Parasitology

cambridge.org/par

Circadian rhythms of trematode parasites: applying mixed models to test underlying patterns

Emily R. Hannon, Dana M. Calhoun, Sindhu Chadalawada and Pieter T.J. Johnson

Ecology and Evolutionary Biology, University of Colorado, Boulder, CO 80309, USA

Research Article

Cite this article: Hannon ER, Calhoun DM, Chadalawada S, Johnson PTJ (2018). Circadian rhythms of trematode parasites: applying mixed models to test underlying patterns. *Parasitology* **145**, 783–791. https://doi.org/10.1017/S0031182017001706

Received: 15 March 2017
Revised: 15 June 2017
Accepted: 31 August 2017
First published online: 16 November 2017

Key words:
Circadian rhythm; circular statistics; cosinor analysis; generalized linear mixed modelling; host–parasite evolution; trematode

Author for correspondence:
Pieter T. J. Johnson, E-mail: pieter.johnson@colorado.edu

Abstract

Circadian rhythms of parasites and their hosts can influence processes such as transmission, pathology and life cycle evolution. For trematode parasites that depend on free-living infectious stages (i.e. cercariae) to move among host species, the timing of parasite release is hypothesized to increase the likelihood of contacting a host. Yet, a persistent challenge in studying such biorhythms involves selection of appropriate analytical techniques. Here, we extend a generalized linear mixed modelling (GLMM) framework to cosinor analyses, thereby allowing flexibility in the statistical distribution of the response variable, incorporation of multiple covariates and inclusion of hierarchical grouping effects. By applying this approach to 93 snails infected with trematode parasites from freshwater pond ecosystems, we detected non-random rhythms in six of eight species, with variation in both the timing of peak cercariae release (between 5:10 and 21:46 h) and its magnitude (between 13 and 386). The use of GLMM yielded more accurate and precise estimates of the cosinor parameters compared with classical least-squares (LS) based on a simulation-based sensitivity analysis. The sensitivity analysis revealed that the amplitude and rhythm-adjusted mean values from the LS models diverged from the true values at some limits. We highlight the importance of novel analytical approaches for evaluating parasite circadian rhythms and investigating their underlying mechanisms.

Introduction

Many parasites and their hosts exhibit circadian rhythms that influence processes such as transmission, pathology and life cycle evolution (Mouahid *et al.* 2012; Khan *et al.* 2015; Théron, 2015). Predictable cycles in the activity or habitat use of hosts offer an opportunity for parasites to increase the probability of successful infection (Aoki *et al.* 2011; Théron, 2015). Within human hosts, for instance, microfilariae of nematodes such as *Wuchereria bancrofti*, *Brugia malayi* and *Loa loa*, aggregate in peripheral blood vessels during periods of peak feeding by their arthropod vectors, presumably maximizing transmission to biting flies and mosquitoes (Hawking, 1967; Dreyer *et al.* 1996; Aoki *et al.* 2011; Craig and Scott, 2014).

Parasites that depend on free-living infectious stages often exhibit circadian rhythms in their emergence from intermediate hosts. For instance, trematode parasites have frequently been used as a model for understanding parasite chronobiology (Lu *et al.* 2009; Théron, 2015; Soldánová *et al.* 2016). Trematodes use free-living infectious stages, such as eggs, miracidia and cercariae, to move among host species (Kuris and Lafferty, 1994; Parker *et al.* 2003). Snails or other molluscs function as first intermediate hosts, from which mobile cercariae emerge and seek out a subsequent host species (Kuris and Lafferty, 1994; Parker *et al.* 2003). Because cercariae are typically short-lived (<24 h), emergence patterns that increase the likelihood of host encounter are selectively favoured (Fried and Ponder, 2003; Karvonen *et al.* 2003). For example, cercariae of human blood flukes (*Schistosoma* spp.), which infect approximately 200 million people worldwide (Fenwick, 2012), emerge *en masse* from aquatic snails at times that parallel peak activity of local villagers (Silva *et al.* 1995; Ahmed *et al.* 2006; Mintsa-Nguéma *et al.* 2014). Intriguingly, parasite emergence patterns vary regionally in association with differences in village-specific behaviours as well as the role of alternative hosts (N'Goran *et al.* 1997; Lu *et al.* 2009; Su *et al.* 2013; Wang *et al.* 2014). Su *et al.* (2013) reported that *S. japonicum* was more likely to emerge nocturnally in regions where rodents served as important reservoir hosts, whereas strains maintained in the laboratory emerge during the day. Similar host-specific timing effects have been demonstrated for trematode parasites of wildlife (Ahmed *et al.* 2006; Steinauer *et al.* 2008; Faltynkova *et al.* 2009; Komatsu *et al.* 2014). Although the precise mechanisms controlling trematode circadian behaviours remain conjectural, environmental cues such as light availability, temperature, salinity and water pressure have been demonstrated to influence cercariae release from snails (Valle *et al.* 1973; Williams *et al.* 1984; Mouritsen, 2002; Sulieman *et al.* 2013; Paull and Johnson, 2014; Wang *et al.* 2015).

A persistent challenge in the study of circadian rhythms involves selection of the appropriate analytical approaches. A commonly applied class of analyses is circular statistics, which

© Cambridge University Press 2017

CAMBRIDGE UNIVERSITY PRESS

involve observations of directional vectors around a unit circle (Su *et al.* 2013; Théron, 2015). While circular statistics recognize the continuity of temporal data, their application to parasitological circadian rhythms is challenging for several reasons. Assigning each parasite to a specific angle assumes they are completely independent observations, even when dealing with pulses of parasites that emerge from a host within the same temporal window. Circular statistics also tend to weight parasite species that produce large numbers of infectious stages more heavily. Moreover, null data (times when parasites are not released) are ignored, despite their clear biological significance to circadian patterns. A more suitable approach for parasite circadian rhythms is therefore cosinor analysis, which involves modelling temporal data with cosine functions to identify the rhythm-adjusted mean value [Midline Estimate Statistic of Rhythm (MESOR)], peak emergence values (amplitude) and the time at which peak values occur (acrophase) (Refinetti *et al.* 2007; Cornelissen, 2014). Nonetheless, while numerous statistical programs and packages can implement cosinor regression (e.g. Lee Gierke *et al.* 2013; Barnett *et al.* 2014; Sachs, 2014; Revelle, 2016), these options vary considerably in features and flexibility. While parasitological data often follow a Poisson or negative binomial distribution (Greene, 2008), most analytical options assume a Gaussian distribution of the response, which may be difficult to approximate even following transformation. Similarly, because observations often come from multiple individual hosts or treatments, it is frequently necessary to include both covariates (e.g. experimental condition) as well as grouping terms (e.g. random intercept terms for the identity of a specific host or trial).

In the current study, we extend a generalized linear mixed modelling (GLMM) framework to cosinor analyses and then apply this approach to evaluate the circadian rhythms of eight trematode species from freshwater pond ecosystems. GLMMs allow for incorporation of different error distributions, such as binomial (e.g. infected *vs* uninfected), Poisson (e.g. count data with or without an offset term to account for varying sampling areas) and negative binomial distributions, which often better approximate the overdispersed counts inherent to parasite infection (Shaw *et al.* 1998; Bolker *et al.* 2009; Lee *et al.* 2012). Moreover, GLMMs include fixed as well as random effects (random intercept or slope terms), which help to fit hierarchically structured data and improve coefficient estimates. Thus, they recognize and allow for correlated observation to be included without introducing problems associated with pseudoreplication. By applying GLMMs to counts of cercariae collected over 24 h periods from multiple trematode species, we: (1) determined whether each parasite exhibited predictable, non-random circadian emergence patterns; (2) characterized the timing and magnitude of peak emergence (including comparison of single- and multiple cosinor models); and (3) used a simulation-based sensitivity analysis to compare the parameter estimates with the classical least-squares (LS) approach. We conclude by discussing future directions in parasite chronobiology with an emphasis on potential factors affecting the timing of cercariae release and their physiological mechanisms.

Materials and methods

Snail collection

During the summers of 2011, 2012 and 2016, we collected rams horn snails (*Helisoma trivolvis*) from pond ecosystems in the eastern San Francisco Bay region of California, USA, including Contra Costa, Alameda and Santa Clara counties. We collected snails by hand, through dipnet surveys and during seine hauls (see Johnson *et al.* 2013). After collection, we acclimated snails to treated (UV-sterilized, carbon-filtered, dechlorinated) tapwater over the course of 6 h before placing snails individually into 50 mL centrifuge tubes. Tubes were checked every 3–6 h for free-swimming cercariae; when detected, cercariae were identified to morphotype using Schell (1985). Whenever possible, we supplemented this coarse-level identification to finer taxonomic resolution using primary literature, molecular analysis and experimental infections (Johnson, 1968; Kanev *et al.* 1995; Johnson *et al.* 2013). In total, we isolated 93 infected snails: 30 snails infected with *Ribeiroia ondatrae* (Cathaemasiidae), 25 with *Echinostoma* sp. (likely *Echinostoma trivolvis*, Echinostomatidae), 16 with *Cephalogonimus* sp. (Cephalogonimidae), three with *Alaria marcianae* (Diplostomatidae), three with *Zygocotyle* sp. (Paramphistomoidea), 11 with *Australapatemon* sp. (Diplostomatidae), two with an unidentified Echinostomatidae (magnacauda morphotype) and three with an unidentified Schistosomatidae (brevifurcate-apharyngeate morphotype) (see Table 1). Infected snails were maintained at 21 °C with a 14:10

Table 1. Results of generalized linear mixed modelling approach to cosinor analysis of freshwater trematode cercariae collected in California

Trematode taxon	Snails (*N*)	Second intermediate host(s) (and citation)	Acrophase (time of peak emergence)	Amplitude (number of cercariae above baseline)	MESOR (baseline trend for emergence)
Ribeiroia ondatrae	30	Amphibians and fishes[a]	21:46 ± 00:01	13 ± 1	33 ± 1
Echinostoma sp.	25	Amphibians and snails[a]	9:44 ± 00:34	105 ± 27	57 ± 1
Cephalogonimus sp.	16	Amphibians[a]	9:21 ± 00:14	52 ± 3	58 ± 15
Australapatemon sp.	10	Leeches[b]	14:58 ± 00:22	103 ± 7	295 ± 4
Magnacauda morphotype	2	Fishes[a]	5:10 ± 01:39	187 ± 48	56 ± 7
Alaria marcianae	3	Amphibians[c]	12:11 ± 00:14	386 ± 41	146 ± 12
Zygocotyle sp.	2	Vegetation[d]	—	—	—
Brevifurcate-apharyngeate morphotype	3	Migratory birds[e]	—	—	—

For each trematode species, we present the number of infected snails included in trials, the downstream second intermediate hosts used by that parasite and the cosinor parameter estimates for species with a non-random circadian rhythm (mean value ± 1 S.E.).
[a]Orlofske *et al.* 2015.
[b]Davies and de Núñez, 2012.
[c]Johnson, 1968.
[d]Núñez *et al.* 2011.
[e]Abdul-Salam and Sreelatha, 2004.

https://doi.org/10.1017/S0031182017001706 Published online by Cambridge University Press

day:light cycle to mimic summer conditions in the region. Snails were fed *ad libitum* a mixture of Tetramin™ (fish food), agar and calcium.

Circadian emergence patterns

Over the course of 93 trials, we quantified the number of cercariae released by each infected snail over 24 h in 2 h intervals. This assumes that the circadian rhythm of trematodes in emerging from snails is 24 h, consistent with the previous research; if the rhythm is longer than 24 h, however, this method would not capture it. We placed each snail individually into a 50 mL centrifuge tube; after 2 h, snails were carefully transferred to a new vial with clean water and the previous vial was labelled, capped and refrigerated (4 °C). This process was repeated every 2 h for 24 h. Trials followed a natural light:dark cycle (light from 6:00 to 20:00 h), during which snails were not fed. We counted released cercariae using an Olympus SZX10 stereodissection microscope (Olympus Corporation, Tokyo, Japan). Every effort was made to minimize disturbances to the snails or alteration of their environmental cues during the transfer process. When necessary to aid in cercariae counting, vials were left at 4 °C for 24 h to limit their swimming activity prior to counting.

Statistical analysis

To determine the rhythmicity of cercariae release, we implemented a single-component cosinor analysis (meaning only one emergence peak was analysed) within a GLMM framework. As outlined in Refinetti *et al.* (2007) and Cornelissen (2014), the general equation for a single cosinor is given by:

$$Y(t) = M + A\cos\left(\frac{2\pi t}{\tau} + \varphi\right) + e(t) \quad (1)$$

Key parameters include the period, τ (24 h for circadian rhythms, which is assumed to be constant), the MESOR, M (which in this case is rhythm-adjusted mean number of cercariae released across all observations); the amplitude, A (effectively the peak emergence value above and beyond the MESOR); and the acrophase, φ (the time at which peak emergence occurs). Prior knowledge of the period is required for cosinor analysis and was limited to 24 h in the current study. This equation can be rewritten into a LS regression model using the trigonometric angle sum identity, simplifying the equation to:

$$Y(t) = M + \beta x + \gamma z + e(t) \quad (2)$$

where

$$\beta = A\cos(\varphi),\ \gamma = -A\sin(\varphi),\ x = \cos(2\pi t/\tau),\ z = \sin(2\pi t/\tau)$$

Once β and γ are determined through optimization techniques (i.e. LS, maximum likelihood), the acrophase and amplitude can be extracted using the equations:

$$A = (\beta^2 + \gamma^2)^{1/2} \quad (3)$$

$$\varphi = \arctan\left(-\frac{\gamma}{\beta}\right) + K\pi \quad (4)$$

where K is an integer.

Cosinor analyses can include single or multiple component analyses (Cornelissen 2014). A multiple component analysis adds two or more cosine terms and takes the form

$$Y(t) = M + A_1\cos\left(\frac{2\pi t}{\tau_1} + \varphi_1\right) + A_2\cos\left(\frac{2\pi t}{\tau_2} + \varphi_2\right) + \cdots + e(t) \quad (5)$$

which is simplified into

$$Y(t) = M + \beta_1 x_1 + \gamma_1 z_1 + \beta_2 x_2 + \gamma_2 z_2 + \cdots + e(t) \quad (6)$$

Equations (3) and (4) from the above can be used to calculate acrophase and amplitude for each cosine line added. The benefit of adding multiple components is to account for the possibility of having multiple peaks in a circadian rhythm.

To incorporate the cosinor method into a GLMM framework, we used the R package 'glmmADMB' (Skaug *et al.* 2016) to build models of cercariae release per snail as a function of time. Our data were modelled as a negative binomial distribution with a log-link, although we compared among alternative distributions (e.g. Poisson and Gaussian) using the difference in the Akaike Information Criterion (delta AIC) (Burnham and Anderson, 2004; Symonds and Moussalli, 2011). We also evaluated whether to include a zero-inflation term to account for 'extra' zeroes in the dataset beyond what would be expected with a Poisson or negative binomial distribution (Zuur *et al.* 2009). As fixed effects, we included both cosine and sine terms as functions of time, which we first converted into radians with equation (3). Parasite species identity was incorporated as both a main (fixed) effect and as an interaction effect with the sine and cosine terms, thereby allowing us to evaluate how parasite species affected the intercept term (here, the MESOR) as well as the amplitude and acrophase. We included host subject (i.e. the specific snail) as a random intercept term to account for the 12 non-independent observations from each snail; alternatively, we also tried adding an offset term of total parasites produced by each snail ($\log_{10}$-transformed), which similarly functioned to capture much of the among-snail variation. This inclusion allowed for variation in the MESOR values among snails infected with the same parasite while still harnessing the full dataset to estimate amplitude and acrophase (as opposed to running separate analyses for each individual snail, which both limits the power of the dataset to detect overall trends and inflates the number of statistical tests to be performed). Mathematically, the offset term transformed the response from the raw number of cercariae released per 2 h to the proportion released relative to the total released over 24 h.

Before extracting estimates of the MESOR, amplitude and acrophase from each model, the use of a negative binomial distribution required exponentiation of the model as follows:

$$\ln(Y) = M + \beta x + \gamma z \quad (7)$$

which simplifies to

$$Y(t) = e^M e^{\beta x} e^{\gamma z} \text{ or } Y(t) = e^M e^{A\cos((2\pi t/\tau)+\varphi)} \quad (8)$$

Equations (3) and (4) from above can be used to calculate the parameters into response units with a few adjustments:

$$MESOR = e^M \quad (9)$$

$$A = e^M(e^{(\beta^2+\gamma^2)^{1/2}} - 1) \quad (10)$$

With the multiplicative nature of exponential functions, the $(e^{(\beta^2+\gamma^2)^{1/2}} - 1)$ term represents the proportion by which the

https://doi.org/10.1017/S0031182017001706 Published online by Cambridge University Press

MESOR increased during peak release, as opposed to the addition of a constant. Here, the amplitude is defined as the increase in cercariae during peak emergence relative to the MESOR [equation (10)]. Because we used a GLMM framework, parameters were estimated using maximum-likelihood estimation (MLE) rather than LS. Since both amplitude and acrophase were calculated using equations with estimates β and γ, which each had their own error [equations (3) and (4)], we performed a Monte Carlo simulation to account for error propagation and calculate the standard errors of the MESOR, amplitude and acrophase. Error propagation is a problem that occurs when multiplying estimated values – each of which have their own error – to obtain desired parameters. The Monte Carlo simulation uses the variance–covariance matrix for each parameter estimate to randomly generate a distribution of 10 000 values, from which the mean estimate and its standard deviation are calculated (Farrance and Frenkel, 2014; Harding *et al.* 2014). Finally, we tested residuals from each model for temporal autocorrelation using autocorrelation function plots and compared whether a single component or multiple component cosinor model offered a better fit to the data. Statistical significance was determined by a P value of <0.05 for the coefficient of the sin or cosine term. Significance for either term indicates a non-random rhythm in parasite emergence.

Model evaluation

To validate our methods, we conducted a simulation study to compare the effectiveness of using a generalized linear model (GLM) in comparison with classical LS regressions. We focused on the effect of using an alternative error distribution and did not include a simulation with random effects, hence the simple GLM model in the simulation study. In R 3.2.4 (R Core Team, 2016) we used the 'rnbinom' function to create random datasets using a range of 'true' values for each parameter (MESOR: range 2–200; amplitude: range 2–200; acrophase: range 0–24 h). For each parameter, we simulated 100 data points over the given range at 25 different combinations of the remaining two parameters for a total of 2500 simulated data points. From there, a GLM and a LS model were created using the 'glm.nb' function from package 'MASS' (Ripley *et al.* 2016) and the 'lm' function, respectively (R Core Team, 2016). The parameters were extracted from these models using the above equations, and Monte Carlo simulation techniques were used to obtain confidence intervals for the extracted parameters. For example, values for the acrophase were extracted from 100 simulated time points between 0 and 24 h for all pairwise combinations of MESOR with values 40, 80, 120, 160 and 200 and amplitude values of 40, 80, 120, 160 and 200 (see Fig. 1 and supplementary Fig. A.1–A.2). This process was completed for all three parameters: acrophase, MESOR and amplitude. These values were chosen to broadly encompass the range of values observed empirically among these parasites. To compare between the LS and GLM, we calculated the number of estimates that contained the true value within the confidence interval and the distance between the true and estimated value. We calculated the line of best-fit using the 'lm' function in R to determine the slope and intercept between the estimated and true values; the closer the slope is to 1 and the intercept to 0, the less biased the estimates are. To evaluate the effects of the negative binomial distribution parameter (θ) and sample size (N) on parameter estimates, we created two additional 100 point datasets using fixed values of 150 for MESOR, 100 for amplitude, and 12:00 for acrophase, while varying sample size between 1 and 100 and θ between 0.2 and 10.

Results

Circadian rhythms of aquatic trematodes

Our analysis strongly supported the existence of non-random, circadian rhythms in the emergence of trematode cercariae, but the patterns of release varied among species (Fig. 1). The best-fitting model incorporated a negative binomial distribution (dispersion parameter, $\theta = 104.71$), fixed effects for parasite species, cosine and sine terms, and their interaction with parasite species (all P values for interactions were <0.05 except those for *Zygocotyle* sp. and brevifurcate-apharyngeate cercaria, which were >0.05), as well as a random intercept term for snail identity (380 point increase in AIC when random intercept removed, 102.8 point increase in AIC with offset instead of random intercept). Because the interaction terms were often significant, we subsequently performed a GLMM cosinor analyses with each of the parasite species individually and compared their parameter estimates. Non-random circadian rhythms – as evidenced by the presence of significant cosine or sine terms – were evident for *R. ondatrae*, *Echinostoma* sp., *Cephalogonimus* sp., *A. marcianae*, the magnacauda morhotype and *Australapatemon* spp., while no evidence of a consistent emergence pattern for *Zygocotyle* sp. or the brevifurcate-apharyngeate morphotype (although these two species also had among the lowest sample sizes) (Table 1 and Fig. 1). Residuals from the global model showed no evidence of temporal autocorrelation. The use of an offset for total cercariae released (instead of a random intercept term for snail identity) was favoured for all parasite species except *R. ondatrae*, for which the random intercept model performed better (delta AIC = 104.3), and *A. marcianae*, for which models were equivalent. Six snails that had <10 total cercariae emerge over the 24 h period were excluded from analyses.

Among trematode taxa, we observed considerable variation in all three of the cosinor parameters (Table 1). MESOR estimates ranged from 33 (*A. marcianae*) to 295 (magnacauda morphotype), with most values about 55 (Table 1); amplitudes ranged from 13 (*R. ondatrae*) to 386 (*A. marcianae*). *Australapatemon* sp. released the most cercariae per 24 h period (average 3062 ± 617 S.E.) followed by *A. marcinae* (average 2428 ± 763 S.E.), both of which significantly exceeded total cercariae counts by intermediate emergence patterns (*Echinostoma* sp., *R. ondatrae* and *Cephalogonimus* sp.) and the lowest levels of release associated with *Zygocotyle* sp. and magnacauda (GLM with negative binomial, McFadden's pseudo $R^2 = 0.32$, followed by Tukey's comparisons). For the subset of measured snails ($n = 48$), snail size had an additional, positive effect on the number of cercariae released per 24 h (coefficient for snail size = 0.14 ± 0.057, $P = 0.014$; McFadden's pseudo $R^2 = 0.35$). Among the diurnal species, *Echinostoma* sp. and *Cephalogonimus* sp. emerged earlier in the light cycle (estimated acrophase of 9:44 and 9:21, respectively), *A. marcianae* emerged around mid-day (12:11), and *Australapatemon* spp. emerged in the afternoon (~14:58). For nocturnal species, *R. ondatrae* and magnacauda emerged early in the dark phase (21:46) and late (5:10), respectively (although replication for magnacauda was low).

Model performance

Relative to the LS approach, the GLM method with a negative binomial distribution yielded more accurate estimates of the 'true' parameters from simulated datasets. For both MESOR and amplitude, the GLM approach improved accuracy (difference between estimated and true values) (Figs. 2 and 3, and Supplementary Figs. A.1–A.3). The average difference between the true and the estimated values of MESOR was 10.3 for GLM and 19.5 for LS estimates (t-test, $P < 0.0001$). With GLM, the

https://doi.org/10.1017/S0031182017001706 Published online by Cambridge University Press

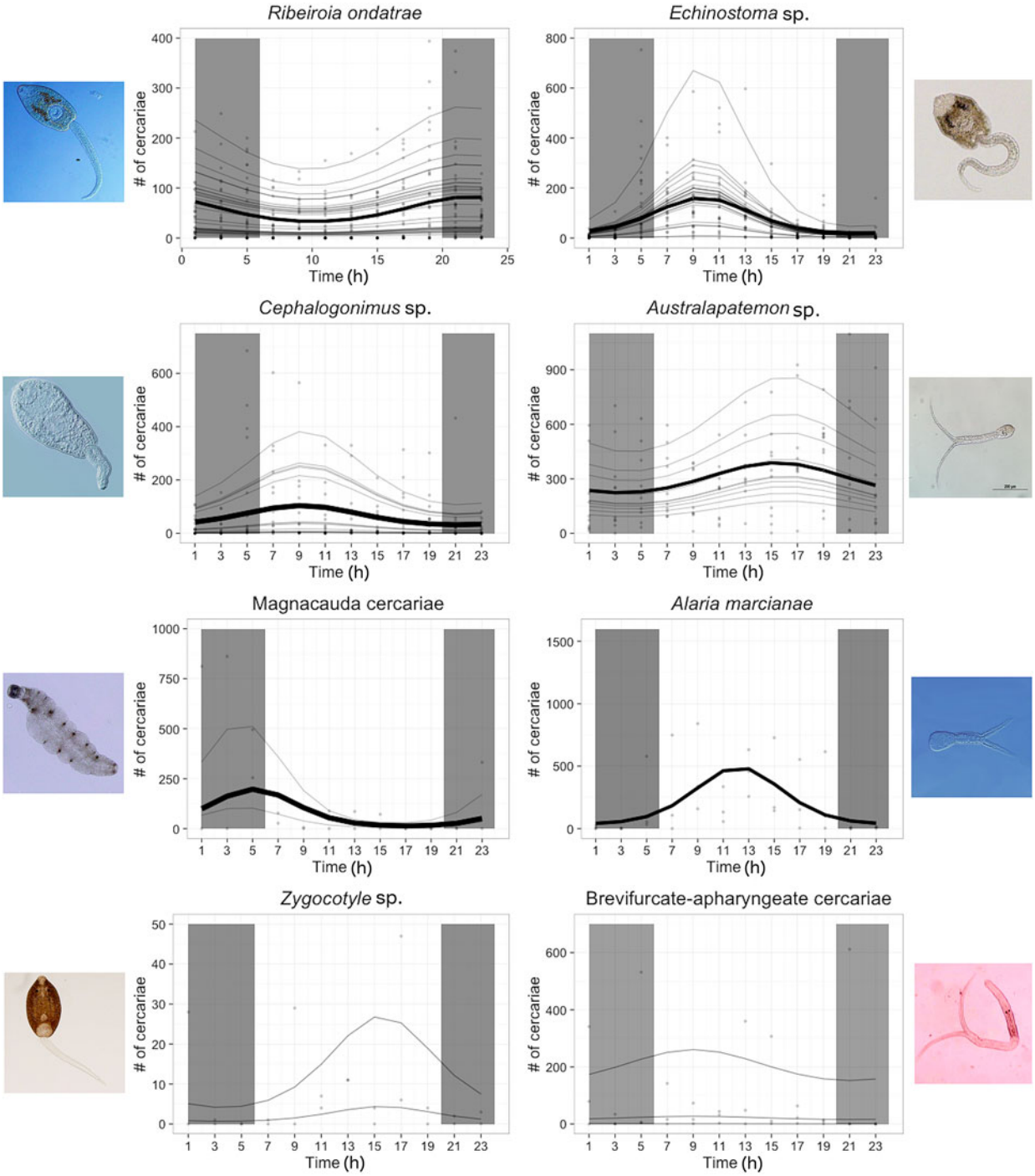


Fig. 1. Circadian rhythm of eight species of trematode parasites found in pond ecosystems. For each parasite, bold lines represent the overall, among-snail trend, faint grey lines are snail-specific estimates, and points are individual observations of cercariae counts within a 2 h period. The shaded vertical parts of the graph represent the hours the snails were in darkness (2000 to 0600). Images of each cercaria are presented on the outer margins of each plot.

95% confidence intervals around the estimated MESOR were also 43.2% more likely than the LS approach to include the true values. Likewise, for amplitude, the difference between the estimated and true values averaged 34.9 using GLM and 37.4 for LS (*t*-test, $P = 0.005$). The confidence interval from the GLM was also 2× more likely to include the true amplitude relative to the LS model.

Estimation of the acrophase showed the least variation between methods, with an average difference between the true and estimated values of 0.88 h using the GLM and 0.92 h with LS (Figs. 2 and 3, and Supplementary Figs. A.1–A.3). Additionally, estimates for the acrophase clustered closest to the 1:1 line when the MESOR:amplitude ratio was low for both methods (A.1). The LS approach exhibited more bias in parameter estimates based on the best-fit lines, the slope and intercept for which should approach 1 and 0, respectively. The LS method tended to underestimate the amplitude as the true amplitude increased, with a slope of 0.47 compared with 0.97 for GLM. Conversely, LS overestimated the MESOR with an intercept of 24.0 and a slope of 0.93, while the GLM line had an intercept of 0.0 and a slope of 0.99 (Fig. 2 and Supplementary Fig. A.2). The GLMM models did not show the same biases for the acrophase parameter estimates (Fig. 2 and Supplementary Figs. A.1–A.2). Overall acrophase estimates from both methods become more variable with larger MESORs (Supplementary Fig. A.2). Increases in the negative binomial dispersion parameter (θ) or sample size (N) tended to decrease variance around the central tendency, suggesting that

https://doi.org/10.1017/S0031182017001706 Published online by Cambridge University Press

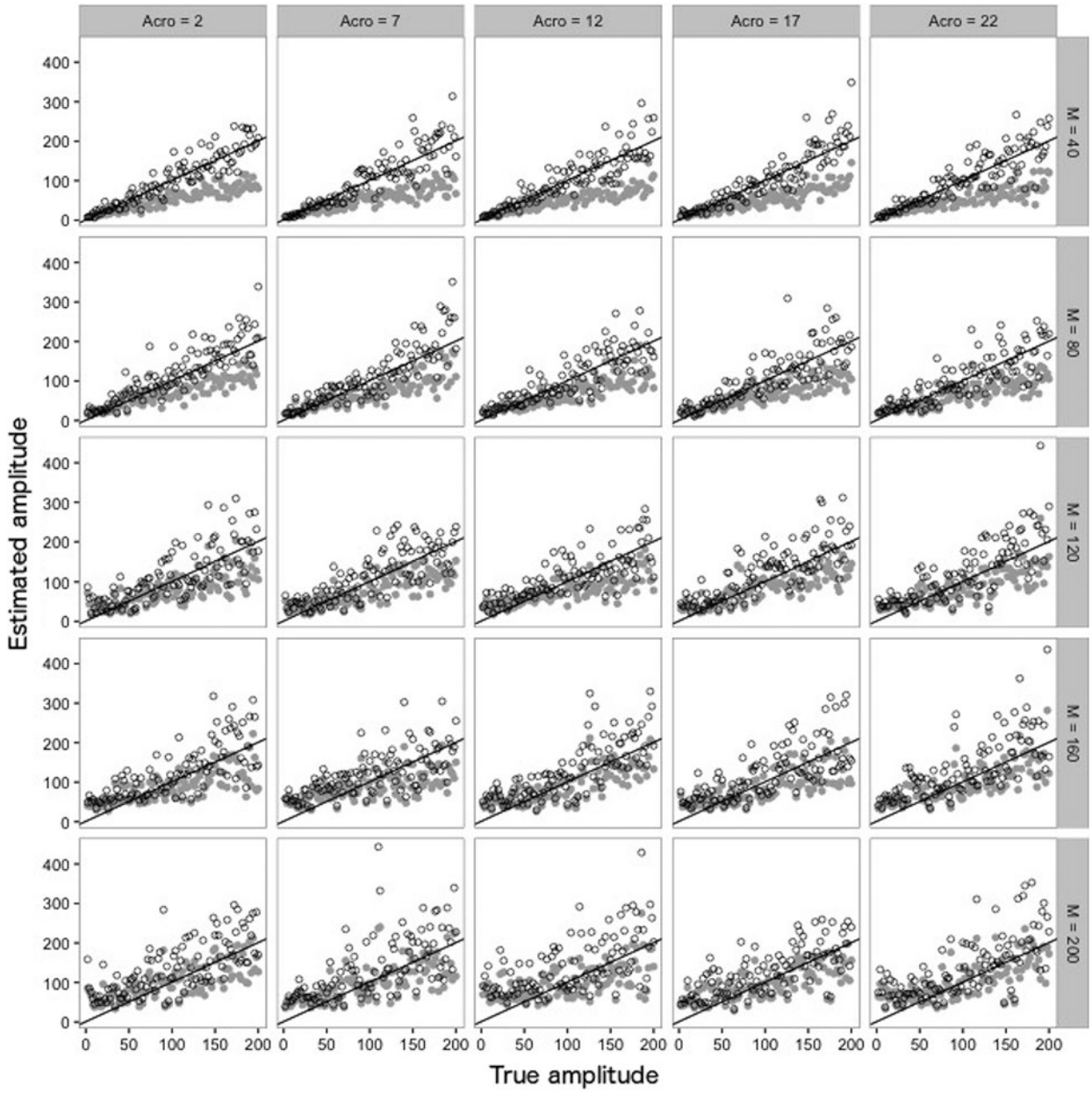


Fig. 2. A comparison of the 'true' and estimated values of the Amplitude parameter from the cosinor analysis between the traditional LS method and the GLM approach. The filled grey circles are the LS estimates, while the open circles are the GLM estimates. Each plot represents a different simulation with specified values of the MESOR (M) and the Acrophase (Acro). Each combination included 100 simulations drawn from a negative binomial distribution.

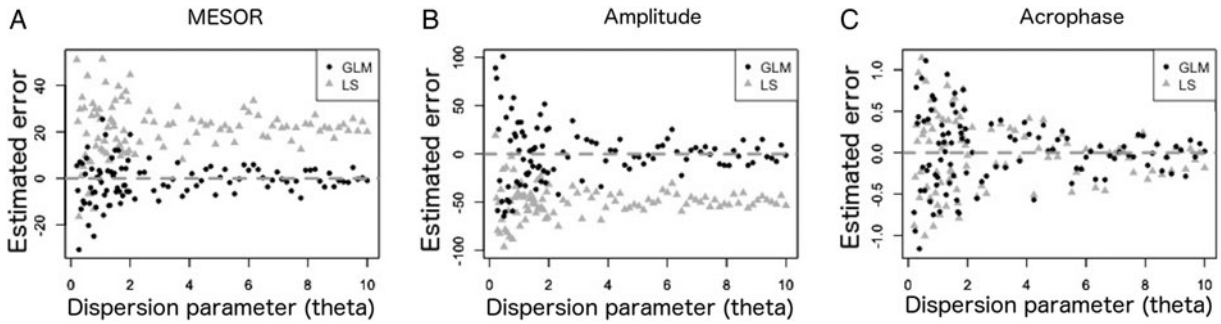


Fig. 3. Sensitivity of cosinor parameter estimates to variation in the negative binomial dispersion parameter (θ, which varied from 0 to 10). Each plot reflects the difference (estimated error) between the 'true' and estimated values from either LS or GLM for (A) MESOR, (B) Amplitude and (C) Acrophase. Dark circles and light triangles represent estimates from the GLM and LS approaches, respectively. The dashed horizontal line references the point at which the estimated and 'true' values are identical (difference of zero).

https://doi.org/10.1017/S0031182017001706 Published online by Cambridge University Press

models will perform better with less extreme overdispersion (larger θ values) and larger datasets (Fig. 3 and Supplementary Fig. A.3).

Discussion

In the summer of 1920, William Cort observed that the number of cercariae released from freshwater snails varied dramatically among trematode species, in response to temperature, and even among individual snails (Cort, 1922). Yet he further noted specific diurnal rhythms for each parasite, commenting that "…their constancy even over a considerable period is very striking." Subsequent studies have consistently reported that many trematode species exhibit predictable circadian rhythms, in some cases showing population-level variation in timing that corresponds with local patterns of downstream host activity (Ahmed *et al.* 2006; Lu *et al.* 2009; Théron, 2015). However, an ongoing challenge has been matching the forms of data typically generated with analytical techniques that can appropriately answer researchers' outlined questions. By using a cosinor analysis implemented within a generalized linear mixed modelling framework, we detected consistent circadian rhythms in six of eight tested species (*R. ondatrae*, *Echinostoma* sp., *Cephalogonimus* sp., *Australapatemon* spp., magnacauda morphotype and *Alaria marciniae*). Among the six species, chronotypes varied in all three parameters: acrophase (timing of peak parasite release, range: 5:10–21:46 h), amplitude (the increase from the MESOR of parasite release at the peak emergence, range: 13–386) and MESOR (range: 11–295). Both diurnal (*Echinostoma* sp., *Cephalogonimus* sp., *Australapatemon* spp. and *A. marcianae*) and nocturnal (*R. ondatrae* and magnacauda morphotype) emergence patterns were observed. No circadian rhythms were detected in *Zygocotyle* sp. or the brevifurcate-apharyngeate morphotype, although these two parasites had among the lowest sample sizes, so these results should be interpreted with caution.

Using a GLMM-based approach to cosinor analysis of circadian rhythms in the release of trematode infectious stages offers several benefits over previous methods. Data from parasitic circadian rhythm studies have key features that must be addressed: (1) the data are circular rather than linear, (2) the measurement of interest often involves count-type data with varying levels of overdispersion (Greene, 2008), (3) observations are autocorrelated in space (host) or time, and (4) study designs often include a combination of fixed and random effects. Classical methods such as circular statistics (Su *et al.* 2013; Théron, 2015) – while well suited for the cyclical data inherent to biorhythms – do not track periods of parasite inactivity or facilitate concurrent testing of multiple covariates. More specialized approaches, such as the cosinor method (Favre *et al.* 1997; Refinetti *et al.* 2007), do not easily allow for hierarchical nesting in data (such as when the release of many parasites is recorded from replicate hosts) and generally assume a Gaussian error distribution. Integrating cosinor analysis into a GLMM framework helped address each of these issues, allowing for flexible error distributions (Poisson or negative binomial with zero-inflation), incorporation of multiple covariates (parasite species, sine and cosine terms, as well as their interactions with parasite identity), and the use of random-effects or grouping variables (in this case, individual snail identity to account for correlated observations from the same host). According to a simulation-based sensitivity analysis, implementation of a GLM framework outperformed a standard LS approach, yielding parameter estimates that were more accurate and more precise relative to the known range of parameters used in the simulation.

Application of this novel modelling approach also creates new opportunities to investigate the biological drivers underlying parasite circadian rhythms. Correspondence between parasite emergence time and peak host activity is often hypothesized to increase the likelihood of transmission into downstream hosts (Combes, 2001; Lu *et al.* 2009; Théron, 2015). Our results provide mixed support for this hypothesis. While the nocturnal rhythm of *R. ondatrae* corresponds with the nocturnal activity patterns reported for some tadpoles (Oishi *et al.* 2004; Fraker, 2008), the trematodes *Echinostoma* sp., *Cephalogonimus* sp. and *A. marcianae* all exhibited diurnal rhythms for peak emergence, despite infecting the same secondary intermediate hosts (Table 1). *Echinostoma* sp. can also infect pulmonate snails as second intermediate hosts, which can exhibit both diurnal and nocturnal behaviour depending on their social interactions (Kavaliers, 1981; Keeler and Huffman, 2009). Cercariae of the *Zygocotyle* sp., which showed no consistent circadian rhythm in emergence, encyst directly on rocks or vegetation and may therefore be less constrained by host activity (Lang, 1968), although other threats such as UVB exposure in shallow waters could select toward nocturnal emergence.

While more research on the activity patterns of the downstream hosts or substrates for these trematodes is necessary to critically test the 'transmission hypothesis', alternative explanations for the timing of emergence warrant further exploration, particularly in light of the broad variation in circadian rhythms among trematodes that all use similar intermediate hosts. Factors such as UV radiation, salinity, tidal cycles and the activity of non-host taxa may also contribute to the observed cycles of parasite emergence (Valle *et al.* 1973; Williams *et al.* 1984; Fingerut *et al.* 2003; Sulieman *et al.* 2013; Wang *et al.* 2015). For example, growing evidence suggests that cercariae are vulnerable to predation by aquatic taxa such as small fishes, cnidarians, molluscs and aquatic insect larvae (Thieltges *et al.* 2008; Kaplan *et al.* 2009; Johnson *et al.* 2010; Orlofske *et al.* 2012). In an examination of the digestive tracts of field caught estuarine fish, Kaplan *et al.* (2009) detected cercariae in three of 61 fish, indicating that parasite consumption occurs in natural environments as well as under experimental conditions. Thus, emergence during periods of predator inactivity could increase the chance of successfully finding a host, whereas high overlap with predators could effectively suppress infection success. Orlofske *et al.* (2015) showed that while vertebrate and invertebrate predators were both effective at consuming cercariae in laboratory trials, the risk of consumption depended on traits of the parasite and the environmental context. Working with three of the same trematodes studied here, the authors found that species with larger cercariae – such as *R. ondatrae* – were especially vulnerable to small fishes and larval damselfly predators, whereas those with small (e.g. *C. americanus*) to intermediate (e.g. *E. trivolvis*) experienced correspondingly lower predation. However, predation risk further depended on light availability: trials conducted in the dark significantly reduced the number of cercariae consumed by visual predators, which is especially relevant for trematode species that naturally emerge at night (e.g. *R. ondatrae*) (Orlofske *et al.* 2015).

Thus, despite the accumulating evidence in support of predictable biological rhythms for trematode cercariae emergence, many outstanding questions persist about the cues that induce release of infectious stages, how these patterns vary among species or populations, and even the mechanisms through which parasites track time, either directly or indirectly (Théron, 2015). Such observations emphasize the need for more comparative research among trematode species or populations to evaluate variation in circadian rhythms and how they correlate with host usage, total release count and vulnerability to predation or physicochemical conditions, such as UVB exposure, salinity and water movements. Ultimately, development of flexible analytical techniques, such as those presented here, represent a useful prerequisite toward reliably characterizing parasite chronobiology and isolating the

https://doi.org/10.1017/S0031182017001706 Published online by Cambridge University Press

ecological and evolutionary factors underlying it. Moreover, while the current application was to parasitological data, the methods developed here can be extended to a wide range of systems with circadian data. Many datasets in chronobiology involve non-normally distributed data, including count data and binary outcomes, and it is often desirable to account for nested sources of non-independence, such as repeated observations on a subject, clustered locations in space, or animals in a particular group. Previous analytical methods often invoked multiple steps to first test for the presence of rhythmicity and then a second test to compare among subjects or between groups. The use of cosinor GLMM combines these into a single step and is thus suitable for testing circular data across biological as well as non-biological datasets.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0031182017001706

Acknowledgements. We thank the following individuals for assistance in laboratory data collection: K. Leslie, J. Jenkins, T. Riepe, A. Oweimrin, S. Anderson and D. Rose. We also thank T. McDevitt-Galles and members of the field crews for snail collection. K. Prior, S. Reece and an anonymous reviewer gave input helping in revising the manuscript.

Financial support. This study was supported through funds from the National Science Foundation (grant no. DEB1149308), National Institutes of Health (grant no. R01GM109499) and the David and Lucile Packard Foundation.

References

Abdul-Salam J and Sreelatha BS (2004) Description and surface topography of the cercaria of *Austrobilharzia* sp. (Digenea: Schistosomatidae). *Parasitology International* **53**, 11–21.

Ahmed AAM, Ibrahim NA and Idris MA (2006) Laboratory studies on the prevalence and cercarial rhythms of trematodes from *Bulinus truncates* and *Biomphalaria pfeifferi* snails from Khartoum State, Sudan. *Sultan Qaboos University Medical Journal* **6**, 65–69.

Aoki Y, Fujimaki Y and Isao T (2011) Basic studies on filarial and filiariasis. *Tropical Medicine and Health* **39**(Suppl. 2), 51–55.

Barnett AG, Baker PJ and Dobson AJ (2014) Season: analyzing seasonal data R functions. Version 0.3-5. Available at https://cran.r-project.org/web/packages/season/season.pdf.

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH and White JS (2009) Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution* **24**, 127–135.

Burnham KP and Anderson DR (2004) Multimodel inference – understanding AIC and BIC in model selection. *Sociological Methods & Research* **33**, 261–304.

Combes C (2001) *Parasitism*. Chicago, IL: University of Chicago Press.

Cornelissen G (2014) Cosinor-based rhythmometry. *Theoretical Biology & Medical Modelling* **11**, 1–24.

Cort WW (1922) A study of the escape of cercariae from their snail hosts. *Journal of Parasitology* **8**, 177–184.

Craig JM and Scott AL (2014) Helminths in the lungs. *Parasite Immunology* **36**, 463–474.

Davies D and de Núñez MO (2012) The life cycle of *Australapatemon magnacetabulum* (Digenea: Strigeidae) from northwestern Argentina. *Journal of Parasitology* **98**, 778–783.

Dreyer G, Pimentael A, Medeiros Z, Béliz F, Moura I, Coutinho A, Andrade LD, Rocha A, Silva LM and Piessens WF (1996) Studies on the periodicity and intravascular distribution of *Wuchereria bancrofti* microfilariae in paired samples of capillary and venous blood from Recife, Brazil. *Tropical Medicine & International Health* **1**, 264–272.

Faltynkova A, Karvonen A, Jyrkka M and Valtonen ET (2009) Being successful in the world of narrow opportunities: transmission patterns of the trematode *Ichthyocotylurus pileatus*. *Parasitology* **136**, 1375–1382.

Farrance I and Frenkel R (2014) Uncertainty in measurement: a review of Monte Carlo simulation using Microsoft Excel for the calculation of uncertainties through functional relationships, including uncertainties in empirically derived constants. *Clinical Biochemist Reviews* **35**, 37–61.

Favre TC, Bogéa T, Rotenberg L, Silva HS and Pieri OS (1997) Circadian rhythms in the cercarial emergence of *Schistosoma mansoni* by *Biomphalaria tenagophila* at outdoors: a comparative study with *Biomphalaria glabrata*. *Biological Rhythm Research* **28**, 348–357.

Fenwick A (2012) The global burden of neglected tropical diseases. *Public Health* **126**, 233–236.

Fingerut JT, Zimmer CA and Zimmer RK (2003) Patterns and processes of larval emergence in an estuarine parasite system. *The Biological Bulletin* **205**, 110–120.

Fraker ME (2008) The influence of the circadian rhythm of green frog (*Rana clamitans*) tadpoles on their antipredator behavior and the strength of the nonlethal effects of predators. *The American Naturalist* **171**, 545–552.

Fried B and Ponder EL (2003) Effects of temperature on survival, infectivity and in vitro encystment of the cercariae of *Echinostoma caproni*. *Journal of Helminthology* **77**, 235–238.

Greene W (2008) Functional forms for the negative binomial model for count data. *Economics Letters* **99**, 585–590.

Harding B, Tremblay C and Cousineau D (2014) Standard errors: a review and evaluation of standard error estimators using Monte Carlo simulations. *The Quantitative Methods for Psychology* **10**, 107–123.

Hawking F (1967) The 24-hour periodicity of microfilariae: biological mechanisms responsible for its production and control. *Proceedings of the Royal Society of London B* **169**, 59–76.

Johnson AD (1968) Life history of *Alaria marcianae* (La Rue, 1917) Walton, 1949 (Trematoda: Diplostomatidae). *Journal of Parasitology* **54**, 324–332.

Johnson PTJ, Dobson A, Lafferty KD, Marcogliese DJ, Memmott J, Orlofske SA, Poulin R and Thieltges DW (2010) When parasites become prey: ecological and epidemiological significance of eating parasites. *Trends in Ecology & Evolution* **25**, 362–371.

Johnson PTJ, Preston DL, Hoverman JT and Richgels KLD (2013) Biodiversity reduces disease through predictable changes in host community competence. *Nature* **494**, 230–234.

Kanev I, Fried B, Dimitrov V and Radev V (1995) Redescription of *Echinostoma trivolvis* (Cort, 1914) (Trematoda: Echinostomatidae) with a discussion on its identity. *Systematic Parasitology* **32**, 61–70.

Kaplan AT, Rebhal S, Lafferty KD and Kuris AM (2009) Small estuarine fishes feed on large trematode cercariae: lab and field investigations. *Journal of Parasitology* **95**, 477–480.

Karvonen A, Paukku S, Valtonen ET and Hudson PJ (2003) Transmission, infectivity and survival of *Diplostomum spathaceum* cercariae. *Journal of Parasitology* **127**, 217–224.

Kavaliers M (1981) Circadian and ultradien activity rhythms of a freshwater gastropod, *Helisoma trivolvis*: the effects of social factors and eye removal. *Behavioral and Neural Biology* **32**, 350–363.

Keeler SP and Huffman JE (2009) Echinostomes in the second intermediate host. In Fried B and Toledo R (eds). *The Biology of Echinostomes*, New York, NY: Springer Science, pp. 61–87.

Khan AM, Dutta P, Das S, Pathak AK, Sarmah P, Hussain ME and Mahanta J (2015) Microfilarial periodicity of *Wuchereria bancrofti* in Assam, Northeast India. *Journal of Vector Borne Diseases* **52**, 208–212.

Komatsu S, Daisuke K, Paller BGV and Uga S (2014) Dynamics of *Centrocestus armatus* transmission in endemic river in Hyogo Prefecture, Japan. *Tropical Medicine and Health* **42**, 35–42.

Kuris AM and Lafferty KD (1994) Community structure: larval trematodes in snail hosts. *Annual Review of Ecology and Systematics* **25**, 189–217.

Lang BZ (1968) The life cycle of *Cephalogonimus americanus* Stafford 1902 (trematoda: Cephalogonimidae). *Journal of Parasitology* **54**, 945–949.

Lee JH, Han G and Giuliano AR (2012) Analysis of overdispersed count data: application to the human papillomavirus infection in men (HIM) study. *Epidemiology & Infection* **140**, 1087–1094.

Lee Gierke C, Cornelissen G and Lindgren J (2013) CAT: Chronomics Analysis Toolkit. University of Minnesota. Available at http://564394709114639785.weebly.com/installing-cat.html.

Lu D, Wang T, Rudge JW, Donnelly CA, Fang G and Webster JP (2009) Evolution in a multi-host parasite: chronobiological circadian rhythm and population genetics of *Schistosoma japonicum* cercariae indicates contrasting definitive host reservoirs by habitat. *International Journal of Parasitology* **39**, 1581–1588.

Mintsa-Nguéma R, Mone H, Ibikounle M, Mengue-Ngou-Milama K, Kombila M and Mouahid G (2014) Cercarial emergence pattern of *Schistosoma haematobium* from Libreville, Gabon. *Parasitology* **21**, 1–5.

Mouahid G, Idris MA, Verneau O, Théron A and Shaban MMA (2012) A new chronotype of *Schistosoma mansoni*: adaptive significance. *Tropical Medicine & International Health* **17**, 727–732.

https://doi.org/10.1017/S0031182017001706 Published online by Cambridge University Press

Mouritsen KN (2002) The *Hydrobia ulvae-Maritrema subdolum* association: influence of temperature, salinity, light, water-pressure and secondary host exudates on cercarial emergence and longevity. *Journal of Helminthology* **76**, 341.

N'Goran E, Brémond P, Sellin E, Sellin B and Théron A (1997) Intraspecific diversity of *Schistosoma haematobium* in West Africa: chronobiology of cercarial emergence. *Acta Tropica* **66**, 35–44.

Núñez MO, Davies D and Spatz L. (2011) The life cycle of *Zygocotyle lunata* (Trematoda, Paramphistomoidea) in the subtropical region of South America. *Revista Mexicana de Biodiversidad* **82**, 581–588.

Oishi T, Nagai K, Harada Y, Naruse M, Ohtani M, Kawano E and Tamotsu S (2004) Circadian rhythms in amphibians and reptiles: ecological implications. *Biological Rhythm Research* **35**, 105–120.

Orlofske SA, Jadin RC, Preston DL and Johnson PTJ (2012) Parasite transmission in complex communities: predators and alternative hosts alter pathogenic infections in amphibians. *Ecology* **93**, 1247–1253.

Orlofske SA, Jadin RC and Johnson PTJ (2015) It's a predator-eat-parasite world: how characteristics of predator, parasite and environment affect consumption. *Oecologia* **178**, 537–547.

Parker GA, Chubb JC, Ball MA and Roberts GN (2003) Evolution of complex life cycles in helminth parasites. *Nature* **425**, 480–484.

Paull SH and Johnson PTJ (2014) Experimental warming drives a seasonal shift in the timing of host-parasite dynamics with consequences for disease risk. *Ecology Letters* **17**, 445–453.

R Core Team (2016) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at https://www.R-project.org/.

Refinetti R, Cornelissen G and Halberg F (2007) Procedures for numerical analysis of circadian rhythms. *Biological Rhythm Research* **38**, 275–325.

Revelle W (2016) *Psych: Procedures for Personality and Psychological Research*. Evanston, Illinois, USA: Northwestern University. Version 1.5.8. Available at http://CRAN.R-project.org/package=psych.

Ripley B, Venables B, Bates DM, Hornik K, Gebhardt A and Firth D (2016) Package 'MASS'. Version 7.3-45. CRAN repository. Available at https://cran.r-project.org/web/packages/MASS/MASS.pdf.

Sachs M (2014) Cosinor: Tools for estimating and predicting the cosinor model. Version 1.1. Available at https://CRAN.R-project.org/package=cosinor.

Schell SC (1985) *Handbook of Trematodes of North America North of Mexico*. Moscow, ID: Idaho Research Foundation.

Shaw DJ, Grenfell BT and Dobson AP (1998) Patterns of macroparasite aggregation in wildlife host populations. *Parasitology* **117**, 597–610.

Silva HS, Rotenberg L, Bogéa T, Favre TC and Pieri O (1995) Longitudinal study of circadian rhythms in the cercarial emergence of *Schistosoma mansoni* from *Biomphalaria glabrata*. *Memórias do Instituto Oswaldo Cruz* **90**, 459–461.

Skaug H, Fournier D, Bolker B, Magnusson A and Nielsen A (2016) Generalized linear mixed models using 'AD Model Builder'. Version 0.8.3.3. Available at http://glmmadmb.r-forge.r-project.org.

Soldánová M, Selbach C and Sures B (2016) The early worm catches the bird? Productivity and patterns of *Trichobilharzia szidati* cercarial emission from *Lymnaea stagnalis*. *PLoS One* **11**, e0149678.

Steinauer ML, Mwangi IN, Maina GM, Kinuthia JM, Mutuku MW, Agola EL, Mungai B, Mkoji GM and Loker ES (2008) Interactions between natural populations of human and rodent schistosomes in the Lake Victoria region of Kenya: a molecular epidemiological approach. *PLoS Neglected Tropical Diseases* **2**, e222.

Su J, Zhou F and Lu D (2013) A circular analysis of chronobiology of *Schistosoma japonicum* cercarial emergence from hilly areas of Anhui, China. *Experimental Parasitology* **135**, 421–425.

Sulieman Y, Pengskul T and Guo Y (2013) Development and effects of *Schistosoma japonicum* (Trematoda) on its intermediate host, *Oncomelania hupensis* (Gastropoda). *Iranian Journal Parasitology* **8**, 212–218.

Symonds MRE and Moussalli A (2011) A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology* **65**, 13–21.

Théron A (2015) Chronobiology of trematode cercarial emergence: from data recovery to epidemiological, ecological and evolutionary implications. *Advances in Parasitology* **88**, 123–164.

Thieltges DW, Jensen KT and Poulin R (2008) The role of biotic factors in the transmission of free-living endohelminth stages. *Journal of Parasitology* **135**, 407–426.

Valle C, Pellegrino J and Alvarenga N (1973) Rhythmic emergence of *Schistosoma mansoni* cercariae from *Biomphalaria glabrata*: influence of temperature. *Journal of the São Paulo Institute of Tropical Medicine* **15**, 195–201.

Wang C, Lu D, Guo C, Li Y, Gao Y, Bian C and Su J (2014) Compatibility of *Schistosoma japonicum* from the hilly region and *Oncomelania hupensis hupensis* from the marshland region within Anhui, China. *Parasitology Research* **113**, 4477–4484.

Wang S, Zhu Y, Ge Q, Yang M, Huang J, Huang W, Zhuge H and Lu D (2015) Effect of photoperiod change on chronobiology of cercarial emergence of *Schistosoma japonicum* derived from hilly and marshy regions of China. *Experimental Parasitology* **159**, 227–232.

Williams CL, Wessels WS and Gilbertson DE (1984) Comparison of the rhythmic emergence of *Schistosoma mansoni* cercariae from *Biomphalaria glabrata* in different lighting regimens. *Journal Parasitology* **70**, 450–452.

Zuur AF, Leno EN, Walker NJ, Saveliev AA and Smith GM (2009) *Mixed Effects Models and Extensions in Ecology with R*. New York, NY: Springer.

https://doi.org/10.1017/S0031182017001706 Published online by Cambridge University Press