Characterizing parasite generalism illuminates patterns of host-parasite associations

Park, A.W.
Farrell, M.J.
Schmidt, J.P.
Huang, S.
Dallas, T.A.
Pappalardo, P.
Drake, J.M.
Stephens, P.R
Poulin, R.
Nunn, C.L.
Davies, T.J.

Abstract

The distribution of parasites among mammalian hosts is complex and represents a differential ability or opportunity to explore the host tree of life. Using data on over 1400 parasite species known to each infect between 1 and 81 mammal host species, we characterize the generalism of parasites using standard effect sizes for three metrics describing host relatedness: average pairwise host phylogenetic distance (PD), maximum PD and degree of aggregation of hosts (the tendency for the hosts of a parasite to cluster within the phylogenetic tree). We explore variation in these metrics in terms of parasite taxonomy and transmission mode. Of the multi-host parasites, the majority are generalist in the sense that the average relatedness of their mammalian host species is roughly equivalent to that expected from a random sample of mammals, with a minority associated with host sets that are more related than expected by chance. The degree of host relatedness is affected by both parasite taxonomy and transmission mode; bacteria and arthropod parasites are typically the most generalist, helminths are intermediate and viruses and protozoa are the most specialist on average. However, characterizing variation within and between groups is important; for example, viruses are more specialist than helminths on average, but the virus group has a larger range on the specialism-generalism continuum, such that the virus group contains some of the most generalist parasites. While not as strong as parasite taxonomy, transmission mode also influences the degree of parasite generalism, with closer relatedness of hosts exhibited for parasites that rely on close contact or complex life-cycle transmission involving intermediate hosts, compared to those with environmental or vector-borne transmission. Both the span of a parasite across the host tree and the tendency to infect clusters of related hosts vary among parasites. For example, all bacteria have large spans, equivalent to random host sets. However, they exhibit the least propensity to infect host clusters within this span. Instead, they appear to typically infect relatively unrelated host species. The taxonomy and transmission modes of parasites with only one known host species are broadly reflective of multi-host parasites, with rarely single-host viruses being a notable exception. Lastly, a host species' evolutionary distinctiveness is a weak predictor of the number of parasite species asscociated with it. Collectively, this macroecological perspective helps to generate a broad picture of how certain types of parasite and transmission mode are differentially linked with the tendency to associate with multiple host species, captured by a set of complementary metrics defining the dimensions of parasite generalism.

Keywords

Parasite, multi-host, generalism, transmission mode, phylogenetic, macroecology

Introduction

A generalist parasite has a minimal definition: *able to infect at least two host species*. However, for an equivalent number of host species, parasites may vary in their ability to infect distantly related hosts, differentiating their potential to explore the space of host species along with their ability to infect target species, including humans. The underlying reasons for observing relatedness of hosts of a given parasite include biogeography and co-evolutionary history [1], translating to opportunities for infection, quality resources provided by a host, and limited costs in terms of immune defense and disease-induced mortality. Mechanisms that may allow parasites to overcome host species barriers include rapid evolution and antigenic plasticity [2], novel spatial translocation of host or parasite [3], and evolution of host tolerance [4]. Consequently, it is challenging to predict the potential for parasites to acquire hosts that are distantly or closely related to their existing hosts, in spite of the utility for understanding the distribution of parasites and predicting future zoonoses.

Previous research has defined parasite generalism in alternative ways, including satisfying the criteria of infecting humans and animals [2] as well as using host taxonomies to characterize the degree of parasite generalism [5,6]; for example, classifying parasites by their ability to infect hosts only within a species, or genus, or family, etc. Here, we build on methods to generate continuous measures of generalism [7] using a detailed mammalian host phylogeny to define parasite generalism by the patterns of relatedness of the host sets of parasites. The study is unparalled in the number and diversity of host and parasite species considered together. Known associations between hosts and parasites come from the global mammal parasite database (GMPD) [8,9], from which this study focuses on 430 mammal species representing primates ($n = 178$), carnivores ($n = 111$) and ungulates ($n = 109$). The approach allows both comparison among parasites and calibration against a null model [10]. From each host set, we calculate the standard effect size of mean pairwise PD between host species, a measure that has previously been used successfully to explain parasite community similarity in primates [11]. In addition, we calculate standard effect sizes for both the maximum PD of host sets and a similar metric that captures the tendency for hosts to be clustered in the phylogenetic tree [12].

We determine that average relatedness of hosts varies with parasite type, defined both taxonomically and by transmission mode, as has been suggested in earlier studies based on categorical definitions of parasite generalism [2,6]. The speciose study, which includes over 1500 parasite species, enables generalism to be characterized further. All taxonomic and transmission mode groupings contain parasites that have large spans in the host tree, measured by standard effect size of maximum PD among hosts. Bacteria, along with environmentally-transmitted and complex life-cycle parasites show the least propensity to infect phylogenetically clustered host groups. Single host parasites ($n = 653$) are almost as common as multi-host parasites ($n = 749$) and show similar distributions among parasite taxonomic and transmission mode groupings, with the notable exception of viruses, which are more likely to infect multiple host species. We observe that the number of parasite species known to infect a host species is not well explained by host evolutionary distinctiveness.

Macroecological approaches to analyzing large host-parasite datasets, such as the one presented here, provide a backdrop against which to consider individual case studies, faciltate an appreciation and understanding of robust patterns of host-parasite associations and extend the concept of parasite generalism to include phenomena such as parasites associating with multiple 'pockets' of related hosts. Further, the study lays out parasite group-level differences in generalism, including within-group variation, which helps navigate to sources of variation that may ultimately be explained by environmental, physiological and behavioral heterogeneity [13–16].

Materials and Methods

Parasite records of occurrence with terrestrial mammals were obtained from GMPD [8,9]. These records include latin binomials and taxonomic classification for host and parasite species, and transmission mode for the majority of parasite species (~80%). Transmission modes were assigned in GMPD based on an extensive literature review [8,9], from which we retain the four heierarchically-dominant transmission modes: close

contact, complex life-cycle, environmental and vector-borne (other transmission modes are recorded in GMPD, such as soil-transmission, but these belong to one of the four main transmission modes). GMPD data were manipulated in the R programming environment [17] to establish the number and identity of parasite species per host species and the number and identity of host species per parasite species.

Additionally, the mammal phylogenetic supertree [18] was used to obtain the PD between all pairs of hosts. For the host set of each parasite species, the mean pairwise PD between hosts was calculated, as was its standard effect size [19]. The latter was obtained using the R package *Picante* [20], and a null model in which the community data matrix (host-parasite occurrences) was randomized with the independent swap algorithm [10] maintaining species occurrence frequency and sample species richness. This measure captures the average relatedness of the host species of a given parasite species on a standardized scale for comparison. Two further standard effect size metrics were calculated on each parasite species' host set in the same way: maximum PD of any two hosts in a parasite's set, and the ratio of the mean minimum PD to the maximum PD. The former metric provides a standardized measure of each parasite's span across the host phylogenetic tree, while the latter provides a measure of the tendency for the host species to exhibit an aggreagated or clumped distribution within the span [12]. This is achieved by calculating the PD of each host species to its nearest host, averaging this across all hosts and scaling by the span. The standard effect size calculations return a z-score and a p-value, where z-scores below -1.96 (specialist) and above +1.96 (generalist) are considered signifcantly different from the null expectation, assuming the null model generates a normally-distributed set of scores [21].

Host evolutionary distinctiveness, measured as millions of years of evolutionary change, was established from the mammal supertree data [18] using the *evol.distinct* function in the R package *Picante* [20], utilizing the *equal splits* routine in which shared branches are apportioned equally among descendant lineages [22]. Each hosts's terminal branch length was also recorded as an alternate measure of evolutionary distinctiveness.

Results

Across all parasites there is a trend of infecting hosts that are more related than expected by chance, indicated by negative z-scores for mean pairwise PD between hosts (Fig. 1, global median z-score=-1.16). However, while 96% of parasites exhibit negative z-scores (trending to more specialist than expected by chance), only 15% of parasites have significantly negative z-scores. There are no parasites that have significantly positive z-scores.

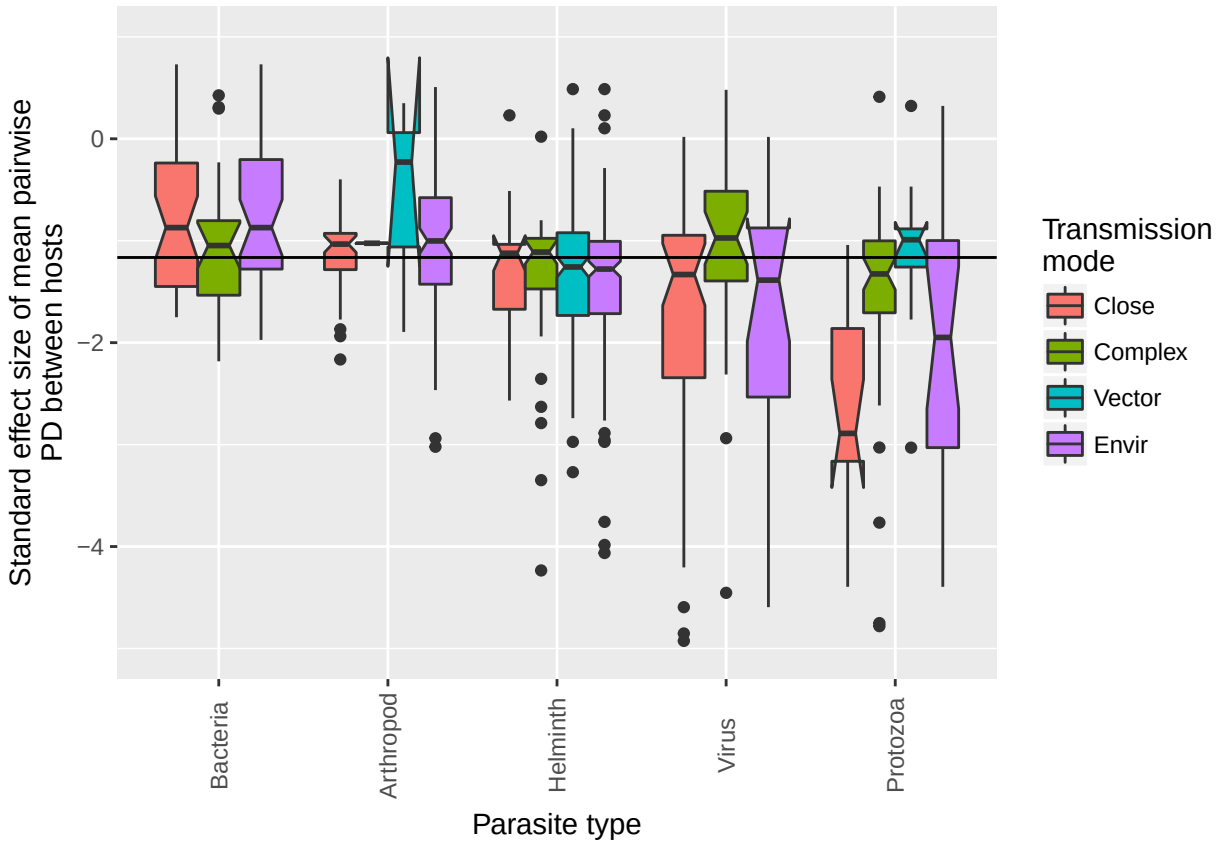


Figure 1. Boxplot of z-scores in each parasite group stratified by transmission mode. Positive z-scores indicate that a parasite species tends to infect host species that are more distantly related than expected by chance (under the null model). Alternatively, negative z-scores indicate the opposite. In terms of deviation from the null model, 15% of z-scores are significantly negative (none of the positive z-scores are statistically significant). Solid horizontal line indicates the global median z-score.

The proportion of these significantly negative z-scores varies across parasite type, with smallest to largest proportions corresponding to bacteria, arthropods, helminths, protozoa and viruses (test for equality of proportions, $p < 0.001$, Fig. 2 - top panel). The proportion of significantly negative z-scores is not significantly different across parasite transmission modes (test for equality of proportions, p=0.77, Fig. 2 - bottom panel).

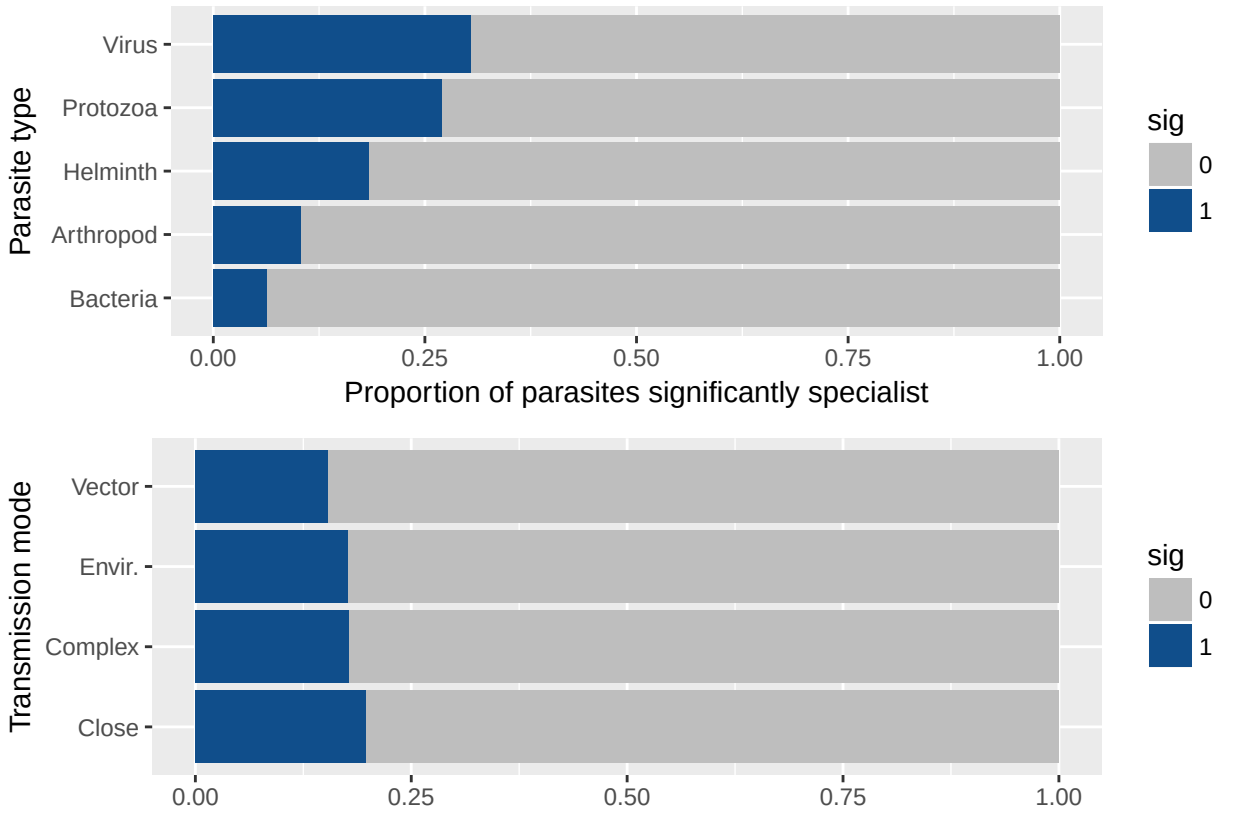


Figure 2. Proportion of parasites with host sets in which the hosts are more related than expected by chance, grouped by parasite type (top panel) and transmission mode (bottom panel)

However, across the full range of z-scores, each of the two covariates (parasite type and transmission mode) and their interaction explain differences between mean values (ANOVA: Parasite type - $p < 0.001$, Transmission mode - $p = 0.003$, Interaction - p<0.001, supplementary material).

Parasite type by transmission mode interactions are particularly driven by protozoa and viruses, and close contact and vector-borne transmission (Fig. 3); close contact transmission is associated with atypical extreme specialism in these groups and vector-borne transmission is associated with atypical generalism in these same groups.

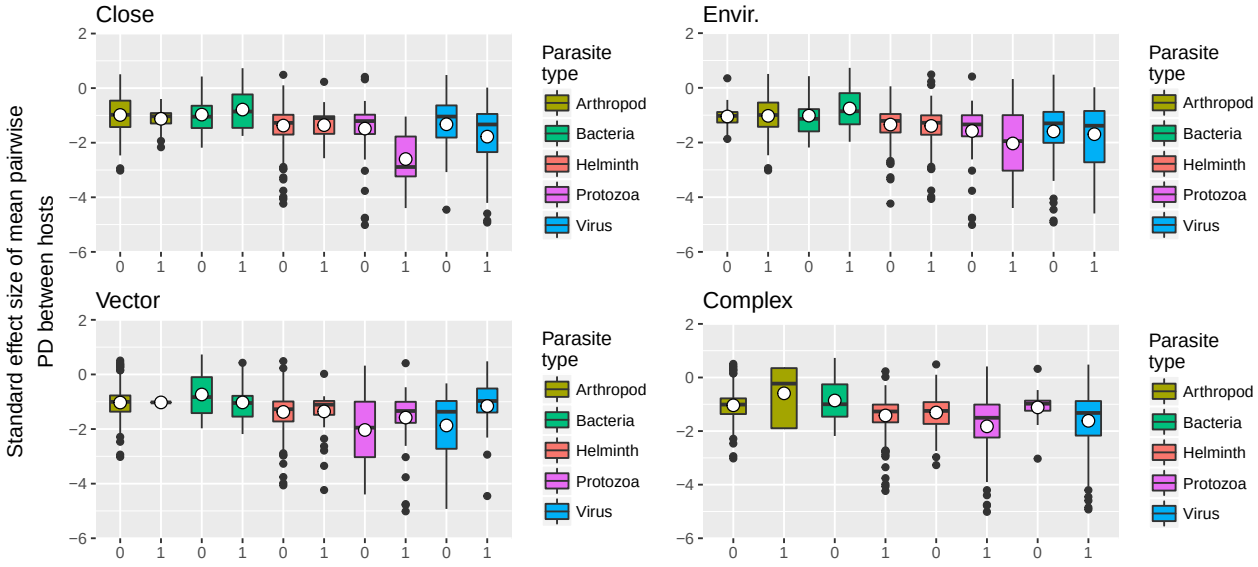


Figure 3. Standard effect size of mean pairwise PD between hosts stratified by parasite type (colored bars) and presence (1) or absence (0) of each transmission mode (named subplots), indicating interactions between parasite type (particularly protozoa and viruses) and transmission mode (particularly close contact and vector-borne transmission)

In comparing the most abundant parasite group (helminths) with the group containing the greatest number of specialists (viruses) it is notable that viruses exhibit much wider variation in degree of generalism (Fig. 4), containing many extreme specialists (large, negative z-scores) but also many relative generalists (positive z-scores). By contrast, helminths are more consistent in their degree of generalism, and contain relatively few species that are at either extreme of the generalism-specialism continuum.

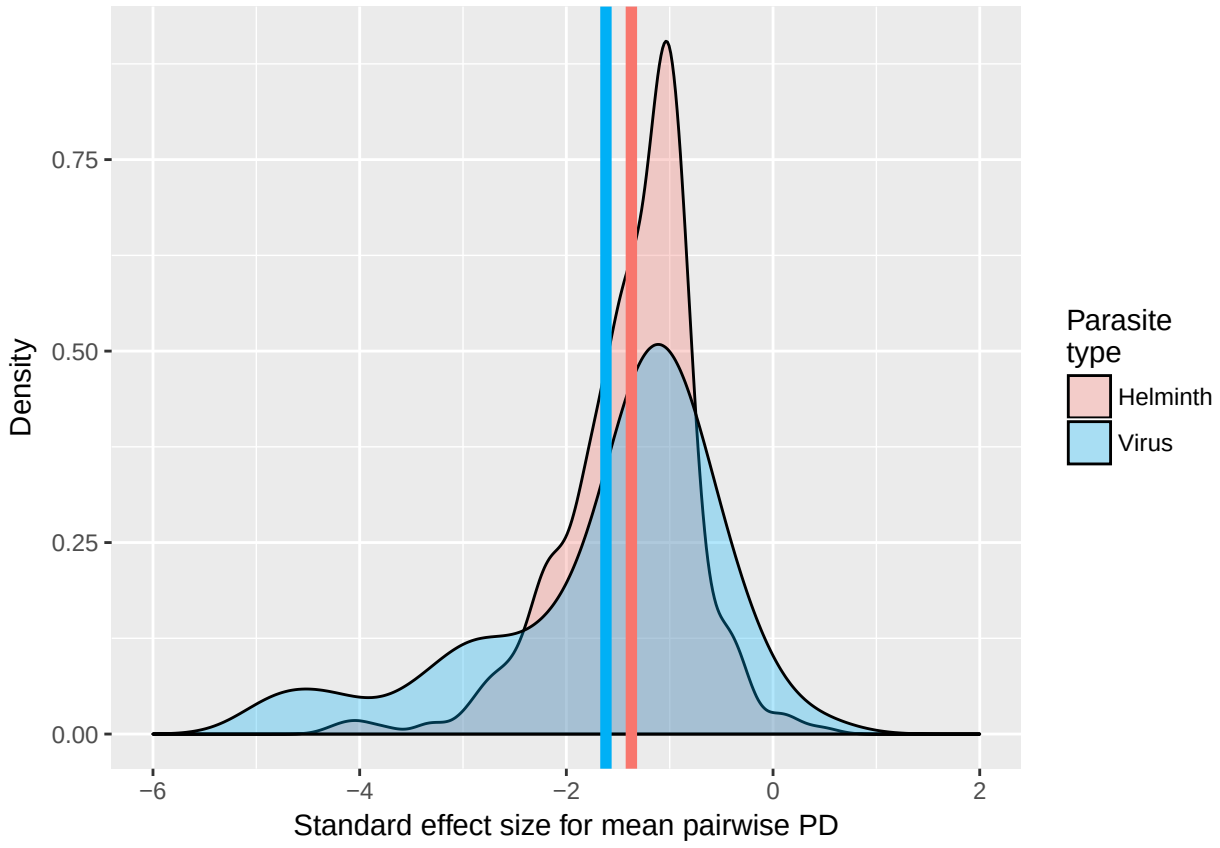


Figure 4. Probability density function of z-scores for mean pairwise PD between hosts of helminths (red) and viruses (blue). Helminths are more generalist on average (group mean values denoted by color-coded vertical lines), but the virus group contains more extreme parasites in terms of both specialism and generalism.

The standard effect size of maximum PD in a parasite's host set provides a comparative measure of the span that a parasite exhibits across the host phylogenetic tree (Fig. 5). In addition, the standard effect size of the ratio of mean minimum to maximum PD measures the tendency for hosts of a parasite to aggregate in the mammal phylogeny. All parasite taxonomic groups exhibit variation in both these metrics. Protozoa, virus and helminth groups contain several parasite species whose span is significantly smaller than expected by chance (Fig. 5 - left of boxes in subplots, see supplementary material for histograms). Viruses and protozoa additionally contain some species whose hosts are more aggregated than expected by chance (Fig. 5 - below boxes in subplots, see supplementary material for histograms).

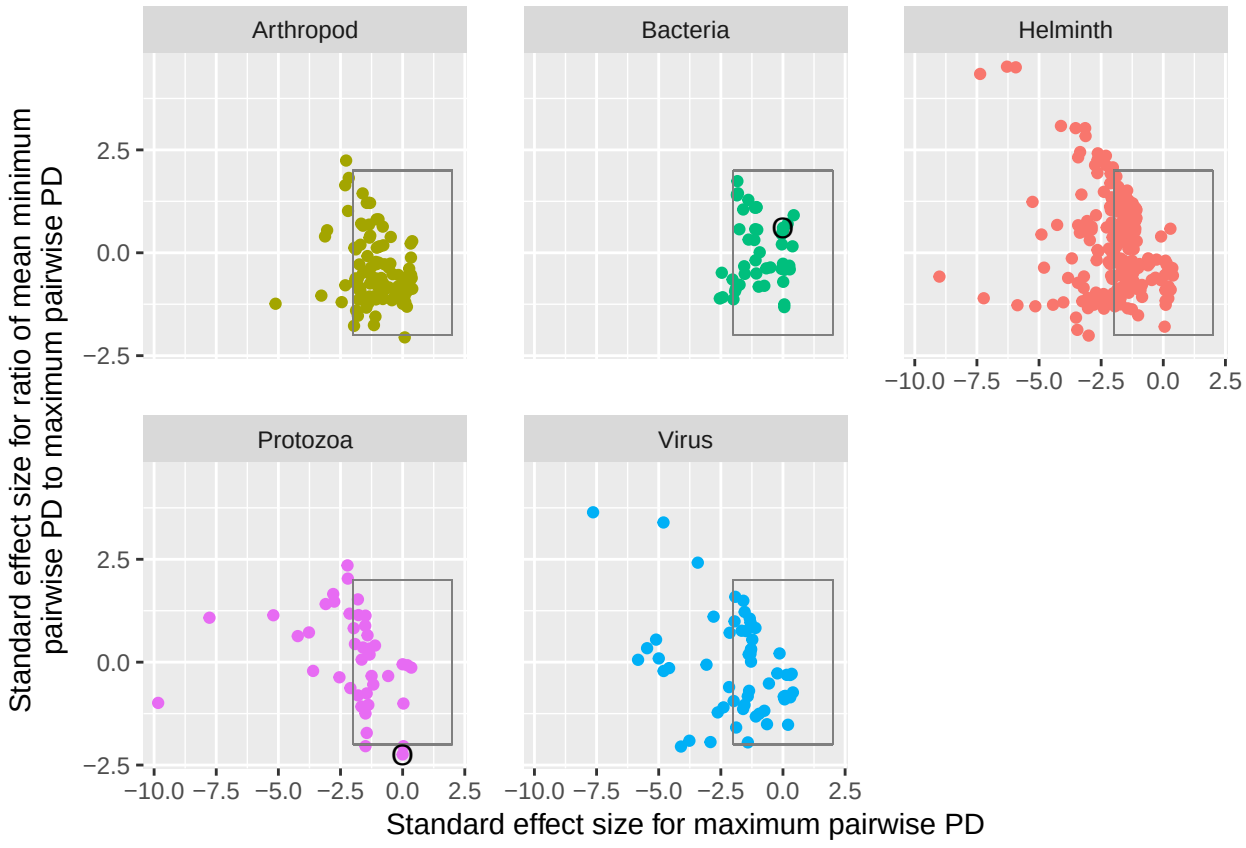


Figure 5. Scatter plot showing x: standard effect size of maximum pairwise PD of a parasite's host set and y: standard effect size of the ratio of mean minimum pairwise PD to maximum pariwise PD, colored by parasite type. The x-axis provides a comparative measure of a parasite's span across the host phylogenetic tree, and the y-axis provides a comparative measure of the aggregation of a parasite's hosts in the phylogenetic tree. Circled bacterial and protozoan parasites are *Escherichia coli* and *Trypanosoma cruzi*, respectively. These parasites have a similar span ($x \approx 0$), but very different host aggregation patterns (*E. coli* randomly distributed host species, *T. cruzi* clumpily distributed host species). Boxes encompassing $-2 < x, y < 2$ contain parasites whose z-scores are not statistically different from null model scores (random host sets).

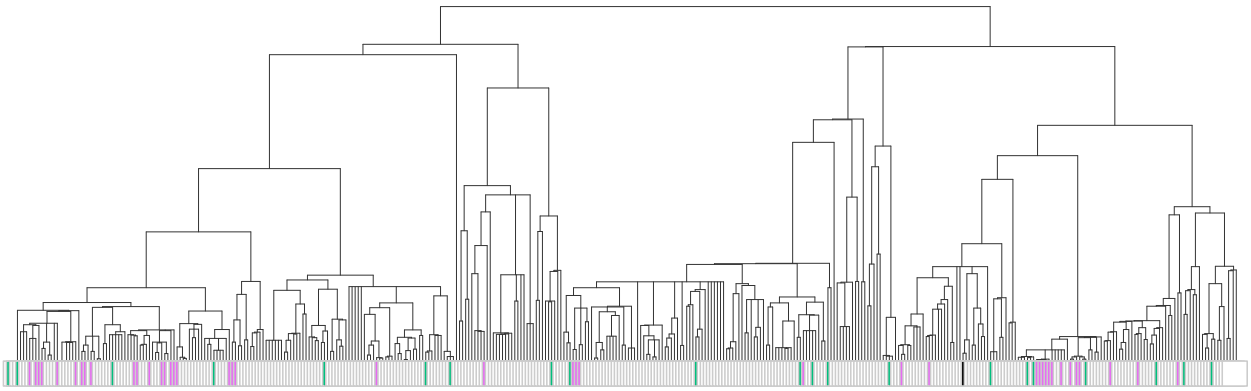


Figure 6. Examples of clumpy vs. non-clumpy distributions of parasites of similar span in the host phylogenetic tree. The parasites are *Escherichia coli* and *Tryapnosoma cruzi*, colored according to their groupings in Fig. 5.

The main analysis is centered on parasite species that infect at least two host species. Including parasite species for which only one terrestrial mammal host is identified in the database provides an opportunity to

compare single-host parasites and multi-host parasites (Fig. 7). Quantitatively, single host parasites ($n = 653$) are almost as common as multi-host parasites ($n = 749$), and broadly reflect the composition of parasite taxonomy and transmission modes observed in multi-host parasites. Two notable exceptions are the virus group, which is under-represented among single-host parasites in general, and protozoan complex life-cycle parasites, which are more commonly single-host than multi-host.

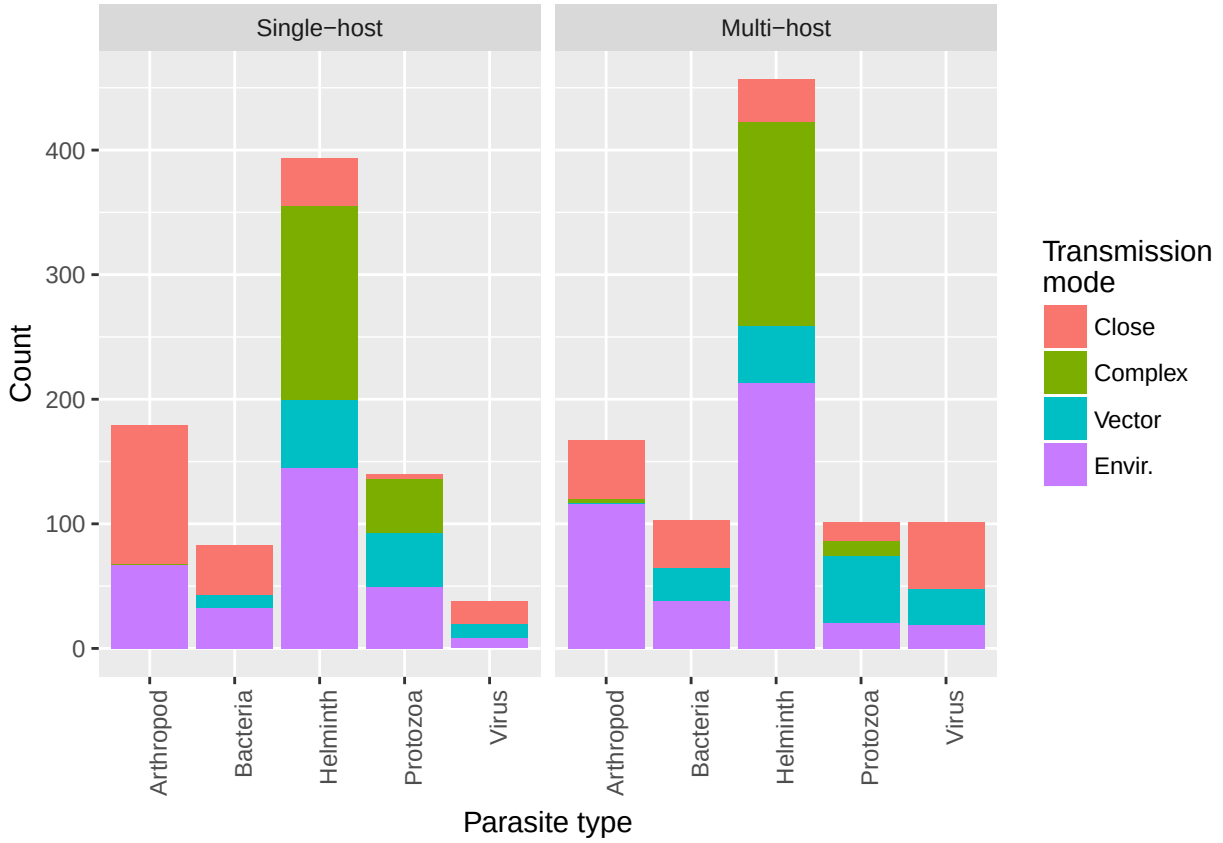


Figure 7. Number of single-host and multi-host (>1 host species) parasites as a function of parasite type and transmission mode. Parasites with more than one reported transmission mode (343/1402) are counted in each valid transmission mode category.

One potential explanation for the number of parasite species associating with a host is the amount of time the host species has evolved since common ancestry. However, neither the host's evolutionary distinctiveness [22] nor terminal branch length (Fig. 8 - top and bottom panels, respectively) are strongly associated with the number of known parasite species (and this is true when inspecting each host order separately - supplementary material).

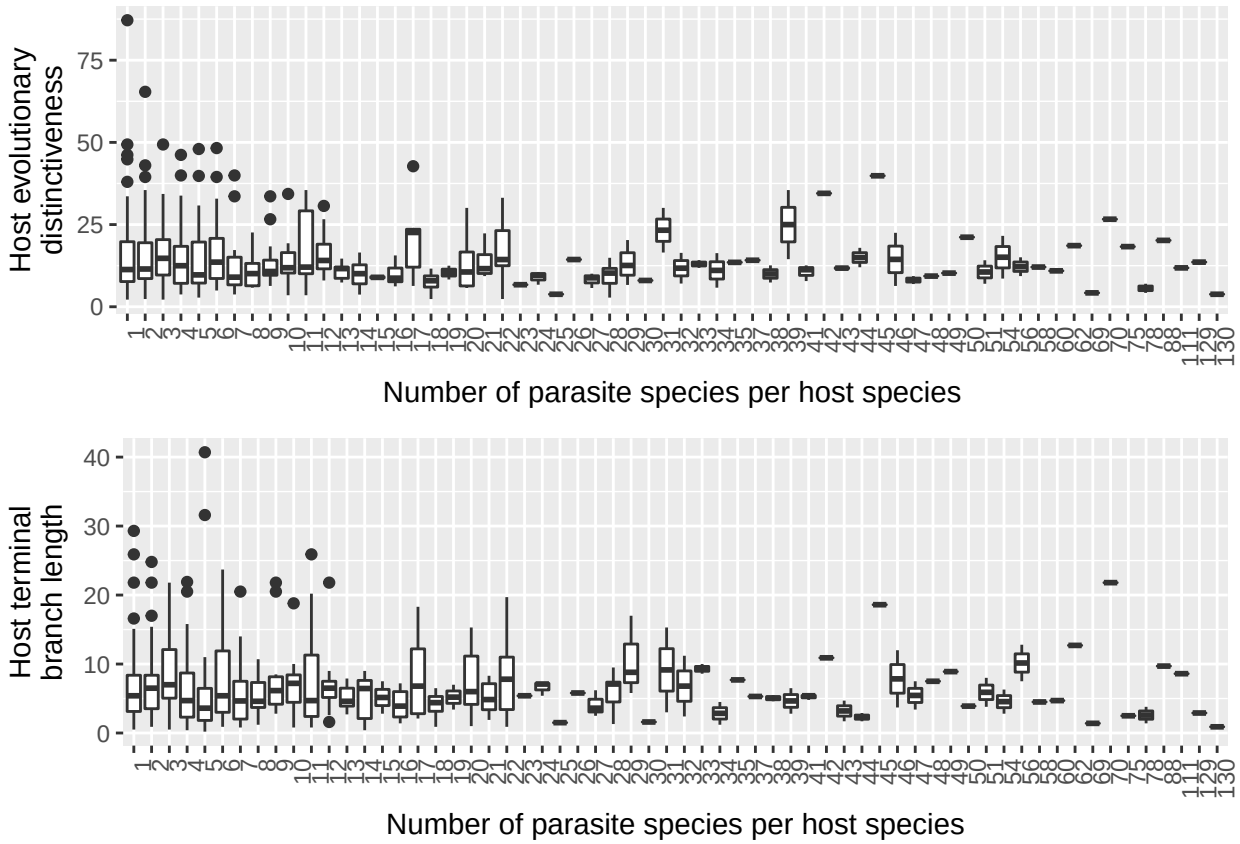


Figure 8. Box-and-whisker plots showing the relationship between number of parasite species per host species and host evolutionary distinctiveness (top panel) and host terminal branch length (bottom panel).

The number of parasite species per host species is over dispersed (variance to mean ratio ~25), meaning that a small number of host species are associated with a large number of parasite species. Accordingly, a negative binomial generalized linear model was used to predict the number of parasite species by each measure of evolutionary distinctivess (and, optionally, the number of host records in GMPD, which was used to control for sampling bias). Only the model with evolutionary distinctiveness (but not the number of host recrods) showed a significant relationship (supplementary material).

Discussion

We have shown that most multi-host parasites are relatively generalist. However, among the ~750 multi-host parasite species considered here, some infect host species that are comparatively closely related. Of the five parasite groups, bacteria are the most generalist on average and protozoan and viral parasite groups contain the most specialist parasites. While transmission mode naturally impacts the opportunity for parasites to encounter novel host species, it was less influential in determining parasite generalism than parasite taxonomy. However, there were interactions between parasite type and transmission mode; both protozoan and viral parasites exhibit specialism when undergoing close contact transmission, in agreement with primate parasite research [6]. Additionally, protozoan parasites that are environmentally transmitted are often more specialist than expected by chance, whereas vector-borne protozoa are typically generalist.

Previous research based on taxonomic definitions of generalism have tended to suggest that viruses and protozoa are relatively generalist and helminths relatively specialist [6]. The differences between those results and the ones we present here lie in the definition of generalism, which can exagerate rare but large host species jumps. We found, for example, that while viruses often infect closely related hosts (shown by large negative z-scores), they are also a group that contains several parasites whose hosts are distantly related. In

concert with purely taxonomic definitions of generalism, such as the ability to infect hosts of different orders, examining the standard effect sizes of mean hostPD, span and clustering complements this approach with continuous, standardized measures that facilitate comparison across parasite species and with null models.

While the underlying data represent known host-parasite occurrences, they are not necessarily indicative of parasite fitness. Parasites may jump between host species, establishing in each, e.g. Rabies virus [23], or maintain themselves in some reservoir host species and spill over to others, e.g. West Nile virus [24]. Consequently, it is difficult to infer future potential for novel host acquisition from existing data. Plausibly, parasites with rapid evolution may be both good adaptors to, and explorers of, the space of host species, as is indicated by the virus group which exhibits clustering of host species in the mammal phylogeny, which is indicative of taking occasional leaps to novel host species and subsequently colonizing closely related host species.

Host species geographical ranges have previously been shown to be one of the strongest predictors of viral parasite sharing among primates [25], and are only implicitly included here due to the non-independence of range overlap and phylogenetic relatedness [1]. Explicit inclusion of geography is a promising line of macroecological inquiry.

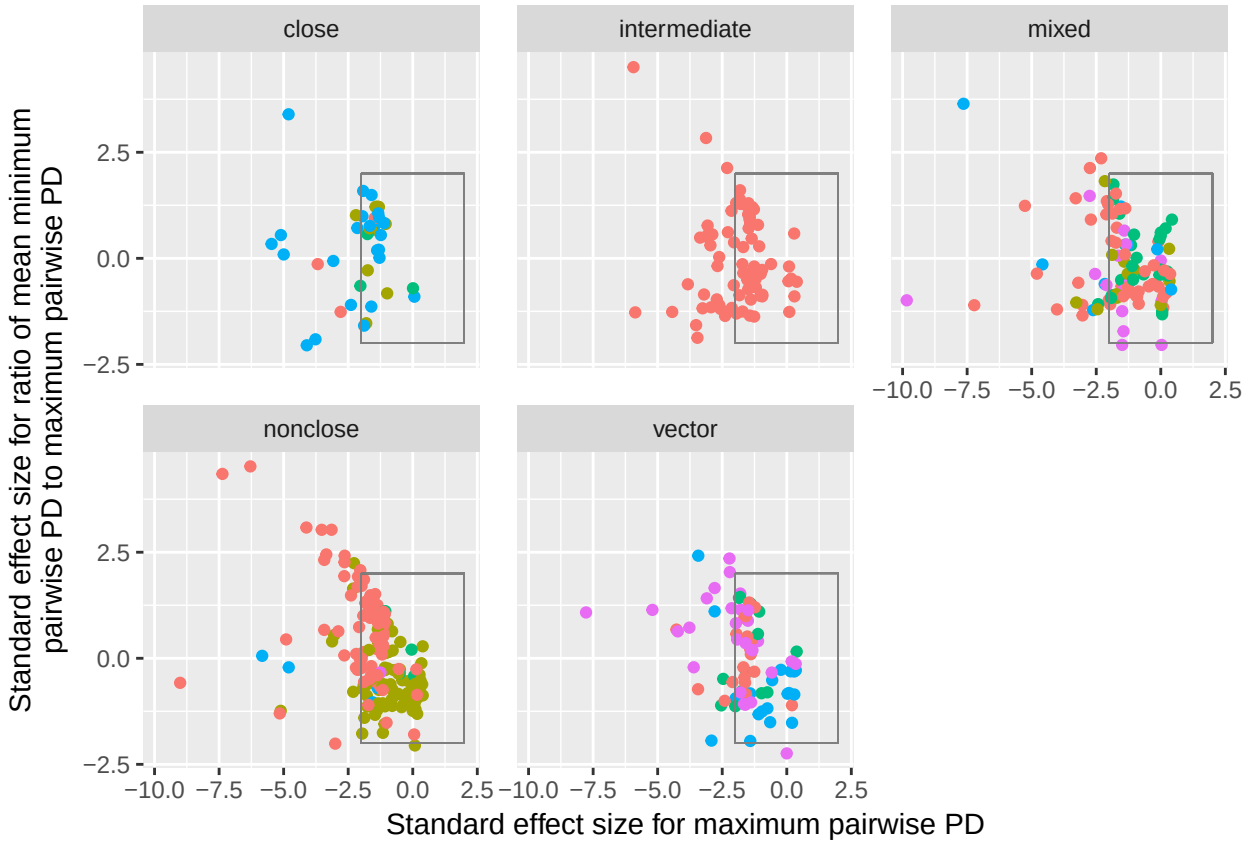
Supplementary material

Parasite taxonomy, transmission mode and their interaction contribute to variation in mean pairwise PD of host sets of each parasite

```
#simple test to motivate investigation of factors and their interaction
basic.aov<-aov(lm(nri.flat$mpd.obs.z~nri.flat$para.type*nri.flat$tmode))
summary(basic.aov)
```

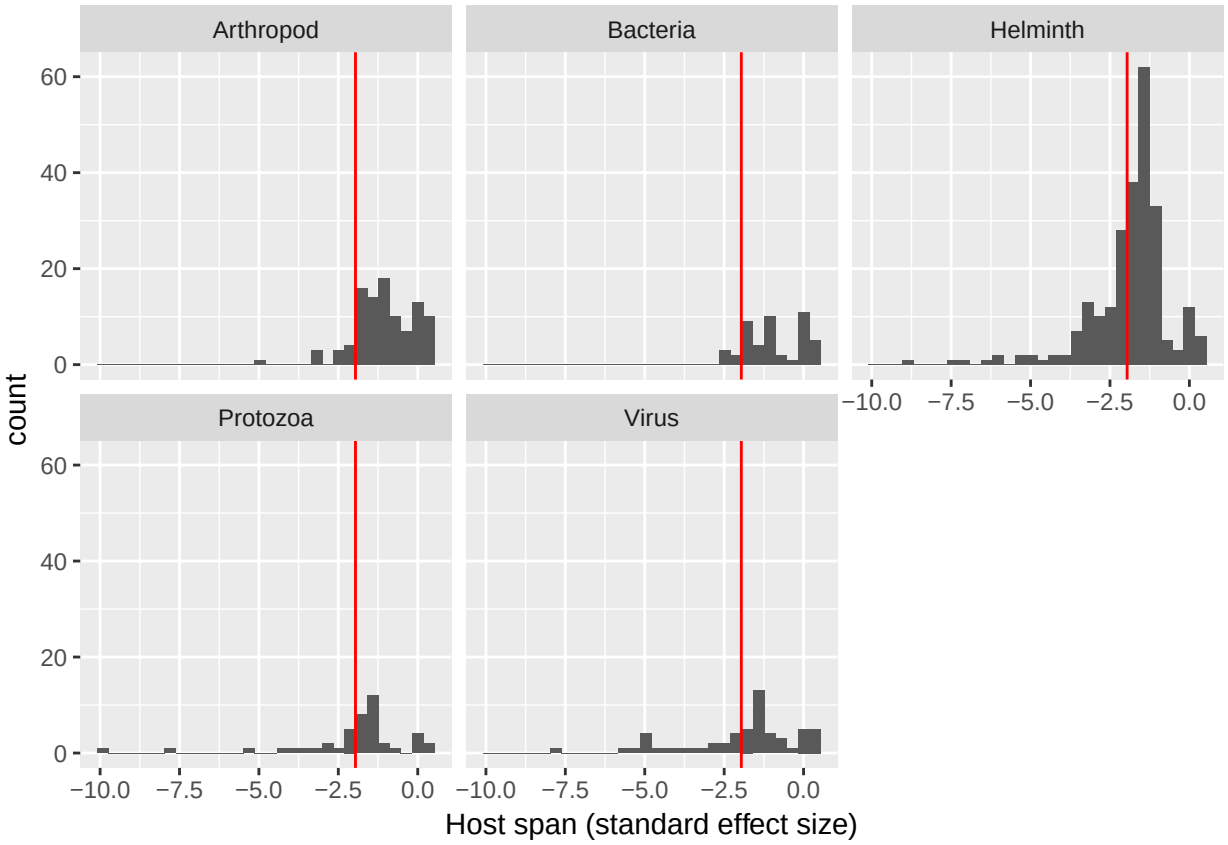
```
##                                   Df Sum Sq Mean Sq F value   Pr(>F)
## nri.flat$para.type                 4   64.7  16.184  27.292  < 2e-16 ***
## nri.flat$tmode                     3    8.5   2.831   4.774 0.002621 **
## nri.flat$para.type:nri.flat$tmode 10   20.8   2.075   3.500 0.000152 ***
## Residuals                        911  540.2   0.593
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Parasite transmission mode as a predictor of span and aggregation of hosts in the mammalian phylogeny

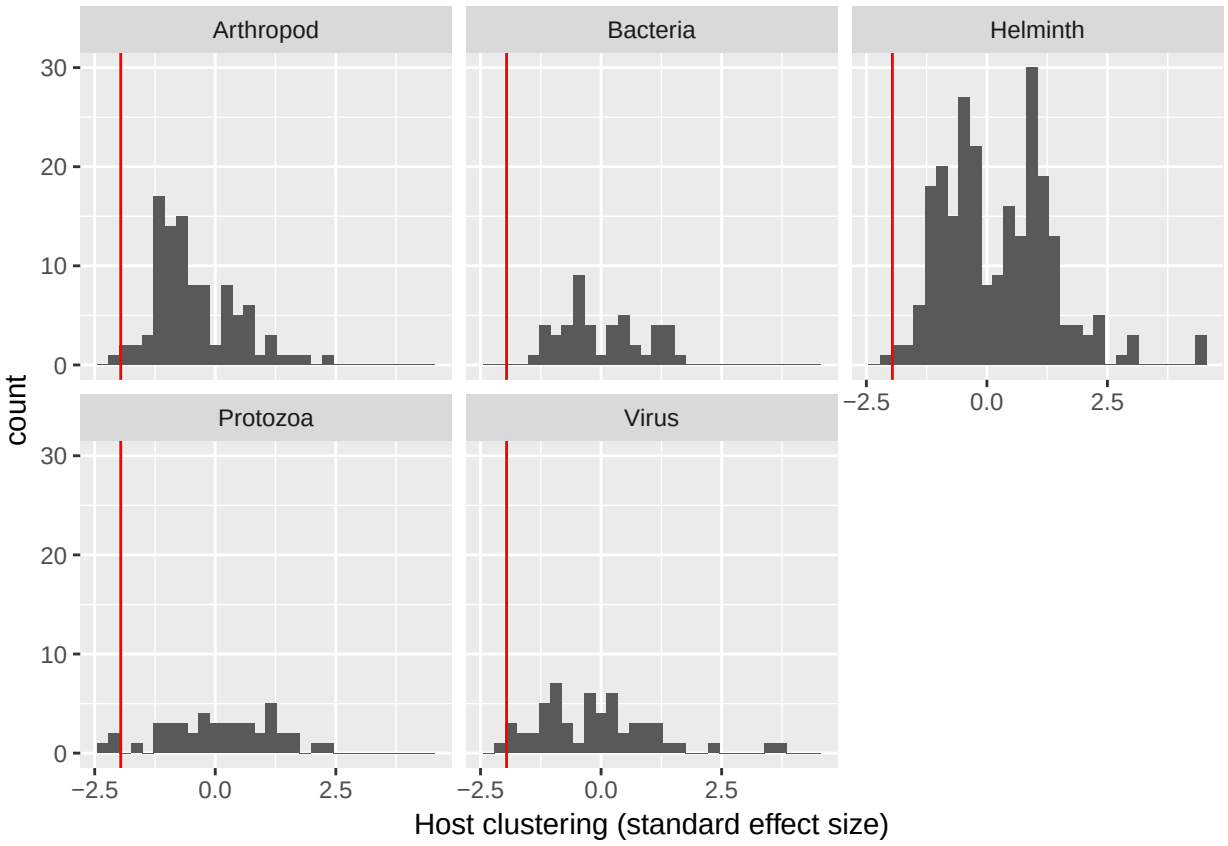


Each transmission mode group contains a majority of parasites species whose standard effect sizes for span and aggregation are not different those of equivalent random host sets. All groups contain some parasites that have a smaller span than expected by change and contain some parasites that exhibit patterns of host species phylogenetic aggregation. Patterns are relatively similar across all groups.

Host span and clutering tendency as function of transmission mode



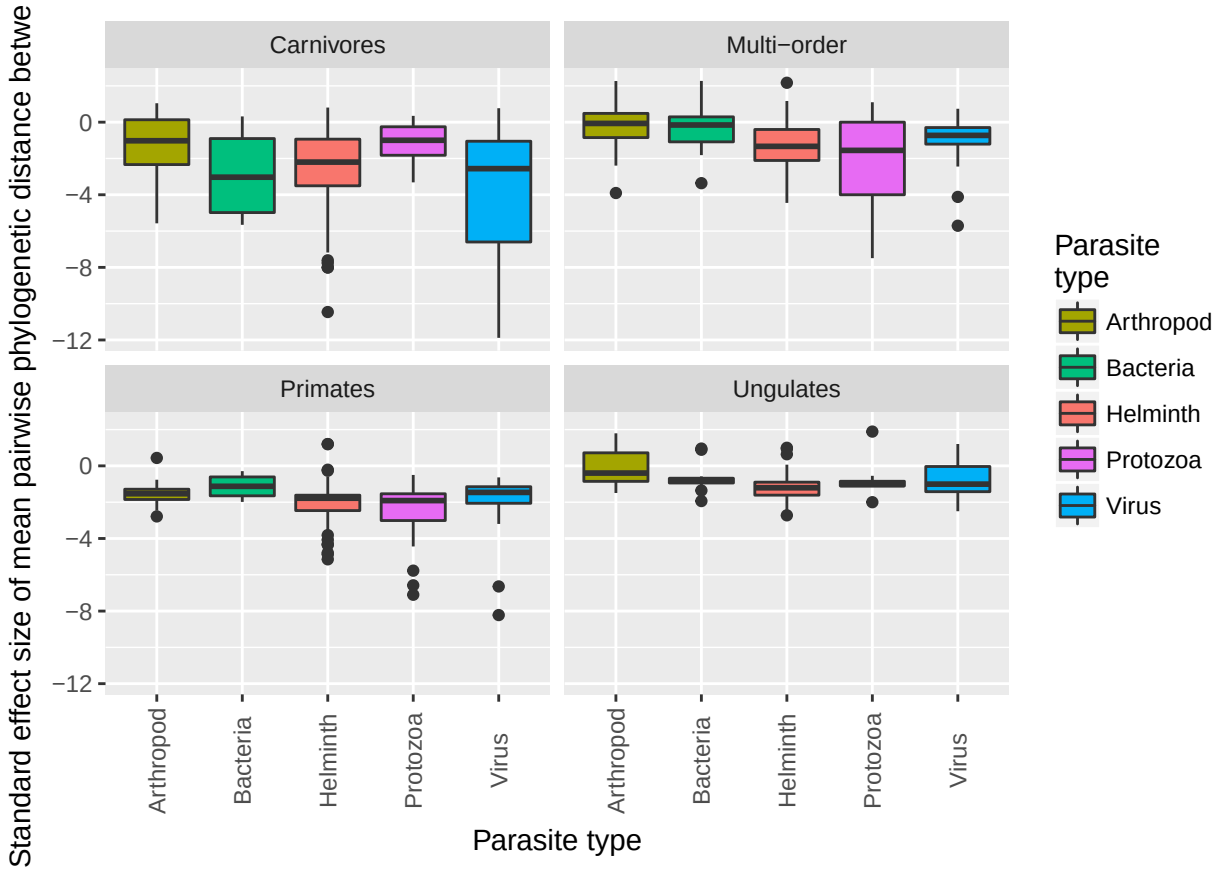
Components of distributions left of vertical red lines have smaller spans than expected by chance.



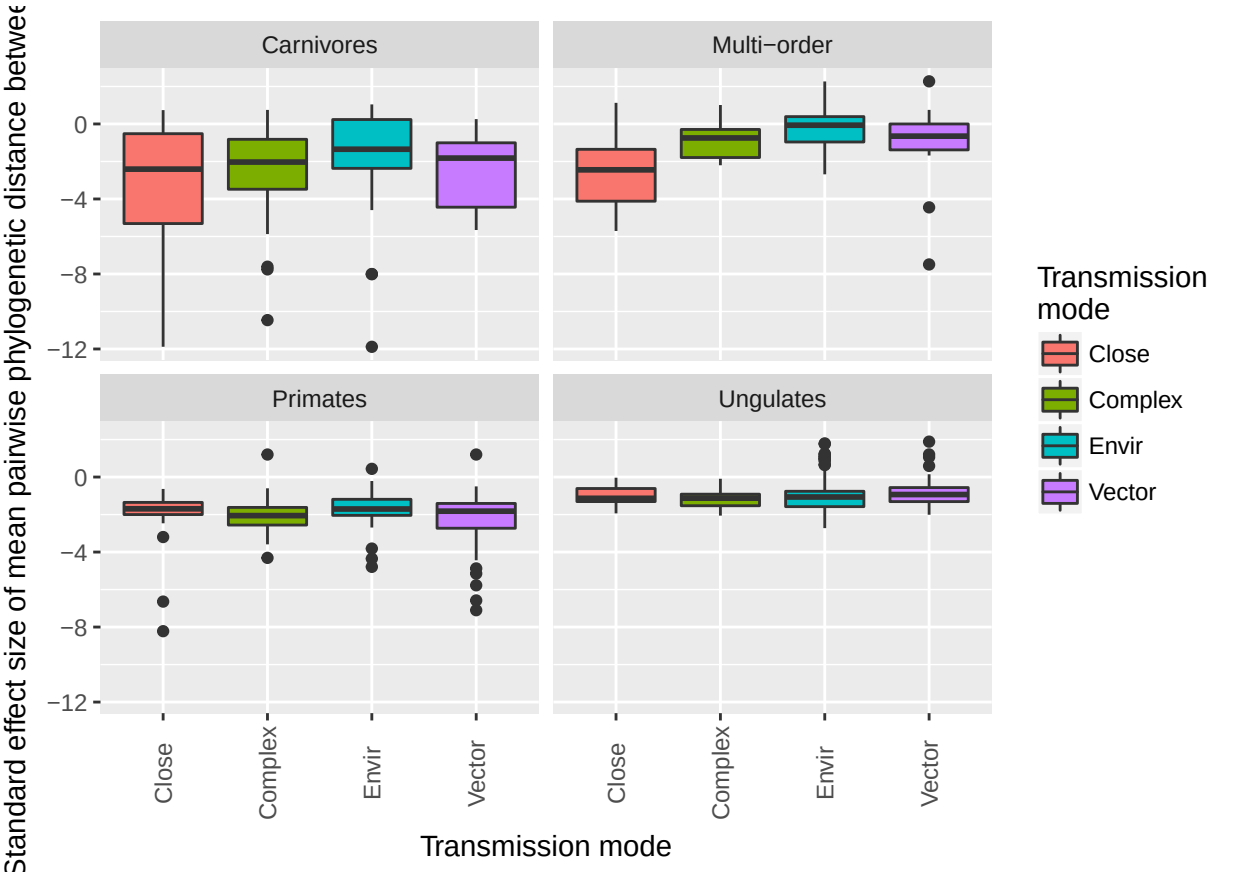
Components of distributions left of vertical red lines exhibit greater host clustering than expected by chance.

Effects of host orders on relatedness of a parasite's host set

Several parasites exclusively infect hosts within a host order. Grouping host species by order allows us to inspect how results on generalism are driven by the non-monophyletic structure of a tree with three host orders. Here, the null models are constructed in a similar way to the main text, except that random mammalian host species are selected according to the host taxonomic bias of the parasite. For example, a parasite known to infect 5 carnivores, 5 primates and 0 ungulates would have its random sets constrcuted from 5 randomly selected carnivores and 5 randomly selected primates. The plot separates parasites that infect hosts of multiple orders, carnivores only, primates only and ungulates only.



The patterns for standard effect size of mean pairwise PD of all host species is broadly reflected in the patterns of specific host orders. For viruses infecting only one host order, the carnivore host group is particularly associated with specialist viruses.



Parasites affecting only primates and ungulates do not exhibit differences in specialism/generalism due to parasite transmission mode. The carnivore-exclusive parasites show a similar trend to the main result (Figs. 1 & 2).

Negative binomial modeling of parasite species richness

Summary of negative binomial GLMs with predictor variables selected from: evolutionary distinctiveness (ed), terminal branch length (tbl) and number of host records in GMPD (n.records), and response variable number of parasite species per host species (n.para)

% latex table generated in R 3.3.1 by xtable 1.8-2 package % Sat Feb 4 15:42:12 2017

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.7623	0.0991	27.88	0.0000
ed	-0.0152	0.0055	-2.77	0.0056

% latex table generated in R 3.3.1 by xtable 1.8-2 package % Sat Feb 4 15:42:12 2017

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.6508	0.0917	28.92	0.0000
tbl	-0.0156	0.0104	-1.50	0.1340

% latex table generated in R 3.3.1 by xtable 1.8-2 package % Sat Feb 4 15:42:12 2017

% latex table generated in R 3.3.1 by xtable 1.8-2 package % Sat Feb 4 15:42:12 2017

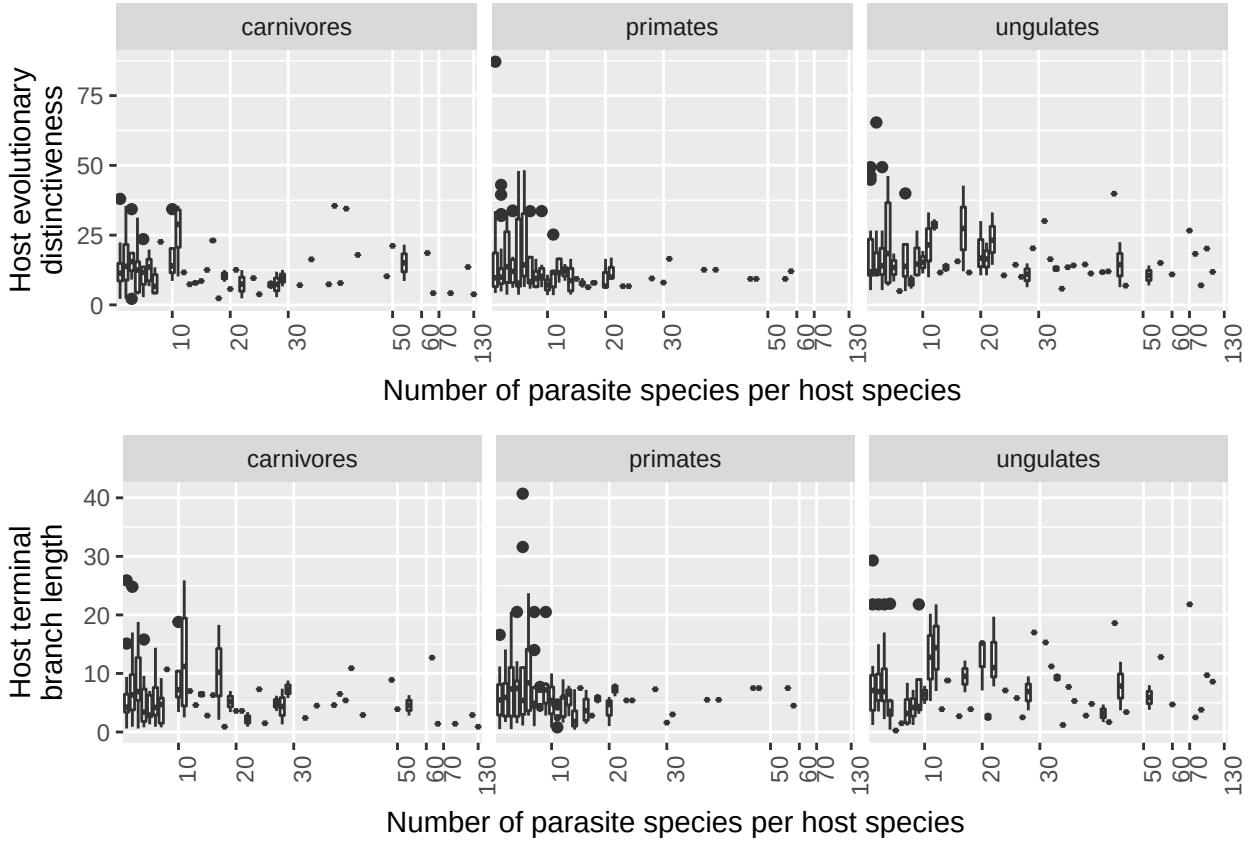
For models using host evolutionary distinctiveness or terminal branch length (with and without number of host records, to control for potential sampling bias), to predict the number of parasite species per host

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.0995	0.0881	23.82	0.0000
ed	-0.0062	0.0047	-1.32	0.1852
n.records	0.0020	0.0001	19.93	0.0000

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.0425	0.0814	25.11	0.0000
tbl	-0.0052	0.0089	-0.59	0.5581
n.records	0.0020	0.0001	20.05	0.0000

species, the main predictor variable is only significant for evolutionary disctinctiveness without controlling for number of host records.

Lack of association between host evolutionary distinctiveness and number of parasite species is exhibited across host orders.



When separately inspecting the relationship between evolutionary distinctiveness and number of parasite species per host species within host orders (carnivores, primates, ungulates) there is no significant relationship in any host order.

References

1. Poulin, R. & Keeney, D. B. 2008 Host specificity under molecular and experimental scrutiny. *Trends in parasitology* **24**, 24–8.

2. Woolhouse, M. E. J. 2002 Population biology of emerging and re-emerging pathogens. *Trends Microbiol*

Trends Microbiol **10**, S3–S7.

3. Farrell, M. J., Berrang-Ford, L. & Davies, T. J. 2013 The study of parasite sharing for surveillance of zoonotic diseases. *Environmental Research Letters* **8**, 015036.

4. Raberg, L., Graham, A. L. & Read, A. F. 2009 Decomposing health: tolerance and resistance to parasites in animals. *Philosophical Transactions of the Royal Society B: Biological Sciences* **364**, 37–49.

5. Poulin, R. & Mouillot, D. 2003 Parasite specialization from a phylogenetic perspective: a new index of host specificity. *Parasitology* **126**, 473–480.

6. Pedersen, A. B., Altizer, S., Poss, M., Cunningham, A. A. & Nunn, C. L. 2005 Patterns of host specificity and transmission among parasites of wild primates. *International journal for parasitology* **35**, 647–57.

7. Poulin, R., Krasnov, B. R. & Mouillot, D. 2011 Host specificity in phylogenetic and geographic space. *Trends in parasitology* **27**, 355–61.

8. Nunn, C. L. & Altizer, S. M. 2005 The global mammal parasite database: An online resource for infectious disease records in wild primates. *Evolutionary Anthropology: Issues, News, and Reviews* **14**, 1–2.

9. Stephens, P. R. et al. 2017 Global Mammal Parasite Database version 2.0. *Ecology* **In Review**.

10. Gotelli, N. J. 2000 Null model analysis of species co-occurrence patterns. *Ecology* **81**, 2606–2621.

11. Cooper, N., Griffin, R., Franz, M., Omotayo, M., Nunn, C. L. & Fryxell, J. 2012 Phylogenetic host specificity and understanding parasite sharing in primates. *Ecology letters* **15**, 1370–7.

12. Cooper, N., Rodríguez, J. & Purvis, A. 2008 A common tendency for phylogenetic overdispersion in mammalian assemblages. *Proceedings of the Royal Society of London B: Biological Sciences* **275**.

13. Ezenwa, V. O. 2004 Host social behavior and parasitic infection: a multifactorial approach. *Behavioral Ecology* **15**, 446–454. (doi:10.1093/beheco/arh028)

14. Lee, K. A., Wikelski, M., Robinson, W. D., Robinson, T. R. & Klasing, K. C. 2008 Constitutive immune defences correlate with life-history variables in tropical birds. *Journal of Animal Ecology* **77**, 356–363. (doi:10.1111/j.1365-2656.2007.01347.x)

15. Poulin, R. & Forbes, M. R. 2012 Meta-analysis and research on host–parasite interactions: past and future. *Evolutionary Ecology* **26**, 1169–1185. (doi:10.1007/s10682-011-9544-0)

16. Stein, A., Gerstner, K. & Kreft, H. 2014 Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters* **17**, 866–880. (doi:10.1111/ele.12277)

17. R Core Team 2016 R: A Language and Environment for Statistical Computing.

18. Bininda-Emonds, O. R. P. et al. 2007 The delayed rise of present-day mammals. *Nature* **446**, 507–12.

19. Webb, C. O., Ackerly, D. D., McPeek, M. A. & Donoghue, M. J. 2002 Phylogenies and community ecology. *Annual Review of Ecology and Systematics* **33**, 475–505.

20. Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomberg, S. P. & Webb, C. O. 2010 Picante: {R} tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–1464.

21. Ulrich, W. & Gotelli, N. J. 2013 Pattern detection in null model analysis. *Oikos* **122**, 2–18. (doi:10.1111/j.1600-0706.2012.20325.x)

22. Redding, D. W. & Mooers, A. Ø. 2006 Incorporating Evolutionary Measures into Conservation Prioritization. *Conservation Biology* **20**, 1670–1678. (doi:10.1111/j.1523-1739.2006.00555.x)

23. Holmes, E. C. 2009 *The Evolution and Emergence of RNA Viruses*. OUP Oxford.

24. Reisen, W. K. 2013 Ecology of West Nile virus in North America. *Viruses* **5**, 2079–105.

25. Davies, T. J. & Pedersen, A. B. 2008 Phylogeny and geography predict pathogen community similarity in wild primates and humans. *Proceedings. Biological sciences / The Royal Society* **275**, 1695–701.