships for snail-borne parasitic diseases such as schistosomiasis.

First and foremost, the precision of the results presented here is of course crucially dependent on the quality and resolution of the data input. A major challenge arises from the fact that snail and parasitological surveys were conducted independently of each other (as is often the case). This independence may cause misalignment in space and time between the snail and schistosomiasis data. There are several ways to try and deal with such misalignment. Here, we chose to
aggregate the snail survey data into grid cells of 1° resolution, matching that of available mammal richness data. It can be argued that this scale may be too coarse to resolve prevalent snail- and mammal richness patterns and to detect significant associations with infection risk. Ideally, a finer-scale landscape of snail and mammal richness information with which to pair the prevalence data at the school locations should be used. However, the scarcity of snail survey locations does not allow this. Alternatively, species distribution modelling could be deployed to develop individual snail species distribution maps and stack these to produce a finer scale richness map. This is not an unusual approach in macroecology (Algar et al. 2009); however, we believe the biases and uncertainties that can arise from this approach (Guisan and Rahbek 2011) are likely to outweigh the advantages in this case. The coarser scale grid approach furthermore has the advantage of being comparable to the resolution of mammal richness data (the best available). Importantly, the grid approach also allows us to account for undersampling in the snail data (which is not an uncommon phenomenon in freshwater snail surveys). Ideally, the immediate vicinity of each of the 153 schools in the present should have been exhaustively sampled for freshwater snail species, to ensure a better alignment of the data in space and time. In future schistosomiasis surveys we thus strongly encourage more collaboration between snail ecologists and health scientists in large-scale investigations, to allow more powerful tests of the decoy effect hypothesis.

Secondly, schistosomiasis is known to be a multi-factorial disease, and the focal human prevalence patterns (here used as a proxy for actual infection risk) are of course highly influenced by a myriad of potentially interacting factors, of which it was only possible to adjust for some with the data at hand. Likewise, regional patterns of species richness are a consequence of many interacting factors, such as productivity, competition, geographical area, topographic heterogeneity, historical or evolutional development, regional species dynamics, environmental variables, and human activity (Currie 1991, Rahbek and Graves 2001, Jetz and Rahbek 2002, Hawkins et al. 2003, Rosenzweig et al. 2008). It is thus possible, that similar factors may be important for sustained disease transmission as well as the generation or maintenance of high levels of species richness, and that any observed correlations simply mirror some common underlying environmental drivers.

Finally, we used a rather simple measure of diversity: species richness. Other studies on the dilution and decoy effect have indicated that the strength of the effect in any given disease will depend as much on the identity and specific traits of diluting organisms as on species richness (Logiudice et al. 2008). Likewise, the relative abundance of dilution or decoy hosts relative to focal hosts plays an important role (Upatham and Sturrock 1973), indicating that studies have to consider both the role of species traits (e.g., species identity) as well as the density of the diluting organisms (e.g., species evenness). However, such measures are very rarely collected in large-scale epidemiological surveys and no reliable details of such measures were available for this study.

So what datasets would we ideally want to more rigorously test the decoy effect in schistosomiasis (and other snail-borne infections) at a scale relevant for public health (i.e., control programmes) and conservation? For schistosomiasis, we suggest that we must start to collect data on biodiversity that are both spatially and temporally related to the relevant measures of disease risk. This entails data on the composition and richness of both the freshwater snail and non-human mammal communities in the vicinity of the actual transmission sites, as well as data on the infection levels in humans and snails and potential reservoir hosts (the latter very rarely reported). Ideally, collection of these large-scale data should be combined with mechanistic on-site experiments, linking field and experimental data to investigate the influence of community structural aspects on parasite infection (see for instance Johnson et al. 2013). Aside from the theoretical benefits that may emerge from an approach that investigates the interplay of a decoy effect with other biotic and abiotic factors at the scale of ‘real’ ecosystems, it will be crucial in order to anticipate the
effects of climate and other global environmental changes on human parasites and their hosts in general.

Collecting this type of data is of course resource-demanding. However, given that elimination of schistosomiasis is now a declared global health goal (WHO, 2012), and although morbidity control is likely to remain the main strategy in most places in Africa (Rollinson et al. 2013), it is worthwhile to invest in the collection of such data to obtain a greater understanding of the role of biodiversity in schistosomiasis transmission. This understanding may be a crucial piece in the puzzle needed to bring transmission levels from low to zero.

Acknowledgements
The Vector Control Division, Ministry of Health, Uganda, is warmly thanked for helping with snail survey arrangements and for logistic assistance during snail sampling and for making the national epidemiological data available. The study was supported by grants from ‘WWF Verdensnaturfonden/Nuvo Nordisk ‘Biodiversitetslegatet’, the Natural History Museum, University of Copenhagen and the Danish Development Research Council. C. Rahbek and A.S. Stensgaard are thankful to the Danish National Research Foundation for its support of the Center for Macroecology, Evolution and Climate (grant number DNRF96). T. K. Kristensen, N.B. Kabaterine acknowledges financial support by the EU-funded CONTRAST project FP6-STREP-2004-INCO-DEV Project no. 032203.

References


Upatham, E.S. (1973) Field investigations on the capability of miracidia of Schistosoma mansoni (Sambon) to infect Biomphalaria glabrata (Say) under field-simulated conditions in St Lucia, West-Indies. Journal of Helminthology, 46, 277–282.


Submitted: 16 March 2014
First decision: 23 May 2014
Accepted: 6 October 2015
Edited by Susanne Fritz