Referee reports:

Referee: 1

Comments to Author(s)

Walker et al. present a modelling framework to analyze how body size may affect the evolution of generalism in host-macroparasite relationship, i.e. whether host generalism is more like to occur for large body size hosts vs small body size.  The authors analyzed also an extensive dataset of 23,331 unique host-parasite associations between 8,846 parasite species and 4,237 fish host from the Fish Parasite Ecology Database and contrasted theoretical predictions with empirical patterns under a number of alternative assumptions about number of specialist parasite species, factors regulating host abundances, etc .

I commend the authors for having put together such an impressive dataset and I believe that there is a value in using simple models of disease dynamics to understand general patterns. Anyway, I believe that the statements made by the authors about possible patterns between host body size and a number of metrics to measure host generalism are poorly supported by their theoretical analysis and even by the data. I’ll illustrate here below why I think this is the case.

This work is clearly focused on macro-parasite dynamics, whether endoparasites or ectoparasites. Macro-parasites in general share two distinct features (along with many others).

First, their distribution in the host population is usually highly aggregated, with the majority of the hosts harboring few or no parasites and few harboring a lot. Second, I believe that aside from few specific cases, coinfection by more than one parasitic species is the norm, rather than the exception. For these reasons, seminal work by Anderson and May (1991) have suggested that they should be modelled in a different way than micro-parasites.  I thus wonder why the author used a SI-P modelling framework – which is typically used in the case of pathogens (see for instance Tien and Earn 2010 Multiple Transmission Pathways and Disease Dynamics in a Waterborne Pathogen Model, Bulletin of Mathematical Biology 72: 1506–1533) - instead of the classic Anderson and May (1978) macro-parasite approach. The implications of using one or the others are not marginal: maybe the threshold conditions at invasion of a healthy population (as expressed by the \*basic\* reproduction number) are slightly different between basic micro- (SI) and macro-parasitic models, but the threshold condition of an already parasitized host population depends, among other things, by the degree of aggregation of the first parasite in the population (the clumping parameter of the Anderson and May models) and by how much the first parasite depresses the host population. On the contrary, the authors use a classic compartmental model for the host that can be either susceptible or infected when it acquires parasites from contact with free living infectious stages.  The authors do not provide any explanation of why they chose this modelling framework.

*Before addressing this particular comment, we would like to thank you for your obvious careful consideration of this manuscript, and for providing us with such detailed comments and criticisms. You have identified a number of important issues regarding the model development and data analysis, and we feel that the manuscript is much improved by our addressing these issues.*

*To this specific comment: we were aware of Anderson and May’s classic macroparasite model. It was not used in this study for a number of reasons. First, the Anderson and May model assumes that all hosts are susceptible to infection by all parasites. In the context of this study, where we are considering whether the invasion of a novel mutant will succeed or fail, that assumption implies that the only restriction placed on invasion success by the resident parasite is through the effect of the resident parasite on total host population size. In reality, of course, there are mechanisms that reduce the probability that another parasite will be able to establish in an already infected host, even for macroparasitic infections. For example, concomitant immunity triggered by adult worms can prevent the establishment of larval worms (e.g. Brown and Grenfell 2001). Using the SI-P model allows us to control the degree to which coinfection is possible to investigate how that influences the invasion success of the mutant generalist (see more discussion of coinfection below).*

*Second, as the reviewer points out, the Anderson and May does allow for aggregation, but in a way that is somewhat unsatisfying, because it is phenomenological and because aggregation only influences the mortality rate of the parasite. In reality, parasite distributions are aggregated because of some interaction between between-host epidemiological processes and within-host immunological processes. Introducing multiple classes of infected host (e.g., super-shedders) might be an efficient way to handle this within our chosen epidemiological framework, but is beyond the scope of this work. Moreover, hosts that have very high parasite burdens should also shed more parasites into the environment, which the Anderson and May model does not consider. Increased parasite shedding would potentially offset the increased parasite mortality due to the death of hosts with high burden. Without including shedding benefits, the effect of aggregation on the invasion success of a mutant is obvious from the outset – aggregation will reduce the abundance of parasites and increase the abundance of hosts, thereby making invasion easier.*

*Third, and related to the above point, the Anderson and May model deals with parasite shedding in an elegant way, but not one that is easily adapted to this study. They assume a per-parasite shedding rate of , so that the total rate that parasites are produced is*

*where is the number of parasites, is the probability a host is harboring parasites, and is the number of hosts. Thus is the average number of parasites per host, or , so the total shedding rate is just . With this formulation, there is no way of accounting for the fact that a larger host, infected with the same parasite as a smaller host, should shed more (other than allowing the per-capita shedding rates to differ, but that is how we are accounting for the cost of generalism). Moreover, the expected differences in shedding across hosts of different body size is one of the more important mechanisms by which host body size influences the evolution of host range in this model (and in the verbal models developed by other authors). Allowing burden to differ across hosts is a common criticism of the use of SI-type models for macroparasites, but one our models avoid because of the inclusion of allometric scaling.*

*Hopefully this is a satisfying explanation for why we did not use the Anderson and May (1978) macroparasite model. We have included an abbreviated form of these comments into both the model development and Discussion, so that modeling-savvy readers such as yourself will better understand our justification for using this model.*

Also, Walker et al.’s basic model (the one presented in the main body of the ms) does not account for co-infection, which is the common case for macroparasites (actually, they talked about it, but as a case analyzed in the appendix and, if I am correct, under the assumption that host populations are constant and not affected by parasitic load). I am not sure where co-infection might have implications on generalism: my verbal argument is that a host infected by one parasite species is not removed from the compartment of susceptible hosts and it can still be infected by a parasite of a different species (in addition to more parasite of the same species), even though the presence of one parasite does have an effect on the abundance of the host at the diseased equilibrium. Whether and how this has implications on the effect of host body size on evolution of generalism, I do not know, but I believe that co-infection shall be the central case (not a case study in the appendix), and I completely support authors’ argument to move away from verbal models, thus I believe that this is a relevant question that needs to be investigated with the proper modelling tools.

*This is a really excellent point. We have extensively revised the model analysis to focus on the effect of coinfection on the evolution of generalism. In fact, we have completely rewritten the model derivation to focus on deriving an expression for , the invasion fitness of a generalist parasite, when it is invading a community with an arbitrary number of hosts and where coinfection is permitted.*

Also, the modelling approach ignores a significant body of literature (especially by Poulin and by Morand, among the others) providing empirical evidence that basic life history traits of macroparasites scale allometrically with parasite body size.  For instance, Skorping et al. (1991) and Morand  et al. (1996, Functional Ecology) presented data supporting the hypothesis that large body size parasites are more fecund than small body size parasites. By definition, per-capita fecundity has a direct influence on individual fitness. Therefore, parasite body size might have profound implications on parasite ability to invade a host population, coexist with or displace competing parasite species.  In addition, Morand et al. (1996) on Oecologia showed that parasites body size also scales allometrically with host body size. Therefore, there are so many variables explicitly associated with host or parasite body size, that I wonder whether Walker et al’s model. is the right modelling framework to address their overarching question of parasite specialization vs generalism.

*Thank you for highlighting these references. It was on our minds that parasite body size might also play an important role in shaping the evolution of host range, a point we discussed in the Discussion section However, because our dataset did not have information on parasite body size, we chose to not include that in the model development. We still feel that there is considerable value in the modeling exercise, even lacking an incorporation of parasite body size, because it provides a theoretical framework for future studies to incorporate other mechanisms. However, we have added discussion of this in the Discussion section.*

I also have concern about the data analysis. When I look at Fig. 1 and 2, I do not see any clear relationship between the host body length and the dependent variables: some relationship may be statistical significant (such as between max host lengths and degree or SPD, or SES.PD  in Fig.1) because of the large sample sizes (and the authors shall add the degree of freedom to all the tables with the results of the statistical analyses) but the overall picture remain fuzzy and, if anything, it shows that there generalism might be more prevalent at intermediate (and in some cases) small body sizes. The only cases in which some more significant pattern may emerge that it is not an artifact created by large numbers is the case of CV and SPD for both directly and trophically transmitted parasites.

*We agree, and have reported, that the relationships observed in the data vary based on the particular length metric and generalism metric that are compared. However, we disagree with the statement that a significant pattern is an artifact of large sample sizes, though of course the large sample sizes may allow us to detect relationships with small effect sizes. The effect sizes are reported in the tables, along with the confidence intervals and test statistic, and we have added the residual degrees of freedom from each test to the tables as suggested.*

Aside from this case, I wonder whether there are no better statistics (maybe quantile regression?) or statistical approach accounting for multiple explanatory covariates and possible interactions among them to disentangle the cloudy set of points depicted in Fig.s 1 & 2.

*We have added the results of a multivariate regression with interactions in the supplementary material (Appendix B). This extends the results of interactions presented beyond the original example in the temperature/geographic range section of the results, which explores the interactions between the independent variables in the case of SES-PD. The points in Figures 1 and 2 look cloudy because of the large sample sizes, so that the density of points is not readily visible, therefore the univariate regression lines are plotted on top to show the statistical relationship.*

In the case of the relationship between CV and SPD, I did not see any convincing explanation of why this should be the case, i.e. why the difference between the body sizes of two host needs to be large for generalism to occur.

*The relationship between the variation in host body size and generalism is presented in the model predictions section, where we stated:*

First, marine organisms, even just fish, might have a very different body shape (a barracuda or shark vs a sun fish). The underlying metabolic theory of ecology is based on body mass, not length, so why the authors don’t used body mass instead of length as independent variable, at least for endoparasites?

*We agree that body length and body mass are not perfectly correlated. However, there was not consistent information on host body mass available in the fish species database, so we present this as a proxy. We added a statement to the discussion section to remind readers of this potential limitation.*

Second, I would refer to Ro (eq. 1) for what it is, i.e. the effective reproduction number (in contrast to the \*basic\* reproduction number). Note that this is different from the basic reproduction number of any single parasite invading a population of healthy hosts at their carrying capacity.  I think that this is relevant because I believe that the \*basic\* reproduction number of the specialist parasite (which should be equal to lambda1/mu1 \* beta\*K1/(beta\*K1 + gamma)) is also an increasing function of host body size. This is because, in the case of an endoparasite, the effect of body size on lambda is going to cancel out with that on K1 (as lambda increases as W^3/4 while K1 decreases as W^-3/4), whereas both parameters at the denominator decrease with host body size - even more true for ectoparasites (please, correct me if I am wrong). Therefore, for \*any\* parasite it is easier to invade a population of large body size hosts than a population of small body size hosts. If, according to eq. 1, a generalist parasite might be less likely to establish in a population of small body hosts, the same is true also for a specialist parasite. If this is the case (i.e. if I am correct), it is crucial to check whether the effective reproduction number for the generalist parasites increases with body size faster than the basic reproduction number for the specialist parasites If they increase at the same pace (unlikely, but to be checked), it could be theoretically possible that the fraction of generalist vs specialist parasites remain roughly constant, whereas the sheer number of parasitic species (whether specialist or generalist) increases with host body size. My later statement (which should be a logical consequence of Ro increasing with host body size for both specialist and generalist parasites) does not seems to be supported, at least at first sight, by the data reported in Fig.s 1 & 2 – the authors can easily check whether this is the case (or whether the hypothesis I just presented is flawed).

*In the revision we refer to the fitness of the mutant as Rm, rather than R0, to avoid this confusion and clarify what, exactly it represents. We will use this terminology throughout our comment here.*

*The R0 you define above is the R0 for a specialist parasite infecting a population of fully susceptible hosts (). As you correctly point out, this is also an increasing function of body size. You could compare this R0 expression against the R0 of a generalist parasite invading a fully susceptible population, in which case the fitness of the generalist is . You could compare these two R0 expressions to determine when the generalist would have a higher fitness than a specialist when invading a fully susceptible population. However, this doesn’t seem to be what you are suggesting, and we are unsure there is any biologically meaningful conclusion that could be drawn from a comparison of the R0 of a specialist parasite invading a fully susceptible population and the Rm of a generalist parasite invading a population where specialist parasites are present. A more meaningful comparison might be between the Rm we calculate and the Rm of a specialist parasite attempting to invade a system where a generalist parasite is already present. We explored this comparison for one case, but the results were analytically intractable (comparing the two Rm values would require extensive numerical exploration). Given the complexity of the analysis we have already carried out, we felt that this was too much for this study.*

Third, what about the effect of host body size on other model parameters? For instance, the authors assume that the transmission rate beta is constant. Now, beta - a composite parameters accounting for both host mobility and probability of getting infected when encountering a free living infectious stage P – might actually be also a function of host body size, as discussed in a number of papers (see for instance Kelt and Van Vuren 1999 Ecology, Kelt and Van Vuren 2001 AmNat, Mysterud et al. 2001 Oecologia, Ottaviani et al. 2004 Journal of Zoology, Bordes et al. 2009 AmNat). The above cited papers are mostly about terrestrial animals, but a relationship between home range and host body size for marine organisms is presented in McCauley et al. (Science 2015, Fig.4) and Tamburello et al. (AmNat 2015). Therefore, I am not sure whether the assumption of a constant beta in supported by the data.

*This is another excellent point for future work. Here we have assumed that the parasite controls the contact process (e.g., the parasite actively seeks out hosts), so in that context the transmission rate might more reasonably be considered independent of host body size. If the contact process depends on host behavior, then that would be entirely appropriate. For example, for many parasites, the host acquires infections during feeding. In such a case there is a direct connection between transmission rates and foraging rates, which are often size-dependent. We have added discussion of this point in the Discussion section.*

Page 7 Lines 11-12: “Taking account of other metabolic constraints on parasite growth [22]…” This sentence is quite vague. In addition, if I remember well, Hechinger 2013 (and not 2012, as cited in the appendix) presents a couple of alternative hypotheses and it might be useful to refer precisely under what specific assumption (limited energy vs limited space) the 3/4 and 5/12 exponents originate from. This is a very crucial point as it links the parasite-centered approach of Anderson & May classic macro-parasite models with the compartmental SI approach used by Walker et al.  IN fact, according to Morand et al. (1996, Oecologia), larger hosts may harbor in general larger and thus more fecund parasites. This is not explained in the text, while I believe it is important.

*We have expanded here to give a more complete explanation of why abundance should scale with ¾ and 5/12 exponents. Thank you for catching the error in citation date.*

Here below a list of additional points:

Please, number all equations (such as those at page 5 and 8)

*We have numbered equations throughout the manuscript.*

Page 7 Line 24: “Thus, generalism is more likely to evolve for endoparasites than ectoparasites.” is this sentence ultimately supported by the data in the present or in other studies?

*As described in the final subsection of the results, we were unable to address this hypothesis because in the database endo vs ecto parasites are almost perfectly confounded with transmission route.*

Page 7, line 27-28. The authors do not explain why the secondary host shall have a body size smaller than the primary one. What about if f>1?

*Mathematically, there was no reason not to consider f > 1, so we have removed this statement.*

Page 7 Lines 36-37:  explicitly state that the RHS is always positive and thus dRo/dt>0 for any body size (it is stated in the appendix but not in the main body of the ms).

Page 7 line 41-42: add \*we expect\* to <…there will be a positive…>. I understand that you want to outline the  hypothesis, but it is a bit strange stylistically to use bold font in a ms.

*Added “…the model predicts that…” for clarity in this sentence, and removed bold font emphasis from the manuscript.*

Page 7, Line 54-55: remove “strong”, unless you can explain how you disentangle between a weak and a strong relationship

*Removed the word strong.*

Page 8, line 21-22. It is explained only in the table 1 (page 21) that “infection site” is used to tease apart endo vs ecto-parasite. When I read it the first time, I thought at intestine, muscle, brain, eyes, gills, etc.   so, an short explanation in the text might be useful.

*Added a parenthetical definition of infection site to clarify how we use the term in this paper.*

Page 8, line 49-50 <… the expression simplifies…> sounds a bit as geek jargon.

Page 8, lines 57-58: <What these analyses reveal is that, for direct life cycle parasites, the effect of host body size is almost always to increase the value of Ro, thereby making it easier for generalists to invade.>. As I argued above, I believe that Ro increases with body size also in the case of a specialist parasites. If I am correct, it is probably how fast one increases with respect to the other that might cast light on whether generalism is (relatively speaking) more frequent than specialization for host with large body size.

*As noted above, while you are mathematically correct, we are not convinced that the comparison is meaningful. It is, however, clear from our analyses that increasing host body size will make it easier for the generalist to invade.*

Page 10, lines 14-22: it is a good idea in general to use a latitudinal criterion to check whether there is a temperature signature. Yet, in the case of coast waters, upwelling zone (characterized by low temperatures) might occur at quite low latitudes, the same where we can have warm subtropical waters. Aside from latitudinal segregation, did the authors check whether they have fish from upwelling systems in their database?

*Unfortunately we were not able to identify the particular locations where the fish were found, so this is a limitation of our data analysis. We have added a statement to the discussion to identify this limitation.*

Page 8 24-28: in addition to GLM, did the authors run the analysis for multiple covariates and their possible interactions?

*We have added the results of multivariate analyses in Supplementary Tables 1-4 in Appendix C.*

There are some differences in the statistical outcome when using maximum vs mean body size of the host, I am not sure that the theoretical models can tease them apart (i.e. project different patterns depending whether we use one or the other) and I cannot make my mind why this should be different, what is the ecological explanation of why different pattern emerges when using maximum instead of mean body size – I do not think that the authors present any convincing explanation about it.

*Added a sentence to the discussion to point out that the difference in results for mean and maximum host length could be an artifact of the distribution of measured host lengths.*

Likewise, for the significant relationship between generalism and the coefficient of variation in host body size, I do not find any explanation (or attempt thereof) why having two or more hosts of remarkably different body size shall foster generalism… (by the way, the abstract states that generalism is more likely to evolve <when variation in host body size is small>, doesn’t this contradict the positive relationship between generalism and the coefficient of variation? I hope I didn’t completely miss the point, if so, I apologize with the authors).

*Thank you for pointing this out. We now realize there was a typo in the abstract and have corrected it to state that generalism is more likely when variation in host body size is large. This result comes from the result*

*Since increasing f increases , as the two hosts become increasingly divergent in size, and the coefficient of variation will increase.*

Page 12, line 22-23  <We then use evolutionary invasion analysis…>. Maybe it is a philosophical issue, but this seems to me a bit overstated through the ms: the authors use an ecological competition approach assuming that there exist already two parasite species, one specialist and one generalist, they did not use game theoretical approaches or quantitative genetic models (such as in Dayand Proulx 2004 AmNat. A general theory for the evolutionary dynamics of virulence, 163:E40–E62), so I will tune down this through the ms and refer explicitly what they really used.

*The approach that we used is taken from the adaptive dynamics literature (e.g., Geritz et al. 1998), and is typically referred to as an evolutionary invasion analysis. However, your point is well-taken that we are not really searching for evolutionary stable strategies, so we have removed the “evolutionary” from descriptions of our approach, which we feel still well-qualifies as an “invasion analysis.”*

It is odd that mathematical formulas are reported toward the end of the discussion/conclusion (page 15). It is a kind of difficult to identify the take home message, with so many differences, and nuance both in the theoretical and statistical analyses.

Fig. 1 < For G, size of points scale with number of points having that value. >. Why there are a number of points having the same value? Are repeated measures of the same host-parasite association? Is so, isn’t this the case of pseudo-replication?

*As in the other plots, there is likely to be overlap between points due to the large sample size and the limited nature of the possible values for each point, but because only two values of G are possible it is especially hard to see the distribution of points along a line, so the point sizes were scaled to show density of the points. This works because there are several parasite species that have the same mean or max values of host length, and this has been clarified in the figure captions.*

Appendix A

I commend the authors for having provided such a great deal of detail in the Appendix A, it is very useful and I prize such an effort. Yet, I would help the reader to navigate through the monster appendix by adding a table of content – in fact, I found a bit confusing reading in the main body of the ms as well as in the introduction of Appendix A that the authors examined 4 specific cases, whereas the Appendix presents 10 cases.  I would number all the figures in Appendix A, add legends, properly label the axes of each figure, report then unit of measure when applicable and include an inset for the line color/style (when more than one line is drawn in the figure).

Thought not sure, it looks to me that the scripts reported in the Appendix have been generated in Wolfram Mathematica (or Maple?) but I don’t think the authors ever explicitly stated so (nor reported whatever other software package they used or why they used that specific syntax).  The scripts should be thoroughly commented, possibly line by line. Also, I believe that the script shall be either complete of any single line to derive the same results and the graphs included in the Appendix or, if not possible or convenient, the author shall add the complete script as on-line supplementary material. In this way, it should be easier for the reader to cut & paste the code in the appropriate environment and engage in small “variations on the theme” presented by the authors, for instance recycle part of the code to assess the effect of including co-infection. I think that the notation used in the script (for parameters’ name) is in some cases different from that used in the main body of the ms which can also be confusing.

*We have made extensive revisions the Appendices in response to these comments. Appendix A is now broken into two clear parts. The first part deals with the derivation of Eqn. 12 in the main text. The second part deals with all of the cases presented in Table 1 in the main text. Appendix B deals with the model of trophic transmission. We have extensively commented throughout the manuscript and will make the original Mathematica scripts available as an online supplement as well.*

Page 6 in the appendix <dR/dT=… is always negative> does not this depend upon the value of “a”. For small “a”, isn’t this derivate positive?

*This has been rewritten in the new version of Appendix A.*

I am sorry that I have not been able to provide a more supportive review and I apologize with the authors for any error or misinterpretation of their work.

Referee: 2

Comments to Author(s)

The authors investigated the evolution of parasite host specificity in relation to host body size. They combined a mathematical modeling approach and a comparative analysis test. They investigated how host body size and environmental temperature affect the evolution of parasite specialization. On the basis of a literature review they made the prediction about the evolution of parasite specialization in relation to host body size and temperature. They tested the results of their model using a large data base of 20,000 fish-macroparasite associations. For this, they also computed measurements of parasite specialization.

Using their model, they found that there should be few generalist parasites in very small bodied or very large bodied hosts and that generalism should be associated to colder environments.

The model using allometric relationships is well constructed as well the comparative analyses. Indeed, the authors may refer to Morand & Poulin (2002), where similar allometric relationships were used (i.e. body size/mortality rate) (Body size-density relationships and species diversity in parasitic nematodes: patterns and likely processes. Evolutionary Ecology Research 4: 951-961).

I appreciated how the model is described with helpful tables. The discussion is very well developed, by analyzing critically the results of both the analytical models and the comparative analyses, discussing potential weaknesses and offering some research avenues.

This paper is of major importance and would be a benchmark in the understanding of parasite specialization.  I recommend its acceptance in Philosophical Transactions of the Royal Society.