# Predicting epidemic size and disease evolution in response to environmental change


## Sam Paplauskas*

**Professor Matthew Tinsley[1], Dr Brad Duthie[2], Dr Stephen Thackeray[3]**

*My email: sam.paplauskas@stir.ac.uk

[1]Primary supervisor, [2]Co-supervisor, [3]External supervisor

[1-2]Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling FK9 4LA, United Kingdom

[3]UK Centre for Ecology and Hydrology (UKCEH), Lancaster, United Kingdom



Evolutionary Ecology, Host-Parasite Interactions, Epidemiology

Submitted on June 2024

## Biography

Sam Paplauskas studied Biological Sciences and obtained a first-class masters from the University of Sheffield. His PhD took place in Stirling, where he worked with Dr Stuart Auld for the first two years of study, before having a change of primary supervisor to Professor Matthew Tinsley. He currently lives in Coventry, where he lives with his family, and hopes to obtain a JSPS short-term fellowship after finishing publishing the work from his thesis.


Photo of Sam Paplauskas.

## Acknowledgements

Thanks to my initial supervisor, Dr **Stuart Auld,** for the opportunity to work together during the start of my PhD. I hope you are doing well in your new job!

Special thanks go to **Louise Boyle**, my university support worker, who guided me back to work after almost two years of absent leave.

## Abstract

Epidemics pose a major health risk to human, animal and plant life both domestically, in agricultural populations, and in the wild. To maintain global food security, biodiversity in the wild and human health, there is an urgent need for improved epidemic forecasting in response to broad environmental change. Most research concerned with this task is based on assessing individual epidemic size for a particular host-parasite interaction. However, in most cases, host populations experience recurrent epidemics that vary in size and severity through time, with shared characteristics among the diseases spread by different parasite species. In addition, there is a well-established link between environmental factors and disease transmission. Therefore, I propose a conceptual 'Disease Cycle' model to link the size of past and future epidemics. After highlighting the gaps in the current literature, I investigate some of the missing links in this theoretical model. Using a combination of real-world coevolution experiments, mathematical modelling of an infectious disease, and meta-analysis, I find: i) the amount of variation in host-parasite coevolutionary trajectories that is explained by the environment (chapter 3), ii) the effect of host-population genetic diversity on the variability in metrics of parasite success (chapter 4), (iii) the extent to which local hosts are affected by migrant competition (chapter 5) and iv) the additional accuracy that is gained by using replicate populations to forecast disease (chapter 6). Overall, I find strong support for certain links in the Disease Cycle, such as the effect of host population genetic diversity on future epidemic size, but there are others which require further study to understand the generality of this eco-evolutionary concept of disease epidemics.

# Contents*

*Please use CTRL click to follow links

# List of figures

201

202 # List of tables



203

204

205 # List of boxes



206

207

## Statement of Ethical Approval

We confirm that all of the experimental and field work methods employed in this study were reviewed and approved by the university's institutional review committee and all animals were cared for in accordance with institutional and national guidelines.

## Publications & pre-prints from this PhD

**Paplauskas, S**., Brand, J., & Auld, S. K. J. R. (2021). Ecology directs host–parasite coevolutionary trajectories across *Daphnia*–microparasite populations. *Nature Ecology & Evolution*, 5(4), 480–486. https://doi.org/10.1038/s41559-021-01390-7

**Paplauskas, S**., Duthie, B., & Tinsley, M. C. (2024). The effect of host population genetic diversity on the variation in metrics of parasite success. *BioRxiv*.

# 1. Thesis introduction

My PhD has focused on predicting epidemic size and disease evolution in response to environmental change using a theoretical 'Disease Cycle' model to link past and future epidemics in combination with empirical experiments involving the natural coevolution of a model *Daphnia* host – parasite system.

## 1.1 A theoretical 'Disease Cycle' model

Outbreaks of infectious disease threaten species and community levels of biodiversity (Altizer et al., 2003; Schmeller et al., 2020), both wild and crop systems (Newton et al., 2011; Strange & Scott, 2005) and pose a major risk to humans through the emergence of highly virulent zoonotic diseases (Jones et al., 2008; Schmeller et al., 2020). Although there are shared characteristics among diseases and most systems experience repeated epidemics that vary in size or severity over time (Altizer et al., 2006), most of our understanding of what drives variation in patterns of disease severity is drawn from studying separate host-parasite associations (Brockhurst & Koskella, 2013) and individual epidemic size (Miller, 2012).

Since host population genetic diversity can limit the spread of disease (King & Lively, 2012), and changes in both host and parasite diversity depend on the mode and pace of coevolutionary dynamics (Brockhurst & Koskella, 2013), it follows that the size of any contemporary outbreak is the product of previous patterns of host-parasite (co)evolution and genetic diversity from past infections. In addition, as we are currently living in an era of broad environmental change, and there is a well-established link between ambient temperature and disease transmission (Lafferty & Mordecai, 2016), there is an urgent need to better understand how we can effectively forecast disease in a changing world.

To address this knowledge gap, I propose a theoretical 'Disease Cycle' model to link past and future epidemic size (Fig. 1.1). After compiling a review of the Disease Cycle from previously published articles (Chapter two), I found consistent evidence for some aspects of the Disease Cycle (such as the mean reduction in parasite spread in high versus low diversity host populations) and less for others (such as the relationship between epidemic size and the strength of antagonistic selection). Therefore, one of the main objectives of my PhD research was to address some of the knowledge gaps in theoretical Disease Cycle model within each chapter of my

thesis. This involved using a combination of experimental coevolution using a model *Daphnia* host - parasite system and mathematical models to forecast future epidemics. The specific research questions addressed in each subsequent chapter are discussed in the following sections.

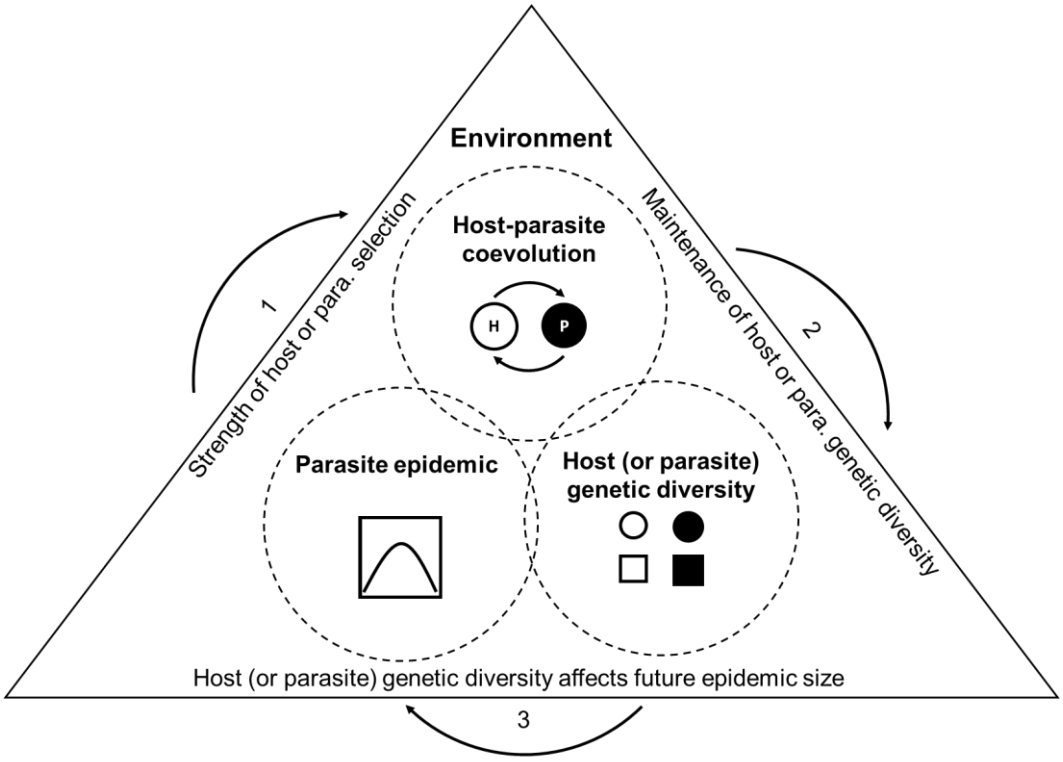



Figure 1.1. A **Disease Cycle** concept for linking the size of past and future epidemics. The proposed link between each component of the model (dashed circles) is shown by a numbered arrow (1-3). Specifically, I make the following predictions; 1) Epidemic size determines the strength of parasite (or host) mediated selection relative to other (a)biotic variables, 2) The tempo and mode of host-parasite co-evolution, which may be linked to the underlying model of host-parasite infection genetics (Agrawal & Lively, 2002), determines how the level of host (or parasite) population genetic diversity changes over time (Brockhurst & Koskella, 2013) and 3) The level of host (or parasite) population genetic diversity determines future epidemic size. Previous studies have shown how host populations with higher levels of genetic diversity have a smaller mean epidemic size (Ekroth et al., 2019; Gibson & Nguyen, 2021), but it is unclear how this combines with the corresponding level of genetic diversity in the parasite population to affect the variability in future epidemic size. Each link in the Disease Cycle is set within the context of environmental change (triangle).

## 1.2 Co-evolutionary trajectories in 'real-world' environments

Although there have been many laboratory-based measurements of the magnitude and direction of host-parasite co-evolution (coevolution *potential*), to what extent these patterns of host-parasite co-evolution translate over to 'real-world' environments is not entirely clear (coevolution *realised*). In addition, whether coevolution is repeatable remains a generally unanswered question in Evolutionary Biology. Therefore, I measured the extent to which environmental differences between populations with a shared ancestral origin followed similar coevolutionary trajectories. Ordinarily, natural populations vary so much that it difficult to examine the repeatability of host-parasite interactions, but the ability of *Daphnia* to produce parthenogenic clones means that starting populations were identical, which allowed me to pose the following questions:

1. What is the pattern of host evolution of resistance, parasite evolution of infectivity, and coevolution (i.e., the extent to which the parasite population non-additively evolved in response to a changed complement of host genotypes)?
2. How much of this change in host resistance, parasite infectivity and coevolution is driven by the environment?
3. Overall, are host, parasite and both host and parasite patterns of coevolution repeatable?

## 1.3 Is there really a conventional 'monoculture effect' beyond agriculture?

So-called 'conventional wisdom' would have us believe that low levels of population genetic diversity in non-plant populations, usually increase the risk of infectious disease epidemics, which is sometimes referred to as a 'monoculture effect'. This is because the susceptibility of low diversity crop mixtures to epidemics of disease, such as the devastation of crop monocultures that are entirely composed of a single species or cultivar, has been well-established in the plant literature for many years. Recent attempts to qualify the generality of this disease-diversity relationship beyond agriculture have focused on studying the mean, rather than the variability of metrics of parasite success. By re-analysing their meta-analytical data, I ask the following questions:

1. What is the general effect of host population genetic diversity on not only the mean, but also the variability of parasite success?

2. Does this effect vary between parasite specialists and generalists, as well as parasite populations with different levels of genetic diversity?
3. Overall, is this consistent with my proposed diversity-uncertainty model?

## 1.4 Parasite-mediated competition in non-locally adapted host populations

Local adaptation is a powerful evolutionary force, whereby the individuals within a population adapt to their local environment by evolving traits that increase their fitness in that environment relative to others. How variation in the competitive ability of local host populations is affected by patterns of local adaptation to the abiotic environment is poorly understood. To test whether host populations are better adapted to their local environment than migrants, and how a general parasite exposure can mediate their competitive interactions, I compared the reproductive output of adult hosts in a series of reciprocal transplant experiments, involving home, away and mixed host groups in either the presence or absence of a shared (ancestral) parasite, among 12 replicate *Daphnia* host – parasite pond populations. Specifically, I asked:

1. What is the pattern of host local adaptation?
2. Do immigrants suffer from competition with resident hosts?
3. Overall, is there a parasite-mediated cost of competition with residents for immigrants?

## 1.5 Quantity has a quality all of its own for predicting epidemic size

Most researchers forecast disease in a single population using long-term historical data from that population. However, long-term data is not always available and instead it might be possible to borrow data from similar populations to forecast future epidemic size for a given population. We might further increase epidemic forecasting accuracy by weighting the contribution of individual epidemics to the future epidemic forecast based on their environmental similarity to a focal population. Therefore, I use a range of approaches to forecasting future epidemic size based on historical data collected from 20 semi-natural pond populations of a model *Daphnia* host - parasite system across four years (total of 80 epidemics). Specifically, I ask the following questions:

1. Are forecasts of future epidemic size from models trained on multiple populations more accurate than those trained only on the target population?

2. Are forecasts of future epidemic size from ARIMA and regression models more accurate than benchmark models?
3. Overall, can replicate populations across space and their corresponding variation in environmental conditions increase epidemic size forecast accuracy?

## 1.6 Natural experimental coevolution of *Daphnia* – parasite systems as a useful model for research

In the following section, I provide a brief introduction to the *Daphnia* host – parasite system used to study disease evolution in the wild in subsequent chapters. Specifically, I discuss the costs and benefits of *Daphnia* – parasite systems as model for my research on predicting epidemic size and disease evolution in 'real-world' environments.

### 1.6.1 Why *Daphnia* hosts are a useful model for (co)evolution research

To what extent *D. magna* is a unique model organism versus a good representation of other non-vertebrate (or even vertebrate) host species is not entirely objective (Ebert, 2008).

*D. magna* (Fig. 1.2) are small, pond-dwelling organisms and, together with *D. pulex*, are the most well studied of species of this genus. They tend to occur mostly in freshwater, but also brackish, throughout the globe, and in particular Western Europe (Fig. 1.3). Some advantages of studying this model system include how easy they are to culture for scientific study and a well-documented host-parasite ecology (Ebert, 2005). The benefits of performing evolutionary studies with this system is that they are able to evolve rapidly in response to parasite-mediated selection (Paplauskas et al., 2021) and, most of all, have the ability to reproduce asexually via parthenogenesis (a form of asexual reproduction where virgin females give birth to daughters, Fig. 1.4). Therefore, this means that ancestral genotypes can be maintained in isolation and compared to evolved genotypes in a so-called 'time-shift' experiment (Brockhurst & Koskella, 2013).

Figure 1.2. Photo of female *Daphnia magna* susceptible to (left) and infected by (right) *Pasteuria ramosa* (scale bar 1mm (Ebert, 2008).

365


Figure 1.3. Examples of freshwater and brackish habitats *D. magna* live in. See the figure legend in (Ebert, 2022) for description of each letter.

366

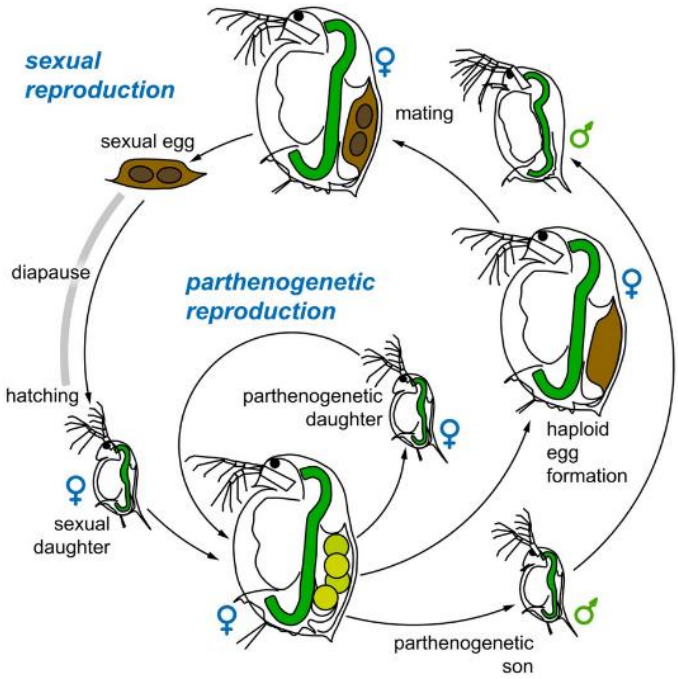



Figure 1.4. *Daphnia* facultative reproduction and life-cycle (Ebert, 2022).

Other examples of model host species used in coevolutionary studies include snails (Koskella & Lively, 2007, 2009), *C. elegans* (Papkou et al., 2019; Schulenburg & Müller, 2004; Schulte et al., 2011) and bacteria infected by phage parasites (Brockhurst et al., 2007; Castledine et al., 2022; Gómez & Buckling, 2011; Koskella, 2013; Koskella & Brockhurst, 2014; Lopez Pascua et al., 2012) (for a review (Brockhurst & Koskella, 2013)).

Potential disadvantages to the *D. magna* model host organism include:

1) The unusual foraging behaviour responsible for causing primary infections (where individuals pick up infections from rummaging around in the substrate) of its environmentally transmitted parasite, *Pasteruia ramosa* (see 1.6.2 Why *Pasteuria* parasites are a useful model for (co)evolution research).
2) Its unique mode of reproduction (which can also be very beneficial).

Although there are legitimate concerns about the generality of *Daphnia* experiments due to these disadvantages, they are outweighed by the considerable benefits to host-parasite research.

### 1.6.2 Why *Pasteuria* parasites are a useful model for (co)evolution research

*Pasteuria ramosa* is a sterilising obligate parasite of *Daphnia*, with the *D. magna* host being its most popular target (Fig. 1.5). It is commonly used in studies of host-parasite coevolution, such as for the investigation of the genetic basis of infection as part of a matching-allele model (Bento et al., 2017a), due to its well-defined genetically determined stepwise infection process (Duneau et al., 2011; Luijckx et al., 2012, 2013a).


Figure 1.5. *Pasteuria ramosa* as a model parasite. a) Healthy (left) and *Pasteuria ramosa* infected (right) adult *D. magna.* b) Transmission stage of the parasite (spores). Attachment of the parasite to the c) oesophagus and d) hindgut of *D. magna* adults. All photos courtesy to (Ebert, 2022).

The parasite also has a strong impact on host fitness; it eventually kills the host as well as sterilising the host (Ebert, 2008). However, this is just as much of an advantage, in terms of having a strong disease phenotype, as it is a disadvantage. The extremely virulent nature of the parasite may be incomparable to other systems. In addition, as mentioned above, the ability of *Pastueria* to produce resting stages means that *Daphnia* primary infections are caused by their contact with these dormant parasite spores in pond sediments – which is an unorthodox mode of transmission.

Despite this potential confounding characteristic of *Pasteuria* transmission, it also provides the unique opportunity to study historical patterns of host-parasite coevolution. Since *Daphnia* can produce sexual resting stages too, this means that

both host and parasite can be resurrected from pond 'sediment cores' (Decaestecker et al., 2007).

In common with certain other *Daphnia* parasite species, such as *Spirobacillus cienkowskii* (Ebert, 2008), the infection of hosts caused by *P. ramosa* can be identified visually (Fig. 1.2 and 1.5a). In addition, in common with other model parasite species used in coevolutionary time-shift experiments (Brockhurst & Koskella, 2013), *P. ramosa* transmission stages can be kept in evolutionary stasis under freezing conditions, so that ancestral strains of the parasite can be compared to their contemporaries.

### 1.6.3 Costs and benefits of experiments in so-called 'real-world' environments

Mesocosms (Odum, 1984), or semi-natural environments, are a useful tool for studying ecological and evolutionary responses to climate-change (Stewart et al., 2013). They are a fundamental part of aquatic ecological experimentation (Spivak et al., 2011) and allow the replication of laboratory studies whilst maintaining some kind of ecological realism.

Mesocosms differ to microcosms by definition of their size, which includes enclosures from 1 to several thousands of litres (Stewart et al., 2013), but also through utilising natural, rather than artificially generated, abiotic conditions (Wijngaarden et al., 2005) (Fig. 1.6).

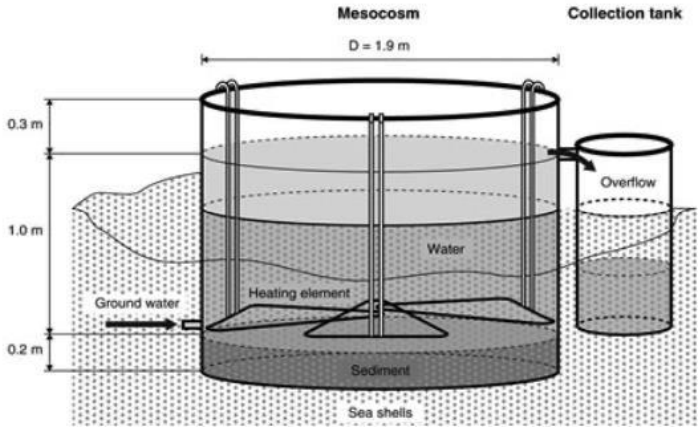



Figure 1.6. Mesocosm experiment for freshwater climate change (Lake Mesocosm Warming Experiment (LMWE), AQUACOSM, Denmark, 2003-2024+). The tank volume is 2.8m$^3$.

Laboratory studies for experimental (co)evolution rose to prominence as a replacement for the first wave of empirical coevolution research, which was predominantly observational and field based (Ehrlich & Raven, 1964; Janzen, 1966). They were useful for providing evidence of reciprocal antagonistic coevolutionary interactions (Kawecki & Ebert, 2004), which were otherwise both attributable to extraneous sources of variation (Brockhurst & Koskella, 2013). However, progressively more studies are returning to the field to study these 'real-world' environments (see reviews by (Brockhurst & Koskella, 2013; Koskella & Brockhurst, 2014)).

The main advantage of these experiments is also their biggest limitation. Since the same level of replication, control and tractability can usually only be achieved under laboratory conditions, there is a resulting trade-off between uncovering general evolutionary mechanisms and understanding how they apply in complex natural environments *sensu* (Scheinin et al., 2015). Others criticise mesocosm experiments as being unrealistic simplifications with limited relevance to natural ecosystems (for a review (Stewart et al., 2013), but see Box 1.1).

**Box 1.1. A response to critics**

In chapter two, I found that mesocosm environments (biotic and abiotic factors collectively referred to by ecology) were significantly involved in directing *Daphnia magna* host-parasite (co)evolutionary trajectories (Paplauskas et al., 2021). This study, made in answer to a call for more ways of measuring the strength of coevolution in the wild (Week & Nuismer, 2019), used the aforementioned mesocosm approach for experimental coevolution.

Week and Nuismer (2019) commented on the fact that time-shift experiments had been broadly implemented in systems where experimental evolution was a tractable approach, but they had not yet yielded a quantitative assessment of the strength of coevolution (Koskella 2014; Blanquart & Gandon 2013; Gaba & Ebert 2009). In addition, due to the constraints that can be imposed on coevolution by natural conditions, we propose that it is equally important to measure the strength of coevolution in both controlled, laboratory based environments and natural ones (Brockhurst & Koskella, 2013; Koskella & Brockhurst, 2014).

For example, unlike coevolution in the lab, which is characterized by an increase in both host resistance and parasite infectivity over time (*sensu* arms-race dynamics (Buckling & Rainey, 2002)), coevolution in soil mesocosms (or technically microcosms, see earlier definition; see 1.6.3 Costs and benefits of experiments in so-called 'real-world' environments) led to greater resistance to contemporary, rather than past or future, parasites (*sensu* fluctuating selection dynamics) in a bacteria-host-bacteriophage-parasite interaction (Gómez & Buckling, 2011). In the same host-parasite association, fluctuating selection dynamics switch back to arms-race dynamics under a mixing treatment (Gómez et al., 2014). Another seminal coevolutionary experiment showed that the evolution of resistance in populations of *D. magna* infected with a fungal parasite under natural conditions were associated with life-history costs (Zbinden et al., 2008).

## 1.7 Thesis structure

In the remainder of the thesis, each chapter provides a more detailed introduction to the focal study, a description of the full methodology, the key findings and a discussion. In chapter seven, the results from each study are discussed in the context of the wider literature, integrated into an evaluation of the theoretical Disease Cycle model and I make some suggestions for future research.

## 1.8 References


Agrawal, A., & Lively, C. M. (2002). Infection genetics: Gene-for-gene versus matching-alleles models and all points in between. *Evolutionary Ecology Research*.

Altizer, S., Dobson, A., Hosseini, P., Hudson, P., Pascual, M., & Rohani, P. (2006). Seasonality and the dynamics of infectious diseases. *Ecology Letters*, *9*(4), 467–484. https://doi.org/10.1111/j.1461-0248.2005.00879.x

Altizer, S., Harvell, D., & Friedle, E. (2003). Rapid evolutionary dynamics and disease threats to biodiversity. *Trends in Ecology & Evolution*, *18*(11), 589–596. https://doi.org/10.1016/J.TREE.2003.08.013

Bento, G., Routtu, J., Fields, P. D., Bourgeois, Y., Du Pasquier, L., & Ebert, D. (2017). The genetic basis of resistance and matching-allele interactions of a host-parasite system: The *Daphnia magna-Pasteuria ramosa* model. *PLOS Genetics*, *13*(2), e1006596. https://doi.org/10.1371/journal.pgen.1006596

476 Brockhurst, M. A., & Koskella, B. (2013). Experimental coevolution of
477 species interactions. *Trends in Ecology & Evolution*, *28*(6), 367–375.
478 https://doi.org/10.1016/J.TREE.2013.02.009

479 Brockhurst, M. A., Morgan, A. D., Fenton, A., & Buckling, A. (2007).
480 Experimental coevolution with bacteria and phage. The *Pseudomonas fluorescens*-
481 Φ2 model system. In *Infection, Genetics and Evolution* (Vol. 7, Issue 4).
482 https://doi.org/10.1016/j.meegid.2007.01.005

483 Buckling, A., & Rainey, P. B. (2002). Antagonistic coevolution between a
484 bacterium and a bacteriophage. *Proceedings of the Royal Society B: Biological*
485 *Sciences*, *269*(1494), 931–936. https://doi.org/10.1098/rspb.2001.1945

486 Castledine, M., Sierocinski, P., Inglis, M., Kay, S., Hayward, A., Buckling,
487 A., & Padfield, D. (2022). Greater phage genotypic diversity constrains arms-race
488 coevolution. *Frontiers in Cellular and Infection Microbiology*, *12*.
489 https://doi.org/10.3389/fcimb.2022.834406

490 Decaestecker, E., Gaba, S., Raeymaekers, J. A. M., Stoks, R., Van
491 Kerckhoven, L., Ebert, D., & De Meester, L. (2007). Host–parasite 'Red Queen'
492 dynamics archived in pond sediment. *Nature*, *450*(7171), 870–873.
493 https://doi.org/10.1038/nature06291

494 Duneau, D., Luijckx, P., Ben-Ami, F., Laforsch, C., & Ebert, D. (2011).
495 Resolving the infection process reveals striking differences in the contribution of
496 environment, genetics and phylogeny to host-parasite interactions. *BMC Biology*, *9*,
497 11. https://doi.org/10.1186/1741-7007-9-11

498 Ebert, D. (2005). *Ecology, Epidemiology, and Evolution of Parasitism in*
499 Daphnia (S. Paliwoda, Ed.).

500 Ebert, D. (2008). Host–parasite coevolution: Insights from the *Daphnia*–
501 parasite model system. *Current Opinion in Microbiology*, *11*(3), 290–301.
502 https://doi.org/10.1016/J.MIB.2008.05.012

503 Ebert, D. (2022). *Daphnia* as a versatile model system in ecology and
504 evolution. *EvoDevo*, *13*(1), 16. https://doi.org/10.1186/s13227-022-00199-0

505 Ehrlich, P. R., & Raven, P. H. (1964). Butterflies and plants: a study in
506 coevolution. *Evolution*, *18*(4), 586–608. https://doi.org/10.1111/j.1558-
507 5646.1964.tb01674.x

Ekroth, A. K. E., Rafaluk-Mohr, C., & King, K. C. (2019). Host genetic diversity limits parasite success beyond agricultural systems: a meta-analysis. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1911), 20191811. https://doi.org/10.1098/rspb.2019.1811

Gibson, A. K., & Nguyen, A. E. (2020). Does genetic diversity protect host populations from parasites? A meta-analysis across natural and agricultural systems. *Evolution Letters*, *5*(1). https://doi.org/10.1002/evl3.206

Gómez, P., Ashby, B., & Buckling, A. (2015). Population mixing promotes arms race host-parasite coevolution. *Proceedings of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rspb.2014.2297

Gómez, P., & Buckling, A. (2011). Bacteria-phage antagonistic coevolution in soil. *Science*, *332*(6025), 106–109. https://doi.org/10.1126/science.1198767

Janzen, D. H. (1966). Coevolution of mutualism between Ants and Acacias in central America. *Evolution*, *20*(3), 249. https://doi.org/10.2307/2406628

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, *451*(7181), 990–993. https://doi.org/10.1038/nature06536

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*(12), 1225–1241. https://doi.org/10.1111/j.1461-0248.2004.00684.x

King, K. C., & Lively, C. M. (2012). Does genetic diversity limit disease spread in natural host populations? *Heredity*, *109*(4), 199–203. https://doi.org/10.1038/hdy.2012.33

Koskella, B. (2013). Phage-mediated selection on microbiota of a long-lived host. *Current Biology*, *23*(13), 1256–1260. https://doi.org/10.1016/J.CUB.2013.05.038

Koskella, B., & Brockhurst, M. A. (2014). Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiology Reviews*, *38*(5), 916–931. https://doi.org/10.1111/1574-6976.12072

Koskella, B., & Lively, C. M. (2007). Advice of the rose: experimental coevolution of a trematode parasite and its snail host. In *Source: Evolution* (Vol. 61, Issue 1).

https://www.jstor.org/stable/pdf/4621270.pdf?refreqid=excelsior%3Ae1a6629ec3688f6fbdaa3c80f4874723

Koskella, B., & Lively, C. M. (2009). Evidence for negative frequency-dependent selection during experimental coevolution of a freshwater snail and a sterilizing trematode. *Evolution*, *63*(9), 2213–2221. https://doi.org/10.1111/j.1558-5646.2009.00711.x

Lafferty, K. D., & Mordecai, E. A. (2016). The rise and fall of infectious disease in a warmer world. *F1000Research*, *5*. https://doi.org/10.12688/f1000research.8766.1

Lopez Pascua, L., Gandon, S., & Buckling, A. (2012). Abiotic heterogeneity drives parasite local adaptation in coevolving bacteria and phages. *Journal of Evolutionary Biology*, *25*(1), 187–195. https://doi.org/10.1111/j.1420-9101.2011.02416.x

Luijckx, P., Fienberg, H., Duneau, D., & Ebert, D. (2012). Resistance to a bacterial parasite in the crustacean *Daphnia magna* shows Mendelian segregation with dominance. *Heredity*, *108*(5), 547–551. https://doi.org/10.1038/hdy.2011.122

Luijckx, P., Fienberg, H., Duneau, D., & Ebert, D. (2013). A matching-allele model explains host resistance to parasites. *Current Biology*, *23*(12), 1085–1088. https://doi.org/10.1016/j.cub.2013.04.064

Miller, J. C. (2012). A note on the derivation of epidemic final sizes. *Bulletin of Mathematical Biology*, *74*(9), 2125–2141. https://doi.org/10.1007/s11538-012-9749-6

Newton, A. C., Johnson, S. N., & Gregory, P. J. (2011). Implications of climate change for diseases, crop yields and food security. In *Euphytica* (Vol. 179, Issue 1). https://doi.org/10.1007/s10681-011-0359-4

Odum, E. P. (1984). The mesocosm. *BioScience*, *34*(9), 558–562. https://doi.org/10.2307/1309598

Papkou, A., Guzella, T., Yang, W., Koepper, S., Pees, B., Schalkowski, R., Barg, M. C., Rosenstiel, P. C., Teotónio, H., & Schulenburg, H. (2019). The genomic basis of red queen dynamics during rapid reciprocal host–pathogen coevolution. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(3), 923–928. https://doi.org/10.1073/pnas.1810402116

Paplauskas, S., Brand, J., & Auld, S. K. J. R. (2021). Ecology directs host–parasite coevolutionary trajectories across *Daphnia*–microparasite populations. *Nature Ecology & Evolution*, *5*(4), 480–486. https://doi.org/10.1038/s41559-021-01390-7

Scheinin, M., Riebesell, U., Rynearson, T. A., Lohbeck, K. T., & Collins, S. (2015). Experimental evolution gone wild. *Journal of the Royal Society, Interface*, *12*(106). https://doi.org/10.1098/rsif.2015.0056

Schmeller, D. S., Courchamp, F., & Killeen, G. (2020). Biodiversity loss, emerging pathogens and human health risks. In *Biodiversity and Conservation*. https://doi.org/10.1007/s10531-020-02021-6

Schulenburg, H., & Müller, S. (2004). Natural variation in the response of *Caenorhabditis elegans* towards *Bacillus thuringiensis*. *Parasitology*, *128*(4), 433–443. https://doi.org/10.1017/S003118200300461X

Schulte, R. D., Makus, C., Hasert, B., Michiels, N. K., & Schulenburg, H. (2011). Host-parasite local adaptation after experimental coevolution of *caenorhabditis elegans* and its microparasite *bacillus thuringiensis*. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1719), 2832–2839. https://doi.org/10.1098/rspb.2011.0019

Spivak, A. C., Vanni, M. J., & Mette, E. M. (2011). Moving on up: can results from simple aquatic mesocosm experiments be applied across broad spatial scales? *Freshwater Biology*, *56*(2), 279–291. https://doi.org/10.1111/j.1365-2427.2010.02495.x

Stewart, R. I. A., Dossena, M., Bohan, D. A., Jeppesen, E., Kordas, R. L., Ledger, M. E., Meerhoff, M., Moss, B., Mulder, C., Shurin, J. B., Suttle, B., Thompson, R., Trimmer, M., & Woodward, G. (2013). Mesocosm Experiments as a Tool for Ecological Climate-Change Research. *Global Change in Multispecies Systems: Part 3*, 71–181. https://doi.org/10.1016/B978-0-12-417199-2.00002-1

Strange, R. N., & Scott, P. R. (2005). Plant disease: A threat to global food security. *Annual Review of Phytopathology*, *43*(1), 83–116. https://doi.org/10.1146/annurev.phyto.43.113004.133839

Week, B., & Nuismer, S. L. (2019). The measurement of coevolution in the wild. *Ecology Letters*, *22*(4), 717–725. https://doi.org/10.1111/ele.13231

Wijngaarden, R. P. A. Van., Brock, T. C. M., & Brink, P. J. Van Den. (2005). Threshold levels for effects of insecticides in freshwater ecosystems: A review. *Ecotoxicology*, *14*(3), 355–380. https://doi.org/10.1007/s10646-004-6371-x

Zbinden, M., Haag, C. R., & Ebert, D. (2008). Experimental evolution of field populations of *Daphnia magna* in response to parasite treatment. *Journal of Evolutionary Biology*, *21*(4), 1068–1078. https://doi.org/10.1111/j.1420-9101.2008.01541.x

## 2. A theoretical ‘Disease Cycle’ model to link past and future epidemics

## 2.1 Introduction

Outbreaks of infectious diseases pose a major threat to biodiversity, agriculture and human health (Altizer et al., 2003; Jones et al., 2008; Schmeller et al., 2020). For any one host population, the various effects of disease can include reduced host genetic diversity, depressed population size and, in some cases, complete extinction (Alan Pounds et al., 2006; Boots & Sasaki, 2002; Vredenburg et al., 2010) These negative effects are also often exacerbated by anthropogenic selection pressures associated with urbanisation, intensive agriculture and human-induced climate change (Engebretsen et al., 2019; Price et al., 2019; White & Razgour, 2020). To protect populations in an era of broad environmental change, we require disease control strategies. The effective design of such strategies relies on (1) a detailed understanding of the various drivers of disease and (2) some capacity to predict outbreaks in the future. Understanding and forecasting any one disease is, however, fraught with challenges.

These challenges stem from two important complexities associated with each disease system. First, transmission itself is typically a multistep process, comprising pathogen contact with the host, entry to the host, various interactions with the host immune system, within-host proliferation and onward transmission (McCallum et al., 2017). Crucially, environmental variation can affect each of these steps (Duneau et al., 2011). For example, higher temperatures generally cause an increase in parasite growth rates, survival and vector competence (Dohm et al., 2002; Ohm et al., 2018; Piotrowski et al., 2004; Poulin, 2006) but these responses vary due to individual differences in thermal biology (Koprivnikar & Poulin, 2009; Mordecai et al., 2019; Poulin, 2006). In addition, covariation among various components of infection can lead to counter-intuitive effects on disease in the future (Fels & Kaltz, 2006; Paaijmans et al., 2012; Paull, Lafonte and Johnson, 2012; Lafferty & Mordecai, 2016; Shocket et al., 2019). For example, higher temperatures cause increased exposure to pathogen infectious stages in a *Daphnia*-parasite system (by speeding up host foraging rate), but reduce within-host parasite growth above a certain threshold once infection has occurred (Shocket et al., 2019). This results in fewer parasite transmission stages for onward transmission and thus potentially smaller epidemics in the future.

Second, any particular host-pathogen relationship is part of a much wider, more complex, ecological arena where other interactions such as competition or predation

can play a greater role in shaping host and pathogen populations (Bowers et al., 1994; Duffy et al., 2012; Gutierrez et al., 2022; Hall et al., 2005, 2009; Ibelings et al., 2004; Paplauskas et al., 2021; Thieltges et al., 2008). In multi-host systems, despite a strong dilution effect, where the presence of compatible hosts that are less susceptible to infection (often termed, more 'unsuitable') reduces overall epidemic size, competition between different hosts can potentially lead to complex and varied disease outcomes (Cáceres et al., 2014) and in one study this lead to an increase in host density and overall epidemic size (Hall, Becker, et al., 2009). In populations with multiple parasites, there can be competition between parasites within the host which can determine their reproductive success (Refardt, 2011) and in some cases leads to the evolution of higher virulence (De Roode et al., 2005) Predation can affect host and pathogen populations in many different ways (Duffy et al., 2019), most notably, selective predation of infected individuals can reduce overall epidemic size (Gutierrez et al., 2022). Moreover, wider ecological shifts can alter the relative size and severity of disease outbreaks from each pathogen. For example, predation of buffalo with heavy tick infestations led to unusually high levels of parasitic infection in Serengeti lions which resulted in a high mortality rate due to the immunosuppressive effects of a coincident canine distemper virus (Munson et al., 2008). In another example, 'sloppy' (messy) predation of *Daphnia* host individuals by *Chaoborus* phantom midge larvae, which results in indirect release of parasite transmission stages, has the capacity to mediate the abundance of different parasites by releasing faster-growing spores from infected individuals (Auld, Hall, et al., 2014).

As previously described in the Thesis Introduction (chapter one), there are numerous factors, other than just the environment, which can affect disease as an additional or principal driver of transmission. This includes the concept of epidemics as drivers of host-parasite co-evolution, host-parasite coevolution mediated-changes in genetic diversity and the effects of host (or parasite) population-level genetic diversity on future epidemic size. Therefore, it follows that any given disease outbreak is a product of both past and present parasite transmission.

Here, I present a simple conceptual model, the Disease Cycle, that bridges the evolutionary ecology of past and future disease outbreaks in a variable world (Fig. 1.1). I review research relevant to each aspect of the Disease Cycle framework and evaluate how environment-mediated selection could influence different components

of the model to affect disease over time. My primary aim is to provide a framework for future modelling approaches that embrace epidemic disease as a recurrent episodic process and help better inform the forecasting and management of disease control strategies.

## 2.2 Epidemics as drivers of host-parasite coevolution

Epidemics occur when the number of hosts infected with a particular pathogen increases rapidly over a short period of time with respect to the usual baseline (endemic) prevalence (Dicker, 2006), and are implicated as engines of evolutionary change in numerous disease systems (Altizer et al., 2003; Auld & Brand, 2017a; Thrall et al., 2012). However, quantifying the specific relationships between epidemic size or severity and the underlying host-parasite (co)evolutionary change across replicated natural populations is a complex and delicate task. Epidemics are population-level expressions of individual-level infections. Each infection is a phenotype that is shaped by the environment and both host and parasite traits such as resistance and infectivity. Therefore, epidemics are both multivariate and multiscale in nature, and vary in magnitude both within and across disease systems (Altizer et al., 2006; Penczykowski et al., 2016).

There are a number of key factors which determine epidemic size, such as the host-pathogen contact rate. This primarily depends on the mode of transmission, which is driven by either the density or frequency of infected hosts. If pathogen transmission depends on host density, then the change in the number of infected hosts in a population is equal to:

$$dI/dt = \beta SI$$

where I is the number of infected individuals, t is time, $\beta$ is the transmission rate and S is the number of susceptible individuals. This means that transmission of the pathogen increases with host density (linearly or non-linearly) and is referred to as density-dependent transmission. In comparison, if pathogen transmission depends on the frequency of infected hosts then the change in the number of infected is equal to:

$$dI/dt = \beta SI/N$$

where N is the population size. This is termed frequency dependent transmission.

Another major driver of epidemic size is environmental change. How organisms interact with their biotic (living) and abiotic (non-living) environment shapes the size and severity of future epidemics by affecting a number of different processes, including host supply (Begon et al., 2009) parasite load (Civitello et al., 2015), host-parasite encounter rates (Hall, Becker, et al., 2009; Strauss et al., 2018), and transmission rate (Shocket et al., 2018). In particular, changes in temperature can dramatically alter parasite growth rates and transmission (Mordecai et al., 2019; Piotrowski et al., 2004; Poulin, 2006).

### 2.2.1 What defines an epidemic?

Disease systems vary considerably in what constitutes a rapid and large increase in infected hosts, i.e. the threshold for an epidemic (Reliefweb, 2008) making it difficult to compare across systems. The absence of a standard measure of epidemic size means that studies use a variety of different measures to describe epidemic size, including peak, mean, or integrated parasite prevalence (Fig. 2.1). Various measures of epidemic size will differ in how they predict important ecological or evolutionary processes in any one host-pathogen system. Similarly, disease systems vary in how they define the severity of an outbreak, which can be measured in terms of the overall impact on host health and fitness and may also account for epidemic size, although this is not a measure of severity *per se*, and is therefore closely tied to parasite virulence, which is the reduction in host fitness caused by infection (Read, 1994). For example, proliferative kidney disease of salmonid fish is caused by a highly virulent parasite and often mortalities reach as high as 95-100% (Hedrick et al., 1993), whereas host abundance shrinks by 20-40% due to mycoplasmal conjunctivis affecting passerine birds, which is commonly regarded as another devastating parasite (Hochachka & Dhondt, 2000). There is, however, considerable merit in placing different disease systems on an equal footing, because it will allow us to make comparisons across disease systems; this will enable us to use knowledge of well-understood host-pathogen systems to understand (and potentially predict) the behaviours of other, less well-known systems (Han et al., 2020).

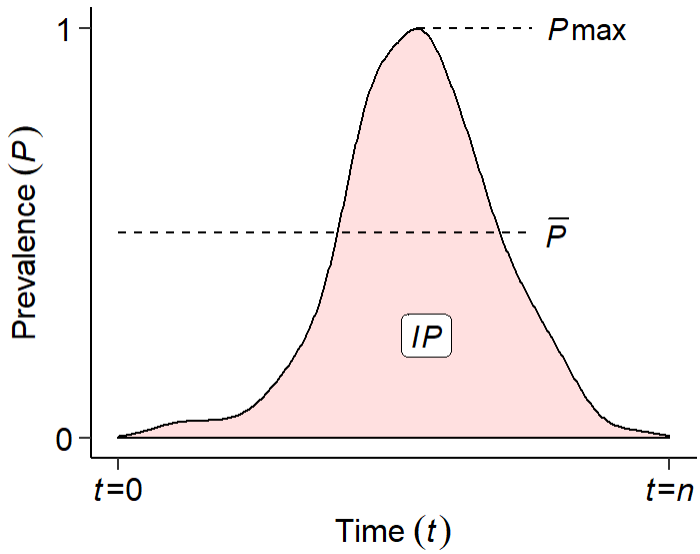



Figure 2.1. Various measures of epidemic size. Epidemic size is described using various measures of parasite prevalence (proportion of infected hosts); $Pmax$ is the peak prevalence, $\overline{P}$ is the mean prevalence and $IP$ is the integrated prevalence (i.e. parasite prevalence over time, equal to the shaded area under the epidemic curve and calculated as $IP = \int_{t=0}^{t=n} f(t)dt$).

### 2.2.2 Epidemics as engines for change

Epidemics act as engines for rapid co-evolutionary change. This is because parasite-mediated selection drives the evolution of hosts and vice-versa, host-mediated selection drives the evolution of parasites. Generally, hosts evolve higher resistance (Ameline et al., 2021, 2022; Auld & Brand, 2017a; Duffy & Forde, 2009; Duffy & Sivars-Becker, 2007; Duncan et al., 2006; Gómez & Buckling, 2011; Ibrahim & Barrett, 1991; Miller & Vincent, 2008; Paplauskas et al., 2021; Thrall et al., 2012; Zbinden et al., 2008) and parasites evolve higher infectivity (Auld, Wilson, et al., 2014; Auld & Brand, 2017a; Gómez & Buckling, 2011; Paplauskas et al., 2021; Thrall et al., 2012). Also, parasites evolve towards greater virulence (Auld & Brand, 2017a), while hosts evolve to reduce the fitness impacts of parasite virulence (Zbinden et al., 2008). One study even found the evolution of parasites in response to the changing complement of host genotypes (i.e. coevolution, Paplauskas et al., 2021).

Although hosts and parasites generally evolve either higher resistance or infectivity in response to an epidemic, this is not always the case. Sometimes theory predicts the evolution of more susceptible hosts (Boots et al., 2009; Boots & Haraguchi, 1999;

Bowers et al., 1994; Duffy & Forde, 2009; Koskella, 2018). For example, theory can predict the evolution of greater host susceptibility when selection directly favours reproduction over resistance. During small epidemics of parasites with low virulence, the benefits of higher fecundity outweigh higher resistance (Donnelly et al., 2015). For intermediate-sized epidemics, the survival benefits of resistance begin to outweigh the benefits of higher fecundity (Donnelly et al., 2015) and for large epidemics, higher fecundity and reduced resistance is most favorable again because the prevalence is so high that the survival benefits of resistance are vastly reduced (Donnelly et al., 2015). This has been supported by empirical assessment using a *Daphnia*-parasite system (Walsman et al., 2023)

Ecological context can also influence epidemic size and the evolution of host susceptibility in *Daphnia*. For example, epidemics are smaller in lakes with low productivity and high predation so hosts evolve higher fecundity and lower resistance, whereas epidemics are larger in lakes with high productivity and low predation so hosts evolve lower fecundity and higher resistance (Duffy et al., 2012).

There are several other examples of increased host susceptibility following an epidemic (Auld & Brand, 2017a; Mitchell et al., 2004; Parker, 1991; Strauss et al., 2017; Thrall et al., 2012), but there are few examples of decreased parasite infectivity (but see Boots and Mealor, 2007 for a decrease in parasite infectivity during experimental coevolution). This is most likely a reflection of host and parasite generation times, which are much shorter for parasites and so they are expected to be better adapted more often than hosts (Schmid-Hempel, 2011).

### 2.2.3 Does epidemic size determine the strength of selection?

In theory, the size and severity of epidemics determine the level of selection on host and parasite populations. We conducted a meta-analysis to examine the relationship between epidemic size and host-parasite coevolution (Box 2.1, Fig. 2.2). We expected that changes in host resistance, parasite infectivity and coevolution would increase with epidemic size as the strength of host and parasite-mediated selection would also increase.

## Box 2.1. Meta-analysis data collection

A meta-analysis was performed on studies of host-parasite coevolution and epidemic size. Specifically, we pooled the data from different studies which measured the change in host resistance and parasite infectivity from infection assays involving an experimental time-shift, which compared ancestral and evolved hosts to the ancestral parasite or *vice-versa*, which compared ancestral and evolved parasites to the ancestral host, or, as in one study, compared evolved hosts to the ancestral and evolved parasite, and linked this to the change in genotype frequency data where available and the size of epidemics, defined as rapid increases in the proportion infected over a relatively short period, measured as integrated prevalence (proportion infected over time in days), to perform our own analysis (*sensu* Curran & Hussong, 2009).

Relevant studies were searched for using Google scholar on 7th of March 2023. The search terms included “epidemic size” AND (“host evolution” OR “evolution of hosts” OR “parasite evolution” OR “evolution of parasites”), “rapid” AND “coevolution*” and “epidemic” AND “daphnia”, which returned approximately 275,000 results. However, preliminary analysis showed that most of these studies were not relevant, so only the first 50 from each search term were used for subsequent analysis (total = 150). Analysis of titles and abstracts indicated that 101 of these might include the appropriate data. Reading these studies in full showed this data was available for 10 of them and was extracted either from plots, using Plot Digitizer (http://plotdigitizer.sourceforge.net), or calculated from the raw data.

Also, for an additional comparison, we took the data on the change in transmission rate from the evolution of parasites in response to a changing complement of host genotypes (i.e. coevolution) from Paplauskas et al., 2021 and plotted this against epidemic size.

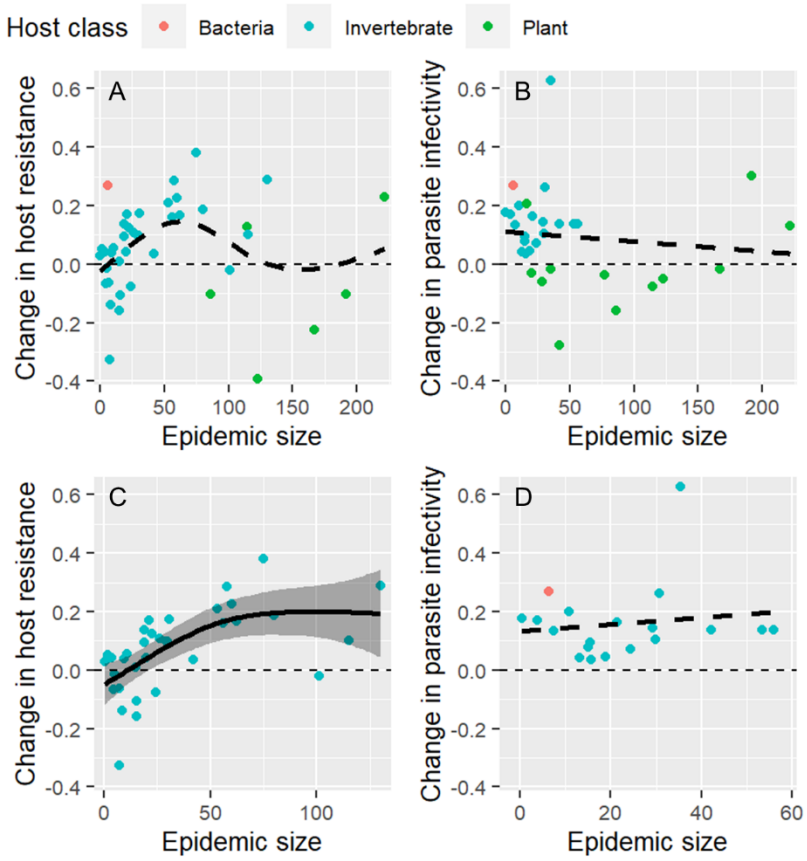



Figure 2.2. The relationship between epidemic size and host-parasite coevolution. The first two panels show (A) change in host resistance and (B) parasite infectivity in response to epidemic size across all three host classes. The next two panels show the relationship between epidemic size and either (C) the change in invertebrate host resistance or (D) the change in parasite infectivity (of invertebrate and bacterial hosts). The colour of the points indicates the host class. The thin dashed line is a reference point for positive and negative change. The thick solid and dashed black lines show the significant ($P<0.05$) and non-significant relationships between host-parasite (co)evolution and epidemic size respectively. Shaded bands denote 95% confidence intervals.

Contrary to our hypothesis, the relationship between either the change in host resistance or parasite infectivity and epidemic size was not significant across all three bacteria, invertebrate and plant host classes (generalised additive model [GAM]: $F=2.50$, $P=0.06$; Fig. 2.2A and linear model [LM]: $t=-0.65$, $P=0.52$; Fig. 2.2B). However, the relationship between either the change in host resistance or parasite infectivity and epidemic size seemed to vary with the host class. Specifically, change in host resistance initially increased with epidemic size for the bacteria and invertebrate host classes and then returned to zero for the plant host class. In

817 comparison, the negative relationship between change in parasite infectivity and
818 epidemic size seemed to be driven by the plant host class.
819
820 When examining a reduced dataset of only the invertebrate host class, the
821 relationship between the change in host resistance and epidemic size was significant
822 (GAM: $F=7.475$, $P<0.001$; Fig. 2.2C). This showed that host resistance increased
823 with epidemic size and then plateaued. In comparison, parasite infectivity increased
824 with epidemic size across a reduced dataset including the bacteria and host classes,
825 but this was not significant (LM: $t=0.59$, $P=0.57$; Fig. 2.2D).
826
827 Although there was clear no relationship between change in host resistance and
828 epidemic size, the differences in the change in host resistance across the three host
829 classes reflects host generation times. There was a large increase in bacterial
830 resistance for a relatively small epidemic because of their short generation times,
831 whereas invertebrate resistance increased more steadily with epidemic size due to
832 intermediate generation times and changes in plant resistance were much more
833 variable because of their long generation times. Some of the hosts evolved higher
834 susceptibility, particularly in the plant host class. This seemingly counter-intuitive
835 pattern of non-adaptive evolution has previously been shown in an annual legume
836 (Parker, 1991) and could be attributable to negative frequency dependent selection
837 (Thrall et al., 2012). Similarly, this could explain why so many of the parasites of plant
838 hosts were found to have evolved lower infectivity in our meta-analysis, but it is often
839 assumed that many plant host-parasite systems coevolve through directional
840 selection (e.g. Zhong et al., 2016).
841
842 The difference in the results between the full and reduced datasets reflects the
843 asymmetry in host-parasite coevolution. Specifically, when examining the reduced
844 datasets, the host resistance increased significantly with epidemic size because
845 parasite-mediated selection increases with the proportion of infected individuals,
846 whereas the change in parasite infectivity was always positive and not significantly
847 associated with epidemic size because parasites are expected to die if they fail to
848 infect (Salathé et al., 2008), so they are under stronger selection to infect than the
849 host is to resist regardless of epidemic size. The change in invertebrate resistance
850 plateaus at larger epidemics which is probably because there is limited genetic
851 variation for resistance, despite stronger parasite-mediated selection.
852

Two studies which measured change in transmission rate were not included in the meta-analysis because they either used the same data as another study already included in the meta-analysis (Paplauskas et al., 2021) or because change in transmission rate data could not be directly compared to change measured from infection assays (Strauss et al., 2017). One of these studies supports the earlier results, showing that epidemics tend to increase or decrease the transmission rate owing to either host or parasite evolution (Paplauskas et al., 2021), whereas the other study shows something different (Strauss et al., 2017), but this is possibly because epidemics were very small (integrated prevalence < 11). A theoretical study which focused on changes in transmission rate found that both costs associated with resistance and ecological context, in terms of nutrient availability, can drive the evolution of greater host susceptibly (Walsman et al., 2023).

One study also measured the relationship between change in parasite virulence or host susceptibility to it and epidemic size (Auld & Brand, 2017a). Parasites evolved to produce more spores regardless of epidemic size, whereas host susceptibility to parasite virulence increased with epidemic size. Again, this reflects the asymmetry of host-parasite coevolution. Another study found that parasite virulence measured in terms of host lifespan and number of clutches did not change over the course of an epidemic, but the parasite evolved to produce fewer spores (Gowler et al., 2022). It was suggested that the reduction in spore yield could have been a result of tradeoffs associated with parasite growth.

Most significantly, another study found that sexual recombination results in genetic slippage, genetic change in the direction contrary to selection (Lynch & Deng, 1994), which restores host susceptibility in natural populations following bouts of parasite-mediated selection (Ameline et al., 2022). This has the potential to disrupt the disease cycle, as this weakens the link between past and future epidemics, but the production of sexual resting stages which avoid selection by parasites until the following season is a unique phenomenon which is unlikely to be replicated in other organisms.

### 2.2.4 Host-parasite coevolution in nature

The context in which host-parasite coevolution occurs affects how accurately selection is measured. We can determine the *potential* for infectious disease epidemics to select on host and parasite populations using laboratory experiments

(Strauss et al., 2017; Walsman et al., 2023), but the extent to which epidemics drive *actual* host and parasite evolutionary change can only be measured in natural or semi-natural environments where other forces of selection are at play (Paplauskas et al., 2021). The importance of studying host-parasite coevolution in natural environments is reflected in the increasing number of studies published on this topic (Fig. 2.3). This allows us to measure not only the strength, but also the direction of selection. For example, a study in replicate populations of *Daphnia* showed that each population followed a unique coevolutionary trajectory, but the level of divergence between populations from a shared ancestral origin could be explained by differences in environmental conditions (Paplauskas et al., 2021).

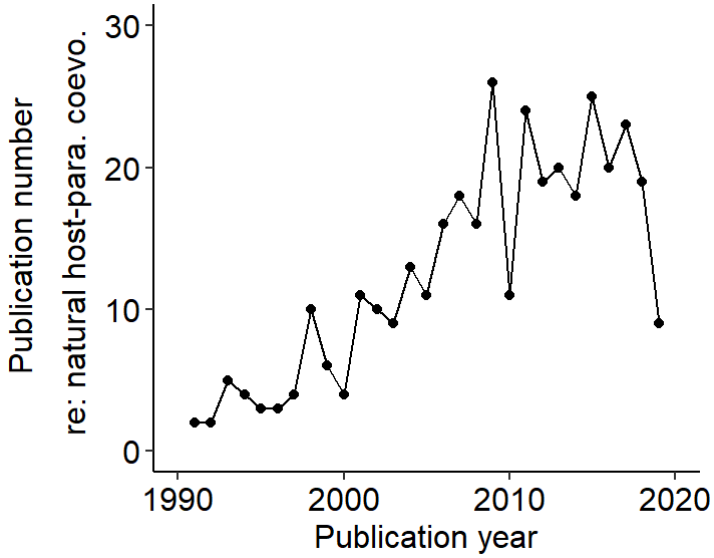



Figure 2.3. The amount of papers published on natural host-parasite coevolution by publication year. On 17th February 2020, the Thomas Reuter's Web of Science portal was used to perform the analysis based on the following search terms, TOPIC: (host* parasite* coevolution natural) NOT TOPIC: ("natural selection"). There was a total of 363 records across all fields shown. The word 'natural' was commonly used to describe essential features of the study design such as natural environments (Gómez & Buckling, 2011), epidemics (Thrall et al., 2012) and populations (Hite et al., 2017).

## 2.3 How does the mode of coevolution shape host and parasite genetic diversity?

For many host-parasite systems, the nature of selection depends on the infection genetics of the system and shapes both host and parasite genetic diversity (Fig. 2.4).

A low level of genetic specificity (e.g. the gene-for-gene model, Thompson & Burdon, 1992; Sasaki, 2000), where parasites can infect multiple hosts and hosts can resist multiple parasites, leads to directional selection for the evolution of increased host resistance and parasite infectivity through a series of selective sweeps, which is referred to as arms-race dynamics (ARD) and decreases genetic diversity over time (Buckling & Rainey, 2002; Obbard et al., 2011). In comparison, a high level of genetic specificity, where infection depends on matching host and parasite genotypes (e.g. the matching allele model, Luijckx et al., 2013; Bento et al., 2017), drives negative frequency dependent selection, where parasite-mediated selection against common hosts causes parasite genotype frequencies to track host genotype frequencies over time, which can be called fluctuating selection dynamics (FSD, Levin, 1988; Koskella & Lively, 2009) or Red Queen dynamics (RQD, Van Valen, 1973; Decaestecker et al., 2007) and maintains genetic diversity.

The tempo of coevolution depends on the nature of selection. ARD should generally lead to a slower rate of coevolution as directional selection strips genetic variation from populations (Anderson et al., 2017; Elena et al., 1996), but many studies of arms-races come from bacteria-phage populations where the rate of coevolution is already high (Brockhurst et al., 2003, 2007; Buckling & Rainey, 2002; Paterson et al., 2010). According to the Red Queen hypothesis, the reciprocal nature of selection between hosts and parasites should accelerate evolutionary rates through the need for continual adaptation and counter-adaptation. Empirical studies in snail-trematode and *Daphnia*-parasite systems suggest this may be the case by showing rapid coevolution between hosts and parasites (Decaestecker et al., 2007; Koskella & Lively, 2009), but a comparison to evolutionary rates in hosts and parasites when evolved in isolation would help to confirm this (Paterson et al., 2010).

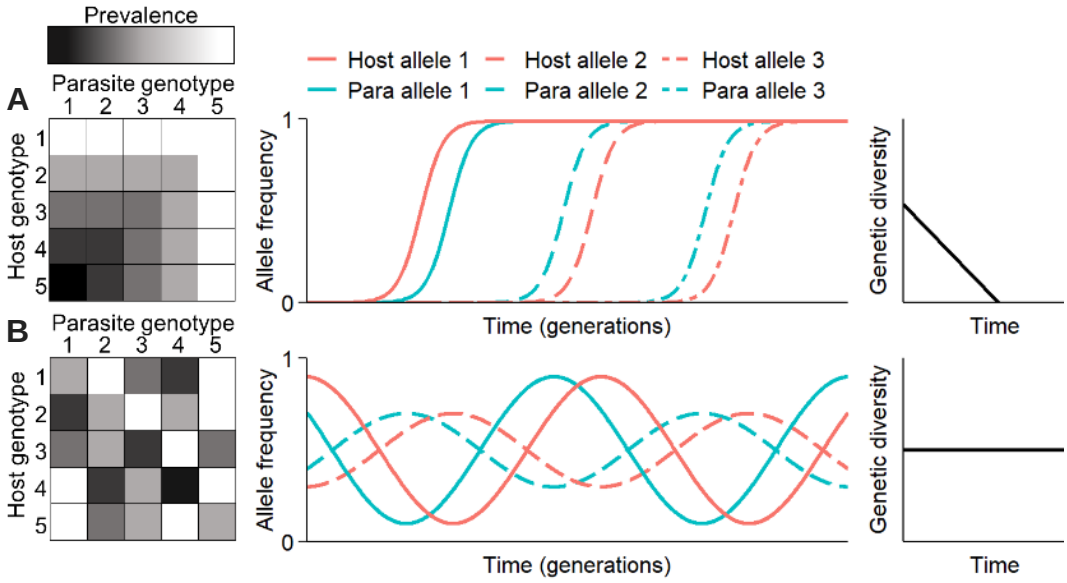



Figure 2.4. The relationship between infection genetics, coevolutionary dynamics and change in genetic diversity. The tables with coloured squares represent the outcomes from two hypothetical cross-infection experiments, where samples of host and parasites genotypes from the same population have been crossed using a factorial design and the proportion of hosts that became infected (infection prevalence) was measured for all possible pairwise combinations of host and parasite genotypes. In population A, there is a low level of genetic specificity that drives arms-race dynamics (ARD) and this leads to the loss of genetic diversity over time. In contrast, there is a high level of genetic specificity in population B that drives fluctuating selection dynamics (FSD) and genetic diversity is maintained over time. The following abbreviation was used; Parasite (Para).

### 2.3.1 A coevolutionary continuum

ARD and RQD are two ends of a coevolutionary continuum (Agrawal & Lively 2002, Engelstädter et al., 2009). Different host-parasite systems will vary in where they fall along this continuum. In reality, extreme cases may not even exist (Luijckx et al., 2013b; Schmid-Hempel et al., 1999; Thompson & Burdon, 1992) and there is some evidence for other types of parasite-mediated selection, including directional selection for increased host susceptibility (Duffy & Forde, 2009), stabilising selection (which favours an intermediate level of host resistance, Duffy & Forde, 2009) and disruptive selection (which favours highly resistant and highly susceptible host genotypes, Duffy & Forde, 2009). Several studies have shown that coevolving host and parasite populations can experience multiple modes of selection (Frickel et al., 2016; A. R. Hall et al., 2011; Masri et al., 2015; Papkou et al., 2019) and the mode

of coevolution can vary between populations of functionally similar species (Betts et al., 2014) and even replicate populations (Kortright et al., 2022). However, we know relatively little of how this continuum is influenced by other factors, such as environmental variation.

Studies performed in more realistic environments show the potentially significant impact that natural conditions have on the mode of coevolution. For example, experimental coevolution of a bacteria-phage system is known to follow ARD under controlled conditions (Gómez & Buckling, 2011), but in soil microcosms it follows FSD. Changes in the environment, such as higher nutrient availability and population mixing, can drive shifts from FSD back to ARD (Gómez et al., 2014; Lopez Pascua et al., 2014). Similarly, mixing outdoor pond populations of *Daphnia* disrupts FSD and causes adaptation of parasites to hosts of intermediate frequency (Auld & Brand, 2017a). The temporal nature of the environmental change can matter too. In bacteria-phage populations, rapidly fluctuating environments constrain coevolutionary arms races by impeding selective sweeps (Harrison et al., 2013) and temperature fluctuations drive host and pathogen populations into and out of coevolutionary cold and hot spots (Duncan et al., 2017).

Variation in the biotic environment, in terms of the presence of microbiota, coinfections and parasite diversity will also influence coevolutionary dynamics. For example, in a recent study of nematodes colonized by protective bacteria, there was reduced dominance of fluctuating selection dynamics in protected compared to unprotected host populations (Rafaluk-Mohr et al., 2022). For coinfections, where a host is infected with multiple parasites, theory predicts enhanced fluctuating selection dynamics when they increase fitness costs, but this depends on parasite characteristics, such as fecundity and virulence (Seppälä et al., 2020). However, we propose that the extent to which coinfections change the mode of coevolution may depend on the level of parasite genetic diversity between infections. If groups of similar parasites cluster together within hosts, coinfecting parasites will select against similar host genotypes and RQD dynamics will still occur. If there is no clustering of parasite genotypes within hosts, there will be low genetic specificity, hosts will be selected for general resistance and ARD will dominate. However, more empirical studies are required to test this hypothesis.

As for parasite diversity, one study found that increases in parasite diversity drove shifts in the mode of selection from fluctuating (Red Queen) dynamics to predominately directional (arms race) dynamics (Betts et al., 2018). In another study, phage populations evolved in isolation with bacteria showed increased phage infectivity and bacterial resistance through time, but two phage genotypes did not lead to an increase in bacterial resistance. This was most likely due to the inability of bacteria to evolve resistance to both phages via the same mutations and suggests that increasing initial parasite genotypic diversity can give parasites an evolutionary advantage that arrests long-term coevolution (Castledine et al., 2022).

Furthermore, different stages of the infection process, which comprises multiple steps, could be subject to different selection dynamics (Agrawal & Lively, 2003; Duneau et al., 2011a; Fenton, Antonovics & Brockhurst, 2012). For example, certain stages of the infection process are more likely to require specific matching between host and parasite genotypes, such as host cell recognition, location of target tissues and attachment of microparasites to hosts, and therefore we would expect these traits to be governed by FSD. In contrast, host exploitation (Fenton, 2012), spore activation and host entry may require a low level of genetic specificity and therefore we would expect these traits to be governed by ARD. Also, preinfeciton may facilitate the subsequent penetration of hosts by other parasites, driving lower specificity (Gopko et al., 2018).

## 2.4 The effect of host (or parasite) population-level genetic diversity on future epidemic size

Previously referred to only as ‘conventional wisdom’ (sensu (King & Lively, 2012)), the generality of the effect of low genetic diversity on the propensity for host populations, such as crop fields composed of a single species (monocultures), to experience larger or more severe parasite outbreaks (referred to as the ‘monoculture effect’ (Browning & Frey, 1969)) beyond agriculture was only recently studied (Ekroth et al., 2019; Gibson & Nguyen, 2021).

Despite a lack of studies measuring integrated epidemic size, rather than various other metrics of parasite success in terms of disease spread (such as snapshot prevalence, that only captures the proportion infected at a single point in time, or mean prevalence, etc.), which would have enabled the precise quantification of the increase in epidemic size linked to an increased level of host, or parasite, population-

level genetic diversity (the third link in the Disease Cycle), the importance of their work shows that host population genetic diversity does indeed have a 'conventional' effect on mean parasite success (but see chapter four (Paplauskas et al., 2024)). However, their rationale for why there a clearly defined relationship between the host population genetic diversity and epidemic size is not present in the current literature is not clear (Ekroth et al., 2019).

Primarily, they suggested that there could be variation in host density across populations from different studies, arising from potentially reduced host range due to habitat fragmentation, which could make it difficult to separate the relative effects of host density and population genetic diversity on disease (Ekroth et al., 2019; King & Lively, 2012). However, a study in bumblebees found that increased genetic diversity reduced disease prevalence and the effect of genetic diversity was much larger than colony density (Parsche & Lattorff, 2018). Although uncontrolled host density may be a possible reason why a compelling diversity-disease relationship is lacking in animal host studies, there are other, potentially more compelling, reasons why this could be the case. For example, the principal idea cited in the past is that the virulence and the presence of an infection depends on how disease interacts with other stressors, such as abiotic aspects of the environment (temperature, resource availability, etc.) and therefore, these additional stressors drive variation in how host population genetic diversity influence parasite infection success (O'Brien & Evermann, 1988). Alternatively, I suggest host genetic diversity may be lower due to parasite-mediated selection, rather than inbreeding, and therefore we might expect greater resistance (assuming that the chance of a host becoming infected relies on a combination of specific and non-specific factors). However, even more significant is that incomparable measures of host population diversity seem to be employed across different studies. For example, a reduction in population-level host genetic diversity as a result of inbreeding (Acevedo-Whitehouse et al., 2003) is very different to a reduction caused by hunting (O'Brien et al., 1985; Roelke et al., 1993) or habitat fragmentation (Belasen et al., 2019). This is because hunting reduces genetic diversity by imposing strong directional selection for morphological (Pigeon et al., 2016) and behavioural (Leclerc et al., 2019) traits or by significantly reducing population size (Allendorf et al., 2008), whereas inbreeding leads to a reduction in genetic diversity by mainly increasing homozygosity (Charlesworth & Meagher, 2003) and habitat loss (or fragmentation) increases the spatial separation between different sub-populations (Cushman, 2006; Leidner & Haddad, 2011) and potentially

may lead to reductions in gene flow and the overall genetic diversity (Aguilar et al., 2008; Frankham, 2005; Honnay & Jacquemyn, 2007).

### 2.4.1 Parasite diversity

Although it has received less attention than the level of host diversity, the level of parasite diversity is another key factor which influences the spread of disease. Theory predicts that evolution in a diverse parasite population leads to epidemiological feedbacks and when parasite-mediated selection is strong, this facilitates the spread of disease (Lively, 2016). Empirical studies tend to focus on the effect of parasite diversity on individual infections (Davies et al., 2002; De Roode et al., 2005). There have been relatively few studies of the effect of parasite diversity on population-level measures of disease. Since disease risk is based on some level of specificity between hosts and parasites, we would expect parasites with higher diversity to spread more rapidly through a host population due to the increased likelihood of encountering a host they are adapted to. One study which measured population-level effects of disease found that the effect of host genetic diversity on the spread of disease depends on the level of genetic diversity in the parasite population. They found that parasite prevalence increased with the number of parasite strains and host monocultures exposed to several parasite strains had higher mean parasite prevalence and higher variance than polycultures (Ganz & Ebert, 2010). Other studies suggest that parasites may also facilitate one another by compromising the host immune system (Karvonen et al., 2011). However, more studies are needed in other disease systems to better understand the generality of these results.

### 2.4.2 The identity of host and parasite genotypes

Another factor which influences the spread of disease that has received relatively little attention is the identity of the host and parasite genotypes. Controlled laboratory experiments have shown that the identity of the host and/or pathogen genotype(s) explain much of the variation in the likelihood of infection (over 44% in the *Daphnia magna-Pasteuria ramosa* freshwater host-pathogen system: Vale et al., 2009 Heredity), and that these effects of genotype can further interact with environmental variables in many host-pathogen systems (Echaubard et al., 2014; Lazzaro et al., 2008; Meixner et al., 2014; Vale & Little, 2009).

### 2.4.3 The genetic basis for infection

A fundamental knowledge gap is that we often do not know which traits, or genes, underlie host and parasite diversity for resistance and infectivity (Ebert, 2018; Ebert & Fields, 2020). However, recent studies have begun to address this gap. For example, a study of coevolution in the nematode, *Caenorhabditis elegans*, and its bacterial parasite has shown genomic changes in a parasite toxin gene in response to selection (Papkou et al., 2019). In another study, coevolution in a bacteria-phage community drove the diversification of CRISPR immunity (Guillemet et al., 2022). Lastly, there has been strong evidence for a gene governing infectivity which provides a molecular basis for study of Red Queen dynamics in the *Daphnia* model system (Andras et al., 2020). Future work should aim to continue uncovering the diverse range of traits for which is there is variation in host resistance and parasite infectivity to answer questions such as; How many genes are involved in host–parasite interactions, and how are they organized in the genome (Ebert, 2018)? How do they interact, and how specific are these interactions (Ebert, 2018)? What form of selection operates on the genes (Ebert, 2018)?

## 2.5 Summary

As a result of infection, disease can have several negative impacts on host populations, including reduced genetic diversity, depressed population size and complete extinction (Alan Pounds et al., 2006; Boots & Sasaki, 2002; Vredenburg et al., 2010). To protect populations in an era of broad environmental change, we require disease control strategies, and the effective design of such strategies relies on a detailed understanding of the various drivers of disease and some capacity to predict outbreaks in the future.

We presented a simple conceptual model, the Disease Cycle, to bridge the gap between the evolutionary ecology of past and future disease outbreaks in a variable world. First, we considered epidemics as drivers of host-parasite coevolution. Epidemics generally increase host resistance and parasite infectivity, and the strength of parasite-mediated selection depends on epidemic size. In comparison, the lack of any relationship between parasite evolution and epidemic size reflects the asymmetry of coevolution. The shift towards studies of coevolution in natural environments reflects the importance of measuring the extent to which epidemics drive actual coevolutionary change. Previous research has focused on parasite

rather than host-mediated selection so more theoretical and empirical studies are required to address this gap.

Second, we considered how the mode of coevolution shapes host and parasite genetic diversity. A low level of genetic specificity leads to arms-race dynamics and the loss of genetic diversity over time, whereas a high level of genetic specificity leads to red-queen dynamics and maintains genetic diversity over time. In reality, these represent two ends of a coevolutionary continuum and where a particular interaction falls along this continuum depends on both biotic and abiotic features of the environment, and the specific stage of infection considered. Short-term studies of coevolution, such as those using bacteria and phages, often show rapid changes in host resistance and parasite infectivity over relatively short time-scales (Brockhurst et al., 2003, 2007; Buckling & Rainey, 2002; Paterson et al., 2010), but the extent to which these findings represent non-model organisms, which possess a much lower potential for evolution, is uncertain. More studies in non-model organisms are required to demonstrate the potential for coevolution to drive rapid, short-term change. On the other hand, long-term studies of coevolutionary responses are relatively rare and tend to focus on host plant-pathogen associations (Soubeyrand et al., 2009; Thrall et al., 2012; Susi and Laine, 2015; Ericson, Müller and Burdon, 2017; but see Dewald-Wang et al., 2022). To what extent coevolutionary dynamics are observable over the short-term (single epidemic) compared to the long term (multi-epidemic) still remains uncertain.

Third and finally, we considered how host and parasite genetic diversity affect future epidemic size. Plant populations with higher genetic diversity are at less risk of the more harmful effects of disease. Although the generality of this relationship outside agricultural systems is unclear, recent evidence suggests that genetic diversity also protect animals from disease. On the other hand, theoretical and empirical evidence suggests that parasite diversity generally increases disease risk. The identity of the host and parasite genotypes is also important. Fundamental knowledge gaps include how genetic diversity affects variation in the level of disease and which traits underlie host and parasite diversity for resistance and infectivity.

Despite the potential for the Disease Cycle model to provide a theoretical framework to link the size of past and future epidemics, I acknowledge that 1) this mainly applies to microparasites (bacteria, viruses etc.) versus macroparasites (nematodes, etc.)

due to the ability of microparasites to induce a rapid increase in the number of infected individuals over a short space of time (such that it meets a threshold for an 'epidemic', Hudson et al., 2002), and 2) this mainly applies to invertebrate versus vertebrate hosts (i.e. those that have innate (Little et al., 2003) versus acquired immunity (Babayan et al., 2011)). In the latter case, this is because vertebrate acquired immunity is a fundamental mechanism that determines infection rate. In support of this, studies in natural host-parasite associations, such as wild rodents and their suite of parasites species (including nematodes, viruses and blood-borne bacteria, etc.), show that antibodies and coinfection drive variation in parasite burdens (Clerc et al., 2018). In addition to these considerations, there may be times when a cycle of host-parasite coevolution is overshadowed by the interactions between host and parasite ecology. In this sense, there may be times at which the ecological theatre matters more than the (co)-evolutionary play (in the sense of (Hutchinson, 1965)). Indeed, contemporary research shows that within-host interactions are often crucial for determining the fitness and transmissibility of co-infecting parasites (Pedersen & Fenton, 2007).

Overall, we hope that this model could provide a framework for future modelling approaches that embrace epidemic disease as a recurrent episodic process and help better inform the forecasting and management of disease control strategies.

## 2.6 References


Acevedo-Whitehouse, K., Gulland, F., Greig, D., & Amos, W. (2003). Disease susceptibility in California sea lions. *Nature*, 422(6927), 35. https://doi.org/10.1038/422035a

Aguilar, R., Quesada, M., Ashworth, L., Herrerias-Diego, Y., & Lobo, J. (2008). Genetic consequences of habitat fragmentation in plant populations: Susceptible signals in plant traits and methodological approaches. In *Molecular Ecology* (Vol. 17, Issue 24). https://doi.org/10.1111/j.1365-294X.2008.03971.x

Alan Pounds, J., Bustamante, M. R., Coloma, L. A., Consuegra, J. A., Fogden, M. P. L., Foster, P. N., La Marca, E., Masters, K. L., Merino-Viteri, A., Puschendorf, R., Ron, S. R., Sánchez-Azofeifa, G. A., Still, C. J., & Young, B. E. (2006). Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature*, *439*(7073), 161–167. https://doi.org/10.1038/nature04246

Allendorf, F. W., England, P. R., Luikart, G., Ritchie, P. A., & Ryman, N. (2008). Genetic effects of harvest on wild animal populations. In *Trends in Ecology and Evolution* (Vol. 23, Issue 6). https://doi.org/10.1016/j.tree.2008.02.008

Altizer, S., Dobson, A., Hosseini, P., Hudson, P., Pascual, M., & Rohani, P. (2006). Seasonality and the dynamics of infectious diseases. *Ecology Letters*, *9*(4), 467–484. https://doi.org/10.1111/j.1461-0248.2005.00879.x

Altizer, S., Harvell, D., & Friedle, E. (2003). Rapid evolutionary dynamics and disease threats to biodiversity. *Trends in Ecology & Evolution*, *18*(11), 589–596. https://doi.org/10.1016/J.TREE.2003.08.013

Ameline, C., Bourgeois, Y., Vögtli, F., Savola, E., Andras, J., Engelstädter, J., & Ebert, D. (2021). A two-locus system with strong epistasis underlies rapid parasite-mediated evolution of host resistance. *Molecular Biology and Evolution*, *38*(4). https://doi.org/10.1093/molbev/msaa311

Ameline, C., Voegtli, F., Andras, J., Dexter, E., Engelstädter, J., & Ebert, D. (2022). Genetic slippage after sex maintains diversity for parasite resistance in a natural host population. In *Sci. Adv* (Vol. 8). https://www.science.org

Anderson, T. J. C., Nair, S., McDew-White, M., Cheeseman, I. H., Nkhoma, S., Bilgic, F., McGready, R., Ashley, E., Phyo, A. P., White, N. J., & Nosten, F. (2017). Population parameters underlying an ongoing soft sweep in southeast asian malaria parasites. *Molecular Biology and Evolution*, *34*(1). https://doi.org/10.1093/molbev/msw228

Andras, J. P., Fields, P. D., Pasquier, L. Du, Fredericksen, M., & Ebert, D. (2020). Genome-wide association analysis identifies a genetic basis of infectivity in a model bacterial pathogen. *Molecular Biology and Evolution*, *37*(12). https://doi.org/10.1093/molbev/msaa173

Auld, S. K. J. R., & Brand, J. (2017). Environmental variation causes different (co) evolutionary routes to the same adaptive destination across parasite populations. *Evolution Letters*, *1*(5), 245–254. https://doi.org/10.1002/evl3.27

Auld, S. K. J. R., Hall, S. R., Ochs, J. H., Sebastian, M., & Duffy, M. A. (2014). Predators and patterns of within-host growth can mediate both among-host competition and evolution of transmission potential of parasites. *American Naturalist*, *184*(SUPPL. 1), S77–S90. https://doi.org/10.1086/676927

Auld, S. K. J. R., Wilson, P. J., & Little, T. J. (2014). Rapid change in parasite infection traits over the course of an epidemic in a wild host-parasite population. *Oikos*, *123*(2), 232–238. https://doi.org/10.1111/j.1600-0706.2013.00720.x

Babayan, S. A., Allen, J. E., Bradley, J. E., Geuking, M. B., Graham, A. L., Grencis, R. K., Kaufman, J., McCoy, K. D., Paterson, S., Smith, K. G. C., Turnbaugh, P. J., Viney, M. E., Maizels, R. M., & Pedersen, A. B. (2011). Wild immunology: converging on the real world. *Annals of the New York Academy of Sciences*, *1236*(1), 17–29. https://doi.org/10.1111/j.1749-6632.2011.06251.x

Begon, M., Telfer, S., Smith, M. J., Burthe, S., Paterson, S., & Lambin, X. (2009). Seasonal host dynamics drive the timing of recurrent epidemics in a wildlife population. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1662), 1603–1610. https://doi.org/10.1098/rspb.2008.1732

Belasen, A. M., Bletz, M. C., Leite, D. da S., Toledo, L. F., & James, T. Y. (2019). Long-term habitat fragmentation is associated with reduced MHC IIB diversity and increased infections in amphibian hosts. *Frontiers in Ecology and Evolution*, *6*(JAN). https://doi.org/10.3389/fevo.2018.00236

Bento, G., Routtu, J., Fields, P. D., Bourgeois, Y., Du Pasquier, L., & Ebert, D. (2017). The genetic basis of resistance and matching-allele interactions of a host-parasite system: The *Daphnia magna-Pasteuria ramosa* model. *PLOS Genetics*, *13*(2), e1006596. https://doi.org/10.1371/journal.pgen.1006596

Betts, A., Kaltz, O., & Hochberg, M. E. (2014). Contrasted coevolutionary dynamics between a bacterial pathogen and its bacteriophages. *Proceedings of the National Academy of Sciences*, *111*(30), 11109–11114. https://doi.org/10.1073/pnas.1406763111

Boots, M., Best, A., Miller, M. R., & White, A. (2009). The role of ecological feedbacks in the evolution of host defence: What does theory tell us? *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2008.0160

Boots, M., & Haraguchi, Y. (1999). The evolution of costly resistance in host-parasite systems. *Source: The American Naturalist*, *153*(4), 359–370. https://doi.org/10.1086/303181

Boots, M., & Mealor, M. (2007). Local interactions select for lower pathogen infectivity. *Science*. https://doi.org/10.1126/science.1137126

Boots, M., & Sasaki, A. (2002). Parasite-driven extinction in spatially explicit host-parasite systems. *American Naturalist*, *159*(6), 706–713. https://doi.org/10.1086/339996

Bowers, R. G., Boots, M., & Begon, M. (1994). Life-history trade-offs and the evolution of pathogen resistance: Competition between host strains. *Proceedings of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rspb.1994.0122

Brockhurst, M. A., Morgan, A. D., Fenton, A., & Buckling, A. (2007). Experimental coevolution with bacteria and phage. The *Pseudomonas fluorescens*-Φ2 model system. In *Infection, Genetics and Evolution* (Vol. 7, Issue 4). https://doi.org/10.1016/j.meegid.2007.01.005

Brockhurst, M. A., Morgan, A. D., Rainey, P. B., & Buckling, A. (2003). Population mixing accelerates coevolution. *Ecology Letters*, *6*(11). https://doi.org/10.1046/j.1461-0248.2003.00531.x

Browning, J. A., & Frey, K. J. (1969). Multiline cultivars as a means of disease control. *Annual Review of Phytopathology*, *7*(1), 355–382. https://doi.org/10.1146/annurev.py.07.090169.002035

Buckling, A., & Rainey, P. B. (2002). Antagonistic coevolution between a bacterium and a bacteriophage. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1494), 931–936. https://doi.org/10.1098/rspb.2001.1945

Cáceres, C. E., Davis, G., Duple, S., Hall, S. R., Koss, A., Lee, P., & Rapti, Z. (2014). Complex *Daphnia* interactions with parasites and competitors. *Mathematical Biosciences*, *258*. https://doi.org/10.1016/j.mbs.2014.10.002

Charlesworth, D., & Meagher, T. R. (2003). Effects of inbreeding on the genetic diversity of populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1434). https://doi.org/10.1098/rstb.2003.1296

Civitello, D. J., Penczykowski, R. M., Smith, A. N., Shocket, M. S., Duffy, M. A., & Hall, S. R. (2015). Resources, key traits and the size of fungal epidemics in *Daphnia* populations. *Journal of Animal Ecology*, *84*(4), 1010–1017. https://doi.org/10.1111/1365-2656.12363

Clerc, M., Devevey, G., Fenton, A., & Pedersen, A. B. (2018). Antibodies and coinfection drive variation in nematode burdens in wild mice. *International Journal for Parasitology*, *48*(9–10), 785–792. https://doi.org/10.1016/j.ijpara.2018.04.003

Curran, P. J., & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods*, *14*(2). https://doi.org/10.1037/a0015914

Cushman, S. A. (2006). Effects of habitat loss and fragmentation on amphibians: A review and prospectus. *Biological Conservation*, *128*(2). https://doi.org/10.1016/j.biocon.2005.09.031

Davies, C. M., Fairbrother, E., & Webster, J. P. (2002). Mixed strain schistosome infections of snails and the evolution of parasite virulence. *Parasitology*, *124*(1). https://doi.org/10.1017/S0031182001008873

De Roode, J. C., Pansini, R., Cheesman, S. J., Helinski, M. E. H., Huijben, S., Wargo, A. R., Bell, A. S., Chan, B. H. K., Walliker, D., & Read, A. F. (2005). Virulence and competitive ability in genetically diverse malaria infections. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(21). https://doi.org/10.1073/pnas.0500078102

Decaestecker, E., Gaba, S., Raeymaekers, J. A. M., Stoks, R., Van Kerckhoven, L., Ebert, D., & De Meester, L. (2007). Host–parasite 'Red Queen' dynamics archived in pond sediment. *Nature*, *450*(7171), 870–873. https://doi.org/10.1038/nature06291

Dewald-Wang, E. A., Parr, N., Tiley, K., Lee, A., & Koskella, B. (2022). Multiyear time-shift study of bacteria and phage dynamics in the phyllosphere. *American Naturalist*, *199*(1). https://doi.org/10.1086/717181

Dicker, R. C. (2006). Principles of epidemiology in public health practice. *Cdc*.

Dohm, D. J., O'Guinn, M. L., & Turell, M. J. (2002). Effect of environmental temperature on the ability of *Culex pipiens* (Diptera: Culicidae) to transmit West Nile virus. *Journal of Medical Entomology*, *39*(1). https://doi.org/10.1603/0022-2585-39.1.221

Donnelly, R., White, A., & Boots, M. (2015). The epidemiological feedbacks critical to the evolution of host immunity. *Journal of Evolutionary Biology*, *28*(11). https://doi.org/10.1111/jeb.12719

Duffy, M. A., Cáceres, C. E., & Hall, S. R. (2019). Healthy herds or predator spreaders? Insights from the plankton into how predators suppress and spread disease. In *Wildlife Disease Ecology: Linking Theory to Data and Application*. https://doi.org/10.1017/9781316479964.016

Duffy, M. A., & Forde, S. E. (2009). Ecological feedbacks and the evolution of resistance. *Journal of Animal Ecology*, *78*(6), 1106–1112. https://doi.org/10.1111/j.1365-2656.2009.01568.x

Duffy, M. A., Ochs, J. H., Penczykowski, R. M., Civitello, D. J., Klausmeier, C. A., & Hall, S. R. (2012). Ecological context influences epidemic size and parasite-driven evolution. *Science*, *335*(6076), 1636–1638. https://doi.org/10.1126/science.1215429

Duffy, M. A., & Sivars-Becker, L. (2007). Rapid evolution and ecological host-parasite dynamics. *Ecology Letters*, *10*(1), 44–53.

Duncan, A. B., Dusi, E., Jacob, F., Ramsayer, J., Hochberg, M. E., & Kaltz, O. (2017). Hot spots become cold spots: coevolution in variable temperature environments. *Journal of Evolutionary Biology*, *30*(1), 55–65. https://doi.org/10.1111/jeb.12985

Duncan, A. B., Mitchell, S. E., & Little, T. J. (2006). Parasite-mediated selection and the role of sex and diapause in *Daphnia* . *Journal of Evolutionary Biology*, *19*(4), 1183–1189. https://doi.org/10.1111/j.1420-9101.2006.01085.x

Duneau, D., Luijckx, P., Ben-Ami, F., Laforsch, C., & Ebert, D. (2011). Resolving the infection process reveals striking differences in the contribution of environment, genetics and phylogeny to host-parasite interactions. *BMC Biology*, *9*, 11. https://doi.org/10.1186/1741-7007-9-11

Ebert, D. (2018). Open questions: what are the genes underlying antagonistic coevolution? *BMC Biology*, *16*(1), 114. https://doi.org/10.1186/s12915-018-0583-7

Ebert, D., & Fields, P. D. (2020). Host–parasite co-evolution and its genomic signature. In *Nature Reviews Genetics* (Vol. 21, Issue 12). https://doi.org/10.1038/s41576-020-0269-1

Echaubard, P., Leduc, J., Pauli, B., Chinchar, V. G., Robert, J., & Lesbarrères, D. (2014). Environmental dependency of amphibian-ranavirus genotypic interactions: Evolutionary perspectives on infectious diseases. *Evolutionary Applications*, *7*(7). https://doi.org/10.1111/eva.12169

Ekroth, A. K. E., Rafaluk-Mohr, C., & King, K. C. (2019). Host genetic diversity limits parasite success beyond agricultural systems: a meta-analysis. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1911), 20191811. https://doi.org/10.1098/rspb.2019.1811

Elena, S. F., Cooper, V. S., & Lenski, R. E. (1996). Punctuated evolution caused by selection of rare beneficial mutations. *Science*, *272*(5269). https://doi.org/10.1126/science.272.5269.1802

Engebretsen, S., Engø-Monsen, K., Frigessi, A., & De Blasio, B. F. (2019). A theoretical single-parameter model for urbanisation to study infectious disease spread and interventions. *PLoS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1006879

Engelstädter, J., & Bonhoeffer, S. (2009). Red queen dynamics with non-standard fitness interactions. *PLoS Computational Biology*, *5*(8), e1000469. https://doi.org/10.1371/journal.pcbi.1000469

Ericson, L., Müller, W. J., & Burdon, J. J. (2017). 28-year temporal sequence of epidemic dynamics in a natural rust–host plant metapopulation. *Journal of Ecology*. https://doi.org/10.1111/1365-2745.12720

Fels, D., & Kaltz, O. (2006). Temperature-dependent transmission and latency of *Holospora undulata*, a micronucleus-specific parasite of the ciliate *Paramecium caudatum*. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1589). https://doi.org/10.1098/rspb.2005.3404

Frankham, R. (2005). Genetics and extinction. In *Biological Conservation* (Vol. 126, Issue 2). https://doi.org/10.1016/j.biocon.2005.05.002

Frickel, J., Sieber, M., & Becks, L. (2016). Eco-evolutionary dynamics in a coevolving host-virus system. In *Ecology Letters* (Vol. 19, Issue 4). https://doi.org/10.1111/ele.12580

Ganz, H. H., & Ebert, D. (2010). Benefits of host genetic diversity for resistance to infection depend on parasite diversity. In *Source: Ecology* (Vol. 91, Issue 5). https://www-jstor-org.ezproxy.stir.ac.uk/stable/pdf/25661177.pdf?refreqid=excelsior%3A5c0f30d81cb9568626bbc1823fa625af

Gibson, A. K., & Nguyen, A. E. (2020). Does genetic diversity protect host populations from parasites? A meta-analysis across natural and agricultural systems. *Evolution Letters*, *5*(1). https://doi.org/10.1002/evl3.206

Gómez, P., Ashby, B., & Buckling, A. (2015). Population mixing promotes arms race host-parasite coevolution. *Proceedings of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rspb.2014.2297

Gómez, P., & Buckling, A. (2011). Bacteria-phage antagonistic coevolution in soil. *Science*, *332*(6025), 106–109. https://doi.org/10.1126/science.1198767

Gopko, M., Chowdhury, M. M. R., & Taskinen, J. (2018). Interactions between two parasites of brown trout ( *Salmo trutta* ): Consequences of preinfection. *Ecology and Evolution*, *8*(20), 9986–9997. https://doi.org/10.1002/ece3.4406

Gowler, C. D., Essington, H., O'Brien, B., Shaw, C. L., Bilich, R. W., Clay, P. A., & Duffy, M. A. (2022). Virulence evolution during a naturally occurring parasite outbreak. *Evolutionary Ecology*. https://doi.org/10.1007/s10682-022-10169-6

Guillemet, M., Chabas, H., Nicot, A., Gatchich, F., Ortega-Abboud, E., Buus, C., Hindhede, L., Rousseau, G. M., Bataillon, T., Moineau, S., & Gandon, S. (2022). Competition and coevolution drive the evolution and the diversification of CRISPR immunity. *Nature Ecology and Evolution*, *6*(10). https://doi.org/10.1038/s41559-022-01841-9

Gutierrez, S. O., Minchella, D. J., & Bernal, X. E. (2022). Survival of the sickest: selective predation differentially modulates ecological and evolutionary disease dynamics. *Oikos*, *2022*(9). https://doi.org/10.1111/oik.09126

Hall, A. R., Scanlan, P. D., Morgan, A. D., & Buckling, A. (2011). Host-parasite coevolutionary arms races give way to fluctuating selection. *Ecology Letters*, *14*(7), 635–642. https://doi.org/10.1111/j.1461-0248.2011.01624.x

Hall, S. R., Becker, C. R., Simonis, J. L., Duffy, M. A., Tessier, A. J., & Cáceres, C. E. (2009). Friendly competition: evidence for a dilution effect among competitors in a planktonic host–parasite system. *Ecology*, *90*(3), 791–801. https://doi.org/10.1890/08-0838.1

Hall, S. R., Duffy, M. A., & Cáceres, C. E. (2005). Selective predation and productivity jointly drive complex behavior in host-parasite systems. *American Naturalist*, *165*(1). https://doi.org/10.1086/426601

Harrison, E., Laine, A.-L., Hietala, M., & Brockhurst, M. A. (2013). Rapidly fluctuating environments constrain coevolutionary arms races by impeding selective sweeps. *Proc. R. Soc. B*, *280*(1764), 20130937. https://doi.org/10.1098/RSPB.2013.0937

Hite, J. L., Penczykowski, R. M., Shocket, M. S., Griebel, K. A., Strauss, A. T., Duffy, M. A., Cáceres, C. E., & Hall, S. R. (2017). Allocation, not male resistance, increases male frequency during epidemics: a case study in facultatively sexual hosts. *Ecology*. https://doi.org/10.1002/ecy.1976

Honnay, O., & Jacquemyn, H. (2007). Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conservation Biology*, *21*(3). https://doi.org/10.1111/j.1523-1739.2006.00646.x

Hudson, P. J., Rizzoli, A., Grenfell, B. T., Heesterbeek, H., & Dobson, A. P. (2002). *The Ecology of Wildlife Diseases*. Oxford University Press Oxford. https://doi.org/10.1093/oso/9780198506201.001.0001

Hutchinson, G. E. (1965). *The ecological theater and the evolutionary play*. Yale University Press.

Ibelings, B. W., De Bruin, A., Kagami, M., Rijkeboer, M., Brehm, M., & Donk, E. Van. (2004). Host parasite interactions between freshwater phytoplankton and chytrid fungi (Chytridiomycota). *Journal of Phycology*, *40*(3), 437–453. https://doi.org/10.1111/j.1529-8817.2004.03117.x

Ibrahim, K. M., & Barrett, J. A. (1991). Evolution of mildew resistance in a hybrid bulk population of barley. *Heredity*, *67*(2), 247–256. https://doi.org/10.1038/hdy.1991.86

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, *451*(7181), 990–993. https://doi.org/10.1038/nature06536

Karvonen, A., Rellstab, C., Louhi, K. R., & Jokela, J. (2011). Synchronous attack is advantageous: Mixed genotype infections lead to higher infection success in trematode parasites. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1726). https://doi.org/10.1098/rspb.2011.0879

King, K. C., & Lively, C. M. (2012). Does genetic diversity limit disease spread in natural host populations? *Heredity*, *109*(4), 199–203. https://doi.org/10.1038/hdy.2012.33

Koprivnikar, J., & Poulin, R. (2009). Interspecific and intraspecific variation in cercariae release. *Journal of Parasitology*, *95*(1). https://doi.org/10.1645/GE-1582.1

Kortright, K. E., Chan, B. K., Evans, B. R., & Turner, P. E. (2022). Arms race and fluctuating selection dynamics in *Pseudomonas aeruginosa* bacteria coevolving with phage OMKO1. *Journal of Evolutionary Biology*, *35*(11), 1475–1487. https://doi.org/10.1111/jeb.14095

Koskella, B. (2018). Resistance gained, resistance lost: An explanation for host-parasite coexistence. *PLoS Biology*, *16*(9). https://doi.org/10.1371/journal.pbio.3000013

Koskella, B., & Lively, C. M. (2009). Evidence for negative frequency-dependent selection during experimental coevolution of a freshwater snail and a sterilizing trematode. *Evolution*, *63*(9), 2213–2221. https://doi.org/10.1111/j.1558-5646.2009.00711.x

Lafferty, K. D., & Mordecai, E. A. (2016). The rise and fall of infectious disease in a warmer world. *F1000Research*, *5*. https://doi.org/10.12688/f1000research.8766.1

Lazzaro, B. P., Flores, H. A., Lorigan, J. G., & Yourth, C. P. (2008). Genotype-by-environment interactions and adaptation to local temperature affect immunity and fecundity in Drosophila melanogaster. *PLoS Pathogens*. https://doi.org/10.1371/journal.ppat.1000025

Leclerc, M., Zedrosser, A., Swenson, J. E., & Pelletier, F. (2019). Hunters select for behavioral traits in a large carnivore. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-48853-3

Leidner, A. K., & Haddad, N. M. (2011). Combining measures of dispersal to identify conservation strategies in fragmented landscapes. *Conservation Biology*, *25*(5). https://doi.org/10.1111/j.1523-1739.2011.01720.x

Levin, B. R. (1988). Frequency-dependent selection in bacterial populations. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. https://doi.org/10.1098/rstb.1988.0059

Little, T. J., O'Connor, B., Colegrave, N., Watt, K., & Read, A. F. (2003). Maternal transfer of strain-specific immunity in an invertebrate. *Current Biology*, *13*(6), 489–492. https://doi.org/10.1016/S0960-9822(03)00163-5

Lively, C. M. (2016). Coevolutionary epidemiology: Disease spread, local adaptation, and sex. *American Naturalist*, *187*(3). https://doi.org/10.1086/684626

Lopez Pascua, L., Hall, A. R., Best, A., Morgan, A. D., Boots, M., & Buckling, A. (2014). Higher resources decrease fluctuating selection during host-parasite coevolution. *Ecology Letters*, *17*(11), 1380–1388. https://doi.org/10.1111/ele.12337

Luijckx, P., Fienberg, H., Duneau, D., & Ebert, D. (2013). A matching-allele model explains host resistance to parasites. *Current Biology*, *23*(12), 1085–1088. https://doi.org/10.1016/j.cub.2013.04.064

Lynch, M., & Deng, H. W. (1994). Genetic slippage in response to sex. *American Naturalist*, *144*(2). https://doi.org/10.1086/285673

Masri, L., Branca, A., Sheppard, A. E., Papkou, A., Laehnemann, D., Guenther, P. S., Prahl, S., Saebelfeld, M., Hollensteiner, J., Liesegang, H., Brzuszkiewicz, E., Daniel, R., Michiels, N. K., Schulte, R. D., Kurtz, J., Rosenstiel, P., Telschow, A., Bornberg-Bauer, E., & Schulenburg, H. (2015). Host–pathogen coevolution: The selective advantage of *Bacillus thuringiensis* virulence and its cry toxin genes. *PLoS Biology*, *13*(6). https://doi.org/10.1371/journal.pbio.1002169

Meixner, M. D., Francis, R. M., Gajda, A., Kryger, P., Andonov, S., Uzunov, A., Topolska, G., Costa, C., Amiri, E., Berg, S., Bienkowska, M., Bouga, M., Büchler, R., Dyrba, W., Gurgulova, K., Hatjina, F., Ivanova, E., Janes, M., Kezic, N., Wilde, J. (2014). Occurrence of parasites and pathogens in honey bee colonies used in a European genotype environment interactions experiment. *Journal of Apicultural Research*, *53*(2). https://doi.org/10.3896/IBRA.1.53.2.04

Miller, M. P., & Vincent, E. R. (2008). Rapid natural selection for resistance to an introduced parasite of rainbow trout. *Evolutionary Applications*, *1*(2). https://doi.org/10.1111/j.1752-4571.2008.00018.x

Mitchell, S. E., Read, A. F., & Little, T. J. (2004). The effect of a pathogen epidemic on the genetic structure and reproductive strategy of the crustacean *Daphnia magna*. *Ecology Letters*. https://doi.org/10.1111/j.1461-0248.2004.00639.x

Mordecai, E. A., Caldwell, J. M., Grossman, M. K., Lippi, C. A., Johnson, L. R., Neira, M., Rohr, J. R., Ryan, S. J., Savage, V., Shocket, M. S., Sippy, R., Stewart Ibarra, A. M., Thomas, M. B., & Villena, O. (2019). Thermal biology of mosquito-borne disease. *Ecology Letters*, *22*(10), 1690–1708. https://doi.org/10.1111/ele.13335

Munson, L., Terio, K. A., Kock, R., Mlengeya, T., Roelke, M. E., Dubovi, E., Summers, B., Sinclair, A. R. E., & Packer, C. (2008). Climate extremes promote fatal co-infections during canine distemper epidemics in African lions. *PLoS ONE*, *3*(6). https://doi.org/10.1371/journal.pone.0002545

Obbard, D. J., Jiggins, F. M., Bradshaw, N. J., & Little, T. J. (2011). Recent and recurrent selective sweeps of the antiviral RNAi gene Argonaute-2 in three

species of *Drosophila*. *Molecular Biology and Evolution*, *28*(2), 1043–1056. https://doi.org/10.1093/molbev/msq280

O'Brien, S. J., & Evermann, J. F. (1988). Interactive influence of infectious disease and genetic diversity in natural populations. In *Trends in Ecology and Evolution* (Vol. 3, Issue 10). https://doi.org/10.1016/0169-5347(88)90058-4

O'Brien, S. J., Roelke, M. E., Marker, L., Newman, A., Winkler, C. A., Meltzer, D., Colly, L., Evermann, J. F., Bush, M., & Wildt, D. E. (1985). Genetic basis for species vulnerability in the Cheetah. *Science*, *227*(4693), 1428–1434. https://doi.org/10.1126/science.2983425

Ohm, J. R., Baldini, F., Barreaux, P., Lefevre, T., Lynch, P. A., Suh, E., Whitehead, S. A., & Thomas, M. B. (2018). Rethinking the extrinsic incubation period of malaria parasites. In *Parasites and Vectors* (Vol. 11, Issue 1). https://doi.org/10.1186/s13071-018-2761-4

Paaijmans, K. P., Blanford, S., Chan, B. H. K., & Thomas, M. B. (2012). Warmer temperatures reduce the vectorial capacity of malaria mosquitoes. *Biology Letters*, *8*(3). https://doi.org/10.1098/rsbl.2011.1075

Papkou, A., Guzella, T., Yang, W., Koepper, S., Pees, B., Schalkowski, R., Barg, M. C., Rosenstiel, P. C., Teotónio, H., & Schulenburg, H. (2019). The genomic basis of red queen dynamics during rapid reciprocal host–pathogen coevolution. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(3), 923–928. https://doi.org/10.1073/pnas.1810402116

Paplauskas, S., Brand, J., & Auld, S. K. J. R. (2021). Ecology directs host–parasite coevolutionary trajectories across *Daphnia*–microparasite populations. *Nature Ecology & Evolution*, *5*(4), 480–486. https://doi.org/10.1038/s41559-021-01390-7

Paplauskas, S., Duthie, B., & Tinsley, M. C. (2024). The effect of host population genetic diversity on variation in metrics of parasite success. *BioRxiv*. https://doi.org/https://doi.org/10.1101/2024.05.28.596150

Parker, M. A. (1991). Nonadaptive evolution of disease resistance in an annual legume. *Evolution*, *45*(5), 1209–1217. https://doi.org/10.1111/j.1558-5646.1991.tb04387.x

Parsche, S., & Lattorff, M. H. G. (2018). The relative contributions of host density and genetic diversity on prevalence of a multi-host parasite in bumblebees.

*Biological Journal of the Linnean Society*, *125*(4). https://doi.org/10.1093/biolinnean/bly151

Paterson, S., Vogwill, T., Buckling, A., Benmayor, R., Spiers, A. J., Thomson, N. R., Quail, M., Smith, F., Walker, D., Libberton, B., Fenton, A., Hall, N., & Brockhurst, M. A. (2010). Antagonistic coevolution accelerates molecular evolution. *Nature*, *464*(7286), 275–278. https://doi.org/10.1038/nature08798

Paull, S. H., Lafonte, B. E., & Johnson, P. T. J. (2012). Temperature-driven shifts in a host-parasite interaction drive nonlinear changes in disease risk. *Global Change Biology*, *18*(12). https://doi.org/10.1111/gcb.12018

Pedersen, A. B., & Fenton, A. (2007). Emphasizing the ecology in parasite community ecology. *Trends in Ecology & Evolution*, *22*(3), 133–139. https://doi.org/10.1016/j.tree.2006.11.005

Penczykowski, R. M., Laine, A.-L., & Koskella, B. (2016). Understanding the ecology and evolution of host-parasite interactions across scales. *Evolutionary Applications*, *9*(1), 37–52. https://doi.org/10.1111/eva.12294

Pigeon, G., Festa-Bianchet, M., Coltman, D. W., & Pelletier, F. (2016). Intense selective hunting leads to artificial evolution in horn size. *Evolutionary Applications*, *9*(4). https://doi.org/10.1111/eva.12358

Piotrowski, J. S., Annis, S. L., & Longcore, J. E. (2004). Physiology of *Batrachochytrium dendrobatidis*, a chytrid pathogen of amphibians. *Mycologia*, *96*(1). https://doi.org/10.1080/15572536.2005.11832990

Poulin, R. (2006). Global warming and temperature-mediated increases in cercarial emergence in trematode parasites. *Parasitology*, *132*(1). https://doi.org/10.1017/S0031182005008693

Price, S. J., Leung, W. T. M., Owen, C. J., Puschendorf, R., Sergeant, C., Cunningham, A. A., Balloux, F., Garner, T. W. J., & Nichols, R. A. (2019). Effects of historic and projected climate change on the range and impacts of an emerging wildlife disease. *Global Change Biology*. https://doi.org/10.1111/gcb.14651

Rafaluk-Mohr, C., Gerth, M., Sealey, J. E., Ekroth, A. K. E., Aboobaker, A. A., Kloock, A., & King, K. C. (2022). Microbial protection favors parasite tolerance and alters host-parasite coevolutionary dynamics. *Current Biology*, *32*(7). https://doi.org/10.1016/j.cub.2022.01.063

Read, A. F. (1994). The evolution of virulence. *Trends in Microbiology*, *2*(3), 73–76. https://doi.org/10.1016/0966-842X(94)90537-1

Refardt, D. (2011). Within-host competition determines reproductive success of temperate bacteriophages. *ISME Journal*, *5*(9). https://doi.org/10.1038/ismej.2011.30

Reliefweb. (2008). Glossary of Humanitarian Terms. In *Reliefweb*.

Roelke, M. E., Martenson, J. S., & O'Brien, S. J. (1993). The consequences of demographic reduction and genetic depletion in the endangered Florida panther. *Current Biology*, *3*(6), 340–350. https://doi.org/10.1016/0960-9822(93)90197-V

Salathé, M., Kouyos, R. D., Regoes, R. R., & Bonhoeffer, S. (2008). Rapid parasite adaptation drives selection for high recombination rates. *Evolution*, *62*(2). https://doi.org/10.1111/j.1558-5646.2007.00265.x

Sasaki, A. (2000). Host-parasite coevolution in a multilocus gene-for-gene system. *Proceedings of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rspb.2000.1267

Schmeller, D. S., Courchamp, F., & Killeen, G. (2020). Biodiversity loss, emerging pathogens and human health risks. In *Biodiversity and Conservation*. https://doi.org/10.1007/s10531-020-02021-6

Schmid-Hempel, P. (2011). Evolutionary parasitology: The integrated study of infections, immunology, ecology and genetics. In *Trends in parasitology*. https://doi.org/10.1016/j.pt.2011.03.004

Schmid-Hempel, P., Puhr, K., Kruger, N., Reber, C., & Schmid-Hempel, R. (1999). Dynamic and genetic consequences of variation in horizontal transmission for a microparasitic infection. *Evolution*. https://doi.org/10.2307/2640779

Seppälä, O., Lively, C. M., & Jokela, J. (2020). Coinfecting parasites can modify fluctuating selection dynamics in host–parasite coevolution. *Ecology and Evolution*, *10*(18). https://doi.org/10.1002/ece3.6373

Shocket, M. S., Magnante, A., Duffy, M. A., Cáceres, C. E., & Hall, S. R. (2019). Can hot temperatures limit disease transmission? A test of mechanisms in a zooplankton–fungus system. *Functional Ecology*, *33*(10), 2017–2029. https://doi.org/10.1111/1365-2435.13403

Shocket, M. S., Vergara, D., Sickbert, A. J., Walsman, J. M., Strauss, A. T., Hite, J. L., Duffy, M. A., Cáceres, C. E., & Hall, S. R. (2018). Parasite rearing and

infection temperatures jointly influence disease transmission and shape seasonality of epidemics. *Ecology*, *99*(9), 1975–1987. https://doi.org/10.1002/ecy.2430

Soubeyrand, S., Laine, A. L., Hanski, I., & Penttinen, A. (2009). Spatiotemporal structure of host-pathogen interactions in a metapopulation. *American Naturalist*. https://doi.org/10.1086/603624

Strauss, A. T., Bowling, A. M., Duffy, M. A., Cáceres, C. E., & Hall, S. R. (2018). Linking host traits, interactions with competitors and disease: Mechanistic foundations for disease dilution. *Functional Ecology*, *32*(5), 1271–1279. https://doi.org/10.1111/1365-2435.13066

Strauss, A. T., Hite, J. L., Shocket, M. S., Cáceres, C. E., Duffy, M. A., & Hall, S. R. (2017). Rapid evolution rescues hosts from competition and disease but—despite a dilution effect—increases the density of infected hosts. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1868), 20171970. https://doi.org/10.1098/rspb.2017.1970

Susi, H., & Laine, A.-L. (2015). The effectiveness and costs of pathogen resistance strategies in a perennial plant. *Journal of Ecology*, *103*(2), 303–315. https://doi.org/10.1111/1365-2745.12373

Thieltges, D. W., Bordalo, M. D., Caballero Hernández, A., Prinz, K., & Jensen, K. T. (2008). Ambient fauna impairs parasite transmission in a marine parasite-host system. *Parasitology*, *135*(9). https://doi.org/10.1017/S0031182008004526

Thompson, J. N., & Burdon, J. J. (1992). Gene-for-gene coevolution between plants and parasites. *Nature*, *360*(6400), 121–125. https://doi.org/10.1038/360121a0

Thrall, P. H., Laine, A.-L., Ravensdale, M., Nemri, A., Dodds, P. N., Barrett, L. G., & Burdon, J. J. (2012). Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. *Ecology Letters*, *15*(5), 425–435. https://doi.org/10.1111/j.1461-0248.2012.01749.x

Vale, P. F., & Little, T. J. (2009). Measuring parasite fitness under genetic and thermal variation. *Heredity*. https://doi.org/10.1038/hdy.2009.54

Van Valen, L. (1973). A new evolutionary law. *Evol. Theory*, *1*, 1–30.

Vredenburg, V. T., Knapp, R. A., Tunstall, T. S., & Briggs, C. J. (2010). Dynamics of an emerging disease drive large-scale amphibian population

extinctions. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(21), 9689–9694. https://doi.org/10.1073/pnas.0914111107

Walsman, J. C., Duffy, M. A., Cáceres, C. E., & Hall, S. R. (2021). “Resistance is futile”: Weaker selection for resistance by abundant parasites increases prevalence and depresses host density. *American Naturalist*, 201(6), 864-879.

White, R. J., & Razgour, O. (2020). Emerging zoonotic diseases originating in mammals: a systematic review of effects of anthropogenic land-use change. In *Mammal Review*. https://doi.org/10.1111/mam.12201

Zbinden, M., Haag, C. R., & Ebert, D. (2008). Experimental evolution of field populations of Daphnia magna in response to parasite treatment. *Journal of Evolutionary Biology*, *21*(4), 1068–1078. https://doi.org/10.1111/j.1420-9101.2008.01541.x

Zhong, Z., Norvienyeku, J., Chen, M., Bao, J., Lin, L., Chen, L., Lin, Y., Wu, X., Cai, Z., Zhang, Q., Lin, X., Hong, Y., Huang, J., Xu, L., Zhang, H., Chen, L., Tang, W., Zheng, H., Chen, X., Wang, Yanli, Lian, B., Zhang, L., Tang, H., Lu, G., Ebbole, D. J., Wang, B., Wang, Z. (2016). Directional selection from host plants is a major force driving host specificity in *Magnaporthe* species. Scientific Reports, 6(1), 25591. https://doi.org/10.1038/srep25591

# 3. Ecology directs host–parasite coevolutionary trajectories across *Daphnia*–microparasite populations

This paper has been published in *Nature Ecology and Evolution*.

Conceptualization: Dr Stuart Auld; data curation: Dr Stuart Auld; formal analysis: Sam Paplauskas, and Dr Stuart Auld; funding acquisition: Dr Stuart Auld; investigation: Sam Paplauskas, J.B., and Dr Stuart Auld; methodology: Sam Paplauskas, June Brand, and Dr Stuart Auld; supervision: Dr Stuart Auld; writing original draft: Sam Paplauskas, and Dr Stuart Auld; writing, review and editing: all authors.

## 3.1 Abstract

Host-parasite interactions often fuel coevolutionary change. However, parasitism is one of a myriad of possible ecological interactions in nature. Biotic (*e.g.,* predation) and abiotic (*e.g.,* temperature) variation can amplify or dilute parasitism as a selective force on hosts and parasites, driving population variation in (co)evolutionary trajectories. We dissected the relationships between wider ecology and coevolutionary trajectory using 16 ecologically complex *Daphnia magna-Pasteuria ramosa* ponds seeded with an identical starting host (*Daphnia*) and parasite (*Pasteuria*) population. We show, using a time-shift experiment and outdoor population data, how multivariate biotic and abiotic ecological differences between ponds caused coevolutionary divergence. Wider ecology drove variation in host evolution of resistance, but not parasite infectivity; parasites subsequently coevolved in response to the changing complement of host genotypes, such that parasites adapted to historically resistant host genotypes. Parasitism was a stronger interaction for the parasite than for its host, likely because the host is the principal environment and selective force, whereas for hosts, parasite-mediated selection is one of many sources of selection. Our findings reveal the mechanisms through which wider ecology creates coevolutionary hotspots and coldspots in biologically realistic arenas of host-parasite interaction, and sheds light on how the ecological theatre can affect the (co)evolutionary play.

## 3.2 Introduction

Parasites are a strong selective force acting on host populations, and *vice versa* (Paterson et al., 2010; Schulte et al., 2010), fuelling rapid cycles of adaptation and counter-adaptation in terms of host resistance and parasite capacity to infect (Decaestecker et al., 2007; Gómez & Buckling, 2011; Koskella & Lively, 2009; Schulte et al., 2010). These coevolutionary processes can have profound effects on disease outbreaks. For example, whether the host or the parasite is ahead in the coevolutionary process can, in part, affect whether epidemics are emerging (Refardt & Ebert, 2007) or in decline (Duffy et al., 2009). A key aim of evolutionary ecologists is to understand the extent to which coevolution is: (1) a deterministic process with repeated, predictable outcomes that are either hard-wired or shaped by measurable abiotic and biotic ecological variation; and (2) a stochastic process driven by unpredictable events.

Ecological variation is known to have strong effects on coevolution (Springer, 2007; Tack et al., 2015; Wolinska & King, 2009). However, dissecting host-parasite coevolution in biologically realistic settings is fraught with difficulty, and much of our understanding of coevolution therefore comes from laboratory experiments that eliminate ecological complexity. This experimental control comes at a cost to biological realism, because parasitism is just one of many ecological interactions that hosts experience in the wild; predation, competition *etc.*, and abiotic variables such as temperature are already known to either amplify or diminish host evolutionary responses to parasite-mediated selection (Auld, Hall, et al., 2014; Auld & Brand, 2017a; Decaestecker et al., 2007; Duffy et al., 2012; Su & Boots, 2017; Wright et al., 2016). By contrast, we expect parasite evolution, particularly for obligate endoparasites, to be driven primarily by shifts in host-mediated selection caused by changes in host genotype frequencies (Auld & Tinsley, 2015), because hosts insulate their endoparasites from the wider environment. These asymmetries in host and parasite responses to reciprocal selection could create discrepancies between coevolution observed in the laboratory and in the natural arena.

We quantified how coevolutionary trajectories varied among 16 biologically realistic pond populations of *Daphnia magna* and its sterilizing bacterial endoparasite, *Pasteuria ramosa.* Each pond was initiated with an identical suite of *Daphnia* genotypes and the same starting population and dose of *Pasteuria* transmission spores, and the densities of healthy and parasite-infected were then monitored

weekly over the course of each pond epidemic. At the end of the epidemic, *Daphnia* were sampled to determine the change in genotype frequencies and additional infected *Daphnia* were sampled to obtain parasite isolates from each pond. We subsequently conducted a time-shift experiment where we exposed replicates of the original twelve *Daphnia* genotypes to either the ancestral parasite used to initiate the pond populations, or to parasite isolates collected from each pond at the end of the epidemic.

By combining data from the time-shift experiment with changes in relative genotype frequencies, we dissected, for each pond, the effects of the three components of host-parasite coevolution on the change in parasite transmission rate over the course of the season: host evolution of resistance, parasite evolution of infectivity, and coevolution (*i.e.,* the extent to which the parasite population non-additively evolved in response to a changed complement of host genotypes). When host genotypes that were resistant to the ancestral parasite increased in frequency within a population, that host population evolved host resistance; when a parasite sample collected at the end of the season caused more infections than the ancestral parasite when exposed to the panel of host genotypes, that parasite population evolved increased infectivity; and when a parasite sample collected at the end of the season became proportionately more infectious to host genotypes that were resistant to the ancestral parasite, that parasite population coevolved in response to the changing complement of host genotypes.

## 3.3 Results and Discussion

### 3.3.1 Coevolutionary trajectories varied among ponds

Whilst the ponds had the same starting populations of hosts and parasites, each pond experienced its own natural temperature profile (with significant variation across ponds), and half underwent an experimental manipulation of within-population flux (mixing) that simulated extreme precipitation events. We recorded the natural variation in 10 biotic and abiotic ecological variables over the season: temperature, pH, dissolved oxygen, chlorophyll, nitrate, and total dissolved salt, parasite prevalence, predator density and adult host density. This allowed us to examine the role of ecological variation early in the season in driving coevolutionary divergence.

We found that each pond population followed its own coevolutionary trajectory (with respect to changes in parasite transmission rate). This was driven by variation in all three coevolutionary axes: host evolution, parasite evolution and coevolution (Fig. 3.1a-c). We uncovered asymmetry in the magnitude of host and parasite evolution: parasite populations evolved more in their capacity to infect the ancestral host population than their corresponding hosts evolved capacity to resist the ancestral parasite population (paired $t$ = -3.25, $P$ = 0.005; Fig. 3.1a). We also found a strong positive relationship between the change in host resistance and coevolution, *i.e.,* a change in transmission rates due to a shifting complement of host genotypes ($r_s$ = 0.69, $P$ = 0.004; Fig. 3.1b): over the course of the season, parasites became disproportionately better at infecting those host genotypes that were previously resistant at the beginning of the season (host genotypes that had become more common), and also disproportionately poorer at infecting host genotypes that were previously susceptible at the beginning of the season (host genotypes that had become rarer). By contrast, there was a lack of relationship between the change in parasite infectivity and coevolution ($r_s$ = 0.39, $P$ = 0.135; Fig. 3.1c). These findings are consistent with the idea that ecological interactions above and beyond parasitism can select on hosts, but do not act on the host insulated parasites; shifts in host genotype frequencies instead drive parasite genetic change *via* coevolution. Whereas, for ectoparasites, which live on the host exterior, wider ecological conditions are known to shape the evolution of virulence(Cardon et al., 2011; Mahmud et al., 2017).

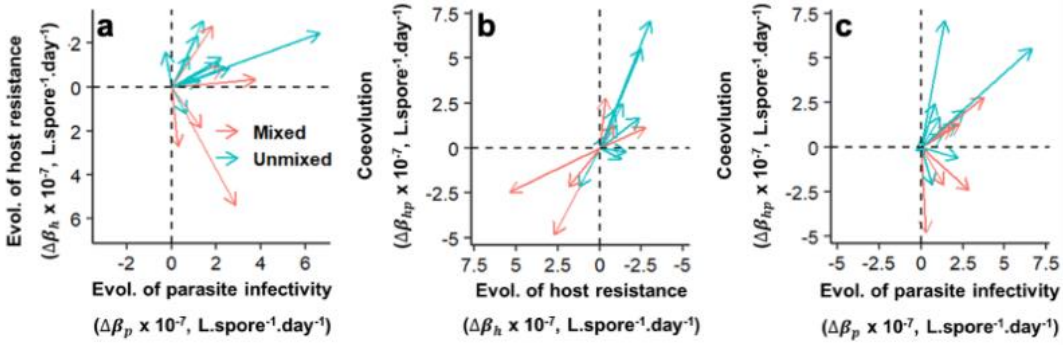



**Figure 3.1. Coevolutionary trajectories vary across populations.** Vectors show pairwise relationships between **a** change in transmission rate due to host evolution of resistance ($\Delta\beta_h$) and change in transmission rate due to parasite evolution of infectivity ( $\Delta\beta_p$), **b** host evolution of resistance ($\Delta\beta_h$) and non-additive change in transmission rate due to coevolution ($\Delta\beta_{hp}$) and **c** parasite evolution of infectivity ( $\Delta\beta_p$) and coevolution ($\Delta\beta_{hp}$). Populations were identical pre-epidemic

(vector tails) and by the end of the epidemic phenotypes had diverged due to variation in evolutionary trajectories (vector heads, open arrowheads). Red arrows denote populations that underwent the mixing treatment and blue arrows denote populations that remained unmixed.

### 3.3.2 Ecology drives variation in coevolution

Initial inspection of the ten ecological variables in isolation revealed that the mixing treatment had no effect on nine of the ten ecological variables, but that it was associated with lower total adult host densities (see Table S3.1). This supports the idea that the mixing treatment affected the ecology of the system primarily by reducing host densities directly; indeed, it is known that sediment suspension can interfere with *Daphnia* filter feeding, reducing population growth and the consumption of algae (Arruda et al., 1983) (see later results). Higher temperatures and lower chlorophyll concentration, dissolved oxygen and pH were each associated with the evolution of host resistance, but none of the ecological variables were associated with parasite evolution or coevolution (see Table S3.2).

In comparison to the initial inspection of mixing treatment and its effect on the ecological variables measured, a more holistic multivariate analysis uncovered a much more interesting story. A Principal Components Analysis of the biotic and abiotic variables (Fig. S3.1) revealed considerable ecological variation among populations, with the first and second PC axes explaining 36.0% and 21.6% of that variation. The main factors driving variation in unmixed populations were mean temperature and host density, whereas several factors explained variation in mixed populations: chlorophyll, predator density, oxygen, pH and nitrate. There was a strong positive relationship between $\delta_{eco}$the pairwise Mahalanobian distances between populations in multivariate space for ecological variation, and $\delta_{coevo}$, the pairwise Mahalanobian distances for coevolutionary net change (Fig. 3.2: Mantel $r$ = 0.36, $P$ = 0.029). Populations that were more ecologically different from each other had more divergent coevolutionary trajectories. Both theory (Mostowy & Engelstädter, 2011) and empirical data (reviewed in (Wolinska & King, 2009)) have previously shown how host and parasite genotypes can differentially respond to particular environmental variation to create (co)evolutionary hotspots and coldspots (Thompson, 2005); these results show how such environmental variables can act in concert to mediate coevolution.

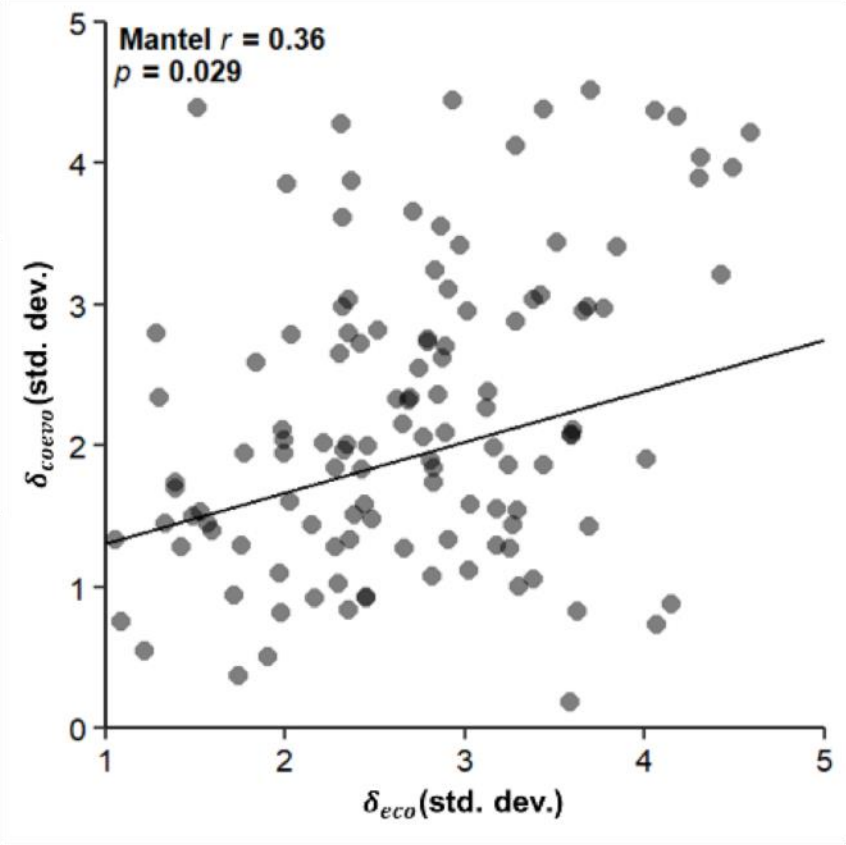



**Figure 3.2. Pairwise ecological differences explain population divergence in coevolutionary trajectory.** Relationship between pairwise population distances (measured as Mahalanobis distances) for ecology (across PC1-PC4, $\delta_{eco}$) and net coevolutionary trajectory (combining the three axes of host evolution, parasite evolution, coevolution, $\delta_{\mathrm{coevo}}$). Pairwise differences are measured in standard deviations of the total variation.

### 3.3.3 Ecology affects host evolution, with consequences for coevolution

The next step was to dissect precisely how ecological variation and coevolutionary change were linked. Using Structural Equation Modelling (SEM; Fig. S3.2), we tested which of two credible scenarios better explained the relationship between ecological and coevolutionary variation among populations (Fig. 3.3). Scenario 1 (SEM1) proposed that mixing affected ecology (measured as PC1), that ecology directly affected host evolution, parasite evolution and coevolution, and that parasite evolution also separately affected coevolution. Scenario 2 (SEM2) was similar, except it proposed that ecology did not affect coevolution directly; here ecological effects on coevolution were mediated by both host evolution and parasite evolution

(see methods section for details). Whilst both SEM1 and SEM2 both provided adequate fit to the data (SEM1: Fisher's C = 19.80, D.F. = 12, *P* = 0.071, BIC = 64.16; SEM2: Fisher's C = 12.66, D.F. = 12, *P* = 0.394, BIC = 57.02), SEM2 was the better performing model (ΔBIC = 7.14), demonstrating that there was greater support for the scenario where ecological effects on coevolution were mediated by both host evolution and parasite evolution.

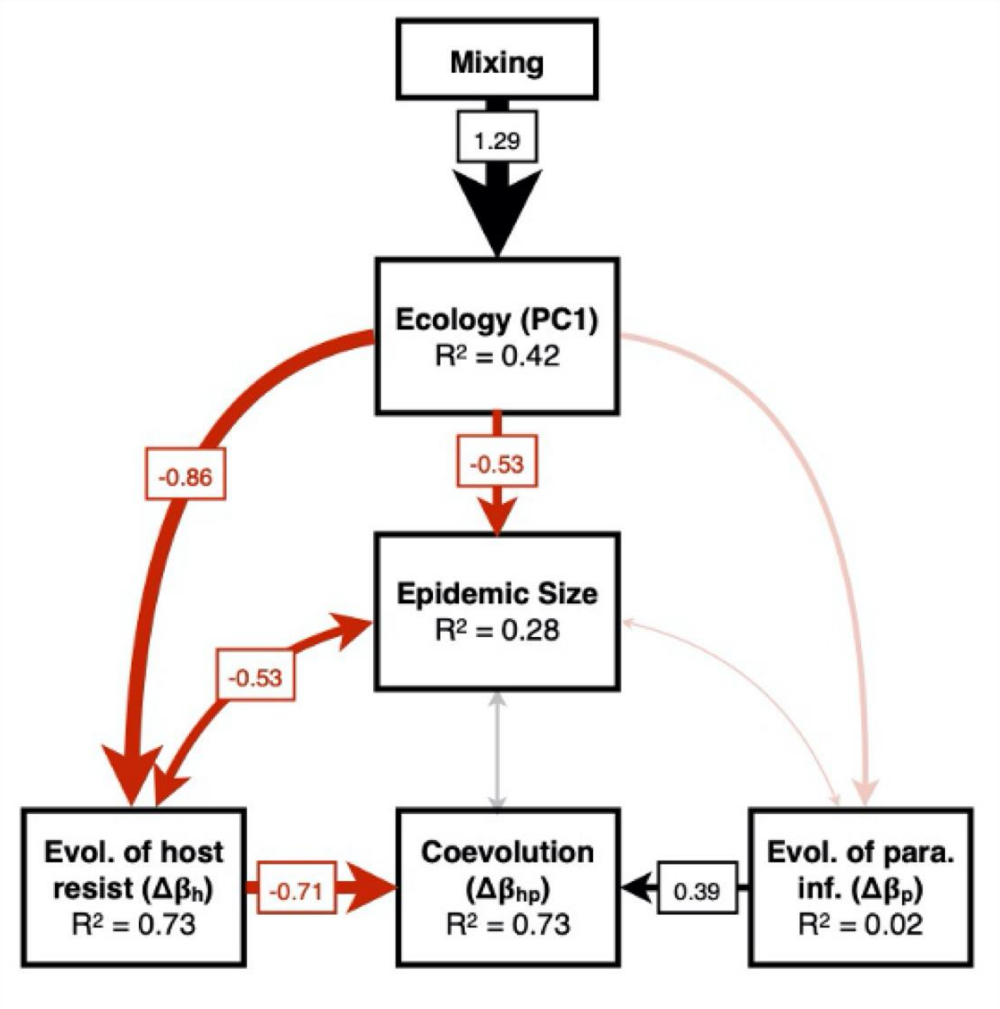



**Figure 3.3. Wider ecology drives coevolution through its effects on host evolution.** Path diagram for SEM2 showing how ecology drives coevolution. Arrows represent unidirectional (single arrowhead) or bidirectional (double arrowheads) relationships. Black arrows denote positive relationships, red arrows negative ones. Significant ($p$<0.05) and non-significant relationships are represented by solid and partially transparent arrows respectively. The arrow width of significant relationships is scaled according to the standardised regression coefficient shown in the small boxes (see also Fig. 3.4, Table S3.1). Note that negative values of $\Delta\beta_h$ represent evolution of host resistance.

Analysis of SEM2 revealed that ecological conditions, as expressed by PC1, were significantly different between mixed and unmixed populations (Fig. 3.3; Fig. 3.4a; Table S3.1), and that epidemic size was negatively associated with this measure of ecological variation (Fig. 3.4b; Table S3.1), such that epidemics were larger in populations that were warmer, had lower chlorophyll concentrations, lower pH and lower predator densities. Epidemic size was associated with the evolution of host resistance (reduced transmission rate) (Fig. 3.4c; Table S3.1), but there was no compelling evidence for an association between epidemic size and parasite infectivity (Fig. 3.4d; Table S3.1), or coevolution (Fig. 3.4e; Table S3.1). Ecology was also directly associated with evolution of host resistance (Fig. 3.4f; Table S3.1), but not parasite infectivity (Fig. 3.4g; Table S3.1). Finally, the ability to examine partial residuals after controlling for other variables (a major advantage of the SEM approach) allowed us to uncover that coevolution was positively associated with both the evolution of host resistance (Fig. 3.4h; Table S3.1) and the evolution of parasite infectivity (Fig. 3.4i; Table S3.1).

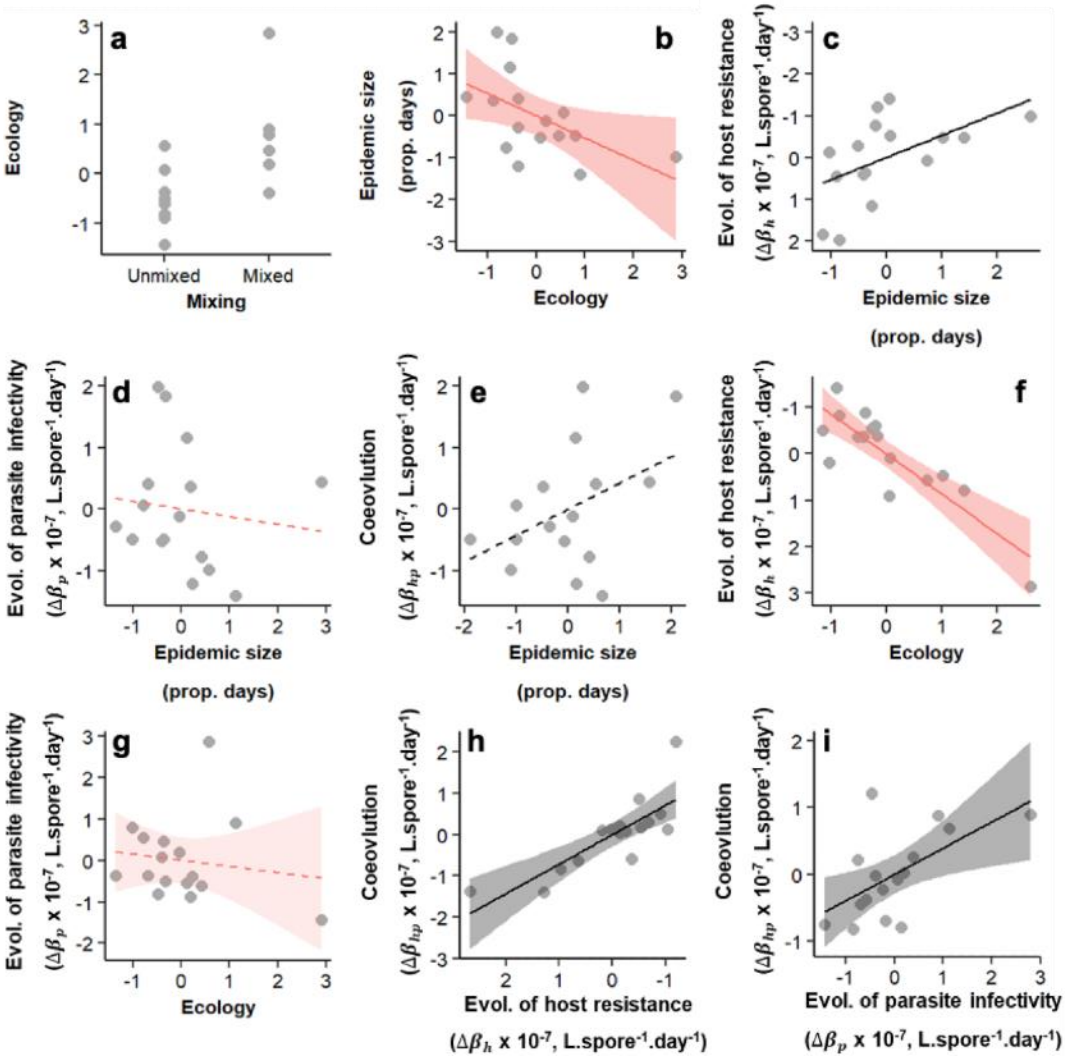


**Figure 3.4. Ecological, epidemiological and coevolutionary relationships across populations.** Relationships between variables from SEM2 **a-i**. Colours show positive (black) and negative (red) relationships, and bands denote 95% CIs. Note that negative values of $\Delta\beta_h$ represent evolution of host resistance. Significant ($p$>0.05) and non-significant relationships are indicated by solid and dashed lines respectively.

These separate effects of epidemic size and wider ecology on host (but not parasite) evolution provide two principal insights. They add support our assertion that hosts are subject to a wide range of selective pressures due to both parasite-mediated selection from disease epidemics and from wider ecology, whereas the parasite's insulation within the host environment and the obligate nature of its relationship with the host ensures the host is the principal agent of selection (hence the relationship between host evolution and coevolution). They also raise the intriguing hypothesis that epidemic size and wider ecology (driven in part by mixing treatment) pull two separate levers to drive host evolution of resistance. First, larger epidemics could have exerted greater parasite-mediated selection for host resistance (Duffy et al., 2012). Second, populations with greater PC1 values, *i.e.*, lower predation and higher temperatures and thus higher *Daphnia* reproductive rate), had high population densities (Brett, 1992)·(Goss & Bunting, 1983), and therefore likely had a greater capacity to respond to any parasite-mediated selection. This may have fuelled coevolution, driving the divergence in coevolutionary trajectories we see in Fig. 3.1.

The next step is to explain the relationships between host evolution, parasite evolution and coevolution. Previous work demonstrated the Matching Allele Model (MAM) best describes the infection genetics of the *Daphnia-Pasteuria* system (Bento et al., 2017b; Decaestecker et al., 2007; Luijckx et al., 2013a): alleles conferring parasite ability to infect one host genotype often preclude it from infecting other different host genotypes (Auld & Brand, 2017a). However, MAM in its purest sense requires just one susceptible host genotype for every infectious parasite genotype (Grosberg, 2000), but in the *Daphnia-Pasteuria* system, parasite genotypes commonly infect >1 host genotypes and also vary in the number of host genotypes each parasite can infect (Luijckx et al., 2013b). This deviation from MAM could potentially explain why coevolution was positively associated with the evolution of host resistance and, to a lesser extent, parasite infectivity (Fig. 3.4h,i; Table S3.1): parasite populations that were more infectious to the ancestral complement of hosts

were also better at infecting the new complement of hosts, and hosts that got better at resisting the ancestral parasite also got better at resisting the evolved parasite. Reciprocal selection could have acted in two ways. First, general selection could have favoured parasite genotypes that infect the broadest range of host genotypes (and *vice versa* for resistance in host genotypes), and second, specific selection could have separately favoured parasite genotypes that could infect host genotypes that had become particularly common (again, *vice versa* for resistance in hosts genotypes).

## 3.4 Conclusion

These results demonstrate that even in seemingly noisy environments, coevolution was still largely driven by deterministic, ecologically-mediated processes. Individual biotic and abiotic variables gave us a small glimpse of how wider ecology shaped coevolution. It was only after viewing multiple ecological variables from a multivariate perspective that we were able to observe that the ecological theatre determined the (co)evolutionary play in a measurable understandable way (*sensu* Hutchinson, 1965). Recent work has demonstrated that quantitative differences among qualitatively similar environments can explain evolutionary divergence among stickleback populations (Stuart et al., 2017); we show the same is true for more complex host-parasite coevolution, and that knowledge of multiple ecological conditions could help us predict the distribution of coevolutionary hotspots and coldspots (Thompson, 2005).

## 3.5 Acknowledgments

We thank Matthew Tinsley, Stephen Thackeray and the Stirling Eco-Evo group for comments on this manuscript.

## 3.6 References (main text)

Arruda, J. A., Marzolf, G. R., & Faulk, R. T. (1983). The role of suspended sediments in the nutrition of zooplankton in turbid reservoirs. *Ecology*, 64(5), 1225–1235. https://doi.org/10.2307/1937831

Auld, S. K. J. R., & Brand, J. (2017). Environmental variation causes different (co) evolutionary routes to the same adaptive destination across parasite populations. *Evolution Letters*, *1*(5), 245–254. https://doi.org/10.1002/evl3.27

Auld, S. K. J. R., Hall, S. R., Ochs, J. H., Sebastian, M., & Duffy, M. A. (2014). Predators and patterns of within-host growth can mediate both among-host

competition and evolution of transmission potential of parasites. *American Naturalist*, *184*(SUPPL. 1), S77–S90. https://doi.org/10.1086/676927

Auld, S. K. J. R., & Tinsley, M. C. (2015). The evolutionary ecology of complex lifecycle parasites: Linking phenomena with mechanisms. *Heredity*, *114*(2), 125–132. https://doi.org/10.1038/hdy.2014.84

Bento, G., Routtu, J., Fields, P. D., Bourgeois, Y., Du Pasquier, L., & Ebert, D. (2017). The genetic basis of resistance and matching-allele interactions of a host-parasite system: The *Daphnia magna*-*Pasteuria ramosa* model. *PLOS Genetics*, *13*(2), e1006596. https://doi.org/10.1371/journal.pgen.1006596

Brett, M. T. (1992). *Chaoborus* and fish-mediated influences on *Daphnia longispina* population structure, dynamics and life history strategies. *Oecologia*, *89*(1), 69–77. https://doi.org/10.1007/BF00319017

Cardon, M., Loot, G., Grenouillet, G., & Blanchet, S. (2011). Host characteristics and environmental factors differentially drive the burden and pathogenicity of an ectoparasite: A multilevel causal analysis. *Journal of Animal Ecology*, *80*(3), 657–667. https://doi.org/10.1111/j.1365-2656.2011.01804.x

Decaestecker, E., Gaba, S., Raeymaekers, J. A. M., Stoks, R., Van Kerckhoven, L., Ebert, D., & De Meester, L. (2007). Host–parasite 'Red Queen' dynamics archived in pond sediment. *Nature*, *450*(7171), 870–873. https://doi.org/10.1038/nature06291

Duffy, M. A., Hall, S. R., Cáceres, C. E., & Ives, A. R. (2009). Rapid evolution, seasonality, and the termination of parasite epidemics. *Ecology*, *90*(6), 1441–1448. https://doi.org/10.1890/08-1130.1

Duffy, M. A., Ochs, J. H., Penczykowski, R. M., Civitello, D. J., Klausmeier, C. A., & Hall, S. R. (2012). Ecological context influences epidemic size and parasite-driven evolution. *Science*, *335*(6076), 1636–1638. https://doi.org/10.1126/science.1215429

Gómez, P., & Buckling, A. (2011). Bacteria-phage antagonistic coevolution in soil. *Science*, *332*(6025), 106–109. https://doi.org/10.1126/science.1198767

Goss, L. B., & Bunting, D. L. (1983). *Daphnia* development and reproduction: Responses to temperature. *Journal of Thermal Biology*, *8*(4), 375–380. https://doi.org/10.1016/0306-4565(83)90025-6

Grosberg, R. K. (2000). Mate selection and the evolution of highly polymorphic self/nonself recognition genes. *Science*, *289*(5487), 2111–2114. https://doi.org/10.1126/science.289.5487.2111

Hutchinson, G. E. (1965). The ecological theater and the evolutionary play. *Yale University Press*.

Koskella, B., & Lively, C. M. (2009). Evidence for negative frequency-dependent selection during experimental coevolution of a freshwater snail and a sterilizing trematode. *Evolution*, *63*(9), 2213–2221. https://doi.org/10.1111/j.1558-5646.2009.00711.x

Luijckx, P., Fienberg, H., Duneau, D., & Ebert, D. (2013). A matching-allele model explains host resistance to parasites. *Current Biology*, *23*(12), 1085–1088. https://doi.org/10.1016/j.cub.2013.04.064

Mahmud, M. A., Bradley, J. E., & MacColl, A. D. C. (2017). Abiotic environmental variation drives virulence evolution in a fish host–parasite geographic mosaic. *Functional Ecology*, *31*(11), 2138–2146. https://doi.org/10.1111/1365-2435.12921

Mostowy, R., & Engelstädter, J. (2011). The impact of environmental change on host-parasite coevolutionary dynamics. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1716), 2283–2292. https://doi.org/10.1098/rspb.2010.2359

Paterson, S., Vogwill, T., Buckling, A., Benmayor, R., Spiers, A. J., Thomson, N. R., Quail, M., Smith, F., Walker, D., Libberton, B., Fenton, A., Hall, N., & Brockhurst, M. A. (2010). Antagonistic coevolution accelerates molecular evolution. *Nature*, *464*(7286), 275–278. https://doi.org/10.1038/nature08798

Refardt, D., & Ebert, D. (2007). Inference of parasite local adaptation using two different fitness components. *Journal of Evolutionary Biology*, *20*(3), 921–929. https://doi.org/10.1111/j.1420-9101.2007.01307.x

Schulte, R. D., Makus, C., Hasert, B., Michiels, N. K., & Schulenburg, H. (2010). Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(16), 7359–7364. https://doi.org/10.1073/pnas.1003113107

Springer, Y. P. (2007). Clinical resistance structure and pathogen local adaptation in a serpentine flax-flax rust interaction. *Evolution*, *61*(8), 1812–1822. https://doi.org/10.1111/j.1558-5646.2007.00156.x

Stuart, Y. E., Veen, T., Weber, J. N., Hanson, D., Ravinet, M., Lohman, B. K., Thompson, C. J., Tasneem, T., Doggett, A., Izen, R., Ahmed, N., Barrett, R. D. H., Hendry, A. P., Peichel, C. L., & Bolnick, D. I. (2017). Contrasting effects of environment and genetics generate a continuum of parallel evolution. *Nature Ecology and Evolution*, *1*(6), 1–7. https://doi.org/10.1038/s41559-017-0158

Su, M., & Boots, M. (2017). The impact of resource quality on the evolution of virulence in spatially heterogeneous environments. *Journal of Theoretical Biology*, *416*, 1–7. https://doi.org/10.1016/j.jtbi.2016.12.017

Tack, A. J. M., Laine, A.-L., Burdon, J. J., Bissett, A., & Thrall, P. H. (2015). Below-ground abiotic and biotic heterogeneity shapes above-ground infection outcomes and spatial divergence in a host-parasite interaction. *New Phytologist*, *207*(4), 1159–1169. https://doi.org/10.1111/nph.13408

Thompson, J. N. (2005). *The Geographic Mosaic of Coevolution*. University of Chicago Press.

Wolinska, J., & King, K. C. (2009). Environment can alter selection in host–parasite interactions. *Trends in Parasitology*, *25*(5), 236–244. https://doi.org/10.1016/J.PT.2009.02.004

Wright, R. C. T., Brockhurst, M. A., & Harrison, E. (2016). Ecological conditions determine extinction risk in co-evolving bacteria-phage populations. *BMC Evolutionary Biology 2016 16:1*, *16*(1), 227. https://doi.org/10.1186/s12862-016-0808-8

## 3.7 Methods

### 3.7.1 Pond experiment

The pond experiment was used to test how epidemic size varied across populations that were initiated with the same suite of hosts and parasites, but experienced biologically realistic variation in biotic and abiotic ecological variables. Additionally, healthy and infected hosts were sampled at the end of the season in order to quantify the change in relative host genotype frequencies across populations and provide parasite samples for the time shift experiment.

To start with, replicate lines of the 12 genotypes of *Daphnia magna* were maintained in the laboratory in a state of clonal reproduction for three generations to reduce variation due to maternal effects. There were five replicates per genotype; each replicate consisted of five *Daphnia* kept in 200 mL of artificial medium (Klüttgen et al., 1994) modified using 5% of the recommended SeO2 concentration (Ebert et al., 1998). Replicate jars were fed 5.0 ABS of *Chlorella* vulgaris algal cells per day (where ABS is the optical absorbance of 650 nm white light by the *Chlorella* culture). *Daphnia* medium was changed three times per week and three days prior to the start of the pond experiment. On the day that the pond experiment commenced, 1–3 day

old offspring were pooled according to host genotype. Ten offspring per genotype were randomly allocated to each of the 16 ponds (giving a total of 120 *Daphnia* per pond). From preliminary work, we knew that the 12 genotypes used in our pond and laboratory experiments were a representative sample of parasite resistance profiles observed in the source population. The proportion of *Daphnia* that became infected with the ancestral mastermix *Pasteuria* after 48h exposure to 2 x $10^5$ spores ranged from 0 to 0.75 depending on genotype, with a mean of 0.27.

Each pond consisted of a 0.65 m tall 1000 Liter PVC tank filled with rainwater. The ponds were set to different depths into the ground and experienced different temperature profiles (Auld & Brand, 2017b). In addition, six of the ponds experienced a weekly mixing treatment where mixed ponds were stirred once across the middle and once around the circumference with a 0.35 $m^2$ paddle submerged halfway into the pond (the exception to this was on the first day of the experiment, when all ponds experienced the mixing treatment to ensure hosts and parasites were distributed throughout the ponds).

The experimental coevolution began on the 2nd April 2015 (Julian day 98), when 120 *Daphnia* (10 *Daphnia* x 12 genotypes) and 1 x $10^8$ *Pasteuria* spores from the ancestral mastermix were added to each of the 16 ponds. The ancestral mastermix comprised *Pasteuria ramosa* spores propagated using 21 separate *Daphnia* genotypes exposed to sediment from their original pond (Kaimes, Scottish Borders, UK (Auld & Brand, 2017b)).

Between the 2nd April and the 17th November 2015, we measured key abiotic and biotic ecological variables on a weekly basis. Temperature, pH, dissolved oxygen (%), chlorophyll (µg. $L^{-1}$), nitrate (mg.$L^{-1}$) and total dissolved salt (mg.$L^{-1}$) were recorded using an Aquaread AP-5000 probe (Aquaread, Broadstairs, Kent, UK). Host density ($L^{-1}$), parasite prevalence and predator density ($L^{-1}$) were determined using standard sampling procedures (Auld & Brand, 2017b).

Twenty-thirty *Daphnia* were sampled from each pond for genotyping after peak epidemic (17th November 2015; Julian Day 321). The DNA extraction and microsatellite genotyping process is described in full in (Auld & Brand, 2017a). Microsatellite genotyping was used to identify the twelve unique multilocus *Daphnia,* and thus track the change in relative genotype frequencies between the beginning

of the experiment (when all genotypes were at equal frequencies) and the end of the experiment. The relative genotype frequencies were used as a measure of relative genotype fitness within each pond. Finally, we sampled 90 infected hosts from each of the 16 ponds, which were homogenised and pooled into three replicate isolates per pond (30 infected *Daphnia* per isolate).

### 3.7.2 Time shift experiment

The time shift experiment was used to understand host and parasite evolution over the course of the epidemic. Specifically, the same panel of host genotypes used to initiate the pond populations was exposed to either the ancestral parasite, or to parasite samples collected from each population at the end of the epidemic, following a fully factorial design.

We established maternal lines for each of the 12 *Daphnia* genotypes used in the pond experiment. There were three replicates per genotype; each replicate consisted of eight adult animals in 100ml of artificial media. The *Daphnia* were fed 0.5 ABS chemostat-grown *Chlorella vulgaris* algae per *Daphnia* per day. Jars were incubated at 20°C on a 12L:12D light cycle, and their media was changed three times per week. Offspring from early instars were taken from the second brood for use in the time shift assay.

The experimental design consisted of a factorial manipulation of the 12 host genotypes and parasite samples collected from each pond ($n$ = 16) plus the original (ancestral) parasite mixed isolate used to seed the populations. There were three independent replicate parasite isolates collected from each pond and a further three replicate isolates of the ancestral parasite (17 parasite treatments; three replicates per treatment). On the day of treatment exposure, neonates from each maternal line were assigned to experimental jars (8 per jar, in 100ml of artificial media) and allocated to parasite treatments following a split-clutch design. There was a total of 612 experimental jars (4896 *Daphnia*). Each jar received a dose of $2 \times 10^5$ *Pasteuria* spores and kept under identical conditions as the maternal lines. After 48 hours exposure to the *Pasteuria* spores, the experimental *Daphnia* were transferred into fresh media. The infection status of each *Daphnia* was determined by eye 25 days post exposure.

Using the results of these infection experiments for each host-parasite combination, we calculated transmission rate ($\beta$, L spore$^{-1}$ day$^{-1}$) using the following equation:

$$\beta = -\frac{1}{Z_0 \cdot t} \cdot \ln\left(\frac{S_t}{S_0}\right) \quad (1)$$

where $Z_0$ is the starting density of spores, $t$ is the duration of the trial exposure, $S_t$ is the density of uninfected hosts at the end of the exposure and $S_0$ is the initial density of hosts.

### 3.7.3 Dissection of host-parasite (co)evolution

By combining transmission rate data from the time shift experiment with relative genotype frequency data from the pond experiment, we dissected the various host and parasite contributions towards the evolution of transmission rate.

To achieve this, we calculated the change in parasite transmission rate over the course of the season and its three contributory components (eq. 2): change in parasite transmission rate due to evolution of host resistance to the ancestral parasite (hereafter, change in host resistance, $\Delta\beta_h$), change in parasite transmission rate due to evolution of parasite infectivity to a set of reference hosts (hereafter, change in parasite infectivity, $\Delta\beta_p$), change in parasite transmission rate due to evolution of parasite infectivity to the evolved host population (non-additive coevolution and hereafter, coevolution, $\Delta\beta_{hp}$).

$$\Delta\beta = \Delta\beta_h + \Delta\beta_p + \Delta\beta_{hp} \quad (2)$$

We used two essential pieces of information to determine how host evolution, parasite evolution and coevolution contributed to changes in overall transmission rate for each population: the change in the relative frequency of each host genotype within each population during the course of the pond experiment; and the difference in the susceptibility of these genotypes relative to the ancestral parasite mix used to seed the populations and the parasite samples collected at the end of the epidemic. First, we calculated the relative frequency of each genotype within each pond at the end of the epidemic. This was done as follows:

$$\bar{w}_{h,t} = P_{h,t} \cdot n_h \quad (3)$$

where $P_{h,t}$ is the frequency of host genotype $h$ at time $t$, and $n_h$ is the total number of host genotypes used to seed the population (in this case, $n_h = 12$). The coevolution experiment started at $t = 0$, when all hosts had a genotype frequency of 1, and ended at $t = 1$.

Then for each population, we calculated the overall change in mean transmission rate. This was done by determining the change in parasite transmission rate for each host genotype between the end of epidemic parasite samples and the ancestral parasite sample, and weighting by the change in host genotype frequency to calculate a mean for each population:

$$\Delta\beta = \frac{1}{n_h} \cdot \sum_h \left( \left( \beta_{h,t=1} \cdot \overline{w}_{h,t=1} \right) - \beta_{h,t=0} \right) \quad (4)$$

where $\beta_{h,t}$ is the transmission rate of each host genotype.

Next, we calculated the mean change in transmission rate due to population-level evolution of host resistance to the ancestral parasite ($\Delta\beta_h$) by calculating the mean resistance to the ancestral parasite weighted by the change in host relative genotype frequency for each population (eq. 5) and the mean change in transmission rate due to parasite evolution in the capacity to infect the ancestral host population ($\Delta\beta_p$, eq. 6).

$$\Delta\beta_h = \frac{1}{n_h} \cdot \sum_h \left( \left( \beta_{h,t=0} \cdot \overline{w}_{h,t=1} \right) - \beta_{h,t=0} \right) \quad (5)$$

$$\Delta\beta_p = \frac{1}{n_h} \cdot \sum_h \left( \beta_{h,t=1} - \beta_{h,t=0} \right) \quad (6)$$

Finally, we calculated mean change in transmission rate due to host-parasite coevolution (*i.e.*, the non-additive component of disease evolution, $\Delta\beta_{hp}$) using eq. 2.

To visualise how changes in host resistance, parasite infectivity and coevolution covaried, we made bivariate plots of $\Delta\beta_h$, $\Delta\beta_p$ and $\Delta\beta_{hp}$ using vectors.

### 3.7.4 Quantifying ecological variation among ponds

We calculated mean values (and also variance for temperature) for each of the 10 ecological variables over the early half of the epidemic season (over twelve sampling dates; Julian days 106-200). Initially, we tested the effects of mixing treatment and then fitted separate linear models to examine the relationships between these ten variables and each of $\Delta\beta_h$, $\Delta\beta_p$ and $\Delta\beta_{hp}$; we evaluated the statistical significance of these relationships after applying a sequential Holm-Bonferroni adjustment for multiple comparisons (Holm, 1979). Next, we conducted a Principal Components

Analysis (using the R function *princomp* (R Core Team, 2019)) on the ten biotic and abiotic environmental variables to generate a multivariate measure of ecological variation across the pond populations (Fig. S3.1). We identified the first four principal components as the minimum number of principal components necessary for explaining over 80% of the combined variation, following standard practice (Brereton & Lloyd, 2016), and used these in subsequent analyses. For outlier detection, we calculated the squared Mahalanobian distances of each population from the mean and compared these values to the critical threshold for Mahalanobis' distance based on a $\chi^2$ distribution, with a critical $\alpha$ value of 0.05. We found that all populations were below the threshold value for outlier detection and thus all of populations were retained.

### 3.7.5 Testing for associations between ecological variation and (co)evolutionary trajectories

We conducted two separate analyses to test for relationships between variation in disease coevolutionary trajectories and wider ecological variation. First, we tested whether pairwise differences in ecological conditions among populations were associated with pairwise differences in disease coevolutionary trajectories. We calculated population differences in ecological conditions ($\delta_{eco}$), made up of the first four principal components (over 80% of combined variation), using the Mahalanobian distances between all of the possible pairwise comparisons of populations and the R package StatMatch v1.3.0 (D'Orazio, 2019). We then calculated the overall multivariate distances for net disease coevolution ($\delta_{coevo}$), *i.e.,* differences in change in parasite transmission rates as a composite for differences across three dimensions: host evolution, parasite evolution and coevolution. We then tested for a relationship between $\delta_{eco}$ and $\delta_{coevo}$ using a Mantel test fitted using the *ecodist* package (Goslee & Urban, 2007).

Second, we used Structural equation modelling (SEM) to dissect the various relationships between ecological variation, epidemic size and the components of coevolution. This was done using the *piecewiseSEM* package v2.0.2 in R (Lefcheck, 2016). SEM allows the evaluation of different causal pathways between variables, and therefore can evaluate support for alternative mediating variables that produce similar associations. We specified two global SEMs (see Fig. S3.2, Table S3.3) with the following variables; mixing, ecological variation (PC1 of the previously described PCA), epidemic size, change in host resistance ($\Delta\beta_h$), change in parasite infectivity

($\Delta\beta_p$) and coevolution ($\Delta\beta_{hp}$). The hypothetical causal relationships between the variables included in these SEMs are outlined below:

*Mixing:* Mixing was an experimental treatment whereby six of the sixteen populations were stirred on a weekly basis. We predicted that this would have a significant effect on the ecological variables. For example, our previous work has shown that mixing significantly changes *Daphnia* host population densities and affects epidemic size (Auld & Brand, 2017b).

*Ecology:* Ecological variation was represented by the first principal component (PC1), which explained 36.0 % of the overall variation, extracted from the PCA of the multiple environmental variables measured during the pond experiment. PC1 was mainly associated with low mean temperature, high chlorophyll concentrations and high predator density. The positive effects of temperature and negative effects of predation on parasite prevalence have been well documented in *Daphnia* disease systems (Auld, Wilson, et al., 2014; Auld & Brand, 2017b; Duffy et al., 2012; Shocket et al., 2018). Therefore, we predicted that our measure of ecological variation would be negatively associated with epidemic size and would be associated with the components of transmission rate evolution (changes in host resistance, parasite infectivity and coevolution).

*Epidemic size:* Epidemic size (integrated parasite prevalence, calculated by integrating the area under the time series of empirically determined prevalence for each mesocosm) could potentially be both a cause and a consequence of host evolution, parasite evolution and coevolution. There is ample evidence from previous studies that epidemics exert parasite-mediated selection and can cause the evolution of host resistance (Auld et al., 2013; Duncan et al., 2006; Laine, 2006; Lohse et al., 2006), and that rapid host evolution of resistance can bring epidemics to an end (Duffy & Sivars-Becker, 2007). Given the bi-directional relationship between these variables we expected that there would be covariation between epidemic size and changes in host resistance, parasite infectivity and coevolution, but made no prediction about the direction of causality.

*Change in host resistance ($\Delta\beta_h$), parasite infectivity ($\Delta\beta_p$), and coevolution ($\Delta\beta_{hp}$):* We developed two SEMs to test between two hypothetical relationships between epidemic size, ecology and different aspects of disease evolution. Hypothesis one is

that ecology directly drives both epidemic size and all three components of disease evolution (Fig. S3.2). Hypothesis two is that ecology affects epidemic size, host evolution of resistance and parasite evolution of infectivity, but that decreases in host resistance (*i.e.,* increased transmission rate) should negatively affect coevolution and increases in parasite infectivity should positively affect coevolution. Following our prediction that the wider environment has a greater impact on hosts compared to parasites, we expected that there would be asymmetry in the strength of the relationship between these different components of evolution with coevolution, such that hosts significantly affect coevolution more than parasites.

After fitting the two SEMs, we tested which provided the superior fit using Bayesian Information Criterion (BIC). We chose BIC over Akaike's Information Criterion (AIC) and AIC corrected for small sample sizes (AICc) because BIC has been shown to better predict model performance when there is unobserved heterogeneity in the data (Brewer et al., 2016), which seems highly likely in both our genotype frequency and ecological variable data. We then conducted Fisher's C tests (Shipley's tests of directed separation (Shipley, 2000) on the best-fitting model to discover potentially relevant relationships that had been excluded from the model. Finally, in order to achieve greater statistical power to test the significance of each of the proposed relationships, we divided the best performing global SEM into two submodels. It should be noted that the parameter estimates for each of the unidirectional relationships in the submodels was identical to the corresponding parameter estimates in the global model.

**Data availability:** All data is available on dryad doi:10.5061/dryad.qv9s4mwd6.

**Code availability:** All companion code is available on Dryad: doi:10.5061/dryad.qv9s4mwd6. As we are actively researching these datasets, we kindly ask that researchers contact us if they are planning to use the data for reasons other than reproducing the findings of our paper.

## 3.8 References (methods)

Klüttgen, B., Dülmer, U., Engels, M. & Ratte, H. . ADaM, an artificial freshwater for the culture of zooplankton. *Water Res.* **28**, 743–746 (1994).

Ebert, D., Zschokke-Rohringer, C. D. & Carius, H. J. Within–and between–population variation for resistance of *Daphnia magna* to the bacterial endoparasite *Pasteuria ramosa*. *Proc. R. Soc. London. Ser. B Biol. Sci.* **265**, 2127–2134 (1998).

Auld, S. K. J. R. & Brand, J. Simulated climate change, epidemic size, and host evolution across host–parasite populations. *Glob. Chang. Biol.* **23**, 5045–5053 (2017).

Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* (1979).

R Core Team. R: A language and environment for statistical computing. (2019).

Brereton, R. G. & Lloyd, G. R. Re-evaluating the role of the Mahalanobis distance measure. *J. Chemom.* **30**, 134–143 (2016).

D'Orazio, M. StatMatch: Statistical matching or data fusion. (2019).

Goslee, S. C. & Urban, D. L. The ecodist package for dissimilarity-based analysis of ecological data. *J. Stat. Softw.* **22**, (2007).

Lefcheck, J. S. piecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics. *Methods Ecol. Evol.* **7**, 573–579 (2016).

Auld, S. K. J. R., Wilson, P. J. & Little, T. J. Rapid change in parasite infection traits over the course of an epidemic in a wild host-parasite population. *Oikos* **123**, 232–238 (2014).

Shocket, M. S. *et al.* Parasite rearing and infection temperatures jointly influence disease transmission and shape seasonality of epidemics. *Ecology* **99**, 1975–1987 (2018).

Duncan, A. B., Mitchell, S. E. & Little, T. J. Parasite-mediated selection and the role of sex and diapause in *Daphnia*. *J. Evol. Biol.* **19**, 1183–1189 (2006).

Auld, S. K. J. R. *et al.* Variation in costs of parasite resistance among natural host populations. *J. Evol. Biol.* **26**, 2479–2486 (2013).

Laine, A.-L. Evolution of host resistance: looking for coevolutionary hotspots at small spatial scales. *Proceedings. Biol. Sci.* **273**, 267–73 (2006).

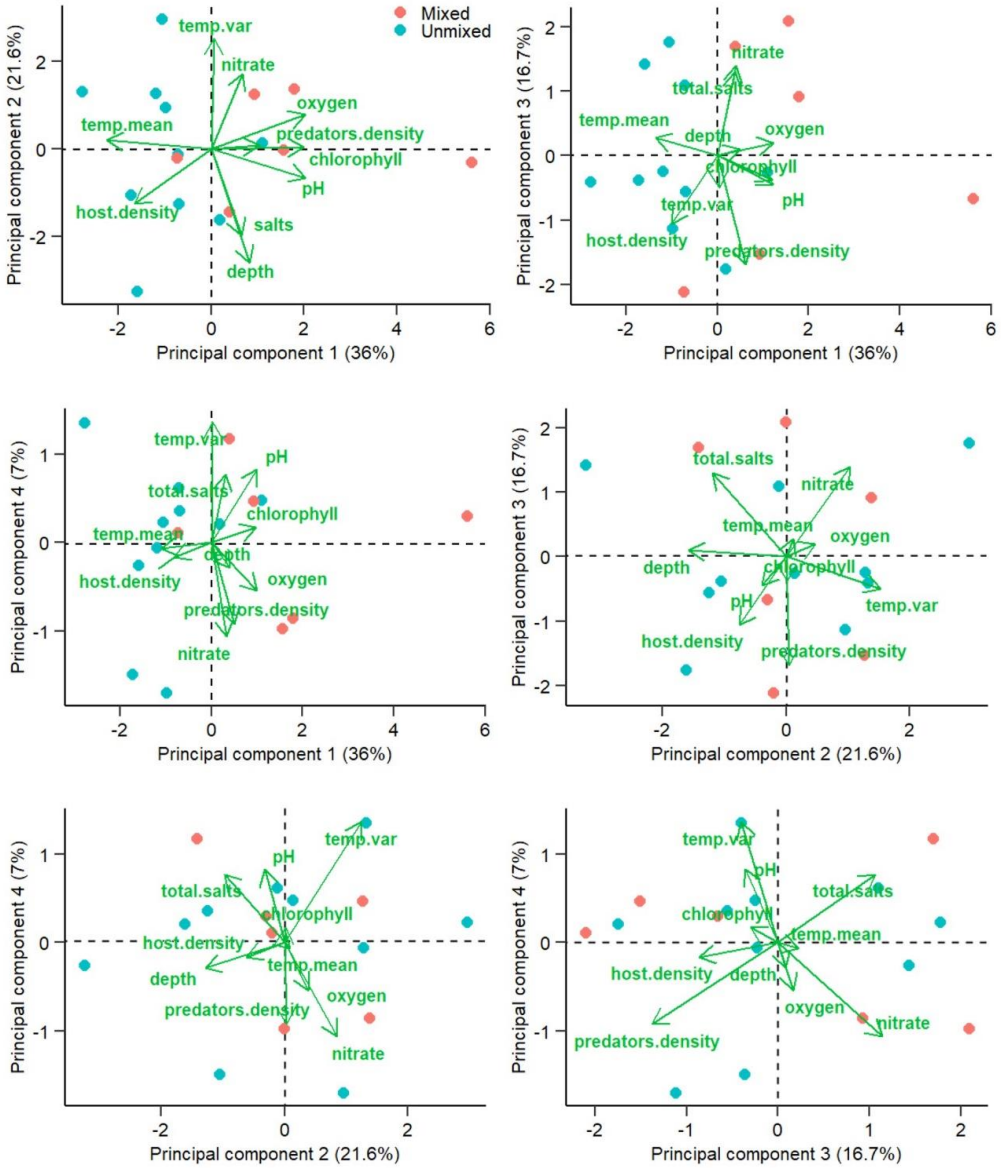
Lohse, K., Gutierrez, A. & Kaltz, O. Experimental evolution of resistance in *Paramecium Caudatum* against the bacterial parasite *Holospora Undulata*. *Evolution (N. Y).* **60**, 1177 (2006).

Duffy, M. A. & Sivars-Becker, L. Rapid evolution and ecological host-parasite dynamics. *Ecol. Lett.* **10**, 44–53 (2007).

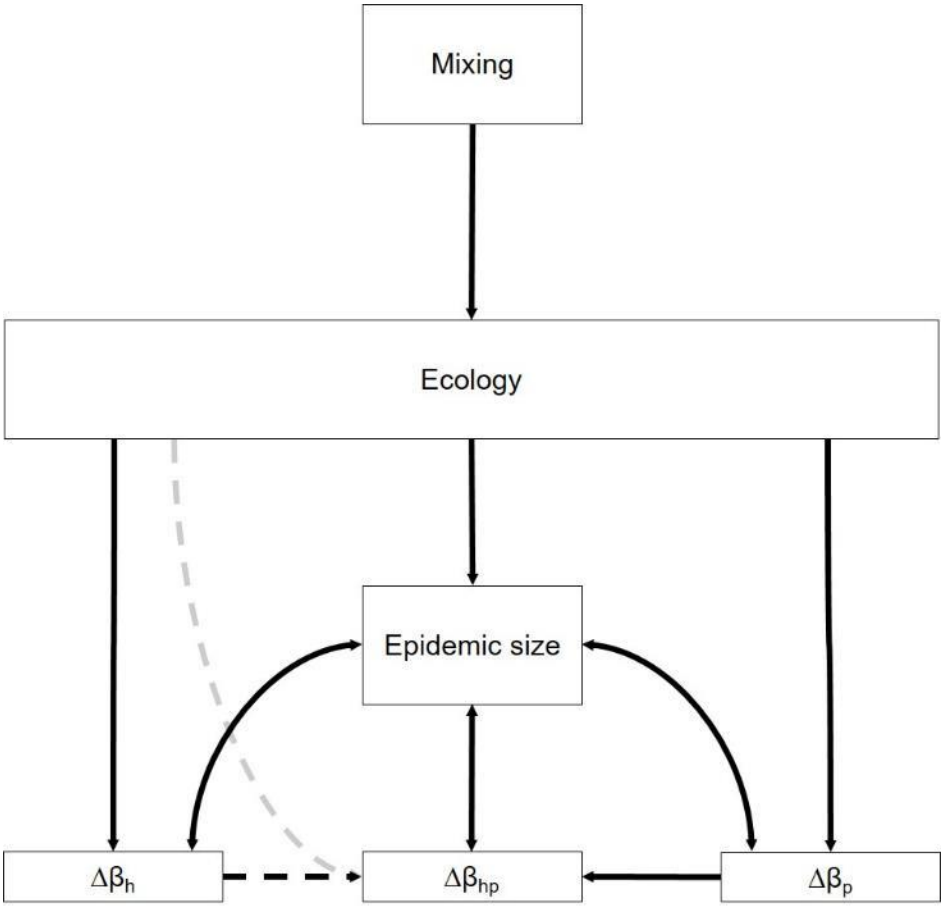
Brewer, M. J., Butler, A. & Cooksley, S. L. The relative performance of AIC, AIC C and BIC in the presence of unobserved heterogeneity. *Methods Ecol. Evol.* **7**, 679–692 (2016).

Shipley, B. A new inferential test for path models based on directed acyclic graphs. *Struct. Equ. Model. A Multidiscip. J.* **7**, 206–218 (2000).

## 3.9 Supplementary information




**Supplementary Figure S3.1.** The composition of principal components in terms of the environmental factors observed. The percentage of overall variance explained by each principal component is shown in brackets. Population environments are represented by the points and these have been coloured according to the mixing treatment, including mixed (red) and unmixed (blue) populations. Variable loadings (i.e. the composition of principal components in terms of the environmental factors observed) are indicated by the green arrows for the following abiotic factors; temperature (mean; temp.mean, variance; temp.var (°C)), pH, total dissolved salts (total.salts ($mg.L^{-1}$)), dissolved oxygen (oxygen (%)), water depth (m) and biotic factors; chlorophyll (µg. $L^{-1}$), nitrate ( $mg.L^{-1}$), adult host density (host.density, ($L^{-1}$)) and predators density (predators.density, ($L^{-1}$)).




**Supplementary Figure S3.2.** Path diagram representing structural equation models for the effects of mixing and ecology on epidemic size and changes in host resistance ($\Delta\beta_h$), parasite infectivity ($\Delta\beta_p$) and coevolution ($\Delta\beta_{hp}$). Large boxes represent measured variables. Arrows represent unidirectional (single arrowhead) or bidirectional (double arrowheads) relationships among variables. There are two different versions of the model and either relationships are specified in both (solid arrows) or only one of the model versions (dashed arrows). In the first model, there is a relationship between change in host resistance and coevolution, whereas in the second model there is a relationship between the environment and coevolution (partially transparent arrow).

2326

**Supplementary table S3.1.** Effect of mixing treatment on each of the ten biotic and abiotic ecological variables. The p-value in bold is significant.

| | Mean in group mixed | Mean in group unmixed | DF | *t* | *P* |
|---|---|---|---|---|---|
| Chlorophyll | 28.95 | 23.10 | 5.37 | 1.56 | 0.17 |
| Water depth | 0.43 | 0.41 | 10.48 | 0.30 | 0.77 |
| Diss. Oxygen | 106.98 | 99.97 | 7.76 | 1.82 | 0.11 |
| Nitrate | 170.08 | 158.91 | 10.11 | 0.57 | 0.58 |
| pH | 8.70 | 8.34 | 12.73 | 2.09 | 0.06 |
| Predator density | 0.09 | 0.06 | 8.72 | 1.37 | 0.21 |
| Adult density | 106.25 | 194.45 | 10.57 | -2.94 | **0.01** |
| Total diss. salt | 61.33 | 57.65 | 9.63 | 0.98 | 0.35 |
| Temp (mean) | 13.66 | 14.30 | 6.26 | -2.07 | 0.08 |
| Temp (var) | 8.36 | 8.29 | 13.24 | 0.16 | 0.87 |

**Supplementary table S3.2.** General linear models testing the univariate relationships between each of the ten ecological variables and host evolution of resistance, parasite evolution of infectivity and host-parasite coevolution. P-values in bold remain significant following a sequential Holm-Bonferroni adjustment[32] for multiple testing.

| | Intercept | (s.e.) | Coefficient | (s.e.) | *P* |
|---|---|---|---|---|---|
| $\Delta\beta_h$ | | | | | |
| Chlorophyll | -6.93E-07 | 1.76E-07 | 2.56E-08 | 6.79E-09 | **0.0020** |
| Water depth | -1.62E-07 | 1.75E-07 | 2.80E-07 | 3.97E-07 | 0.4923 |
| Oxygen | -2.33E-06 | 5.65E-07 | 2.23E-08 | 5.50E-09 | **0.0012** |
| Nitrate | -2.83E-07 | 2.68E-07 | 1.46E-09 | 1.61E-09 | 0.3792 |
| pH | -3.39E-06 | 9.63E-07 | 3.95E-07 | 1.13E-07 | **0.0037** |
| Predator density | -2.40E-07 | 1.15E-07 | 2.75E-06 | 1.45E-06 | 0.0797 |
| Adult density | 2.29E-07 | 1.24E-07 | -1.70E-09 | 7.08E-10 | 0.0309 |
| Total diss. salt | -2.84E-07 | 5.01E-07 | 4.05E-09 | 8.44E-09 | 0.6383 |
| Temp (mean) | 3.95E-06 | 9.61E-07 | -2.84E-07 | 6.83E-08 | **0.0010** |
| Temp (var) | -3.56E-07 | 5.47E-07 | 3.74E-08 | 6.54E-08 | 0.5767 |
| $\Delta\beta_p$ | | | | | |
| Chlorophyll | 2.37E-07 | 1.81E-07 | -1.77E-09 | 6.95E-09 | 0.8027 |
| Water depth | 3.73E-07 | 1.18E-07 | -4.32E-07 | 2.68E-07 | 0.1296 |
| Oxygen | 1.33E-07 | 3.36E-07 | 5.77E-10 | 5.86E-09 | 0.9229 |
| Nitrate | 1.92E-07 | 1.99E-07 | 5.46E-12 | 1.20E-09 | 0.9964 |
| pH | 5.24E-07 | 9.46E-07 | -3.91E-08 | 1.11E-07 | 0.7308 |
| Predator density | 2.63E-07 | 9.09E-08 | -9.91E-07 | 1.15E-06 | 0.4019 |
| Adult density | 1.56E-07 | 1.06E-07 | 2.27E-10 | 6.04E-10 | 0.7132 |
| Total diss. salt | 7.20E-07 | 3.36E-07 | -8.93E-09 | 5.66E-09 | 0.1372 |
| Temp (mean) | 1.51E-09 | 1.04E-06 | 1.36E-08 | 7.37E-08 | 0.8562 |
| Temp (var) | -4.13E-07 | 3.65E-07 | 7.28E-08 | 4.36E-08 | 0.1174 |
| $\Delta\beta_{hp}$ | | | | | |
| Chlorophyll | 5.92E-07 | 3.05E-07 | -2.00E-08 | 1.17E-08 | 0.1101 |
| Water depth | 4.05E-07 | 2.20E-07 | -7.67E-07 | 4.99E-07 | 0.1468 |
| Oxygen | 1.76E-06 | 1.02E-06 | -1.63E-08 | 9.93E-09 | 0.1233 |
| Nitrate | 2.82E-07 | 3.65E-07 | 1.21E-09 | 2.19E-09 | 0.5904 |

| | | | | | |
|---|---|---|---|---|---|
| pH | 3.09E-06 | 1.56E-06 | 3.88E-07 | 1.84E-07 | 0.0748 |
| Predator density | 2.36E-07 | 1.67E-07 | -2.12E-06 | 2.10E-06 | 0.3293 |
| Adult density | -1.57E-07 | 1.84E-07 | 1.50E-09 | 1.05E-09 | 0.1741 |
| Total diss. salt | 1.28E-06 | 5.94E-07 | -2.03E-08 | 9.99E-09 | 0.0617 |
| Temp (mean) | -3.21E-06 | 1.71E-06 | 2.34E-07 | 1.21E-07 | 0.0738 |
| Temp (var) | -1.49E-07 | 7.36E-07 | 2.82E-08 | 8.81sE-08 | 0.7533 |

2339

2340

# 4. The effect of host population genetic diversity on variation in metrics of parasite infection success

This chapter is the most up to date version of the pre-print which has been published on *BioRxiv*.

All authors discussed the results and contributed to the final manuscript. Sam Paplauskas performed the data collection, analysed the data and wrote the manuscript. Dr Brad Duthie and Professor Matthew Tinsley contributed to the final version of the manuscript and supervised the project.

## 4.1 Abstract

Conventional wisdom suggests that populations with lower levels of genetic diversity are at a greater risk of the more harmful effects of disease. However, previous attempts to qualify this proposition have focused on measuring the mean, rather than the variability, in metrics of parasite infection success. Since the ability of host population genetic diversity to limit the spread of disease requires some specificity between hosts and parasites, and the benefits of host population genetic diversity in resistance to infection may depend on the respective parasite population genetic diversity, we propose a diversity-uncertainty model which predicts that the mean and variability in parasite success depend on a combination of parasite host range and parasite population genetic diversity. By re-analyzing a dataset combining 48 studies collected by previous meta-analyses, we show that the effect of host population genetic diversity reduces the mean infection success of single-host, but not host generalist, parasites. We find evidence for our original hypothesis that the variability of parasite success depends on a combination of host population genetic diversity, parasite population genetic diversity and host range. Together, these results challenge conventional wisdom and have important implications for how genetic diversity can be better managed in host populations.

## 4.2 Introduction

It is commonly believed that host populations with lower genetic diversity are at a greater risk of experiencing higher parasite success (i.e. disease (King & Lively, 2012)). This refers to the population-level prevalence (proportion of infected hosts), virulence (parasite-induced loss of fitness) or parasite load (average parasites per host (Hamilton, 1987; O'Brien & Evermann, 1988; Sherman et al., 1988)).

Previous studies of the generality of this proposed 'conventional wisdom' (King & Lively, 2012), have often focused on measuring the mean, rather than the variability, of parasite success (Ekroth et al., 2019; Gibson & Nguyen, 2021). This is surprising, considering the importance of parasitic extremes, in terms of epidemics and whether they cause mass extinction (Alan Pounds et al., 2006; De Castro & Bolker, 2004), the predictably of recurrent bouts of disease across years and the repeatability of disease experiments in general. As a result, the relationship between host diversity and variability in parasite success is poorly understood (Gibson, 2022). However, it is central to our ability to protect against future emerging diseases (Altizer et al., 2006).

The implications of host community, species or genetic diversity on infectious diseases is often referred to as 'disease dilution' (Johnson et al., 2015; Keesing et al., 2006, 2010; Keesing & Ostfeld, 2021; Ostfeld & Keesing, 2012), the diversity-disease hypothesis (Altermatt & Ebert, 2008a; Johnson et al., 2012; Mihaljevic et al., 2014) or the monoculture effect (Browning & Frey, 1969; Elton, 1958; Garrett & Mundt, 1999; Leonard, 1969; van der Plank, 1963). This can be caused by an increase in individual host susceptibility (Coltman et al., 1999), or a variety of population-level effects such as reducing the rate of encounter between susceptible and infectious individuals (encounter reduction), reducing the probability of transmission given an encounter (transmission reduction), decreasing the density of susceptible individuals (susceptible host regulation), increasing the recovery rate (recovery augmentation), or increasing the death rate of infected individuals (infected host mortality) (for a review, see (Keesing et al., 2006)).

Although the exact mechanism is unclear, the negative relationship between host population genetic diversity and disease spread is often attributed to encounter reduction (Anderson et al., 1986). Specifically, assuming that there is some level of matching (or genetic specificity (Schmid-Hempel & Ebert, 2003)) required for a

successful infection to occur (*sensu* matching alleles model [MAM] (Agrawal & Lively, 2002)), there should be a lower chance of a parasite genotype encountering a susceptible host genotype as it spreads through a more diverse host population. Since the strength of genetic specificity varies across different host-parasite systems (Agrawal & Lively, 2002), we might expect that the effect of host population diversity on parasite success depends on the level of specificity for infection.

In theory, the effects of host population diversity on parasite success may also depend on the level of parasite diversity (Boomsma, 1996; Van Baalen & Beekman, 2006). For example, if there is a high level of genetic specificity for infection (*sensu* MAM), then we might expect host populations composed of a single genotype to be entirely susceptible to a single parasite genotype, which is much more likely to occur in a population with a high level of parasite diversity (Boomsma, 1996; Van Baalen & Beekman, 2006). One empirical study in a *Daphnia* host-parasite system found that the benefits of host genetic diversity for resistance to infection were reliant on a high level of parasite diversity (Ganz & Ebert, 2010).

Therefore, if we assume that there is a high level of genetic specificity for infection (*sensu* MAM) and both host and parasite populations are characterized by either high or low levels of genetic diversity, we can predict the following patterns for both the mean and variability in parasite success (Fig. 4.1):

A) Low host x low parasite population genetic diversity (Fig. 4.1A): We predict that there will be a high level of variability in parasite success, due to the host population being composed entirely of susceptible, or resistant, host genotypes, and an intermediate level of mean parasite success (determined by the overall frequency of resistant cf. susceptible populations).

B) High host x low parasite population genetic diversity (Fig. 4.1B): We predict that there will be a low level of both mean parasite success and variability in parasite success, due to the consistency of hosts to resist infection through a reduced encounter rate with matching parasite genotypes.

C) Low host x high parasite population genetic diversity (Fig. 4.1C): We predict that there will be a high level of mean parasite success and a low level of variability in

parasite success, due to the consistency of parasite transmission through an enhanced encounter rate with matching host genotypes.

D) High host x high parasite population genetic diversity (Fig. 4.1D): We predict that there will be an intermediate level of both mean parasite success and variability in parasite success, due to the diverging effects of host and parasite genetic diversity leading to an inconsistent encounter rate between matching host and parasite genotypes.

Collectively, these predictions form our 'diversity-uncertainty' model for predicting the mean and variability of parasite success across populations with different levels of host and parasite diversity. This builds on previous work (Bensch et al., 2021), which focused on the relationship between population diversity and variability in parasite-induced host mortality and pathogen abundance for only three out of the four possible combinations in Figure 4.1, without also acknowledging the influence of the genetic specificity for infection on these hypotheses.

To test our diversity-uncertainty model, we examine the relationship between host population genetic diversity, parasite population genetic diversity and variability in parasite success for different levels of a proxy for genetic specificity using meta-analysis. After confirming the results of previous studies (Ekroth et al., 2019; Gibson & Nguyen, 2021), which found a significant difference in mean parasite success between various host populations with high versus low genetic diversity, we then extend their analysis to a study of variability using a suite of different moderator variables.

In particular, we compare the difference in the variability of parasite success between host populations with high versus low genetic diversity using a combination of host range and parasite population genetic diversity variables. Since the underlying genetic model of infection is known for only a small number of host-parasite systems (e.g. *Daphnia-Pasteuria* (Pepijn et al, 2013)), we instead used parasite host range as a proxy for the genetic specificity of each host-parasite system. We characterised parasites with a host range of one species by a matching-alleles model (Agrawal & Lively, 2002) and parasites with a host range of more than one species by a gene-for-gene model of infection genetics (Agrawal & Lively, 2002). The reasoning behind this was that tightly knit host-parasite coevolution (sensu a matching-alleles model

of infection genetics) would be more likely for highly specific interactions between host and parasite genotypes (Schmid-Hempel & Ebert, 2003), which might be expected for specialist, rather than generalist parasites. On the other hand, we do not make any predictions about the mean level of, or level of variability in, parasite success for systems with a low genetic specificity for infection.

Overall, we find that the relationship between host genetic diversity and both the mean level of, and level of variability in, parasite success depends on a combination of host range and parasite genetic diversity.

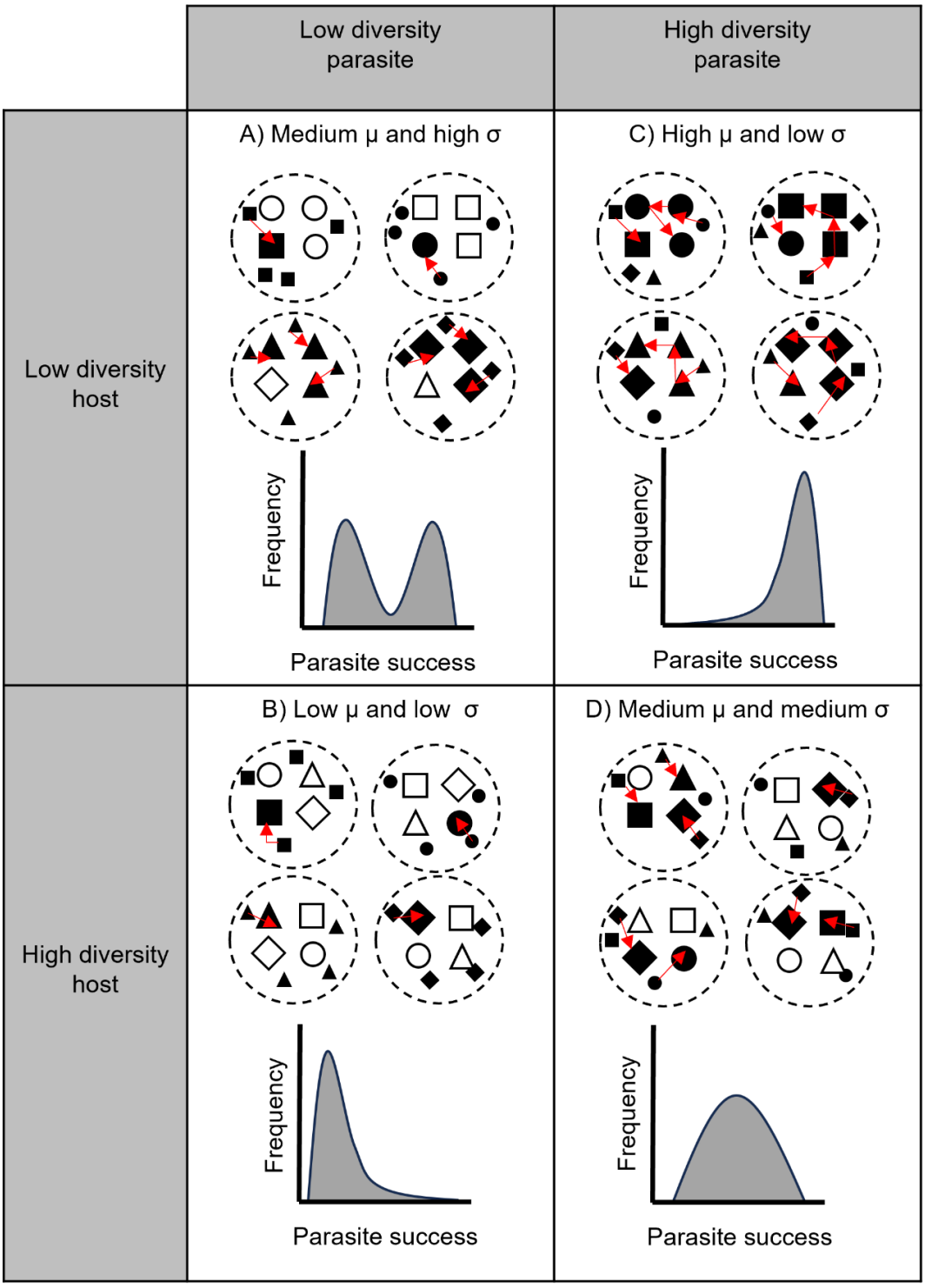



Figure 4.1. A hypothetical 'Epidemic Diversity' model for the combined relationship between host and parasite population genetic diversity and either the mean (μ) or the variability (σ) in parasite success. There are four hypothetical populations for each combination of host and parasite population genetic diversity (dashed circles). The level of population genetic diversity is indicated by the number of unique host and parasite genotypes (large and small shapes respectively) and is

the same in each replicate population. The colour of hosts indicates their infection status, such that susceptible hosts are white and infected hosts are black, whereas the parasite is always the same colour (also black). Parasite transmission can only occur between matching host and parasite genotypes (shapes) and is indicated by the red arrows. The resulting frequency distributions of parasite infection success for each set of replicate populations is shown at the bottom of each plot. Notably, this hypothetical model only applies for host--parasite systems that have a high level of genetic specificity for infection (i.e. matching-allele versus gene-for-gene infection genetics (Agrawal & Lively, 2002)).

## 4.3 Methods

### 4.3.1 Summary

We combined the data from two previous meta-analyses (Ekroth et al., 2019; Gibson & Nguyen, 2021) that used the standardised mean difference (SMD) to calculate the significance of the relationship between host genetic diversity and metrics of mean parasite infection success. These data were mainly from studies where the metrics of parasite infection success were measured from replicate host populations that were already classified qualitatively as having either a ‘high’ or ‘low’ level of genetic diversity (27 out a total of 48 independent studies), including studies of replicate host populations with a high versus low level of inbreeding (e.g. Baer & Schmid-Hempel, 1999), different combinations of host genotypes (e.g. Altermatt & Ebert, 2008) or comparisons between wild host populations exposed to different selection regimes (e.g. a population bottleneck, random genetic drift, e.g. Hale & Briskie, 2007). However, there were some studies that measured the relationship between metrics of parasite infection success and a continuous measure of host population genetic diversity (for the absence of any significant correlation between metrics of mean parasite infection success and host population genetic diversity, involving all of the studies from this subset of data, see Gibson & Nguyen, 2021), and therefore these data were binned into ‘high’ and ‘low’ categories (as mentioned above). We used the data combined from the two previous meta-analyses (Ekroth et al., 2019; Gibson & Nguyen, 2021) for further study of the factors influencing mean parasite success, then we assessed how host population genetic diversity influenced variation in parasite success by calculating the log coefficient of variation ratio (lnCVR). This variation was quantified between experimental replicates in a laboratory

environment, between multiple natural host populations with similar genetic diversity, or sometimes between repeated measures of single populations along a time series.

### 4.3.2 Data collection

The data collection for each comparison of a group of high versus low genetic diversity populations, which was later used in calculating effect sizes, involved five main steps (Fig. 4.1):

1) First, we combined the list of studies from (Gibson & Nguyen, 2021) and (Ekroth et al., 2019), removed any duplicate studies and added the data used to calculate the effect size, SMD, and its sampling variance in the original studies (mean, standard deviation, sample size), the metric of parasite success and the unique study, experiment and replicate identifiers used to account for the non-independence of separate effect sizes. We did not use the parasite success data from (Ekroth et al., 2019) because the original data extracted from each study was missing from the online supplementary material, meaning we were unable to check the data accuracy during validation (step 3); for these studies we extracted the replicate or population summaries from the original papers ourselves after step 3 and recalculated the mean, standard deviation, sample size etc..

The data used to calculate Fisher's z (an effect size for the difference between two correlation coefficients) for the observational field studies from (Gibson & Nguyen, 2021), was not in the correct format to calculate either SMD or lnCVR. Therefore, we did not include this information (from multiple populations with a continuous measure of genetic diversity) at this stage and instead extracted the data from the original publications ourselves and recalculated it during steps 4 and 5. Also, we excluded any studies on plants (wild or agricultural) because a more detailed analysis of the plant literature would require a separate review.

2) Second, we amended the inclusion criteria used in the original meta-analyses (Ekroth et al., 2019; Gibson & Nguyen, 2021) (Table S4.1) and removed any studies, experiments or comparisons which did not meet these criteria:

(i) 'Parasite success', which we define as the ability of a parasite to spread among hosts (transmission rate, infection rate, prevalence), replicate on / within hosts

(macro / microparasite load, disease severity) or kill hosts (virulence i.e. host survival / mortality rate) was measured among replicate populations across time or space.

(ii) Parasite success data was collected from two or more host populations with a difference in genetic diversity assessed by metrics such as: individual inbreeding status (inbred versus outbred), genotypic diversity (high versus low) or heterozygosity (high vs low).

(iii) Genetic diversity was quantified at the level of the host population, rather than for community-level diversity.

(iv) The study focused on an animal (or bacterial) host species.

(v) The study does not re-analyze the data from a previously published study.

(vi) The parasite success data was not replicated simply by using an alternate way of measuring host population diversity.

(vii) Figures required to extract parasite success data were clearly legible.

3) Third, we checked the accuracy of the data from the excel spreadsheets used to calculate the summary of the parasite data for each group of host populations with either high or low genetic diversity from the online data supplied by one of the previous meta-analyses (Gibson & Nguyen, 2021) and corrected these in the fourth step of the data collection before including them in our analysis. The different types of error made by the previous meta-analysis included (i) 27 comparisons that did not match the published raw data (available in the main text or online or in the supplementary material of each publication), (ii) 32 comparisons where effect sizes were calculated wrong and (iii) 10 comparisons which had not been transferred into the final metadata file correctly. There was one study which we could not check, because the original data was sent by personal communication from (King et al., 2011) to (Gibson & Nguyen, 2021), nevertheless we included it in our analysis.

4) Fourth, for those studies or comparisons we had excluded (due to missing data or data errors) we extracted the data from the main text or supplementary files by going back to the original publications (we used PlotDigitizer (https://plotdigitizer.com/) to

extract the information for any figures). In addition to the comparisons removed in the third step of data collection, this also included (i) 26 studies that, despite meeting our inclusion criteria, were removed because they were either missing the replicate-level raw data (Ekroth et al., 2019) or they were observational field studies based on multiple populations with a continuous measure of genetic diversity (Gibson & Nguyen, 2021) and (ii) additional data for 3 comparisons (Agha et al., 2018; Baer & Schmid-Hempel, 2001; Giese & Hedrick, 2003) that were not made in the original analysis by (Gibson & Nguyen, 2021).

We also collected information on 10 different moderators (see Table 1) by standardizing or recoding existing moderator variables used by (Gibson & Nguyen, 2021), including host range, parasite diversity, metric of parasite success, host species, parasite type, source of host genetic diversity, scale of host diversity, mode of host reproduction, whether the parasite induces host mortality and whether the study was performed in a laboratory environment. Parasite diversity was not quantified as a continuous variable in the original studies, nor was it examined as part of the original experiment in most cases. Therefore, we binned parasite diversity into 'high' or 'low' groups depending on the following reasoning; if the isolate was collected from a natural population for a lab study, if the data was from an observational or experimental field study, or if more than one genotype had been identified (but this only applied to a small number of studies); low parasite diversity was specified if it was a laboratory strain, or only one genotype had been identified (but again, this only applied to a small number of studies). Where the information on these moderator variables was not already available from the supplementary material of (Gibson & Nguyen, 2021) and was not available in the published article, we performed an online search to determine characteristics.

5) Fifth, we calculated the mean, standard deviation and sample size for each comparison of high versus low genetic diversity groups of host populations for the data we extracted in the fourth step of data collection (Gibson & Nguyen, 2021). For certain studies, we calculated a pooled measure of the mean metric of parasite infection success for each group of high and low genetic diversity host populations, along with a pooled standard deviation and a pooled sample size. This included (i) studies based on multiple populations with one or more continuous measures of genetic diversity (Dagan et al., 2013; Dionne et al., 2009; Ellison et al., 2011; S. G. Field et al., 2007; Giese & Hedrick, 2003; King et al., 2011; Kyle et al., 2014; Loiseau

et al., 2011; Meagher, 1999; Neumann & Moritz, 2000; Parsche & Lattorff, 2018; Pierce et al., 2014; Puurtinen et al., 2004; Queirós et al., 2016; Rahn et al., 2016; Savage et al., 2015; Trouvé et al., 2003; Velavan et al., 2009; Whitehorn et al., 2011, 2014; Whiteman et al., 2006), for which the most appropriate measure of population-level genetic diversity was used (e.g. a measure of population-level genetic diversity based on Hardy-Weiberg equilibrium) and two separate groups of host and low diversity host populations were made with the same number of host populations in each group and (ii) studies with multiple groups of either high or low diversity host populations that shared the same corresponding (or so-called 'reference') group (Agha et al., 2018; Schmidt et al., 2011). In addition, host survival was converted into host mortality in some studies to reflect our definition of parasite infection success (see step two of data collection). Overall, this fifth step of data collection involved calculating parasite success data for 130 comparisons.

After finishing all five steps of data collection, there was enough parasite success data to calculate both the SMD and lnCVR for 211 non-independent comparisons of high versus low genetic diversity groups of host populations.

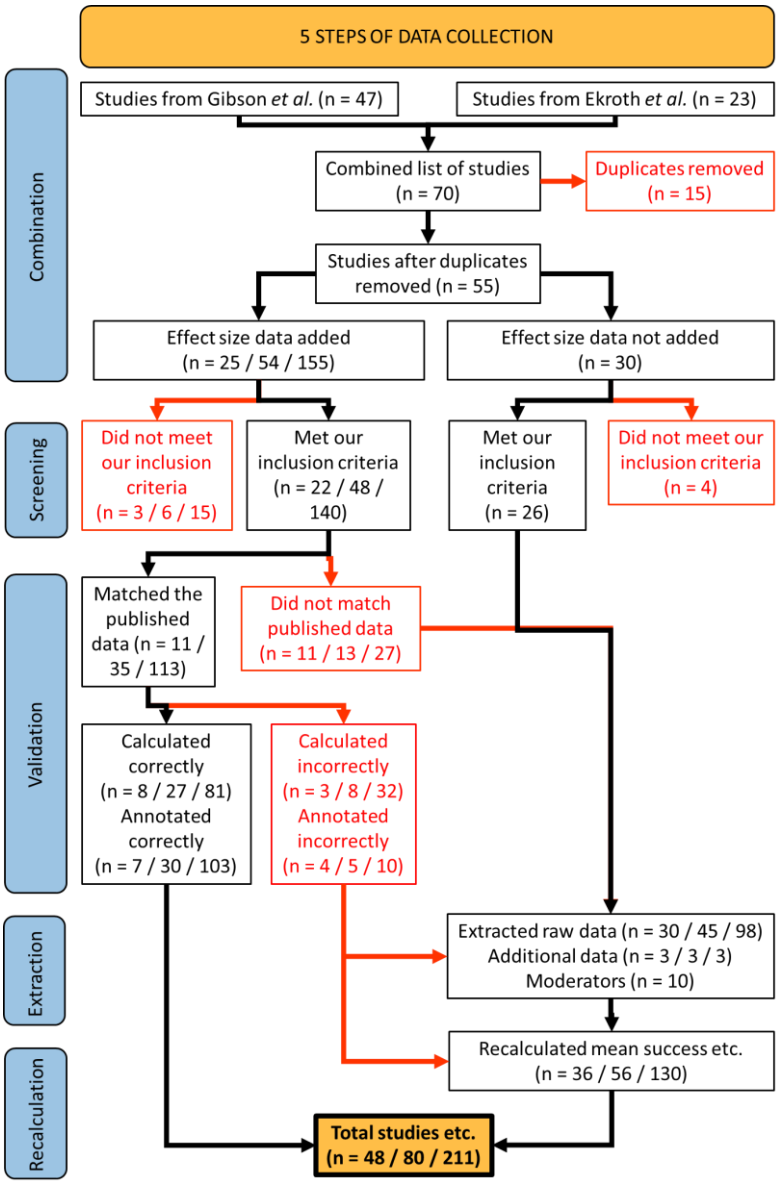



Figure 4.2 How data was collected for each comparison of a parasite infection success metric between high versus low genetic diversity groups of host populations, which was later used in calculating effect sizes. n = number of studies / experiments / comparisons (the multi-level structure of the data make it appear that some sums are incorrect). Adapted from the preferred reporting items for systematic reviews and meta-analyses (PRIMSA) statement (Page et al., 2021).

### 4.3.3 Calculation of effect sizes (SMD, lnCVR, lnRR, lnVR)

We calculated two main effect sizes (standardized mean difference [SMD] and log coefficient of variation ratio [lnCVR]) and (to test the robustness of our results) two additional effects sizes (log response ratio [lnRR] and log variability ratio [lnVR]). All

effect size calculations and subsequent calculations were performed in R v4.3.2 (R Core Team, 2023).

*Main effect sizes* - We chose SMD and lnCVR as our main effects sizes because they compare the difference in either the mean or variability of two groups whilst accounting for certain factors: (i) SMD measures the mean difference between two groups (high versus low genetic diversity) in terms of standard deviations (Borenstein et al., 2009; Field & Gillett, 2010), so it can be used to compare metrics measured on very different scales (prevalence, load and virulence) (Higgins et al., 2024); it also corrects for small sample sizes, which is a common feature of ecological studies (Jennions, 2003). (ii) lnCVR measures the ratio of variability between two groups adjusted for the size of the group means (Nakagawa et al., 2015) and as a result, accounts for the possibility that the magnitude of the variability may scale with the mean, as is the case for many types of count data (such as parasite load) that follow a Poisson distribution.

*Alternative effect sizes* - We calculated lnRR and lnVR as alternatives to our main effect sizes, and although they did not account for all of the same factors, these provided a separate way of assessing effects of host population genetic variation on the mean and variability of parasite success (Nakagawa et al., 2015, 2023). By comparing the two sets of effect sizes we assessed the robustness of our results (Koricheva & Gurevitch, 2014).

Before calculating our effect sizes, we added a small value (0.001) to the mean and standard deviation in parasite success for each pair of control and treatment groups to ensure log values were calculated correctly. For consistency, we calculated SMD and its sampling variance from the formula derived from the supplementary material of (Gibson & Nguyen, 2021), whereas we calculated all variability effect sizes and their sampling variances using the code from (Nakagawa et al., 2015). We calculated lnRR using the escalc function from the metafor package v4.4.0 (Viechtbauer, 2010). To account for comparisons based on shared controls, we calculated the variance-covariance matrix for each effect size, using the make_VCV_matrix function from the metaAidR package v0.0.0.9000 (Lagisz et al., 2024).

### 4.3.4 Publication bias

Before analyzing the data fully, we calculated the overall effect sizes for SMD and lnCVR and tested for any potential publication bias using funnel plots and Egger's regression (Sutton, 2009).

Meta-analytic models were fitted to the data using the rma.mv function from the metafor package v4.4.0 (Viechtbauer, 2010). We included fixed effects for each type of effect size, the variance-covariance matrix of sampling errors, standard random effects for study and host genus, and correlated random effects for comparisons taken from the same experiment. Standard random effects for study and host genus were used to account for the possibility of non-independence between experiments originating from the same study and potential correlations between effects from closely related host species. Similarly, correlated random effects were used to account for potential non-independence of comparisons taken from the same experiment (multiple timepoints for a single comparison of control and treatment groups, or effect sizes from the same group of populations based on different measures of parasite success).

Funnel plots were used to identify whether published effect sizes were evenly distributed around model means by examining how outcomes varied as a function of their precision (standard error). This was achieved from a visual inspection of these plots and statistical evaluation using Egger's regression.

### 4.3.5 Meta-analysis of overall data

To test whether there was a significant difference in the mean parasite success (SMD) or variability in parasite success (lnCVR) between host populations with high versus low genetic diversity, we fitted mixed effects meta-analytic models. All of the models used in this paper were based on the same structure as those used for testing the presence of publication bias.

### 4.3.6 Context dependence

*Partial moderator analysis* - To test if the overall effect of host population genetic diversity on the mean and variability in parasite success depended on an interaction between host range and parasite genetic diversity, we introduced an interaction term

for these two moderators in our original meta-analytic models. Therefore, we could compare:

1) High versus low single-host parasite population genetic diversity.
2) High versus low multi-host parasite population genetic diversity.

We compared the significance level of each individual predictor within the model, as well as the contrasts between them using the *glht* function from the *multcomp* package v1.4.25 (Hothorn et al., 2008).

*Full moderator analysis* – To test our additional hypotheses (Table 4.1) for the eight remaining moderator variables, we modelled each moderator separately with its own individual mixed effects model. Before running the models, we removed redundant moderator categories with a limited sample size, such as transmission or infection rate (versus prevalence) and disease severity (versus load) for the metric of parasite success, and prokaryotic (versus vertebrate or invertebrate) for host species.
We compared the significance level of each individual predictor within the model, as well as the contrasts between them using ANOVA with a correction for multiple comparisons (Holm's method).

Table 4.1. Hypotheses for the influence of additional moderator variables on the nature of the effect of host population genetic diversity on mean and variability in parasite success.

| Moderator | Hypothesis |
|---|---|
| Metric of parasite success | Our study of 'parasite success' combined data of several types (eg prevalence, virulence, infection load). Using this moderator, we tested if the effects of host population genetic diversity differed between these different metrics. |
| Host type | The effect of host population genetic diversity may be influenced by the specificity of genetic interactions between host and parasite. These genetic interactions are thought to be more specific in invertebrates than in vertebrates (Dybdahl et al., 2014), therefore we tested for inconsistency of the host population genetic diversity effect in these two groups. |
| Parasite type | Microparasites and macroparasites tend to have contrasting infection biology: microparasite infections are often short-lived, whereas macroparasite infections can be long-lasting due to |

| | |
|---|---|
| | parasite abilities to circumvent host immune responses (Sorci, 2014). These differences might drive variation in the impact of host population genetic diversity. Therefore, we tested for inconsistency of the host population genetic diversity effect in these two groups. |
| Source of host genetic diversity | Studies typically investigate the impact of host genetic diversity by either (i) inbreeding lineages to create a comparison between inbred and outbred populations, (ii) using a suite of wildtype genotypes for controlled experiments with either low genetic diversity or high genetic diversity, or (iii) sampling organisms from the wild from populations that have been characterised as having different levels of genetic diversity. We used this moderator to test if these different sources of genetic diversity affected the influence of host population genetic diversity. |
| Scale of host diversity | Host populations were predetermined as having either high or low diversity (discrete) or we separated them into such categories as part of our data collection (Fig. 2, step 5) because the authors used multiple populations with a continuous measure of diversity. We used this moderator to test if this feature of how studies were designed had an effect on the influence of host population genetic diversity. |
| Mode of host reproduction | Host species reproduced sexually, asexually or using a combination of the two (i.e. facultatively sexual, such as *Daphnia*). We used this moderator to test if these different modes of host reproduction affected the influence of host population genetic diversity. |
| Host mortality? | The range of parasites studied can be further categorised by whether or not infection typically kills the host (which may be proxy for virulence). We used this moderator to test if differences in the virulent effects of parasitism affected the influence of host population genetic diversity. |
| Laboratory? | We used this moderator to test if the difference in study setting (laboratory versus field) affected the influence of host population genetic diversity. |

### 4.3.7 Sensitivity analysis of overall effects

To test the robustness of our results for the combined (overall) dataset, we performed a series of 'leave-one-out' sensitivity analyses. This involved the iterative exclusion of either one independent comparison (i.e. treatments with shared controls were considered grouped together into a single comparison) or study at a time.

## 4.4 Results

### 4.4.1 Absence of publication bias

Our dataset contained 211 estimates of the effect that changes in host population genetic diversity have on parasite success; we assessed this effect on both mean parasite success (SMD) and the variability in parasite success (lnCVR). Visual inspection of funnel plots for the effect of host population diversity on mean parasite success (Fig. 4.3A) and its effect on the variability in parasite success (Fig. 4.3B), showed no evidence for publication bias. More stringent evaluation showed that there was no correlation between the size of the effects themselves and their standard error (Egger's test for both SMD and lnCVR: R = 0.06, 95% CI [-0.24, 0.37], p = 0.67 and R = -0.03, 95% CI [-0.38, 0.32], p = 0.86, respectively).

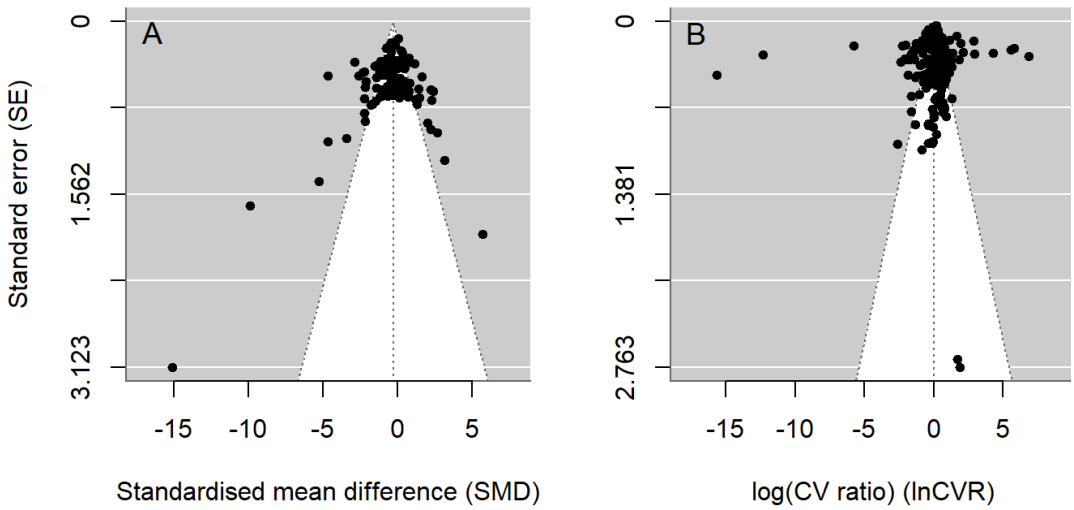



Figure 4.3. Testing for publication bias: the distribution of published effect sizes for our meta-analysis as a function of their precision (standard error). The x-axis in both plots shows effects of an increase in host population genetic diversity (high

vs low) on A) the mean difference in parasite success (SMD) and B) the ratio of variation in parasite success (lnCVR). Model means and their 95% confidence intervals are shown by the dashed black lines.

### 4.4.2 Evenly distributed animal host and parasite taxa

Our dataset included a diverse range of hosts and parasites, including 31 unique host genera, 60 unique parasite genera and 71 unique combinations of host-parasite genera (or 92 unique species combinations).

Most unique host taxa in our dataset were animals (invertebrate and vertebrate genera / species), with only two unique non-animal (prokaryotic) host species (Table 4.2). However, there was an even distribution of the unique parasite taxa across the combination of all unique host taxa.

Table 4.2. The number of unique host and parasite combinations and how evenly they are distributed across different taxonomic groups. The number of unique combinations of host and parasite genera and species is shown by the first two numbers separated by a backslash (genera / species) and the number of studies they correspond to in parentheses. The total number of studies (58) is higher than the total number of studies in our dataset (48), because there were some studies with multiple comparisons of unique host and parasite combinations. The colour coding is based on the number of unique combinations of host and parasite genera.

| | | Host taxon | | | |
|---|---|---|---|---|---|
| | | Prokaryote | Invertebrate | Vertebrate | Total |
| Parasite taxon | Animal | 0 / 0 (0) | 12 / 14 (6) | 10 / 13 (8) | 22 / 27 (14) |
| | Bacteria | 0 / 0 (0) | 2 / 2 (2) | 13 / 16 (5) | 15 / 18 (7) |
| | Fungi | 1 / 1 (1) | 15 / 18 (16) | 1 / 1 (1) | 17 / 20 (18) |
| | Protozoa | 0 / 0 (0) | 7 / 12 (11) | 3 / 3 (2) | 10 / 15 (13) |
| | Virus | 1 / 1 (1) | 3 / 7 (2) | 4 / 4 (4) | 7 / 12 (7) |
| | Total | 1 / 2 (2) | 39 / 53 (37) | 31 / 37 (20) | 71 / 92 (59) |

### 4.4.3 Host population genetic diversity has an overall negative effect on mean parasite success

Averaging over the whole data set, there was a significant effect of host population genetic diversity on mean parasite success (SMD = -0.29, 95% CI = [-0.57, -0.02], n = 211; Fig. 4.4A); higher levels of host population genetic diversity were associated with lower mean parasite success. However, across the whole data set, there was no effect of host population genetic diversity on the variability of parasite success (lnCVR = 0.02, 95% CI = [-0.30, 0.35], n = 211; Fig. 4.4B).

In these analyses the residual variation (heterogeneity) in the data for both the difference in the mean and the variability of parasite success was high ($I^2$ = 84.0% & 82.0% respectively). Most of this variation was explained by the effect of study (84.0% & 80.7%) and only a small amount was explained by host genus (0.0% & and 3.3%).

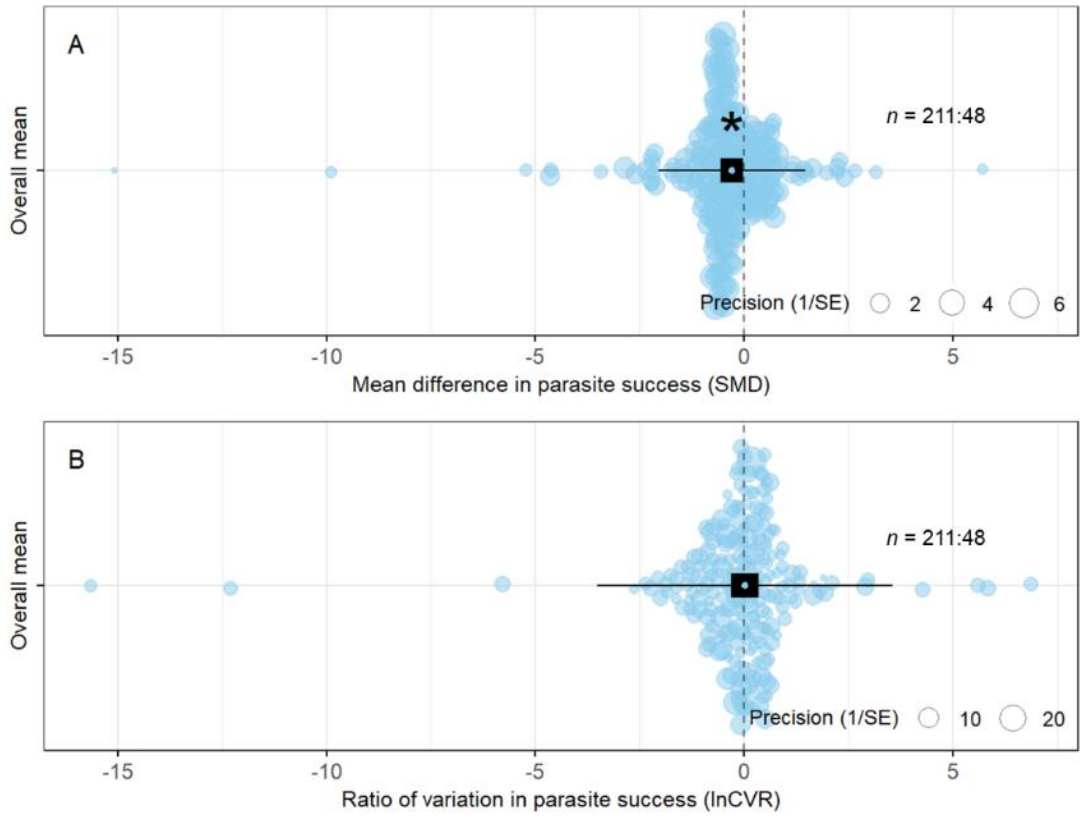



Figure 4.4. The overall effect of host population genetic diversity on the mean and variability in parasite success. The x-axis in each plot shows the effect that an increase in host population genetic diversity had on either A) the mean parasite success (SMD) or B) the variability in parasite success (lnCVR). The dashed line

indicates an effect size of zero where host population genetic diversity has no influence. Model means are shown with 95% confidence intervals (black rectangles and prediction intervals (thin black lines). Circles show individual effect sizes and are scaled according to the inverse of their standard error. n = sample size of the data (the number of effect sizes : the number of studies). The asterisk shows that the model means is significantly different from zero ($p < 0.05$). Forest plot alternatives are shown in the online supplementary material (Fig. S4.1).

### 4.4.4 Impacts of host population genetic variation on parasite success differ between multi-host and single-host parasites

Next, we investigated whether the effect of host population genetic variation on parasite success was influenced by two fundamental characteristics of the parasite: the host-specificity of the parasite and the likely genetic diversity of the parasite population studied.

In contrast to the overall effect of host population genetic diversity on mean parasite success, which was significantly negative (see above), separating the effects of host population genetic diversity by a combination of parasite population genetic diversity and host range showed that the effect of host population genetic diversity on mean parasite success was only significant for single host parasites (Fig. 4.5A). In contrast, there was no significant evidence of an effect of host-population genetic diversity on the mean success of multi-host parasites (Fig. 4.5A).

In addition, although there was no overall effect of host population genetic diversity on the variability in parasite success (see above), there was a significant difference in the effect of host population genetic diversity on the variability in the success of singe-host parasites with low versus high population genetic diversity (glht: $p = 0.03$; Fig. 4.5B). Specifically, increased host population genetic diversity lead to either an increase (lnCVR = -0.54, Fig. 4.5B) or decrease (lnCVR = 0.61, Fig. 4.5B) in the variability of the success of single-host parasites when their own population genetic diversity was either high or low.

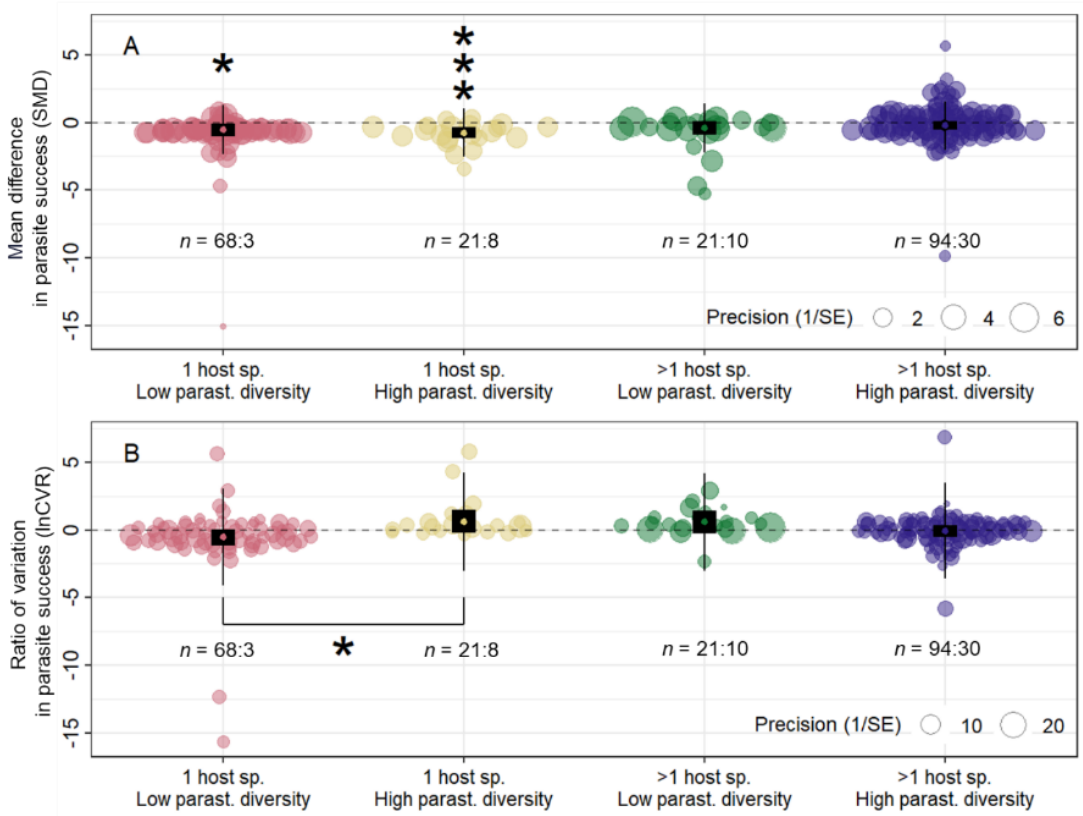



Figure 4.5. The influence of host range and parasite population genetic diversity on the effect of host population genetic diversity on the mean and variability in parasite success. The x-axis in each plot shows the effect of an increase in host population genetic diversity on either A) mean parasite success (SMD) or B) variability in parasite success (lnCVR). The dashed line indicates an effect size of zero where there is no influence of host population genetic diversity on parasite success. Model means are shown with 95% confidence intervals (black rectangles) and prediction intervals (thin black lines). Individual effect sizes (circles) are scaled according to the inverse of their standard error. n = sample size of the data (the number of effect sizes : the number of studies). The significance level of individual model means, as well as any pairwise contrasts, is indicated by one ($p < 0.05$) or three ($p < 0.001$) asterisks.

### 4.4.5 Context-dependent effect of host population genetic diversity on parasite success

We investigated how eight other aspects of study design (see hypotheses in Table 4.1) influenced the effect of host population genetic diversity on the mean and variability in parasite success (Fig. 4.6)

We found that the effect of host population genetic diversity was significantly negative on the mean success of microparasites (Fig. 4.6C), for inbred versus outbred hosts (Fig. 4.6D), for sexually reproducing hosts (SMD = -0.40, 95% CI = [-0.73, -0.07], p = 0.02, Fig. 4.6F), parasites which caused host mortality (SMD = -0.34, 95% CI = [-0.68, -0.00], p = 0.05, Fig. 4.6G) and non-lab based studies (SMD = -0.33, 95% CI = [-0.65, -0.01], p = 0.04, Fig. 4.6H). For the effect of host population genetic diversity on the variability in parasite success, we found that this was significantly negative for asexually reproducing hosts (Fig. 4.6N).

In addition, comparisons between specific levels of these moderators showed a highly significant difference in the mean difference in parasite success between micro- and macroparasites (QM = 13.2, df = 1, $p < 0.001$, Fig. 4.6) and also a significant difference in the ratio of variability in parasite success between both sexual hosts and either asexual (QM = 6.40, df = 1, p = 0.01, Fig. 4.6N) or facultatively sexual hosts (QM = 5.53 df = 1, p = 0.02, Fig. 4.6N).

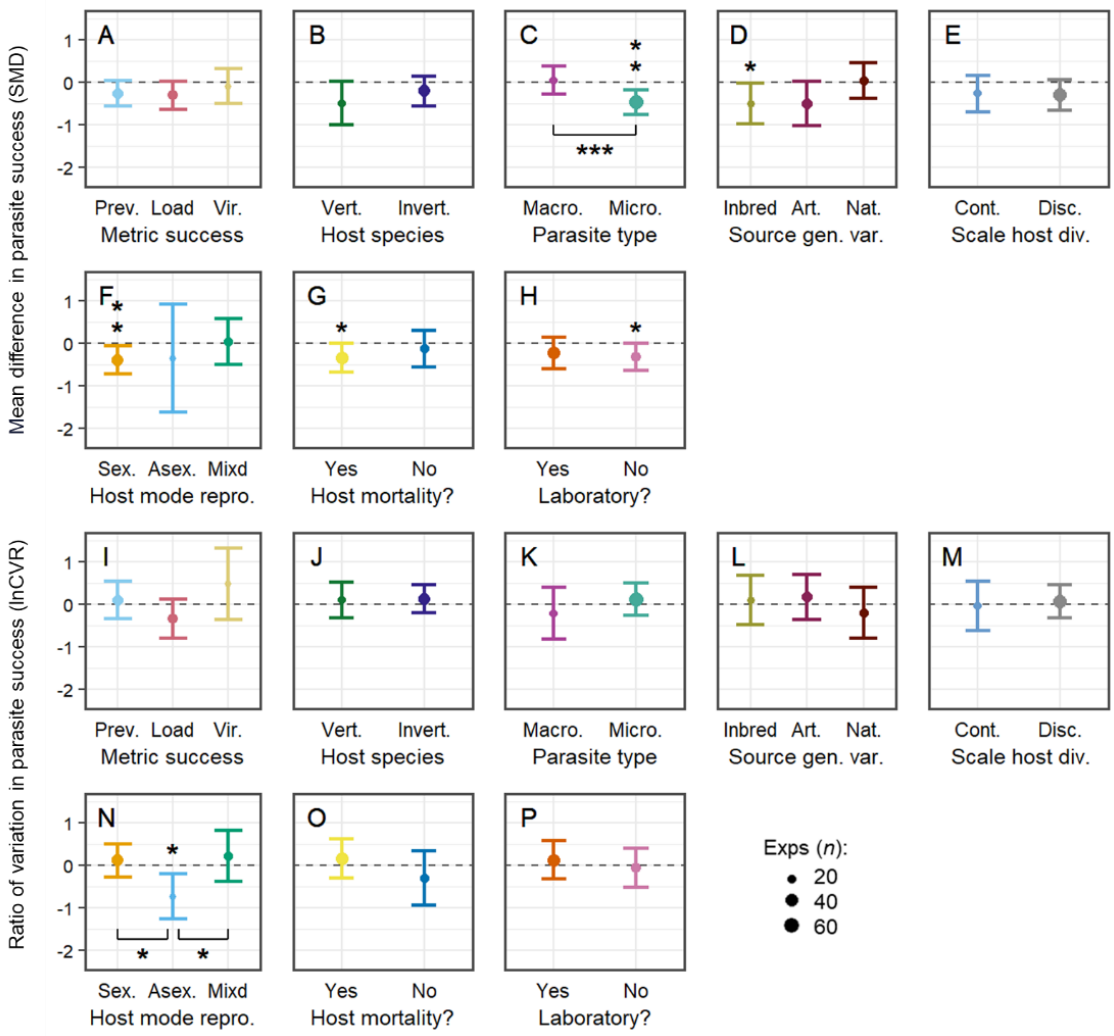



Figure 4.5. The context-dependence of the effect of host population genetic diversity on the mean and variability in parasite success. The y-axis in each plot shows the effect of an increase in host population genetic diversity on either the difference in mean parasite success (SMD) (panels A-H), or the difference in the variability in parasite success (lnCVR) (panels I-P). Model means are shown with 95% confidence intervals and are scaled according to the number of experiments. The dashed line indicates an effect size of zero. The significance level of individual model means, as well as any pairwise contrasts, is indicated by one ($p < 0.05$), two ($p < 0.01$) or three ($p < 0.001$) asterisks. The following abbreviations are used; Prev. (Prevalence), Vir. (Virulence), Vert. (Vertebrate), Invert. (Invertebrate), Macro. (Macroparasite), Micro. (Microparasite), Source gen. var. (Source of host genetic diversity), Art. (Artificial), Nat. (Natural), Scale host div. (Scale of host diversity), Cont. (Continuous), Disc. (Discrete), Host mode repro. (Mode of host reproduction), Sex. (Sexual), Asex. (Asexual), Mixd (Mixed).

2835

### 4.4.6 Our results are robust to leaving data out, but require the right 'mean' effect size

To test the robustness our of results, we performed a suite of 'leave-one-out' sensitivity analyses and remodelled our parasite success data with an alternative set of effect sizes.

The suite of sensitivity analyses showed that the results of our main effects were not dependent on the inclusion of a particular study (Table 4.3., Fig. S4.2) or independent comparison in our dataset (Table 4.3, Fig. S4.3). However, they were less robust to using the log response ratio (lnRR) as an alternative effect size to the standardized mean difference (SMD) to measure to effect of host population genetic diversity on mean parasite success. Although both measures showed a negative effect of host population genetic diversity on mean parasite success, the alternate way of measuring this was not significant (lnRR = 0.93, 95% CI = [-0.40, 2.25], n = 211). In comparison, the alternate variability measure, the log variability ratio (lnVR), supported the result of the main effect size (the log coefficient of variation ratio, lnCVR) by showing that there was no significant effect of host population genetic diversity on the variability in parasite success.

Table 4.3. Results of the leave-one-out sensitivity analyses. To test the robustness of the results using our main effect sizes, we re-modelled the data using an iterative exclusion of either one study (Leave1studyout) or one independent comparison (Leave1trtout) and calculated the mean model estimate and the mean p-value across all the models. The following abbreviations are used; ES (effect size), SE (mean standard error across all models), ci.lb and ci.ub (mean lower and upper bounds of 95% confidence intervals across all models respectively).

| Method | ES | Estimate | SE | z-value | p-value | ci.lb | ci.ub |
|---|---|---|---|---|---|---|---|
| Leave1trtout | SMD | -0.29 | 0.14 | -2.07 | 0.04 | -0.56 | -0.02 |
| Leave1trtout | lnCVR | 0.02 | 0.17 | 0.13 | 0.9 | -0.3 | 0.35 |
| Leave1studyout | SMD | -0.29 | 0.14 | -2.05 | 0.04 | -0.57 | -0.01 |
| Leave1studyout | lnCVR | 0.02 | 0.17 | 0.13 | 0.87 | -0.31 | 0.35 |

## 4.5 Discussion

By re-analysing the effect size data from two previous meta-analyses (Ekroth et al., 2019; Gibson & Nguyen, 2021) we show that conventional theory, which suggests that a high level of host population genetic diversity tends to limit the spread of disease in both wild and domestic animal (i.e. non-plant) populations (King & Lively, 2012), is only true some of the time. In fact, we show that the specific effect of host population genetic diversity on either metrics of the mean or variability in parasite infection success actually depend on a combination of both the host range of the parasite and its level of population genetic diversity. For instance, a high level of host population genetic diversity tends to limit metrics of mean infection success for specialist, but not generalist, parasites relative to a low level of host population genetic diversity, but also either increases or decreases the variability in metrics of specialist parasite infection success relative to a low level of host population genetic diversity depending on the corresponding level of parasite population genetic diversity. Therefore, the idea that a relatively higher level of host population genetic diversity tends to limit the spread of disease, and thus epidemic size, is not necessarily best described as ‘conventional’ wisdom.

Our results contrast those from previous meta-analytical studies of the effect of host population genetic diversity on metrics of parasite infection success, which also investigated the host range of the parasite as part of their analysis (Ekroth et al., 2019; Gibson & Nguyen, 2021). As already eluded to above, we found that the host range of the parasite was a significant moderator of the effect of host population genetic diversity on metrics of mean parasite infection success. Specifically, we found that a high level of host population genetic diversity tended to limit metrics of mean infection success for specialist, but not generalist, parasites relative to a low level of host population genetic diversity. One possible explanation for this is the increased statistical power of our study due to a larger number of effect sizes from combining the effect size data from these previous analyses (Gibson, 2022). This supports our original suggestion that parasite host range is closely related to the level of genetic specificity for infection (because specialist parasites are more likely to have evolved highly specific, matching-allele-type interactions between host resistance and parasite infectivity alleles than generalist parasites that are less tightly co-evolved to their range of hosts). It also highlights the susceptibility of host populations with only a small amount of genetic diversity to consistently high levels of infection success by specialist parasites.

Our finding that more diverse host populations tend to have smaller metrics of mean infection success for specialist parasites shows how there is slightly more complexity associated with conventional wisdom than previously thought (King & Lively, 2012). It also has important implications for how the level of host population genetic diversity is managed in species of conservation concern (Meuwissen et al., 2020). For example, one approach to species management may be to prioritise the maintenance or restoration of genetic diversity in host populations threatened by specialist parasite species, or by finding a safe approach for broadening a specialist parasite's host range. For example, the introduction of a novel host or parasite species, as some form of biological control (Stenberg et al., 2021), that can either act as a catalyst for host-mediated parasite evolution of greater generality (Bull et al., 2022) or cause a parasite host shift through direct competition for hosts (for a review, see Bashley, 2015) may broaden the host range of a specialist parasite away from its target host to include a non-target, pest species. In addition, recent empirical work has started to test the theory that high host population genetic diversity (*sensu* 'resource heterogeneity') selects for the evolution, or maintenance, of a broader parasite host range (*sensu* 'niche width', Gibson et al., 2020). Therefore, understanding how host population genetic diversity is linked to the evolution of parasite host range in a number of different host-parasite systems should be a priority for future research.

Again, in contrast to the results of previous meta-analytical studies (Ekroth et al., 2019; Gibson & Nguyen, 2021), we also found that there was a significant difference between the effect of host population genetic diversity on the variability in metrics of infection success for specialist parasites with a high level of parasite population genetic diversity and a low level of parasite population genetic diversity. Specifically, we showed that a high level of both host and parasite population genetic diversity increased the variability in metrics of infection success for specialist parasites relative to host populations with a low level of population genetic diversity, whereas a high level of host population genetic diversity and a low level of parasite population genetic diversity decreased the variability in metrics of infections success for specialist parasites relative to host populations with a low level of population genetic diversity. Although these previous meta-analytical studies focused on the mean, rather than the variability in metrics of parasite infection success (Ekroth et al., 2019; Gibson & Nguyen, 2021), nevertheless the authors of both studies had expected to find a significant effect of parasite population genetic diversity on the relationship

between host population genetic diversity and metrics of mean parasite infection success and were surprised that there was no such significant result (Ekroth et al., 2019; Gibson & Nguyen, 2021). In addition to their reduced statistical power (as mentioned above), one possible reason for this could be that only one out of two of these studies investigated the interaction between different moderators (Gibson & Nguyen, 2021). On the other hand, the difference we observed in the variability in metrics of infection success for specialist parasites between host populations with high parasite population genetic diversity and low parasite population genetic diversity matched the initial predictions we made in our proposed Diversity-Uncertainty theoretical model (Fig. 4.1). This confirms previous theories that the benefits of host population genetic diversity for resistance to disease depend on the corresponding parasite population genetic diversity (Bensch et al., 2021; Boomsma, 1996; Van Baalen & Beekman, 2006).

This idea that the combination of both host and parasite population genetic diversity influence the variability in metrics of parasite infection success has important implications for host-parasite systems in general. As already mentioned previously, not only could such variability in metrics of parasite infection success be important for predicting the occurrence of potentially severe disease epidemics, which could benefit conservation by informing genetic diversity management strategies to prioritise at risk host populations (Meuwissen et al., 2020), but it could also be central to our ability to protect against future emerging diseases (Altizer et al., 2006) and for understanding the extent to which disease experiments are repeatable. For example, our results highlight that host populations with a low level of genetic diversity are particularly susceptible to consistently large disease epidemics caused by specialist parasites with a high level of diversity. Conversely, the inconsistent levels of parasite success predicted for combinations of low host x low parasite and high host x high parasite population genetic diversity suggest that the repeatability of both laboratory and field experiments may be quite low, since they are often characterised respectively by such combinations of host-parasite diversity. Similarly, patterns of future disease occurrence (and emergence) may be more difficult to predict in such systems compared to those with different combinations of diversity.

In addition to our moderator analysis using models with an interaction term, we also investigated the effects of eight other contextual factors to evaluate our list of hypotheses (Table 4.1). These are the same as the moderators used in previous

meta-analyses (Ekroth et al., 2019; Gibson & Nguyen, 2021), but in comparison to the total number of significant effects they observed in their analysis (two), our results show that there were six moderator levels that had significant effects. In particular, the effect of host population genetic diversity on metrics of mean infection success for microparasites was much more negative than for macroparasites. In agreement with our original hypothesis, this suggests that the difference in infection durability between micro- and macroparasites (Sorci, 2014) affects the specificity of their interactions with the host (Schmid-Hempel & Ebert, 2003). Therefore, we suggest that macroparasites, due the longer-lasting nature of their infections (Sorci, 2014), are less tightly coevolved with their hosts and thus have a lower genetic specificity for infection. We also found that there was a significant negative effect of host population genetic diversity on metrics of mean parasite infection success for comparisons of outbred versus inbred hosts. Although such an effect was absent for other host population comparisons, such as between naturally high and low genetic diversity populations of hosts, it was quite similar to the effect for host populations composed of select genotypes. Therefore, this could suggest that experimental manipulations of host population genetic diversity had a stronger effect on metrics of mean parasite infection success than studies using a purely natural source of hosts. However, it is worth noting that this result is somewhat inconsistent with the significantly negative effect of host population genetic diversity on metrics of mean parasite infection success observed for non-laboratory-based studies, for which the opposite effect was observed in one out of the two previous meta-analyses (Ekroth et al., 2019). As such, an alternative explanation would be that the effect of host population genetic diversity on metrics of mean parasite infection success were exacerbated for outbred versus inbred hosts by the increased susceptibility of inbred hosts to disease (Coltman et al., 1999).

Other notable observations from the individual models include a significant negative effect of host population genetic diversity on metrics of mean parasite infection success in sexually reproducing host populations and a significant negative effect of host population genetic diversity on the variability in metrics of parasite infection success for host populations reproducing asexually, which was strongly contrasted against the absence of either a sexually reproducing or facultatively sexually reproducing host. These results suggest that sexual reproduction might contribute to the strength of how population genetic diversity limits disease spread due to greater dissimilarity between genotypes from genetic recombination than achieved by asexual reproduction (Hamilton et al., 1990), but also that asexual reproduction can

lead to greater disparity between the consistency of metrics of parasite infection success of host populations with high versus low genetic diversity than other forms of host reproduction. There was also a significant negative effect of host population genetic diversity on metrics of mean infection success for parasites that typically kill the host. Compared to less harmful parasites, this suggests that virulent parasites could select for higher levels of resistance and greater variation of resistance in the host population (Ekroth et al., 2019).

Despite the potentially exciting nature of our results, there are some additional considerations that should be taken into account. For example, there is a large number of effect sizes (68) for specialist parasites with low population genetic diversity, but most of these actually come from a prokaryotic bacterial host study (Van Houte et al., 2016), rather than a vertebrate or invertebrate host, which is the case for most of our data. In addition, the host range of the parasite may not be a reliable estimate of the genetic specificity for infection. The host range of the parasite was used as a proxy for the genetic specificity for infection, as such a detailed level of information was not available. Therefore, we made the prediction that highly specific interactions between host and parasite genotypes (Schmid-Hempel & Ebert, 2003) would be more likely for tightly coevolving pathogens (i.e. following a MAM of infection, Agrawal & Lively, 2002), as might be expected for specialist, but not generalist parasites. Similarly, the results of our moderator analysis rely on somewhat arbitrary ways of creating data sub-categories. In the case of parasite population genetic diversity, comparing mainly natural versus laboratory strains of parasites could be a poor indication of the effect of parasite population genetic diversity because the exact level of diversity was not actually quantified. In the case of the host range of the parasite, this measure is subjective and based somewhat on an incomplete literature (Hyman & Abedon, 2010).

One other final consideration is that the majority of our data concentrates on the effect of host population genetic diversity on both the mean and variability in metrics of parasite infection success for spatially replicated groups of host populations (but see Hale & Briskie, 2007). Although we might expect the temporal pattern of the effect of host population genetic diversity on metrics of parasite infection success to be similar to that observed across space, we also predict some key differences. For example, recurrent bouts of parasite-mediated directional selection have the ability to reduce host and parasite population genetic diversity over time (Buckling &

Rainey, 2002; Obbard et al., 2011), which could be accompanied by a higher mean and lower variability in metrics of parasite infection success. However, the maintaince of host and parasite genetic diversity over time depends on the precise nature of selection and the underlying host-parasite infection genetics (i.e. a MAM versus a GFG model for genetic specificity, Boots et al., 2014). Although there are some studies which measure metrics of parasite infection success for host populations with different levels of genetic diversity at multiple timepoints (e.g. Altermatt & Ebert, 2008), more studies would be required to provide a comprehensive test of the effect of host population genetic diversity on metrics of parasite infection success over time.

## 4.6 Summary

In this study, we measured the difference in the mean and variability in metrics of parasite infection success between host populations with high versus low genetic diversity. After first challenging so-called 'conventional wisdom' (*sensu* (King & Lively, 2012)) we proposed a Diversity-Uncertainty model to better understand the context around how host population genetic diversity might affect not only the mean, but also the variability in metrics of parasite infection success. We found that host population genetic diversity affected metrics of mean infection success for specialist but not generalist parasites. We also found that the effect of host population diversity on the variability in metrics of parasite infection success depends on a combination of the host range of the parasite and the parasite population diversity, such that there is some evidence for a Diversity-Uncertainty theoretical model, at least for the collection of studies reviewed in this meta-analysis. Additionally, we found that there was a number of other context dependent effects of host population genetic diversity on both the mean and variability in metrics of parasite infection success, such as parasite type. Overall, these findings represent a change of perspective that could help to protect vulnerable host populations by prioritizing how genetic diversity within these populations is managed. Future study of the Diversity-Uncertainty hypothesis across a range of plant host-parasite systems would help generalize these findings.

## 4.7 References

Agha, R., Gross, A., Rohrlack, T., & Wolinska, J. (2018). Adaptation of a chytrid parasite to its cyanobacterial host is hampered by host intraspecific

diversity. *Frontiers in Microbiology*, 9(MAY), 1–10. https://doi.org/10.3389/fmicb.2018.00921

Agrawal, A., & Lively, C. M. (2002). Infection genetics: Gene-for-gene versus matching-alleles models and all points in between. *Evolutionary Ecology Research*.

Alan Pounds, J., Bustamante, M. R., Coloma, L. A., Consuegra, J. A., Fogden, M. P. L., Foster, P. N., La Marca, E., Masters, K. L., Merino-Viteri, A., Puschendorf, R., Ron, S. R., Sánchez-Azofeifa, G. A., Still, C. J., & Young, B. E. (2006). Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature*, *439*(7073), 161–167. https://doi.org/10.1038/nature04246

Altermatt, F., & Ebert, D. (2008). Genetic diversity of *Daphnia magna* populations enhances resistance to parasites. *Ecology Letters*. https://doi.org/10.1111/j.1461-0248.2008.01203.x

Altizer, S., Dobson, A., Hosseini, P., Hudson, P., Pascual, M., & Rohani, P. (2006). Seasonality and the dynamics of infectious diseases. *Ecology Letters*, *9*(4), 467–484. https://doi.org/10.1111/j.1461-0248.2005.00879.x

Anderson, R. M., May, R. M., Joysey, K., Mollison, D., Conway, G. R., Cartwell, R., Thompson, H. V., & Dixon, B. (1986). The invasion, persistence and spread of infectious diseases within animal and plant communities. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *314*(1167), 533–570. https://doi.org/10.1098/rstb.1986.0072

Baer, B., & Schmid-Hempel, P. (2001). Unexpected consequences of polyandry for parasitism and fitness in the bumblebee, *Bombus terrestris*. *Evolution*. https://doi.org/10.1111/j.0014-3820.2001.tb00683.x

Bashey, F. (2015). Within-host competitive interactions as a mechanism for the maintenance of parasite diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1675), 20140301. https://doi.org/10.1098/rstb.2014.0301

Bensch, H. M., O'Connor, E. A., & Cornwallis, C. K. (2021). Living with relatives offsets the harm caused by pathogens in natural populations. *ELife*, *10*. https://doi.org/10.7554/eLife.66649

Boomsma, J. J. (1996). Paternity in eusocial Hymenoptera. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1342), 947–975. https://doi.org/10.1098/rstb.1996.0087

Boots, M., White, A., Best, A., & Bowers, R. (2014). How specificity and epidemiology drive the coevolution of static trait diversity in hosts and parasites. *Evolution*, *68*(6), 1594–1606. https://doi.org/10.1111/evo.12393

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to meta-analysis. In *Introduction to Meta-Analysis*. https://doi.org/10.1002/9780470743386

Browning, J. A., & Frey, K. J. (1969). Multiline cultivars as a means of disease control. *Annual Review of Phytopathology*, *7*(1), 355–382. https://doi.org/10.1146/annurev.py.07.090169.002035

Buckling, A., & Rainey, P. B. (2002). Antagonistic coevolution between a bacterium and a bacteriophage. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1494), 931–936. https://doi.org/10.1098/rspb.2001.1945

Bull, J. J., Wichman, H. A., & Krone, S. M. (2022). Modeling the directed evolution of broad host range Phages. *Antibiotics*, *11*(12), 1709. https://doi.org/10.3390/antibiotics11121709

Coltman, D. W., Pilkington, J. G., Smith, J. A., & Pemberton, J. M. (1999). Parasite-mediated selection against inbred Soay sheep in a free-living, island population. In *Evolution* (Vol. 53, Issue 4). https://about.jstor.org/terms

Dagan, Y., Liljeroos, K., Jokela, J., & Ben-Ami, F. (2013). Clonal diversity driven by parasitism in a freshwater snail. *Journal of Evolutionary Biology*, *26*(11). https://doi.org/10.1111/jeb.12245

De Castro, F., & Bolker, B. (2004). Mechanisms of disease-induced extinction. *Ecology Letters*, *8*(1), 117–126. https://doi.org/10.1111/j.1461-0248.2004.00693.x

Dionne, M., Miller, K. M., Dodson, J. J., & Bernatchez, L. (2009). MHC standing genetic variation and pathogen resistance in wild Atlantic salmon. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1523). https://doi.org/10.1098/rstb.2009.0011

Dybdahl, M. F., Jenkins, C. E., & Nuismer, S. L. (2014). Identifying the molecular basis of host-parasite coevolution: Merging models and mechanisms. *The American Naturalist*, *184*(1), 1–13. https://doi.org/10.1086/676591

Ekroth, A. K. E., Rafaluk-Mohr, C., & King, K. C. (2019). Host genetic diversity limits parasite success beyond agricultural systems: a meta-analysis. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1911), 20191811. https://doi.org/10.1098/rspb.2019.1811

Ellison, A., Cable, J., & Consuegra, S. (2011). Best of both worlds? Association between outcrossing and parasite loads in a selfing fish. *Evolution*, *65*(10), 3021–3026. https://doi.org/10.1111/j.1558-5646.2011.01354.x

Elton, C. S. (1958). *The Ecology of Invasions by Animals and Plants*. Springer US. https://doi.org/10.1007/978-1-4899-7214-9

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, *63*(3). https://doi.org/10.1348/000711010X502733

Field, S. G., Lange, M., Schulenburg, H., Velavan, T. P., & Michiels, N. K. (2007). Genetic diversity and parasite defense in a fragmented urban metapopulation of earthworms. *Animal Conservation*, *10*(2). https://doi.org/10.1111/j.1469-1795.2006.00084.x

Ganz, H. H., & Ebert, D. (2010). Benefits of host genetic diversity for resistance to infection depend on parasite diversity. In *Source: Ecology* (Vol. 91, Issue 5). https://www-jstor-org.ezproxy.stir.ac.uk/stable/pdf/25661177.pdf?refreqid=excelsior%3A5c0f30d81cb9568626bbc1823fa625af

Garrett, K. A., & Mundt, C. C. (1999). Epidemiology in mixed host populations. *Phytopathology*, *89*(11), 984–990. https://doi.org/10.1094/PHYTO.1999.89.11.984

Gibson, A. K. (2022). Genetic diversity and disease: The past, present, and future of an old idea. *Evolution*, *76*(S1), 20–36. https://doi.org/10.1111/evo.14395

Gibson, A. K., Baffoe-Bonnie, H., Penley, M. J., Lin, J., Owens, R., Khalid, A., & Morran, L. T. (2020). The evolution of parasite host range in heterogeneous host populations. *Journal of Evolutionary Biology*, *33*(6), 773–782. https://doi.org/10.1111/jeb.13608

Gibson, A. K., & Nguyen, A. E. (2020). Does genetic diversity protect host populations from parasites? A meta-analysis across natural and agricultural systems. *Evolution Letters*, *5*(1). https://doi.org/10.1002/evl3.206

Giese, A. R., & Hedrick, P. W. (2003). Genetic variation and resistance to a bacterial infection in the endangered Gila topminnow. *Animal Conservation*, *6*(4). https://doi.org/10.1017/S1367943003003445

Hale, K. A., & Briskie, J. V. (2007). Decreased immunocompetence in a severely bottlenecked population of an endemic New Zealand bird. *Animal Conservation*, *10*(1). https://doi.org/10.1111/j.1469-1795.2006.00059.x

Hamilton, W. D. (1987). *Kinship, recognition, disease, and intelligence: constraints of social evolution*. Japan Scientific Societies Press.

Higgins, J. P. T., Thomas J., Chandler J., Cumpston M., Li T., Page M. J., Welch V. A. (2024). *Cochrane Handbook for Systematic Reviews of Interventions* (Version 5.) The Cochrane Collaboration, 2011. www.handbook.cochrane.org

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363. https://doi.org/10.1002/bimj.200810425

Hyman, P., & Abedon, S. T. (2010). Bacteriophage host range and bacterial resistance. *Advances in Applied Microbiology*, *70*, 217–248. https://doi.org/10.1016/S0065-2164(10)70007-1

Jennions, M. D. (2003). A survey of the statistical power of research in behavioral ecology and animal behavior. *Behavioral Ecology*, *14*(3), 438–445. https://doi.org/10.1093/beheco/14.3.438

Johnson, P. T. J., Ostfeld, R. S., & Keesing, F. (2015). Frontiers in research on biodiversity and disease. *Ecology Letters*, *18*(10), 1119–1133. https://doi.org/10.1111/ele.12479

Johnson, P. T. J., Preston, D. L., Hoverman, J. T., Henderson, J. S., Paull, S. H., Richgels, K. L. D., & Redmond, M. D. (2012). Species diversity reduces parasite infection through cross-generational effects on host abundance. *Ecology*, *93*(1), 56–64. https://doi.org/10.1890/11-0636.1

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T., & Ostfeld, R. S. (2010). Impacts of biodiversity on the emergence and transmission of infectious diseases. In *Nature*. https://doi.org/10.1038/nature09575

Keesing, F., Holt, R. D., & Ostfeld, R. S. (2006). Effects of species diversity on disease risk. *Ecology Letters*, *9*(4), 485–498. https://doi.org/10.1111/j.1461-0248.2006.00885.x

Keesing, F., & Ostfeld, R. S. (2021). Dilution effects in disease ecology. *Ecology Letters*, *24*(11). https://doi.org/10.1111/ele.13875

King, K. C., Jokela, J., & Lively, C. M. (2011). Parasites, sex, and clonal diversity in natural snail populations. *Evolution*, *65*(5). https://doi.org/10.1111/j.1558-5646.2010.01215.x

King, K. C., & Lively, C. M. (2012). Does genetic diversity limit disease spread in natural host populations? *Heredity*, *109*(4), 199–203. https://doi.org/10.1038/hdy.2012.33

Koricheva, J., & Gurevitch, J. (2014). Uses and misuses of meta-analysis in plant ecology. *Journal of Ecology*, *102*(4), 828–844. https://doi.org/10.1111/1365-2745.12224

Koskella, B., & Lively, C. M. (2009). Evidence for negative frequency-dependent selection during experimental coevolution of a freshwater snail and a sterilizing trematode. *Evolution*, *63*(9), 2213–2221. https://doi.org/10.1111/j.1558-5646.2009.00711.x

Kyle, C. J., Rico, Y., Castillo, S., Srithayakumar, V., Cullingham, C. I., White, B. N., & Pond, B. A. (2014). Spatial patterns of neutral and functional genetic variations reveal patterns of local adaptation in raccoon (*Procyon lotor*) populations exposed to raccoon rabies. *Molecular Ecology*, *23*(9). https://doi.org/10.1111/mec.12726

Lagisz, M., Senior, A., Nakagawa, S., & Noble, D. (2024). *_metaAidR: Functions and Data for Aiding Meta-analyses in Ecology and Evolution_* (0.0.0.9000).

Leonard, K. J. (1969). Factors affecting rates of stem rust increase in mixed plantings of susceptible and resistant oat varieties cultivars. *Phytopathology*, *59*, 1845–1850.

Loiseau, C., Zoorob, R., Robert, A., Chastel, O., Julliard, R., & Sorci, G. (2011). *Plasmodium relictum* infection and MHC diversity in the house sparrow (*Passer domesticus*). *Proceedings of the Royal Society B: Biological Sciences*, *278*(1709). https://doi.org/10.1098/rspb.2010.1968

Meagher, S. (1999). Genetic diversity and *capillaria hepatica* (nematoda) prevelance in Michigan deer mouse populations. *Evolution*, *53*(4), 1318–1324. https://doi.org/10.2307/2640837

Meuwissen, T. H. E., Sonesson, A. K., Gebregiwergis, G., & Woolliams, J. A. (2020). Management of genetic diversity in the era of genomics. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.00880

Mihaljevic, J. R., Joseph, M. B., Orlofske, S. A., & Paull, S. H. (2014). The scaling of host density with richness affects the direction, shape, and detectability of diversity-disease relationships. *PLoS ONE*, *9*(5), e97812. https://doi.org/10.1371/journal.pone.0097812

Nakagawa, S., Noble, D. W. A., Lagisz, M., Spake, R., Viechtbauer, W., & Senior, A. M. (2023). A robust and readily implementable method for the meta-analysis of response ratios with and without missing standard deviations. *Ecology Letters*, *26*(2), 232–244. https://doi.org/10.1111/ele.14144

Nakagawa, S., Poulin, R., Mengersen, K., Reinhold, K., Engqvist, L., Lagisz, M., & Senior, A. M. (2015). Meta-analysis of variation: ecological and evolutionary applications and beyond. *Methods in Ecology and Evolution*, *6*(2), 143–152. https://doi.org/10.1111/2041-210X.12309

Neumann, P., & Moritz, R. F. A. (2000). Testing genetic variance hypotheses for the evolution of polyandry in the honeybee (*Apis mellifera* L.). *Insectes Sociaux*, *47*(3). https://doi.org/10.1007/PL00001714

Obbard, D. J., Jiggins, F. M., Bradshaw, N. J., & Little, T. J. (2011). Recent and recurrent selective sweeps of the antiviral RNAi gene Argonaute-2 in three species of *Drosophila*. *Molecular Biology and Evolution*, *28*(2), 1043–1056. https://doi.org/10.1093/molbev/msq280

O'Brien, S. J., & Evermann, J. F. (1988). Interactive influence of infectious disease and genetic diversity in natural populations. In *Trends in Ecology and Evolution* (Vol. 3, Issue 10). https://doi.org/10.1016/0169-5347(88)90058-4

Ostfeld, R. S., & Keesing, F. (2012). Effects of host diversity on infectious disease. *Annual Review of Ecology, Evolution, and Systematics*, *43*(1), 157–182. https://doi.org/10.1146/annurev-ecolsys-102710-145022

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, Luke A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Paplauskas, S., Duthie, B., & Tinsley, M. C. (2024). Meta-analysis of host population-level diversity and variation in 'parasite success' [Dataset]. *Dryad*. https://doi.org/https://doi.org/10.5061/dryad.2bvq83bzq

Parsche, S., & Lattorff, M. H. G. (2018). The relative contributions of host density and genetic diversity on prevalence of a multi-host parasite in bumblebees. *Biological Journal of the Linnean Society*, *125*(4). https://doi.org/10.1093/biolinnean/bly151

Pierce, A. A., De Roode, J. C., Altizer, S., & Bartel, R. A. (2014). Extreme heterogeneity in parasitism despite low population genetic structure among monarch butterflies inhabiting the Hawaiian Islands. *PLoS ONE*, *9*(6). https://doi.org/10.1371/journal.pone.0100061

Puurtinen, M., Hytönen, M., Knott, K. E., Taskinen, J., Nissinen, K., & Kaitala, V. (2004). The effects of mating system and genetic variability on susceptibility to trematode parasites in a freshwater snail, *Lymnaea stagnalis*. *Evolution*, *58*(12). https://doi.org/10.1111/j.0014-3820.2004.tb01626.x

Queirós, J., Vicente, J., Alves, P. C., de la Fuente, J., & Gortazar, C. (2016). Tuberculosis, genetic diversity and fitness in the red deer, *Cervus elaphus*. *Infection, Genetics and Evolution*, *43*. https://doi.org/10.1016/j.meegid.2016.05.031

R Core Team. (2023). *_R: A Language and Environment for Statistical Computing_*. R Foundation for Statistical Computing. https://www.r-project.org/

Rahn, A. K., Krassmann, J., Tsobanidis, K., MacColl, A. D. C., & Bakker, T. C. M. (2016). Strong neutral genetic differentiation in a host, but not in its parasite. *Infection, Genetics and Evolution*, *44*. https://doi.org/10.1016/j.meegid.2016.07.011

Savage, A. E., Becker, C. G., & Zamudio, K. R. (2015). Linking genetic and environmental factors in amphibian disease risk. *Evolutionary Applications*, *8*(6). https://doi.org/10.1111/eva.12264

Schmid-Hempel, P., & Ebert, D. (2003). On the evolutionary ecology of specific immune defence. *Trends in Ecology & Evolution*, *18*(1), 27–32. https://doi.org/10.1016/S0169-5347(02)00013-7

Schmidt, A. M., Linksvayer, T. A., Boomsma, J. J., & Pedersen, J. S. (2011). No benefit in diversity? The effect of genetic variation on survival and disease resistance in a polygynous social insect. *Ecological Entomology*, *36*(6). https://doi.org/10.1111/j.1365-2311.2011.01325.x

Sherman, P. W., Seeley, T. D., & Reeve, H. K. (1988). Parasites, pathogens, and polyandry in social Hymenoptera. In *Source: The American Naturalist* (Vol. 131, Issue 4). https://www.jstor.org/stable/pdf/2461747.pdf?refreqid=excelsior%3A10d16706d07e226edd1f3ace6f23a6b4

Sorci, G. (2014). Immunology of parasitism. In *Reference Module in Biomedical Sciences*. Elsevier. https://doi.org/10.1016/B978-0-12-801238-3.00253-1

Stenberg, J. A., Sundh, I., Becher, P. G., Björkman, C., Dubey, M., Egan, P. A., Friberg, H., Gil, J. F., Jensen, D. F., Jonsson, M., Karlsson, M., Khalil, S.,

Ninkovic, V., Rehermann, G., Vetukuri, R. R., & Viketoft, M. (2021). When is it biological control? A framework of definitions, mechanisms, and classifications. *Journal of Pest Science*, *94*(3), 665–676. https://doi.org/10.1007/s10340-021-01354-7

Sutton, A. J. (2009). Publication bias. In *The handbook of research synthesis and meta-analysis.* (2nd ed., pp. 435–452). Russell Sage Foundation.

Trouvé, S., Degen, L., Renaud, F., & Goudet, J. (2003). Evolutionary implications of a high selfing rate in the freshwater snail *Lymnaea truncatula*. *Evolution*, *57*(10). https://doi.org/10.1554/02-452

Van Baalen, M., & Beekman, M. (2006). The costs and benefits of genetic heterogeneity in resistance against parasites in social insects. In *American Naturalist*. https://doi.org/10.1086/501169

van der Plank, J. E. (1963). *Plant diseases: epidemics and control*. Academic Press.

Van Houte, S., Ekroth, A. K. E., Broniewski, J. M., Chabas, H., Ashby, B., Bondy-Denomy, J., Gandon, S., Boots, M., Paterson, S., Buckling, A., & Westra, E. R. (2016). The diversity-generating benefits of a prokaryotic adaptive immune system. *Nature*, *532*(7599). https://doi.org/10.1038/nature17436

Velavan, T. P., Weller, S., Schulenburg, H., & Michiels, N. K. (2009). High genetic diversity and heterogeneous parasite load in the earthworm *Lumbricus terrestris* on a German meadow. *Soil Biology and Biochemistry*, *41*(8). https://doi.org/10.1016/j.soilbio.2009.03.026

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, *36*(3). https://doi.org/10.18637/jss.v036.i03

Whitehorn, P. R., Tinsley, M. C., Brown, M. J. F., Darvill, B., & Goulson, D. (2011). Genetic diversity, parasite prevalence and immunity in wild bumblebees. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1709), 1195–1202. https://doi.org/10.1098/rspb.2010.1550

Whitehorn, P. R., Tinsley, M. C., Brown, M. J. F., Darvill, B., & Goulson, D. (2014). Genetic diversity and parasite prevalence in two species of bumblebee. *Journal of Insect Conservation*, *18*(4). https://doi.org/10.1007/s10841-014-9673-1

Whiteman, N. K., Matson, K. D., Bollmer, J. L., & Parker, P. G. (2006). Disease ecology in the Galápagos Hawk (*Buteo galapagoensis*): host genetic diversity, parasite load and natural antibodies. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1588), 797–804. https://doi.org/10.1098/rspb.2005.3396

## 4.8 Data accessibility

Reviewer URL:
https://datadryad.org/stash/share/E2NqLZ8KL2oYaPQLSYatmNNIeUub3aeC4ExfjJvE5Hw
Data available from the Dryad Digital Repository: doi:10.5061/dryad.2bvq83bzq (Paplauskas et al., n.d.).

## 4.9 Acknowledgements

We are grateful to Assistant Professor Alexander Strauss, Dr Camille Ameline, Dr Peter Thrall, Dr Jason Walsman and Assistant Professor Amanda Kyle Gibson for sharing their raw data and clarifying their methods with us. We also thank Dr Laura Braunholtz, Dr Ka Raines, Dr Luc Bussière and Kyle Morrison for their general statistical and meta-analytical advice.

## 4.10 Supplementary information

Supplementary table S4.1. The difference between our amended study inclusion criteria and the original study inclusion criteria.

| New study inclusion criteria | Original study inclusion criteria | Why changed |
|---|---|---|
| 1) 'Parasite success', which we define as the ability of a parasite to spread among hosts (transmission rate, infection rate, prevalence), replicate on / within hosts (macro / microparasite load, disease severity), or kill hosts (virulence i.e. host survival / mortality rate) was measured among replicate populations across time or space. | - Ekroth, Rafaluk-Mohr and King, 2019: Defined parasite success as any measure of a parasite's ability to proliferate within a host population.<br>- Gibson and Nguyen, 2020: Focused on population-level parasitism, including prevalence, load and virulence. | We combined the two previous versions of the study inclusion criteria to include several different measures of parasite success, which were later used for contextual factor analysis. |

| | | |
|---|---|---|
| 2) Parasite success data was collected from two or more host populations with any comparable difference in genetic diversity, such as the level of relatedness among individuals (inbred versus outbred), genotypic diversity (high versus low) or heterozygosity. | - Ekroth, Rafaluk-Mohr and King, 2019: Data was collected from any study with two distinct populations and any measured difference in diversity.<br>- Gibson and Nguyen, 2020: Collected data for two or more populations. | We collected data from studies of multiple populations with any comparable difference in genetic diversity to increase our sample size and because there was one study with differences in genetic diversity which were not comparable between all pairwise combinations (Baer 2001). |
| 3) Genetic diversity was measured at the host population level and not community diversity or individual-level genetic heterozygosity. | - Ekroth, Rafaluk-Mohr and King, 2019: Used the exact same wording.<br>- Gibson and Nguyen, 2020: Stated that host genetic diversity had to be intra-specific. | We followed both Ekroth, Rafaluk-Mohr and King, 2019 and Gibson and Nguyen 2020 in this criterion. |
| 4) The study focused on an animal (or bacterial) host species. | - Ekroth, Rafaluk-Mohr and King, 2019: Excluded studies of agricultural systems.<br>- Gibson and Nguyen, 2020: Did not specify the study system. | We did not include any studies of non-animal populations, except for prokaryotic bacteria, because a more detailed analysis of the plant literature would require a separate review. |
| 5) The study does not re-analyze the data from a previously published study. | - Both Ekroth, Rafaluk-Mohr and King, 2019 and Gibson and Nguyen, 2020: Did not include this specification. | We included this specification because Ekroth, Rafaluk-Mohr and King, 2019 included data from two different studies by Baer and Schmid-Hempel which were |

| | | |
|---|---|---|
| | | based on the same dataset. |
| 6) The parasite success data was not replicated simply by using an alternate way of measuring host population diversity. | - Both Ekroth, Rafaluk-Mohr and King, 2019 and Gibson and Nguyen, 2020: Did not include this specification. | We included this specification because there two studies included by the previous meta-analyses (Giese 2003 and Puurtinen 2004) which included parasite success data for the same populations with two different measures of genetic diversity, which was a form of pseudoreplication. |
| 7) An attempt to take the parasite success data from clearly illegible figures was not made. | - Both Ekroth, Rafaluk-Mohr and King, 2019 and Gibson and Nguyen, 2020: Did not include this specification. | We included this specification because Gibson and Nguyen, 2020 had collected data from two studies with illegible figures (Agha, 2018 and van Houte et al. 2016. |

3378

3379

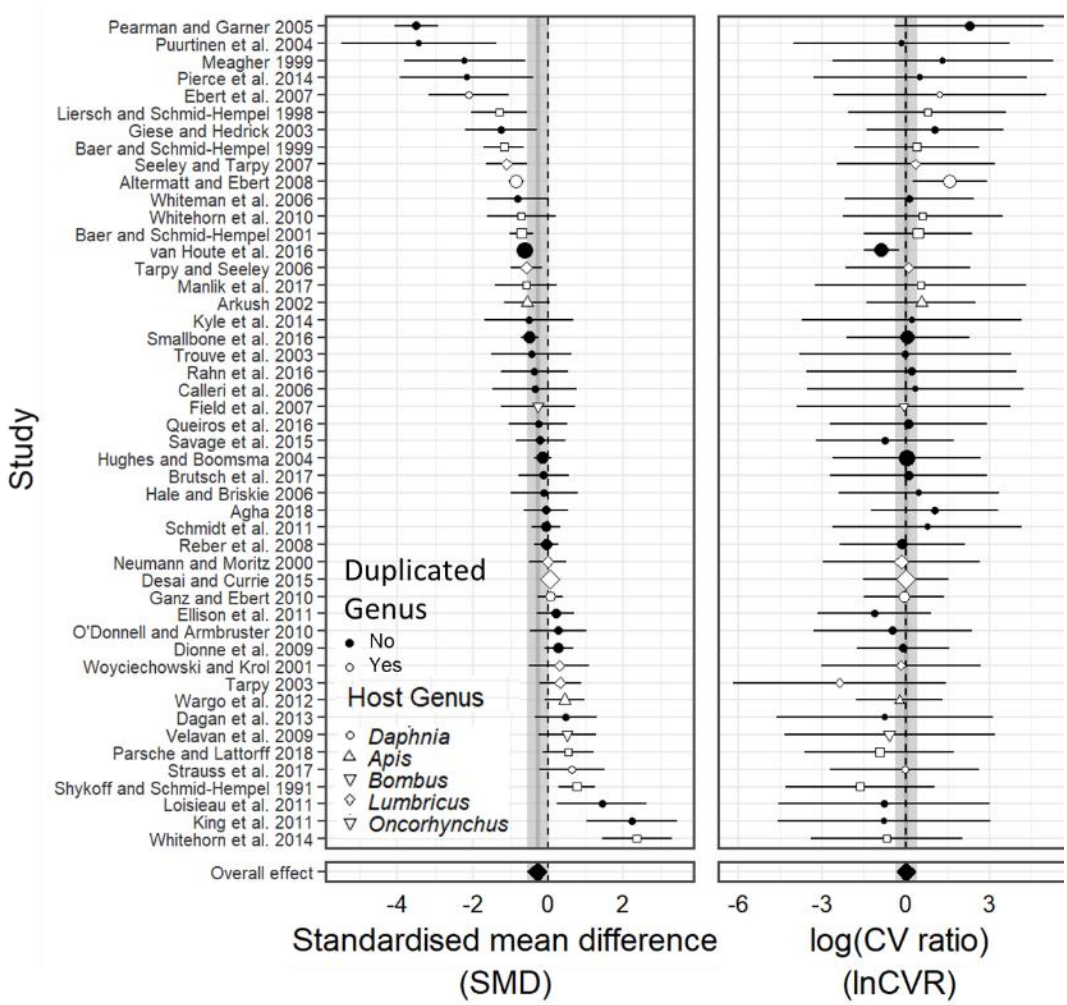



Figure S4.1. Study effects of host population genetic diversity on the mean and variability in parasite success. The x-axis in each plot shows the effect of increasing host population genetic diversity on either A) the difference in mean parasite success (SMD) or B) the difference in the variability in parasite success (lnCVR). Aggregated effects for each study are shown with 95% confidence intervals. Where the same host genus was studied more than once ('Duplicated Genus'), the colour of the points is white, rather than black, and the specific host genus studied is indicated by its shape (there were only five duplicated host genera). Each point is scaled by the amount of weighting they received in an aggregated mixed effects model, whereas the actual analysis was conducted based on the full set of 211 individual data points. The dashed lines indicate an effect size of zero and the overall model means are shown by the solid grey line with 95% confidence intervals bands in light grey.

3380

3381

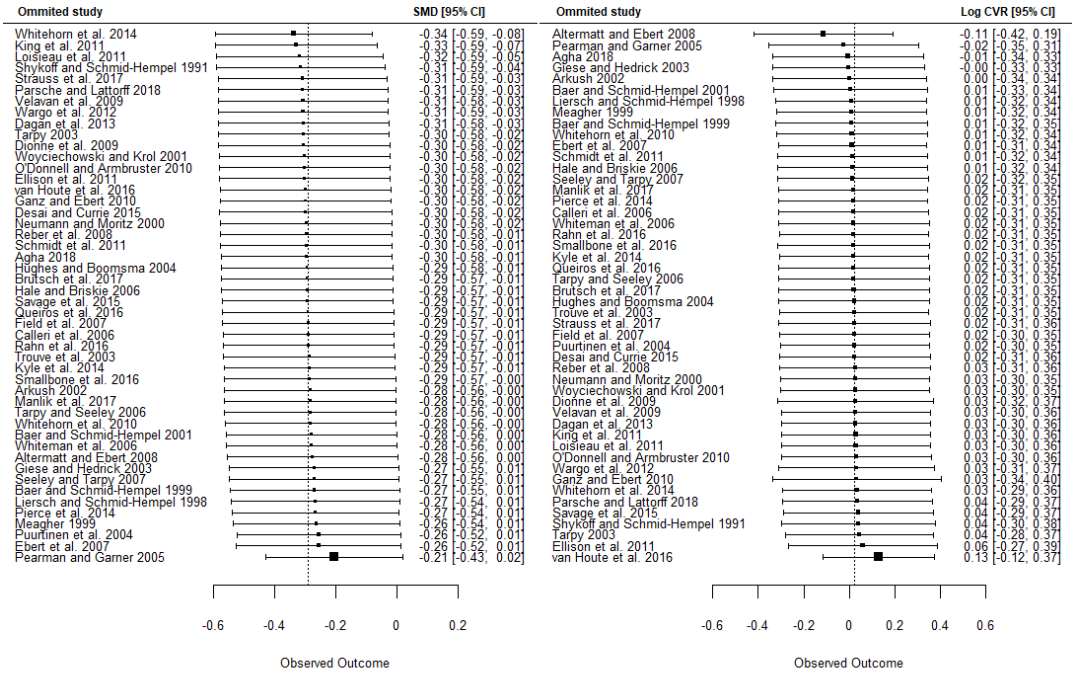



Figure S4.2. The results of the leave-one-study-out method of sensitivity analysis. The x-axis in each plot shows the effect of increasing host population genetic diversity on either A) the difference in mean parasite success (SMD) or B) the difference in the variability in parasite success (lnCVR). The names of the authors and the publication date for the study omitted in each model iteration is shown on the left, with the overall effect size and its confidence interval shown on in the middle. The mean effect size across all models is shown by the vertical line and specific value is shown on the right (with 95% confidence intervals). The size of each point is scaled according to its precision.

3382

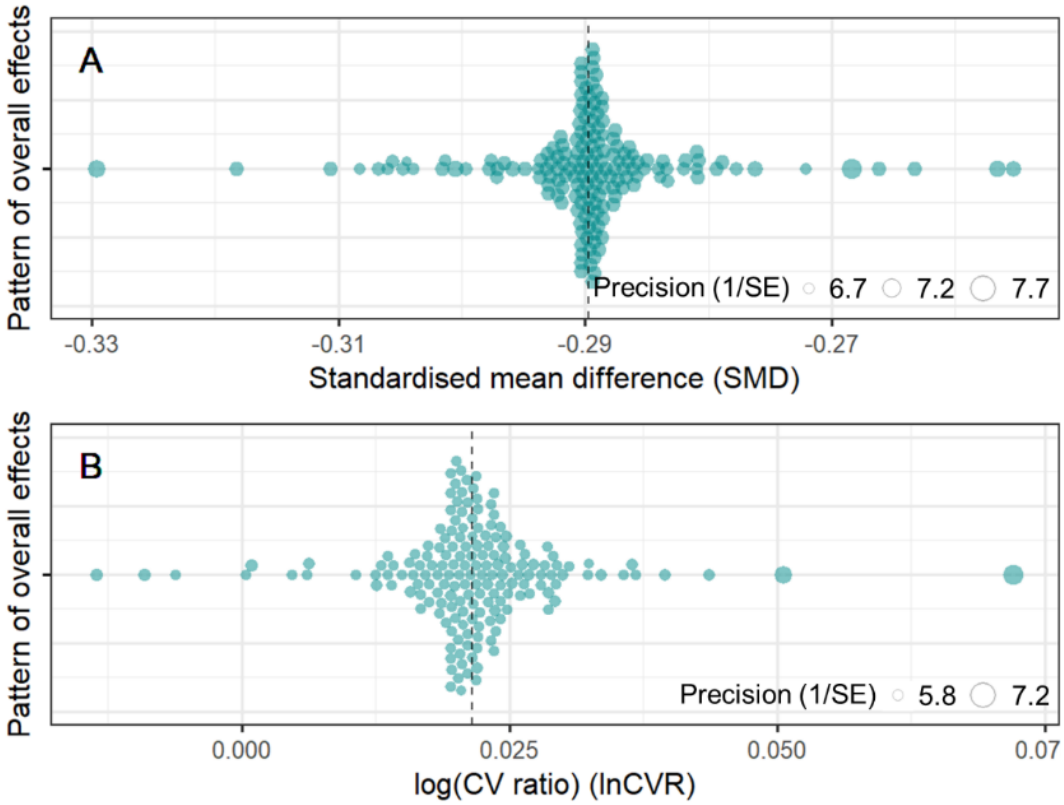



Figure S4.3. The results of the leave-one-independent-comparison-out method of sensitivity analysis visualized using a modified version of an orchard plot. The x-axis in each plot shows the effect of increasing host population genetic diversity on either A) the difference in mean parasite success (SMD) or B) the difference in the variability in parasite success (lnCVR). Unlike traditional orchard plots, which show the distribution of individual effect sizes, the mean effect size for each model iteration is shown by the coloured circles. The size of each point is scaled by its precision (inverse of the standard error). The mean effect size across all models is shown by the dashed line.

3383

3384

## 5. The ability of non-locally adapted hosts to outcompete resident hosts in wild populations

## 5.1 Abstract

Local adaptation is the process whereby the individuals within a population adapt to their local environment by evolving traits that increase their fitness in that environment relative to others. Local adaptation is expected to play a key role in protecting species from climate change and can affect the strength of species interactions, such as competition and parasitism. However, how local adaptation to environmental conditions influences competition between groups of local and migrant individuals from the same species is poorly understood. In addition, how this intra-specific competition is influenced by parasitism is also unclear. Therefore, to address this uncertainty, we performed a host reciprocal transplant experiment across 12 artificial pond populations of a naturally coevolving *Daphnia* host - parasite system. Animals were grouped into separate Home, Away and Mixed field cages within each pond and exposed to an ancestral parasite (with a control group). Specifically, we measured the ability of resident hosts to withstand competition from migrant hosts by comparing host fitness, in terms of the number of offspring, in both home and away environments versus a mixture of animals from different origins. Surprisingly, resident hosts were not locally adapted, and despite quite a large, but not statistically significant, reduction in the fitness of hosts between the mean of the unmixed and mixed categories in the third week of data collection, this was only statistically significant for the parasite-exposed treatment and not the parasite-free control. Therefore, this suggested that the cost of mixed competition for resident hosts was revealed by the addition of the ancestral parasite as a general stressor. The ability of resident hosts that are not locally adapted to outcompete migrant genotypes under parasite exposure may promote gene flow and decrease the size and severity of future disease outbreaks by increasing the capacity of host population genetic diversity to reduce transmission.

## 5.2 Introduction

Local adaptation is the process whereby the individuals within a population adapt to their local environment by evolving traits that increase their fitness in that environment relative to others (Kawecki & Ebert, 2004). Since the natural range of many species, such as insect vectors of disease (Sternberg & Thomas, 2014), are expected to shift in response to a warming climate (Price et al., 2019), and locally maladapted populations are vulnerable to extinction (Bocedi et al., 2013), local adaptation could increase the displacement of local populations by competitive exclusion and potentially play a significant role in how species respond to climate change (Aitken & Whitlock, 2013; Meek et al., 2023; Peterson et al., 2019).

What exactly defines this environment depends on the aspect of local adaptation in question. For example, local adaptation can refer to how well-adapted individuals are to the biotic or abiotic environment, such as predation or temperature and food availability respectively (Blanquart et al., 2013; Kawecki & Ebert, 2004). A special case of local adaptation is found in host-parasite systems, where the parasites are expected to be locally adapted to their hosts most of the time because they generally evolve faster than their hosts (Gandon, 2002; Greischar & Koskella, 2007; Hoeksema & Forde, 2008). This means that in examples of local adaptation, other than those driven by host-parasite antagonistic coevolution, locals are generally expected to be better adapted than immigrants to their local environment (Blanquart et al., 2012; Hereford, 2009; Holt & Gomulkiewich, 1997; Lascoux et al., 2016; Reger et al., 2018). However, in patterns of local adaptation driven by such antagonistic coevolution between hosts and parasites, there is the opposite expectation that immigrants are better adapted than residents to the local environment (Gandon & Nuismer, 2009; Morgan et al., 2005; Refardt & Ebert, 2007; Schulte et al., 2011).

The ability of migrant hosts to outcompete locals may depend on the strength of intra versus inter-population intra-specific competition. This is analogous to the competitive exclusion principle which relies on the strength of intra-specific competition being greater than inter-specific competition for species co-existence (Barabás et al., 2016). For example, even if locals are better adapted to their abiotic environment, the competition for resources between individuals may be so high that it significantly reduces their fitness by intraspecific competition. In this case, despite being less adapted to their new environment, migrants could outcompete locals if they exploit resources differently, so they actually have a higher fitness than locals.

Although this is theoretically possible, it assumes that locals are near carrying capacity for their particular resource exploitation behavior.

The ability of migrant hosts to outcompete locals may also depend on how environmental conditions differentially affect the competitive ability of locals versus migrants. For example, temperature can determine the outcome of intra versus interspecific competition (Ntiri et al., 2016) and heterogeneity of consumable resources, such as the quality, quantity, size and availability of food particles, could facilitate different exploitation strategies (Kolasa & Pickett, 1991). Previous studies have shown that different species, such as *Daphnia*, have a range of these consumption behaviours; body size in different *Daphnia* species affects the maximum size of particle that can be ingested during filter-feeding (Burns, 1968). Correspondingly, aspects of the biotic environment may also influence the outcome of intra versus interspecific competition. For example, there may be predator-mediated competition of their prey (Wilson, 1989) or parasite-mediated competition of their hosts (Orlansky & Ben-Ami, 2023).

In particular, and as already introduced above, locally adapted parasites, to which local hosts are less resistant than migrants, could determine the outcome of host inter-population intra-specific competition. As predicted by general theory (Gandon, 2002), the tendency for parasites to have larger effective population sizes and shorter generation times than hosts means that they are usually able to infect hosts better if they are from their native environment (Greischar & Koskella, 2007; Hoeksema & Forde, 2008). Since hosts are expected to be locally adapted to their abiotic environment, but not to their corresponding parasites, the relative influence of either form of local adaptation on the outcome of host intra-specific competition with migrants is unclear.

It has been suggested that partitioning the relative effects of intraspecific competition and parasitism on host-parasite populations may be too difficult in the wild (Hochberg, 1991). One solution would be to expose hosts originating from different populations to a shared ancestral parasite (similar to experimental coevolution (Brockhurst & Koskella, 2013). Despite not having a long history of coevolution with different host populations, which makes them less likely to be more infectious of local hosts, it would allow us to measure the extent of parasite-mediated intra-specific competition between local and migrant hosts in a natural setting. We might expect

that the effect of infection on a host's fitness (in terms of reproduction) or competitive ability to be resource dependent, such that locally adapted hosts that are better at exploiting a shared resource also have a better condition than immigrants, and therefore a higher fitness.

In this study, I investigated the ability of locally adapted hosts to withstand competition from migrants. I performed a series of reciprocal transplants across 12 outdoor mesocosms to measure the ability of local and migrant adult *Daphnia* to reproduce in the presence or absence of a sterilizing microparasite and evaluated the following hypotheses (i) intra-specific host competition is driven by some sort of resource limitation, (ii) there is local adaptation of home (resident) versus away (foreign) genotypes to abiotic factors, (iii) immigrants suffer from competition with resident hosts and (iv) this cost is exacerbated for hosts artificially exposed to a shared ancestral parasite.

## 5.3 Methods

### 5.3.1 Methods (summary)

To measure how much wild *Daphnia* host genotypes are robust to competition from other non-local genotypes, we performed a series of reciprocal transplant experiments between nine pairs, six of which were not fully independent, of outdoor pond populations of the invertebrate model host (*Daphnia magna*) and its sterilising microparasite (*Pasteuria ramosa)* (Fig. 5.1A and B). We compared how host origin (home, away and mixed) interacted with exposure to ancestral parasite (Fig. 5.1C-F). By assuming that the absence of any cost of competition experienced by either the resident or immigrant host is consistent with no difference between the average host fitness (adult survival and fecundity) from both the home and away groups versus the mixed group, we were able to demonstrate a fitness cost when this value was a non-zero sum.

### 5.3.2 Study species

The experiment focussed on the freshwater micro-crustacean host, *Daphnia magna*, and its sterilising bacterial parasite, *Pasteuria ramosa*. *D. magna* and *P. ramosa* occur together naturally in lakes and ponds throughout Europe (Ebert, 2005). *P. ramosa* infects *Daphnia* by attaching itself to the gut, penetrating the gut wall and

then reproducing once inside the host (Auld et al., 2012; Auld, Hall, et al., 2014; Duneau et al., 2011). It is a highly virulent parasite which severely limits *Daphnia* reproduction and eventually kills the host. *P. ramosa* also causes infected *Daphnia* to turn red from bacterial growth in the haemolymph and to grow significantly larger through gigantism. Upon death, the infected cadaver releases millions of spores which are released into the environment for onward transmission (Ebert et al. 1996).

### 5.3.3 Experimental design

The experiment took place in 12 semi-natural outdoor pond populations (referred to as mesocosms) that had been established in April 2015 as part of a previous long-term research project. Since being established with an identical mix of host genotypes and parasite transmission stages (Auld & Brand, 2017b), differences in the pond environments have driven rapid co-evolution and the populations have subsequently diverged for both host and parasite characteristics (Paplauskas et al., 2021).

They were each allocated a unique identifier and randomly paired to another mesocosm (Fig. 5.1B). Generally, pairs were not made using adjacent mesocosms to avoid the comparison of mesocosms with similar environments (Fig. S5.1). Despite a large number of mesocosms (20 total), there was only a sufficiently large number of healthy (uninfected) *Daphnia* adults to establish three fully independent mesocosm pairs, so an additional three pairs of mesocosms were created by using some of the same ponds within the other three pairs. Therefore, there were nine pond pairs in total, but six of these were not entirely independent.

In each mesocosm, there was a total of six treatment combinations made up of three sources of *Daphnia* (pond of origin), including the local pond of origin (home), the neighbouring pond (away) and mixed (50:50 home and away), and two parasite treatments, including a control (P-) and parasite-exposed group (P+). For example, within pond pair one, there were two ponds; the local pond environment one (E1, Fig. 5.1C and D) was paired to the neighbouring pond environment (E2, Fig. 5.1E and F). In pond environment one, there were three animal sources, including home (E1 *Daphnia* only), away (E2 *Daphnia* only) and mixed (E1, E2), each crossed with a control and parasite exposed treatment. In comparison, in pond environment two (E2), there were the same three animals sources, but labelled accordingly, including

(E2 *Daphnia* only), away (E1 *Daphnia* only) mixed (E1, E2), each crossed with a control and parasite exposed treatment.

Each treatment combination was set-up using a stainless-steel coffee filter basket attached to a polystyrene tile (Fig. 5.2). Following a preliminary test of their wind-resistance, in which tiles were overturned by the strong winds, we weighted each tile down using weight discs attached to large coach bolts. Each control and parasite-exposed set of field cages were kept separate to avoid cross-infection, but the position of each field cage within the floating platform was randomised to avoid any bias. A total of eight healthy adults were added to each field cage at the beginning of the reciprocal transplant experiment.

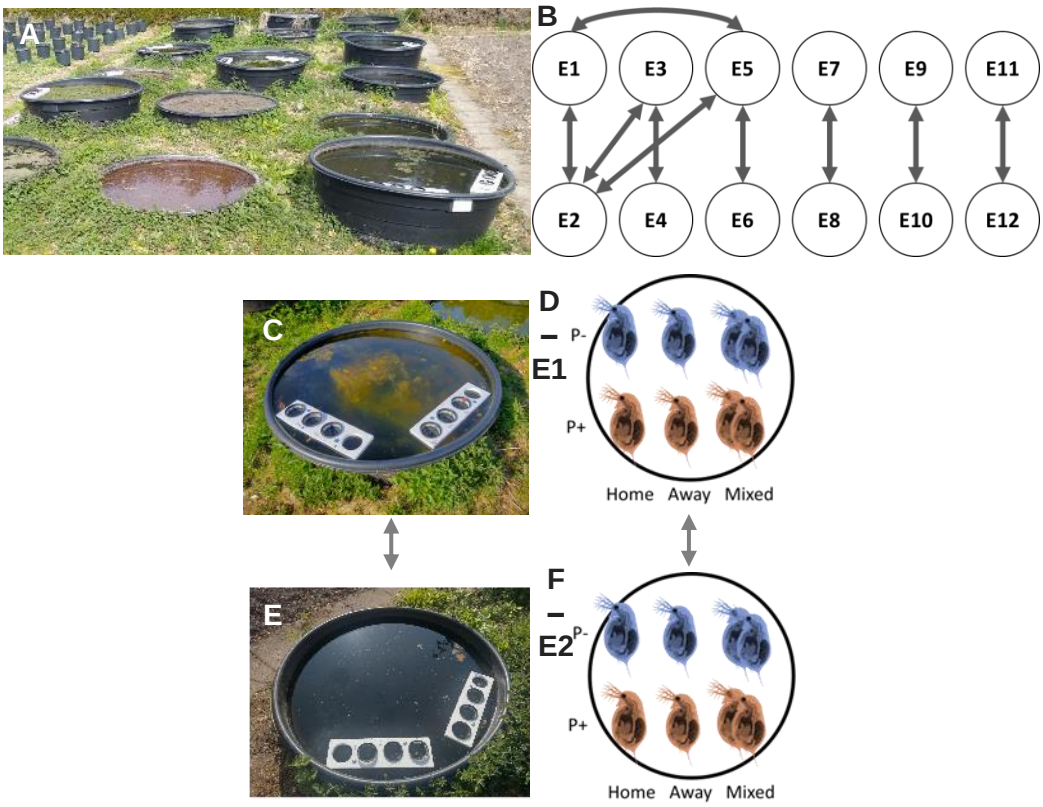



Figure 5.1. Summary of experimental design. Shown are A) photo of field setting and, B) reciprocal design, where M+No. refers to each mesocosm (M, note that there are some non-independent pairs) and C-F) photos of two experimental mesocosms within a single experiment transplant (C and E) and the treatment combinations within each mesocosm (D and F). There are three *Daphnia* origins (home, away, mixed) crossed with either a control (P-) or parasite-exposed (P+) treatment. The blue and red *Daphnia* refer to the control and parasite-exposed

treatment respectively. A total of eight healthy adult females were used to establish each treatment combination at the beginning of the reciprocal transplant experiment. Note that there are four holes in each float, but only three were used (fourth was a back-up if the tile broke).

Figure 5.2. Field cages (coffee filter baskets) used to establish each treatment combination for the transplant experiment. Inspired by (O'Connor et al., 2021).

### 5.3.4 Reciprocal transplant experiment

*Daphnia* were collected from each population by passing a 0.048 $m^2$ pond net across the diameter of each mesocosm (1.51 m) several times and transferred to a plastic tray. Eight uninfected *Daphnia* adults (no observable infection) were transplanted to each field cage from the appropriate origin (eight from home, eight from away or 4 from home and away in mixed treatment). 36 field cages were set up for three mesocosms pairs each day from the 24th to 26th April 2023, in a staggered approach. Although there were 20 mesocosms in total, there were only enough *Daphnia* to establish reciprocal transplants from 12 ponds, which involved the re-implementation of certain mesocosms with each pair (see above).

Frozen parasite transmission stages, which had been produced by propagation of 21 unique *D. magna* clones exposed to sediment from their natural pond (Auld & Brand, 2017b)), sampled as part of a previous experiment from Kaimes farm in at Leitholm (Scottish Borders, UK, geographic coordinates: 2° 20′ 43″ W, 55° 42′ 15″ N) (Auld, Wilson, et al., 2014), were used to apply a heavy dose of parasites (approximately 1 x $10^8$ *Pasteuria* spores) to each field cage that was part of the parasite treatment. This was the same ancestral parasite used to establish the

mesocosm populations in 2015 (Auld & Brand, 2017b). During the initial parasite treatment, the surface temperature of each mesocosm was measured for one minute using a thermometer inserted into a polystyrene buoyancy aid which held in the centre of the pond, that was otherwise unreachable (Fig. 5.3).

Figure 5.3. Thermometer inserted into a polystyrene buoyancy aid.

Subsequent temperature measurements were taken twice a week for three weeks, along with weekly measurements on host demography, including the number of healthy adults, the number of adults carrying ephippia (resting stages), the number of infected adults and the number of offspring. To count the number of animals in each field cage, a laboratory squeeze bottle used to spray pond water from the back of the animals into a translucent plastic tray. The animals were counted visually and then the tray was washed with the pond water and the animals were returned to their corresponding field cage. The order in which animals were counted for each field cage was randomised within and between mesocosms to avoid any bias. Collection of this demographic data each week was staggered across a three-day period to reflect how the experiment was established (see above).

### 5.3.5 Preliminary analysis of host fitness

To determine the ability of resident hosts to withstand competition from migrants, we compared estimates of adult fecundity across treatment groups by calculating the mean change in the number of offspring per adult from the previous week. We did not use the number of offspring per adult in the current week because this reflected the accumulation of offspring over time, which meant that it was not an accurate

reflection of adult fecundity each week (see Fig. S5.3). However, this information was included as Fig. S5.3 to investigate the carrying capacity of the field cages.

One potential limitation of this approach is that the number of adults varied by week which could have been the major driving force behind the observed differences in offspring production. This variation could have been caused by a combination of different factors, such as survival, offspring maturation and mortality (Fig. 5.4). For example, adult mortality prior to the weekly data collection might artificially inflate adult fecundity. However, it is safe to assume that *Daphnia* are more likely to survive and have chance to give birth during any previous week, than to die at the beginning of the week. In week one, this could be explained as part of a stress-induced response to transplanting. Therefore, it might be preferable to calculate the mean change in the number of offspring per adult from the previous week, rather than the current week.

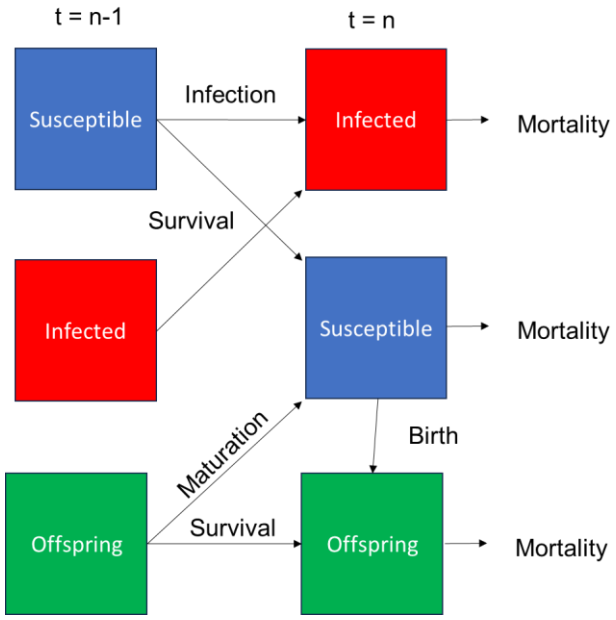



Figure 5.4. Sources of variation in the mean number of adults and offspring between consecutive weeks. For any given week (t = n), the number of susceptible adults (blue squares), infected adults (red squares) and the number of offspring (green squares) are determined by a combination of positive versus negative changes, such as infection, survival, maturation of offspring into adults and birth versus mortality (arrows). Infection tends to completely sterilise the host, with little chance of returning to a susceptible state (Ebert, 2005), and eventually leads to host mortality.

Another considerable advantage of using the mean change in the number of offspring per adult from the previous week, rather than the current week, to calculate a cost of mixed competition, is that there will be immature offspring that look as if they could have given birth during specific points in the experiment. However, if we assume that the generation of *D. magna* is similar in the field to the lab, and occurs every 8-14 days (Ebert, 2005), then the number of adults from counted the previous week will exclude these immature offspring (Fig. 5.5).

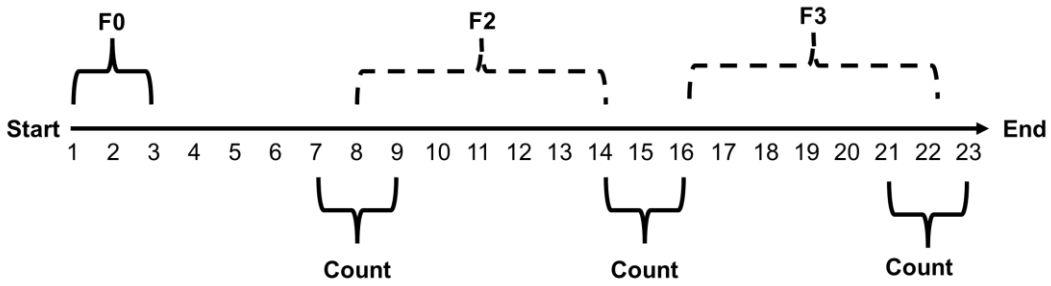



Figure 5.5. The benefit of using the number of adults observed from the previous week to calculate mean change in the number of offspring per adult. The experiment started on day one and ended on day 23 (arrow). Over the first three days of the experiment, the first generation of adults (F0) were added to the field cages in a staggered experimental design. Since the generation time of *Daphnia* is approximately 8-14 days in the laboratory (Ebert, 2005), we predicted that the earliest the F2 generation should have appeared was at day eight. Therefore, despite predicting the presence of immature offspring that would appear as if they could have given birth by this point, the actual number of adults during the counting period would be inaccurate if these immature offspring were included.

Finally, since we did not count any infected *Daphnia* in the parasite-exposed treatment (which is probably due to a high level of density-dependent competition) and it takes a couple of weeks for host sterilisation to set in (Ebert, 2005), we did not need to account for any difference in the number of offspring per infected versus susceptible adult from the previous week.

### 5.3.6 Statistical analysis

To determine whether there was any evidence for intra-specific host competition, that could possibly be driven by some sort of resource competition over food limitation (sensu (Lang, 2013)), and may explain the expected variation in adult

survival and changes in the offspring between treatments, we developed a series of linear mixed-effects models using the *lmer* function from the *lme4* package version in R to compare the change in the number of offspring per adult (from the current and previous week) with the total number of adults for both the control and parasite-exposed treatment each week. The *Daphnia* pond of origin and the current pond environment were treated as random effects.

To determine if there was any local adaptation by the host to the abiotic environment and how this was influenced by parasitism, we compared linear mixed effects models with the same random effects structure as described above, but included an additional interaction term for host origin, to examine the mean change in the number of offspring per adult from the previous week across both different host origins (home, away etc.) and the control versus parasite treatment for each week. As mentioned previously, we did not test either host or parasite local adaptation to one another.

To measure the cost of mixed host competition and how this was influenced by parasitism, we compared the mean change in the number of offspring per adult from the previous week across a combination of the home and away origins with the mixed origin for both the control and parasite-exposed treatment. This used the same model structure as for the test of host local adaptation to environmental conditions.

All analysis was performed in R version 4.4.1.

## 5.4 Results

### 5.4.1 Habitat and species diversity

The replicate populations represented fairly unique environments (Fig 5.6). For example, the water surface was covered in green floating plants in some ponds (Fig. 5.6A) and clear in others (Fig. 5.6B). There was variation in pond colour, from reddish brown (Fig. 5.6C) to bright green (Fig. 5.6D), that corresponded with the relevant field cages in these ponds (Fig. 5.6E and F) and was most likely caused by variation in host density, leading to a red colour in ponds with a high population density and low oxygen concentration (Fig. 5.6G).

There was also a large amount of variation observed in the diversity of pond wildlife (Fig. 5.7). In some cases, there was a large density of *Chaoborus* (Fig. 5.7A) that corresponded with very few *Daphnia*, suggesting that this may have been a key factor in limiting host population size. In other ponds there was a mix of species in variable abundances, including a large amount of *Planorbidae* snails in most ponds (Fig. 5.7B), multiple species of worm (Fig. 5.7E, F) and beetle (Fig. 5.7C, D) and a low abundance of other species (Fig. 5.7H, I, J). In addition, there were some species on the water surface, such as pond skaters (Fig. 5.7G), bees (Fig. 5.7L) and even ducks (Fig. XK).

Figure 5.6. Observed differences in pond environments. A) A pond covered with duckweed algae versus B) a pond with completely clear water. C) A pond with reddish brown water versus D) a pond with bright green water. E) Field cages corresponding to D versus F) field cages corresponding to C. G) The origin of the reddish brown coloured ponds; a sample of pond water with a high density of *Daphnia magna*.

3692

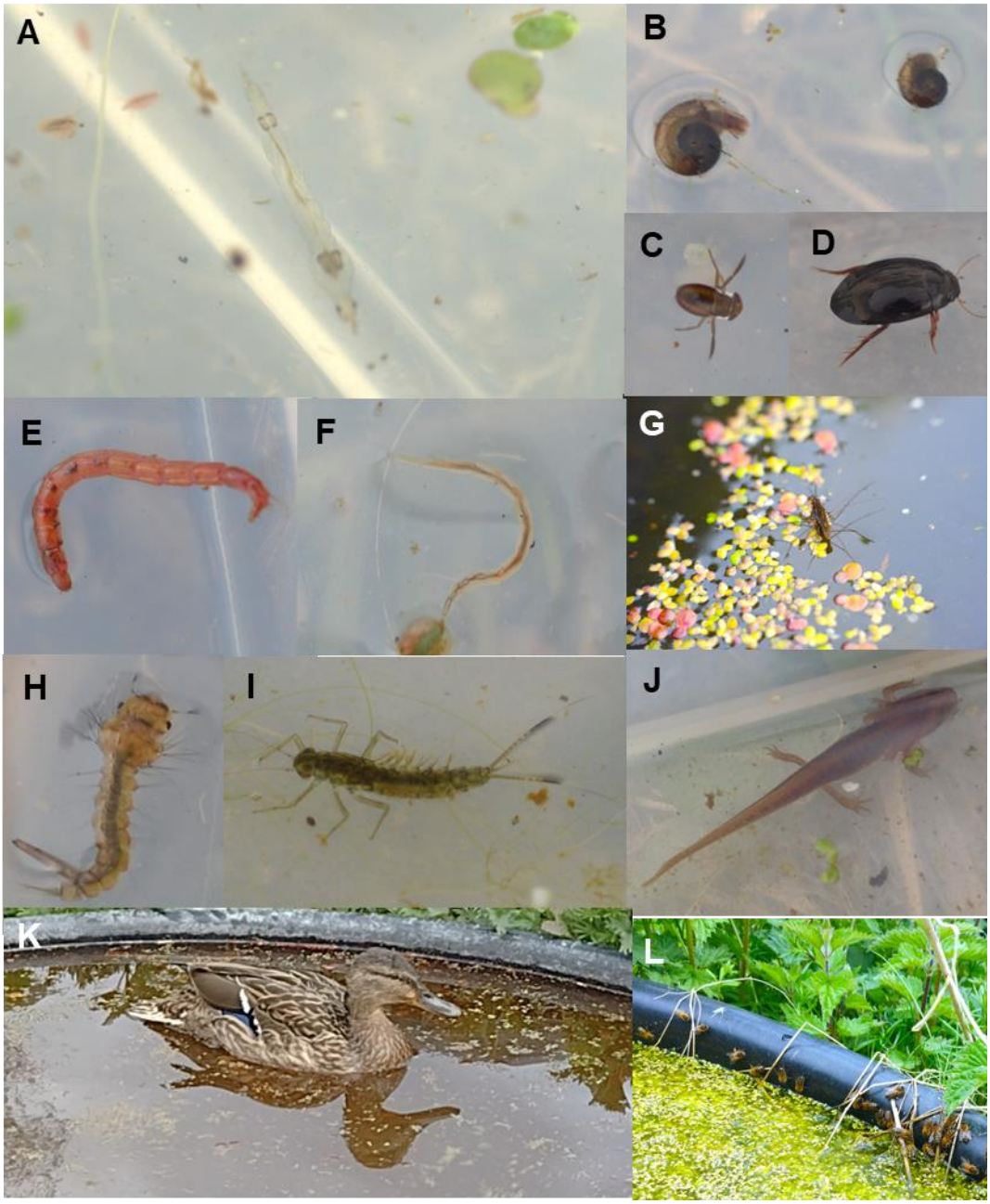



Figure 5.7. The diversity of pond wildlife in the outdoor mesocosms. A) *Daphnia magna* (top left) and phantom midge larva (*Chaoborus*) which is a common predator of *Daphnia*, B) ramshorn snails (*Planorbidae*), C) water boatman

(*Corixidae*), D) great diving beetle (*Dytiscus marginalis*), E) bloodworm (*Glycera*), F) nematode species, G) common pond skater (*Gerris lacustris*), H) mosquito larva (*Culicidae*), I) damselfly larva (*Zygoptera*), J) newt (*Pleurodelinae*), K) mallard (*Anas platyrhynchos*) L) bees (*Anthophila*).

### 5.4.2 The presence of density-dependent competition between adult hosts

To look for evidence of any competition between hosts, regardless of their origin treatment (home, away or mixed), we investigated the relationship between the change in the number of offspring per adult from the current week and the total number (alt. density) of adults from the current week (Fig. 5.8, row one) versus the change in the number of offspring per adult at a time-lag of one week with the total number of adults at a time-lag of one week (Fig. 5.8, row two). We found a consistently negative relationship between the independent and dependent variables in the across weeks, regardless of parasite exposure or not, but this was only significant for some models (Supplementary Table 5.1). This suggests that space or resources are limiting in the field cages, which drives density-dependent adult host fecundity, and in particular, any significant differences observed in the mean change in the number of offspring per adult from the current week, or at a time-lag of one week, among treatments will be due to how competition varies across treatments, as opposed to other extraneous factors.

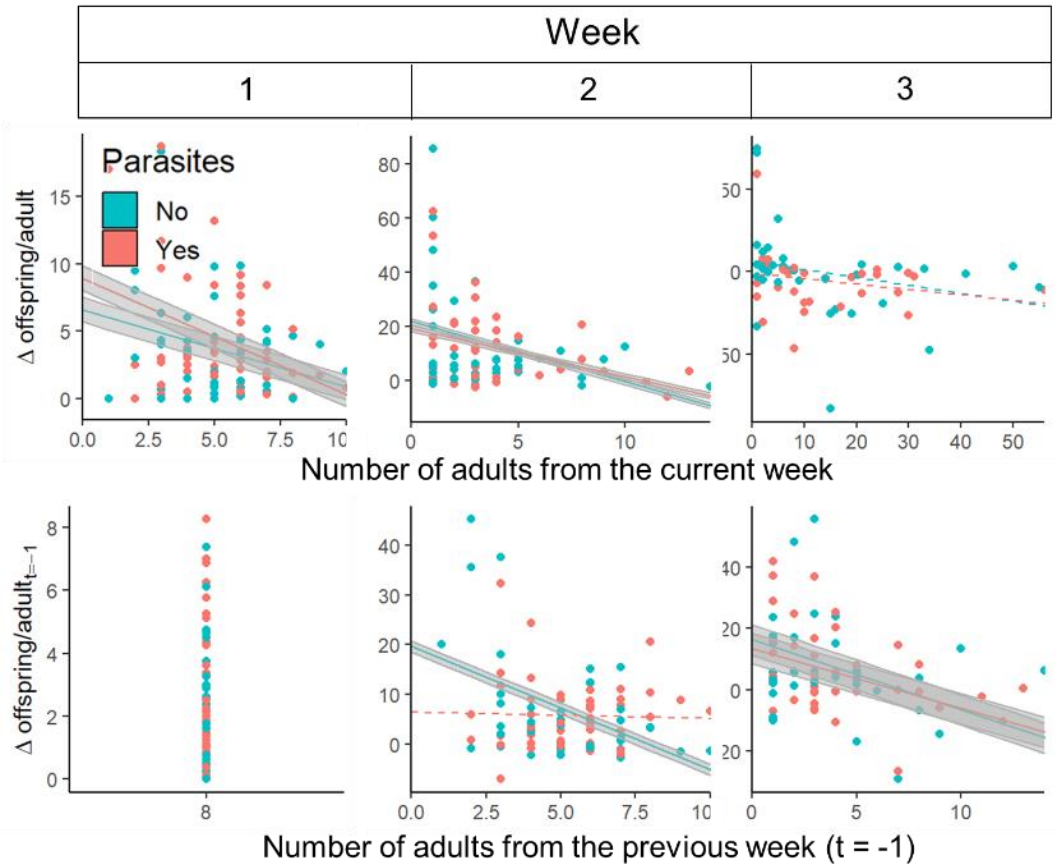



Figure 5.8. The presence of density-dependent competition between adult hosts. The different colours correspond to the control (blue) and parasite-exposed treatment (red). The linear relationship from a mixed effects model is shown for each subset of the data, where the dashed versus solid lines correspond to non-significant versus significant effects. 95% confidence intervals are shown for the significant effects (grey bands). Row one - the consistently negative relationship between change in the number of offspring per adult from the current and the total number of adults from the current week across the entire duration of the experiment (three weeks in total). Row two - the consistently negative relationship between change in the number of offspring per adult at a time-lag of one week and the total number of adults at a time-lag of one week across the entire duration of the experiment (three weeks in total).

3710

### 5.4.3 Temporal variation in total adults and a cost of intra-specific host competition he presence of density-dependent competition between adult hosts

Although the experiment started with the same total number of adults (eight) in each field cage, there was considerable variation in the mean number of adults each week across treatment combinations (Fig. 5.9). There was a steady decline in the mean number of adults across most treatment combinations in weeks one and two (Fig. 5.9A and B), followed by a relatively large increase in week 3 (Fig. 5.9C). These relative differences in the mean number of adults between both weeks one and two versus week three were highly significant (Tukey adjusted estimated marginal means = -4.36 and -5.94; $p = 0.0021$ and $< .0001$ respectively) and probably stemmed from a combination of reduced adult survival and offspring maturation (Fig. 5.4). In addition, there were no significant pairwise differences between either home and away groups from the control and parasite-exposed treatment (Tukey adjusted comparison of all pairwise differences using estimated marginal means; $p > 0.05$, Supplementary table S5.2), or the mean number of adults across both of the home and away groups versus the mixed group (Tukey adjusted comparison of all pairwise differences using estimated marginal means; $p > 0.05$, Supplementary table S5.3). Therefore, this suggests that adult survival is not locally adapted to the abiotic environment and there is no cost associated with host mortality as a result of intra-specific competition.

On the other hand, the mean change in the number of offspring per adult at a time-lag of one week, which is indicate of adult fecundity (but see Fig. 5.4), showed some interesting results. First, despite all pairwise differences being statistically equivalent (Tukey adjusted comparison of all pairwise differences using estimated marginal means; $p > 0.05$, Supplementary table S5.4), the mean change in the number of offspring per adults at a time-lag of one week was consistently higher in the parasite-exposed treatment versus the control in week one (Fig. 5.9D). Second, there was an absence of host local adaptation to the abiotic environment in terms of mean change in offspring per adult at a time-lag of one week (i.e. adult fecundity, Supplementary table S5.4), similar to adult survival (see above) and there was no cost of intra-specific competition in weeks one and two (Fig. 5.9D and E, Supplementary table S5.5). However, this difference in the mean change in the number of offspring per adult at a time-lag of one week for both the home and away groups versus the mixed group was significant for the parasite-exposed treatment in week 3 (Tukey adjusted

estimated marginal mean = 9.77; p = 0.035, Fig. 5.9F). These findings suggest 1) there is short-term parasite-induced fecundity compensation and 2) and there is a cost of intra-specific competition that is exacerbated by parasite-exposure.

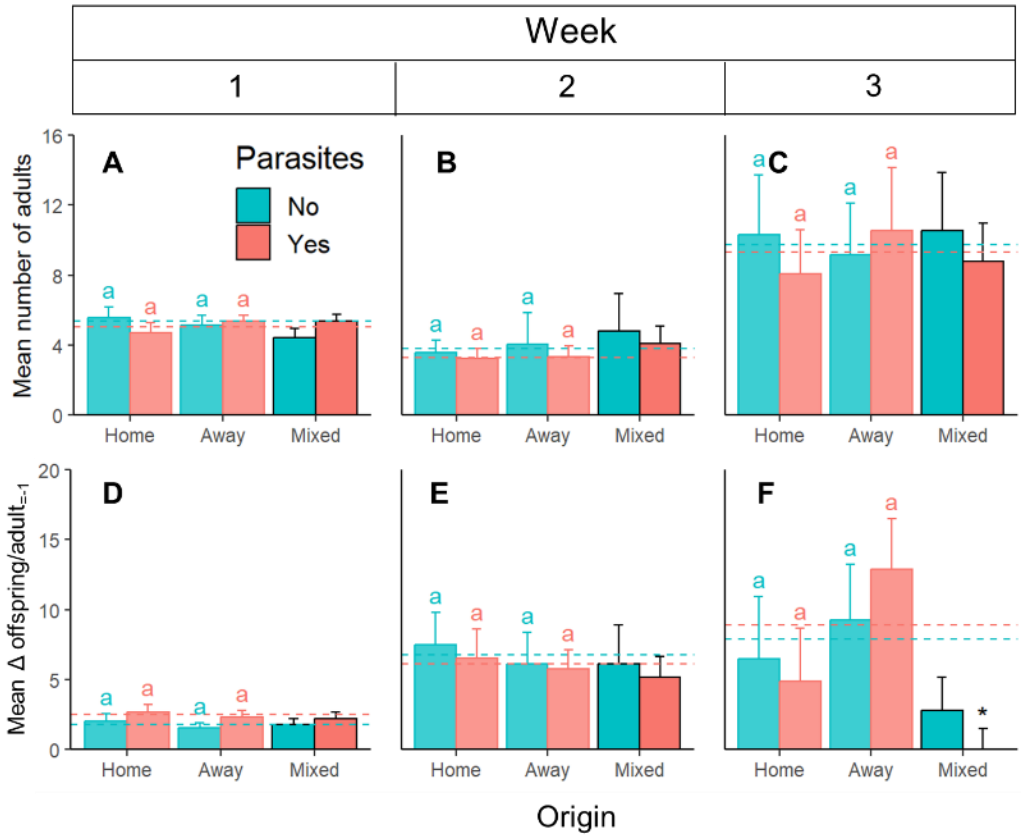



Figure 5.9. Evidence of short-term parasite-induced fecundity compensation, an absence of host local adaptation to the abiotic environment and the ability of non-locally adapted hosts to outcompete resident genotypes. A-C) Variation in the mean number of adults each week, driven by both adult survival and offspring maturation. D-F) Variation in the mean change in the number of offspring per adult at a time-lag of one week. Since the change in offspring is not based on the number of adults from the current week (i.e. directly above the plot), the change in the offspring corresponds to the plot diagonally above it on the left-hand side (which means that the change in offspring at week one is based on the eight adults from the initial set-up of the experiment). The source of the *Daphnia* in each basket (Origin) is shown on the x-axis and the colour of the bars indicates the control (blue) and parasite exposed (red) treatments. Home and away groups are statistically equivalent (all pairwise differences; p > 0.05), so they share the letter 'a'. The mean of both the home and away groups (dashed lines) versus the mixed group was significantly different for the change in the number of offspring per adult

at a time-lag of one week for the parasite-exposed treatment in week three (single asterisk; $p < 0.05$). See Fig. S5.3 for additional analysis of the number of offspring.

To determine the precise cost of this mixed competition in week three, and who is the most likely winner in this scenario, we compared the observed mean change in the number of offspring per adult at a time-lag of one week (i.e. adult fecundity) for the separate home and away groups compared to the expected proportion of the mixed group that was composed of animals from the home and away origins (50:50). This showed that overall cost of this mixed competition in week three for the parasite-exposed treatment was a 9.19 reduction in the mean change in the number of offspring per adult at a time-lag of one week between both the home and away groups compared to only the mixed group (Supplementary Table S5.6). Assuming that this overall reduction was equally distributed across animals of both home and away origin within the parasite-exposed mixed treatment, relative to what was observed in the corresponding parasite-exposed unmixed home and away groups (Supplementary Table S5.6; example C), then the expected mean change in the number of offspring per adult at a time-lag of one week is net positive for migrants (0.72) and net negative for locals (-1.04), which means that migrants are expected to outcompete locals.

We also found that the variation in the total number of adults in each treatment combination, caused by adult mortality and offspring maturation (Fig. 5.4), was not responsible for the differences observed between the mean change in number of offspring per adult at a time-lag of one week (Fig. S5.3).

## 5.5 Discussion

To measure whether hosts were adapted to their local abiotic environment and how this might interact with parasitism to influence inter-population competition between local and migrant hosts, we performed a series of reciprocal transplant experiments in replicate populations of the naturally coevolving model *Daphnia* host – parasite system.

First, in support of our original hypothesis, we found that there was indeed evidence for some sort of resource competition, as shown by the density-dependent change in the number of offspring per adult at a time-lag of one week (i.e. adult fecundity), which could potentially drive any variation subsequently observed in host competitive

ability across the experimental treatment combinations. We anticipate this competition could either be some form of scramble (also termed 'exploitation') competition over a shared resource with limited availability, such as food, or it may be some kind of interference competition, where *Daphnia* compete directly with one another (Lang, 2013).

It is more likely that this is some form of resource competition, as there is a large amount of heterogeneity in the *Daphnia* food base (algae species), in terms of size, quality, quantity and availability (Kolasa & Pickett, 1991). Analogous to examples of interspecific resource competition between *Daphnia* with different body sizes (Kreutzer & Lampert, 1999), which affects the rate at which they filter food out of the environment (Burns, 1968; Porter et al., 1983), there may be similar patterns of competition between hosts from different origins. In support of this idea, there is genetic variation in *Daphnia* feeding behaviours within populations, but these are related to only infected hosts (Pfenning-Butterworth et al., 2023). Ideally, to demonstrate the extent to which the competition observed in this study was driven by variation in such feeding behaviours, we would have measured whether these traits had diverged in the replicate populations.

In comparison to the support found for resource competition between hosts, and in contrary to all of the other three remaining hypotheses, we found no evidence for host local adaptation to the environment and that migrant hosts were able to compete with locals in terms of offspring birth rate and survival in the first two weeks of the experiment, such that they eventually incurred a cost of competition that was shared by locals, and exacerbated by exposure to an ancestral parasite. This is potentially indicative of the ability of non-locally adapted hosts outcompete residents under a high level of resource competition together with a general stressor, such as parasitism.

The lack of host local adaptation to the abiotic environment was consistent across all three weeks of the experiment. This was surprising, given the absence of any gene flow between separate mesocosms, which is a major factor driving patterns of local adaptation across most wild systems (Kawecki & Ebert, 2004). *Daphnia* populations are usually distinctly separated from each other as they are limited to standing water bodies, so there are moderate amounts of gene flow (Ebert, 2022). Therefore, in combination with their short generation times, populations have the

ability to evolve rapidly in response to environmental conditions. Indeed, the variation in pond diversity we observed, along with the differences observed in environmental conditions between ponds from a previous study (Paplauskas et al., 2021), would suggest there is significant potential for host local adaptation to the abiotic environment. However, this could have not been observed due to being masked by genetic drift (Gandon & Nuismer, 2009), as a result of variation in population density across replicate ponds.

It was also surprising that we found migrants had the ability to outcompete residents. This was observed in the third week of the experiment, where both the control and parasite exposed treatments suffered a cost of mixed competition. This cost was not apparent in the first two weeks of the experiment, where the overall population density was lower in each treatment combination, so the strength of intra-specific resource competition would have been lower too. However, this cost of mixed competition was only significant for the parasite exposed treatment of the third week. This suggests that the parasites treatment, which was composed of a mixture of shared ancestral parasites used to establish the replicate mesocosm environment as part of a much earlier experiment (Auld & Brand, 2017a, 2017b), and was therefore not locally adapted to the resident hosts in each mesocosm, exacerbated the high level of resource competition by acting as a general stressor.

One possible explanation for why parasites that were acting as a general stressor, in combination with resource competition, may have impacted on resident host fitness more than immigrants, is because there may have been a relative difference in the strength of intra versus inter-population intra-specific competition (see introduction). However, this would have been more relevant if there had been local adaption to the abiotic environment. Alternatively, a more compelling reason is that there may have been a differential response to the environment in the exploitative behaviour of local and migrant hosts (see introduction). As described above, infected hosts can demonstrate ‘sickness behaviours’ which affect their resource consumption (Pfenning-Butterworth et al., 2023). However, one study of parasite-mediated interspecific competition driven by these sickness behaviors in *Daphnia* species actually found that they promoted species coexistence (Orlansky & Ben-Ami, 2023). Assuming interspecific competition can be considered analogous to interpopulation (local versus migrant) intra-specific competition, we show a different

version of events whereby parasite-mediated competition enhances the cost of competition with migrant conspecifics.

We also found that there was parasite-induced fecundity compensation of hosts in week one. This result has been observed in a previous lab study of *Pasteuria*-infected *D. magna* (Vale & Little, 2012), but has not been confirmed in the wild. This previous study compared the number of offspring produced in the first clutch of infected hosts before sterilization to an unexposed control group, but did not measure the number of offspring in subsequent clutches due to natural variation between the timing of infection and the subsequent onset of sterilisation (Vale & Little, 2012). Therefore, although *D. magna* are completely sterile between approximately 5-15 days after an initial infection (Ebert, 2005), it is not clear how long fecundity compensation lasts.

Surprisingly, we counted only a very small number of infected *Daphnia* across all replicate field cages, despite exposing them to a very high dose of parasite transmission spores. Two possible reasons for why this happened are that 1) the effect of resource competition combined with parasite infection may have lead to high mortality of infected hosts and 2) the hosts were able to resist infection in the parasite treatment, as they had evolved resistance to this ancestral parasite. The latter is supported by a previous study, which found that most host populations evolved resistance to this same parasite mix (Paplauskas et al., 2021), but it is unclear how long this resistance would have been maintained as coevolution in *Daphnia* species is generally driven by negative frequency-dependent selection (Luijckx et al., 2013). This means that hosts that are uncommon are resistant to parasites until they begin to increase in frequency and they are subject to parasites becoming more infectious of them, at the cost of other hosts, and they quickly become susceptible once again (Brockhurst & Koskella, 2013).

The main reason for treating hosts with a shared, ancestral parasite from a previous experiment (Auld & Brand, 2017a, 2017b), rather than relying on natural infection, which would have been very low at the time of year the experiment was conducted, was to use parasites as a general stressor. This was because there was no indication of whether there would have been a significant amount of competition between *Daphnia* (over food) within experimental replicates (i.e. field cages), especially considering the fact that there were only eight adults in each field cage used to

establish the experiment. Therefore, we had intended to use parasites as a general stressor to reveal the effect of competition of host local adaptation to the abiotic environment, but we found no evidence for such local adaptation.

Since our previous study of host-parasite coevolution in the replicate pond populations focused on the extent to which variation in pond environments could explain variation in the direction of host-parasite coevolutionary trajectories (Paplauskas et al., 2021), rather than testing whether hosts and parasites from different ponds were locally adapted to one another, and we were unable to measure the infection rate across *Daphnia* populations from different origins in this experiment, there might be an opportunity for further research to measure how variation among the environment of replicate populations drives patterns of both host and parasite local adaptation.

## 5.6 References


Aitken, S. N., & Whitlock, M. C. (2013). Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics* (Vol. 44). https://doi.org/10.1146/annurev-ecolsys-110512-135747

Auld, S. K. J. R., & Brand, J. (2017a). Environmental variation causes different (co) evolutionary routes to the same adaptive destination across parasite populations. *Evolution Letters*, *1*(5), 245–254. https://doi.org/10.1002/evl3.27

Auld, S. K. J. R., & Brand, J. (2017b). Simulated climate change, epidemic size, and host evolution across host–parasite populations. *Global Change Biology*, *23*(12), 5045–5053. https://doi.org/10.1111/gcb.13769

Auld, S. K. J. R., Edel, K. H., & Little, T. J. (2012). The cellular immune response of daphnia magna under host-parasite genetic variation and variation in initial dose. *Evolution*, *66*(10). https://doi.org/10.1111/j.1558-5646.2012.01671.x

Auld, S. K. J. R., Hall, S. R., Ochs, J. H., Sebastian, M., & Duffy, M. A. (2014). Predators and patterns of within-host growth can mediate both among-host competition and evolution of transmission potential of parasites. *American Naturalist*, *184*(SUPPL. 1), S77–S90. https://doi.org/10.1086/676927

Auld, S. K. J. R., Wilson, P. J., & Little, T. J. (2014). Rapid change in parasite infection traits over the course of an epidemic in a wild host-parasite population. *Oikos*, *123*(2), 232–238. https://doi.org/10.1111/j.1600-0706.2013.00720.x

Barabás, G., J. Michalska-Smith, M., & Allesina, S. (2016). The effect of intra- and interspecific competition on coexistence in multispecies communities. *The American Naturalist*, *188*(1), E1–E12. https://doi.org/10.1086/686901

Blanquart, F., Gandon, S., & Nuismer, S. L. (2012). The effects of migration and drift on local adaptation to a heterogeneous environment. *Journal of Evolutionary Biology*, *25*(7), 1351–1363. https://doi.org/10.1111/j.1420-9101.2012.02524.x

Blanquart, F., Kaltz, O., Nuismer, S. L., & Gandon, S. (2013). A practical guide to measuring local adaptation. *Ecology Letters*, *16*(9), 1195–1205. https://doi.org/10.1111/ele.12150

Bocedi, G., Atkins, K. E., Liao, J., Henry, R. C., Travis, J. M. J., & Hellmann, J. J. (2013). Effects of local adaptation and interspecific competition on species' responses to climate change. *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/nyas.12211

Brockhurst, M. A., & Koskella, B. (2013). Experimental coevolution of species interactions. *Trends in Ecology & Evolution*, *28*(6), 367–375. https://doi.org/10.1016/J.TREE.2013.02.009

Burns, C. W. (1968). The relationship between body size of filter-feeding *Cladocera* and the maximum size of particle ingested. *Limnology and Oceanography*, *13*(4), 675–678. https://doi.org/10.4319/lo.1968.13.4.0675

Duneau, D., Luijckx, P., Ben-Ami, F., Laforsch, C., & Ebert, D. (2011). Resolving the infection process reveals striking differences in the contribution of environment, genetics and phylogeny to host-parasite interactions. *BMC Biology*, *9*, 11. https://doi.org/10.1186/1741-7007-9-11

Ebert, D. (2005). *Ecology, Epidemiology, and Evolution of Parasitism in* Daphnia (S. Paliwoda, Ed.).

Ebert, D. (2022). *Daphnia* as a versatile model system in ecology and evolution. *EvoDevo*, *13*(1), 16. https://doi.org/10.1186/s13227-022-00199-0

Gandon, S. (2002). Local adaptation and the geometry of host-parasite coevolution. *Ecology Letters*, *5*(2), 246–256. https://doi.org/10.1046/j.1461-0248.2002.00305.x

Gandon, S., & Nuismer, S. L. (2009). Interactions between genetic drift, gene flow, and selection mosaics drive parasite local adaptation. *American Naturalist*, *173*(2), 212–224. https://doi.org/10.1086/593706

Greischar, M. A., & Koskella, B. (2007). A synthesis of experimental work on parasite local adaptation. *Ecology Letters*, *10*(5), 418–434. https://doi.org/10.1111/j.1461-0248.2007.01028.x

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. In *American Naturalist*. https://doi.org/10.1086/597611

Hochberg, M. E. (1991). Population dynamic consequences of the interplay between parasitism and intraspecific competition for host-parasite systems. *Oikos*, *61*(3), 297. https://doi.org/10.2307/3545237

Hoeksema, J. D., & Forde, S. E. (2008). A meta-analysis of factors affecting local adaptation between interacting species. *American Naturalist*, *171*(3), 275–290. https://doi.org/10.1086/527496

Holt, R. D., & Gomulkiewich, R. (1997). How does immigration influence local adaptation? A reexamination of a familiar paradigm. In *American Naturalist* (Vol. 149, Issue 3, pp. 563–572). https://doi.org/10.1086/286005

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*(12), 1225–1241. https://doi.org/10.1111/j.1461-0248.2004.00684.x

Kolasa, J., & Pickett, S. T. A. (Eds.). (1991). *Ecological heterogeneity* (Vol. 86). Springer New York. https://doi.org/10.1007/978-1-4612-3062-5

Kreutzer, C., & Lampert, W. (1999). Exploitative competition in differently sized *Daphnia* species: A mechanistic explanation. *Ecology*, *80*(7), 2348. https://doi.org/10.2307/176915

Lang, J. M. (2013). Species interactions and competition. *Nature Education*, *4*(4).

Lascoux, M., Glémin, S., & Savolainen, O. (2016). Local adaptation in plants. In *Encyclopedia of Life Sciences* (pp. 1–7). Wiley. https://doi.org/10.1002/9780470015902.a0025270

Luijckx, P., Fienberg, H., Duneau, D., & Ebert, D. (2013). A matching-allele model explains host resistance to parasites. *Current Biology*, *23*(12), 1085–1088. https://doi.org/10.1016/j.cub.2013.04.064

Meek, M. H., Beever, E. A., Barbosa, S., Fitzpatrick, S. W., Fletcher, N. K., Mittan-Moreau, C. S., Reid, B. N., Campbell-Staton, S. C., Green, N. F., & Hellmann, J. J. (2023). Understanding local adaptation to prepare populations for climate change. *BioScience*, *73*(1). https://doi.org/10.1093/biosci/biac101

Morgan, A. D., Gandon, S., & Buckling, A. (2005). The effect of migration on local adaptation in a coevolving host-parasite system. *Nature*, *437*(7056), 253–256. https://doi.org/10.1038/nature03913

Ntiri, E. S., Calatayud, P.-A., Van Den Berg, J., Schulthess, F., & Le Ru, B. P. (2016). Influence of temperature on intra- and interspecific resource utilization within a community of Lepidopteran maize stemborers. *PLOS ONE*, *11*(2), e0148735. https://doi.org/10.1371/journal.pone.0148735

O'Connor, M., Sadler, D. E., Brunner, F. S., Reynolds, A., White, N., Price, S., & Plaistow, S. J. (2021). Measuring *Daphnia* life history in the wild: The efficacy of individual field cages. *Ecology and Evolution*, *11*(23). https://doi.org/10.1002/ece3.8326

Orlansky, S., & Ben-Ami, F. (2023). The parasites of my rival are my friends. *Frontiers in Microbiology*, *14*. https://doi.org/10.3389/fmicb.2023.1135252

Paplauskas, S., Brand, J., & Auld, S. K. J. R. (2021). Ecology directs host–parasite coevolutionary trajectories across *Daphnia*–microparasite populations. *Nature Ecology & Evolution*, *5*(4), 480–486. https://doi.org/10.1038/s41559-021-01390-7

Peterson, M. L., Doak, D. F., & Morris, W. F. (2019). Incorporating local adaptation into forecasts of species' distribution and abundance under climate change. In *Global Change Biology* (Vol. 25, Issue 3). https://doi.org/10.1111/gcb.14562

Pfenning-Butterworth, A. C., Vetter, R. E., & Hite, J. L. (2023). Natural variation in host feeding behaviors impacts host disease and pathogen transmission potential. *Ecology and Evolution*, *13*(3). https://doi.org/10.1002/ece3.9865

Porter, K. G., Feig, Y. S., & Vetter, E. F. (1983). Morphology, flow regimes, and filtering rates of *Daphnia*, *Ceriodaphnia*, and *Bosmina* fed natural bacteria. *Oecologia*, *58*(2), 156–163. https://doi.org/10.1007/BF00399211

Price, S. J., Leung, W. T. M., Owen, C. J., Puschendorf, R., Sergeant, C., Cunningham, A. A., Balloux, F., Garner, T. W. J., & Nichols, R. A. (2019). Effects of historic and projected climate change on the range and impacts of an emerging wildlife disease. *Global Change Biology*. https://doi.org/10.1111/gcb.14651

Refardt, D., & Ebert, D. (2007). Inference of parasite local adaptation using two different fitness components. *Journal of Evolutionary Biology*, *20*(3), 921–929. https://doi.org/10.1111/j.1420-9101.2007.01307.x

Reger, J., Lind, M. I., Robinson, M. R., & Beckerman, A. P. (2018). Predation drives local adaptation of phenotypic plasticity. *Nature Ecology & Evolution*, *2*(1), 100–107. https://doi.org/10.1038/s41559-017-0373-6

Schulte, R. D., Makus, C., Hasert, B., Michiels, N. K., & Schulenburg, H. (2011). Host-parasite local adaptation after experimental coevolution of *caenorhabditis elegans* and its microparasite *bacillus thuringiensis*. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1719), 2832–2839. https://doi.org/10.1098/rspb.2011.0019

Sternberg, E. D., & Thomas, M. B. (2014). Local adaptation to temperature and the implications for vector-borne diseases. In *Trends in Parasitology* (Vol. 30, Issue 3). https://doi.org/10.1016/j.pt.2013.12.010

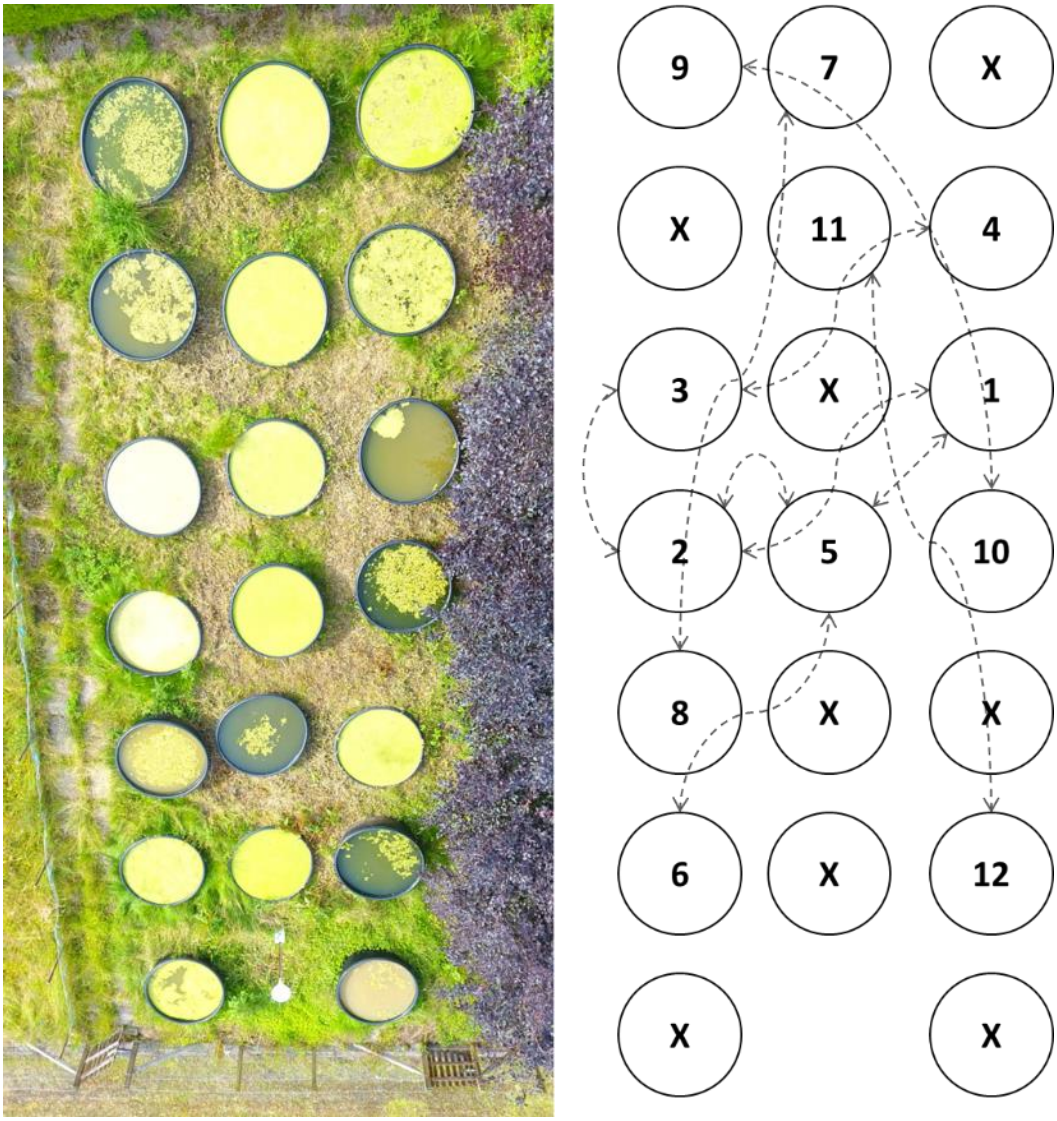
Vale, P. F., & Little, T. J. (2012). Fecundity compensation and tolerance to a sterilizing pathogen in *Daphnia*. *Journal of Evolutionary Biology*, *25*(9), 1888–1896. https://doi.org/10.1111/j.1420-9101.2012.02579.x

Wilson, W. H. (1989). Predation and the mediation of intraspecific competition in an infaunal community in the Bay of Fundy. *Journal of Experimental Marine Biology and Ecology*, *132*(3), 221–245. https://doi.org/10.1016/0022-0981(89)90129-9

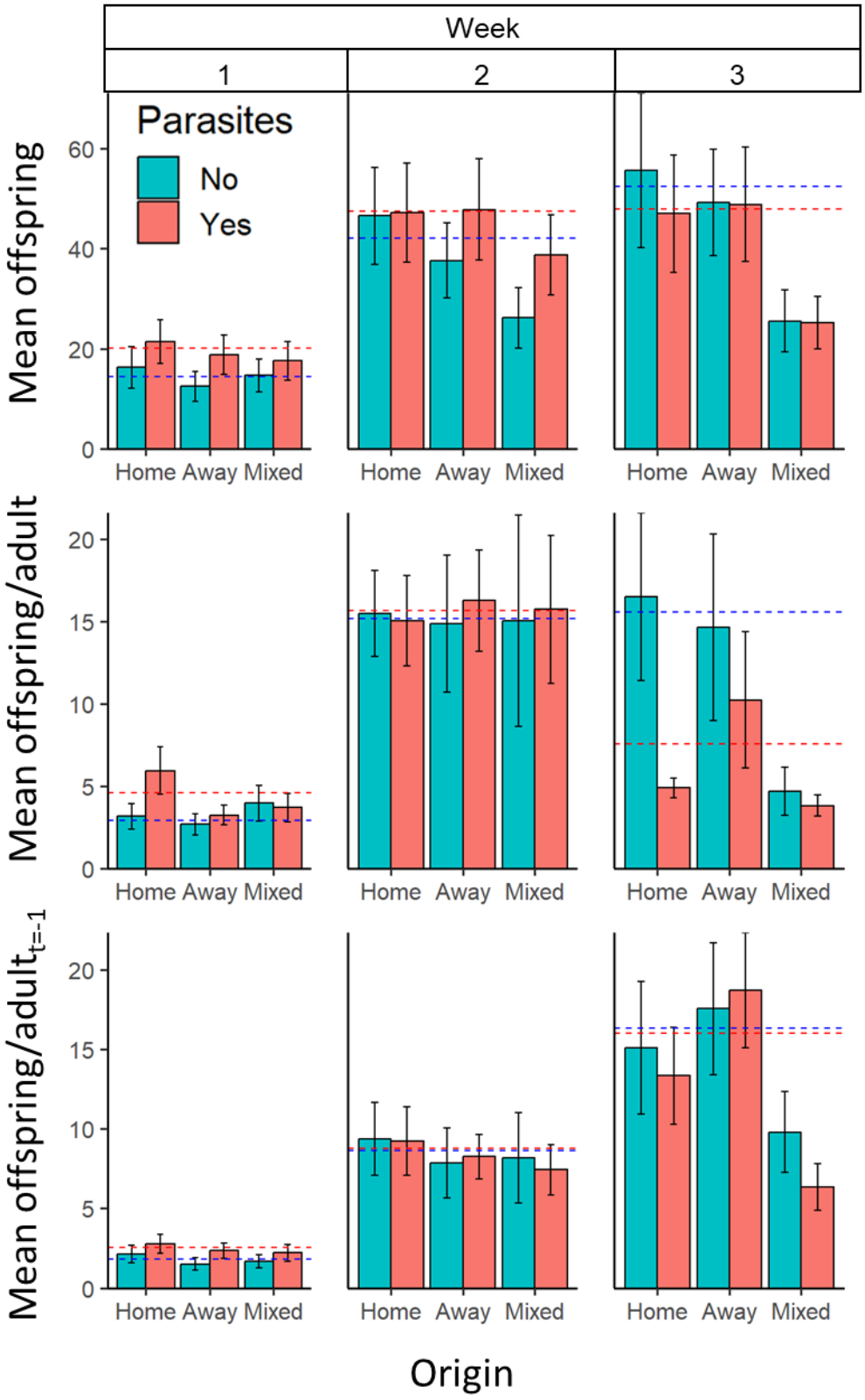
## 5.7 Acknowledgements

Thanks first go to Dr Stuart Auld for his project development and his visiting international student for conducting a preliminary experiment. Second, extra special thanks must go to Dr Peter Hunter, for the opportunity to use his equipment, and to his lab group for their extensive help in chlorophyll data collection and analysis, including the lab technician Matthew Blake and senior technician Ian Washbourne (who both ran the HPLC and analyzed the output).

## 5.8 Supplementary information




**Supplementary figure S5.1.** The relative location of each paired mesocosm in the reciprocal transfer experiment. Aerial photograph of the mesocosms used for the reciprocal transplant experiment (left) and a characterization of the selected comparisons (right). Each arrow represents a pond pair (reciprocal transfer).




**Supplementary figure S5.2.** Additional analysis of the number of offspring. The source of the *Daphnia* in each basket (Origin) is shown on the x-axis and the colour of the bars indicates the control (blue) and parasite exposed (red) treatments. Home and away groups are statistically equivalent (all pairwise differences; p >

0.05), so they share the letter ‘a’. The mean of both the home and away groups (dashed lines) is shown for comparison against the mixed group.

**Supplementary table S5.1.** Linear model coefficients for the relationship between the change in the number of offspring per adult from the current week and the total number of adults from the current week (data with no time-lag) versus the change in the number of offspring per adult at a time-lag of one week with the total number of adults at a time-lag of one week (data with a time-lag). The treatment identifier refers to the control (P-) and parasite-exposed treatment (P+). Significant p-values are highlighted in bold ($p < 0.05$ or lower).

| Data lag | Week | Treatment | Coef. | SE | t-value | p-value |
|---|---|---|---|---|---|---|
| No | 1 | P- | -0.57 | 0.24 | -2.37 | **0.022** |
| No | 1 | P+ | -0.86 | 0.30 | -2.86 | **0.006** |
| No | 2 | P- | -2.22 | 0.92 | -2.42 | **0.025** |
| No | 2 | P+ | -1.78 | 0.67 | -2.64 | **0.012** |
| No | 3 | P- | -0.48 | 0.30 | -1.60 | **0.001** |
| No | 3 | P+ | -0.33 | 0.23 | -1.40 | 0.806 |
| Yes | 1 | P- | NA | NA | NA | NA |
| Yes | 1 | P+ | NA | NA | NA | NA |
| Yes | 2 | P- | -2.48 | 0.55 | -4.49 | 0.120 |
| Yes | 2 | P+ | -0.12 | 0.50 | -0.25 | 0.172 |
| Yes | 3 | P- | -2.28 | 0.77 | -2.95 | **0.006** |
| Yes | 3 | P+ | -1.94 | 0.71 | -2.75 | **0.009** |

**Supplementary table S5.2.** Model results for test of host local adaptation to the abiotic environment in terms of the mean number of adults each week. Test statistics are shown for a post-hoc analysis of linear mixed effects models using Tukey-adjusted pairwise comparisons of estimated marginal means (EMM). The comparison identifier refers to the origin combined with control (P-) and parasite-exposed treatment (P+).

| Response | Week | Comparison | EMM | SE | df | p-value |
|---|---|---|---|---|---|---|
| Adults | 1 | Away P- vs Home P- | -0.39 | 0.64 | 50.47 | 0.929 |
| Adults | 1 | Away P- vs Away P+ | -0.22 | 0.64 | 50.47 | 0.985 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Adults | 1 | Away P- vs Home P+ | 0.41 | 0.65 | 51.12 | 0.921 |
| Adults | 1 | Home P- vs Away P+ | 0.17 | 0.64 | 50.47 | 0.994 |
| Adults | 1 | Home P- vs Home P+ | 0.80 | 0.65 | 51.12 | 0.610 |
| Adults | 1 | Away P+ vs Home P+ | 0.63 | 0.65 | 51.12 | 0.764 |
| Adults | 2 | Away P- vs Home P- | 0.45 | 1.18 | 45.31 | 0.981 |
| Adults | 2 | Away P- vs Away P+ | 0.71 | 1.15 | 44.17 | 0.926 |
| Adults | 2 | Away P- vs Home P+ | 0.77 | 1.18 | 45.31 | 0.915 |
| Adults | 2 | Home P- vs Away P+ | 0.26 | 1.18 | 45.31 | 0.996 |
| Adults | 2 | Home P- vs Home P+ | 0.32 | 1.20 | 45.34 | 0.993 |
| Adults | 2 | Away P+ vs Home P+ | 0.06 | 1.18 | 45.31 | 1.000 |
| Adults | 3 | Away P- vs Home P- | -1.23 | 3.69 | 43.12 | 0.987 |
| Adults | 3 | Away P- vs Away P+ | -1.27 | 3.66 | 42.32 | 0.985 |
| Adults | 3 | Away P- vs Home P+ | 1.20 | 3.76 | 43.24 | 0.989 |
| Adults | 3 | Home P- vs Away P+ | -0.04 | 3.74 | 43.26 | 1.000 |
| Adults | 3 | Home P- vs Home P+ | 2.43 | 3.82 | 43.38 | 0.920 |
| Adults | 3 | Away P+ vs Home P+ | 2.47 | 3.81 | 43.31 | 0.915 |

**Supplementary table S5.3.** Model results for test of host local adaptation to the abiotic environment in terms of the mean change in the number of offspring per adult at a time-lag of one week. Test statistics are shown for a post-hoc analysis of linear mixed effects models using Tukey-adjusted pairwise comparisons of estimated marginal means (EMM). The comparison identifier refers to the origin combined with control (P-) and parasite-exposed treatment (P+).

| Response | Week | Comparison | EMM | SE | df | p-value |
|---|---|---|---|---|---|---|
| Offspring | 1 | Away P- vs Home P- | -0.47 | 0.50 | 48.88 | 0.786 |
| Offspring | 1 | Away P- vs Away P+ | -0.79 | 0.50 | 48.88 | 0.406 |
| Offspring | 1 | Away P- vs Home P+ | -1.16 | 0.52 | 49.47 | 0.124 |
| Offspring | 1 | Home P- vs Away P+ | -0.32 | 0.50 | 48.88 | 0.921 |
| Offspring | 1 | Home P- vs Home P+ | -0.69 | 0.52 | 49.47 | 0.547 |
| Offspring | 1 | Away P+ vs Home P+ | -0.37 | 0.52 | 49.47 | 0.892 |
| Offspring | 2 | Away P- vs Home P- | -0.99 | 1.99 | 43.57 | 0.959 |
| Offspring | 2 | Away P- vs Away P+ | 0.65 | 1.93 | 42.73 | 0.987 |
| Offspring | 2 | Away P- vs Home P+ | 0.34 | 1.99 | 43.57 | 0.998 |
| Offspring | 2 | Home P- vs Away P+ | 1.64 | 1.95 | 43.40 | 0.833 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Offspring | 2 | Home P- vs Home P+ | 1.34 | 1.98 | 43.39 | 0.906 |
| Offspring | 2 | Away P+ vs Home P+ | -0.31 | 1.95 | 43.40 | 0.999 |
| Offspring | 3 | Away P- vs Home P- | 3.25 | 5.96 | 35.82 | 0.947 |
| Offspring | 3 | Away P- vs Away P+ | -2.50 | 5.87 | 35.59 | 0.974 |
| Offspring | 3 | Away P- vs Home P+ | 5.19 | 6.25 | 37.36 | 0.840 |
| Offspring | 3 | Home P- vs Away P+ | -5.75 | 5.58 | 35.09 | 0.733 |
| Offspring | 3 | Home P- vs Home P+ | 1.94 | 6.01 | 37.13 | 0.988 |
| Offspring | 3 | Away P+ vs Home P+ | 7.69 | 5.83 | 35.37 | 0.557 |

**Supplementary table S5.4.** Model results for the cost of migrant competition with locals in terms of the mean number of adults. Test statistics are shown for a post-hoc analysis of linear mixed effects models using Tukey-adjusted pairwise comparisons of estimated marginal means (EMM). All of the comparisons were made between the mean of both the home and away groups (Comb.) versus the mixed group. The treatment (Trt) refers to the control (P-) and parasite-exposed treatment (P+).

| Response | Week | Comparison | Trt | EMM | SE | df | p-value |
|---|---|---|---|---|---|---|---|
| Adults | 1 | Comb, vs Mixed | P- | 0.92 | 0.49 | 53.66 | 0.257 |
| Adults | 1 | Comb. vs Mixed | P+ | -0.31 | 0.49 | 53.66 | 0.925 |
| Adults | 2 | Comb. vs Mixed | P- | -1.09 | 1.32 | 49.66 | 0.841 |
| Adults | 2 | Comb. vs Mixed | P+ | -0.88 | 1.32 | 49.66 | 0.908 |
| Adults | 3 | Comb. vs Mixed | P- | -1.09 | 2.57 | 47.58 | 0.974 |
| Adults | 3 | Comb. vs Mixed | P+ | -0.19 | 2.66 | 47.87 | 1.000 |

**Supplementary table S5.5.** Model results for the cost of migrant competition with locals in terms of the mean change in the number of offspring per adult at a time-lag of one week. Test statistics are shown for a post-hoc analysis of linear mixed effects models using Tukey-adjusted pairwise comparisons of estimated marginal means (EMM). All of the comparisons were made between the mean of both the home and away groups (Comb.) versus the mixed group. The treatment (Trt) refers to the control (P-) and parasite-exposed treatment (P+). The only significant p-value is shown in bold for the parasite-exposed treatment in week three ($p < 0.05$).

| Response | Week | Comparison | Trt | EMM | SE | df | p-value |
|---|---|---|---|---|---|---|---|
| Offspring | 1 | Comb, vs Mixed | P- | -0.04 | 0.41 | 52.94 | 1.000 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Offspring | 1 | Comb. vs Mixed | P+ | 0.29 | 0.41 | 52.94 | 0.894 |
| Offspring | 2 | Comb. vs Mixed | P- | 0.06 | 2.01 | 48.54 | 1.000 |
| Offspring | 2 | Comb. vs Mixed | P+ | 0.77 | 1.98 | 48.38 | 0.980 |
| Offspring | 3 | Comb. vs Mixed | P- | 5.60 | 3.45 | 40.04 | 0.376 |
| Offspring | 3 | Comb. vs Mixed | P+ | 9.76 | 3.44 | 39.98 | **0.035** |

**Supplementary table S5.6.** All possible explanations for the significant difference in the mean change in the number of offspring / $adult_{t=-1}$ between both the Home and Away treatments versus the Mixed treatment under parasite exposure. Both observed and expected values for the mean change in the number of offspring / $adult_{t=-1}$ are shown. The observed values for the Mixed (M), Home (H) and Away (A) treatments are shown, along with the mean of the Home and Away treatments (H+A/2). The Cost / $adult_{t=-1}$ refers to the cost of mixed competition in terms of the difference between the mean change in the number of offspring / $adult_{t=-1}$ for the Mixed versus the mean of the Home and Away treatments (H+A/2). Assuming that one half of the Mixed treatment is made up of local *Daphnia* adults and the other half is made up of migrant *Daphnia* adults, we calculated the expected mean change in the number of offspring / $adult_{t=-1}$ owing to either the local (H (H+A/2) or migrant *Daphnia* adults (A (H+A/2). The expected cost of mixed competition for both the local (Cost (M H) and migrant *Daphnia* adults present in the Mixed treatment (Cost (M A), and the combined total of these costs, are also shown. A) Only away loses: The mean change in the number of offspring / $adult_{t=-1}$ for local *Daphnia* is the same in the mixed treatment as it is in the (unmixed) home treatment, but lower in the mixed treatment compared to the (unmixed) away treatment for migrant *Daphnia*. Therefore, the net mean change in the number of offspring / $adult_{t=-1}$ for away *Daphnia* under parasite exposure in week 3 is negative, and positive for local *Daphnia*, so locals win. B) Only home loses: The mean change in the number of offspring / $adult_{t=-1}$ for migrant *Daphnia* is the same in the mixed treatment as it is in the (unmixed) home treatment, but lower in the mixed treatment compared to the (unmixed) away treatment for local *Daphnia*. Therefore, the net mean change in the number of offspring / $adult_{t=-1}$ for local *Daphnia* under parasite exposure in week 3 is negative, and positive for migrant *Daphnia*, so migrants win. C) Both home and away lose, but away wins overall: The mean change in the number of offspring / $adult_{t=-1}$ for both local and migrant *Daphnia* is lower in the mixed treatment compared to the (unmixed) home and away treatments as a proportion of their individual mean change in the number of offspring / $adult_{t=-1}$ divided by a combination of the two.

Therefore, despite a significant reduction in the fecundity of both local and migrant *Daphnia*, the net mean change in the number of offspring / adult$_{t=-1}$ for local *Daphnia* under parasite exposure in week 3 is slightly negative, and slightly positive for migrant *Daphnia*, so migrants just win out (N.B. there is some uncertainty in this result as the error associated with each expected mean is likely to be overlapping with zero).

| Observed | | | | | |
|---|---|---|---|---|---|
| **Example** | **Mixed (M)** | **Home (H)** | **Away (A)** | **H+A/2** | **Cost/adult$_{t=-1}$** |
| A | -0.3 | 4.9 | 12.9 | 8.9 | 9.2 |
| B | -0.3 | 4.9 | 12.9 | 8.9 | 9.2 |
| C | -0.3 | 4.9 | 12.9 | 8.9 | 9.2 |
| **Expected** | | | | | |
| **Example** | **H (H+A/2)** | **A (H+A/2)** | **Cost (M H)** | **Cost (M A)** | **Total cost** |
| A | 2.4 | -2.7 | 0.0 | 9.2 | 9.2 |
| B | -6.8 | 6.5 | 9.2 | 0.0 | 9.2 |
| C | -1.0 | 0.7 | 3.5 | 5.7 | 9.2 |

## Fig. S5.3: Variation in fecundity is not linked to the total number of adults

Assuming that reproduction and population growth rate did not exceed the carrying capacity in each field cage, to allow a direct comparison of adult host fecundity using the mean change in the number of offspring, rather than the mean change in the number of offspring per adult, either with or without a time-lag of one week, we would have required a constant total number of adults in each field cage. However, due to variation in adult host survival and offspring maturation, it was possible that the variation observed adult fecundity could have been driven by changes in the total number of adults each week. Therefore, to investigate the relationship between adult fecundity and the total number of adults, we performed a series of correlations using linear mixed effects models (Supplementary Figure S5.2).

We found that there was not a consistent pattern across all three weeks between the change in the number of offspring and the total number of adults from either the current or previous week. Although testing each week separately, we found a

4156 significant positive relationship between change in the number of offspring and the
4157 total number of adults from the current week in week two for both the control and
4158 parasite-exposed treatment ($t > 1.96$ for both the parasite control and treatment), this
4159 does not affect either of the interesting results from our main analysis, including A)
4160 the parasite-induced fecundity compensation observed in week one and B) the cost
4161 of mixed competition in terms of offspring in the parasite-exposed treatment in week
4162 three.
4163

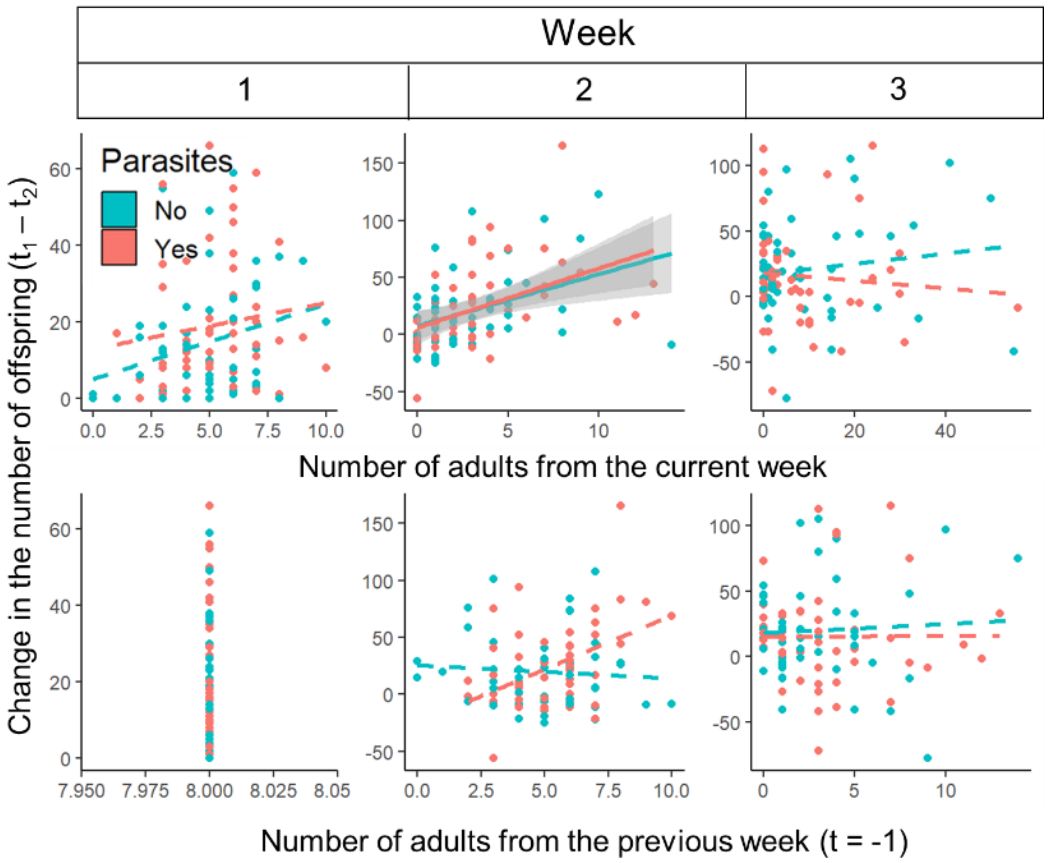



**Figure S5.3.** Variation in fecundity is not linked to the total number of adults. The different colours correspond to the control (blue) and parasite-exposed treatment (red). The linear relationship from a mixed effects model is shown for each subset of the data, where the dashed versus solid lines correspond to non-significant versus significant effects. 95% confidence intervals are shown for the significant effects (grey bands). Row one - the relationship between change in the number of offspring and the total number of adults from the current week. Row two - the relationship between change in the number of offspring and the total number of adults at a time-lag of one week.

4164
4165

## 6. Borrowing data from other populations to forecast epidemic size

## 6.1 Abstract

A key challenge for disease ecology is predicting the size of epidemics. Most models forecast disease in a single population using long-term historical data from that population. However, long-term data is not always available and a possible alternative is to borrow data from multiple similar populations to forecast disease for a population of interest. One step further is to weight the contribution of epidemics to the forecast based on their similarity to the focal population. In this study, we use data from twenty populations of the freshwater crustacean *Daphnia magna* and its sterilizing bacterial parasite *Pasteuria ramosa* tracked over four epidemic seasons (a total of 80 epidemics) to predict future epidemics. We evaluate single population, multiple average population and multiple weighted average population approaches for training three suites of forecast model: seasonal naïve, auto-regressive integrated moving average and time series regression models. We found that forecast accuracy depended on both the type of training data and the choice of forecast model, but models trained on data from multiple populations consistently outperformed those trained on single population data. Our study demonstrates the benefit of using a collection of similar populations to forecast disease for a focal population which has limited data.

## 6.1 Introduction

Epidemic size is a key metric of infectious disease and can be defined as the proportion of individuals within a population that are infected at any given time (disease prevalence), or across multiple timepoints (mean or integrated disease prevalence, see chapter two, Fig. 2.1). In the wild, epidemics occur periodically and largely predictably, but their precise magnitude often varies depending on wider environmental factors such as temperature (Altizer et al., 2006).

Predicting the precise magnitude of future epidemics is very difficult for two main reasons. First, since disease prevalence is calculated as the number of infected individuals as a proportion of the overall population size, it is the product of two varying measurements: the total number of infected and healthy hosts. This means that the variation in each separate measurement contributes to the total variation in disease prevalence. Second, there is substantial spatiotemporal variation in disease prevalence across different populations over time (Altizer et al., 2004; Aznar et al., 2015; Cáceres et al., 2006; Carlsson-Granér & Thrall, 2002; Ericson et al., 1999; Montano et al., 2016; Thrall et al., 2012; Vergara et al., 2013). For example, peak prevalence of *Metschnikowia bicuspidata* in populations of *Daphnia dentifera* varies from 0% to more than 60% across lakes (Penczykowski et al., 2016), and peak cowpox prevalence varies from 9% to >30% in field voles over the course of a season (Begon et al., 2009). The large amount of spatiotemporal variation in disease prevalence makes it difficult to identify common drivers of epidemic size across both within and between host-parasite systems.

Although producing epidemic forecasts is challenging, there are opportunities for us to use our understanding of environmental variation to better forecast disease in focal populations. Temperature is easy to measure, and variation in temperature is associated with patterns of disease prevalence in a range of host-parasite systems (Alonso et al., 2011; Auld & Brand, 2017b; Beckley et al., 2016; Bravo et al., 2020; Groner et al., 2018, 2021; Krauer et al., 2021; Ruiz-Moreno et al., 2012; Schaaf et al., 2017; Susi et al., 2017; Swinford & Anderson, 2021; Thoirain et al., 2007). For example, an approach to disease forecasting using time-series analysis can be used to incorporate information on environmental conditions. This includes autoregressive integrated moving average (ARIMA) and time-series regression models. ARIMA models make inferences based on underlying patterns of temporal autocorrelation, and despite previously having been used in disease forecasting without the addition

of environmental data (Allard, 1998; Helfenstein, 1991), they can easily be adjusted to incorporate seasonality into epidemic predictions (Hyndman & Athanasopoulos, 2021). For time-series regression, which naturally rely on the effect of predictor variables (Hyndman & Athanasopoulos, 2021), they have often been used to forecast cases of vector-borne viruses, such as malaria and dengue, using environmental factors, such as temperature and rainfall (Gao et al., 2012; Hii et al., 2012; Hu et al., 2006).

In addition, there is the potential to use information from a group of similar populations to forecast disease in a population of interest. For some infectious disease systems, we have a lot of data and a long-term dataset for a single population of interest. Whereas, for others, we have little data available for the focal population, but data from various other populations that might vary in their similarity to the focal population. To reflect these differences in data availability, it would be possible to compare models trained only on individual populations (single population models), models which exclude the focal population and are trained on averages from the other remaining populations (average population models) and models which use weighted averages based on similarity to the focal population in terms of environmental temperature (weighted average population models).

It might also be better to forecast other additional components of disease rather than just disease prevalence. Studies of infectious diseases in the wild usually use disease prevalence as a measure of epidemic size (Jennelle et al., 2007). However, predicting the density of infected hosts (incidence) and healthy hosts over time might be preferable for some systems. For example, the risk of infection to certain vector-borne diseases is driven by infected vector density (Pepin et al., 2012) and conservation may only be interested in predicting the number of healthy hosts.

Here, we used epidemic and temperature data collected from 20 replicated semi-natural *Daphnia*-parasite pond populations over four seasons (80 epidemics) to predict three variables over time (disease prevalence, infected host density and the density of healthy hosts) using three sets of training data (single population, average population and population data weighted by temperature similarity) to train three suites of forecast models (benchmark, ARIMA and regression). Additional regression models were built to compare the use of photoperiod data for predicting different response variables. We expected that the models trained on multiple populations

would perform better than the models trained on only the target population due to the inclusion of a larger amount of data and that the ARIMA and regression models would outperform the benchmark models due to their ability to model more complex time series patterns. Therefore, our two main hypotheses were (i) the models trained on multiple populations would perform better than the models trained on only the target population and (ii) the ARIMA and regression models would outperform the benchmark models. However, our results were nuanced: we found that the performance of models varied according to the type of data used to train the models and the class of forecast model used.

## 6.2 Methods

### 6.2.1 Study system

In this study, we focused on the *Daphnia magna-Pasteuria ramosa* host-parasite system. *D. magna* is a small freshwater crustacean and naturally occurs with the obligate sterilizing bacterial micro-parasite, *P. ramosa.* In the wild, *D. magna* populations experience regular epidemics of *P. ramosa* on an annual basis. Initially, hosts become infected when they ingest parasite spores from the pond sediment during filter feeding and epidemics begin as host densities peak in spring (Ebert, 2005). Parasite prevalence fluctuates throughout the summer and declines in the autumn, with parasites often disappearing completely in winter due to a drop in host density.

### 6.2.2 Pond experiment

To start with, replicate lines of the 12 genotypes of *Daphnia magna* were maintained in the laboratory in a state of clonal reproduction for three generations to reduce variation due to maternal effects. There were five replicates per genotype; each replicate consisted of five *Daphnia* kept in 200 ml of artificial medium (Klüttgen et al., 1994) modified using 5% of the recommended SeO2 concentration(Ebert et al., 1998). Replicate jars were fed 5.0 ABS of *Chlorella* vulgaris algal cells per day (where ABS is the optical absorbance of 650 nm white light by the *Chlorella* culture). *Daphnia* medium was changed three times per week and three days prior to the start of the pond experiment. On the day that the pond experiment commenced, 1–3 day old offspring were pooled according to host genotype. Ten offspring per genotype

were randomly allocated to each of the 20 ponds (giving a total of 120 *Daphnia* per pond).

Each pond consisted of a 0.65 m tall 1000 litre PVC tank filled with rainwater. The ponds were set to different depths into the ground and experienced different temperature profiles (Auld & Brand, 2017b). In addition, six of the ponds experienced a weekly mixing treatment where mixed ponds were stirred once across the middle and once around the circumference with a 0.35 $m^2$ paddle submerged halfway into the pond (the exception to this was on the first day of the experiment, when all ponds experienced the mixing treatment to ensure hosts and parasites were distributed throughout the ponds).

The experiment began on the 2nd April 2015 (Julian day 98), when 120 *Daphnia* (10 *Daphnia* x 12 genotypes) and 1 x $10^8$ *Pasteuria* spores from the mastermix were added to each of the 20 ponds. The mastermix comprised *Pasteuria ramosa* spores propagated using 21 separate *Daphnia* genotypes exposed to sediment from their original pond (Kaimes, Scottish Borders, UK, Auld & Brand, 2017). Seasonal epidemics were tracked for the next four years between April 2015 and November 2018. This involved weekly measurements of parasite prevalence, the number of diseased adults, the number of healthy adults and pond temperature between either April or May and November each year (see Chapter 3, Auld & Brand, 2017b Paplauskas et al., 2021).

### 6.2.3 Format of the time series

For each population, there were multiple time series data collected across four years for the following variables, temperature, the number of healthy adults and the number of diseased adults. Disease prevalence, healthy adult density and diseased adult density were calculated from the original time series data. Density data were all natural logarithm transformed prior to model construction.

The set of multiple time series for each population was collectively referred to as a multivariate time series (MTS). The first three years of data were used to train the forecast models (hereafter referred to as training data) and were evaluated against the final year of data (hereafter referred to as the test data).

All of the MTS consisted of weekly data based on the UK convention where week of the year was represented as a decimal number (00–53) using Monday as the first day of week (and typically with the first Monday of the year as day 1 of week 1). Means were taken of weeks with multiple recorded values. Missing values were added via linear interpolation so that the time series length each year, referred to as the frequency, was constant across the first three years, 31 weeks, except where there were three or more missing values in a row in which case NA values were used. The final year had a frequency of 26 weeks.

### 6.2.4 Model construction

Three types of training data were used for three distinct modelling frameworks (see later) and were used to predict disease prevalence, log diseased adult density and log healthy adult density. The first two sets of training data, the mean and temperature weighted mean, borrowed data from other populations to make predictions about disease in the focal population, whereas the third set of training data was taken from the test population and acted as a control by only using historical data from the focal population. These three classes of training data were created as follows.

The calculation of the mean and weighted mean data consisted of two steps. First, the time series data for disease prevalence, log diseased adult density and log healthy adult density for all 20 populations was divided into all possible combinations of 19 training populations and 1 test population (Fig. 6.1A). Using different sets of training and test data was a form of cross-validation; different sets of data were used to evaluate the accuracy of forecast models for a number of different focal populations. Second, the mean training data was calculated across each set of training populations. Similar to the mean training data, the weighted mean datasets were calculated using the following equation:

$$\frac{\sum wx}{\sum w}$$

where $w$ is the weighting calculated as the similarity of the average temperature each year between the training and test populations and $x$ is the disease prevalence, log diseased adult density or log healthy adult density at each week. Where the difference between the training and test population was zero, 0.1 was used to avoid producing ‘not a real number’ when calculating the similarity. For the time series

regression models, the mean and weighted mean temperature and photoperiod were also calculated for use as independent variables.

The mean and weighted mean training data underwent a final calculation depending on the type of forecast used. For the first set of forecasts, each epidemic was treated as separate from one another to calculate a global average of the mean and weighted mean data (Fig. 6.1B). For the second and third set of forecasts, epidemics from the same population were grouped into time series to model patterns across years (Fig. 6.1C).

The third set of training data came from the test population data. For the first set of forecast models, this consisted of the third year of epidemic data. For the second and third set of forecast models, this consisted of the first three years of epidemic data.

For each set of training data, there was the corresponding test data. This consisted of the fourth year of epidemic data and was used as validation of forecast accuracy. Finally, the MTS containing the training and test data for all three suites of forecast models were converted into time series objects.

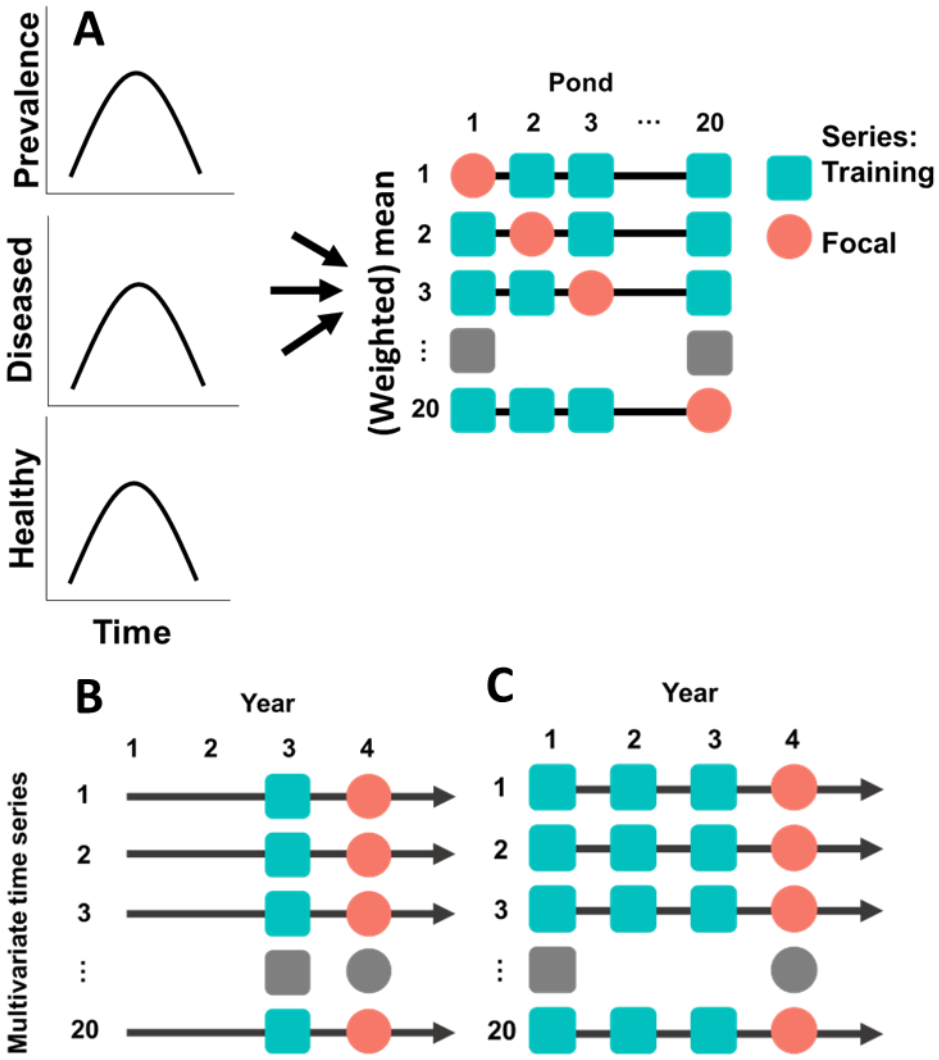



Figure 6.1. Model construction. A) Calculation of mean and weighted mean training data. The time series data for all 20 populations (ponds) were divided into all possible combinations of 19 training populations and 1 test population. The mean training data was calculated across each set of training populations, whereas the weighted mean training data used only the five most similar populations to the test population. B) The final multivariate time series (MTS) for the benchmark models. For each population, there were time series for the three classes of training data and the corresponding test data. The mean and temperature weighted mean training data consisted of a global average calculated across the first three years of epidemic data (blue squares). For the test population data, only the third year of epidemic data was used as training data (blue squares) and the fourth year was used to evaluate the accuracy of the benchmark forecasts (red circles). C) The final multivariate time series (MTS) for the ARIMA and

regression models. For each population, there were time series for the three classes of training data and the corresponding test data. The training data consisted of the first three years epidemic data (blue squares) and the test data consisted of the fourth year of epidemic data (red circles). For both B and C, each MTS had a frequency of 31 weeks in the first three years and a frequency of 26 in the final year. The direction of time is represented by black arrows. The grey shapes represent skipped MTS.

### 6.2.5 Forecasting

As mentioned above, each set of training data was combined with three suites of forecasting models, including a benchmark, an autoregressive integrated moving average (ARIMA) and a time-series regression-based model. The first group of models was the benchmark group that provided a baseline comparison for the accuracy of more complex models. This benchmark was produced using a Seasonal Naïve forecasting technique, which is a type of time-series analysis that is quite basic, as it assumes that the forecast is directly equal the same observed value as in the previous season (Hyndman, 2021), but works remarkably well for many economic and financial time series (Hyndman, 2021). The second group of models included the ARIMA models, which were based on a linear combination of past values of the variable ('autocorrelation'), past forecast errors ('moving average') and extended to incorporate the seasonality of the time-series data (seasonal autoregressive integrated moving average, often referred to as 'SARMIA', Hyndman, 2021). The third group of forecasting models used time-series regression to predict the time series for (x) by assuming that it had a linear relationship with temperature (sensu a predictor variable, y, Hyndman, 2021). The temperature values of the test population from the fourth year of epidemic data were used as 'future' values for the time series regression group of models.

For the benchmark models, the mean and weighted mean training data was used differently to the training data which came from the test population. Specifically, the mean and weighted mean training data used the global average of epidemic data, whereas the training data which came from the test population used only the third year of epidemic data. For the ARIMA models, forecasts were produced using the auto.arima function in R which selected the model with the lowest corrected Akaike information criterion (AICc, Hyndman, 2021).

For the time series regression models, forecasts were produced using temperature, Julian Day (a proxy for photoperiod) and a combination of the two as predictor variables. Both temperature and photoperiod are important in influencing epidemics. Photoperiod captures the seasonality of the data which is explicitly modelled in the benchmark and ARIMA models. However, seasonality is not central to a standard regression model in the way it is in an ARIMA, so we therefore modelled photoperiod independently in the regression models.

For regression models using temperature as a predictor, mean and temperature weighted mean disease prevalence, log diseased adult density and log healthy adult density were fitted against mean temperature, whereas the test population data were fitted against test population temperature. For regression models using Julian Day as a predictor, Julian Day was treated as a polynomial. For regression models using log healthy and log diseased adult density as predictors, the mean, weighted mean and test population data was fitted against mean, weighted mean and test population log diseased and log healthy adult host density respectively. Multiple regression was used for models with a combination of temperature and photoperiod as predictors. Test population temperature, photoperiod, log diseased and log healthy adult density from the fourth year of epidemic data were used as ‘future’ values for regression forecast models.

For all three suites of forecasting models, a forecast horizon of 26 weeks was used to match the time series frequency in the final year of data. 17 forecasts were made for each of the mean, temperature weighted mean and test population data, excluding years in which there was no epidemic. Forecast accuracy was determined using root mean squared error (RMSE), which measured the difference between predicted and observed values. For the ARIMA and time series regression models, the forecasts predicted some negative values on a normal scale and some extremely small values on a log scale. These were removed by zero-bounding the disease prevalence and effectively zero-bounding the log diseased and log healthy adult density (log(0.01)-bound). This involved adding the difference between zero and the minimum disease prevalence value to the forecast or adding the difference between log(0.01) and the minimum log diseased or log healthy adult density to the forecast. The average adjusted RMSE was calculated for each set of training data.

### 6.2.6 Analysis of model error

Significant differences between mean model errors were calculated using estimated marginal means (EMMs) from generalised least squares (GLS) models accounting for unequal variances among the three classes of training data. Two sets of three EMMs were performed in total, one for each of the three forecast variables and one for each set of comparisons. The first three EMMs calculated the significance of differences between models trained on different data grouped by the type of forecast model. The second three EMMs calculated the significance of differences between forecast models grouped by training data.

Also, GLS was used to test the significance of the association between the mean error of forecasts and the average temperature in the fourth year of the focal population for each type of forecast model.

## 6.3 Results

As expected, there was a clear seasonal pattern in temperature across the 20 populations and across all years (Fig. 6.2). This showed that temperature increased from approximately 10°C at the beginning of each year, peaked at approximately 15 to 20°C around the halfway point of in the season and then declined back to approximately 10°C by the end of each season. There was substantial variation in disease prevalence across the 20 populations and across all years (Fig. 6.2). Further inspection shows this variation in disease prevalence was driven principally by variation in log diseased adult density; log healthy adult density was more consistent across epidemics. Healthy adult density consistently peaked at around 10 weeks.

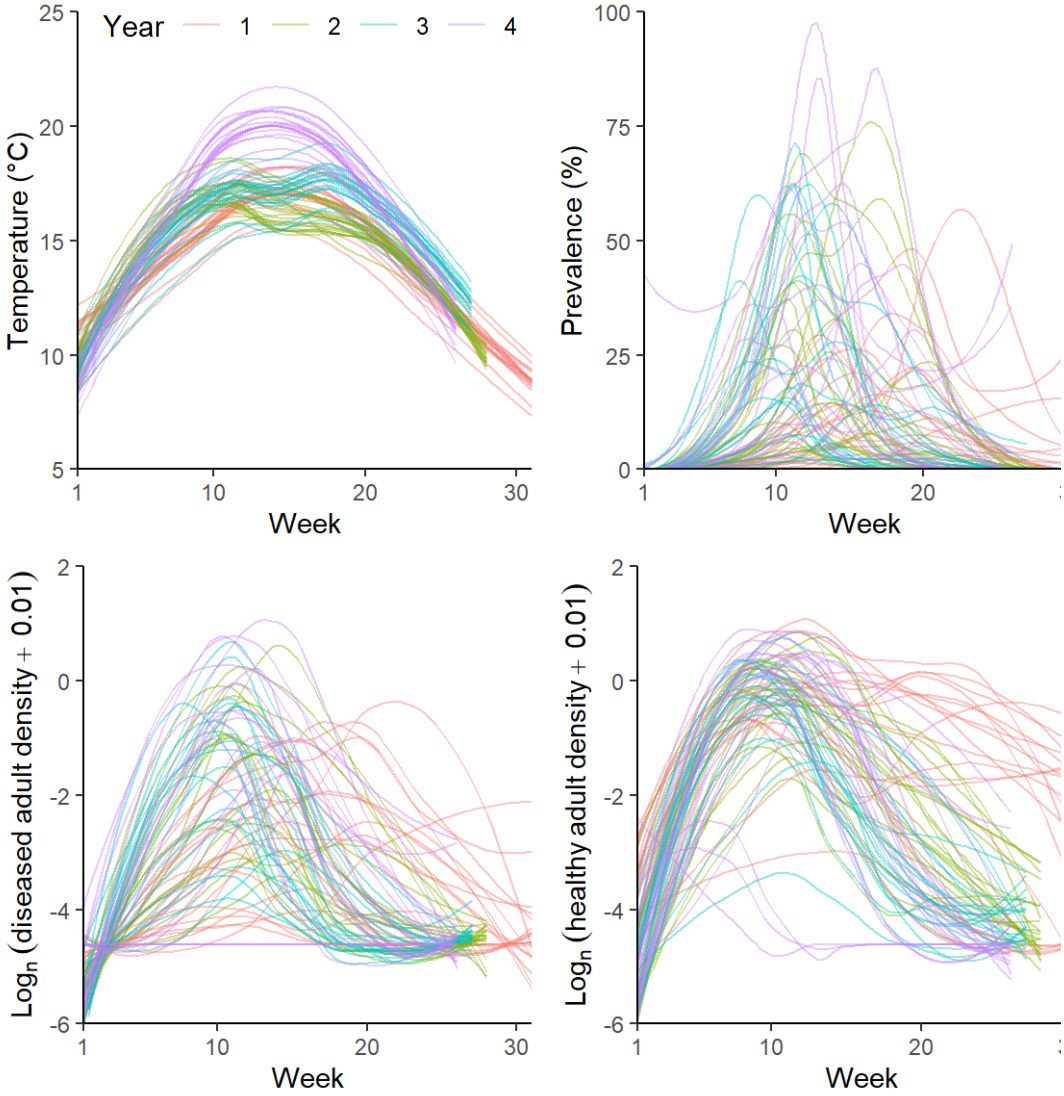



Figure 6.2. Plots A-D show temperature, infection prevalence, log diseased adult density and log healthy adult density over time respectively. Lines are loess fits for all populations. Prevalence was zero-bound by performing loess on a $\log(1+x)$ transformation and then back-transforming the result using the inverse, $\exp(x)-1$. Years one to four are indicated by the colour of the lines.

There was significant variation in forecast accuracy between ponds (Fig. 6.3). For the first suite of forecasts, benchmark forecasts, where the forecasts were equal to the values observed in the previous season, the mean and temperature weighted mean forecasts of disease prevalence were more similar to the test data than the test population forecasts in pond one and two, but not pond three. For benchmark forecasts of log diseased and log healthy adult density, there was less variation among the different classes of forecast in terms of similarity to the test data.

Similar to the first suite of forecasts, the mean and temperature weighted mean forecasts of disease prevalence for the second suite of forecasts, ARIMA forecasts, were more similar to the test data than the test population forecasts in pond one and two, but not pond three. The automated function which was used to develop the ARIMA forecasts selected a structure equivalent to the benchmark forecasts in some cases, such as for the test population forecasts, except for the test population forecast of disease prevalence in pond two. Similar to the first suite of forecasts, the ARIMA forecasts of log diseased and log healthy adult density showed a small amount of variation among the different classes of forecast in terms in similarity to the test data.

For the third suite of forecasts, regression forecasts, which used temperature as a predictor, in comparison to the first and second suite of forecasts, the test population forecasts of disease prevalence were similar to the mean and temperature weighted mean forecasts. However, similar to the first and second suite of forecasts, the regression forecasts of log diseased and log healthy adult density showed a small amount of variation among the different classes of forecast in terms in similarity to the test data.

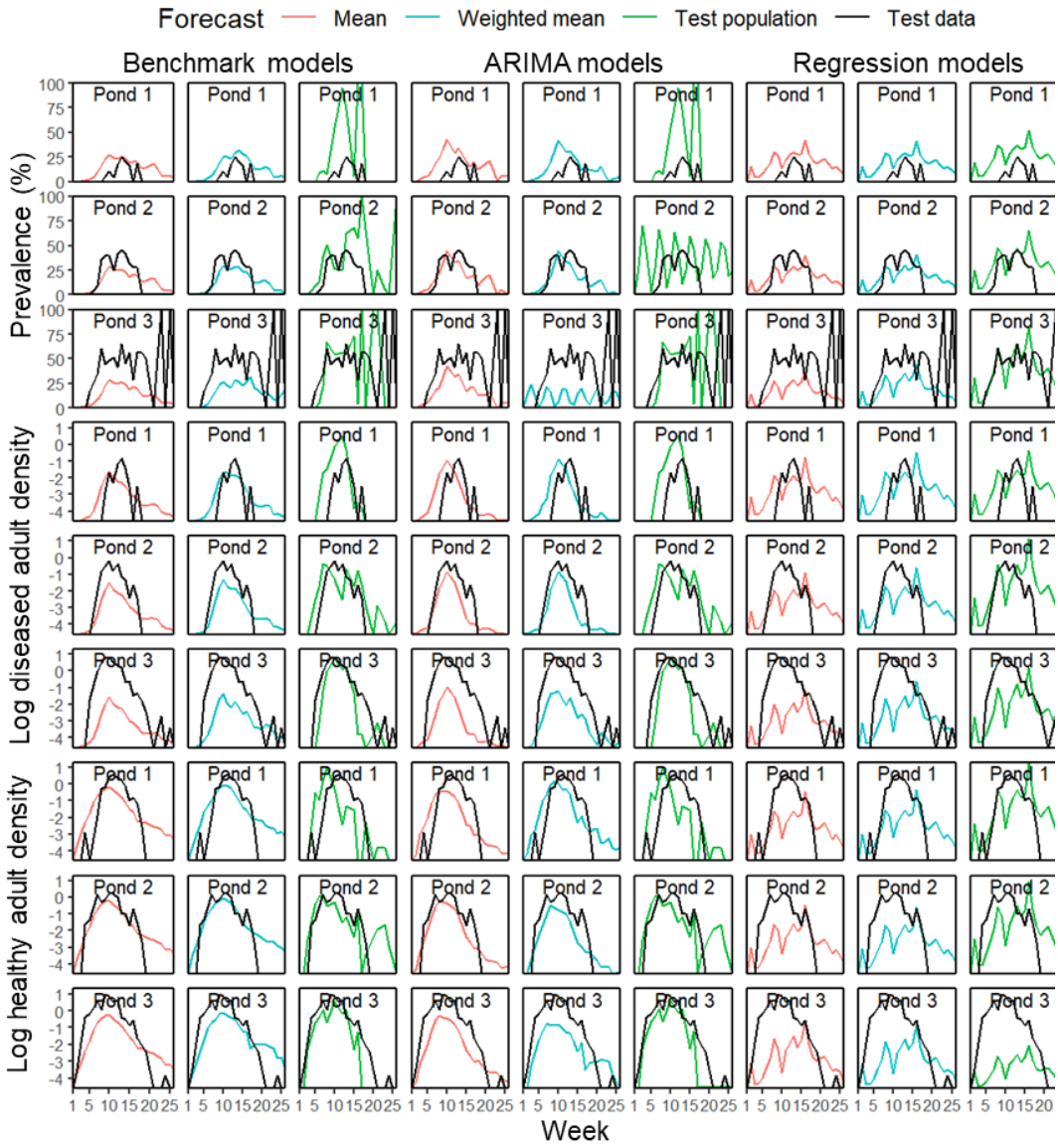



Figure 6.3. Variation in forecast accuracy between three example ponds. For each forecast variable, disease prevalence, log diseased adult density and log healthy adult density, there were three suites of model, benchmark, where the forecast was equal to the values observed in the previous season, auto-regressive integrated moving average (ARIMA) and regression models which used temperature as a predictor. For each model, there were three classes of training data which produced three separate forecasts, the mean, temperature weighted mean and test population forecasts which are indicated by the red, blue and green lines respectively, as well as the corresponding test data which was used to evaluate the accuracy of forecasts as indicated by the black lines. The text shows the different pond numbers.

*Forecasting infection prevalence*

The mean error for forecasts of disease prevalence, incidence and healthy host density varied according to the model and the training data (Fig. 6.4). Mean error for forecasts of disease prevalence was not significantly different between the three suites of forecast models trained on either the mean or the weighted mean data. In contrast, the mean error of regression models trained on only the test population data was significantly lower than the benchmark and ARIMA models trained on the same data. The mean error for forecasts of incidence was not significantly different between models, but the range of error was very low for the regression models. However, the mean error for forecasts of healthy host density was significantly different between models. Specifically, the mean error for benchmark and ARIMA models was significantly lower than the mean error for the regression models. In addition, the range of error was very low for the forecast models of healthy host density.

*Forecasting density of infected hosts (infection incidence)*

Second, comparisons of models trained on different sets of data were made. For forecasts of disease prevalence, the mean error of benchmark and ARIMA models trained on either the mean or the weighted mean data were significantly lower than those trained on only the test population data. In contrast, there was no significant difference in the mean error between regression models trained on different data. For forecasts of incidence, there was no significant difference between the mean error of benchmark and regression models. However, the mean error of the ARIMA models trained on the mean and weighted mean data was significantly lower than those trained on only the test population data.

*Forecasting healthy host density*

For forecasts of healthy adult density, there was no significant difference between the mean error of benchmark models. In comparison, the mean error of ARIMA and regression models trained on the mean and weighted mean data was significantly lower than those trained on only the test population.

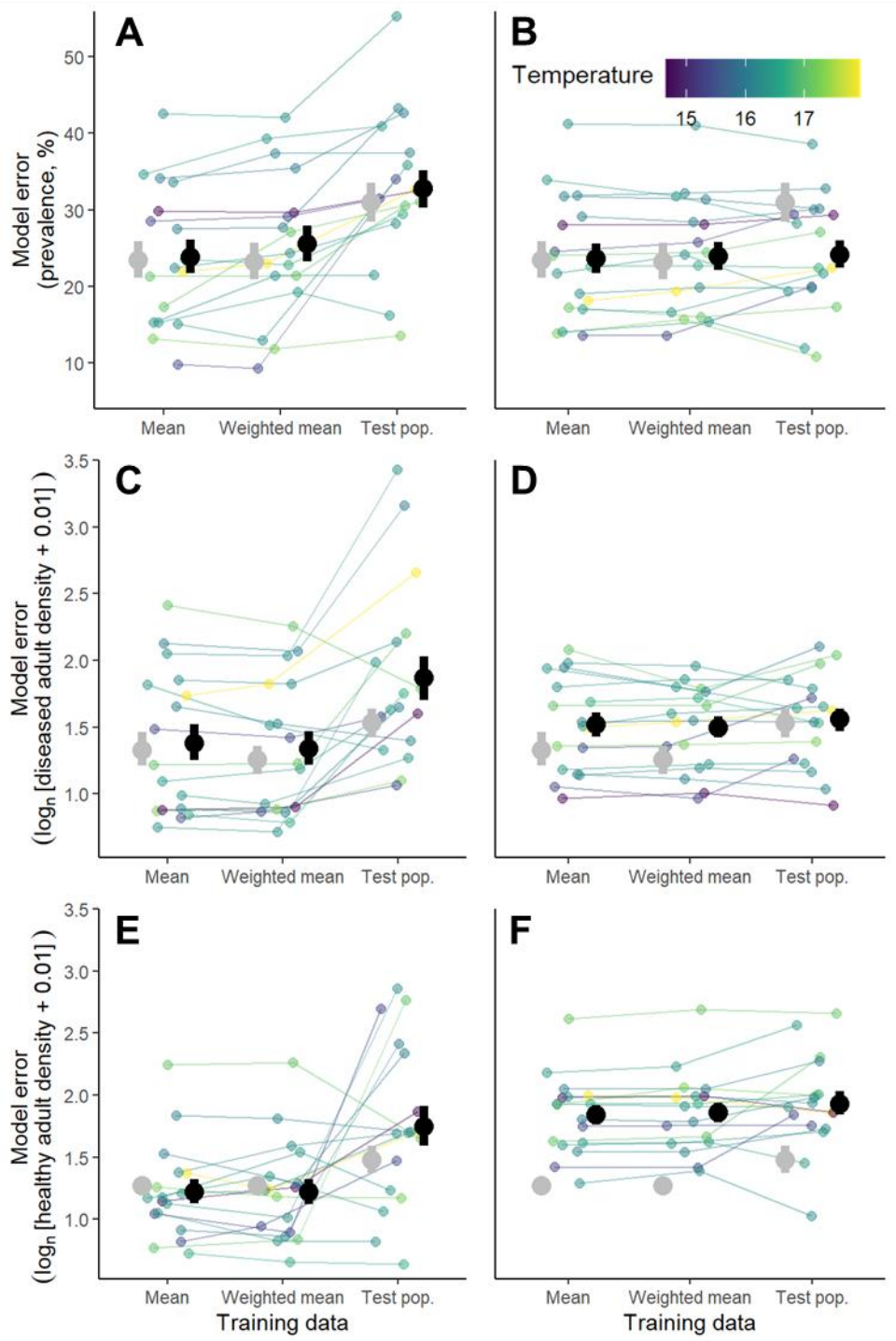



Figure 6.4. Model error for forecasts of disease prevalence, log diseased adult density and log healthy adult density. The mean model error across all 20 populations (large points), with error bars indicating the standard error, and the individual model error for each population (small points), with the colour scale indicating the average temperature of the focal population in year four and with lines connecting forecasts with the same population of interest, are shown for three sets of models, benchmark (grey points), ARIMA (black points in the first column

of panels; A, C, E) and regression (black points in the second column of panels; B, D, F) and three sets of training data, the mean (Mean), temperature weighted mean (Weighted mean) and test population (Test pop.) data. Model error is root mean squared error.

## 6.4 Discussion

A key aim of epidemiology is to predict the timing and size of epidemics. This is challenging because of natural variation in epidemics caused by environmental and ecological factors, such as temperature (Alan Pounds et al., 2006; Descloux et al., 2012; El-Sayed & Kamel, 2020; Zell, 2004) and epidemic termination due to rapid evolution of host resistance (Duffy et al., 2012). Most studies tend to focus on predicting disease in a single population based on historical data from that population. We asked whether it is possible to borrow data from other populations to predict future epidemics, including when weighting the influence of populations based on their environmental similarity to the population of interest. We found that models trained on mean and ecologically weighted mean data often performed better than those trained only on the focal population data.

Overall, we found that disease prevalence, log diseased adult density and log healthy adult density are difficult to predict. This was reflected in both the variation in the range of model errors and the relatively large mean errors. In our first set of results, which tested the significance of differences between models trained on different data grouped by the type of forecast model, the mean error for forecast models of disease prevalence based on historical data and trained on the mean and temperature weighted mean data was significantly lower than the mean error for the same models trained on only the test population data. Therefore, in agreement with our hypothesis, we demonstrated the potential for forecasting disease prevalence in populations where there is no historical data by using data from replicate populations across space. This approach can be easily generalised to other systems where there is a lack of research about forecasting disease in the wild.

In our second set of results, which tested the significance of differences between forecast models grouped by training data, we found that the mean error of models which were trained on the test population data and used future temperature values to predict disease prevalence was significantly lower than models which used historical data. This demonstrates the benefit of data-rich systems for predicting

disease from only one population. However, for forecasts of log diseased and log healthy adult density, the models which used future temperature values performed the same as or worse than the models which used historical data.

Also, we found a significant association between the mean error for the benchmark forecasts of log diseased adult density and the average temperature of the focal population, as well as between the mean error for all of the regression models and the average temperature of the focal population.

Previous basic ecological models of disease have focused on Daphnia-parasite systems using a small number of populations and a few select ecological drivers, such as host density and temperature (Duffy et al., 2005; Duffy & Sivars-Becker, 2007; Hall, Becker, et al., 2009; Hall, Knight, et al., 2009). These studies have been useful in understanding how ecological drivers of disease affect the timing of epidemics, but they are not forecasts of disease prevalence and the emergent models are often highly specific and thus lack generality. In this study, we performed forecasts of disease prevalence, log diseased adult density and log healthy adult density using limited data from across a group of spatially explicit populations with ecologically realistic variation. We found that models trained on mean and ecologically weighted mean data often performed better than those trained only on the focal population data and this approach can be easily generalized to other systems.

A potential shortcoming of using mean and weighted mean data from multiple populations to forecast disease rather than data from a single population is the trade-off between being roughly accurate most of the time with being highly accurate some of the time. However, the spread of model errors shows that this is not the case because they are roughly the same between models trained on the mean and temperature weighted mean data compared to the test population data.

The mean error for the benchmark forecasts of log diseased adult density was significantly associated with the average temperature of the focal population, but not disease prevalence or log healthy adult density. Previous findings show that there is a strong relationship between temperature and the spread of disease. There are temperature-dependent effects on host and parasites in the *Daphnia-Pasteuria* system (Allen & Little, 2011; Vale et al., 2008) and biologically reasonable increases

in environmental temperature can cause larger epidemics (Auld & Brand, 2017b). Interestingly, the absence of any significant relationship between the mean error for benchmark forecasts of disease prevalence or healthy adult density with the average temperature of the focal population indicates that there are varying effects of temperature between diseased and healthy hosts.

Other forecasting models that are based on time series data are available including models which use artificial intelligence (Chimmula & Zhang, 2020; Lalmuanawma et al., 2020; Yang et al., 2020). However, the results of these models can be difficult to interpret. Although benchmark forecasts are commonly outperformed by these more complex forecasting techniques (Abbasimehr & Paki, 2021; Baquero et al., 2018; Perone, 2022), our study shows that simple benchmark forecasting techniques can still produce the most accurate results.

Future work should focus on two fronts. First, future work should focus on performing the forecasting techniques used in this study in other systems where there is a lack of research about forecasting disease in the wild. Secondly, future work should focus on understanding how ecological and environmental drivers of disease can affect the size and shape of epidemics, for example by the termination of epidemics. Traditionally, the termination of parasite epidemics has been attributed to the depletion of susceptible hosts as a result of mortality or acquired immunity (Anderson & May, 1978; Kermack & McKendrick, 1927) and more recently due to rapid host evolution (Duffy & Sivars-Becker, 2007; Gandon et al., 2016). Most studies which investigate the ecological and environmental drivers of disease focus on changes in parasite transmission (e.g. Shocket et al., 2018), rather than the size and shape of epidemics. In turn, a better understanding of how ecological and environmental drivers affect the size and shape of epidemics will contribute to refining forecasts of disease.

## 6.5 Acknowledgements

Thanks to Dr Stuart Auld for their co-conceptualization and feedback on the writing. Thanks to Dr Stephen Thackeray for their comments on an early draft. Thanks to June Brand for their extensive data collection.

## 6.6 References


Abbasimehr, H., & Paki, R. (2021). Prediction of COVID-19 confirmed cases combining deep learning methods and Bayesian optimization. *Chaos, Solitons and Fractals*, 142. https://doi.org/10.1016/j.chaos.2020.110511

Alan Pounds, J., Bustamante, M. R., Coloma, L. A., Consuegra, J. A., Fogden, M. P. L., Foster, P. N., La Marca, E., Masters, K. L., Merino-Viteri, A., Puschendorf, R., Ron, S. R., Sánchez-Azofeifa, G. A., Still, C. J., & Young, B. E. (2006). Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature*, *439*(7073), 161–167. https://doi.org/10.1038/nature04246

Allard, R. (1998). Use of time-series analysis in infectious disease surveillance. In *Bulletin of the World Health Organization* (Vol. 76, Issue 4).

Allen, D. E., & Little, T. J. (2011). Dissecting the effect of a heterogeneous environment on the interaction between host and parasite fitness traits. *Evolutionary Ecology*, *25*(2). https://doi.org/10.1007/s10682-010-9413-2

Alonso, D., Bouma, M. J., & Pascual, M. (2011). Epidemic malaria and warmer temperatures in recent decades in an East African highland. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1712). https://doi.org/10.1098/rspb.2010.2020

Altizer, S., Dobson, A., Hosseini, P., Hudson, P., Pascual, M., & Rohani, P. (2006). Seasonality and the dynamics of infectious diseases. *Ecology Letters*, *9*(4), 467–484. https://doi.org/10.1111/j.1461-0248.2005.00879.x

Altizer, S., Hochachka, W. M., & Dhondt, A. A. (2004). Seasonal dynamics of *mycoplasmal conjunctivitis* in eastern North American house finches. *Journal of Animal Ecology*. https://doi.org/10.1111/j.0021-8790.2004.00807.x

Anderson, R. M., & May, R. M. (1978). Regulation and stability of host-parasite population interactions: I. Regulatory processes. *The Journal of Animal Ecology*, *47*(1). https://doi.org/10.2307/3933

Auld, S. K. J. R., & Brand, J. (2017). Simulated climate change, epidemic size, and host evolution across host–parasite populations. *Global Change Biology*, *23*(12), 5045–5053. https://doi.org/10.1111/gcb.13769

Aznar, M. N., Linares, F. J., Cosentino, B., Sago, A., La Sala, L., León, E., Duffy, S., & Perez, A. (2015). Prevalence and spatial distribution of bovine brucellosis in San Luis and La Pampa, Argentina. *BMC Veterinary Research*, *11*(1). https://doi.org/10.1186/s12917-015-0535-1

Baquero, O. S., Santana, L. M. R., & Chiaravalloti-Neto, F. (2018). Dengue forecasting in São Paulo city with generalized additive models, artificial neural

networks and seasonal autoregressive integrated moving average models. *PLoS ONE*, *13*(4). https://doi.org/10.1371/journal.pone.0195065

Beckley, C. S., Shaban, S., Palmer, G. H., Hudak, A. T., Noh, S. M., & Futse, J. E. (2016). Disaggregating tropical disease prevalence by climatic and vegetative zones within tropical west Africa. *PLoS ONE*, *11*(3). https://doi.org/10.1371/journal.pone.0152560

Bravo, F., Sidhu, J., Bernal, P., Bustamante, R., Condie, S., Gorton, B., Herzfeld, M., Jimenez, D., Mardones, F., Rizwi, F., & Steven, A. (2020). Hydrodynamic connectivity, water temperature, and salinity are major drivers of piscirickettsiosis prevalence and transmission among salmonid farms in Chile. *Aquaculture Environment Interactions*, *12*. https://doi.org/10.3354/aei00368

Cáceres, C. E., Hall, S. R., Duffy, M. A., Tessier, A. J., Helmle, C., & MacIntyre, S. (2006). Physical structure of lakes constrains epidemics in *Daphnia* populations. *Ecology*, *87*(6), 1438–1444. http://www.ncbi.nlm.nih.gov/pubmed/16869418

Carlsson-Granér, U., & Thrall, P. H. (2002). The spatial distribution of plant populations, disease dynamics and evolution of resistance. *Oikos*, *97*(1). https://doi.org/10.1034/j.1600-0706.2002.970110.x

Descloux, E., Mangeas, M., Menkes, C. E., Lengaigne, M., Leroy, A., Tehei, T., Guillaumot, L., Teurlai, M., Gourinat, A. C., Benzler, J., Pfannstiel, A., Grangeon, J. P., Degallier, N., & de Lamballerie, X. (2012). Climate-based models for understanding and forecasting dengue epidemics. *PLoS Neglected Tropical Diseases*, *6*(2). https://doi.org/10.1371/journal.pntd.0001470

Duffy, M. A., Hall, S. R., Tessier, A. J., & Huebner, M. (2005). Selective predators and their parasitized prey: Are epidemics in zooplankton under top-down control? *Limnology and Oceanography*. https://doi.org/10.4319/lo.2005.50.2.0412

Duffy, M. A., Ochs, J. H., Penczykowski, R. M., Civitello, D. J., Klausmeier, C. A., & Hall, S. R. (2012). Ecological context influences epidemic size and parasite-driven evolution. *Science*, *335*(6076), 1636–1638. https://doi.org/10.1126/science.1215429

Duffy, M. A., & Sivars-Becker, L. (2007). Rapid evolution and ecological host-parasite dynamics. *Ecology Letters*, *10*(1), 44–53.

Ebert, D. (2005). *Ecology, Epidemiology, and Evolution of Parasitism in* Daphnia (S. Paliwoda, Ed.).

Ebert, D., Zschokke-Rohringer, C. D., & Carius, H. J. (1998). Within–and between–population variation for resistance of *Daphnia magna* to the bacterial

endoparasite *Pasteuria ramosa*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*(1410), 2127–2134. https://doi.org/10.1098/rspb.1998.0549

El-Sayed, A., & Kamel, M. (2020). Climatic changes and their role in emergence and re-emergence of diseases. In *Environmental Science and Pollution Research* (Vol. 27, Issue 18). https://doi.org/10.1007/s11356-020-08896-w

Ericson, L., Burdon, J. J., & Müller, W. J. (1999). Spatial and temporal dynamics of epidemics of the rust fungus *Uromyces valerianae* on populations of its host Valeriana salina. *Journal of Ecology*, *87*(4). https://doi.org/10.1046/j.1365-2745.1999.00384.x

Gandon, S., Day, T., Metcalf, C. J. E., & Grenfell, B. T. (2016). Forecasting epidemiological and evolutionary dynamics of infectious diseases. *Trends in Ecology & Evolution*, *31*(10), 776–788. https://doi.org/10.1016/j.tree.2016.07.010

Gao, H. W., Wang, L. P., Liang, S., Liu, Y. X., Tong, S. L., Wang, J. J., Li, Y. P., Wang, X. F., Yang, H., Ma, J. Q., Fang, L. Q., & Cao, W. C. (2012). Change in rainfall drives malaria re-emergence in Anhui province, China. *PLoS ONE*, *7*(8). https://doi.org/10.1371/journal.pone.0043686

Groner, M. L., Eisenlord, M. E., Yoshioka, R. M., Fiorenza, E. A., Dawkins, P. D., Graham, O. J., Winningham, M., Vompe, A., Rivlin, N. D., Yang, B., Burge, C. A., Rappazzo, B., Gomes, C. P., & Drew Harvell, C. (2021). Warming sea surface temperatures fuel summer epidemics of eelgrass wasting disease. *Marine Ecology Progress Series*, *679*. https://doi.org/10.3354/meps13902

Groner, M. L., Shields, J. D., Landers, D. F., Swenarton, J., & Hoenig, J. M. (2018). Rising temperatures, molting phenology, and epizootic shell disease in the American lobster. *American Naturalist*, *192*(5). https://doi.org/10.1086/699478

Hall, S. R., Becker, C. R., Simonis, J. L., Duffy, M. A., Tessier, A. J., & Cáceres, C. E. (2009). Friendly competition: evidence for a dilution effect among competitors in a planktonic host–parasite system. *Ecology*, *90*(3), 791–801. https://doi.org/10.1890/08-0838.1

Hall, S. R., Knight, C. J., Becker, C. R., Duffy, M. A., Tessier, A. J., & Cáceres, C. E. (2009). Quality matters: Resource quality for hosts and the timing of epidemics. *Ecology Letters*, *12*(2), 118–128. https://doi.org/10.1111/j.1461-0248.2008.01264.x

Helfenstein, U. (1991). The use of transfer function models, intervention analysis and related time series methods in epidemiology. *International Journal of Epidemiology*, *20*(3). https://doi.org/10.1093/ije/20.3.808

Hii, Y. L., Zhu, H., Ng, N., Ng, L. C., & Rocklöv, J. (2012). Forecast of Dengue incidence using temperature and rainfall. *PLoS Neglected Tropical Diseases*, *6*(11). https://doi.org/10.1371/journal.pntd.0001908

Hu, W., Tong, S., Mengersen, K., & Oldenburg, B. (2006). Rainfall, mosquito density and the transmission of Ross River virus: A time-series forecasting model. *Ecological Modelling*, *196*(3–4), 505–514. https://doi.org/10.1016/j.ecolmodel.2006.02.028

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: principles and practice*. OTexts.

Jennelle, C. S., Cooch, E. G., Conroy, M. J., & Senar, J. C. (2007). State-specific detection probabilities and disease prevalence. *Ecological Applications*, *17*(1), 154–167. https://doi.org/10.1890/1051-0761(2007)017[0154:SDPADP]2.0.CO;2

Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *115*(772), 700–721. https://doi.org/10.1098/rspa.1927.0118

Klüttgen, B., Dülmer, U., Engels, M., & Ratte, H. T. (1994). ADaM, an artificial freshwater for the culture of zooplankton. *Water Research*, *28*(3), 743–746. https://doi.org/10.1016/0043-1354(94)90157-0

Krauer, F., Viljugrein, H., & Dean, K. R. (2021). The influence of temperature on the seasonality of historical plague outbreaks. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1954). https://doi.org/10.1098/rspb.2020.2725

Montano, S., Giorgi, A., Monti, M., Seveso, D., & Galli, P. (2016). Spatial variability in distribution and prevalence of skeletal eroding band and brown band disease in Faafu Atoll, Maldives. *Biodiversity and Conservation*, *25*(9). https://doi.org/10.1007/s10531-016-1145-3

Paplauskas, S., Brand, J., & Auld, S. K. J. R. (2021). Ecology directs host–parasite coevolutionary trajectories across *Daphnia*–microparasite populations. *Nature Ecology & Evolution*, *5*(4), 480–486. https://doi.org/10.1038/s41559-021-01390-7

Penczykowski, R. M., Laine, A.-L., & Koskella, B. (2016). Understanding the ecology and evolution of host-parasite interactions across scales. *Evolutionary Applications*, *9*(1), 37–52. https://doi.org/10.1111/eva.12294

Pepin, K. M., Eisen, R. J., Mead, P. S., Piesman, J., Fish, D., Hoen, A. G., Barbour, A. G., Hamer, S., & Diuk-Wasser, M. A. (2012). Geographic variation in the

relationship between human lyme disease incidence and density of infected host-seeking *Ixodes scapularis* nymphs in the Eastern United States. *American Journal of Tropical Medicine and Hygiene*, *86*(6). https://doi.org/10.4269/ajtmh.2012.11-0630

Perone, G. (2022). Using the SARIMA model to forecast the fourth global wave of cumulative deaths from COVID-19: Evidence from 12 hard-hit big countries. *Econometrics*, *10*(2). https://doi.org/10.3390/econometrics10020018

Ruiz-Moreno, D., Willis, B. L., Page, A. C., Weil, E., Cróquer, A., Vargas-Angel, B., Jordan-Garza, A. G., Jordán-Dahlgren, E., Raymundo, L., & Harvell, C. D. (2012). Global coral disease prevalence associated with sea temperature anomalies and local factors. *Diseases of Aquatic Organisms*, *100*(3). https://doi.org/10.3354/dao02488

Schaaf, C. J., Kelson, S. J., Nusslé, S. C., & Carlson, S. M. (2017). Black spot infection in juvenile steelhead trout increases with stream temperature in northern California. *Environmental Biology of Fishes*, *100*(6). https://doi.org/10.1007/s10641-017-0599-9

Shocket, M. S., Vergara, D., Sickbert, A. J., Walsman, J. M., Strauss, A. T., Hite, J. L., Duffy, M. A., Cáceres, C. E., & Hall, S. R. (2018). Parasite rearing and infection temperatures jointly influence disease transmission and shape seasonality of epidemics. *Ecology*, *99*(9), 1975–1987. https://doi.org/10.1002/ecy.2430

Susi, H., Thrall, P. H., Barrett, L. G., & Burdon, J. J. (2017). Local demographic and epidemiological patterns in the *Linum marginale - Melampsora lini* association: a multi-year study. *Journal of Ecology*, *105*(5), 1399–1412. https://doi.org/10.1111/1365-2745.12740

Swinford, J. L., & Anderson, J. D. (2021). Prevalence of Black Gill (*Hyalophysa lynni*) in White Shrimp *Litopenaeus setiferus* and Brown Shrimp *Farfantepenaeus aztecus* along the Texas Gulf Coast. *Marine and Coastal Fisheries*, *13*(3). https://doi.org/10.1002/mcf2.10153

Thoirain, B., Husson, C., & Marçais, B. (2007). Risk factors for the Phytophthora-induced decline of alder in northeastern France. *Phytopathology*, *97*(1). https://doi.org/10.1094/PHYTO-97-0099

Thrall, P. H., Laine, A.-L., Ravensdale, M., Nemri, A., Dodds, P. N., Barrett, L. G., & Burdon, J. J. (2012). Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. *Ecology Letters*, *15*(5), 425–435. https://doi.org/10.1111/j.1461-0248.2012.01749.x

Vale, P. F., Stjernman, M., & Little, T. J. (2008). Temperature-dependent costs of parasitism and maintenance of polymorphism under genotype-by-environment interactions. *Journal of Evolutionary Biology*. https://doi.org/10.1111/j.1420-9101.2008.01555.x

Vergara, D., Lively, C. M., King, K. C., & Jokela, J. (2013). The geographic mosaic of sex and infection in lake populations of a New Zealand snail at multiple spatial scales. *American Naturalist*, *182*(4). https://doi.org/10.1086/671996

Zell, R. (2004). Global climate change and the emergence/re-emergence of infectious diseases. *International Journal of Medical Microbiology, Supplement*, *293*(37). https://doi.org/10.1016/S1433-1128(04)80005-6

# 7. Thesis discussion

In the following discussion, the results of each research chapter (3-6) are contextualised with the current state of knowledge and the original questions or hypotheses outlined previously (see chapter one; Thesis introduction). How the results of each chapter addressed some of the gaps identified in the proposed Disease Cycle model (see chapter two; Literature Review) are presented in a final integrated discussion together with recommendations for the direction of future research.

## 7.1 Chapter three: Ecology directs host–parasite coevolutionary trajectories across *Daphnia*–microparasite populations

In chapter three (published by 2021 in *Nat. Eco. & Evo.*), I found the precise level of variation in host-parasite coevolutionary trajectories in wild populations that could be explained by environmental factors, such as temperature and food availability, in a 'world-first' study of its kind. In agreement with our original hypothesis, this means that there is some level of repeatability in host-parasite coevolutionary trajectories in the wild, despite their ecological 'noise'. Another exciting result that I found was that the environment influenced coevolution indirectly through changes in host resistance, rather than parasite infectivity, in the replicate pond populations. This supports my original hypothesis that endoparasites would be less affected by external abiotic factors compared to their hosts (alt. ectoparasites (Cardon et al., 2011; Mahmud et al., 2017) because they are insulated from the wider environment.

Together the results of this study demonstrate that the ability of host-parasite (*potential*) coevolution measured in a laboratory-based environment to translate into (*observed* or *realized*) natural (i.e. 'real-world') environments depends on the strength of host and parasite-mediated selection relative to other biotic and abiotic factors. Regarding the first link in the Disease Cycle model, I confirm that epidemics can exert strong parasite-mediated selection, which can interact with other ecological (i.e. 'environmental') factors, to drive host-parasite coevolutionary asymmetry.

As a closing remark, if there was one criticism of the experiment, I would have liked to publish the analysis of not only the direction, but also the magnitude of change and the distance between phenotypic endpoints (Adams & Collyer, 2009; Bolnick et

al., 2018), to demonstrate how these other aspects of phenotypic trajectories are dependent (or not) on an interaction with the environment (Fig. S7.1).

## 7.2 Chapter four: The effect of host population genetic diversity on metrics of parasite infection success

In chapter four, I investigated the third link in the Disease Cycle, which focused on the effect of host population genetic diversity on epidemic size. By re-analysing the meta-analytical data collected by two previous studies (Ekroth et al., 2019; Gibson & Nguyen, 2021), I found that the effect of host population genetic diversity on metrics of parasite infection success was not as straightforward as previous studies would have us believe. Although previous studies have shown a 'conventional' effect of host population genetic diversity in limiting the metrics of mean parasite infection success, I found that this is an over-generalisation. In actual fact, host population genetic diversity limits metrics of mean infection success for specialist, but not generalist, parasites (Paplauskas et al., 2024). This challenges conventional theory (King & Lively, 2012) and has large implications for how genetic diversity is managed in wild host communities. For example, we ought to be prioritising host populations with low genetic diversity that are susceptible to specialist parasites for management of genetic diversity (Meuwissen et al., 2020), or genetic restoration (Whiteley et al., 2015).

I also found support for my proposed diversity uncertainty model, which predicts a complex interaction between the effect of host population genetic diversity on both the mean and variability in metrics of parasite infection success with both parasite host range and parasite population genetic diversity. This result further challenges conventional theory, which is focused on the relationship between host population genetic diversity and epidemic size (King & Lively, 2012), by identifying how the variability in metrics of parasite success e.g. epidemic size, are determined by a combination of host and parasite disease traits, such as parasite host range and parasite population genetic diversity. Perhaps a measure of the variability in parasite success is better suited to investigating the relationship between host population genetic diversity and epidemic size, rather than a measure of mean parasite success, as the definition of an epidemic is a relatively large increase in the proportion of infected individuals over time above the threshold for endemic-level disease (Dicker, 2006). In other words, it constitutes an unusually large amount of change (i.e. *variation*) in the proportion of diseased hosts within a population. In

comparison, the concept of mean parasite success may be more relevant to small, but significant, shifts in the endemic level of disease.

Some of the limitations of this study, shared by previous research (Ekroth et al., 2019; Gibson & Nguyen, 2021), is the difficulty in communicating what the difference is between a high versus low genetic diversity population. This can refer to studies which have (i) inbred lineages to create a comparison between inbred and outbred populations (Baer & Schmid-Hempel, 2001), (ii) used a suite of wildtype genotypes for controlled experiments with either low genetic diversity or high genetic diversity (Florian Altermatt & Ebert, 2008), (iii) sampled organisms from the wild from populations that have been characterised as having different levels of genetic diversity (Tarpy & Seeley, 2006) or (iv) quantified a continuous measure of genetic diversity, such as heterozygosity (Ellison et al., 2011). The inconsistency between these different metrics of host population genetic diversity, combined with a lack of understanding as to what 'parasite infection success' actually refers to, limits the ability to quantify exactly how host population genetic diversity leads to some kind of tractable change in future epidemic size (third link in the Disease Cycle). Although a previous meta-analysis went some way to quantifying the effect of host population genetic diversity on the spread of disease, by showing the reduction in metrics of mean parasite infection success between host populations with high versus low genetic diversity was approximately 20% for non-crop hosts and 50% for crop hosts (Gibson & Nguyen, 2021), there is potential confusion over what the distinction between high and low population genetic diversity is.

This lack of a quantitative estimate of the effect of host population genetic diversity on the spread of disease has been studied previously as part of the concept of a 'diversity threshold' (Lively, 2010). This research was motivated by the idea that parasites might be eliminated by increasing host genetic diversity above a certain level. By simulating hosts with two resistance loci and up to three alleles (total of nine genotypes), the author found that, despite the positive effect of increasing population size on $R_0$, doubling host population size and increasing the number of genotypes by four times decreases $R_0$ below zero. I proposed one possible solution to this problem would be for future studies to utilize a standardized measure of epidemic size, such as integrated disease prevalence (which is the proportion of infected individuals within in a population over time).

Another potentially confounding factor, which may limit the ability of this study to identify the real relationship between the effect of host population genetic diversity on metrics of infection parasite success, is what the shape is of the host and parasite genetic diversity distributions. Since the ability of host population genetic diversity to affect metrics of parasite infection success relies on matching host and parasite genotypes (Schmid-Hempel & Ebert, 2003), a key question becomes are these distributions symmetrical? For example, host populations with a low level of genetic diversity may have undergone balancing selection, and suffered from a loss of extreme phenotypes. Alternatively, host populations may have experienced genetic diversity loss through direction selection, leading to a loss of one group of extreme genotypes, but not the other. This is implicated with the history of antagonistic selection between the host and parasite populations (i.e. selective sweeps versus negative frequency dependent selection that either erode or maintain genetic diversity over time (see Fig. 2.4)) and the reason for genetic diversity loss e.g. hunting versus inbreeding (see chapter two, 2.4). Finally, the particular model of infection genetics that describes a given host-parasite system (i.e. matching-alleles vs gene-for-gene (Agrawal & Lively, 2002)) and the corresponding infection specificity (Schmid-Hempel & Ebert, 2003) will influence how important the symmetry in host and parasite genetic diversity distributions is. If there is a high level of specificity for infection (i.e. matching-alleles) then we would expect symmetry to be important, whereas if there is a low level of specificity for infection (i.e. gene-for-gene) then we would not.

One parting comment on the potential limitation of using the log coefficient of variation ratio (lnCVR) over an alternative effect size (such as the log variability ratio, lnVR) is that it is not possible to model a mean-variance relationship between the standardized mean difference (SMD) and lnCVR (Supplementary figure S7.2). This is because the calculation of both SMD and lnCVR involves using the mean of each control (high diversity) and treatment (low diversity) within a comparison. To the best our knowledge, this issue has not been encountered before in previous research. To enable a direct comparison of the effect of host population genetic diversity on the mean and variability in metrics of parasite infection success between host populations with high versus low genetic diversity, we would require a much larger amount of data on the specific sampling distribution which most accurately reflect the true probability term for a positive occurrence in each metric of parasite success (transmission, parasite load and virulence). For example, if we were interested in simulating epidemic size (proportion of infected hosts) as a metric of parasite

infection success, this would require an estimate for the binomial probability term used to define whether any given susceptible host is infected within a population. Once this data is available, we would then be able to test the extent to which any regression model of true (observed) effect sizes deviates from a simulated dataset. Producing such a detailed background dataset of sampling distributions would require an enormous amount of empirical work in different host-parasite systems.

## 7.3 Chapter five: The ability of non-locally adapted hosts to outcompete resident hosts in wild populations

In chapter five, I found the ability of non-locally adapted hosts to outcompete resident host genotypes under parasite exposure. This could affect future epidemic size in host populations in various ways. For example, the inability of residents hosts to withstand migrant invasion means that gene flow, and the accompanying overall level of host population genetic diversity, could potentially increase and lead to small average epidemic size in the future (Paplauskas et al., 2024). Alternatively, the susceptibility of residents to invasion by migrant hosts may limit future mean epidemic size increasing the turnover rate of local populations. Since fundamental local adaptation theory predicts that parasites will be less well-adapted to non-local hosts, as local hosts are often trapped on the losing side of a cycle of antagonistic coevolution (Gandon, 2002), newly founded migrant host populations could be more resistant to disease.

Similarly, the lack of host local adaptation to the abiotic environment observed in the experimental populations suggests that local populations are also at risk of displacement by migrants undergoing range shifts in response to climate change (Price et al., 2019). Although in comparison to interspecific competition, rather than inter-population competition between local and migrant conspecifics, in exploratory experiments varying the strength of adaptation and competition, one study found that competition actually reduced the level of population genetic diversity in competing species, leading to a reduction in the rate of range change (Bocedi et al., 2013). However, in accordance with the results of my study, weak selection on local adaptation resulted in the tracking of cooler-adapted phenotypes away from an expanding range margin and therefore a loss of warmer-adapted phenotypes.

As alluded to in the discussion section of chapter five, there might be an opportunity for future research to study patterns of parasite local adaptation in the replicate pond populations (Supplementary S7.3). In comparison to my previous research, which focused on the extent to which variation in pond environments could explain variation in the direction of host-parasite coevolutionary trajectories (Paplauskas et al., 2021), a test of parasite local adaptation would focus on whether parasites from different ponds were better adapted to local versus away hosts (Gandon & Nuismer, 2009). Also, this would offer the opportunity to examine the environmental drivers of parasite local adaptation patterns, which is a common goal of local adaptation experiments (Blanquart et al., 2013; Kawecki & Ebert, 2004).

## 7.4 Chapter six: Borrowing data from other populations to forecast epidemics

In chapter six, I found that data from replicate *Daphnia* host populations could be used to improve forecast accuracy relative to using a single population. Specifically, other than the regression models that used predicted temperature values to forecast disease prevalence, where the mean accuracy between single and multi-population models was equivalent, the mean accuracy of disease prevalence forecasts based on both the mean and temperature weighted mean data was significantly higher than the mean accuracy of those same models trained on data from a single population. This was consistent with the benchmark models, which showed similar accuracy to the other more complex forecasting approaches, and for environmentally (temperature) weighted versus standard means, which were also equivalent. Therefore, despite the lack of forecast accuracy gained by using the similarity in environmental conditions between replicate populations, this study demonstrated how data from separate populations can be used to predict future epidemic size, rather than relying on several years of historical data from a single population.

Overall, a major application of this work could be in a new method that can be used for quickly predicting the size of future epidemics of emerging or novel wildlife diseases. For a great many disease systems, forecasting future epidemic size relies on extensive historical data from a single population. However, since there is no such historical data for emerging diseases, it might instead be possible to forecast the magnitude of a future emerging disease epidemic by establishing a group of replicate populations across space, similar to the method used in this study, as a trade-off against time. However, the extent to which this method can be employed in a non-model system, such as vertebrates or plant host – parasite systems, that may require a large amount of habitat space, is not well-understood. Similar to predicting epidemic size for emerging or novel disease, although the epidemic forecasting models trained on my temperature-weighted mean were not significantly better than those using a standard mean, the same method could theoretically be applied to predict epidemic size in response to climate change.

As previously mentioned, although I found that weighting the training datasets by their (temperature) similarity to the focal population, one possible direction for future work is to measure precisely how much additional accuracy is (or is not) gained by adding further variables. For example, it would be useful to know how much a

multivariate measure of environment similarity (e.g. using Mahalanobis distance in the sense of (Paplauskas et al., 2021) improves forecast accuracy relative to a more limited dataset with only a single variable, and whether it is worth all of the extra effort in collecting the data in the first place. This idea of how much data a model really needs in order to avoid a diminishing return, is a fundamental unanswered question in statistics generally (McCrea et al., 2023; Simmonds et al., 2020).

## 7.5 Impact of my proposed Disease Cycle

The primary objective of this thesis was to evaluate the current support for the Disease Cycle model proposed in chapter one, intended to provide a theoretical framework to link the size of past and future epidemics of disease (Fig. 1.1), and address any knowledge gaps that would require filling to complete the cycle.

Following my review of the current literature in chapter two, I identified a few areas for future research. This included (i) measuring the extent to which epidemics drive realised coevolutionary change (versus its potential) in 'real-world' environments (first link in the Disease Cycle), and (ii) the extent to which host population genetic diversity limits the mean and affects the variability in future epidemic size (third link in the Disease Cycle). In chapter three, using parasite infectivity data and change in host genotype frequencies within replicate host-parasite populations, I was able to demonstrate the precise amount of variation in coevolutionary trajectories that was driven by a mixture of biotic and abiotic environmental conditions (Paplauskas et al., 2021). In chapter four, by re-analysing the meta-analytical data collected by two previous studies (Ekroth et al., 2019; Gibson & Nguyen, 2021), I found that the effect of host population genetic diversity on the mean and variability metrics of parasite infection success depended on a combination of parasite specificity and parasite population genetic diversity (Paplauskas et al., 2024). Specifically, I found that (a) host population genetic diversity limited the metrics of mean infection success for specialist, but not generalist, parasites and (b) there was support for a diversity uncertainty model that predicts a complex interaction between the effect of host population genetic diversity on the variability in metrics of parasite infection success with both parasite host range and parasite population genetic diversity.

These findings from chapters three and four have helped to fill major gaps in the Disease Cycle. In comparison, the findings from chapters five and six have addressed more specific topics within the overall Disease Cycle perspective. In

chapter five, I found the ability of non-locally adapted hosts to outcompete resident hosts in the wild. This has implications for the extent of gene flow between host subpopulations, which may increase in response to successful migrant competition, and therefore decrease the size and severity of future disease outbreaks by increasing the level of host population genetic diversity (Paplauskas et al., 2024). In chapter six, I borrowed data from other populations to forecast epidemic size and found that it increases forecast accuracy relative to single population models. This was an actual demonstration of how the size of past epidemics can be used to predict the size of future epidemics.

In realisation of a somewhat narrow focus of my preceding work, there is a substantial knowledge gap which remains in the second link in the Disease Cycle, between the effect of coevolution on the amount of host and parasite population-level genetic diversity. In particular, how the tempo and mode of coevolution affects the maintenance of host and parasite population genetic diversity is still limited to model systems, such as *Daphnia* host – parasite systems (Bento et al., 2017a). More studies in non-model organisms, such as long-lived vertebrates or plants, for which the genetic basis for infection is less well-understood (Brockhurst & Koskella, 2013; Schmid-Hempel & Ebert, 2003), are required to confirm the generality of current theory. In addition, more long-term studies of coevolution are required to understand the extent to which coevolutionary dynamics are maintained over time (but see (Soubeyrand et al., 2009; Thrall et al., 2012; Susi and Laine, 2015; Ericson, Müller and Burdon, 2017; but see Dewald-Wang et al., 2022)).

Moving forward, I propose that future studies utilize standardized measures of epidemic size, such as integrated disease prevalence (which is the proportion of infected in a population over time, see chapter two Fig. 2.1) to understand how it shapes patterns of host and parasite population genetic diversity. Ideally, future research should focus on developing a true measure of epidemic severity which combines size (transmission) and virulence. This would further help to quantify the realised versus potential ability of epidemics to drive evolutionary change.

Another potential direction for future research is to build deterministic forecast models for future epidemic size which account for the strength of host-parasite selection, the tempo and mode of coevolution and population-level genetic diversity (the three major axes of the Disease Cycle, Fig. 1.1). These sorts of processes are

not part of current epidemic modelling research, but could be useful for disease systems where there is strong host and parasite-mediated reciprocal selection. This would be particularly relevant for invertebrates and other non-vertebrate systems, where there is innate (Little et al., 2003) versus acquired immunity (Babayan et al., 2011). However, this approach to epidemic modelling would perhaps require a require a large amount of data that is currently unavailable, such as (i) a general estimate of the potential for epidemics to drive host-parasite evolutionary change, (ii) system-specific rates of host evolution of resistance and parasite evolution of infectivity and (iii) a measure of host (and parasite) population genetic diversity, to parameterise these Disease Cycle processes within a mathematical model. In addition, maybe prediction is not as useful as prevention in some disease systems. For example, for human hosts, we may not require a high level of forecasting accuracy because after an initial outbreak occurs, the priority quickly shifts to intervention (i.e. development of a vaccine, such as for Covid-19 (Moghadas et al., 2021), and for other (vertebrate) host systems, acquired immunity has the ability to break the cycle of disease by disrupting the link between the size of past and future epidemics.

Overall, the concept of a Disease Cycle offers a new, coevolutionary perspective on epidemics in a range of host-parasite systems. This includes microparasites, whose infections of host populations are characterised by epidemics (Hudson et al., 2002), and host species with innate rather than acquired immunity, including plants, fungi, prokaryotes and invertebrates (Janeway et al., 2001). The intended use of the model is to provide a theoretical link between the size and severity of past and future epidemics within a broad context of environmental change, that can be applied to a wide range of host-parasite systems, to better understand the underlying coevolutionary processes that cause epidemic size to vary across time and space. This theoretical Disease Cycle model can be further evaluated and reinforced by future empirical studies.

## 7.6 References

Adams, D. C., & Collyer, M. L. (2009). A general framework for the analysis of phenotypic trajectories in evolutionary studies. *Evolution*, 63(5), 1143–1154. https://doi.org/10.1111/j.1558-5646.2009.00649.x

Agrawal, A., & Lively, C. M. (2002). Infection genetics: Gene-for-gene versus matching-alleles models and all points in between. *Evolutionary Ecology Research*.

Babayan, S. A., Allen, J. E., Bradley, J. E., Geuking, M. B., Graham, A. L., Grencis, R. K., Kaufman, J., McCoy, K. D., Paterson, S., Smith, K. G. C., Turnbaugh, P. J., Viney, M. E., Maizels, R. M., & Pedersen, A. B. (2011). Wild immunology: converging on the real world. *Annals of the New York Academy of Sciences*, *1236*(1), 17–29. https://doi.org/10.1111/j.1749-6632.2011.06251.x

Baer, B., & Schmid-Hempel, P. (2001). Unexpected consequences of polyandry for parasitism and fitness in the bumblebee, Bombus terrestris. *Evolution*. https://doi.org/10.1111/j.0014-3820.2001.tb00683.x

Bento, G., Routtu, J., Fields, P. D., Bourgeois, Y., Du Pasquier, L., & Ebert, D. (2017). The genetic basis of resistance and matching-allele interactions of a host-parasite system: The *Daphnia magna-Pasteuria ramosa* model. *PLOS Genetics*, *13*(2), e1006596. https://doi.org/10.1371/journal.pgen.1006596

Blanquart, F., Kaltz, O., Nuismer, S. L., & Gandon, S. (2013). A practical guide to measuring local adaptation. *Ecology Letters*, *16*(9), 1195–1205. https://doi.org/10.1111/ele.12150

Bocedi, G., Atkins, K. E., Liao, J., Henry, R. C., Travis, J. M. J., & Hellmann, J. J. (2013). Effects of local adaptation and interspecific competition on species' responses to climate change. *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/nyas.12211

Bolnick, D. I., Barrett, R. D. H., Oke, K. B., Rennison, D. J., & Stuart, Y. E. (2018). (Non)Parallel Evolution. *Annual Review of Ecology, Evolution, and Systematics*, *49*(1), 303–330. https://doi.org/10.1146/annurev-ecolsys-110617-062240

Brockhurst, M. A., & Koskella, B. (2013). Experimental coevolution of species interactions. *Trends in Ecology & Evolution*, *28*(6), 367–375. https://doi.org/10.1016/J.TREE.2013.02.009

Cardon, M., Loot, G., Grenouillet, G., & Blanchet, S. (2011). Host characteristics and environmental factors differentially drive the burden and pathogenicity of an ectoparasite: A multilevel causal analysis. *Journal of Animal Ecology*, *80*(3), 657–667. https://doi.org/10.1111/j.1365-2656.2011.01804.x

Dicker, R. C. (2006). Principles of Epidemiology in Public Health Practice. *Cdc*.

Ekroth, A. K. E., Rafaluk-Mohr, C., & King, K. C. (2019). Host genetic diversity limits parasite success beyond agricultural systems: a meta-analysis. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1911), 20191811. https://doi.org/10.1098/rspb.2019.1811

Ellison, A., Cable, J., & Consuegra, S. (2011). Best of both worlds? Association between outcrossing and parasite loads in a selfing fish. *Evolution*, *65*(10), 3021–3026. https://doi.org/10.1111/j.1558-5646.2011.01354.x

Altermatt, F., & Ebert, D. (2008). Genetic diversity of *Daphnia magna* populations enhances resistance to parasites. *Ecology Letters*. https://doi.org/10.1111/j.1461-0248.2008.01203.x

Gandon, S. (2002). Local adaptation and the geometry of host-parasite coevolution. *Ecology Letters*, *5*(2), 246–256. https://doi.org/10.1046/j.1461-0248.2002.00305.x

Gandon, S., & Nuismer, S. L. (2009). Interactions between genetic drift, gene flow, and selection mosaics drive parasite local adaptation. *American Naturalist*, *173*(2), 212–224. https://doi.org/10.1086/593706

Gibson, A. K., & Nguyen, A. E. (2020). Does genetic diversity protect host populations from parasites? A meta-analysis across natural and agricultural systems. *Evolution Letters*, *5*(1). https://doi.org/10.1002/evl3.206

Hudson, P. J., Rizzoli, A., Grenfell, B. T., Heesterbeek, H., & Dobson, A. P. (2002). *The Ecology of Wildlife Diseases*. Oxford University Press, Oxford. https://doi.org/10.1093/oso/9780198506201.001.0001

Janeway, C. A., Travers, P. J., Walport, M., & Shlomchik, M. J. (2001). *Immunobiology* (Fifth edit). Garland Science.

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*(12), 1225–1241. https://doi.org/10.1111/j.1461-0248.2004.00684.x

King, K. C., & Lively, C. M. (2012). Does genetic diversity limit disease spread in natural host populations? *Heredity*, *109*(4), 199–203. https://doi.org/10.1038/hdy.2012.33

Little, T. J., O'Connor, B., Colegrave, N., Watt, K., & Read, A. F. (2003). Maternal transfer of strain-specific immunity in an invertebrate. *Current Biology*, *13*(6), 489–492. https://doi.org/10.1016/S0960-9822(03)00163-5

Lively, C. M. (2010). The effect of host genetic diversity on disease spread. *The American Naturalist*, *175*(6), E149–E152. https://doi.org/10.1086/652430

Mahmud, M. A., Bradley, J. E., & MacColl, A. D. C. (2017). Abiotic environmental variation drives virulence evolution in a fish host–parasite geographic mosaic. *Functional Ecology*, *31*(11), 2138–2146. https://doi.org/10.1111/1365-2435.12921

McCrea, R., King, R., Graham, L., & Börger, L. (2023). Realising the promise of large data and complex models. *Methods in Ecology and Evolution*, *14*(1), 4–11. https://doi.org/10.1111/2041-210X.14050

Meuwissen, T. H. E., Sonesson, A. K., Gebregiwergis, G., & Woolliams, J. A. (2020). Management of genetic diversity in the era of genomics. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.00880

Moghadas, S. M., Vilches, T. N., Zhang, K., Wells, C. R., Shoukat, A., Singer, B. H., Meyers, L. A., Neuzil, K. M., Langley, J. M., Fitzpatrick, M. C., & Galvani, A. P. (2021). The impact of vaccination on Coronavirus Disease 2019 (COVID-19) outbreaks in the United States. *Clinical Infectious Diseases*, *73*(12), 2257–2264. https://doi.org/10.1093/cid/ciab079

Paplauskas, S., Brand, J., & Auld, S. K. J. R. (2021). Ecology directs host–parasite coevolutionary trajectories across *Daphnia*–microparasite populations. *Nature Ecology and Evolution*, *5*(4), 480–486. https://doi.org/10.1038/s41559-021-01390-7

Paplauskas, S., Duthie, B., & Tinsley, M. C. (2024). The effect of host population genetic diversity on variation in metrics of parasite success. *BioRxiv*. https://doi.org/https://doi.org/10.1101/2024.05.28.596150

Price, S. J., Leung, W. T. M., Owen, C. J., Puschendorf, R., Sergeant, C., Cunningham, A. A., Balloux, F., Garner, T. W. J., & Nichols, R. A. (2019). Effects of historic and projected climate change on the range and impacts of an emerging wildlife disease. *Global Change Biology*. https://doi.org/10.1111/gcb.14651

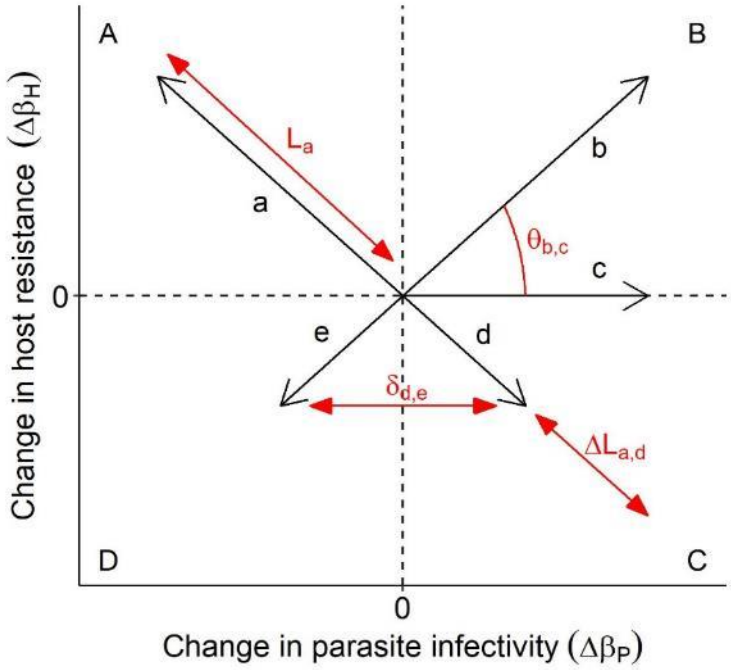
Schmid-Hempel, P., & Ebert, D. (2003). On the evolutionary ecology of specific immune defence. *Trends in Ecology & Evolution*, *18*(1), 27–32. https://doi.org/10.1016/S0169-5347(02)00013-7

Simmonds, E. G., Jarvis, S. G., Henrys, P. A., Isaac, N. J. B., & O'Hara, R. B. (2020). Is more data always better? A simulation study of benefits and limitations of integrated distribution models. *Ecography*, *43*(10), 1413–1422. https://doi.org/10.1111/ecog.05146

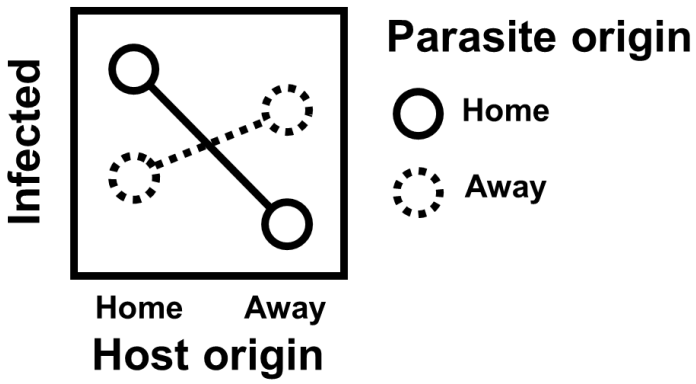
Tarpy, D. R., & Seeley, T. D. (2006). Lower disease infections in honeybee (*Apis mellifera*) colonies headed by polyandrous vs monandrous queens. *Naturwissenschaften*, *93*(4). https://doi.org/10.1007/s00114-006-0091-4

Whiteley, A. R., Fitzpatrick, S. W., Funk, W. C., & Tallmon, D. A. (2015). Genetic rescue to the rescue. *Trends in Ecology & Evolution*, *30*(1), 42–49. https://doi.org/10.1016/j.tree.2014.10.009

## 7.7 Supplementary information

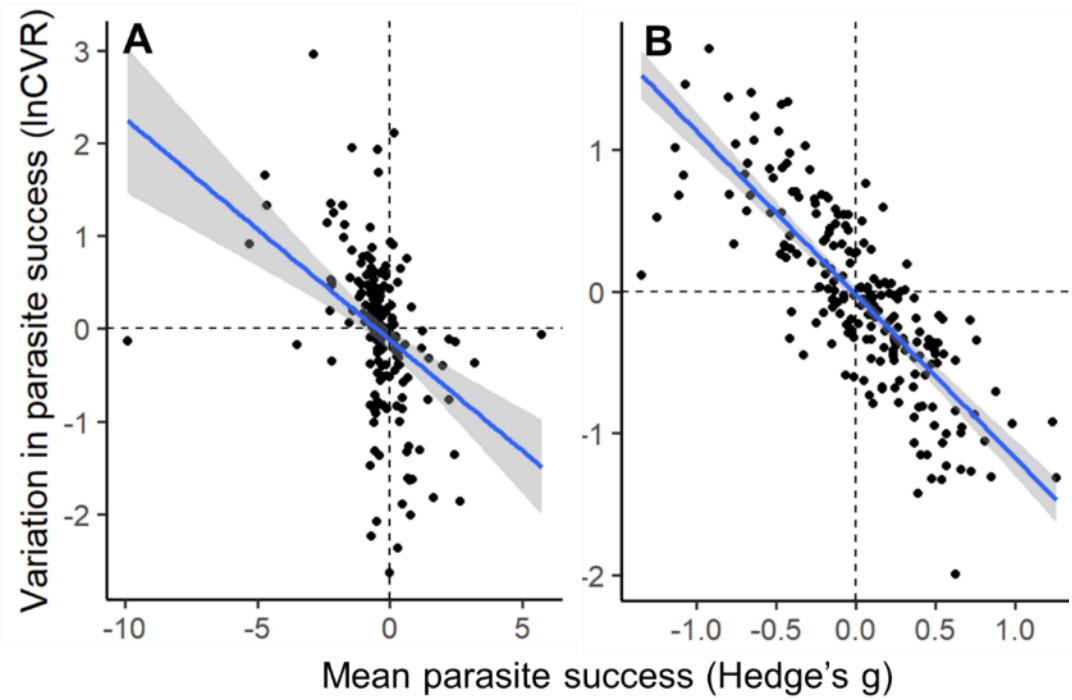



Supplementary figure S7.1. Demonstrative host-parasite coevolutionary trajectories and their measurement in different ways (L, ΔL, θ and δ). Each quarter of the plotting area shows a different (co)evolutionary relationship, where negative values represent a loss of either host resistance (y-axis) or parasite infectivity (x-axis), and positive values of change in host or parasite disease traits the opposite (uppercase letters A-D). No change in either host or parasite (co)evolutionary trajectories is indicated at zero (dashed line in the middle of each antagonist's axis). The vectors (phenotypic, so-called trajectories), correspond to the solid arrows. They share a common origin (vector tails) and divergent end positions (open arrows). There are five cases of host-parasite (co)evolution (lowercase a-e) and four different units for measuring phenotypic trajectories (annotations in red). The magnitude of evolution is represented by L, vector length (e.g. $L_a$ represents the magnitude of evolution trajectory a). The difference in L (ΔL) represents the difference in the length of phenotypic trajectories, such as a and d ($\Delta L_{a,d}$). The relative contribution of each axis to the combined host-parasite vector is shown by θ, which is the difference in the direction (or angle) between two vectors, such as a and b ($\theta_{b,c}$). The degree of dissimilarity between evolved phenotypes (open arrowheads) is indicated by the distance between vector endpoints, such between endpoints of vectors d and e ($\delta_{d,e}$). Adapted from Stuart et al. (2017).




Supplementary figure S7.2. Pairwise comparison of the mean proportion of infected hosts as part of a hypothetical local adaptation experiment in the experimental pond populations. The home and away origin of the parasite is indicated by the solid and dashed lines respectively.

5266




Supplementary figure S7.3. The mean-variance relationship between the difference in the mean (standardized mean difference [SMD] or "Hedge's g") and variability in parasite success (log coefficient of variation ratio [lnCVR]). A) Effect sizes from the actual data and B) effect sizes from the simulated data. Assuming that there were approximately ten replicates per study in the actual effect size data, simulated data were randomly generated from a normal distribution for each high versus low population genetic diversity comparison for the calculation of simulated effect sizes. The significance of the mean-variance relationships for each dataset

was tested using linear modelling and is shown by the blues lines with 95% confidence interval bands (linear model coefficient = -0.24 and -1.15; SE = 0.04 and 0.06; $p < 0.001$ and $p < 0.001$ for the real and fake dataset respectively).

5267