March 2019

Genome Size and Host Specialization in Parasites

Nicholas Theodore Ogburn
University of South Florida, Partmoose.partsquid@gmail.com

Genome Size and Host Specialization in Parasites

by

Nicholas Theodore Ogburn

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science
Department of Integrative Biology
College of Arts and Sciences
University of South Florida

Major Professor: Marc J. Lajeunesse, Ph.D.
Stephen M. Deban, Ph.D.
Kathleen M. Scott, Ph.D.

Date of Approval:
January 17, 2019

Keywords: genome size evolution, host use, parasitism

Personal Acknowledgments

I thank Marc Lajeunesse for helping to move me and this project forward through both injury and emergency, and being a tireless investigator and mentor through thick and thin. I also thank Daniel “Jake” Zydek and Merideth Krause, for being the best, most insightful colleagues that I could ask to share a workspace. Many a good idea came from our discussions, and many bad ideas were avoided due to the sober and well-considered counsel of our group. Finally, I thank the Integrative Biology Department, for being a rich repository of knowledge for any and all arcane questions that arose during this project.

Table of Contents

List of Tables

List of Figures

Abstract

In parasites, there are several examples of changes in genome size linked to a parasitic lifestyle—with some species having greatly reduced or expanded genome sizes relative to free-living non-parasitic relatives. What is unknown is whether there is correlated evolution between genome size and host specialization, and whether there is a generalizable framework in predicting genome size evolution in parasites using their genetic architecture and host use ecology. Here, I tested whether genome size of 96 eukaryotic parasites across a wide variety of taxa correlates with host specialization, quantified by the number and phylogenetic relatedness of host species they parasitize. I did not find that genome size and host specialization shared a correlated phylogenetic history; however, I did find that ectoparasites tended to have larger genomes then endoparasites, and that parasitic fungi had more host species then either animal or protozoan parasites. Although no clear trends in the evolution of genomes and host specificity were observed among parasites, my study was significantly limited by gaps in both genome size and host range availability. Future research should seek to address these gaps, as well as improve taxonomic coverage of data, so that trends in the evolution of parasite genome architecture could be adequately tested and delineated.

Introduction

There is remarkable variation in genome size among parasites (Konstantinidis and Tiedje 2004, Morrison et al. 2007, Combes 2011), and several studies have linked this variation with evolutionary transitions to parasitism and patterns of host use. In some lineages, a shift to parasitism can coincide with a change in genome size—with parasites having smaller genomes than free-living relatives (Anderson and May 1982, Poulin 1998), but there are notable exceptions (Wicke et al. 2013, Sundberg and Pulkinnen 2015). Other studies also report significant reductions of genome size among highly specialized parasites with few host species (Keeling and Slamovitz 2005)—such as among intestinal cell specialists of nematodes (Cuomo et al. 2012). Conversely, genomes can also expand when parasites add host new species to their host range (Poulin 1998, McBride 2007). These include large increases in genome size in microsporidian parasites of arthropods (Pan et al. 2013), and among wheat pathogens, where the acquisition of new genes is linked with large increases in host number and phylogenetic diversity due to the parasitism of novel host group such as grasses (Poppe 2015, Spanu et al. 2010).

Here, I explore the link between host specialization and genome size with the goals to explain (1) the vast heterogeneity in genome sizes found among parasites, and (2) provide insight on the genetic basis for why some parasites can infect a diversity of hosts. To achieve these goals, I searched the literature for reports of genome size of parasites and used phylogenetic comparative analyses (Felsenstein 1985; Harvey and

Pagel 1991) to correlate genome size with estimates of host specialization. Before describing these analyses, I first summarize general mechanisms of genome size evolution, and then place in context these mechanisms for genome size evolution in parasites. In particular, I seek to emphasize a unique tension in genome size evolution in parasites: reductions in the metabolic toolkit due to sequestering resources from hosts and expansions to the host-use toolkit due to parasitizing diverse host species.

General mechanisms for genome size evolution

Changes in genome size in parasites, as in other organisms, can occur through several well-studied mechanisms. Genome size can increase by gene duplication, polyploidization, the copying of transposable elements, and, in extreme cases, via whole-genome duplication. These increases can allow new gene functions to develop and evolve (Mayr 1963), such as when gene duplication allows for one gene copy to accumulate random mutations without resulting in deleterious effects. Polyploidization and whole-genome duplication can also lead to new gene functions, and may result in instantaneous speciation (Ohno et al. 1968, Sundberg and Pulkkinen 2015). Other mechanisms can also either increase or reduce to genome size; such as in unequal crossing over events, where misaligned homologous chromosomes undergo an unequal exchange of DNA during meiosis (Sturtevant 1925, Silver et al. 1980).

Explanations for between-species variability in genome size typically focus on mechanisms that limit or control genome size evolution (Gregory 2010). One hypothesized mechanism is the “mutational equilibrium” model, where specific rates of DNA accumulation and deletion (which vary across taxa) interact to keep genome sizes

at a species-specific equilibrium (Lozovskaya et al. 1999, Petrov 2002, Cavalier-Smith 2005). Others have argued that there is an optimal ratio between DNA content (both by number of base pair and by nuclear volume) and cell volume; although there has yet been a well-supported mechanism for why such an optimum would exist (Gregory 2001, Beaulieu et al. 2008). Finally, there is also the hypothesis that genome size may be constrained by the metabolic costs of replicating large amounts of DNA; here, although the frequency of deletion events can vary among species (Petrov 2002, Cavalier-Smith 2005), selection towards metabolic efficiency would favor genome reductions and smaller genomes (Morrison et al. 2007, Spanu et al. 2010).

Competing mechanism for genome size evolution among parasites

Exploring variation in genome size among parasites also can provide key insight into these mechanisms of genome size evolution. For example, often due to selection for metabolic efficiency, small genomes evolve when the derived metabolic toolkit is rendered unnecessary (Pagel and Johnstone 1992, Morrison et al. 2007, Tsai et al. 2017). These types of reductions are observed in many highly specialized parasites with few closely-related hosts, and often coincide with loss of functions that have been rendered obsolete, via the direct sequestering of metabolic resources from the host (Katinka et al. 2001, Lynch and Connery 2003, Keeling and Slamovitz 2005, Spanu et al. 2010, Tsai et al. 2013). A classic example is found among obligate parasitic plants—where small plastid size has repeatedly evolved due to the transition from autotrophic to heterotrophic lifestyles and the presumed relaxed selection to maintain and keep redundant photosynthetic abilities (Westwood et al. 2010, Wicke et al. 2013).

In contrast, there are also many examples of parasites increasing genome size via the addition of new DNA or genes (Poulin 1998, McBride 2007, Spanu et al. 2010, Pan et al. 2013). In some parasites, gene duplication can experience positive selection to enhance infection ability and to counter evolving host defenses (e.g., Spanu et al. 2010). Other genome expansions that can enhance infectivity and survivorship on hosts include integration of host-derived transposable elements (Pan et al. 2013), or the horizontal transfer of genes from other pathogens or symbionts (Konstantinidis and Tiedje 2004). In general, expansions appear to coincide with acquiring new hosts, and such expansions are often seen as key driving forces of speciation among parasites, either by causing speciation by whole-chromosome duplication (Poppe 2015), or by allowing populations to become reproductively isolated due to evolving under novel host conditions (Sundberg and Hulkkinen 2015, Alfsnes et al. 2017, Kuchiki et al. 2017).

These examples emphasize that selection on genomic toolkits for efficient metabolism and diverse host use may be competing mechanisms for genome size evolution in parasites—with each optimizing either reductions or expansions, respectively (Katinka et al. 2001, Keeling and Slamovitz 2005). Here I test this potential link between genome architecture and host use by testing whether genome size correlates with host specialization. I also take advantage of the broad diversity of modes of parasitism and patterns of host specialization, to separate instances where the competing tension between metabolic efficiency and host use can be imbalanced. I quantify host specialization as host number (i.e. host range or host species richness) and host phylogenetic diversity (i.e., a quantitative measure of the degree to which host species are phylogenetically related; Faith 1992). These two metrics quantify different

dimensions of host specialization that can reveal key differences among parasites and how genetic architecture can impact patterns in host use. Here, a correlation between genome size and host range could exist if underlying genetic architecture supports a large host-use toolkit for the use of many host species; however, if host species are closely-related, then a large toolkit may not be necessary as genomic tools could be applied to multiple environmental conditions (e.g., related host species are similar and represents comparable environmental conditions due to their shared evolutionary history; Futuyma and Moreno 1988, Cavalier-Smith 2005, Kikuchi et al. 2017). Finally, different modes of parasitism may experience different selection for reduced metabolic efficiency. Here, by parsing analyses among endo- and ecto-parasites, I can test the prediction that endoparasites may be more prone to genome reductions due to loss of behavioral, structural, or metabolic traits as a consequence of transitioning to a closed, resource rich host environment (Anderson and May 1982, Sundberg and Pulkkinen 2015).

Methods

Genome sizes of parasites were collected from the *Animal Genome Size Database* (Gregory 2012), Hou and Lin's (2012) large genome database, as well as queries to the *NCBI genome database* (Geer et al. 2010) with the following search terms: "eukaryote parasite", "prokaryote parasite", and "prokaryote pathogen". Genome size estimates were reported in either picograms of weighed DNA, C-value (corresponding to 1 = 1 picogram of DNA), megabases of DNA, or numbers of base pairs. To create a common currency of DNA size, all DNA content was converted into megabases (106 base pairs, Mbp) with the following conversion ratio: 1 pg = 978 Mbp for measurements in pictograms (pg), and standardized the DNA amount to what is found in diploid cells (i.e. C-value estimates from haploid cells were doubled). Further, these data originated from a large diversity of protocols for quantifying DNA amounts (i.e., biochemical analysis, flow cytometry, Fuelgen image analysis densiometry, direct genome sequencing).

Inclusion of genome size data were significantly limited by the availability of host range information for parasites. Host ranges (e.g., total number of host species recorded) for each parasite with genome data was determined using *Web of Science* and *Google Scholar* searches with the following terms: "([parasite binomial name] OR [parasite common name])", which was modified per search, AND ("host range" OR "host number" OR specialization OR "host use"). I primarily aimed to find publications/records that explicitly reported the total number of host species or preferably, a list of host species. If there were multiple records of host ranges for parasites species or taxonomic

groups, only the records derived from the lowest possible explicit taxonomic ranking were included, as host ranges based on higher taxonomic rankings tend to be coarser estimates of host specialization. However, in many instances, host ranges were only available as the number of taxonomic families used or other counts of Linnean rakings. Finally, for plant-utilizing fungi, host-use records were collected from *the Systematic Mycology and Microbiology Laboratory Fungal Database* (Farr and Rossman 2014). In total, I found 96 parasitic species with both genome size and host range data. Two of the 96 parasite species had host range data based on counts of families parasitized, and 22 were based on counts of genera parasitized.

The phylogenetic diversity of hosts within a host range was also estimated for each parasite species. Unfortunately, because host species lists are typically unavailable for a given host range, estimates of host phylogenetic diversity could only be estimated for 32% of parasites with genomes size and host range data (N = 31 of 96 parasite species). A composite phylogenetic tree of all host species (N = 1611) was constructed using *TimeTree* (Hedges et al. 2015), and this tree was used to estimate the mean phylogenetic diversity (MPD) of hosts within a parasite's host range—a large tree with all hosts species was necessary to standardized phylogenetic branch-length distances across MPDs. Not all host species were available in the *TimeTree* database; in these cases, the most closely-related species was used as a surrogate in that tree, following recommendations by Wiens (2006) and Combes (2011). To calculate MPD, I first subsetted the overall host phylogeny to only include the recorded hosts for a parasites given host range, and then calculated MPD as the mean pairwise phylogenetic branch-length distance between all hosts within the subsetted tree

(following Faith 1992, Clarke and Warwick 1999, Webb et al. 2000). These MPD are in units of millions of years ago (myo) and were calculated in *R* using the *APE* package (v. 5.0; Paradis et al. 2004).

Further, along with Linnean rankings (e.g., taxonomic Order), I classified each parasite as either obligate (i.e., could not complete life cycle without a host) or facultative (i.e., could complete life cycle without a host), and whether they were ectoparasitic eukaryotes, endoparasitic eukaryotes, or fungi. Fungi were grouped separately due to the difficulty in categorizing their parasitic and mutualistic interactions—which is dependent on the pairing of plant–fungi species (Combes 2001).

Finally, phylogenetic general least-squares regressions (PGLS) were used to test the relationship between host range (or host phylogenetic diversity) and genome size among parasites. Specifically, PGLS were performed using the *gls()* function from the *NLME* package in *R* (v. 3.1-131.1; Pinheiro et al. 2007), that assumed a maximum-likelihood optimized estimate of Pagel's λ (Pagel 1999), which is used to scale the expected covariance between pairs of species under Brownian motion evolution. This phylogenic correlation model was implemented using the *APE* package for *R* (Paradis et al. 2004). This PGLS model applied a phylogenetic hypothesis of the evolutionary history of parasites that was constructed with topologies and divergence times reported in *TimeTree* (Kumar et al. 2016; Figure 1). Phylogenetic means were conducted by excluding predictors in PGLS models (i.e., intercept only models), and conventional means excluded the phylogenetic correlational structure from the PGLS model (i.e., reducing to a standard linear regression model). Multiple regression versions of PGLS analyses, such as the phylogenetic ANOVA and ANCOVA, were implemented by

including factors (and interaction terms) to test for differences among species categorical groups such as endo- and eco-parasites, or among different approaches used to determine genome size (e.g., flow cytometry, densiometry). Ninety-five percent confidence intervals (95% CI) around PGLS regression lines were estimated following Smaers and Rohlf (2016) using a modified version of the *gls.ci*() function of the *EVOMAP* package for *R* (v. 0.0.0.9000; Smaers, J. and Mongle, C.: https://github.com/JeroenSmaers/evomap). Finally, prior to analyses, both genome sizes and host ranges were natural log transformed.

Results

Overall trends in genome size and host specialization among parasites

In total, 96 eukaryotic species of parasite had data on both genome size and host range, and the diversity of these parasites was broad and included prokaryotic intracellular pathogens to avian brood parasites (see Figure 1).

Genome sizes ranged considerably across parasites (4.11 to 7335 Mbp; Table 1), and in aggregate, strong phylogenetic non-independence in genome sizes tended to overestimate overall size among parasites (Pagel's $\lambda = 0.998$; Figure 2a). There were also significant differences in genome sizes among different modes of parasitism (conventional ANOVA: $F = 789.9$, d.f. = 2, $p < 0.001$; phylogenetic ANOVA: $F = 15.8$, d.f. = 2, $p < 0.001$; Figure 2a), where ectoparasites tended to have greater genome sizes then either endoparasites (conventional $t = -11.3$, $p < 0.001$; phylogenetic $t = -3.1$, $p = 0.0022$) or parasitic fungi (conventional $t = -12.3$, $p < 0.001$; phylogenetic $t = -2.5$, $p = 0.0144$). Further, there were also differences among approaches used to determine genome size (conventional ANOVA: $F = 517.4$, d.f. = 3, $p < 0.001$; phylogenetic ANOVA: $F = 15.8$, d.f. = 3, $p < 0.001$; Figure 2b), where direct sequencing methods tended to yield smaller genomes sizes than either biochemical analyses (conventional $t = 4.2$, $p < 0.001$; phylogenetic $t = 4.6$, $p < 0.001$), densitometry (conventional $t = 9.5$, $p < 0.001$; phylogenetic $t = 3.37$, $p = 0.0011$), or flow cytometry (conventional $t = 9.6$, $p < 0.001$; phylogenetic $t = 3.44$, $p = 0.0009$).

There was also broad host use among these parasites, both in host range, ranging from 1 to 15,400 host species, and host phylogenetic diversity, ranging from 0 (only 1 host species) to 1,594 million years in mean pairwise phylogenetic distance. Despite the significant phylogenetic signal of host range among parasites (Pagel's λ = 0.345), overall, this signal was less important in modelling phylogenetic non-independence among host specialization then genome sizes. However, average host ranges differed among modes of parasitism (conventional ANOVA: F = 49.7, d.f. = 3, p < 0.001; phylogenetic ANOVA: F = 8.5, d.f. = 3, p < 0.001; Figure 2c), where parasitic fungi had greater host ranges then endoparasites (conventional t = -3.8, p < 0.001; phylogenetic t = -3.1, p = 0.0024), but not ectoparasites (conventional t = -4.6, p < 0.001; phylogenetic t = -1.6, p = 0.1093).

Correlated history between genome size and host specialization

It is inconclusive whether genome size and host range share a correlated evolutionary history among parasites (Pagel's λ = 0.398; PGLS slope t = 1.838, p = 0.0692; N = 96; Figure 3a), despite a slight trend of parasites with many host species having larger genomes (intercept = 0.41 ± 0.98 SE, slope = 0.35 ± 0.19 SE). As a sensitivity analysis, parasitic taxa with host ranges of one (i.e., very specialized taxa) were excluded; however, there was no overall effect in the removal of specialist parasites with no variability in host ranges (Pagel's λ = 0.427; PGLS slope t = 1.576, p = 0.1199; N = 66; Figure 3b). In addition, a phylogenetic ANCOVA did not find evidence for differences in the relationship between genome size and host range among endoparasites, ectoparasites, and fungi (Figure 3e; parasitic mode: F = 7.11, d.f. = 2, p = 0.0014;

genome size: F = 1.66, d.f. = 1, p = 0.2012; genome size × parasitic mode: F = 2.07, d.f. = 2, p = 0.1325). Excluding the non-significant interaction term did not change these results (parasitic mode: F = 5.07, d.f. = 2, p = 0.0081; genome size: F = 3.16, d.f. = 1, p = 0.0788). Separately within modes of parasitism, there was no relationship between genome size and host range for endoparasites (Figure 3e; intercept = -3.41 ± 3.0 SE, slope = 0.77 ± 0.43 SE; ANCOVA PGLS slope t = 1.79, p = 0.0766), ectoparasites (intercept = 1.06 ± 1.04 SE, slope = 0.03 ± 0.25 SE; ANCOVA PGLS slope t = 0.15, p = 0.8777), or fungi (intercept = -0.57 ± 2.06 SE, slope = 1.06 ± 0.54 SE; ANCOVA PGLS slope t = 1.963, p = 0.0528).

There is a shared correlated evolutionary history between genome size and the average phylogenetic diversity of hosts (PGLS slope t = -4.98, p < 0.001; N = 31; Figure 3c)—where parasites with smaller genomes parasitized more distantly related host species. However, the residual variability among genome size and host phylogenetic diversity is negatively phylogenetically correlated (Pagel's λ = -0.233), indicating here 0 that closely related parasites tended to be more different then distantly related parasites. Finally, there was no correlated evolutionary history between host range and host phylogenetic distance (Figure 3d; Pagel's λ = 0.667; PGLS slope t = -1.29, p = 0.2085; N = 31).

Discussion

I did not find a clear correlation between genome size and host range among parasites. However, I did find marginal evidence that larger genomes can be associated with parasites having a narrow group of phylogenetic-related hosts, and that endoparasites can have smaller genomes than ectoparasites. Finally, I must caution that my study had poor taxonomic coverage, where for example all but one of the endoparasites were nematodes, and below I outline how this poor taxonomic coverage has hindered progress in testing the link between host specificity and the genomes of parasites.

There were several limitations to my comparative study. First, the number of parasite species compared is relatively small and composed of few closely related groups (e.g., *Amblyomma* ticks; see Figure 1). Higher representation among distantly related parasite groups is key to testing predictors of parasite genome evolution, as it would provide more convincing evidence for strong convergent evolution for (1) when taxa transition to a parasitic life-style, and (2) the common environmental conditions experienced when either few or many host species are parasitized (Lajeunesse and Forbes 2002). Given the several hundreds of independent transitions from free-living to parasitism recorded across a broad diversity of taxa (Poulin and Randhawa 2015), the opportunity to test trends genome size evolution is high; however, genome size estimates, as well as host range data, are scant and often poor quality.

For example, genome size estimates were only available for a limited number of species (see Table 1), and of those available, many could not be included in my study

due to not finding matching host range data. Further, when both genome sizes and host range data were available, the quality and precision varied. It is well-documented that host ranges tend to be incomplete due to the limited time and resources needed to build exhaustive host use records. Consequently, host ranges tend to be underestimated (Anderson and May 1982, Dallas et al. 2017) and prone to taxonomic sampling bias (Colwell and Futuyma 1971, Chacoff et al. 2012). Poor taxonomic sampling, in turn, negatively impacts estimates of host phylogenetic diversity—as the host species composition of small host ranges will over- or under-estimate the relatedness of hosts parasitized (Gilbert et al. 2012). The high phylogenetic signal of host ranges among related parasites also negatively impacts comparative analyses, since a larger and more phylogenetically independent groups are needed to compensate the increased Type I and Type II statistical errors associated with highly phylogenetically conserved traits (see Lajeunesse and Fox 2015). These challenges with host range and host phylogenetic diversity data are unfortunate, as it is agreed that they can be high-quality surrogates for estimating phenotypic differences among parasites that determine host-use success (Gomez et al. 2010).

There is also varying reliability of genome size estimation (see Figure 2b). For example, I found that direct sequencing of DNA tended to yield smaller genome sizes than biochemical analysis, densiometry, or flow cytometry. It is understood that data originating from flow cytometry is limited given that it is a comparison to a “standard” known genome size, and does not produce quantification of DNA independent of comparison to known quantities (Arumuganathan and Earle 1991, Doležel and Bartoš 2005). Consequently, flow cytometry data are usually coarse and relatively low

precision. Densitometry is also limited by poor repeatability and high heterogeneity in results (Križman et al. 2009). Finally, too few species in my study had genome sizes estimated via biochemical analysis (N = 2; Besansky 1992); therefore, no meaningful conclusions could be drawn from these data. One suggestion to improve genome size estimates is to take the average of multiple published values, and then use these in comparative analyses (Nayfach and Pollard 2015); however, again many of the parasite species included in my study are highly understudied and have only a single estimate available.

I also found that endoparasites tended to have smaller genomes then ectoparasites. The larger mean genome size among ectoparasites can be explained by the high number of complex traits needed that are independent of host parasitism but useful to live outside their hosts—such as predation-avoidance traits, flight, mate finding, and temperature heterogeneity. Likewise, smaller genomes are also consistent with the prediction that endoparasites experience relaxed selection for gene functions rendered obsolete (Anderson and May 1982, Hafner and Nadler 1988, Combes 2001, Sundberg and Pulkkinnen 2015). However, it is important to note that my ectoparasite data were composed entirely of multicellular organisms; multicellular species tend to have increased genome sizes compared to unicellular organisms (Lynch and Connery 2003).

The lack of ample comparative data, along with the poor quality of what is available, can make it difficult to explain why I found a null relationship between genome size and host specificity among parasites. The goal was to use host specialization to help identify the signature of selection on genome size evolution—that is, to help

distinguishing instances with relaxed selection and genome reductions (i.e., functional loss in tool-box), or adaptive evolution in broad host use and genome expansion (i.e., functional gains in tool-box). However, other factors should also be explored to help explain variation in genome size of parasites and how host specialization can moderate genome evolution. First, there are ample examples of genome size reduction related to an evolutionary shift to parasitism (Anderson and May 1982, Poulin 1998), though this does not necessarily provide answers as to why such transitions impact the genetic architecture of organisms. The time since the evolution of parasitism (evolutionary origin) may be one factor influencing genome size and/or host range; for instance, platyhelmenthes developed parasitism extremely early in their evolutionary history, and the parasites in this phylum display remarkably narrow host ranges compared to other parasites (Littlewood et al. 1999). Second, I recommend examination of another potential predictor a parasite species – geographic distribution. This reflects overall opportunity to parasitism multiple and diverse host species, since geography scales positively with the richness and diversity of taxa (Bottini et al. 2000). Thirdly, the evolutionary consequences of host specialization are not uniform across parasites; some aphid species experience significant performance loss on an established host species if the parasite expands its host range (Straub et al. 2011), while others do not find such tradeoffs in host use (Drovetski et al. 2014). These studies suggest that genetic trade-offs in host use, due to specializing on few host species, are uncommon (Futuyma and Moreno 1986), and that underlying genetic framework regulating host use may be less important then ecological opportunities to parasitize (Zydek and Lajeunesse, in prep). Future research should aim to integrate these potential predictors

of genome size evolution among parasites—especially when data on genome size of a richer diversity of organisms becomes more prevalent in the literature.

References

Alfsnes, K., Leinaas, H. P., and Hessen, D. O. 2017. Genome size in arthropods; different roles of phylogeny, habitat and life history in insects and crustaceans. Ecology and Evolution 7: 5939-5947.

Anderson, R. M., and R. May. 1982. Coevolution of hosts and parasites. Parasitology 85: 411-426.

Arumuganathan, K. and E. D. Earle. 1991. Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter 9: 208-218.

Beaulieu, J. M., I. J. Leitch, S. Patel, A. Pendharkar, and C. A. Knight. 2008. Genome size is a strong predictor of cell size and stomatal density in angiosperms. New Phytologist 179: 975-986.

Besansky, N.J. and J. R. Powell. 1992. Reassociation kinetics of Anopheles gambiae (Diptera: Culicidae) DNA. Journal of Medical Entomology 29:125-128.

Bottini, M.C.J., E. J. Greizerstein, M. B. Aulicino and L. Poggio. 2000. Relationships among genome size, environmental conditions and geographical distribution in natural populations of NW Patagonian species of *Berberis* (Berberidaceae). Annals of Botany 86: 565-573.

Cavalier-Smith, T. (2005). Economy, speed and size matter: evolutionary forces driving nuclear genome miniaturization and expansion. Annals of Botany 95: 147-175.

Clarke, K. R., and R. M. Warwick. 1999. The taxonomic distinctness measure of biodiversity: weighting of step lengths between hierarchical levels. Marine Ecology Progress Series 184: 21-29.

Combes, C. 2001. Parasitism: The Ecology and Evolution of Intimate Interactions. University of Chicago Press, Chicago, Illinois, USA.

Cuomo, C. A., C. A. Desjardins, M. A. Bakowski, J. Goldberg, A. T. Ma, J. J. Becnel, E. S. Didier, L. Fan, D. I. Hieman, J. Z. Levin, S. Young, Q. Zhang, and E. R. Troemel. 2012. Microsporidian genome analysis reveals evolutionary strategies for obligate intracellular growth. Genome Research 12: 2478-2488.

Dallas, T., S. Huang, C. Nunn, A. W. Park and J. M. Drake. 2017. Estimating parasite host range. Proceedings of the Royal Society B 284: 1250.

Doležel, J. and J. A. N. Bartoš. 2005. Plant DNA flow cytometry and estimation of nuclear genome size. Annals of Botany 95: 99-110.

Drovetski, S.V., S. A. Aghayan, V. A. Mata, R. J. Lopes, N. A. Mode, J. A. Harvey and G. Voelker. 2014. Does the niche breadth or trade-off hypothesis explain the abundance–occupancy relationship in avian Haemosporidia?. Molecular Ecology 23: 3322-3329.

Faith, D. P. 1992. Conservation evaluation and phylogenetic diversity. Biological Conservation 61: 1-10.

Farr, D.F., and Rossman, A.Y. Fungal Databases, U.S. National Fungus Collections, ARS, USDA. Retrieved November 14, 2017, from https://nt.ars-grin.gov/fungaldatabases/

Futuyma, D.J. and G. Moreno. 1988. The evolution of ecological specialization. Annual Review of Ecology and Systematics 2: 207-233.

Geer, L.Y., A. Marchler-Bauer, R. C. Geer, L. Han, J. He, S. He, C. Liu, W. Shi, and S. H. Bryant. 2010. The NCBI BioSystems database. Retrieved November 14, 2017 from https://www.ncbi.nlm.nih.gov/

Gregory, T.R. 2001. Animal genome size database. Retrieved August 5, 2015 from http://www.genomesize.com/

Hafner, M., and S.A. Nadler. 1988. Phylogenetic trees support the coevolution of parasites and their hosts. Nature 332: 258-259.

Hedges, S. B., J. Marin, M. Suleski, M. Paymer, and S. Kumar. 2015. Tree of Life Reveals Clock-Like Speciation and Diversification. Molecular Biological Evolution 32: 835-845.

Hou, Y. and S. Lin. 2009. Distinct gene number-genome size relationships for eukaryotes and non-eukaryotes: gene content estimation for dinoflagellate genomes. PLoS One 4: e6978.

Katinka, M. D., S. Duprat, E. Cornillot, G. Méténier, F. Thomarat, G. Prensier, V. Barbe, E. Peyretaillade, P. Brotier, P. Wincker, F. Delbac, H. el Alaoui, P. Peyret, W. Saurin, M. Gouy, J. Weissenbach, and C.P. Vivarès. 2001. Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi. Nature 414: 450-453.

Keeling, P. J., and C.H. Slamovits. 2005. Causes and effects of nuclear genome reduction. Current Opinion in Genetics & Development 15: 601-608.

Kikuchi, T., Eves-van den Akker, S., and Jones, J. T. 2017. Genome Evolution of Plant-Parasitic Nematodes. Annual Review of Phytopathology 55: 333-354.

Konstantinidis, K. T., and J.M Tiedje. 2004. Trends between gene content and genome size in prokaryotic species with larger genomes. Proceedings of the National Academy of Sciences of the United States of America 101: 3160-3165.

Križman, M., Jakše, J., Prošek, M., Baričevič, D. and Javornik, B. 2009. Densitometric DNA analysis in agarose electrophoretic gels. JPC-Journal of Planar Chromatography-Modern TLC 22: 167-170.

Lapp, N.A. and A.C. Triantaphyllou 1972. Relative DNA content and chromosomal relationships of some *Meloidogyne*, *Heterodera*, and *Meloidodera* spp. (Nematoda: Heteroderidae). Journal of Nematology 4: 287-291.

Littlewood, D.T.J., K. Rohde, R. A. Bray and E. A. Herniou. 1999. Phylogeny of the Platyhelminthes and the evolution of parasitism. Biological Journal of the Linnean Society 68: 257-287.

Lozovskaya, E. R., D. Nurminsky, D. A. Petrov, and D. L. Hartl. 1999. Genome size as a mutation-selection-drift process. Genes and Genetic Systems 74: 201-207.

Lynch, M. and J. S. Conery. 2003. The origins of genome complexity. Science 302: 1401-1404.

McBride, C. S. 2007. Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia*. Proceedings of the National Academy of Sciences of the United States of America 104: 4996–5001.

Morrison, H.G., A.G. McArthur, F.D. Gillin, S.B. Aley, R.D .Adam, G.J. Olsen, A.A. Best, W.Z. Cande, F. Chen, M.J. Cipriano and B.J. Davids. 2007. Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. Science 317: 1921-1926.

Ohno, S., U. Wolf, and N. B. Atkin. 1968. Evolution from fish to mammals by gene duplication. Hereditas 59: 169-187.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20: 289-290.

Pagel, Mark, and R. A. Johnstone. 1992. Variation across species in the size of the nuclear genome supports the junk-DNA explanation for the C-value paradox. Proceedings of the Royal Society of London B 249: 119-124.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature 401: 877-884

Pan, Guoqing, J. Xu, T. Li, Q. Xia, S Liu, G. Zhang, S. Li, C. Li, H. Liu, L. Yang and T. Liu. 2013. Comparative genomics of parasitic silkworm microsporidia reveal an association between genome expansion and host adaptation. BMC Genomics 14: 186.

Petrov, D.A., 2002. Mutational equilibrium model of genome size evolution. Theoretical Population Biology 61: 531-544.

Poppe, S., L. Dorscheimer, P. Happel, and E. H. Stukenbrock. 2015. Rapidly Evolving Genes Are Key Players in Host Specialization and Virulence of the Fungal Wheat Pathogen Zymoseptoria tritici (Mycosphaerella graminicola). PLOS Pathogens 11: e1005055.

Poulin, R. 1998. Evolutionary Ecology of Parasites. Chapman and Hall, London

Pinheiro, José, D. Bates, S. DebRoy, and D. Sarkar. 2014. nlme: linear and nonlinear mixed effects models. R package version 3.1-117. R foundation for statistical computing, Vienna.

Silver, L. M., M. White, and K. Artzt. 1980. Evidence for unequal crossing over within the mouse T/t complex. Proceedings of the National Academy of Sciences 77: 6077-6080.

Spanu, P. D., J. C. Abbott J. Amselem, T. A. Burgis, D. M. Soanes, K. Stüber, E.V.L. van Themaat, J.K.M. Brown, S.A. Butcher, S.J. Gurr, M.C. Lebrun, C.J. Ridout, P. Shulze-Lefurt, N.J. Talbot, N. Ahmadinejad, C. Amertz, G.R. Barton, M. Benjidia, and R. Reinhardt. 2010. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science 330: 1543-1546.

Straub, C.S., Ives, A.R. and Gratton, C. 2011. Evidence for a trade-off between host-range breadth and host-use efficiency in aphid parasitoids. American Naturalist 177: 389-395.

Sturtevant, A. H. 1925. The effects of unequal crossing over at the Bar locus in *Drosophila*. Genetics 10: 117.

Sundberg, L. R., and Pulkkinen, K. 2015. Genome size evolution in macroparasites. International Journal for Parasitology 45: 285-288.

Tsai, I. J., M. Zarowiecki, N. Holroyd, A. Garciarrubio, A. Sanchez-Flores, K.L Brooks, A. Tracey, R.J. Bobes, G. Fragoso, E. Sciutto, M. Aslett, H. Beasley, H.M. Bennett, J. Cai, F. Camicia, R. Clark, M. Cucher, N. De Silva, T. A. Day, P. Deplazes, K. Estrada, C. Fernandez, P.W.H. Holland, J. Hou, S. Hu, T. Huckvale, S.S. Hung, L. Kamenetzky, J.A. Keane, F. Kiss, U. Kozioi, O. Lampert, K. Liu, X. Luo, Y. Luo, N. Macchiaroli, S. Nichol, J. Paps, J. Parkinson, N. Pouchkina-Stantcheva, N. Riddiford, M. Rozenzvit, G. Salinas, J.D. Wasmuth, M. Zamanian, Y. Zheng, X. Cai, X. Soberón, P.D. Olson, J.P. Laciette, K. Brehm,

and M. Berriman. 2013. The genomes of four tapeworm species reveal adaptations to parasitism. Nature 496: 57-63.

Vilhar, B., J. Greilhuber, J.D. Koce, E.M. Temsch, and M. Dermastia. 2001. Plant genome size measurement with DNA image cytometry. Annals of Botany 87: 719-728.

Webb, C. O., D. D. Ackerly, M. A. McPeek, and M. J. Donoghue. Phylogenies and community ecology. Annual Review of Ecology and Systematics 33: 475-505.

Westwood, J. H., J. I. Yoder, and M. P. Tinko. 2010. The evolution of parasitism in plants. Trends in Plant Science 15: 227-235.

Wicke, S., K. F. Müller, C. W. de Pamphilis, D. Quandt, N. J. Wickett, Y. Zhang, S. S. Renner, and G. M. Schneeweiss. 2013. Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. Plant Cell 25: 3711-3725.

Wiens, J. J. 2006. Missing data and the design of phylogenetic analyses. Journal of Biomedical Informatics 39: 34-42.

Woolhouse, M.E., L.H. Taylor, and D.T. Haydon. 2001. Population biology of multihost pathogens. Science 292: 1109-1112.

Webb, C., D. Ackerly, M. McPeek, and M. Donoghue. 2002. Phylogenies and community ecology. Annual Review of Ecology and Systematics 33:475-505.

Genome size and host range references

Adams, A. S., F. O. Aylward, S. M. Adams, N. Erbilgin, B. H. Aukema, C. R. Currie, G. Suen, and K. F. Raffa. 2013 Mountain pine beetles colonizing historical and naive host trees are associated with a bacterial community highly enriched in genes contributing to terpene metabolism. Applied and environmental microbiology: AEM-00068.

Aksoy, S., M. Berriman, N. Hall, M. Hattori, W. Hide, and M.J. Lehane. 2005. A case for a *Glossina* genome project. Trends in Parasitology 21: 107-111.

Besansky, N.J. and J.R. Powell. 1992. Reassociation kinetics of *Anopheles gambiae* (Diptera: Culicidae) DNA. Journal of Medical Entomology 29: 125-128.

Clausen, P-H., F. H. Leendertz, A. Blankenburg, U. Tietjen, D. Mehlitz, I. Sidibe, and B. Bauer. 1999 A drug incubation glossina infectivity test (DIGIT) to assess the susceptibility of Trypanosoma congolense bloodstream forms to trypanocidal drugs (Xenodiagnosis). Acta tropica 72: 111-117.

Dekker, T., W. Takken, and A. H. Braks Marieta. 2001 Innate preference for host-odor blends modulates degree of anthropophagy of Anopheles gambiae sensu lato (Diptera: Culicidae). Journal of Medical Entomology 38: 868-871.

Farr, D.F., and Rossman, A.Y. 2006. Fungal Databases, U.S. National Fungus Collections, ARS, USDA. Retrieved December 3, 2016, from https://nt.ars-grin.gov/fungaldatabases/

Favret, C., and D. J. Voegtlin. 2004 Speciation by host-switching in pinyon Cinara (Insecta: Hemiptera: Aphididae). Molecular phylogenetics and evolution 32, no. 1: 139-151.

Finston, T.L., P.D.N. Hebert, and R.B. Foottit. 1995. Genome size variation in aphids. Insect Biochemistry and Molecular Biology 25: 189-196.

Foottit, R. G., S. E. Halbert, G. L. Miller, E. Maw, and L. M. Russell. 2006. Adventive aphids (Hemiptera: Aphididae) of America North of Mexico. Proceedings of the Entomological Society of Washington 108: 583.

Geraci, N.S., S. J. Johnston, P. J. Robinson, S. K. Wikel, and C. A. Hill. 2007. Variation in genome size of argasid and ixodid ticks. Insect Biochemistry and Molecular Biology 37: 399-408

Ghedin, E., S. Wang, D. Spiro, E. Caler, Q. Zhao, J. Crabtree, J. E. Allen, A. L. Delcher, D. B. Guilliano, D. Miranda-Saaverdra, S. V. Angiuoli, T. Creasy, P. Amedeo, B. Haas, N. M. El-Sayed, J. R. Wortman, T. Feldblyum, L. Talion, M Schatz, M. Shumway, H. Koo, S. L. Salzberg, S. Schobel, M. Pertea, M. Pop, Q. White, G. J. Barton, C. K. Carlow, M. J. Crawford, J. Daub, M. W. dimmic, C. F. Estes, J. M. Foster, M. Ganatra, W. F. Greggory, N. M. Johnson, J. Jin, R. Komuniecki, I Korf, S. Kuman, S. Laney, B. W. Li, T. H. Lindblom, S. Sustigman, D. Ma, C. V. Maina, D. M. Martin, J. P. McCarter, L. McReynolds, M. Mitreva, Nutman, T. B., J. Parkinson, J. M. Peregrin-Alvarez, C. Poole, Q. Ren, L. Saunders, A. E. Slunder, K. Smith, M. Slanke, T. R. Unnasch, J. Ware, A.D. Wei, G. Weil, D. J. Williams, Y. Zhang, S. A. Williams, C. Fraser-Liggett, B. Stalke, M. L, Blaxter and

A. L. Scott. 2007. Draft genome of the filarial nematode parasite *Brugial malayi*. Science 317: 1756-60.

Goody, J.B. 1957. Laboratory method for work with plant and soil nematodes. Ministry of Agric., Fisheries and Food: Tech. Bull. 2, London, England, pp. 44

Gregory, T.R.,P. Nathwani, T.R. Bonnett, and D.P.W. Huber. 2013. Sizing up arthropod genomes: an evaluation of the impact of environmental variation on genome size estimates by flow cytometry and the use of qPCR as a method of estimation. Genome 56: 505-510.

Hale, F. A. White Pine Aphid. Agricultural Extension Service, University of Tennessee.

Hirshman, H. and A. C. Triantaphyllou. 1967. Helicotylenchus dihystera. Nematologica 13: 558-574

Jones, F. G. W. 1956. Soil Population of Beet Eelworm (Heterodera Schactii schm)in relation to cropping: II. Microplot and field plot results." Annals of applied Biology 44: 25-56.

Riggs, R. D., M. L. Hamblen, and L. Rakes. 1981 Infra-species variation in reactions to hosts in Heterodera glycines populations." Journal of Nematology 13: 179.

Leroy, S., S. Bouamer, S. Morand, and M. Fargette. 2007. Genome size of plant-parasitic nematodes. Nematology 9: 449-450.

Johnston, J.S., L.D. Ross, L. Beani, D.P. Hughes, and J. Kathirithamby. 2004. Tiny genomes and endoreduplication in Strepsiptera. Insect Molecular Biology 13: 581-585

Ma, R.Z., W.C. Black, and J.C. Reese. 1992. Genome size and organization in an aphid (*Schizaphis graminum*). Journal of Insect Physiology 38: 161-165.

Manicardi, C.C., E. Galli, A. Malavasi, and A.M. Bonvinci Pogliai. 1995. DNA content in the nurse cell nuclei of viviparous and oviparous females of *Megoura viciae* (Homoptera, Aphididae). Invertebrate Reproduction and Development 28: 1-6.

Mercer, C. F., and D. R. Woodfield. 1986. A survey of root knot and clover cyst nematodes in dry hill country. New Zealand journal of agricultural research 29: 129-135.

Lapp, N.A. and A.C. Triantaphyllou. 1972. Relative DNA content and chromosomal relationships of some *Meloidogyne*, *Heterodera*, and *Meloidodera* spp. (Nematoda: Heteroderidae). Journal of Nematology 4: 287-291.

Ullman, A. J., G. M. Lima, F. D. Guerrero, J. Piesman, and W. C. Black. 2005. Genome size and organization in the blacklegged tick, *Ixodes scapularis* and the Southern cattle tick, *Boophilus microplus*. Insect Molecular Biology 14: 217-222.

Ruehle, J. L. 1968. Plant-parasitic nematodes associated with southern hardwood and coniferous forest trees. Plant Dis. Rep. 52:837-839.

Solomon, J.D. 1986 Early Impact and Control of Aphid (Chaitophorus populicola Thomas) Infestations on Young Cottonwood Plantations in the Mississippi Delta, Report No. SO-326. USDA Forest Service, Southern Forest Experiment Station, New Orleans, Louisiana

Steele, A. E., H. Toxopeus, and W. Heijbroek. 1983. Susceptibility of plant selections to Heterodera schachtii and a race of H. trifolii parasitic on sugarbeet in The Netherlands. Journal of nematology 15: 281.

Straub, C.S., A. R. Ives and C. Gratton. 2011. Evidence for a trade-off between host-range breadth and host-use efficiency in aphid parasitoids. American Naturalist 177: 389-395.

Thindwa, H.P., G.L. Teetes, and J.S. Johnston. 1994. Greenbug DNA content. Southwestern Entomologist 19: 371-378.

Triantaphyllou, A. C. 1973. Gametogenesis and reproduction of Meloidogyne graminis and M. ottersoni (Nematoda: Heteroderidae). Journal of nematology 5: 84.

Waiswa, C., K. Picozzi, E. Katunguka-Rwakishaya, W. Olaho-Mukani, R. A. Musoke, and S. C. Welburn. 2006. Glossina fuscipes fuscipes in the trypanosomiasis endemic areas of south eastern Uganda: apparent density, trypanosome infection rates and host feeding preferences. Acta tropica 99: 23-29.

Whittington, I.D., B. W. Cribb, T.E. Hamwood and J. A. Halliday. 2000. Host-specificity of monogenean (platyhelminth) parasites: a role for anterior adhesive areas? International Journal for Parasitology 30: 305-320.

Table 1. Genome size and host range information of the 96 species of parasites used in this comparative analysis. Host ranges are based on host species counts; however, when based on the number of host genera, these are indicated by *, and when based on counts of Families by **.(see Figure 1 for phylogenetic relationships).

Species	Kingdom	Genome size (Mbp)	Genome size source	Host range	Host range source
Amblyomma americanum	Animalia	3061.14	Geraci et al. (2007)	39	Keirans (1998)
Amblyomma cajenesse	Animalia	2767.74	Geraci et al. (2007)	6	Labruna (2011), Childs (2003), Oliviera (2003), Bishop (1945)
Amblyomma maculatum	Animalia	3158.94	Geraci et al. (2007)	65	Teel et al. (2010)
Rhipicephalus microplus	Animalia	7335	Ullman (2005)	1*	Hoogstraal (1982)
Ixodes scapularis	Animalia	2484.12	Geraci et al. (2007)	125	Keirans (1996)
Ixodes pacificus	Animalia	2107.59	Geraci et al. (2007)	107	Castro (2007)
Adelges cooleyi	Animalia	371.64	Finston (1995)	1*	Annand (1928)
Pineus similis	Animalia	361.86	Finston (1995)	1*	Cumming (1962)
Amphorophora agathonica	Animalia	381.42	Finston (1995)	1*	Lightle (2013)
Mogoura viceae	Animalia	469.44	Manicardi (1995)	2*	Blackman (2006)
Gysoaphis oestlundi	Animalia	870.42	Finston (1995)	1*	Finston (1995)
Macrosiphum californicum	Animalia	498.78	Finston (1995)	1*	Finston (1995)
Macrosiphum euphorbiae	Animalia	391.2	Finston (1995)	10	Finston (1995)
Macrosiphum roseae	Animalia	508.56	Finston (1995)	1*	Foottit et al.(2006) ,Finston (1995)
Schizaphis graminum	Animalia	485.74	Thindwa (1994), Ma (1992), Finston (1995)	15*	Finston (1995)
Aphis pomi	Animalia	479.22	Finston (1995)	5*	Finston (1995)
Tamalia coweni	Animalia	616.14	Finston (1995)	1*	Finston (1995)
Eriosoma americanum	Animalia	322.74	Finston (1995)	68	Finston (1995)
Chaitophorus populicola	Animalia	469.44	Finston (1995)	2*	Solomon (1986)
Cinara strobi	Animalia	694.38	Finston (1995)	1	Hale (2008)
Anopheles gambiae	Animalia	264.06	Besansky (1992)	1	Dekker (2001)
Glossina fuscipes	Animalia	489	Aksoy (2005)	13	Waiswa et al. (2003), Clausen et al. (1998)
Caenocholax fenyesi	Animalia	107.58	Johnston (2004)	2	Kathirithamby (2002)
Xenos vesparum	Animalia	127.14	Johnston (2004)	1	Dapporto (2007)
Dendroctonus ponderosa	Animalia	205.38	Gregory (2013)	1*	Adams et al. (2013)
Heterodera betula	Animalia	97.8	Lapp (1972)	6	Hirschman and Triantaphyllou (1969), Riggs et al. 1992

Heterodera glycines	Animalia	58.68	Lapp (1972)	6	Riggs et al. (1992)
Heterodera schacti	Animalia	58.68	Lapp (1972)	16	Jones (1956), Steele et al. (1983)
Heterodera trifolii	Animalia	97.8	Lapp (1972)	15	Mercer and Woodfield (1986), Steele et al. (1983)
Meliodogyne floridensis	Animalia	58.68	Lapp (1972)	15	Goodey (1956), Ruelhe (1968)
Meliodogyne graminis	Animalia	29.34	Lapp (1972)	5*	Goodey (1956),
Meliodogyne ottersoni	Animalia	29.34	Lapp (1972)	1	Triantaphyllou (1973)
Hemicycliophoras conida	Animalia	19.56	Leroy (2007)	1	Goodey (1956)
Brugia malayi	Animalia	93.659	Ghedin (2007)	3	Edeson (1964), Laing (1960)
Ascaris lumbricoides	Animalia	312.96	Moritz (1976)	1	King et al. (2005)
Ascaris suum	Animalia	244.5	Searcy (1970)	2	Nejsum et al. (2005)
Parascaris univalens	Animalia	2445	Nigon (1955)	1*	Bullini et al. (1986)
Placobdella ornate	Animalia	577.02	Genomesize (2014)	1**	Ryan and Lambert (2005)
Placobdella parasitica	Animalia	616.14	Genomesize (2014)	1**	Ryan and Lambert (2005)
Ophidonais serpentina	Animalia	1496.34	Gregory (2002)	1	Aardalan et al. (2011)
Molothrus ater	Animalia	1378.98	Andrews (2009)	200	Friedman (1963)
Molothrus bonariensis	Animalia	1476.78	Wright (2014)	80	Friedman (1963)
Vidua paradisaea	Animalia	1154.04	Wright (2014)	2	Klein (1998)
Vidua macroura	Animalia	1056.24	Wright (2014)	9	Sorenson (2004)
Chrysococcyx capris	Animalia	1418.1	Wright (2014)	2*	Rothstein (1971)
Clamator jacobinus	Animalia	1369.2	Wright (2014)	2*	Osmaston (1916)
Candida albicans	Fungi	28.5407	Bartelli(2013)	1	Hise et al. (2009)
Eremothecium gossypii	Fungi	8.7424	Hou and Lin (2009)	6	Farr and Rossman (2006)
Aspergillus flavus	Fungi	36.89	Nierman (2015)	120	Farr and Rossman (2006)
Penicillium chrysogenium	Fungi	32.22	Bohm (2015)	52	Farr and Rossman (2006)
Cochliobolus heterostrophus	Fungi	36.46	Geer (2005)	53	Farr and Rossman (2006)
Pyrenophora tritici	Fungi	38	PRJNA29813, PRJNA18815	45	Farr and Rossman (2006)
Alternaria brassicola	Fungi	29.54	PRJNA34523	35	Farr and Rossman (2006)
Phaeosphaeria nodorum	Fungi	37.21	PRJNA21049, PRJNA13754	68	Farr and Rossman (2006)
Cladosporum fulvum	Fungi	29.3849	Kuznetzov and Ivanovsky (2013)	1*	Farr and Rossman (2006)
Zymoseptoria tritici	Fungi	39.69	PRJNA19047	24	Farr and Rossman (2006)
Blumeria graminis	Fungi	118.73	PRJNA28821	436	Farr and Rossman (2006)
Botryotinia fuckeliana	Fungi	42.74	PRJNA15632, PRJNA20061	1049	Farr and Rossman (2006)
Marssonina brunnea	Fungi	51.95	PRJNA215720	1*	Han et al. (2000)
Giberella zeae	Fungi	36.4873	Gale (2005)	95	Farr and Rossman (2006)
Trichoderma atrovirde	Fungi	36.14	PRJNA19867	5	Farr and Rossman (2006)
Trichoderma virens	Fungi	39.02	PRJNA19983	9	Farr and Rossman (2006)
Claviceps purp	Fungi	32.09	PRJEA76493	406	Farr and Rossman (2006)
Claviceps paspali	Fungi	28.97	PRJNA51623	34	Farr and Rossman (2006)
Epichloe glyceriae	Fungi	46.72	PRJNA67247	1	Pann et al. (2003)
Nectria haematococca	Fungi	51.29	PRJNA51499, PRJNA16586	112	Farr and Rossman (2006)
Fusarium pseudograminarium	Fungi	36.93	PRJNA66583	10	Farr and Rossman (2006)
Fusarium fujikuroi	Fungi	43.81	PRJNA171493	45	Farr and Rossman (2006)
Chaetomium globosum	Fungi	34.89	PRJNA16821, PRJNA12795	91	Farr and Rossman (2006)
Grossmania claviger	Fungi	29.79	PRJNA39837	5**	Farr and Rossman (2006)
Magnaporthe grisea	Fungi	41.4959	Hou and Lin (2009)	120	Farr and Rossman (2006)
Gaeumannomyces graminis	Fungi	43.77	PRJNA37931	60	Farr and Rossman (2006)
Ustilago maydis	Fungi	19.719	PRJNA1446	9	Farr and Rossman (2006)
Tremella mesenterica	Fungi	28.64	PRJNA32829	50	Farr and Rossman (2006)
Fomitopora mediterranea	Fungi	63.35	PRJNA56107	2	Farr and Rossman (2006)
Postia placenta	Fungi	90.89	PRJNA19789	7	Farr and Rossman (2006)

Phanerochaete crysosporum	Fungi	29.84	PRJNA135	8	Farr and Rossman (2006)
Dichomitus squalens	Fungi	42.75	PRJNA53511	26	Farr and Rossman (2006)
Moniliophthora perniciosa	Fungi	26.66	Rincones et al. (2008)	19	Farr and Rossman (2006)
Heterobasidion annosum	Fungi	33.65	PRJNA46703	5	Farr and Rossman (2006)
Puccinia graminis	Fungi	88.72	PRJNA66375	560	Farr and Rossman (2006)
Puccinia striiformis	Fungi	130.49	PRJNA176877	153	Farr and Rossman (2006)
Puccinia triticina	Fungi	162.95	PRJNA36323	30	Farr and Rossman (2006)
Mixia	Fungi	13.39	Nishida et al. (2011)	1	Farr and Rossman (2006)
Hamiltosporidium tvaerminnensis	Fungi	13.27	Corradi et al. (2009)	1	Haang(2011)
Nematocida parisii	Fungi	4.11	Cuomo(2012)	1	Tromel (2008)
Trypanosoma brucei	Protist	20.5272	Berriman (2005)	7	Bitter (1988), Matthews (2005), Brun (2010)
Cryptosporidium hominis	Protist	8.171	Hou and Lin (2009)	1	Fayer (2004)
Theileria parva	Protist	8.3476	Gardner(2005)	28	Mbezeni 2013
Theileria equi	Protist	11.67	Kappmeyer(2012)	1*	Englund(2003)
Plasmodium falciparum	Protist	22.8595	Gardner(2002)	4	Olzewski(2009)
Plasmodium knowelesi	Protist	23.46	Pain(2008)	3*	Yakob (2010)
Plasmodium yoelii	Protist	20.087	Carlton(2002)	1	Carlton(2002)
Neurospora caninum	Protist	57.55	PRJNA172444, PRJEA37771	5*	Dubey (1996), Bartley (2004)
Toxoplasma gondii	Protist	63.05	PRJNA61553	15400	Boothroyd (2009)
Entameoba histolytica	Protist	21.5815	Hou and Lin (2009)	1	Wasson and Peper (2000)

Figure captions

Figure 1. Phylogenetic history of 96 parasite species used in PGLS analyses on whether there is a shared correlated history between genome size and host specificity. The topology and branch lengths (mya) of this phylogeny were derived from Hedges et al. (2016).

Figure 2. Conventional (red) and PGLS phylogenetic means (blue) of genome sizes (**a**) across modes of parasitism, (**b**) genome sizes across various methodologies used to determine DNA content, and (**c**) among host ranges of parasites. Error bars are standard errors (SE) derived from PGLS analyses including or excluding the parasite phylogeny. Numbers within parentheses are the number of parasite species (*N*) within groups.

Figure 3. PGLS analyses of the relationship between (**a**) genome size and host range across 96 parasite species, (**b**) genome size and host range when excluding specialists with only one host species as their host range (*N* = 66), (**c**) the relationship between genome size and host phylogenetic diversity estimated as the mean pairwise phylogenetic branch-length distances (MPD) between host species (*N* = 31), (**d**) the relationship between host range and host phylogenetic diversity, and finally (**e**) a PGLS

multiple regression of the relationship between genome size and host range parsed among fungi (red), ectoparasites (green), and endoparasites (blue). Ninety-five percent confidence intervals (95% CI) around PGLS regression lines were estimated following Smaers and Rohlf (2016).

Figure 1.

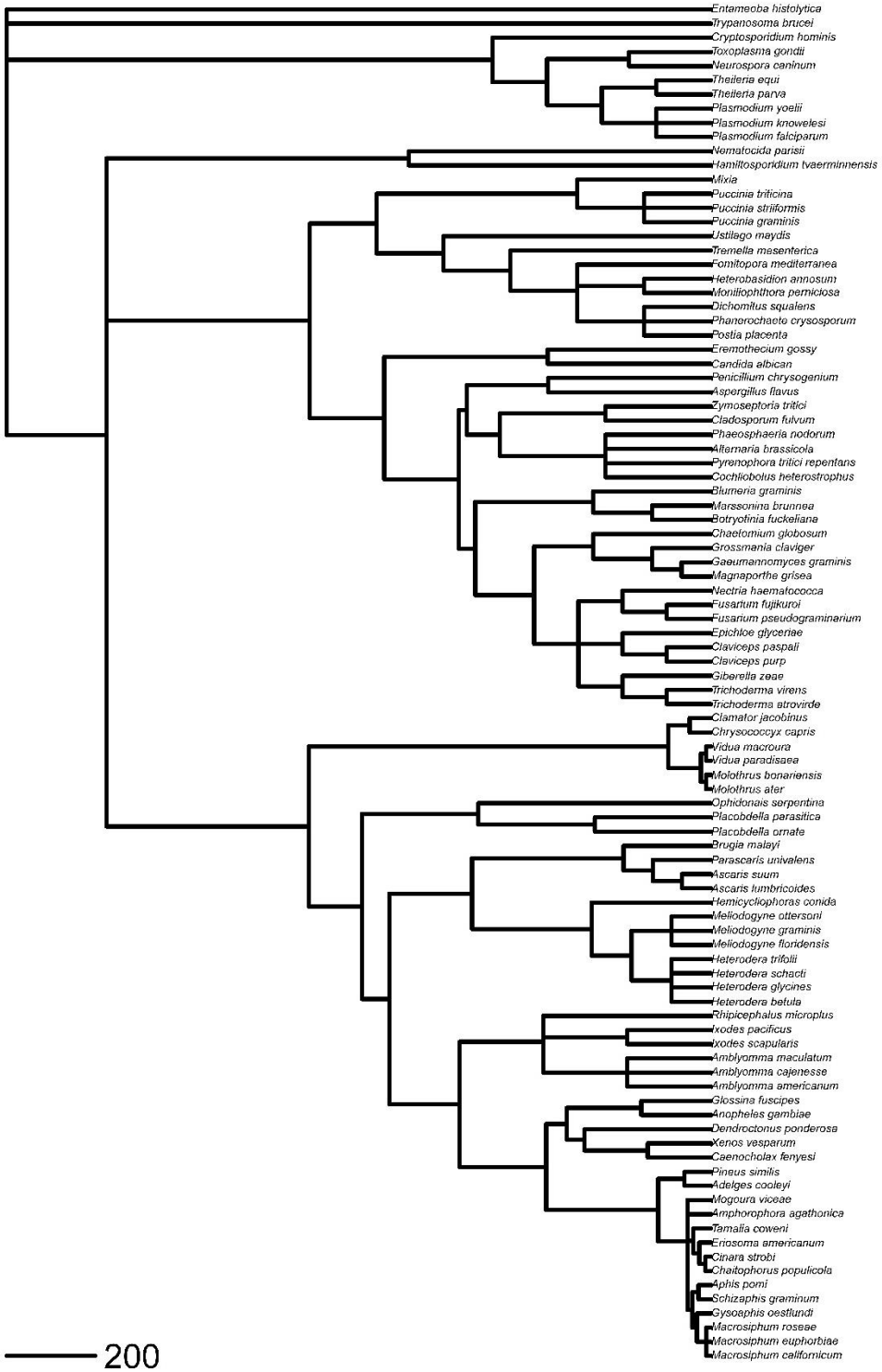

Entamoeba histolytica
Trypanosoma brucei
Cryptosporidium hominis
Toxoplasma gondii
Neurospora caninum
Theileria equi
Theileria parva
Plasmodium yoelii
Plasmodium knowelesi
Plasmodium falciparum
Nematocida parisii
Hamiltosporidium tvaerminnensis
Mixia
Puccinia triticina
Puccinia striiformis
Puccinia graminis
Ustilago maydis
Tremella mesenterica
Fomitopora mediterranea
Heterobasidion annosum
Moniliophthora perniciosa
Dichomitus squalens
Phanerochaete crysosporum
Postia placenta
Eremothecium gossy
Candida albican
Penicillium chrysogenium
Aspergillus flavus
Zymoseptoria tritici
Cladosporum fulvum
Phaeosphaeria nodorum
Alternaria brassicola
Pyrenophora tritici repentans
Cochliobolus heterostrophus
Blumeria graminis
Marssonina brunnea
Botryotinia fuckeliana
Chaetomium globosum
Grossmania claviger
Gaeumannomyces graminis
Magnaporthe grisea
Nectria haematococca
Fusarium fujikuroi
Fusarium pseudograminarium
Epichloe glyceriae
Claviceps paspali
Claviceps purp
Giberella zeae
Trichoderma virens
Trichoderma atrovirde
Clamator jacobinus
Chrysococcyx capris
Vidua macroura
Vidua paradisaea
Molothrus bonariensis
Molothrus ater
Ophidonais serpentina
Placobdella parasitica
Placobdella ornate
Brugia malayi
Parascaris univalens
Ascaris suum
Ascaris lumbricoides
Hemicycliophoras conida
Meliodogyne ottersoni
Meliodogyne graminis
Meliodogyne floridensis
Heterodera trifolii
Heterodera schacti
Heterodera glycines
Heterodera betula
Rhipicephalus microplus
Ixodes pacificus
Ixodes scapularis
Amblyomma maculatum
Amblyomma cajenesse
Amblyomma americanum
Glossina fuscipes
Anopheles gambiae
Dendroctonus ponderosa
Xenos vesparum
Caenocholax fenyesi
Pineus similis
Adelges cooleyi
Mogoura viceae
Amphorophora agathonica
Tamalia coweni
Eriosoma americanum
Cinara strobi
Chaitophorus populicola
Aphis pomi
Schizaphis graminum
Gysoaphis oestlundi
Macrosiphum roseae
Macrosiphum euphorbiae
Macrosiphum californicum
200

Figure 2.

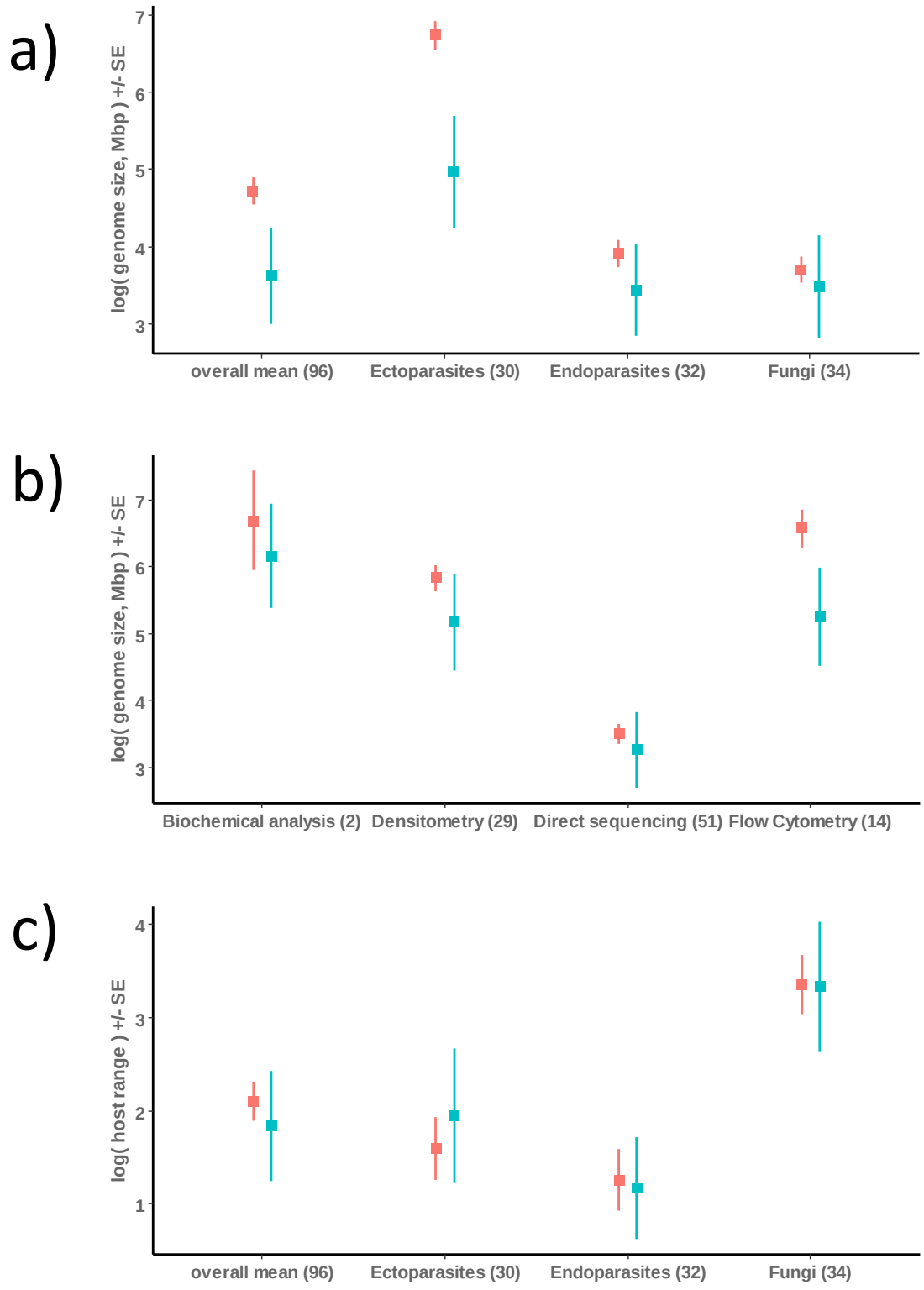
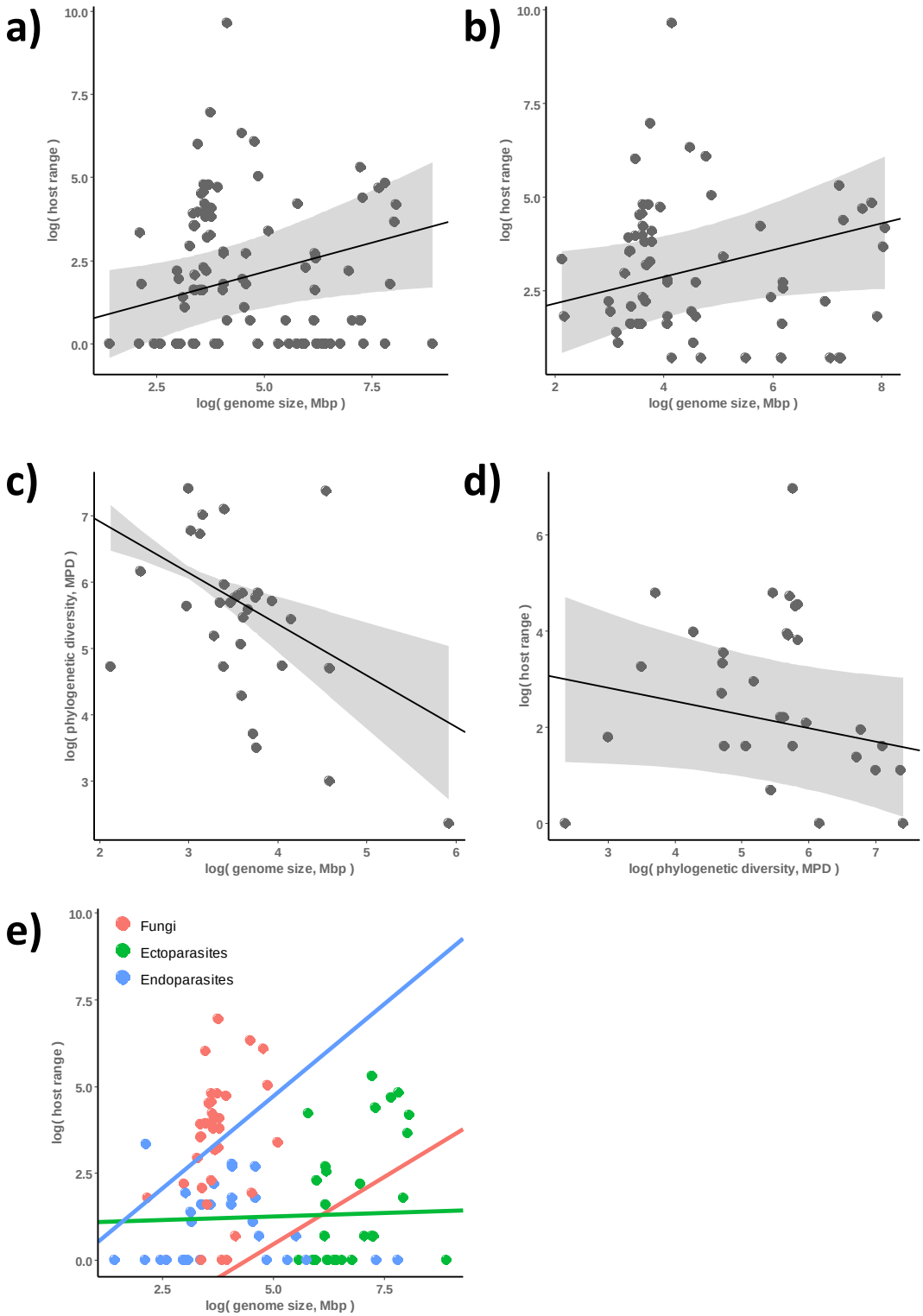

a)
log(genome size, Mbp) +/- SE
7
6
5
4
3
overall mean (96)
Ectoparasites (30)
Endoparasites (32)
Fungi (34)
b)
log(genome size, Mbp) +/- SE
7
6
5
4
3
Biochemical analysis (2)
Densitometry (29)
Direct sequencing (51)
Flow Cytometry (14)
c)
log(host range) +/- SE
4
3
2
1
overall mean (96)
Ectoparasites (30)
Endoparasites (32)
Fungi (34)

Figure 3


a)
b)
c)
d)
e)
log(host range)
log(genome size, Mbp)
log(phylogenetic diversity, MPD)
Fungi
Ectoparasites
Endoparasites