DISS. ETH NO. 26902

PARASITE PING PONG
—
HOST-PARASITE COEVOLUTION, METAPOPULATION STRUCTURE AND GENETIC DIVERSITY

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

ROBERT PASCAL DÜNNER

MSc ETH Biology

Born on 31.07.1987

citizen of Zollikon ZH and Zurich ZH

accepted on the recommendation of

Prof. Dr. Jukka Jokela
Prof. Dr. Mike Boots

2020

Contents

Page No.

Zusammenfassung

Genetische Diversität ist ein allgegenwärtiges Phänomen dessen Existenz Evolutionsbiologen seit Generationen fasziniert. Häufig wird die Koevolution von Wirtsorganismen mit Parasiten als Erklärung für die vorhandene genetische Diversität herangezogen. Von besonderer Bedeutung ist in diesem Zusammenhang die negativ -genotyp-frequenzabhängige Interaktion, welche einen oszillierenden Selektionsdruck erzeugt, der immer wieder andere Genotypen bevorzugt und somit eine hohe genetische Diversität fördert.

Negativ frequenzabhängige Interaktionen von Wirten und Parasiten sind häufig hochdynamische Prozesse und können also solche leicht beeinflusst werden. Ein Pendel kann gut als Beispiel verwendet werden um das schwingen von einer Wirtsgenotypfrequenz zur nächsten zu visualisieren. Genau wie ein bei einem Pendel können schon kleinste Störungen das dynamische Verhalten von Wirts- und Parasitgenotypfrequenzen verändern.

Bevor nun aber Aussagen über das Verhalten eines Pendels gemacht werden können, muss zuerst das Pendel selbst charakterisiert werden. Wie lange ist der Faden, wie schwer der Kopf des Pendels? In Kapitel 1 präsentiere ich eine neue, individuen-basierte Simulation für Wirt-Parasit Koevolution welche es erlaubt die Lebensgeschichte einzelner Individuen hochdetailiert zu simulieren. Populationen können durch das Erzeugen von Gruppen solcher Individuen simuliert werden. Das Verhalten von Populationen ist somit ein Ergebnis einer Vielzahl von Interaktionen zwischen Individuen. Diese Art Simulation erlaubt einen neuen, detaillierten Blick auf die Kontextabhängigkeit von Wirts-Parasit-Interaktionen.

Ein Pendel das an einem kurzen Faden hängt wird anders schwingen als ein Pendel an einem längeren Faden, erst recht wenn sich auch noch der Befestigungspunkt bewegt. Genau so wird auch das dynamische Verhalten von Wirts- und Parasit Genotypfrequenzen in kleineren Populationen anders sein als in grösseren Populationen. Dies wird in Kapitel 2 behandelt, wo ich die Interaktion zwischen zufälligem genetischen Drift und negativ-frequenzabhängiger Selektion beleuchte und auf die Auswirkungen dieser Interaktion auf die Erhaltung der genetischen Diversität eingehe. Die Interaktion kann bei kleinen Populationsgrössen bewirken dass sich das Vorhandensein eines Parasiten negativ auf die Erhaltung der genetischen Diversität auswirkt anstatt diese zu erhalten.

Zwei Pendel deren Schwingungen auch nur minimst verbunden sind, etwa über eine Feder, werden ein ganz anderes Verhalten zeigen als zwei komplett isolierte Pendel. Negativ frequenzabhängige Selektion welche in asymmetrisch verbundenen Metapopulationen stattfindet wird zu einem anderen Resultat führen als negativ frequenzabhängige Selektion in einzelnen, isolierten Populationen. In Kapitel 3 zeige ich Resultate einer Simulation einer Metapopulation mit komplett verbundenen Parasitensubpopulationen und komplett isolierten Wirtssubpopulationen. Ich zeige auf was diese spezielle Metapopulationsstruktur für Auswirkungen auf die Verteilung der genetischen Diversität des Wirtes hat. Nicht nur wird die negativ-frequenzabhängige Selektion innerhalb der Wirtssubpopulationen gestoppt, sie wird auf das Niveau der Metapopulation angehoben. Dies hat zur Konsequenz das sich die Genotypfrequenzen der Wirtssubpopulationen untereinander stark zu Unterscheiden beginnen, bis hin zur kompletten Partitionierung der genetischen Diversität.

Ein Pendel welches im Wasser schwingt und ein Pendel welches in der Luft schwingt zeigen ein klar anderes Verhalten. Dasselbe gilt für Wirts- und Parasit Koevolution unter verschiedenen Umweltbedingungen. Kapitel 4 handelt von einem Experiment mit *Daphnia magna* als Wirt und *Ordospora colligata* als Parasit. In diesem Experiment zeige ich das die Interaktion der *Daphnia*-Klone und der *Ordospora*-Linien in verschiedenen Temperaturen zu unterschiedlichen Ergebnissen führt. Dieser Effekt ist gleichzeitig sehr subtil und hoch-konsequent, da sich die Altersstruktur des Wirtes als Reaktion auf den Parasiten verschiebt. Sollten solche GxGxE Interaktionen verbreitet sein, so könnte dies weitreichende Konsequenzen für unsere Erwartungen an die Wirts-Parasit-Koevolution haben.

Summary

Genetic diversity is a phenomenon whose existence has fascinated generations of evolutionary biologists. Coevolution of hosts and parasites is a plausible explanation for its existence. Of special interest in this respect is negative frequency-dependent selection between host and parasite genotypes (NFDS). NFDS can create fluctuating selection pressure which repeatedly favours different host and parasite genotypes and hence could maintain genetic diversity.

NFDS between hosts and parasites is a dynamic process and as such is susceptible to interference. I like to visualize the dynamic swings of host and parasite genotype frequencies as a pendulum, swinging from one frequency to another. And much like with a pendulum, also small disturbances, can interfere greatly with the dynamics of host and parasite genotype frequencies.
But before statements about the behavior of a pendulum can be made, the pendulum itself must be examined. How long is the string, how heavy the head? In chapter 1 I introduce a novel agent-based time-forward stochastic simulation for host and parasite co-evolution which allows to simulate life histories of individual agents in great detail. Populations can be simulated by generating groups of such agents. Population-level behaviour then emerges from a multitude of individual-level interactions. This type of simulation allows new and detailed insights into the context-dependency of host-parasite coevolution.

A pendulum on a short string swings differently from a pendulum on a longer string. Similarly will the dynamic behaviour of host and parasite genotype frequencies change with population size. This is treated in chapter two, where I explore the consequences of the interaction between random drift and NFDS for the maintenance of genetic diversity. In small populations, NFDS can lead to the degradation of genetic diversity by boosting drift effects instead of maintaining diversity by counteracting drift.

Two pendulums whose swing is ever so lightly connected will display an entirely different dynamic from two isolated pendulums. NFDS that is happening in host and parasite populations that are part of a metapopulation will show a different outcome than NFDS in isolated populations. This I research in chapter 3, where I look at the outcome of NFDS in asymmetrically connected host-parasite metapopulations; and explore the consequence of this for the patterning of genetic diversity. I show that a special metapopulation structure, where parasite subpopulations are highly connected, but host subpopulations remain isolated, can switch of NFDS in host subpopulations. NFDS is instead shifted to the level of the metapopulation. This results in strong genetic divergence of, and the compartmentalisation of host genetic diversity into, host subpopulations.

Pendulums that swing in air and pendulums that swing in water clearly behave differently. So is the expectation for host and parasite co-evolution in different environments. In chapter 4 I show genotype by genoptype by environment interactions in an empirical system and speculate on the consequences of widespread GxGxE interactions on host and parasite co-evolution. I conducted an experiment using *Daphnia magna* as host and *Ordospora colligata* as parasite. I exposed *Daphnia* clones to *Ordospora* lines in different temperatures and found that this leads to a GxGxE interaction in the shift of the host demography. If such GxGxE interactions should be widespread, then this will have consequences on the expectations for the outcomes of host-parasite coevolution.

General introduction

Introduction

Parasites are everywhere. They come in a plethora of forms and functions, but they all have in common that they need to infect a host for their livelihood. Parasitism is probably the "most common animal lifestyle" on the planet (de Meeûs and Renaud 2002, Hechinger and Lafferty 2005, Lafferty et al. 2006, 2008), found on every continent and in almost all ecosystems. Almost all known species, from tiny bacteria to giant blue whales, have one or more parasite species.
Parasites can influence the species composition of ecosystems. They can facilitate coexistence of species by changing the competitive relationship between them upon infection (Park 1948, Barbehenn 1969, Freeland 1983, Morris et al. 2004, Hatcher et al. 2006). Or they can prevent coexistence by having disproportional virulence on some of the involved host species (apparent competition) (Anderson 1972, Cornell 1974, Price et al. 1986, Holt and Lawton 1994).
Parasites can also enable neobiotic expansion, influencing species composition on a global scale (Hatcher et al. 2006, Dunn et al. 2012, Dunn and Hatcher 2015, García-Ramos et al. 2015). Invading species can profit from parasite release when they make the jump to a new ecosystem without bringing their parasites with them (Sánchez et al. 2012). Invading host species can also bring an adapted parasite along with them that is more virulent on native host species (disease-mediated-invasion, DMI) (Strauss et al. 2012).
Interactions of parasites with one or more host species can have widespread carry-over effects on whole ecosystems (Hatcher et al. 2012). Parasites can decrease a host species density (Ebert et al. 2000), alter the interaction of predator and prey species (Su et al. 2015), decrease the efficiency of grazers (Wood et al. 2007), affect food web stability and energy flow (Thompson et al. 2005, Lafferty et al. 2006, 2008, Dunne et al. 2013) or even change behaviour of key ecosystem engineer species (Thomas et al. 1998).

Implicit in those mentioned effects of parasites on ecosystems is that infection with a parasite can decrease the competitive ability of a host (individual, genotype, species, ...) in respect to other hosts (Refardt and Ebert 2012). This change of competitive ranking upon parasite infection can lead to changes in the genotype distribution within host populations that is dependent on the identities of both the host and the parasite (Haag and Ebert 2004). Infection success as well as the outcome of an infection by parasites can be specific on the genotypes of both the host and the parasite (G x G interaction) (Salvaudon et al. 2005, Rauch et al. 2006). Often no host or parasite genotype can be considered universally resistant or virulent (Carius et al. 2001). Two key models for host-parasite specificity have been proposed. The gene-for-gene model is originally based on plant-pathogen interactions and allows for universally resistant hosts or universally infective parasites. In order to successfully establish stable dynamical behaviour of host and parasite genotype frequencies it needs the assumption of cost of resistance or infectivity (Thompson and Burdon 1992, Agrawal and Lively 2002, Thrall 2003, Nuismer 2006). The matching-allele model does not have universally resistant or infective hosts or parasites but genotype-specific interactions, where only matching hosts and parasites can successfully infect. Under the matching-allele model there is no need for cost of resistance or cost of infectivity to create stable dynamic behaviour in genotype frequencies. The matching-allele model is the most widespread infection model used for theory on host-parasite coevolution (Hamilton 1980, Frank 1994, Peters and Lively 1999, Agrawal and Lively 2002, Nuismer 2006).

The genetic specificity of infection and the change of host competitive ability upon infection by a parasite can lead to negative frequency-dependent selection (NFDS) (Clarke 1976).
Under negative frequency-dependency the specific parasite genotype that is compatible with the most common host genotype will have the highest number of successful infections in the host population. The selective advantage will lead to an increase in frequency of this specific parasite genotype. In turn, this will lead to an increasing infection rate in the matching, common host genotype. As infection with a parasite can change competitive ability of the host, this host genotype is then selected against and highly likely to decrease in frequency.
This NFDS leads to common host genotypes being selected against, and conversely, to an advantage for rare host genotypes. As rare host genotypes will be under-infected by their genotype-specific parasite, they enjoy the "rare advantage" and are likely to subsequently increase in frequency in the population.

The Red Queen hypothesis was first introduced by Van Valen (1973) as an argument supporting the observation that the extinction risk of a species is uncorrelated to age of this species. Van Valen interpreted that this means that all species have to continuously adapt to new conditions in order not to go extinct, and drew a parallel to the character of the Red Queen in Lewis Carols novel "Alice through the looking-glass".

This analogy resurfaced later in the context of host-parasite coevolution and the maintenance of sexual reproduction (Bell 1982, Stenseth and Smith 1984).
Clonal organisms can, all other things considered equal, reproduce twice as fast as sexual organisms. This should, in theory, quickly lead to the competitive exclusion of sexual organisms in populations where both coexist. The reason for this is that in clonally reproducing species, females do not have to bother with males in order to reproduce. This was called the "twofold cost of sex" by Maynard Smith (1968) and has bothered evolutionary biologists ever since.
Earlier, Muller (1964) had proposed that mutational meltdown of clonal organisms (Muller's ratchet) could be the reason for the maintenance of sexual reproduction, later reiterated by Felsenstein (1974). The time needed for clonal lines to reach mutational meltdown though was either too long to provide enough disadvantage to clonal lines to maintain sexual lines or needed extremely high mutation rates.
Maynard Smith (1971) and Charlesworth (1976) pointed out that in order for sexuals not to be outcompeted by clones, the sign of epistasis and linkage disequilibrium would need to change repeatedly and quickly. First Jaenike (1978), drawing on Clarke (1976), and then Hamilton (1980) proposed that NFDS by parasites could provide such a fluctuating selection pressure. Finally, Bell (1982) and then Stenseth and Smith (1984) took up the term Red Queen for this hypothesis, as host species continually have to adapt to their parasites in order not to become over-infected, and parasites continually have to adapt to their hosts in order to still be able to infect them in the future. Continuous adaptation is needed by hosts and parasites just to keep up the status quo. NFDS of parasites on host genotypes can lead to cyclical dynamics in host and parasite genotype frequencies. The Red Queen thus was a fitting analogy for the repeated replacement of common host genotypes by formerly rare genotypes without a change in genetic diversity.
The Red Queen hypothesis has since been extensively researched and advanced (Bremermann 1980, Judson 1997, Dybdahl and Storfer 2003, Kouyos et al. 2007, Salathé et al. 2009, Rabajante et al. 2015, 2016, da Silva and Galbraith 2017, Anzia and Rabajante 2018) and has received empirical support as well (Lively 1989, Lively et al. 1990, Dybdahl and Lively 1998, Decaestecker et al. 2007, Jokela et al. 2009, Wolinska and Spaak 2009, Morran et al. 2011, Råberg et al. 2014, Ignacio-Espinoza et al. 2020). At the same time there has been ample criticism, see (Salathé and Bonhoeffer 2008, Lively 2010, Brockhurst et al. 2014) for reviews.

One side-effect of NFDS by parasites is that it may not only maintain sexual reproduction, but also genetic diversity in general (Wright 1948, Haldane 1949, Haldane and Jayakar 1963, Clarke and O'Donald 1964, Gillespie 1975, Clarke 1976, 1979, Levin 1988). NFDS provides a selective advantage to host genotypes at low relative frequency, thus counteracting the diversity degrading effects of drift and directional selection. This effect has also been found in empirical studies that report host populations suffering high levels of parasitism have higher levels of genetic diversity (Gulland et al. 1993, Weeks and Hoffmann 2008, King et al. 2011, Gsell et al. 2013, Dagan et al. 2013, Turko et al. 2018, Kurbalija Novičić et al. 2020), or that the genetic structure of a population under parasite pressure changes in accordance with NFDS (Lively 1992, Jokela et al. 2003, Hall et al. 2011, Paczesniak et al. 2014, Ignacio-Espinoza et al. 2020).

Host-parasite coevolution rarely takes place in singular isolated host-parasite populations, but instead in metapopulations consisting of several, more or less interconnected, host and parasite subpopulations. The structure of the metapopulation can greatly influence the outcome of host-parasite coevolution (Ladle et al. 1993, Thrall et al. 2012, Boëte et al. 2019). Inherent to the concept of the metapopulation is that there are subpopulations which are connected by migration of the host and/or the parasite (Hamilton 2009). Depending on the size distribution of the subpopulations and the connection regime, a few characteristic types of metapopulations have been identified, like the mainland-island or the stepping stone model (Hamilton 2009).
The structure of the migration pattern, especially the amount of connectedness between metapopulations can greatly influence the spread of parasite epidemics through the metapopulation (Watts et al. 2005, Huang et al. 2015).
Migration distance for example can influence virulence evolution of parasites. A model by Boots and Sasaki (1999) showed that when populations become more connected and parasite infection thus takes place over longer distances, the parasite is selected to evolve to higher virulence. The same pattern was later found in an experiment by Boots and Mealor (2007) that showed that a decreased migration distance of the host and hence a more localized interaction with a parasite, led to reduced virulence of the parasite in about eight generations.
The larger the geographic span of a metapopulation, the higher the probability that subpopulations differ in environmental conditions. This could have great influence on host-parasite coevolution in the

metapopulation. Environmental conditions have the potential to change the outcome of host-parasite interactions through context-dependency in a genotype-by-genotype-by-environment kind of fashion (G x G x E) (Mitchell et al. 2005, Wolinska and King 2009, Tadiri et al. 2013).
This geographical context-dependency of coevolution led Thompson (1999) to develop the geographic mosaic theory of coevolution. A metapopulation can have coevolutionary hot-spots and cold-spots. A coevolutionary hot-spot is a subpopulation where the conditions of host and parasite diversity and migration as well as the environment are right for intense coevolutionary interactions. A coevolutionary cold-spot is a subpopulation where those conditions are at a combination that does not allow meaningful coevolutionary interaction (Gomulkiewicz et al. 2000, Barabas et al. 2004, Forde et al. 2004, Thompson 2005).
Such hot-spots of coevolution can lead to local adaptation of host and parasite. Local adaptation is often defined as locally adapted parasite populations having higher infection success on sympatric host populations that on allopatric host populations (Parker 1985, Lively 1989, Ebert 1994, Lively and Jokela 1996, Kaltz and Shykoff 1998, Lively and Dybdahl 2000).
Migration rate is thought to be a key parameter influencing the strength of local adaptation. Moderate migration rates can increase local genetic diversity, thus providing genetic variation upon which selection can act. Lower migration rates turn the focus more on local interactions. Extremely high migration rates swamp any pattern of local adaptation. In general, it is thought that the interaction partner that migrates more is usually locally adapted (Gandon 1996, Gandon and Michalakis 2002).
The migration pattern of both the host and the parasite also determines the amount of genetic diversity that can be found in the metapopulation. Restricted migration, leading to localized interactions, can increase the amount of genetic diversity in the metapopulation compared to cases with high migration of both host and parasite (Thrall and Burdon 2002, Papaïx et al. 2014).

Specific hierarchical migration patterns of host and parasite create a network structure of the metapopulation that can be characterized with tools from graph theory (Keeling 2005, Keeling and Eames 2005, Gilarranz and Bascompte 2012, Dunne et al. 2013, Pilosof et al. 2014). This could allow to identify certain topologies of metapopulations that in themselves have influence on the behaviour of the host-parasite dynamics (Eubank et al. 2004, Siegal et al. 2006, Gilarranz and Bascompte 2012). The idea is that certain topologies could, like biochemical regulatory network motifs, create certain functionalities (Pavlopoulos et al. 2011). Amplifier network structures, for example, can dramatically increase the fixation probabilities of beneficial mutants compared to random network structure (Lieberman et al. 2005).

Taken together, this means that host-parasite coevolution may only be fully understood in the context of metapopulation structure.

This thesis

The central topic of my thesis is host-parasite coevolution, with a special focus on NFDS.
I address the question how the dynamic interaction of host and parasite genotype frequencies is modified by its contextual environment. Environment should be understood very broadly here, as it encompasses both the classical ecological environment as well as the population structure of the host and the parasite. Negative frequency-dependent interactions of hosts and parasites are highly dynamical processes and as such have a great potential to be influenced by interference with other processes.

In chapter 1 I present an agent-based, time-forward, stochastic simulation for host and parasite coevolution (Dünner 2020). The simulation uses agents based on the resource allocation hypothesis, so that the agents face trade-offs of internal energy allocation into different processes (Stearns 1980, Noordwijk and Jong 1986, Perrin and Sibly 1993). Both host and parasite type agents are simulated, and arbitrary infection systems of varying specificity and virulence can be implemented. To be able to answer questions on host-parasite coevolution in metapopulations, distinct host and parasite populations of varying sizes can be simulated. Those populations can then be connected by customizable migration patterns, both for the host and the parasite separately. As the simulation is time-forward, host and parasite dynamics can be examined in great temporal detail. I named the simulation "Digital_Coevolution simulation".

In chapter 2 I used the Digital_Coevolution simulation to explore the interaction of random genetic drift and NFDS. I show that in small host populations, NFDS by parasites increases the strength of genetic drift. This interaction leads to a faster degradation of genetic diversity in host populations that are coevolving with parasites than in host populations where no parasites are present.

In chapter 3 I used the Digital_Coevolution simulation to examine the influence of metapopulation structure on host-parasite coevolution. Specifically, I examined hierarchically structured metapopulations, where the parasite metapopulation consisted of subpopulations that were strongly connected, and the host metapopulation consisted of subpopulations that were isolated. In this special metapopulation structure, host-parasite coevolution leads to the divergence of host subpopulation genotype frequencies, the loss of NFDS within host subpopulations and to a distinct compartmentalisation of host genetic diversity.

In chapter 4 I report on an experiment using *Daphnia magna* as hosts and *Ordospora colligata* as parasites. I challenged monoclonal host populations with single lines of parasites in a factorial design in two environmental conditions. I found marked GxGxE interactions that manifested in changes of the hosts demography.

References

Agrawal, A., and C. M. Lively. 2002. Infection genetics: Gene-for-gene versus matching- alleles models and all points in between. Evolutionary Ecology Research:79–90.

Anderson, R. C. 1972. The ecological relationships of meningeal worm and native cervids in North America. Journal of Wildlife Diseases 8:304–310.

Anzia, E. L., and J. F. Rabajante. 2018. Antibiotic-driven escape of host in a parasite-induced Red Queen dynamics. Royal Society Open Science 5:180693.

Barabas, L., B. Gilicze, F. Takasu, and C. Moskat. 2004. Survival and anti-parasite defense in a host metapopulation under heavy brood parasitism: A source?Sink dynamic model. Journal of Ethology 22.

Barbehenn, K. R. 1969. Host-Parasite Relationships and Species Diversity in Mammals: An Hypothesis. Biotropica 1:29.

Bell, G. 1982. The masterpiece of nature: The evolution and genetics of sexuality. CroomHelm, London, United Kingdom.

Boëte, C., M. Seston, and M. Legros. 2019. Strategies of host resistance to pathogens in spatially structured populations: An agent-based evaluation. Theoretical Population Biology 130:170–181.

Boots, M., and M. Mealor. 2007. Local Interactions Select for Lower Pathogen Infectivity. Science 315:1284–1286.

Boots, M., and A. Sasaki. 1999. "Small worlds" and the evolution of virulence: Infection occurs locally and at a distance. Proceedings of the Royal Society of London. Series B: Biological Sciences 266:1933–1938.

Bremermann, H. 1980. Sex and polymorphism as strategies in host-pathogen interactions. Journal of Theoretical Biology:671–702.

Brockhurst, M. A., T. Chapman, K. C. King, J. E. Mank, S. Paterson, and G. D. D. Hurst. 2014. Running with the Red Queen: The role of biotic conflicts in evolution. Proceedings of the Royal Society B: Biological Sciences 281:20141382.

Carius, H. J., T. J. Little, and D. Ebert. 2001. Genetic variation in a host-parasite association: Potential for coevolution and frequency-dependent selection. Evolution 55:1136–1145.

Charlesworth, B. 1976. Recombination modification in a fluctuating environment. Genetics:15.

Clarke, B. 1976. The ecological genetics of host-parasite relationships. *in* Genetic aspects of host-parasite relationships. Blackwell, London.

Clarke, B. 1979. The Evolution of Genetic Diversity. Proceedings of the Royal Society B: Biological Sciences 205:453–474.

Clarke, B., and P. O'Donald. 1964. Frequency-dependent selection. Heredity:6.

Cornell, H. 1974. Parasitism and Distributional Gaps between Allopatric Species. The American Naturalist 108:880–883.

Dagan, Y., K. Liljeroos, J. Jokela, and F. Ben-Ami. 2013. Clonal diversity driven by parasitism in a freshwater snail. Journal of Evolutionary Biology 26:2509–2519.

da Silva, J., and J. D. Galbraith. 2017. Hill-Robertson interference maintained by Red Queen dynamics favours the evolution of sex. Journal of Evolutionary Biology 30:994–1010.

Decaestecker, E., S. Gaba, J. A. M. Raeymaekers, R. Stoks, L. Van Kerckhoven, D. Ebert, and L. De Meester. 2007. HostParasite "Red Queen" dynamics archived in pond sediment. Nature 450:870–873.

de Meeûs, T., and F. Renaud. 2002. Parasites within the new phylogeny of eukaryotes. Trends in Parasitology 18:247–251.

Dunn, A. M., and M. J. Hatcher. 2015. Parasites and biological invasions: Parallels, interactions, and control. Trends in Parasitology 31:189–199.

Dunn, A. M., M. E. Torchin, M. J. Hatcher, P. M. Kotanen, D. M. Blumenthal, J. E. Byers, C. A. Coon, V. M. Frankel, R. D. Holt, R. A. Hufbauer, A. R. Kanarek, K. A. Schierenbeck, L. M. Wolfe, and S. E. Perkins. 2012. Indirect effects of parasites in invasions. Functional Ecology 26:1262–1274.

Dunne, J. A., K. D. Lafferty, A. P. Dobson, R. F. Hechinger, A. M. Kuris, N. D. Martinez, J. P. McLaughlin, K. N. Mouritsen, R. Poulin, K. Reise, D. B. Stouffer, D. W. Thieltges, R. J. Williams, and C. D. Zander. 2013. Parasites Affect Food Web Structure Primarily through Increased Diversity and Complexity. PLoS Biology 11:e1001579.

Dünner, R. P. 2020. Digital_Coevolution: An agent-based simulation for host-parasite coevolution.

Dybdahl, M. F., and C. M. Lively. 1998. Host-Parasite Coevolution: Evidence for Rare Advantage and Time-Lagged Selection in a Natural Population. Evolution 52:1057.

Dybdahl, M. F., and A. Storfer. 2003. Parasite local adaptation: Red Queen versus Suicide King. Trends in Ecology & Evolution 18:523–530.

Ebert, D. 1994. Virulence and Local Adaptation of a Horizontally Transmitted Parasite. Science 265:1084–1086.

Ebert, D., M. Lipsitch, and K. L. Mangin. 2000. The Effect of Parasites on Host Population Density and Extinction: Experimental Epidemiology with Daphnia and Six Microparasites. The American Naturalist 156:19.

Eubank, S., H. Guclu, V. S. Anil Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. 2004. Modelling disease outbreaks in realistic urban social networks. Nature 429:180–184.

Felsenstein, J. 1974. The evolutionary advantage of recombination. Genetics:737–756.

Forde, S. E., J. N. Thompson, and B. J. M. Bohannan. 2004. Adaptation varies through space and time in a coevolving hostParasitoid interaction. Nature 431:841–844.

Frank, S. A. 1994. Recognition and polymorphism in host-parasite genetics. Philosophical Transactions of the Royal Society B: Biological Sciences:283–293.

Freeland, W. J. 1983. Parasites and the Coexistence of Animal Host Species. The American Naturalist 121:223–236.

Gandon, S. 1996. Local adaptation and gene-for-gene coevolution in a metapopulation model. Proceedings of the Royal Society of London. Series B: Biological Sciences 263:1003–1009.

Gandon, S., and Y. Michalakis. 2002. Local adaptation, evolutionary potential and host-parasite coevolution: Interactions between migration, mutation, population size and generation time: Local adaptation and coevolution. Journal of Evolutionary Biology 15:451–462.

García-Ramos, G., L. A. Dunoyer, K. L. Sasser, and P. H. Crowley. 2015. Evolution of resistance by a native competitor can lead to invasion collapse in disease-mediated invasions. Biological Invasions 17:2863–2879.

Gilarranz, L. J., and J. Bascompte. 2012. Spatial network structure and metapopulation persistence. Journal of Theoretical Biology 297:11–16.

Gillespie, J. H. 1975. Natural Selection for Resistance to Epidemics. Ecology 56:493–495.

Gomulkiewicz, R., J. N. Thompson, R. D. Holt, S. L. Nuismer, and M. E. Hochberg. 2000. Hot Spots, Cold Spots, and the Geographic Mosaic Theory of Coevolution. The American Naturalist 156:156–174.

Gsell, A. S., L. N. de Senerpont Domis, K. J. Verhoeven, E. van Donk, and B. W. Ibelings. 2013. Chytrid epidemics may increase genetic diversity of a diatom spring-bloom. The ISME Journal 7:2057–2059.

Gulland, F. M. D., S. D. Albon, J. M. Pemberton, P. R. Moorcroft, and T. H. Clutton-Brock. 1993. Parasite-associated polymorphism in a cyclic ungulate population. Proceedings of the Royal Society of London. Series B: Biological Sciences 254:7–13.

Haag, C. R., and D. Ebert. 2004. ParasiteMediated selection in experimental metapopulations of *Daphnia Magna*. Proceedings of the Royal Society of London. Series B: Biological Sciences 271:2149–2155.

Haldane, J. B., S. 1949. Disease and evolution. La Ricerca Scientifica Supplemento.

Haldane, J. B., S., and S. D. Jayakar. 1963. Polymorphism due to selection depending on the composition of a population. Journal of Genetics 3:318–323.

Hall, A. R., P. D. Scanlan, A. D. Morgan, and A. Buckling. 2011. Host-parasite coevolutionary arms races give way to fluctuating selection: Bacteria-phage coevolutionary dynamics. Ecology Letters 14:635–642.

Hamilton, M. B. 2009. Population Genetics. Wiley-Blackwell.

Hamilton, W. D. 1980. Sex versus Non-Sex versus Parasite. Oikos 35:282.

Hatcher, M. J., J. T. A. Dick, and A. M. Dunn. 2006. How parasites affect interactions between competitors and predators. Ecology Letters 9:1253–1271.

Hatcher, M. J., J. T. Dick, and A. M. Dunn. 2012. Diverse effects of parasites in ecosystems: Linking interdependent processes. Frontiers in Ecology and the Environment 10:186–194.

Hechinger, R. F., and K. D. Lafferty. 2005. Host diversity begets parasite diversity: Bird final hosts and trematodes in snail intermediate hosts. Proceedings of the Royal Society B: Biological Sciences 272:1059–1066.

Holt, R. D., and J. H. Lawton. 1994. The Ecological Consequences of Shared Natural Enemies. Annual Review of Ecology and Systematics 2:495–520.

Huang, Z. Y., F. van Langevelde, H. H. Prins, and W. F. de Boer. 2015. Dilution versus facilitation: Impact of connectivity on disease risk in metapopulations. Journal of Theoretical Biology 376:66–73.

Ignacio-Espinoza, J. C., N. A. Ahlgren, and J. A. Fuhrman. 2020. Long-term stability and Red Queen-like strain dynamics in marine viruses. Nature Microbiology 5:265–271.

Jaenike, J. 1978. An Hypothesis to account for the maintenance of sex within populations. Evolutionary Theory:191–194.

Jokela, J., M. F. Dybdahl, and C. M. Lively. 2009. The Maintenance of Sex, Clonal Dynamics, and Host-Parasite Coevolution in a Mixed Population of Sexual and Asexual Snails. The American Naturalist 174:S43–S53.

Jokela, J., C. M. Lively, M. F. Dybdahl, and J. A. Fox. 2003. Genetic variation in sexual and clonal lineages of a freshwater snail: GENETIC VARIATION and MAINTENANCE OF SEX. Biological Journal of the Linnean Society 79:165–181.

Judson, O. P. 1997. A Model of Asexuality and Clonal Diversity: Cloning the Red Queen. Journal of Theoretical Biology 186:33–40.

Kaltz, O., and J. A. Shykoff. 1998. Local adaptation in hostParasite systems. Heredity:361–370.

Keeling, M. 2005. The implications of network structure for epidemic dynamics. Theoretical Population Biology 67:1–8.

Keeling, M. J., and K. T. Eames. 2005. Networks and epidemic models. Journal of The Royal Society Interface 2:295–307.

King, K. C., J. Jokela, and C. M. Lively. 2011. Parasites, sex, and clonal diversity in natural snail populations. Evolution 65:1474–1481.

Kouyos, R. D., M. Salathé, and S. Bonhoeffer. 2007. The Red Queen and the persistence of linkage-disequilibrium oscillations in finite and infinite populations. BMC Evolutionary Biology 7:211.

Kurbalija Novičić, Z., A. Sayadi, M. Jelić, and G. Arnqvist. 2020. Negative frequency dependent selection contributes to the maintenance of a global polymorphism in mitochondrial DNA. BMC Evolutionary Biology 20:20.

Ladle, R. J., R. A. Johnstone, and O. P. Judson. 1993. Coevolutionary dynamics of sex in a metapopulation: Escaping the Red Queen. Proceedings of the Royal Society of London. Series B: Biological Sciences 253:155–160.

Lafferty, K. D., S. Allesina, M. Arim, C. J. Briggs, G. De Leo, A. P. Dobson, J. A. Dunne, P. T. J. Johnson, A. M. Kuris, D. J. Marcogliese, N. D. Martinez, J. Memmott, P. A. Marquet, J. P. McLaughlin, E. A. Mordecai, M. Pascual, R. Poulin, and D. W. Thieltges. 2008. Parasites in food webs: The ultimate missing links: Parasites in food webs. Ecology Letters 11:533–546.

Lafferty, K. D., A. P. Dobson, and A. M. Kuris. 2006. Parasites dominate food web links. Proceedings of the National Academy of Sciences 103:11211–11216.

Levin, B. R. 1988. Frequency-dependent selection in bacterial populations. Philosophical Transactions of the Royal Society B: Biological Sciences:459–472.

Lieberman, E., C. Hauert, and M. A. Nowak. 2005. Evolutionary dynamics on graphs 433:5.

Lively, C. M. 1989. Adaptation by a Parasitic Trematode to Local Populations of Its Snail Host. Evolution 43:1663.

Lively, C. M. 1992. Parthenogenesis in a freshwater snail: Reproductive assurance versus parasitic release. Evolution 46:907–913.

Lively, C. M. 2010. A Review of Red Queen Models for the Persistence of Obligate Sexual Reproduction. Journal of Heredity 101:S13–S20.

Lively, C. M., C. Craddock, and R. C. Vrijenhoek. 1990. Red Queen hypothesis supported by parasitism in sexual and clonal fish. Nature 344.

Lively, C. M., and M. F. Dybdahl. 2000. Parasite adaptation to locally common host genotypes. Nature 405:679–681.

Lively, C. M., and J. Jokela. 1996. Clinal variation for local adaptation in a host-parasite interaction. Proceedings of the Royal Society of London. Series B: Biological Sciences.

Maynard Smith, J. 1968. Evolution in Sexual and Asexual Populations. The American Naturalist 102:469–473.

Maynard Smith, J. 1971. What use is sex. Journal of Theoretical Biology.

Mitchell, S. E., E. S. Rogers, T. J. Little, and A. F. Read. 2005. Host-parasite and genotype-by-environment interactions: Temperature modifies potential for selection by a sterilizing pathogen. Evolution 59:70–80.

Morran, L. T., O. G. Schmidt, I. A. Gelarden, R. C. Parrish, and C. M. Lively. 2011. Running with the Red Queen: Host-Parasite Coevolution Selects for Biparental Sex. Science 333:216–218.

Morris, R. J., O. T. Lewis, and H. C. J. Godfray. 2004. Experimental evidence for apparent competition in a tropical forest food web. Nature 428:310–313.

Muller, H. 1964. The relation of recombination to mutational advance. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 1:2–9.

Noordwijk, A. J. van, and G. de Jong. 1986. Acquisition and Allocation of Resources: Their Influence on Variation in Life History Tactics. The American Naturalist 128:137–142.

Nuismer, S. L. 2006. Parasite local adaptation in a geographic mosaic. Evolution 60:24–30.

Paczesniak, D., S. Adolfsson, K. Liljeroos, K. Klappert, C. M. Lively, and J. Jokela. 2014. Faster clonal turnover in high-infection habitats provides evidence for parasite-mediated selection. Journal of Evolutionary Biology 27:417–428.

Papaïx, J., J. J. Burdon, C. Lannou, and P. H. Thrall. 2014. Evolution of Pathogen Specialisation in a Host Metapopulation: Joint Effects of Host and Pathogen Dispersal. PLoS Computational Biology 10:e1003633.

Park, T. 1948. Interspecies Competition in Populations of Trilobium confusum Duval and Trilobium castaneum Herbst. Ecological Monographs 18:265–307.

Parker, M. A. 1985. Local Population Differentiation for Compatibility in an Annual Legume and its Host-Specific Fungal Pathogen. Evolution 39:713.

Pavlopoulos, G. A., M. Secrier, C. N. Moschopoulos, T. G. Soldatos, S. Kossida, J. Aerts, R. Schneider, and P. G. Bagos. 2011. Using graph theory to analyze biological networks. BioData Mining 4:10.

Perrin, N., and R. M. Sibly. 1993. Dynamic Models of Energy Allocation and Investment. Annual Review of Ecology and Systematics:379–410.

Peters, A. D., and C. M. Lively. 1999. The Red Queen and Fluctuating Epistasis: A Population Genetic Analysis of Antagonistic Coevolution. The American Naturalist 154:393–405.

Pilosof, S., M. A. Fortuna, J.-F. Cosson, M. Galan, C. Kittipong, A. Ribas, E. Segal, B. R. Krasnov, S. Morand, and J. Bascompte. 2014. HostParasite network structure is associated with community-level immunogenetic diversity. Nature Communications 5:5172.

Price, P. W., M. Westoby, B. Rice, P. R. Atsatt, R. S. Fritz, J. N. Thompson, and K. Mobley. 1986. Parasite Mediation in Ecological Interactions. Annual Review of Ecology and Systematics 1:487–505.

Rabajante, J. F., J. M. Tubay, H. Ito, T. Uehara, S. Kakishima, S. Morita, J. Yoshimura, and D. Ebert. 2016. Host-parasite Red Queen dynamics with phase-locked rare genotypes. Science Advances 2:e1501548.

Rabajante, J. F., J. M. Tubay, T. Uehara, S. Morita, D. Ebert, and J. Yoshimura. 2015. Red Queen dynamics in multi-host and multi-parasite interaction system. Scientific Reports 5:10004.

Rauch, G., M. Kalbe, and T. B. Reusch. 2006. One day is enough: Rapid and specific hostParasite interactions in a stickleback-trematode system. Biology Letters 2:382–384.

Råberg, L., E. Alacid, E. Garces, and R. Figueroa. 2014. The potential for arms race and Red Queen coevolution in a protist host-parasite system. Ecology and Evolution 4:4775–4785.

Refardt, D., and D. Ebert. 2012. The impact of infection on host competition and its relationship to parasite persistence in a Daphnia microparasite system. Evolutionary Ecology 26:95–107.

Salathé, M., and S. Bonhoeffer. 2008. The effect of opinion clustering on disease outbreaks. Journal of The Royal Society Interface 5:1505–1508.

Salathé, M., R. D. Kouyos, and S. Bonhoeffer. 2009. On the Causes of Selection for Recombination Underlying the Red Queen Hypothesis. The American Naturalist 174:S31–S42.

Salvaudon, L., V. Héraudet, and J. A. Shykoff. 2005. Parasite-host fitness trade-offs change with parasite identity: Genotype-specific interactions in a plant-pathogen system. Evolution 59:2518–2524.

Sánchez, M. I., N. O. Rode, E. Flaven, S. Redón, F. Amat, G. P. Vasileva, and T. Lenormand. 2012. Differential susceptibility to parasites of invasive and native species of Artemia living in sympatry: Consequences for the invasion of A. Franciscana in the Mediterranean region. Biological Invasions 14:1819–1829.

Siegal, M. L., D. E. L. Promislow, and A. Bergman. 2006. Functional and evolutionary inference in gene networks: Does topology matter? Genetica 129:83–103.

Stearns, S. C. 1980. A New View of Life-History Evolution. Oikos 35:266.

Stenseth, N. C., and J. M. Smith. 1984. Coevolution in Ecosystems: Red Queen Evolution or Stasis? Evolution 38:870.

Strauss, A., A. White, and M. Boots. 2012. Invading with biological weapons: The importance of disease-mediated invasions. Functional Ecology 26:1249–1261.

Su, M., C. Hui, and Z. Lin. 2015. Effects of the transmissibility and virulence of pathogens on intraguild predation in fragmented landscapes. Biosystems 129:44–49.

Tadiri, C. P., F. Dargent, and M. E. Scott. 2013. Relative host body condition and food availability influence epidemic dynamics: A *Poecilia Reticulata-Gyrodactylus Turnbulli* host-parasite model. Parasitology 140:343–351.

Thomas, F., F. Renaud, T. de Meeûs, and R. Poulin. 1998. Manipulation of host behaviour by parasites: Ecosystem engineering in the intertidal zone? Proceedings of the Royal Society of London. Series B: Biological Sciences 265:1091–1096.

Thompson, J. N. 1999. Specific Hypotheses on the Geographic Mosaic of Coevolution. The American Naturalist 153:S1–S14.

Thompson, J. N. 2005. Coevolution: The Geographic Mosaic of Coevolutionary Arms Races. Current Biology 15:R992–R994.

Thompson, J. N., and J. J. Burdon. 1992. Gene-for-gene coevolution between plants and parasites. Nature 360:121–125.

Thompson, R. M., K. N. Mouritsen, and R. Poulin. 2005. Importance of parasites and their life cycle characteristics in determining the structure of a large marine food web: *Parasites and Food Web Structure*. Journal of Animal Ecology 74:77–85.

Thrall, P. H. 2003. Evolution of Virulence in a Plant Host-Pathogen Metapopulation. Science 299:1735–1737.

Thrall, P. H., and J. J. Burdon. 2002. Evolution of gene-for-gene systems in metapopulations: The effect of spatial scale of host and pathogen dispersal. Plant Pathology 51:169–184.

Thrall, P. H., A.-L. Laine, M. Ravensdale, A. Nemri, P. N. Dodds, L. G. Barrett, and J. J. Burdon. 2012. Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation: Coevolution in a wild host-pathogen system. Ecology Letters 15:425–435.

Turko, P., C. Tellenbach, E. Keller, N. Tardent, B. Keller, P. Spaak, and J. Wolinska. 2018. Parasites driving host diversity: Incidence of disease correlated with *Daphnia* clonal turnover*. Evolution 72:619–629.

Van Valen, L. 1973. A new evolutionary law. Evolutionary Theory.

Watts, D. J., R. Muhamad, D. C. Medina, and P. S. Dodds. 2005. Multiscale, resurgent epidemics in a hierarchical metapopulation model. Proceedings of the National Academy of Sciences 102:11157–11162.

Weeks, A. R., and A. A. Hoffmann. 2008. Frequency-dependent selection maintains clonal diversity in an asexual organism. Proceedings of the National Academy of Sciences 105:17872–17877.

Wolinska, J., and K. C. King. 2009. Environment can alter selection in hostParasite interactions. Trends in Parasitology 25:236–244.

Wolinska, J., and P. Spaak. 2009. The cost of being common: Evidence from natural *Daphnia* populations. Evolution 63:1893–1901.

Wood, C. L., J. E. Byers, K. L. Cottingham, I. Altman, M. J. Donahue, and A. M. H. Blakeslee. 2007. Parasites alter community structure. Proceedings of the National Academy of Sciences 104:9335–9339.

Wright, S. 1948. On the roles of directed and random changes in gene frequency in the genetics of populations. Evolution 2:279–294.

Chapter 1

Digital_Coevolution
An agent-based simulation for host-parasite coevolution

Authors: Robert P. Dünner[1,2], Jukka Jokela[1,2].

Author affiliations: ETH Zurich, Institute for Integrative Biology[1], Eawag, Department of Aquatic Ecology[2].

Corresponding author: Robert P. Dünner; robert.duenner@eawag.ch, Eawag ECO BU-G-12, Uberlandstrasse 133, 8600 Dubendorf, Switzerland.

Keywords: Agent-based simulation, resource allocation hypothesis, host-parasite coevolution, negative frequency-dependent selection, matching-alleles model, metapopulation structure, genetic drift, maintenance of genetic diversity.

Author contributions: RD conceived the ideas and designed the work, RD authored the simulation, RD and JJ led the writing.

Acknowledgements: I was supported by the adaptation to a changing environment center (ACE) at ETH. I would like to thank Jukka Jokela for his generous support and his inspiring supervision. I would like to thank Kyra Gähwiler, Nadine Tardent, Oliver Dressler and Felix Moerman for helpful comments on the manuscript.

Data accessibility: The simulation is available on GitHub.

Abstract

Host-parasite coevolution is highly context-dependent. Population size, demography, metapopulation structure and the details of both the hosts and the parasites life-histories influence the process of coevolution.
This context-dependency is difficult to be completely accounted for in empirical systems. Agent-based simulations can be used to implement and control multiple aspects of host-parasite coevolution while retaining a degree of "realism". Such simulations can be used for hypothesis testing and to generate expectations for empirical systems.
Here I present the "Digital_Coevolution" simulation, a novel, highly flexible, agent-based, time-forward simulation that I created to explore questions in negative frequency-dependent host-parasite coevolution. It allows the simulation of detailed host and parasite agent life-histories based on resource allocation trade-offs, their population structure, demography and migration patterns. Different infection systems anywhere between gene-for-gene to matching-allele models can be implemented. This makes the "Digital_Coevolution" simulation a powerful tool for research and teaching on the context-dependency of host-parasite coevolution.

Introduction

The growth of a host population and its infection with parasites is a highly dynamic, context-dependent process. Each host is an individual that is born, competes for resources, grows and ideally also reproduces. The fitness of a host individual is determined by its life-history trade-offs and their interactions with its biotic and abiotic environment.
The host might be unlucky and encounter a parasite. Whether or not this encounter leads to an infection depends on many factors. Do the genotypes of the host and the parasite match? How genotype specific and infectious is the parasite? How many parasites were encountered by the host, and how competent are the hosts defenses? Does the host have enough resources to mount an appropriate immune defense? Does the encounter happen in an environmental context beneficial for the host or the parasite? These and many more factors can determine the success of an infection.

The outcome of host-parasite interactions depends on more than just the properties of the individuals themselves. All organisms are part of a population. Some populations might be larger, some smaller, some denser, others sparse, some isolated, others very well connected and part of a larger metapopulation. Those population properties can influence the infection dynamics between hosts and parasites. Larger host populations are less affected by genetic drift and can potentially maintain a higher genetic diversity. In small populations selection by parasites can interfere with drift and increase the speed of degradation of genetic diversity (see chapter 2). Metapopulation structure impacts coevolutionary dynamics as well (Ladle et al. 1993, Lion and Gandon 2015) and influences the genetic structure of both the host and the parasite (Judson 1995, 1997). The number of migrants between subpopulations and the migration distance can, for example, influence virulence evolution (Boots and Sasaki 1999, Boots and Mealor 2007), can enable local adaptation of the parasite on the host (Gandon 1996, 2002, Gandon and Michalakis 2002) or it can select for increased genetic divergence between host subpopulations (see chapter 3).

It is challenging to account for all these individual and population level factors in order to conduct experiments on coevolution. Only few examples of well-controlled systems with different, comparable parameter settings that enable research on coevolution exist in the wild. A famous example would be the Trinidad guppy system (Oosterhout et al. 2006).
The alternative of establishing mesocosms that would allow for more control over the factors influencing infection while keeping a certain level of ecological realism is a logistical challenge in itself and can not be done in many institutions.
Establishing highly controlled laboratory populations is challenging as well. Getting the conditions right to sustain parasite infection in order to be able to do experiments can sometimes absorb more time than the actual experiment itself (personal observation).
Simulations provide an alternative to these challenges. They can present a way to implement theoretical expectations and biological assumptions for difficult-to-manipulate systems. Simulations can help novel insights themselves or can be used to generate hypotheses that are more feasibly testable in natural systems.

Invaluable insights into many areas of host-parasite coevolution have been gained by simulation studies or theoretical models. This includes the first hypotheses on how negative frequency-dependent selection by parasites maintains genetic diversity (Haldane 1949, Haldane and Jayakar 1963, Clarke 1976, 1979) or sexual reproduction (Jaenike 1978, Hamilton 1980, Hamilton et al. 1990). More recent developments aim to understand the process of evolution itself by using "digital organisms", like for example the AVIDA platform (Ofria and Wilke 2004, McKinley et al. 2008). The power and capability of computer simulations have increased steadily over time as hardware became more powerful and software advanced. This increase in computational potential opens the opportunity to embed more and more biological realism into simulations, trying to position simulations closer to well controlled laboratory experiments. All models are wrong, but some are useful (Box 1976, 1979). A simulation can never be as realistic as the real world, as much as a map can never be the real landscape. Nevertheless, one can include plausible details of the hosts and the parasites life cycles into a simulation in order to increase realism. More detailed simulations can potentially give rise to more precise predictions. Similarly, one can draw more information from a more detailed map.

One attempt to capture realism is by the use of agent-based models. These models simulate the behaviour of each individual organism explicitly, allowing for a great amount of detail. Instead of implementing required properties at the population level, those simulations describe rules for individual behaviour and development. All further higher level dynamics, for example population and metapopulation behaviour,

then result from the sum of individual level interactions (for example see (Lenski et al. 1999), (Boëte et al. 2019) or (Bonabeau 2002)).

The simulation that is presented here, called the "Digital_Coevolution" simulation, belongs to the category of agent-based models. It simulates detailed host and parasite agents as well as their interactions. Population level behaviour then emerges from the sum of individual-level interactions in groups of these agents.
The aim of the Digital_Coevolution simulation was to create a tool to study how metapopulation structure influences coevolutionary dynamics between hosts and parasites. Due to the high level of control over the life-histories of both the host and the parasite as well as their interactions, the Digital_Coevolution simulation can also be extended to the study of other questions in host-parasite coevolution, like for example evolution of virulence.

This chapter serves as the documentation for the Digital_Coevolution simulation. It describes the inner workings of and the reasoning behind the simulation and should enable the reader to thoroughly understand and use it. The first part describes the "biological logic" behind the code. It outlines the functions that govern the behaviour of the agents in the simulation. The second part then describes the implementation of the Digital_Coevolution simulation, moving from biological logic towards R logic. Throughout the document, the term agent and individual are used interchangeably.

This document is written in *r markdown* (Xie et al. 2018, Allaire et al. 2020) in R (Team 2020) using R Studio (Team 2019) as IDE. It consists of paragraphs of text explaining the simulation that are immediately followed by chunks of the according, commented R code. The text and the comments in the code do show substantial overlap, trying to reinforce a thorough understanding of what the simulation does. The code is not complete though and serves only as demonstration. If you would like to have a look at the complete source code or even contribute, you are very welcome to do so and invited to visit the Digital_Coevolution GitHub repository of the simulation (Dünner 2020).

Part One: Biology of the Digital_Coevolution simulation

Overview

The Digital_Coevolution simulation models individual host and parasite agents in detail. It has instructions for within-agent life-history as well as for between-agent behaviour. When groups of those agents are simulated, population level behaviour emerges. The simulation is implemented in a time-forward fashion. The state of each individual agent is calculated at discrete time-steps and each individual lives for multiple time-steps. Each time-step is dependent on the state of the simulation at the time-step before, similar to a Markov chain. The description below gives an overview of the dynamics that are evaluated for each individual agent at each time-step. One time-step can be thought of as one "day" in the life of an agent. Not all events do necessarily take place every time-step and nearly all are context-dependent. Not all agents will experience the same events each time-step.

The agents that make up the Digital_Coevolution simulation have internal resource-allocation trade-offs among fitness-correlated life-history traits (Stearns 1980, Noordwijk and Jong 1986, Perrin and Sibly 1993).
Each host agent has a distinct immune genotype and a set of life-history traits that it inherited. Host agents are born as small individuals with enough starting resources for the first day of their life. They immediately start foraging for resources, for which they compete with other host agents. Resources are limited for each population. The outcome of competition for resources between host agents depends on the population density, the host agent size distribution and a stochastic component.
Host agents then use those acquired resources to invest into growth, maintenance or reproduction. Excess resources can be used to build up a "fat storage" for harder times. The older an individual grows, the less resources it invests into growth and the more it invests into reproduction. Reproductive allocation is implemented as iteroparous clonal income breeding (Sibly and Calow 1984, Houston et al. 2007, Stephens et al. 2009) after a maturation threshold. Each agent has to save up resources until the maturation threshold is reached after which it can start reproducing. Each agent can reproduce several times during its lifetime, if resources permit it. The gradual, age-dependent shift of resource investment from growth to reproduction ensures that early in life the resources invested into reproduction are not sufficient to produce offspring. This creates a distinct juvenile phase that is dominated by growth. Later in life the majority of resources are invested into reproduction, and an individual only grows very slowly. This creates an adult life stage that is dominated by reproduction. An individual dies as soon as it runs out of resources (starves) or if it reaches the pre-set age threshold (senescence). Senescence is thus a deterministic event at a certain age of an agent, if that agent has survived up to that age.

Throughout all its life stages, a host agent can encounter parasite agents. Groups of parasite agents are randomly ingested by the host agent with food resources while foraging. An infection may establish within a host individual depending on host and parasite genotype matching and the number of parasite agents ingested. All types of genotype-specific infections can be implemented via an infection matrix. Infections can only be established by one parasite genotype at a time, there are no co-infections possible. Each host individual can potentially ingest several parasite agents each time-step while foraging. Parasite agents that are ingested will be removed from the parasite population, independent of their genotype, their infection success or the infection state of the host.
If an infection does establish within a host agent, it starts small. This means that infections do have a size within the infected host individual, reflecting both the physical space occupied and the resources needed by the host immune system to interact with the infection. Infections within host individuals grow and mature over time until they completely fill out the space that is available for them. This mimics a parasite that infects a certain host tissue, for example the gonads, and whose within host growth is therefore limited. The space that is available for a parasite infection is directly related to the size of the host individual.
Infections steal resources from the host individual that are then no longer available to the host. Virulence is thus implemented as a resource transfer from the host agent towards the parasite infection. The amount of resources that a parasite infection steals is dependent on the size of the infection. A younger, smaller infection will steal fewer resources than a larger, more mature infection. An infection will kill its host agent as soon as the host agent does not have enough resources left to invest into maintenance. This happens when a parasite infection is absorbing a lot of resources and there are not enough resources coming in from foraging or from the fat storage of the host.

Parasite agent reproduction is implemented with the same life-history strategy as that of host agents.

Parasite reproduction only happens within infected host agents, with the infection acting as the reproducing unit. This means that new parasite agents are created by infected host agents and not by free parasite agents. This simulates a parasite behaviour that is inspired by that of micro-parasites. They have no substantial life-history outside of their hosts, and in order to reproduce obligatorily have to infect a host. As soon as an infection has stolen enough resources from the host mature, it will produce a clutch of parasite offspring that is shed without killing the host. Each infected host will therefore repeatedly shed bursts of new parasite agents into the parasite population. The amount of parasite agents produced this way is dependent on the parasite infection size within the host agent, which itself is dependent on the age of the infection and the host size. Mature infections in large host individuals can hence have a significant influence on the parasite population.

The infection spreads horizontally through the host population, as parasite agents are passively ingested by host individuals while foraging. This, together with the non-lethal shedding of new parasite agents by infected host agent, creates a faeco-oral horizontal transmission. Host and parasite agent encounters being random. One should imagine a well-mixed population in terms of host agents, resources and also parasite agents.

There can be more than one population though. The Digital_Coevolution simulation allows to specify any number of discrete populations with any amount of resources per population. Population size cannot be set exactly, but is an emergent property of the simulation. Nevertheless it is strongly linked to the amount of resources that are simulated per population.
Several populations form a metapopulation when there is a distinct migration pattern between those populations. Therefore the Digital_Coevolution simulation allows to specify any type of migration pattern between populations, including source-sink dynamics. The migration structure can be specified for both the host and the parasite separately. This can be used to create complex metapopulation structures through which hosts and parasites, and hence also epidemics, can spread.

Within-agent dynamics

The within-agent dynamics are the heart of the Digital_Coevolution simulation. Here the rules for the life-history of each individual agent are defined. This is where each host agent takes shape as an individual.
The host agents that make up the Digital_Coevolution simulation can best be imagined as small, mobile, aquatic filter feeders, for example zooplankters like *Daphnia* or *Cyclops*. Thinking of the agents as zooplankters can help with imagining some of the biology that has been implemented. Host agents are simply called hosts in the code of the simulation.
During the simulation each agent moves through different life-history stages, struggles for resources and is potentially exposed to parasites. Those parasites are horizontally transmitted, passively dispersed, obligate parasites, which might best be imagined as either bacterial or fungal pathogens, for example microsporidians. Parasite agents are simply called parasites in the code of the simulation.
The following part gives an overview over the different functions, called "dynamics functions", that each govern the behaviour of a different life-history trait. Each of the "dynamics functions" is evaluated once every time-step.

Size

Size is one of the central parameters of the Digital_Coevolution simulation. It is based on the observation that most organisms grow during their lifetime, and that their size defines some of their interactions with their biotic and abiotic environment. In the Digital_Coevolution simulation size influences many life-history traits but especially the interaction with food resources. Larger agents consume more resources (eat more) but also have a larger metabolic requirement (need more food). This influences several other parameters that are linked either proportionally or with absolute thresholds against the resource availability within an agent. The feeding rate of an agent and the amount of ingested parasite spores are directly linked. A larger agent will be exposed to more parasite spores per time-step than a small one as it has a higher feeding rate.
A larger organism will, on the other hand, also be able to produce more offspring per time-step, as it can acquire resources faster and the production of one offspring always costs the same amount of resources.

Size is not implemented as one single "dynamics function", but as a trait (of an agent) that is influencing many "dynamics functions". Size itself is influenced in the metabolism function when resources are allocated between different life-history processes.
Digital_Coevolution agents are born with size = 1 and then invest resources into growth depending on their age. The younger an individual is, the more resources it will invest into growth as opposed to reproduction. The proportion of resources invested into growth declines negatively proportional to numerator 2 with age ($2/AGE$). Numerator 2 causes the resources that are available at the first time-step for growth to be counted twice. This was done to create a pronounced infant growth boost and a juvenile phase without reproduction (see section "Host metabolism").
Size can also be switched off by setting the host.size parameter to "OFF", which then skips the growth step and keeps all individuals at size = 1 at all times. This is especially useful for comparing results to other, less complex simulations.

```
# Depending on the age of the host, resources are funneled more towards reproduction
# or more towards growth.
if(host.size == "ON"){
    Host[Alive.Hosts$Is.Alive, Size :=
           Size + (pmin(Reproduction.Allocation * Size, Resource.Work) * (2 / Age))]
}
```

In the Digital_Coevolution simulation only host agents have an individual size. Parasite agents have no size themselves but the infection within the host agent does. The infection size within the host agent is directly linearly limited by the host agent size. Smaller individuals can only maintain smaller infections.

```
Host[Alive.Hosts$Is.Alive, Infection.Size := Infection.Size * infection.growth.factor]

Host[Alive.Hosts$Is.Alive, Infection.Size :=  pmin(Infection.Size, Size)]
```

Age

All things age. The agents in the Digital_Coevolution simulation should be no exception. There is an age variable built into both host and parasite agents. It is a simple, integer counter whose value increases by 1 each time-step. There is an old-age cut-off, creating senescence, implemented as a threshold value for both hosts and parasite agents. Once an agent reaches this threshold, it is culled and removed from the simulation.

```
# Takes the Age vector and adds 1 everywhere, then truncates by a pre-set threshold.
Host[Alive.Hosts$Is.Alive, Age := Age + 1L]

Host[Alive.Hosts$Is.Alive & Age > age.threshold.host, Alive := 0L]

Parasite[Alive.Parasites$Is.Alive, Age := Age + 1L]

Parasite[Alive.Parasites$Is.Alive & Age > age.threshold.parasite, Alive := 0L]
```

Within host resources

All host agents compete with other host agents for a shared pool of finite resources that renew every time-step. Each host agent will obtain a discrete amount of resources that is dependent on the size of the host agent and the size of the host population respective to the available food resources.
The resource allocation within each agent works as follows. The resource pool within each agent is split in three "containers". One container is for incoming resources gained through foraging ("Resource.In"), one for resources that are currently available for the metabolism ("Resource.Work") and one is for storing excess resources in a fat storage (Resource.Have). Each time-step, resources are moved between these containers.

The resources dynamics take place over three time-steps. Resources that are gained by a host agent through foraging will remain in the "Resource.In" container for the duration of one time-step and only become metabolically available in the next time-step. After one time-step the resources are moved to the "Resource.Work" container. If only few resources are coming in, host agents also have the ability to recruit resources from the "Resource.Have" container to the "Resource.Work" container. From the "Resource.Work" container, resources are distributed to different life-history traits according to the agents resource allocation scheme. Any resources that are left over in the "Resource.Work" container at the beginning of the next time-step are moved into the "Resource.Have" container. The amount of resources that can be stored in the "Resource.Have" container grows with the squared size of the host agent. If excess resources are available but the "Resource.Have" container is full, those resources are lost. Whether or not there are resources remaining in the "Resource.Work" container at the beginning of a time-step is dependent on the settings for the life-history allocations. These settings are explained a bit further below in the "Host metabolism" section.
This three time-step process of resource movement to different containers within host agents was implemented to reflect natural resource dynamics, where consumed resources are not immediately metabolically available but have to be digested first. The combination of the time-lag and the "Resource.Have" container also allows host agents to cope with of variability in resource availability to some extent.
The last step in the function is a starvation process, which culls all individuals that have less than a pre-set metabolic resource threshold available in the "Resource.Work" container.

```
# This moves the leftover resources from the resource.work container to the
# resource.have container. So leftover metabolic resources to the "fat" storage.
Host[Alive.Hosts$Is.Alive, Resource.Have := Resource.Have + Resource.Work]

# The resource.have, so basically the fat content, is size dependant.
# As larger individuals should be able to proportionally store more fat, a qubic
# increase instead of a linear increase with size might be applicable.
Host[Alive.Hosts$Is.Alive, Resource.Have := pmin(Resource.Have, (Size ^ 2))]

# Then the resource.work (metabolic resource) gets filled by the resource.in container
# of the last time-step, creating a delay between feeding and availability of energy.
# This part also lets hosts that get less external resources use some stored resources
```

```
# to fill resource.work.
Host[Alive.Hosts$Is.Alive, Resource.Work :=
       Resource.In + pmin((round(Size) - Resource.In), Resource.Have)]
Host[Alive.Hosts$Is.Alive, Resource.Have :=
       Resource.Have - pmin((round(Size) - Resource.In), Resource.Have)]

# This part removes individuals that have run out of resources (starving).
Host[Alive.Hosts$Is.Alive & Resource.Work < resource.threshold.host, Alive := 0]
```

Host metabolism

The agents that make up the Digital_Coevolution simulation have internal resource allocation trade-offs among fitness correlated life-history traits (Stearns 1980, Noordwijk and Jong 1986, Perrin and Sibly 1993).
Each host has a metabolic resource availability that is a result of the ingested resources a time-step earlier and resources available from the fat storage (see "within host resources" section above). These resources now need to be allocated to different life-history traits. In the case of the Digital_Coevolution simulation, there are three life-history traits that can receive resources.
The "Immune.State" is a variable that summarizes the overall maintenance state of the host including its immune system. It is the baseline investment that has to be made by every organism into maintenance of tissue, body temperature, innate immune response etc. Basically, this is the metabolic cost of being alive. It is the first trait that any agent will have to invest resources in. The amount of resources that need to be invested is dependent on the hosts size and an adjustable immune allocation proportion. The value of "Immune.State" resets every time-step, which means that the host will have to invest resources to immunity every time-step anew.
Next the host will invest into reproduction ("Reproduction.Have"). The amount of resources invested into reproduction is determined by the hosts size and the reproduction allocation setting and changes with age. The same formula that has modified the investment into growth of an individual is also applied here. Younger individuals invest less into reproduction with proportion $2/AGE$. Numerator 2 creates a negative value in the "Reproduction.Have" container at the first time-step that has to be neutralized by subsequent investment. This means that the first batch of offspring costs more resources than following batches. Reproduction allocation was implemented this way to create a distinct juvenile phase without reproduction (see section "Size").
Finally, a host will invest into growth. Depending on the age of the individual, a certain proportion is invested into growth (increasing the value of the "Size" parameter). This proportion is declining proportional to age, so that younger individuals invest proportionally more into growth instead of reproduction (see section "Size"). This size step can also be switched off to create a simulation with agents of constant size.

```
# This part drains resource.work and assigns to immune.state, refreshes every round,
# so no immune build-up.
Host[Alive.Hosts$Is.Alive, Immune.State :=
       pmin(Immune.Allocation * Size, Resource.Work)]

Host[Alive.Hosts$Is.Alive, Resource.Work :=
       Resource.Work - pmin(Immune.Allocation * Size, Resource.Work)]

Host[Alive.Hosts$Is.Alive, Reproduction.Have :=
       Reproduction.Have +
       (pmin(Reproduction.Allocation * Size, Resource.Work) * (1 - (2 / Age)))]

# This part drains resource.work and moves to reproduction.have.
# Depending on the age of the host, resources are funneled more towards reproduction
# or more towards growth.
if(host.size == "ON"){
  Host[Alive.Hosts$Is.Alive, Size :=
         Size + (pmin(Reproduction.Allocation * Size, Resource.Work) * (2 / Age))]
}
```

```
Host[Alive.Hosts$Is.Alive, Resource.Work :=
       Resource.Work - pmin(Reproduction.Allocation * Size, Resource.Work)]
```

Host reproduction

Each host individual first must accumulate resources until it reaches an adjustable threshold in order to mature. After it has reached this threshold, it reproduces as a iteroparous clonal income breeder (Sibly and Calow 1984, Houston et al. 2007, Stephens et al. 2009). The number of offspring produced per agent per batch is dependent on the resources invested in reproduction and an adjustable multiplicative reproduction factor. It is also dependent on the size of the individual, as larger individuals can obtain more resources which then can be invested into reproduction.

Reproduction is implemented in the Digital_Coevolution simulation by first calculating the identities of host agents that have matured. Those host agents are then copied multiple times. The amount of times they are copied is dependent on the amount of resources each individual had accumulated in excess of the maturation threshold and the multiplicative reproduction factor.
The life-history traits are heritable on the "genotype-level" are directly copied from the parent agent. Those include the "Reproduction.Allocation" and "Immune.Allocation" traits that detail which proportion of resources is invested into reproduction and immune system respectively. It includes the immune genotype of the agent, the last immune state of the parent agent and the population assignment of the parent agent. Other traits of an agent are dependent on its individual background and are not heritable. Those traits are reset to the starting conditions in the offspring. Those traits include age, size, infection state and all resource traits ("Resource.In", "Resource.Have", "Resource.Work"). The parent agent remains in the population and looses the resources it has used for reproduction.

```
# Calculate the number of reproducing individuals and how many offspring each
# individual produces.
reproducing.hosts <- Host[, rep(
  .I[Alive.Hosts$Is.Alive & Reproduction.Have > reproduction.threshold.host],
  times =
    round(
      Host[Alive.Hosts$Is.Alive &
             Reproduction.Have > reproduction.threshold.host, Reproduction.Have] -
        reproduction.threshold.host) * reproduction.factor.host)
  )
  ]

# If there are reproducing individuals, proceed, otherwise, skip.
# Open if statement.
if (Host[reproducing.hosts, .N] > 0) {

  # In the subset of empty (dead) host rows, fill host rows according to the
  # reproducing hosts.
  Host[Host[, .I[!Alive.Hosts$Is.Alive][seq(Host[reproducing.hosts, .N])]],
         `:=`
       (Alive = 1L,
         Host.Replicate = Host[reproducing.hosts, Host.Replicate],
         Host.Population = Host[reproducing.hosts, Host.Population],
         Host.Infection.Genotype = NA,
         Age = 1,
         Resource.Have = 1,
         Reproduction.Allocation = Host[reproducing.hosts, Reproduction.Allocation],
         Immune.Allocation = Host[reproducing.hosts, Immune.Allocation],
         Immune.Genotype = Host[reproducing.hosts, Immune.Genotype],
         Resource.In = 1,
         Resource.Work = 0,
         Reproduction.Have = 0,
         Immune.State = Host[reproducing.hosts, Immune.State],
```

```
          Infection.State = 0L,
          Infection.Size = 0,
          Parasite.Resources = 0,
          Host.TempID = NA,
          Size = 1,
          Host.Generation = (Host[reproducing.hosts, Host.Generation] + 1L),
          Origin = Host[reproducing.hosts, Host.Population])
          ]

  # Withdraw the resources used for reproduction.
  Host[Alive.Hosts$Is.Alive & Reproduction.Have > reproduction.threshold.host,
       Reproduction.Have :=
         Reproduction.Have - (round(Reproduction.Have) - reproduction.threshold.host)]

# Close if statement.
}
```

Disease and virulence

Infections have a size within a host agent ("Infection.Size"). The size of an infection defines how many resources are withdrawn from the hosts resource budget ("Resource.Work") towards the infection. Virulence is hence implemented as a resource loss for the host agent. Infections all start at size = 1 and then grow at an exponential rate, mimicking natural infections with micro-parasites. Infections can grow until they fill the space that is available, which is linked to the size of the host individual. The resources that are withdrawn by the infection are used by the parasite to reproduce. Resources stolen by the parasite are stored in the host level variable "Parasite.Resources".

```
# First the infection matures.
Host[Alive.Hosts$Is.Alive, Infection.Size := Infection.Size * infection.growth.factor]

Host[Alive.Hosts$Is.Alive, Infection.Size := pmin(Infection.Size, Size)]

# Then the parasite draws resources dependant on infection size.
Host[Alive.Hosts$Is.Alive, Parasite.Resources :=
       Parasite.Resources + pmin(Resource.Work, Infection.Size * virulence)]

Host[Alive.Hosts$Is.Alive, Resource.Work :=
       Resource.Work - pmin(Resource.Work, Infection.Size * virulence)]
```

Parasite reproduction

The reproduction of the parasite is implemented in the same way as that of the host. Parasite reproduction only happens within infected host agents, with the infection acting as the reproducing unit. This means that the parasite infection within the host agent must accumulate resources until an adjustable threshold value is reached and the infection matures. After that, resources stolen from the host can be used for iteroparous clonal income breeding. As larger infections accumulate resources faster, larger infections will produce more offspring parasite agents per clutch.

The implementation of parasite reproduction is identical to that of host reproduction. The amount of resources that an infection has stolen from the host individual is checked against the parasite maturation threshold to calculate which infection will have a reproduction event. Then the according number of parasite agents is created. The newly created parasites inherit the genotype of their "parent" infection as well as the population assignment. Other values like age are reset to starting values.

```
# The first step selects the subset of the host population that is infected and where
# the infections have accumulated enough resources to reproduce. And then multiplies
# this subset by the reproduction factor.
reproducing.parasites <-
  Host[
```

```
    ,
    rep(
      .I[Alive.Hosts$Is.Alive & Parasite.Resources > reproduction.threshold.parasite],
       times = round(
         Host[Alive.Hosts$Is.Alive &
                Parasite.Resources > reproduction.threshold.parasite,
              Parasite.Resources] - reproduction.threshold.parasite) *
         reproduction.factor.parasite
      )
    ]

# If there are reproducing individuals, proceed, otherwise, skip.
# Open if statement.
if (Host[reproducing.parasites, .N] > 0) {

  # This part selects the "dead" rows in the parasite data.table, then calcualtes how
  # many offsping are produced, and updates the values from the corresponding
  # host data.table
  Parasite[
    Parasite[, .I[!Alive.Parasites$Is.Alive][seq(Host[reproducing.parasites, .N])]],
             `:=`
             (Alive = 1L,
               Parasite.Replicate = Host[reproducing.parasites, Host.Replicate],
               Parasite.Population = Host[reproducing.parasites, Host.Population],
               Parasite.Infection.Genotype =
                 Host[reproducing.parasites, Host.Infection.Genotype],
               Attack.Host.TempID = NA,
               Attack.Host.Genotype = NA,
               Success.Parasite.Infection.Genotype = NA,
               Ingested = 0,
               Age = 1L)
               ]

  Host[Alive.Hosts$Is.Alive & Parasite.Resources > reproduction.threshold.parasite,
       Parasite.Resources :=
         Parasite.Resources -
         (round(Parasite.Resources) - reproduction.threshold.parasite)]

  # Uptdate parasite alive vector
  set(Alive.Parasites, j = "Is.Alive", value = Parasite[, Alive == 1])

# Close if statement.
}
```

Between-agent dynamics

A great benefit of agent-based simulations is that between-agent dynamics can be observed at the interaction level. This simulation was not written with natural populations in mind, but with laboratory populations. A population in the Digital_Coevolution simulation can be imagined as a glass jar with some medium and a few zooplankters in it. Environmental conditions in those glass jars are fairly stable and feeding happens once a day. With limited amounts of medium in a glass jar available to filter-feeding zooplankters, a large enough population can potentially "over-filter" the available media volume. This also affects interactions with parasite agents, as they are suspended in the same media volume as the food particles. As parasites are passively ingested during food consumption, resource dynamics and parasite epidemiological dynamics are closely linked.

Host resources population wide

The host resources are implemented as a batch of resources per population that renew every time-step. The host agents then compete for a portion of those resources. Resources are partitioned out according to a transformed Poisson distribution. Each time-step there is a random Poisson vector calculated with one element per host. As the amount of resources available per host individual is dependent on the population size and the summed-up filtering capacity of the host individuals, the expected mean of the distribution is proportional to the fraction of total filtering capacity per individual. This fraction is calculated by dividing the host individuals size, which equals its filtering capacity, by the sum of sizes over the population, and that fraction is multiplied by the size of the individual again. This means that the population size is density-dependent with a soft upper border.

```
# This part takes all the available resources and redistributes them according to
# host size. It adjusts the resources by relative filtering capacity if the population
# gets less than maximal food.
Host[Alive.Hosts$Is.Alive, Resource.In :=
       (rpois(
         n = .N,
         lambda = (Size * min(1, (resources.host[Host.Population[1]] / sum(Size)))) *
           resource.grain) / resource.grain),
     by = list(Host.Population, Host.Replicate)]

# This part restricts resource.in to Size + 10% in order to avoid unrealistic
# overfeeding.
Host[Alive.Hosts$Is.Alive, Resource.In := pmin(Resource.In, (Size * 1.1))]
```

Infection model

The infection model in the Digital_Coevolution simulation is implemented as a genotype-by-genotype look-up table that defines the infection affinity between the host and parasite genotypes. This can be specified so that the resulting infection model ranges anywhere between a gene-for-gene model to a matching-allele type infection model. Most simulation models of host-parasite coevolution are run as a perfect matching-allele model with complete parasite specificity.

```
# Infection dynamic parameters, row is host, column is parasite.
infection.table <- matrix((1 - parasite.specificity),
                          nrow = host.genotypes,
                          ncol = parasite.genotypes)

diag(infection.table) <- 1
```

Host exposure to parasites

This is the most important part of the simulation. Here the host agents are exposed to parasites. The exposure within populations is completely random, i.e. there is no host seeking behaviour by the parasite and no parasite avoidance by the host. The simulated transmission is faeco-oral, so the transmission is linked to the feeding rate. As the parasite population can be both larger or smaller than the host population, the number of parasites each host is exposed to varies largely. Host agents consume parasite

agents which are then removed from the population, independent of the infection state of the host or the infection success of the parasite agent. Infection success is dose and genotype dependent. Co-infections are not possible.

The implementation starts by calculating for each parasite agent whether or not it has been ingested by a host agent, which is dependent on the size of the host population relative to the available resources (is there over- or underfeeding?). Then each ingested parasite agent gets assigned the specific host agent by which it has been ingested. The assignment is dependent on the size of an individual host agent relative to the summed-up size of the host population. This is because the feeding rate of host agents is directly linked to host agents size. One host agent can ingest several parasite agents. Not all ingestion events lead to a successful infection. Only one parasite can infect the host. The probability for each of the ingested parasite agents to successfully infect is weighted by both the number of parasite agents with a certain genotype and by the matching of the host and parasite genotype in the infection table. With perfect specificity only parasite agents that match the host agents genotype have a chance at establishing an infection. If the specificity of the infection system is not perfect, then there is a dose response in the probability of the infection identity. Finally, host agents that have already been infected in a prior time-step cannot acquire a new infection (vaccination effect), but will still consume and remove parasite agents from the population.

```
# This part sets an identifier to be used later on.
Host[Alive.Hosts$Is.Alive, Host.TempID := 1 : .N]

Host[! Alive.Hosts$Is.Alive, Host.TempID := NA]

  set(Parasite, j = "Ingested", value = 0)
  set(Parasite, j = "Attack.Host.TempID", value = NA)
  set(Parasite, j = "Attack.Host.Genotype", value = NA)
  set(Parasite, j = "Success.Parasite.Infection.Genotype", value = NA)

# This part removes the identifier from those hosts that already are infected,
# creating a vaccination effect.
Host[Alive.Hosts$Is.Alive & Infection.State == 1, Host.TempID := NA]

# This part does assign to each parasite spore if it has been ingested by a host,
# dependant on resource availability, host and parasite population size.
Parasite[Alive.Parasites$Is.Alive,
           Ingested :=
           rbinom(n = .N,
                  size = 1,
                  prob = min(1,
                             (sum(
                               Host[Alive.Hosts$Is.Alive &
                                    Host.Population == Parasite.Population[1] &
                                    Host.Replicate == Parasite.Replicate[1], Size]) /
                              resources.host[Parasite.Population[1]]))),
         by = list(Parasite.Population, Parasite.Replicate)]

# Updating the Alive.Parasites$Is.Ingested vector for faster look-ups.
Alive.Parasites[Alive.Parasites$Is.Alive,
                Is.Ingested := Parasite[Alive.Parasites$Is.Alive, Ingested == 1]]

# This part does assign to each parasite spore the host.tempID it has been
# ingested by.
Parasite[Alive.Parasites$Is.Ingested,
         Attack.Host.TempID :=
           base:::sample(x = Host[Alive.Hosts$Is.Alive &
                                    Host.Population == Parasite.Population[1] &
                                    Host.Replicate == Parasite.Replicate[1],
                                  Host.TempID],
```

```
                          size = .N,
                          replace = TRUE,
                          prob = c(
                            Host[Alive.Hosts$Is.Alive &
                                   Host.Population == Parasite.Population[1] &
                                   Host.Replicate == Parasite.Replicate[1], Size] *
                              (.N / resources.host[Parasite.Population[1]]) *
                              min(1, resources.host[Parasite.Population[1]] /
                                    sum(
                                      Host[Alive.Hosts$Is.Alive &
                                           Host.Population == Parasite.Population[1] &
                                           Host.Replicate == Parasite.Replicate[1],
                                           Size]))
                          )
                        ),
         by = list(Parasite.Population, Parasite.Replicate)]

# This part assigns to each parasite the host genotype it has been ingested by.
Parasite[Alive.Parasites$Is.Ingested, Attack.Host.Genotype :=
           Host[Alive.Hosts$Is.Alive][Attack.Host.TempID, Immune.Genotype]]

# This part calculates which parasite successfully infects, takes into account the
# relative abundance of ingested parasite spores and their genetic specificity.
Parasite[Alive.Parasites$Is.Ingested & ! is.na(Attack.Host.TempID),
         Success.Parasite.Infection.Genotype :=
           sample(c(NA, Parasite.Infection.Genotype),
                  size = 1,
                  prob = c(1,
                           infection.table[Attack.Host.Genotype[1],
                                           Parasite.Infection.Genotype]),
                  replace = TRUE),
         by = list(Attack.Host.TempID, Parasite.Population, Parasite.Replicate)]

# And the last thing to do is to assign the infection genotype back to the host.
infected.hosts <- unique(Parasite[Alive.Parasites$Is.Ingested &
                                    ! is.na(Success.Parasite.Infection.Genotype)],
                         by = "Attack.Host.TempID")$Attack.Host.TempID

infected.hosts.infection.genotypes <-
  unique(Parasite[Alive.Parasites$Is.Ingested &
                    ! is.na(Success.Parasite.Infection.Genotype)],
         by = "Attack.Host.TempID")$Success.Parasite.Infection.Genotype

Host[Host[, .I[Alive.Hosts$Is.Alive]][infected.hosts],
     Host.Infection.Genotype := infected.hosts.infection.genotypes]

# And the very last thing is to update the infection status of the host that got
# assigned a infection.genotype.
Host[Alive.Hosts$Is.Alive & ! is.na(Host.Infection.Genotype) & Infection.State == 0,
     c("Infection.Size", "Infection.State") := 1]

# And kill the parasites that have been ingested.
Parasite[Alive.Parasites$Is.Ingested, Alive := 0]
```

Host resources simulation wide

Resources indirectly control the size of a simulated population. Therefore, the resources can be used together with a migration pattern to create any type of metapopulation. The resources control the

size of the subpopulations, while the migration pattern controls the movement of agents between the subpopulations. The resources that are available per subpopulation are implemented as an integer vector, where the number of vector elements is the number of populations that are going to be simulated, and the value of the vector elements is the size of the simulated population. This way it is possible to simulate a metapopulation that consists of subpopulations of different population sizes.

```
# Here at the same time the number of populations and their size are set. For each
# population define explicitly the amount of resources it receives daily.
resources.host <- c(500, 50, 50)

###
number.populations.host <- length(resources.host)

populations.host <- c(1 : number.populations.host)

starting.population.sizes.host <- ceiling(resources.host / 2)

###
number.populations.parasite< - number.populations.host

populations.parasite <- c(1 : number.populations.parasite)

starting.population.sizes.parasite <- ceiling(resources.host * 10)
```

Between population dynamics

A central feature of the Digital_Coevolution simulation is that it can simulate several interconnected subpopulations of both host and parasite agents. When a host agent population is imagined as a glass jar full of zooplankton, a metapopulation can be imagined as the climate chamber in which the glass jars are standing. Depending on the migration schedule, host or parasite individuals are moved between different glass jars in the climate chamber. Accordingly can host and parasite agents in the Digital_Coevolution simulation migrate between different host and parasite subpopulations. This migration takes place once every time-step (once a day). Like in an experiment, migration can happen at different rates between different populations. Last but not least, those subpopulations can also be of different sizes, i.e. with different amounts of food resources available to them.

Hierarchical metapopulation

Population inter-connectivity is implemented as a population-by-population fully crossed look-up table that defines how many individuals move from one population another, both for the host and the parasite. This allows for the implementation of arbitrary metapopulation structures ranging from only lightly connected large metapopulations to a network of closely knitted subpopulations. Combined with the explicit resource availability per population, this can also create asymmetrical metapopulations, as for example a mainland-island model.
Most importantly, by enabling the user to specify the metapopulation structure for both the host and the parasite separately, situations can be created where either the host or the parasite migrates substantially more. This makes the Digital_Coevolution simulation a valuable tool to investigate questions of host-parasite coevolution in metapopulations.

```
# Migration matrix for the host.
migration.matrix.host <- matrix((host.migration / number.populations.host),
                                nrow = number.populations.host,
                                ncol = number.populations.host)

diag(migration.matrix.host) <-
  diag(migration.matrix.host) + (1 - host.migration)

# Migration matrix for the parasite.
migration.matrix.parasite <-
  matrix((parasite.migration / number.populations.parasite),
                 nrow = number.populations.parasite,
                 ncol = number.populations.parasite)

diag(migration.matrix.parasite) <-
  diag(migration.matrix.parasite) + (1 - parasite.migration)
```

Parasite migration

The “parasite.migration” function allows the parasite to migrate between different populations. It does so simply by picking a random subset of parasites from one population and randomly assigning it to a new population. Migration is thus completely random and independent of any parasite properties. The migration matrix defines the probabilities for each possible migration path.

```
# Parasite migration
# Here a random subset of each parasite population is picked
# and moved to another population. Size of subset and probability
# of destination population are defined in the migration matrix.
parasite.migration.function <- function(){
  Parasite[Alive.Parasites$Is.Alive,
           Parasite.Population :=
             base:::sample(1 : number.populations.parasite,
                                              size = .N,
                                              prob =
                             migration.matrix.parasite[Parasite.Population[1], ],
```

```
                            replace = TRUE),
          by = list(Parasite.Population, Parasite.Replicate)]
  }
```

Host migration

The host migration works analogously to the parasite migration. A subset of the host population is randomly assigned to a new population according to the probabilities from the migration matrix.

```
# Host migration
# Here a random subset of each host population is picked
# and moved to another population. Size of subset and probability
# of destination population are defined in the migration matrix.
host.migration.function <- function(){
  Host[Alive.Hosts$Is.Alive,
       Host.Population :=
         base:::sample(1 : number.populations.host,
                       size = .N,
                       prob = migration.matrix.host[Host.Population[1], ],
                       replace = TRUE),
       by = list(Host.Population, Host.Replicate)]
  }
```

Part Two: Implementation of the Digital_Coevolution simulation

Overview

The Digital_Coevolution simulation has two logical areas. The most important part of the simulation describes all the rules for the behaviour within and between individuals. This "biological logic" is mostly implemented in the functions and parameters that have been described in part one of this document. The second area of the simulation is the implementation of the simulation itself. It specifies how an agent actually is simulated, how time is implemented and how results are recorded. This is described in part two of this document.

The core structure of the Digital_Coevolution simulation is the data.table (Dowle and Srinivasan 2019). A data.table is also a data.frame but much more powerful for data handling. Each row of the data.table contains all information on one individual agent. Each column of the data.table represents one trait, for example the immune genotype, the reproduction allocation proportion, the amount of resources already saved for reproduction or the infection state.
The whole data.table therefore comprises the state of all traits of all agents at a certain point in time. Every time-step each column is updated according to the "dynamics functions" that have been described in part one of this document. This utilizes R's vectorized structure whereby the same operation can be applied to all elements of a column simultaneously in a computationally efficient way. At specified intervals a copy of the data.table is saved, allowing for the analysis of host-parasite coevolutionary dynamics over time.

The Digital_Coevolution simulation currently consists of three to four interdependent R scripts.
The "Digital_Coevolution_Dynamics_Functions.R" script, where the within- and between-agent dynamics as well as the time-forward simulation process is defined.
The "Digital_Coevolution_Parameterspace.R" script, where different within and between-agent parameters can be set. Those parameters for example define the characteristics of agents that are simulated.
The "Digital_Coevolution_Run.R" script, which coordinates the simulation.
The "Digital_Coevolution_User.R" script, which is where the user interacts with and runs the Digital_Coevolution simulation.

In order to be able to run the Digital_Coevolution simulation, simply download or copy the relevant scripts from the GitHub repository to your computer. On the GitHub repository there are detailed instructions on how to install and use the Digital_Coevolution simulation on either normal personal computers or on high performance cluster computers.
The Digital_Coevolution simulation is implemented in R, so a R installation is needed as well. The newest R version for your system can be downloaded from CRAN.

Setup of the Digital_Coevolution simulation

Data structure

The Digital_Coevolution simulation is implemented in a way that each host or parasite agent is represented as a vector. Each element of the vector corresponds to a certain trait of that agent, as for example age or genotype. Those vectors are then stacked to make up the populations. Populations are represented as tables, with each individual agent corresponding to one row. Reading out a column of this table gives a vector that represents the variable/trait state of the whole population at a certain time-point. This allows for the manipulation of all individuals simultaneously if necessary.

As agents can reproduce and also die, the population size will vary throughout the simulation. This means that the number of rows in the data.table will vary throughout the simulation as well. Having data.tables (or worse, data.frames) with varying numbers of rows within a loop is problematic in R, as R automatically creates a copy of the old data.table whenever the number of rows changes. This can slow the code down compared to loops where a data.table of constant size is used. Therefore, the data.table used for the Digital_Coevolution simulation is created to be able to accommodate the largest possible population size of the simulation. Adding a variable that records whether or not the agent that is residing in this row is currently alive or dead allows to subset this larger data.table to just the alive population. When new agents are created (born) they are assigned to a row that has either been empty or that contains an individual with the marker "dead". The rows in the data.table are thus constantly recycled.

This circumvents the growing data.table problem and speeds up the simulation. See below for the first two rows of the host data.table at the initializing state of the simulation.

The Digital_Coevolution simulation relies heavily on the data.table package and hence its syntax (Dowle and Srinivasan 2019). The data.table syntax is analogous to SQL and has three elements: Where, order by / select, and update group / by.

```
##    Alive Host.Replicate Time Host.Population Host.Infection.Genotype Age
## 1:     1              1    0               1                    <NA>   1
## 2:     1              1    0               1                    <NA>   1
##    Resource.Have Reproduction.Allocation Immune.Allocation Immune.Genotype
## 1:             1                   0.35              0.35               1
## 2:             1                   0.35              0.35               5
##    Resource.In Resource.Work Reproduction.Have Immune.State Infection.State
## 1:           1             0                0            0               0
## 2:           1             0                0            0               0
##    Infection.Size Parasite.Resources Host.TempID Size Host.Generation Origin
## 1:              0                  0          0    1               1      1
## 2:              0                  0          0    1               1      1
```

The creation of the first agents

The simulation process is initialized by the creation of the host and parasite agents. This is done by a call to the "individual.creator" function that is defined in the "Digital_Coevolution_Dynamics_Functions" script. It uses variables that are set in either the "Digital_Coevolution_User" script or the "Digital_Coevolution_Parameterspace" script. That includes variables that will regularly be changed, like the number of replicates, and variables that will be changed less often, such as the allocation of resources to reproduction. The "individual.creator" function is automatically invoked when the simulation is started. The "individual.creator" function creates two data.tables, one for the host and one for the parasite.

```
#########################################################################
# Individual.creator.function. Here the host and parasite agents are
# created upon initializing of the simulation.

# Open function.
individual.creator.function <- function(){

# In this step all the data.table space that might be necessary is preallocated
# in an attempt to speed up the simulation.
  preallocation.margin <- 10

  parasite.margin <- 1

  preallocation.length <-
    sum(starting.population.sizes.host * replicates * preallocation.margin)

  preallocation.parasite <-
    sum(starting.population.sizes.parasite * replicates *
          preallocation.margin * parasite.margin)


  if(exists("Host")) {rm(Host, pos = ".GlobalEnv")}

# Here the data.table of the host is created.
  Host <<- data.table(
    Alive = integer(preallocation.length),
    Host.Replicate = integer(preallocation.length),
    Time = integer(preallocation.length),
    Host.Population = integer(preallocation.length),
    Host.Infection.Genotype = factor(NA, levels = c(1 : parasite.genotypes)),
```

```
  Age = integer(preallocation.length),
  Resource.Have = numeric(preallocation.length),
  Reproduction.Allocation = numeric(preallocation.length),
  Immune.Allocation = numeric(preallocation.length),
  Immune.Genotype = factor(sample(c(1 : host.genotypes),
                                  size = preallocation.length,
                                  prob = rep(1 / host.genotypes,host.genotypes),
                                  replace = T),
                           levels = c(1 : host.genotypes)),
  Resource.In = numeric(preallocation.length),
  Resource.Work = numeric(preallocation.length),
  Reproduction.Have = numeric(preallocation.length),
  Immune.State = numeric(preallocation.length),
  Infection.State = integer(preallocation.length),
  Infection.Size = numeric(preallocation.length),
  Parasite.Resources = numeric(preallocation.length),
  Host.TempID = integer(preallocation.length),
  Size = numeric(preallocation.length),
  Host.Generation = integer(preallocation.length),
  Origin = integer(preallocation.length)
)

# Here the starter populations are initialized with values.
# Replicate
Host[, Host.Replicate :=
       c(rep(1 : replicates,
             each = sum(starting.population.sizes.host)),
         integer(preallocation.length - sum(starting.population.sizes.host) *
                   replicates))]

# Alive variable, 1 = alive, 0 = not alive.
Host[Host.Replicate != 0, Alive := 1L]

# Population
Host[Alive == 1, Host.Population :=
       as.integer(rep(1 : number.populations.host,
                      times = starting.population.sizes.host)),
     by = Host.Replicate]

# Age
Host[Alive == 1, Age := 1L]

# Resource.Have
Host[Alive == 1, Resource.Have := 1]

# Resource.In, meaning that the starters start with a full belly.
Host[Alive == 1, Resource.In := 1]

# Reproduction.Allocation
Host[Alive == 1, Reproduction.Allocation := reproduction.allocation]

# Immune.Allocation
Host[Alive == 1, Immune.Allocation := immune.allocation]

# Immune.Genotype, has been set when initializing, needs to be cleaned.
Host[Alive == !1, Immune.Genotype := NA]
```

```
# Infection.Genotype
Host[, Host.Infection.Genotype := NA]

# Size
Host[Alive == 1, Size := 1]

# Host.Generation
Host[Alive == 1, Host.Generation := 1L]

# Origin
  Host[Alive == 1, Origin := Host.Population]

###
# Here a vector that just contains if a host is alive or not is started
# to circumvent all the look-ups.
if(exists("Alive.Host")) {rm(Alive.Host, pos = ".GlobalEnv")}

Alive.Hosts <<- data.table(Is.Alive = Host[, Alive == 1])

# Here the parasite data.table structure is initialized.
if(exists("Parasite")) {rm(Parasite, pos = ".GlobalEnv")}


Parasite <<- data.table(
  Alive = integer(preallocation.parasite),
  Parasite.Replicate = integer(preallocation.parasite),
  Time = integer(preallocation.parasite),
  Parasite.Population = integer(preallocation.parasite),

  Parasite.Infection.Genotype =
    factor(sample(c(1 : parasite.genotypes),
                  size = preallocation.parasite,
                  prob = rep(1 / parasite.genotypes, parasite.genotypes), replace = T),
           levels = c(1 : parasite.genotypes)),

  Attack.Host.TempID = integer(preallocation.parasite),
  Attack.Host.Genotype = factor(sample(c(1 : parasite.genotypes),
                  size = preallocation.parasite,
                  prob = rep(1 / parasite.genotypes, parasite.genotypes), replace = T),
           levels = c(1 : parasite.genotypes)),
  Success.Parasite.Infection.Genotype = factor(NA,
                                               levels = c(1 : parasite.genotypes)),
  Ingested = integer(preallocation.parasite),
  Age = integer(preallocation.parasite)
)


# Here the starter populations for the parasite are filled with values.
Parasite[, Parasite.Replicate :=
           c(rep(1 : replicates, each = sum(starting.population.sizes.parasite)),
             integer(preallocation.parasite -
                       sum(starting.population.sizes.parasite) * replicates))]

# Alive variable, 1 = alive, 0 = not alive.
Parasite[Parasite.Replicate != 0, Alive := 1L]

# Population
Parasite[Alive == 1,
```

```
            Parasite.Population :=
              as.integer(rep(1 : number.populations.parasite,
                             times = starting.population.sizes.parasite)),
            by = Parasite.Replicate]

  # Age
  Parasite[Alive == 1, Age := 1L]

  # Infection.Genotype, has been set when initializing, needs to be cleaned.
  Parasite[Alive != 1, Parasite.Infection.Genotype := NA]

  # Attack.Host.Rownumber
  Parasite[, Attack.Host.TempID := NA]

  # Attack.Host.Genotype
  Parasite[, Attack.Host.Genotype := NA]

  # Here a vector that just contains if a parasite is alive or not is started
  # to circumvent all the look-ups.
  if(exists("Alive.Parasite")) {rm(Alive.Parasite, pos = ".GlobalEnv")}

  Alive.Parasites <<- data.table(Is.Alive = Parasite[, Alive == 1],
                                 Is.Ingested = FALSE)

# Close the individual.creator.function.
}
```

Within-agent parameters

When the first agents are created they need to be assigned trait values. Most within-agent parameters, like the resource allocation proportion or the old-age threshold, can be set in the "Digital_Coevolution_Parameterspace.R" script. They define the life-history of the agents that are simulated. Depending on the settings in the "Digital_Coevolution_Parameterspace.R" script, the agents can for example be parametrized to behave more like a k-strategist or more like a r-strategist. If one for example changes the old-age threshold and the reproductive allocation, that can vastly change the nature of the agents simulated. This high amount of control over the behaviour of the agents allows the Digital_Coevolution simulation to be fine-tuned to match a variety of empirical systems.

```
# Parameter space
# Internal dynamic parameters
# Host internal dynamic parameters
host.size <- "OFF"
age.threshold.host <- 30
resource.threshold.host <- 0.2
reproduction.threshold.host <- 2
reproduction.factor.host <- 4
reproduction.allocation <- 0.35
immune.allocation <- 0.35

# Parasite internal dynamic parameters
parasite.genotypes <- host.genotypes
age.threshold.parasite <- 60
reproduction.threshold.parasite <- 2
reproduction.factor.parasite <- 23

# This factor gives the per time-tep growth of an infection in percent.
infection.growth.factor <- 1.15
```

Between-agent parameters

The Digital_Coevolution simulation allows for a high level of control over the between-agent parameters like the infection system or the migration pattern. This level of control is reached by using look-up tables that specify the outcome of interactions at the per-combination level.
The infection table contains for all combinations of host and parasite genotypes the infection specificity of that interaction. A value of 0 leads to no infection, a value of 1 means full infection potential. Values in between are a reduced infection specificity. A value of 0.5, for example, means that this host and parasite genotype combination has half the infection probability of a fully specific combination. The infection table is implemented as a full matrix, so it allows for the specification of asymmetrical infection patterns, while the default setting is symmetrical.
The migration matrix is implemented similarly. A migration setting of 0 will create no migrants, and a setting of 1 will randomly re-assign all individuals to any of the subpopulations each time-step. Settings between 0 and 1 will randomly assign a proportion of individuals to a new subpopulation. The default migration matrix is also symmetrical. This means all subpopulations are connected with migration of equal strength.

```
# Infection table, row is host, column is parasite.
infection.table <- matrix((1 - parasite.specificity),
                          nrow = host.genotypes,
                          ncol = parasite.genotypes)
diag(infection.table) <- 1

###
# Migration matrix
migration.matrix.host <- matrix((host.migration / number.populations.host),
                                nrow = number.populations.host,
                                ncol = number.populations.host)

diag(migration.matrix.host) <-
  diag(migration.matrix.host) + (1 - host.migration)

migration.matrix.parasite <-
  matrix((parasite.migration / number.populations.parasite),
                               nrow = number.populations.parasite,
                               ncol = number.populations.parasite)

diag(migration.matrix.parasite) <-
  diag(migration.matrix.parasite) + (1 - parasite.migration)
```

Dead or alive

The agents of the Digital_Coevolution simulation can be, much like biological individuals, either dead or alive.
The data.table that contains the agents of the Digital_Coevolution simulation always has the same size for technical reasons (see section “data structure”). This means it contains both currently alive as well as dead individuals. Dead individuals are simply agents from past time-steps whose rows have not yet been recycled. In order to optimize the simulation in terms of computational efficiency, and because dead agents should not interact with alive ones, calculations should only be done with alive agents. This is made possible by using the “Alive” variable, which can take the values 0 = dead and 1 = alive, to subset the data.table. As a subsetting operation always is a comparison of the queried value to all values in the vector that is to be subsetted, it can become computationally demanding. The data.table package already has internal optimisation to make this type of look-up as fast as possible (Dowle and Srinivasan 2019), yet further optimization in execution speed are possible when the comparison operation is circumvented altogether. This is especially valuable when a comparison is done repeatedly.
In the Digital_Coevolution simulation this is done by creating an external vector (or to be more precise, a one-column data.table) that has the same length as the data.table that contains the agents. This external vector contains the information indicating which agent is currently alive in logical form (alive = TRUE). It is essentially a copy of the trait “Alive” in logical form. This vector can now be used to subset

the data.table that contains the agents to only the agents which are alive, without the necessity for a comparison operation.

```
# Here a vector that just contains if a host is alive or not is started to circumvent
# all the look-ups.
Alive.Hosts <<- data.table(Is.Alive = Host[, Alive == 1])

# The Alive.Hosts vector is updated by the set function after a host has been born
# or died.
set(Alive.Hosts, j = "Is.Alive", value = Host[, Alive == 1])

# In the host migration function the Alive.Hosts vector is used to subset the
# data.table to only the alive hosts.
host.migration.function <- function(){
  Host[Alive.Hosts$Is.Alive,
       Host.Population :=
         base:::sample(1 : number.populations.host,
                       size = .N,
                       prob = migration.matrix.host[Host.Population[1], ],
                       replace = TRUE),
       by = list(Host.Population, Host.Replicate)]
  }
```

The simulation process

The dynamics wrapper

The dynamics wrapper is the heart of the Digital_Coevolution simulation. All of the "biological logic" that was explained in part one of this document results in functions. Functions in R are collections of code that are defined under a common name and will be executed once the function is called. In the Digital_Coevolution simulation each of the distinct within- and between-agent dynamics that were explained in part one is defined as a separate function. These functions are called "dynamics functions" throughout the document.
Having all the different rules for dynamic behaviour of the agents defined as separate functions does allow to invoke them separately. That is important, as they will be executed consecutively in the simulation. The order in which the functions are invoked can have a big influence on the "biological logic" of the Digital_Coevolution simulation. As an example, the timing of migration is important. Does migration of parasite agents happen after the parasite reproduces but before there is a new round of hosts being exposed to parasites? Or will the parasites migrate after an exposure event? This simple question of timing will influence whether coevolution has a more local component or a more global component.
Still, every time-step all the functions that govern the behaviour of the agents have to be called once in order to complete the full life-cycle of an agent. To achieve that, and to ensure consistency in the order of function calling, they are wrapped in another function that is called "dynamics.wrapper". All this "dynamics.wrapper" does, is simply calling the "dynamics functions" one after another. One call of the "dynamics.wrapper" function is one time-step in the simulation. Changing the order of the functions in the "dynamics.wrapper" is a simple yet powerful way to change the behaviour of the Digital_Coevolution simulation.

```
# Here all the functions defined above are combined into a wrapper function.
# One call is one time-step. The order of the functions can have an influence on
# the dynamics of the simulation.
dynamics.wrapper <- function(){
  time.function()
  senescence.function()
  host.resource.function()
  infection.function()
  host.exposure.function()
  parasite.reproduction.function()
  metabolism.function()
  host.reproduction.function()
  host.migration.function()
  parasite.migration.function()
}
```

Obtaining Results

A key feature of the Digital_Coevolution simulation is that the dynamic behaviour of each individual host and parasite agent can be examined at each time-step of the simulation. This allows coevolutionary dynamics to be analysed with great temporal resolution.

There is a piece of code in the "Digital_Coevolution_Run.R" script that saves a complete copy of the host and parasite agent population to an output file at specified intervals. This means that a complete snapshot of every single agent with all its internal states like reproduction, infection and resources, gets saved.
The "saving.interval" setting in the "Digital_Coevolution_Run.R" script allows to control the frequency of this data extraction process in time-steps. With a value of 1 every time-step is saved, otherwise every N-th time-step. For long simulations it is worth to consider if saving every time-step is necessary or if less data is sufficient as well.

In the "Digital_Coevolution_Run.R" script there is also an option that allows to get a summarized report instead of a raw copy of the host and parasite agent states. That summary includes the number of host or parasite individuals per genotype per sub- and metapopulation and the number of infections. It lacks some of the details of the raw data, but is a much more convenient data set.

```
# Because it is a time-forward simulation, it is necessary to loop through the
# time-steps. This is done by calling the dynamics.wrapper function within a loop.

for(i in 1 : duration.days){
  dynamics.wrapper()
  # result saving
  if(i %in%
     c(1, seq(from = saving.intervall, to = duration.days, by = saving.intervall))){
    if(raw.results){
      fwrite(Host[Alive.Hosts$Is.Alive], file =
               paste(result.file.location,
                     result.file.name,
                     "_Host_",
                     run.date,
                     "_raw_",
                     ".csv",
                     sep = ""), append = TRUE)
      fwrite(Parasite[Alive.Parasites$Is.Alive], file =
               paste(result.file.location,
                     result.file.name,
                     "_Parasite_",
                     run.date,
                     "_raw_",
                     ".csv",
                     sep = ""), append = TRUE)
    }
# This part is only invoked if the option for a summarized report is enabled.
    if(summarized.results) {
      temp.data.host <- copy(Host)
      temp.data.host[, Virulence := virulence]
      temp.data.host[, Popsize := host.populations[1]]
      temp.data.host[, Random.Drift := random.drift]
      temp.data.host[, Parasite.Connection := parasite.migration]
      temp.data.host[, Host.Connection := host.migration]
      temp.data.host[, Host.Time := Time]

      temp.data.host[, Host.Number.Individuals := .N,
                     by = list(Host.Time, Host.Replicate, Host.Population,
                               Immune.Genotype, Virulence, Popsize, Parasite.Connection,
                               Host.Connection)]
      temp.data.host[, Host.Population.Size := .N,
                     by = list(Host.Time, Host.Replicate, Host.Population, Virulence,
                               Popsize, Parasite.Connection, Host.Connection)]
      temp.data.host[, Host.Number.Individuals.Between := .N,
                     by = list(Host.Time, Host.Replicate, Immune.Genotype, Virulence,
                               Popsize, Parasite.Connection, Host.Connection)]
      temp.data.host[, Host.Population.Size.Between := .N,
                     by = list(Host.Time, Host.Replicate, Virulence, Popsize,
                               Parasite.Connection, Host.Connection)]
      temp.data.host[, Epidemic.Size.Within := sum(Infection.State),
                     by = list(Host.Time, Host.Replicate, Host.Population,
                               Immune.Genotype, Virulence, Popsize, Parasite.Connection,
                               Host.Connection)]
      temp.data.host[, Epidemic.Size.Total := sum(Infection.State),
                     by = list(Host.Time, Host.Replicate, Host.Population, Virulence,
                               Popsize, Parasite.Connection, Host.Connection)]

      #########
```

```
    temp.data.parasite <- copy(Parasite)
    temp.data.parasite[, Virulence := virulence]
    temp.data.parasite[, Popsize := host.populations[1]]
    temp.data.parasite[, Random.Drift := random.drift]
    temp.data.parasite[, Parasite.Connection := parasite.migration]
    temp.data.parasite[, Host.Connection := host.migration]
    temp.data.parasite[, Parasite.Time := Time]

    temp.data.parasite[, Parasite.Number.Individuals := .N,
                  by = list(Parasite.Time, Parasite.Replicate,
                            Parasite.Population, Parasite.Infection.Genotype,
                            Virulence, Popsize, Parasite.Connection,
                            Host.Connection)]
    temp.data.parasite[, Parasite.Population.Size := .N,
                  by = list(Parasite.Time, Parasite.Replicate, Parasite.Population,
                            Virulence, Popsize, Parasite.Connection,
                            Host.Connection)]
    temp.data.parasite[, Parasite.Number.Individuals.Between := .N,
                  by = list(Parasite.Time, Parasite.Replicate,
                            Parasite.Infection.Genotype, Virulence, Popsize,
                            Parasite.Connection, Host.Connection)]
    temp.data.parasite[, Parasite.Population.Size.Between := .N,
                  by = list(Parasite.Time, Parasite.Replicate, Virulence, Popsize,
                             Parasite.Connection, Host.Connection)]

    #########
    temp.data.host[, Total.Parasite.Number.Individuals :=
                     Epidemic.Size.Within +
                     temp.data.parasite[
                       Parasite.Replicate == Host.Replicate[1] &
                         Parasite.Population == Host.Population[1] &
                         Parasite.Time == Host.Time[1] &
                         Parasite.Infection.Genotype == Immune.Genotype[1], .N],
                   by = list(Host.Time, Host.Replicate, Host.Population,
                          Immune.Genotype, Virulence, Popsize, Parasite.Connection,
                          Host.Connection)]
    temp.data.host[, Total.Parasite.Population.Size :=
                     Epidemic.Size.Total +
                     temp.data.parasite[
                       Parasite.Replicate == Host.Replicate[1] &
                         Parasite.Population == Host.Population[1] &
                         Parasite.Time == Host.Time[1], .N],
                  by = list(Host.Time, Host.Replicate, Host.Population, Virulence,
                            Popsize, Parasite.Connection, Host.Connection)]

    temp.data.host[, Total.Parasite.Number.Individuals.Between :=
                     Epidemic.Size.Within +
                     temp.data.parasite[
                       Parasite.Replicate == Host.Replicate[1] &
                         Parasite.Population == Host.Population[1] &
                         Parasite.Time == Host.Time[1] &
                         Parasite.Infection.Genotype == Immune.Genotype[1], .N],
                  by = list(Host.Time, Host.Replicate, Immune.Genotype, Virulence,
                            Popsize, Parasite.Connection, Host.Connection)]
    temp.data.host[, Total.Parasite.Population.Size.Between :=
                     Epidemic.Size.Total +
                     temp.data.parasite[
```

```
                        Parasite.Replicate == Host.Replicate[1] &
                          Parasite.Population == Host.Population[1] &
                          Parasite.Time == Host.Time[1], .N],
                    by = list(Host.Time, Host.Replicate, Virulence, Popsize,
                              Parasite.Connection, Host.Connection)]

      #########
      fwrite(
        unique(temp.data.host[,
                          list(Host.Time, Host.Replicate, Host.Population,
                              Immune.Genotype, Virulence, Popsize, Random.Drift,
                              Parasite.Connection, Host.Connection,
                              Host.Number.Individuals, Host.Population.Size,
                              Host.Number.Individuals.Between,
                              Host.Population.Size.Between, Epidemic.Size.Within,
                              Epidemic.Size.Total, Total.Parasite.Number.Individuals,
                              Total.Parasite.Population.Size,
                              Total.Parasite.Number.Individuals.Between,
                              Total.Parasite.Population.Size.Between, Origin)]),
        file = paste(
          result.file.location,
          result.file.name,
          "_Host_",
          run.date,
          "_summarized_",
          ".csv", sep = ""), append = TRUE)

      fwrite(
        unique(temp.data.parasite[,
                              list(Parasite.Time, Parasite.Replicate,
                                    Parasite.Population, Parasite.Infection.Genotype,
                                    Virulence, Popsize, Random.Drift,
                                    Parasite.Connection, Host.Connection,
                                    Parasite.Number.Individuals,
                                    Parasite.Population.Size,
                                    Parasite.Number.Individuals.Between,
                                    Parasite.Population.Size.Between)]),
        file = paste(
          result.file.location,
          result.file.name,
          "_Parasite_",
          run.date,
          "_summarized_",
          ".csv", sep = ""), append = TRUE)

    }
  }
}
```

Time-stamp

Because the Digital_Coevolution simulation is a time-forward simulation and the result file will contain many agent-states from different time-points, each agent needs a time-stamp as an identifier. This is implemented as an integer counter that increases its value each time-step. It is particularly useful for filtering the results.

```
# Time-stamp, counts the number of time-steps that the simulation
# has been looped over.
```

```
Host[, Time := Time + 1L]
Parasite[, Time := Time + 1L]
```

Time-forward simulation

The simulation is implemented as a time-forward simulation. This means that the simulation loops through discrete time-steps. Each time-step, all dynamics functions are called once and the host and parasite agent data.tables are updated. Then the next time-step is calculated with the updated data.table. This is necessary technically, as the simulation contains a number of stochastic elements. This process is similar to the creation of a Markov-chain.
Time-forward simulations also allow for a very fine-grained analysis of temporal dynamics of host and parasite coevolution, as every time-step the status of the simulation can be fully examined. This is especially valuable for systems where at least part of the dynamics have a time-shift property.
The time-forward property of the Digital_Coevolution is implemented as a for-loop. A for-loop in R calls its arguments for N times consecutively. The number of times over which the for-loop iterates is the number of time-steps that the simulation is run for. This is essentially what runs the Digital_Coevolution simulation.

```
# Now the simulation is run.
# Because it is a time-forward simulation it is necessary to loop through
# the time-steps.
for(i in 1 : duration.days){
  dynamics.wrapper()
}
```

Conclusions

The Digital_Coevolution simulation is a novel agent-based simulation that implements detailed life-histories of host and parasite agents. It can be used to test hypotheses and create theoretical expectations for natural systems. It is positioned to be close to a digital organism, insofar as the agents have detailed, individual life-histories. Compared to many other digital organisms, it lacks the self-modificability, but it includes emergent properties that have not been programmed and are a result of individual interactions. Such emergent properties can lead to unanticipated yet insightful results (Lehman et al. 2018, 2019).

Agent-based simulations and digital organisms are widely used in a diverse array of research topics, such as the origin of life (C G et al. 2017), evolution in small populations (LaBar and Adami 2016, 2017), metapopulations (Fortuna et al. 2013, Boëte et al. 2019), Red Queen Hypothesis (Kidner and Moritz 2015), diversity (Zaman et al. 2011) and mutational robustness (Lenski et al. 1999). They are not only used in biology but have, for example, also been applied in crowd control and economics (Bonabeau 2002). Agent-based models should not be seen as an alternative to analytic models but as a complement (Gräbner et al. 2019). Researching the same questions with several different approaches will only increase the robustness of findings.

The Digital_Coevolution simulation is a work in progress and shows potential to be further modified. The modular approach with the "dynamics functions" enables that additional properties of host and parasite agent interactions can easily be included in the simulation without altering existing code.
A few future additions are already anticipated. Environmental conditions in subpopulations that affect host and parasite competitive rankings could be implemented to allow research on genotype-by-genotype-by-environment interactions in a metapopulation context. Allowing for variance in the inheritance of life-history traits, like reproduction allocation, would open the simulation to research on optimization of life-history traits under different conditions. Adding further agent types, for example predators, would vastly improve the realism of the simulation and allow for different type of interactions between agents like parasites with complex life-cycles (multi-host life-cycles) or predator prey dynamics.

The Digital_Coevolution simulation can not only be used for research purposes but also for teaching. It can showcase how individual level rules and behaviours can lead to population and ecosystem properties. As many students and researchers are already familiar with R, no new program or syntax has to be learned before the Digital_Coevolution simulation can be used. The approachability of R also means that advanced users can modify the simulation themselves, which makes the Digital_Coevolution simulation a highly flexible tool. Further integration with R towards a CRAN package and the implementation of a graphical user interface for parameter settings via the *shiny* package (Chang et al. 2020) could increase usability of the Digital_Coevolution simulation for research and teaching.

References

Allaire, J., Y. Xie, J. McPherson, J. Luraschi, K. Ushey, A. Atkins, H. Wickham, J. Cheng, W. Chang, and W. Iannone. 2020. R markdown: Dynamic Documents for R.

Boëte, C., M. Seston, and M. Legros. 2019. Strategies of host resistance to pathogens in spatially structured populations: An agent-based evaluation. Theoretical Population Biology 130:170–181.

Bonabeau, E. 2002. Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences 99:7280–7287.

Boots, M., and M. Mealor. 2007. Local Interactions Select for Lower Pathogen Infectivity. Science 315:1284–1286.

Boots, M., and A. Sasaki. 1999. "Small worlds" and the evolution of virulence: Infection occurs locally and at a distance. Proceedings of the Royal Society of London. Series B: Biological Sciences 266:1933–1938.

Box, G. 1976. Science and statistics. Journal of the American Statistical Association 71:791–799.

Box, G. 1979. Robustness in the Strategy of Scientific Model Building. Pages 201–236 *in* Robustness in Statistics. Elsevier.

C G, N., T. LaBar, A. Hintze, and C. Adami. 2017. Origin of life in a digital microcosm. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 375:20160350.

Chang, W., J. Cheng, J. Allaire, and Y. Xie. 2020. Shiny: Web Application Framework for R.

Clarke, B. 1976. The ecological genetics of host-parasite relationships. *in* Genetic aspects of host-parasite relationships. Blackwell, London.

Clarke, B. 1979. The Evolution of Genetic Diversity. Proceedings of the Royal Society B: Biological Sciences 205:453–474.

Dowle, M., and A. Srinivasan. 2019. Data.Table: Extension of 'data.Frame'.

Dünner, R. P. 2020. Digital_Coevolution: An agent-based simulation for host-parasite coevolution.

Fortuna, M. A., L. Zaman, A. P. Wagner, and C. Ofria. 2013. Evolving Digital Ecological Networks. PLoS Computational Biology 9:e1002928.

Gandon, S. 1996. Local adaptation and gene-for-gene coevolution in a metapopulation model. Proceedings of the Royal Society of London. Series B: Biological Sciences 263:1003–1009.

Gandon, S. 2002. Local adaptation and the geometry of host-parasite coevolution. Ecology Letters 5:246–256.

Gandon, S., and Y. Michalakis. 2002. Local adaptation, evolutionary potential and host-parasite coevolution: Interactions between migration, mutation, population size and generation time: Local adaptation and coevolution. Journal of Evolutionary Biology 15:451–462.

Gräbner, C., C. S. E. Bale, B. A. Furtado, B. Alvarez-Pereira, J. E. Gentile, H. Henderson, and F. Lipari. 2019. Getting the Best of Both Worlds? Developing Complementary Equation-Based and Agent-Based Models. Computational Economics 53:763–782.

Haldane, J. B., S. 1949. Disease and evolution. La Ricerca Scientifica Supplemento.

Haldane, J. B., S., and S. D. Jayakar. 1963. Polymorphism due to selection depending on the composition of a population. Journal of Genetics 3:318–323.

Hamilton, W. D. 1980. Sex versus Non-Sex versus Parasite. Oikos 35:282.

Hamilton, W. D., R. Axelrod, and R. Tanese. 1990. Sexual reproduction as an adaptation to resist parasites (a review). Proceedings of the National Academy of Sciences 87:3566–3573.

Houston, A. I., P. A. Stephens, I. L. Boyd, K. C. Harding, and J. M. McNamara. 2007. Capital or income breeding? A theoretical model of female reproductive strategies. Behavioral Ecology 18:241–250.

Jaenike, J. 1978. An Hypothesis to account for the maintenance of sex within populations. Evolutionary Theory:191–194.

Judson, O. P. 1995. Preserving genes; a model of the maintenance of genetic variation in a metapopulation under frequency-dependent selection. Genet. Res.:175–191.

Judson, O. P. 1997. A Model of Asexuality and Clonal Diversity: Cloning the Red Queen. Journal of Theoretical Biology 186:33–40.

Kidner, J., and R. F. A. Moritz. 2015. Host-parasite evolution in male-haploid hosts: An individual based network model. Evolutionary Ecology 29:93–105.

LaBar, T., and C. Adami. 2016. Different Evolutionary Paths to Complexity for Small and Large Populations of Digital Organisms. PLOS Computational Biology 12:e1005066.

LaBar, T., and C. Adami. 2017. Evolution of drift robustness in small populations. Nature Communications 8:1012.

Ladle, R. J., R. A. Johnstone, and O. P. Judson. 1993. Coevolutionary dynamics of sex in a metapopulation: Escaping the Red Queen. Proceedings of the Royal Society of London. Series B: Biological Sciences 253:155–160.

Lehman, J., J. Clune, and D. Misevic. 2018. The Surprising Creativity of Digital Evolution. Pages 55–56 *in* The 2018 Conference on Artificial Life. MIT Press, Tokyo, Japan.

Lehman, J., J. Clune, D. Misevic, C. Adami, L. Altenberg, J. Beaulieu, P. J. Bentley, S. Bernard, G. Beslon, D. M. Bryson, P. Chrabaszcz, N. Cheney, A. Cully, S. Doncieux, F. C. Dyer, K. O. Ellefsen, R. Feldt, S. Fischer, S. Forrest, A. Frénoy, C. Gagné, L. L. Goff, L. M. Grabowski, B. Hodjat, F. Hutter, L. Keller, C. Knibbe, P. Krcah, R. E. Lenski, H. Lipson, R. MacCurdy, C. Maestre, R. Miikkulainen, S. Mitri, D. E. Moriarty, J.-B. Mouret, A. Nguyen, C. Ofria, M. Parizeau, D. Parsons, R. T. Pennock, W. F. Punch, T. S. Ray, M. Schoenauer, E. Shulte, K. Sims, K. O. Stanley, F. Taddei, D. Tarapore, S. Thibault, W. Weimer, R. Watson, and J. Yosinski. 2019. The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities. arXiv:1803.03453 [cs].

Lenski, R. E., C. Ofria, T. C. Collier, and C. Adami. 1999. Genome complexity, robustness and genetic interactions in digital organisms. Nature 400:661–664.

Lion, S., and S. Gandon. 2015. Evolution of spatially structured host-parasite interactions. Journal of Evolutionary Biology 28:10–28.

McKinley, P., B. Cheng, C. Ofria, D. Knoester, B. Beckmann, and H. Goldsby. 2008. Harnessing Digital Evolution. Computer 41:54–63.

Noordwijk, A. J. van, and G. de Jong. 1986. Acquisition and Allocation of Resources: Their Influence on Variation in Life History Tactics. The American Naturalist 128:137–142.

Ofria, C., and C. O. Wilke. 2004. Avida; A software platform for research in computational evolutionary biology. Artificial Life:129–229.

Oosterhout, C., D. A. Joyce, S. M. Cummings, J. Blais, N. J. Barson, I. W. Ramnarine, R. S. Mohammed, N. Persad, and J. Cable. 2006. Balancing selection, random genetic drift, and genetic variation at the major histocompatibility complex in two wild populations of guppies (poecilia reticulata). Evolution 60:2562–2574.

Perrin, N., and R. M. Sibly. 1993. Dynamic Models of Energy Allocation and Investment. Annual Review of Ecology and Systematics:379–410.

Sibly, R., and P. Calow. 1984. Direct and absorption costing in the evolution of life cycles. Journal of Theoretical Biology 111:463–473.

Stearns, S. C. 1980. A New View of Life-History Evolution. Oikos 35:266.

Stephens, P. A., I. L. Boyd, J. M. McNamara, and A. I. Houston. 2009. Capital Breeding and Income Breeding: Their Meaning, Measurement, and Worth. Ecology 90:2057–2067.

Team, R. 2019. RStudio: Integrated Development Environment for R. RStudio, Inc., Boston, MA.

Team, R. C. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austia.

Xie, Y., J. Allaire, and G. Grolemund. 2018. R Markdown: The Definitive Guide. Chapman and Hall / CRC.

Zaman, L., S. Devangam, and C. Ofria. 2011. Rapid host-parasite coevolution drives the production and maintenance of diversity in digital organisms. Page 219 *in* Proceedings of the 13th annual conference on Genetic and evolutionary computation - GECCO '11. ACM Press, Dublin, Ireland.

Chapter 2

Negative frequency-dependent selection by parasites boosts genetic drift

Authors: Robert P. Dünner[1,2], Jukka Jokela[1,2].

Author affiliations: ETH Zurich, Institute for Integrative Biology[1], Eawag, Department of Aquatic Ecology[2].

Corresponding author: Robert P. Dünner; robert.duenner@eawag.ch, Eawag ECO BU-G-12, Uberlandstrasse 133, 8600 Dubendorf, Switzerland.

Keywords: Host-parasite coevolution, negative frequency-dependent selection, genetic drift, maintenance of genetic diversity, agent-based simulation.

Author contributions: RD and JJ conceived the ideas and designed the work, RD authored the simulation and collected the data, RD analysed the data, RD and JJ led the writing.

Acknowledgements: RD was supported by the adaptation to a changing environment center (ACE) at ETH. We would like to thank the HPC Team at ETH Zurich for providing the computational resources used in this study. We would like to thank Kyra Gähwiler, Heidi Käch, Hugo Mathe-Hubert and Cansu Cetin for helpful comments on the manuscript.

Data accessibility: Data will be made available through the Dryad Digital Repository upon publication.

Abstract

Theory predicts that negative frequency-dependent selection (NFDS) by parasites promotes maintenance of genetic diversity. We asked how general this prediction is in finite populations that experience high amounts of genetic drift. To address this question, we used an agent-based simulation model based on matching-alleles type host-parasite coevolution. Our results indicate that the presence of coevolving parasites erodes genetic diversity faster than drift, especially in small populations. It does this by increasing the extinction probabilities of both host and parasite genotypes. We propose that the additional variance in genotype frequencies caused by NFDS leads to the faster loss of genotypes. More specifically, in our simulations, parasite genotypes with low virulence tended to go extinct, while parasites with high virulence drove their matching host genotypes to extinction. In both cases this led to faster degradation of genetic diversity than without parasites under drift. These results help to explain the patterns of biodiversity found in small, isolated populations in the wild as well as the difficulties in finding experimental support for host-parasite coevolution in small laboratory populations.

Introduction

Genetic diversity is a phenomenon that fascinated generations of biologists. Already Darwin pointed out in *the Origin of Species* that natural selection will favour individuals (phenotypes and, if heritable, genotypes) that fit the environment well on cost of those who fit less well. Natural selection thus reduces genetic diversity (Fisher 1930). But genetic diversity is also degraded without natural selection. Random genetic drift, caused by stochastic variation in offspring number between genotypes, inevitably leads to extinction of genotypes (Wright 1931). Bremermann (1980) put this very well:

> "If no other forces were acting, most wild-type populations should have lost their diversity long ago..."

Coevolving parasites may be an important selective force that maintains genetic diversity. Negative frequency-dependent selection (NFDS) is known to promote genetic diversity and suggested to be common in host-parasite coevolution (Wright 1948, Haldane 1949, Haldane and Jayakar 1963, Clarke and O'Donald 1964, Gillespie 1975, Clarke 1976, 1979, Bremermann 1980, Levin 1988, Weeks and Hoffmann 2008, Zaman et al. 2011, Kurbalija Novičić et al. 2020). NFDS provides a selective advantage to host genotypes at low relative frequency and vice-versa. The "rare advantage" essentially reduces the probability of genotypes to go extinct. This hypothesis has also received support from some empirical studies, which show that heavily parasitized host populations have higher levels of genetic diversity (Gulland et al. 1993, King et al. 2011, Gsell et al. 2013, Dagan et al. 2013, Turko et al. 2018). Or similarly, that the genetic structure of a population under parasite pressure changes in accordance with NFDS (Lively 1992, Jokela et al. 2003, Hall et al. 2011, Paczesniak et al. 2014, Ignacio-Espinoza et al. 2020).

For parasites to maintain diversity, the selection pressure they exert on host populations needs to overcome drift. The strength of drift is inversely proportional to population size, so drift is a substantial selective force in small populations. Hence it is trivial that below a certain population size, the strength of drift will be stronger than any counteracting selection by parasites favouring rare genotypes. Notably, most of the empirical data comes from systems with very large population sizes, as for example the New Zealand mud snail (Lively 1992, Jokela et al. 2003, King et al. 2011, Dagan et al. 2013, Paczesniak et al. 2014), arthropods (Weeks and Hoffmann 2008), zooplankton (Turko et al. 2018), phytoplankton (Gsell et al. 2013), bacteria (Hall et al. 2011) or marine viruses (Ignacio-Espinoza et al. 2020). Those are all systems where drift can be considered negligible.

Our hypothesis is that in small populations, where drift is strong, the fluctuations in host genotype frequency caused by NFDS may actually contribute to the loss of genotypes by drift instead of preventing it.

We constructed an agent-based, discrete-time, time-forward, stochastic simulation to examine maintenance of genetic diversity by selection through parasites and by drift. We simulated small, finite populations of clonal host individuals and challenged them with genotype specific, clonal parasites with different levels of virulence. The agents were based on the resource allocation hypothesis, allowing for within agent trade-offs and population-wide ecology (e.g. density dependence) (Stearns 1980, Noordwijk and Jong 1986, Perrin and Sibly 1993). We followed genotype frequencies over time and explored the power of NFDS by parasites to maintain genetic diversity.

Our results showed that presence of a parasite does not necessarily maintain genetic diversity in small host populations. Interaction of drift and NFDS by parasites can lead to faster degradation of genetic diversity than under drift alone.

Methods

Simulation

We used the Digital_Coevolution simulation (version: v1.0-thesis) (Dünner 2020) to explore the influence of parasites on maintenance of genetic diversity.
The Digital_Coevolution simulation is an agent-based, discrete-time, time-forward, stochastic simulation written in R (Team 2020) that was created to explore negative frequency-dependent dynamics of host-parasite coevolution.
Agent-based simulations consist of individual entities (agents), which are interacting according to specified individual-level rules. Population-level properties then emerge from the sum of the individual-level interactions of the agents.
The main (host) agents of the Digital_Coevolution simulation are fairly detailed, governed by resource allocation trade-offs, with different energy-requiring metabolic compartments as well as distinct life histories (Stearns 1980, Noordwijk and Jong 1986, Perrin and Sibly 1993). Agents are born, consume resources, age, potentially get infected, reproduce and eventually die. The life-time of each individual can span numerous time-steps of the simulation.

Populations are created by simulating a group of several distinct host agents. Interactions between agents within these groups then give rise to the population level behaviour. Populations can differ in the resource availability per time-step. Within each population the agents have to compete for these limited resources. Chunks of resources are distributed among agents as discrete units, creating variance in the received resources. The average amount of resources per agent is inversely proportional to the number of agents present in the population, but can not reach infinitely small amounts. This creates density dependence in the availability of resources and a soft upper limit for population size. More resources lead to more agents being simulated, but the exact number of agents in a population is an emergent property of the simulation and can not be controlled directly.

Parasite agents are simulated in less detail than host agents. The only dynamic behaviour they exhibit independently from the host is migration. For all other life history traits the parasite agents obligatorily have to infect host agents. This behaviour is inspired by that of microparasites transmitted as passive spores, like microsporidians. Reproduction of parasites is set to take place within host agents, with constant, non-lethal shedding of new infective parasite agents.
The virulence of parasite agents is implemented as a fraction of within host-agent resources that are used by the parasite at each time-step. The host-agent then has less resources available for its own metabolic needs. This allows for the simulation of parasite agents with very different effects on hosts upon infection. Parasite agents with zero virulence, for example, do not withdraw any resources and hence have no effect on the host. Such parasite populations can be used to simulate baseline behaviour of host-agent populations under drift without any further manipulation of the simulation. Parasite agents with a low virulence of a few percent host resource loss lead to mild effects, only manifesting as a decrease in host-agent fecundity. If more resources are withdrawn by the parasite agent, infection by parasites can also be completely sterilizing, as hosts have to invest a certain amount of their internal energy budget into maintenance. So if the parasites absorb the remainder of resources, host agents survive but can not reproduce. The highest virulence settings create lethal parasites, as they quickly absorb the majority of resources and the host agents starve.

Genetic drift is implemented implicitly as a purely stochastic process created by differences in number of offspring between individual host or parasite agents of distinct genotypes. Genotypes of both host and parasite drift to extinction when the last agent of that genotype fails to reproduce. Drift then leads to changes in genotype frequencies and genetic diversity.

For more details on the Digital_Coevolution simulation, please refer to the detailed documentation in chapter 1 and the source code on the GitHub page .

Simulated scenarios

For this study, the Digital_Coevolution simulation was parametrised such that the host agents life-cycles resembled that of zooplankters like *Daphnia* or *Cyclops*. Before host agents started reproducing, they needed to accumulate resources to pass a maturation threshold. After they passed that threshold, they behaved as iteroparous clonal income breeders (Sibly and Calow 1984, Houston et al. 2007, Stephens et al. 2009). The maximum reachable life-span per host agent was set to 30 time-steps. Host agents could not

change in size as the agent growth functionality of the Digital_Coevolution simulation was switched off. We simulated five distinct genotypes of host agents. All host genotypes were set to have identical fitness and only differed in their susceptibility to parasite agents. In other words, they were identical except for the haploid susceptibility locus which had five alleles. The infection system was implemented as a matching-alleles system with perfect specificity. Each host genotype could only be infected by a parasite agent with the same (matching) allele at its infectivity locus.

We simulated scenarios with different levels of parasite virulence and different levels of host populations size.
We simulated five different levels of virulence, 0, 0.2, 0.3, 0.4 and 0.5. We noticed that simulating virulences below 0.2 always resulted in immediate loss of the parasite population as they could not withdraw enough resources for their own reproduction. Therefore, we did not simulate any virulence levels below 0.2 in this study. Likewise, simulating virulences above 0.5 always resulted in crashing of the host population, as these parasites were too virulent to be maintained sustainably in finite host populations. Virulence of 0.2 somewhat reduced the fecundity of host agents, but did not sterilize them. We refer to a virulence of 0.2 as a low virulence. Virulence of 0.5 strongly reduced fecundity of host agents and is referred to as a high virulence. Virulence 0.3 and 0.4 are referred to as low-intermediate and high-intermediate virulence respectively. We used the virulence level of 0 as a drift baseline.

The Digital_Coevolution simulation can be used to simulate an arbitrary number of populations which can be connected by a specific migration pattern. For this study we only simulated singular isolated populations.
We simulated six different resource levels, 25, 50, 100, 200, 400 and 800 units of resources, each of which resulted in a different host population size. The number of host agents within a population at a certain resource level was dependent on that resource level but ultimately an emergent property of the simulation. The realized population size depended on many factors, but especially on the virulence of the parasite. With a parasite agent virulence level of 0, host population sizes fluctuated at around 2.5 times the number of resource units. With a highly virulent parasite agent, host population size fluctuated at around 2 times the number of resource units. The population sizes that we simulated hence reached about 50, 100, 200, 400, 800 and 1600 individuals respectively.
The size of the parasite population was not directly dependent on the resource level, but on the host population size and the parasite virulence. It fluctuated at about 5 to 6 times the host population size. Hence, the average population sizes of the parasite that were simulated were about 275, 550, 1100, 2200, 4400 and 8800 parasite individuals respectively.

The strength of drift is inversely proportional to population size. Small populations (with fewer resources) were simulated for fewer time-steps than large populations. We adjusted the duration of the simulation in time-steps so that at the end of the simulation host populations with 0 virulence parasite agents (drift baseline) had, on average, lost all genetic diversity. This resulted in simulation lengths of 1250, 2500, 5000, 10000, 20000 and 40000 time-steps respectively for the six levels of resources.

In order to have the same amount of data points for all different resource levels, we subsampled the simulation at different time-steps. In shorter runs, the simulation was sampled at narrower time-step intervals than in longer runs, so that we would have 50 data points per replicate for each resource levels. This means we had a reporting window of 25, 50, 100, 200, 400 and 800 time-steps respectively for the six levels of resources.

We ran all combinations of virulence and resource levels, resulting in 30 sets of parameters. For each of those settings we ran 100 replicates.
Thus, in total our analysis spanned 150000 data points. At each of these datapoints we had information on the number of host and parasite agents as well as infection rate per genotype. From this we calculated further measures like genotype frequency and genetic diversity.

What should be kept in mind is that we did include no mechanism that could re-created diversity in our simulation, like for example mutation or re-introduction of host or parasite genotypes through migration. This means any extinction in our simulation was final. Our simulation best describes small well-isolated populations, where a significant rise of diversity by mutation or migration is unlikely. While such isolated small populations might not be the rule in the wild, they will certainly exist.

The simulations were run on the “Euler” high performance cluster computer at ETH Zurich on R version 3.6.0 linked against OpenBLAS available in module gcc/4.8.2. Euler is a hpc cluster running on Linux and uses the IBM load sharing facility (LSF).

Analyses

We were interested in the fate of genetic diversity within singular, isolated host populations over time, under the influence of different virulence levels of parasites. For measuring genetic diversity we used the number of genotypes present in a population.
We visualized the outcome of our simulations by plotting the number of genotypes present within a population over time. We calculated the average number of genotypes and a confidence interval assuming a normal distribution per time-step (Figure 1).
Reduction in host genetic diversity is a step-like process that progresses when host genotypes go extinct. Each level of genetic diversity can be seen as a discrete state (5 genotypes, 4 genotypes and so on). As our data was panel-observed, we do not know the exact time-points of each extinction event, but we do know the state of diversity of each population at every N-th time-step. We used multi-state continuous-time Markov-chains to model this process and estimate a matrix of transition intensities between states of genetic diversity. Transition intensities are related to hazards, as they are the probability of an event happening (a state transition) per unit of time. We used a matrix of all allowed transitions and a prior transition matrix estimated from the frequencies of each transition in our data. We then fitted the Markov-chain using maximum likelihood estimation available in the the *msm* package in R (Jackson 2011). We calculated separate transition-intensity matrices for each level of virulence.
This approach allowed us to compare the transition intensities between virulences by taking the ratio between the drift baseline and any other level of virulence. This "hazard ratio" compares the relative risk of a transition (the extinction of a genotype) happening when a parasite is coevolving with the host population to the risk of the same transition happening under drift alone without a parasite (Figure 2). A hazard ratio of 1 shows that there is no difference in transition intensity between drift and parasite. A hazard ratio higher than one shows that the transition is more likely to occur in the presence of a parasite than under drift alone.

Taken together, the visualisation of genetic diversity in Figure 1 and the hazard ratios from the Markov-chain in Figure 2 allowed us to infer on the maintenance of genetic diversity in coevolving host and parasite populations compared to host populations under drift.
The entire document was written in *rmarkdown* (Xie et al. 2018, Allaire et al. 2020), and all analyses were conducted in R (Team 2020) using RStudio (Team 2019) as IDE, *data.table* for data handling (Dowle and Srinivasan 2019) and *ggplot2* (Wickham 2016) and *RColorBrewer* (Neuwirth 2014) for visualisations.

Results

We found that for most levels of parasite virulence the genetic diversity of the host population degraded faster with coevolving parasites than under drift alone (Figure 1, any line against the red drift baseline). This suggests that in populations smaller than around 400 individuals genetic diversity is lost faster when drift and NFDS interact. Only at high levels of parasite virulence (0.4 and 0.5) in populations larger than 400 individuals was genetic diversity maintained (see for example virulence 0.4 at population size 800). Genetic diversity under drift always degrades to a value of 1, independent of population size. Taken together, this result suggests that parasites can, by adding fluctuations to host genotype frequencies, lower the genetic diversity below what would be expected under drift alone. It also suggests that parasites can nevertheless maintain genetic diversity in large populations when drift is low.

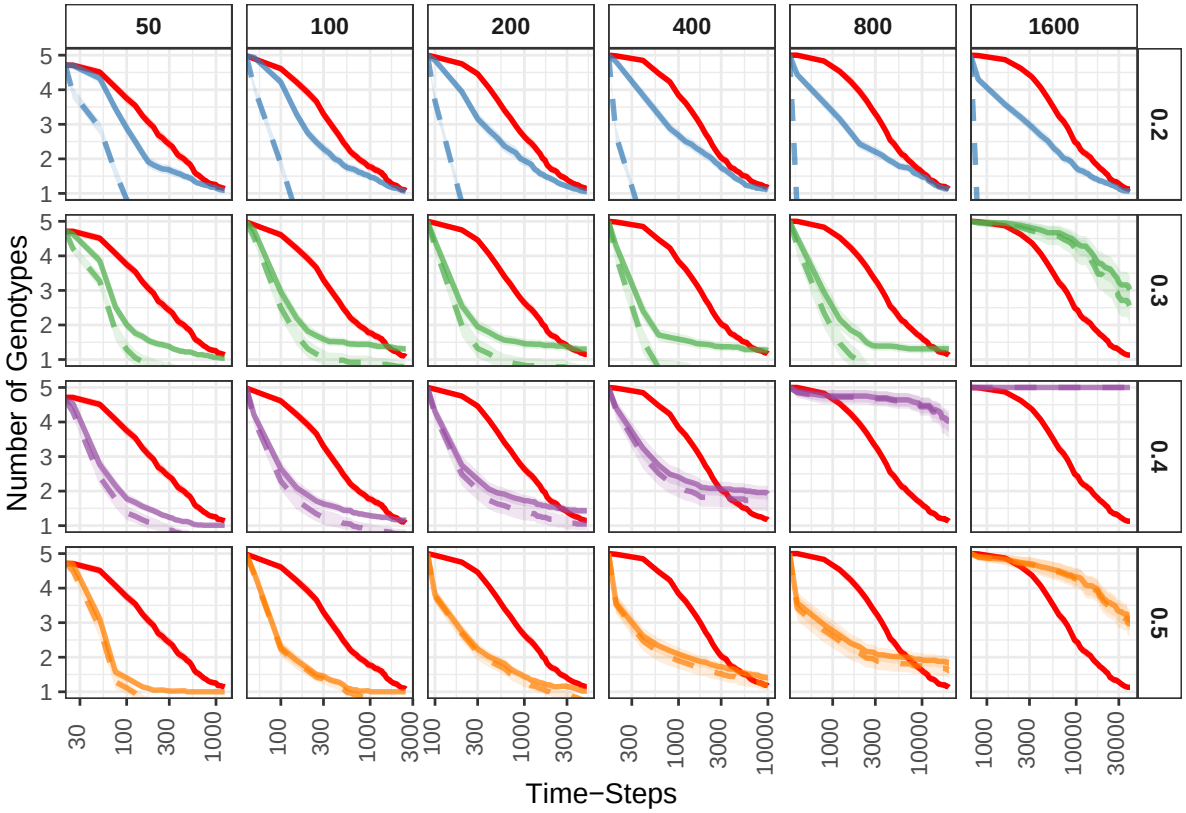

Figure 1: Number of genotypes in a population over time. Colours depict virulence, virulence = 0 is the drift baseline (red line). Solid lines are host genotypes, dashed lines are parasite genotypes. The shaded areas are a 95% confidence interval around the mean. Panels are split by approximate population size horizontally and by virulence levels vertically. Time-steps on the x-axis have been log10 transformed. The lower end of the y-axis has been cut off at 1 genotype, as that is the lowest possible number of genotypes before a population goes extinct.

The hazard ratios calculated from the Markov-chain analysis confirm these findings, as the hazard of losing a genotype was highest when a parasite was present in small populations (see Figure 2). Parasites especially increased the hazard of losing 1 or 2 genotypes (transition of 5:4 and 4:3 genotypes), pointing towards a fast initial decline of diversity when a parasite is present.

It also becomes apparent that not all levels of virulence interact with drift in the same manner. When parasite virulence was low (0.2) or low-intermediate (0.3), the hazard of losing 1 or 2 genotypes was always slightly higher than under drift alone without a parasite (Figure 2). The hazard of losing 3 or 4 genotypes under low virulence parasites was very similar to that under drift (hazard ratio very close to 1).

However, virulence levels high-intermediate (0.4) and high (0.5) showed a very different pattern. Both of these virulence levels had hazard ratios far greater than 1 for host populations smaller than about 400 individuals for all transitions. This means that the hazard of loosing all host genotypes was greatly increased when parasites were present. For small populations virulent parasites strongly degraded genetic

diversity. This effect quickly reversed with increasing host population size. For larger host populations virulent parasites decreased the hazard ratios below 1 for all transitions. This means they reduced the risk of loosing a genotype compared to drift. High-intermediate or high virulence coevolving parasites maintain genetic diversity in large host populations (see bottom two rows of Figure 2).

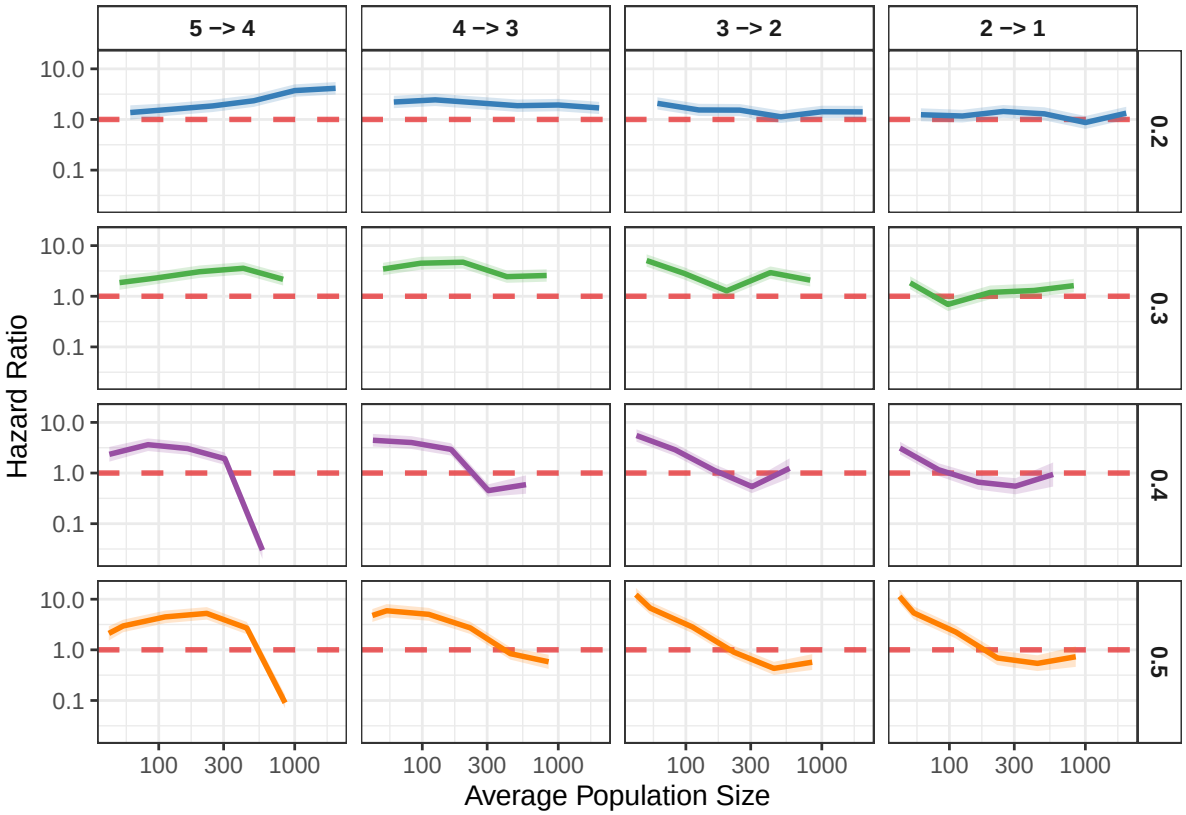

Figure 2: Plot of hazard ratios for the four transitions between states of genetic diversity (5 -> 4 = first genotype of the population lost, 4 -> 3 = second genotype lost, etc.). The hazard ratios compare transition intensity for each genotype loss under a coevolving parasite to the drift baseline. Different colours depict different virulence levels. Shaded areas show a 0.95 % confidence interval. The red dashed line is a hazard ratio of 1, depicting that the transition intensity with a parasite is the same as without a parasite (drift). A hazard ratio below 1 implies NFDS by parasites maintains genetic diversity, values above 1 imply that NFDS by parasites degrades genetic diversity. Panels are split by transition between levels of genetic diversity (horizontally) and virulence (vertically.) The x-axis shows the average realized host population sizes. The x-axis as well as the y-axis have been log10 transformed.

Discussion

Our results showed that in small populations genetic diversity degraded faster when coevolving parasites were present than under drift alone. We found that all simulated levels of parasite virulence increased the hazard of losing genotypes and consequently degraded genetic diversity. This was especially pronounced in small populations of below 400 host individuals, where drift is high. In larger populations, in our case starting with 400 host individuals per population, some levels of virulence lowered the hazard of losing genotypes and maintained a low level of genetic diversity. Only at a high-intermediate level of virulence (0.4) and only in the largest simulated population size of 1600 individuals, could parasites counteract the effect of drift immediately, reduce the hazard of losing genotypes and maintain genetic diversity (Figure 1 and 2).
These results highlight the importance of host population size in evaluating the maintenance of host genetic diversity by parasites.

Drift and NFDS by parasites are two processes that both create fluctuations in host genotype frequencies over time.
The direction of the fluctuations in host genotype frequencies under drift is completely random, while the direction under NFDS is determined by selection by the parasite. The strength of selection by parasites is independent of population size but determined by the genotype frequencies of both the host and the parasite. The strength of drift, on the other hand, is inversely proportional to population size. At a certain population size, the fluctuations in host genotype frequencies that are caused by drift and the fluctuations that are caused by NFDS will have a similar magnitude and can hence interfere with each other.

Interference between drift and NFDS may increase the risk of extinction of rare host genotypes. The fluctuations caused by NFDS bring host genotypes to very low frequencies much faster than drift alone. At low genotype frequencies random fluctuations caused by drift can substantially increase the risk of extinction. This effect is strongest when at low frequency the absolute population size is small as well. Population size per host genotype is critical in determining the strength of genetic drift. The number of individuals per genotype is lower in a population with a high genetic diversity than in a less diverse population of equal size. That is why this mechanism will be strongest in diverse populations and decrease both with increasing population size or decreasing genetic diversity.

Interference of NFDS and drift can also increase the risk of extinction of parasite genotypes instead of host genotypes. This could have even more severe consequences for the maintenance of genetic diversity. Only one parasite genotype needs to go extinct in order to lead to complete degradation of host genetic diversity. If one parasite genotype goes extinct, its matching host genotype is released from parasite selection, gaining an immediate selective advantage against the other remaining host genotypes. A host genotype that has been released from its parasite is hence likely to spread to fixation, which completely degrades the genetic diversity of the population. If more than one parasite genotype is lost by drift, then the host genetic diversity will again be shaped by drift alone.

If the interference of NFDS and drift has led to the extinction of one or more parasite genotypes, the resulting host population would be free of parasites at the end of the simulation. It seems that this is more likely to be happening with less virulent parasites, as we can see host populations without parasites in our results for low and low-intermediate virulences (0.2 and 0.3) (Figure 1).
If the interference of NFDS and drift rather leads to the extinction of host genotypes, then the resulting population would still have the remaining host genotypes and their matching parasites at the end of the simulation. Our results suggest that this is more likely to be happening with more virulent parasites, as we can see host and parasite populations for virulences 0.4 and 0.5 (Figure 1).

We have now proposed two mechanisms on how the interaction of drift with NFDS by parasites can degrade genetic diversity. One of them seems more likely for less virulent parasites, the other for more virulent parasites. It may be that at intermediate levels of virulence lies the "sweet spot" at which genetic diversity is maintained best. This is visible in our data starting at population size 400, and more obviously so at population size 800 and 1600, where genetic diversity is maintained best by a high-intermediate virulence parasite (0.4) (Figure 1). This is in accordance with results from the literature, for example in Rabajante et al. (2015) who found the most stable Red Queen Cycles for intermediate virulence parasites.

Our findings are in congruence with results of studies documenting that when population sizes of host and parasite are allowed to fluctuate, extinction of alleles or genotypes is faster (Gokhale et al. 2013, Schenk et al. 2020). Gokhale et al. (2013) implemented host-parasite coevolution using functions equivalent

to Lotka-Voltera equations. They found that allowing host population size to vary led to increased fixation rates of alleles. Schenk et al. (2020) compared different models of host-parasite coevolution with or without ecological interactions and stochastic elements. They found that when population size was allowed to vary, the time to extinction of genotypes was faster than in models with fixed population sizes. Both studies therefore implicitly confirm the interaction between drift and frequency-dependent selection on maintenance of genetic diversity. Our study further advances those findings by providing details on the possible mechanism of how this could be happening.

We conclude that our results show an important interaction between NFDS and drift, and that this interaction is detrimental for the maintenance of genetic diversity. This should be taken into account when hypotheses relying on NFDS are being tested with small populations that likely experience strong drift. Such populations might be found both in the wild for specialist species with small population ranges, as well as in laboratories, where micro- or mesocosms will only be able to sustain limited population sizes. Our results should also be taken into account for conservation, where remnant populations of endangered species are especially vulnerable to the diversity degrading effect of the interaction of drift and negative frequency-dependent selection.

References

Allaire, J., Y. Xie, J. McPherson, J. Luraschi, K. Ushey, A. Atkins, H. Wickham, J. Cheng, W. Chang, and W. Iannone. 2020. R markdown: Dynamic Documents for R.

Bremermann, H. 1980. Sex and polymorphism as strategies in host-pathogen interactions. Journal of Theoretical Biology:671–702.

Clarke, B. 1976. The ecological genetics of host-parasite relationships. *in* Genetic aspects of host-parasite relationships. Blackwell, London.

Clarke, B. 1979. The Evolution of Genetic Diversity. Proceedings of the Royal Society B: Biological Sciences 205:453–474.

Clarke, B., and P. O'Donald. 1964. Frequency-dependent selection. Heredity:6.

Dagan, Y., K. Liljeroos, J. Jokela, and F. Ben-Ami. 2013. Clonal diversity driven by parasitism in a freshwater snail. Journal of Evolutionary Biology 26:2509–2519.

Dowle, M., and A. Srinivasan. 2019. Data.Table: Extension of 'data.Frame'.

Dünner, R. P. 2020. Digital_Coevolution: An agent-based simulation for host-parasite coevolution.

Fisher, R. A. 1930. The genetical theory of natural selection. Dover Publications, Inc., New York.

Gillespie, J. H. 1975. Natural Selection for Resistance to Epidemics. Ecology 56:493–495.

Gokhale, C. S., A. Papkou, A. Traulsen, and H. Schulenburg. 2013. LotkaVolterra dynamics kills the Red Queen: Population size fluctuations and associated stochasticity dramatically change host-parasite coevolution. BMC Evolutionary Biology 13:254.

Gsell, A. S., L. N. de Senerpont Domis, K. J. Verhoeven, E. van Donk, and B. W. Ibelings. 2013. Chytrid epidemics may increase genetic diversity of a diatom spring-bloom. The ISME Journal 7:2057–2059.

Gulland, F. M. D., S. D. Albon, J. M. Pemberton, P. R. Moorcroft, and T. H. Clutton-Brock. 1993. Parasite-associated polymorphism in a cyclic ungulate population. Proceedings of the Royal Society of London. Series B: Biological Sciences 254:7–13.

Haldane, J. B., S. 1949. Disease and evolution. La Ricerca Scientifica Supplemento.

Haldane, J. B., S., and S. D. Jayakar. 1963. Polymorphism due to selection depending on the composition of a population. Journal of Genetics 3:318–323.

Hall, A. R., P. D. Scanlan, A. D. Morgan, and A. Buckling. 2011. Host-parasite coevolutionary arms races give way to fluctuating selection: Bacteria-phage coevolutionary dynamics. Ecology Letters 14:635–642.

Houston, A. I., P. A. Stephens, I. L. Boyd, K. C. Harding, and J. M. McNamara. 2007. Capital or income breeding? A theoretical model of female reproductive strategies. Behavioral Ecology 18:241–250.

Ignacio-Espinoza, J. C., N. A. Ahlgren, and J. A. Fuhrman. 2020. Long-term stability and Red Queen-like strain dynamics in marine viruses. Nature Microbiology 5:265–271.

Jackson, C. H. 2011. Multi-State Models for Panel Data: The **Msm** Package for *R*. Journal of Statistical Software 38.

Jokela, J., C. M. Lively, M. F. Dybdahl, and J. A. Fox. 2003. Genetic variation in sexual and clonal lineages of a freshwater snail: GENETIC VARIATION and MAINTENANCE OF SEX. Biological Journal of the Linnean Society 79:165–181.

King, K. C., J. Jokela, and C. M. Lively. 2011. Parasites, sex, and clonal diversity in natural snail populations. Evolution 65:1474–1481.

Kurbalija Novičić, Z., A. Sayadi, M. Jelić, and G. Arnqvist. 2020. Negative frequency dependent selection contributes to the maintenance of a global polymorphism in mitochondrial DNA. BMC Evolutionary Biology 20:20.

Levin, B. R. 1988. Frequency-dependent selection in bacterial populations. Philosophical Transactions of the Royal Society B: Biological Sciences:459–472.

Lively, C. M. 1992. Parthenogenesis in a freshwater snail: Reproductive assurance versus parasitic release. Evolution 46:907–913.

Neuwirth, E. 2014. RColorBrewer: ColorBrewer Palettes.

Noordwijk, A. J. van, and G. de Jong. 1986. Acquisition and Allocation of Resources: Their Influence on Variation in Life History Tactics. The American Naturalist 128:137–142.

Paczesniak, D., S. Adolfsson, K. Liljeroos, K. Klappert, C. M. Lively, and J. Jokela. 2014. Faster clonal turnover in high-infection habitats provides evidence for parasite-mediated selection. Journal of Evolutionary Biology 27:417–428.

Perrin, N., and R. M. Sibly. 1993. Dynamic Models of Energy Allocation and Investment. Annual Review of Ecology and Systematics:379–410.

Rabajante, J. F., J. M. Tubay, T. Uehara, S. Morita, D. Ebert, and J. Yoshimura. 2015. Red Queen dynamics in multi-host and multi-parasite interaction system. Scientific Reports 5:10004.

Schenk, H., H. Schulenburg, and A. Traulsen. 2020. How long do Red Queen dynamics survive under genetic drift? A comparative analysis of evolutionary and eco-evolutionary models. BMC Evolutionary Biology 20:8.

Sibly, R., and P. Calow. 1984. Direct and absorption costing in the evolution of life cycles. Journal of Theoretical Biology 111:463–473.

Stearns, S. C. 1980. A New View of Life-History Evolution. Oikos 35:266.

Stephens, P. A., I. L. Boyd, J. M. McNamara, and A. I. Houston. 2009. Capital Breeding and Income Breeding: Their Meaning, Measurement, and Worth. Ecology 90:2057–2067.

Team, R. 2019. RStudio: Integrated Development Environment for R. RStudio, Inc., Boston, MA.

Team, R. C. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austia.

Turko, P., C. Tellenbach, E. Keller, N. Tardent, B. Keller, P. Spaak, and J. Wolinska. 2018. Parasites driving host diversity: Incidence of disease correlated with *Daphnia* clonal turnover*. Evolution 72:619–629.

Weeks, A. R., and A. A. Hoffmann. 2008. Frequency-dependent selection maintains clonal diversity in an asexual organism. Proceedings of the National Academy of Sciences 105:17872–17877.

Wickham, H. 2016. Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wright, S. 1931. Evolution in mendelian populations. Genetics 2.

Wright, S. 1948. On the roles of directed and random changes in gene frequency in the genetics of populations. Evolution 2:279–294.

Xie, Y., J. Allaire, and G. Grolemund. 2018. R Markdown: The Definitive Guide. Chapman and Hall / CRC.

Zaman, L., S. Devangam, and C. Ofria. 2011. Rapid host-parasite coevolution drives the production and maintenance of diversity in digital organisms. Page 219 *in* Proceedings of the 13th annual conference on Genetic and evolutionary computation - GECCO '11. ACM Press, Dublin, Ireland.

Chapter 3

Hierarchical host-parasite metapopulation structure selects for genetic divergence of host subpopulations

Authors: Robert P. Dünner[1,2], Jukka Jokela[1,2].

Author affiliations: ETH Zurich, Institute for Integrative Biology[1], Eawag, Department of Aquatic Ecology[2].

Corresponding author: Robert P. Dünner; robert.duenner@eawag.ch, Eawag ECO BU-G-12, Uberlandstrasse 133, 8600 Dubendorf, Switzerland.

Keywords: Host-parasite coevolution, negative frequency-dependent selection, metapopulation structure, genetic diversity, population divergence, agent-based simulation.

Author contributions: RD and JJ conceived the ideas and designed the work, RD authored the simulation and collected the data, RD analysed the data, RD and JJ led the writing.

Acknowledgements: RD was supported by the adaptation to a changing environment center (ACE) at ETH. We would like to thank the HPC Team at ETH Zurich for providing the computational resources used in this study. We would like to thank Kyra Gähwiler, Moritz Lürig, Natalia Zajac, Natalie Sieber, Elena Gimi and Hugo Mathe-Hubert for helpful comments on the manuscript. Anja Gähwiler kindly designed Figures 1 and 2.

Data accessibility: Data will be made available through the Dryad Digital Repository upon publication.

Abstract

Negative frequency-dependent selection (NFDS) in host-parasite coevolution promotes the maintenance of genetic diversity. Migration patterns and population structure play a key role in determining the dynamic behaviour of coevolving host and parasite genotype frequencies in metapopulations. The structure of the metapopulation can play a key role in determining the dynamic behaviour of coevolving host and parasite genotype frequencies. We are interested in the coveolutionary dynamics of cases where the metapopulation structure of the host and the parasite differ substantially. We focus on the situation where the parasite has a metapopulation consisting of well connected subpopulations and is coevolving with a host whose metapopulation consists of isolated subpopulations.

In this study we used agent-based simulations of NFDS to analyse the influence of parasite metapopulation connectedness on host subpopulation divergence and genetic diversity. Our simulation results suggest that a connected parasite metapopulation coevolving with isolated host subpopulations can shift the focus of coevolution from within the host subpopulations to the overall genotype frequency of the host "metapopulation". This stops NFDS by the parasite on the host genotype frequency within the subpopulations, leading to a loss of host genetic diversity. Even though host genetic diversity is lost within host subpopulations, it is maintained at the level of the metapopulation. This leads to a compartmentalisation of genetic diversity into host subpopulations and consequentially host subpopulation divergence.

Introduction

Negative frequency-dependent selection (NFDS) between hosts and parasites can maintain genetic diversity in clonally reproducing host organisms by providing a selective advantage to locally rare host genotypes (rare advantage) (Wright 1948, Haldane 1949, Haldane and Jayakar 1963, Clarke and O'Donald 1964, Gillespie 1975, Clarke 1976, 1979, Levin 1988).
Host-parasite coevolution rarely takes place in singular isolated host-parasite populations, but instead in metapopulations consisting of several, more or less interconnected, host and parasite subpopulations. The structure of metapopulations strongly influences the outcome of host-parasite coevolution (Ladle et al. 1993, Thrall et al. 2012, Boëte et al. 2019).

Of special interest are asymmetrical metapopulation structures between the host and the parasite, especially when the parasite metapopulation is substantially more interconnected. This means that the parasite subpopulations are connected by a large amount of migration, leading to a high rate of gene flow. The host subpopulations, on the other hand, will have no migration between them and stay isolated in terms of gene flow. In other words, several isolated host subpopulations share a common parasite population. This can, for example, be the case if the parasite is more mobile than the host, like in many plant pathogen systems. It may also be the case if the parasite has a complex life cycle involving several host species, where the final host species has a larger population range than the intermediate host species. Some authors have even suggested that parasites might be migrating more than hosts in general (Gandon 1996).
In all those cases one should find much less population structure in the parasite than in the host. This has been demonstrated in two separate snail-trematode systems in New Zealand (Dybdahl and Lively 1996) and in Finland (Louhi et al. 2010), as well as in a *Daphnia*-microsporidian system in the Czech Republic (Wolinska et al. 2011).

NFDS is caused by the reciprocal coevolutionary responses of the host and the parasite on each others genotype frequencies. If a parasite population spans several separate host populations, those host populations will be perceived as one single host population by the parasite. Consequentially, the parasite is going to coevolve with the average genotype frequency across these host populations. All host populations connected to the same parasite metapopulation will experience the same parasite genotype frequencies and hence also the same parasite selection pressure. As this selection pressure is dependent on the host genotype frequency across all host subpopulations, the local host population genotype frequencies are freed from direct feedback by the parasite. Host genotypes will experience parasite genotype frequencies and hence infection rates that will differ from those that would be expected under a locally coevolving parasite. Host genotypes that are rare in their local population, but whose overall genotype frequency is high will not experience a rare advantage. Instead they will face an infection rate that is higher than warranted by their local frequency and will be more likely to go extinct. On the other hand, a host genotype that is locally frequent but globally rare will not experience as much selection by the parasite locally as would be warranted by the host genotypes local frequency. If the local host genotype frequency and the global host genotype frequency are similar, then the host experiences normal NFDS by the parasite. We hypothesized that this relative difference in selection pressure by the parasite on corresponding host genotypes in separate host subpopulations would lead to genetic divergence between the host subpopulations.

If all host subpopulations experience the same selection pressure by a parasite that is coevolving with the average genotype frequencies across these subpopulations, NFDS within host subpopulations may be lost. Loss of NFDS in host subpopulations immediately opens the question about the fate of genetic diversity. If NFDS by the parasite is switched off within host subpopulations, but instead happens at the metapopulation level, does that mean that host genetic diversity can degrade in the host subpopulations but is maintained in the metapopulation?

To investigate whether the connection pattern of the parasite metapopulation can influence host-parasite coevolution under NFDS, we used an agent-based, discrete-time, time-forward, stochastic simulation with clonal reproduction. We simulated five isolated subpopulations of clonal host individuals consisting of the same five genotypes each, that either coevolved with five unconnected parasite subpopulations, or with one global connected parasite metapopulation.

Specifically, we wanted to test three aspects of host-parasite coevolution under this metapopulation structure.
Will coevolution of isolated host subpopulations with a connected parasite metapopulation instead of

isolated parasite subpopulations lead to increased genetic divergence of host subpopulations?
Is NFDS lost within host subpopulations when they coevolve with a global parasite metapopulation instead of a local parasite subpopulation?
How much host genetic diversity is maintained, and where, when isolated host subpopulations coevolve with a global parasite metapopulation?

Methods

Simulation

To explore the influence of parasite metapopulation connectedness on NFDS and on maintenance of genetic diversity we used the Digital_Coevolution simulation (version: v1.0-thesis). The Digital_Coevolution simulation is an agent-based, discrete-time, time-forward, stochastic simulation with agents that are based on the resource allocation hypothesis, that has been written in R by Robert Dünner (Dünner 2020). The simulation was used with the agent growth functionality switched off. See chapter 1 for a very detailed and chapter 2 for a brief description of the simulation.

We simulated two scenarios that represent the two possible extremes of parasite metapopulation connectedness.
In the global scenario five unconnected host subpopulations that each contain the same five host genotypes coevolve with a parasite metapopulation whose subpopulations are fully connected by high amounts of migration and hence gene flow (Figure 1). The connectedness of the parasite subpopulations was implemented as a repeated random re-assignment of population affiliation. At each time-step, each parasite individual would hence find itself in a random host subpopulation and potentially get consumed or re-assigned to a new population in the next time-step.
The local scenario, on the other hand, consists of five unconnected host subpopulations that each contain the same five host genotypes. Each of those host subpopulations coevolves with one local parasite subpopulation which is not connected to other parasite subpopulations (Figure 2).

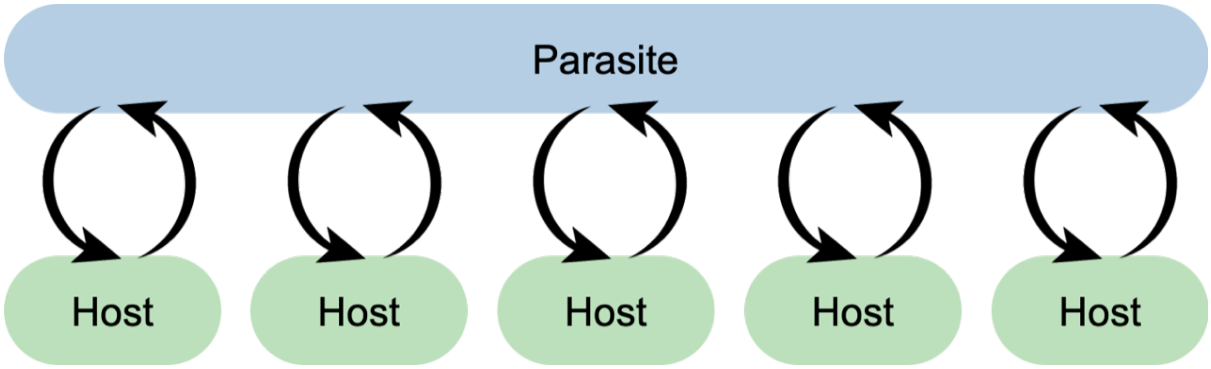

Figure 1: Metapopulation consisting of isolated host subpopulations that coevolve with a connected parasite metapopulation.

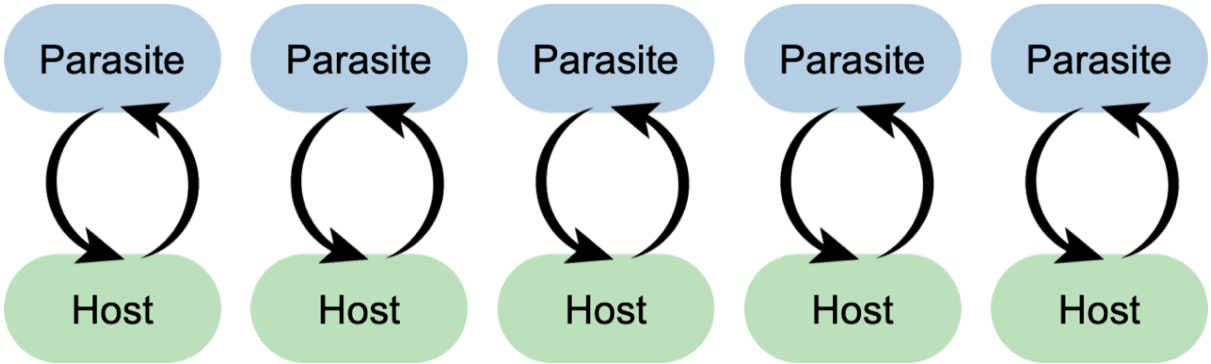

Figure 2: Metapolation consisting of isolated host and parasite subpopulations that coevolve locally.

Host population size is an emergent property of agent based simulations and can not be controlled directly. Population size can be manipulated though by supplying the simulated populations with different levels of resources. We simulated two levels of subpopulation resources, with all subpopulations within a metapopulation having the same resource levels. Low resource levels resulted in small subpopulations of around 100 host individuals which then formed metapopulations of around 500 host individuals. High resource levels resulted in large subpopulations of around 800 host individuals which then formed metapopulations of around 4000 individuals.

The parasite subpopulation sizes also are an emergent property of the simulation and are related to the host population size. The parasite subpopulation infecting a small host subpopulation reached around 550 individuals, leading to a parasite metapopulation size of around 2750 individuals. The parasite subpopulation infecting a large host subpopulation reached around 4400 individuals, leading to a parasite metapopulation size of around 22000 individuals.

All host subpopulations were started with the same five host genotypes at equal frequencies, which means that all subpopulations were identical at the beginning of the simulation. The parasite subpopulations were started with five matching genotypes at equal frequencies. The infection system was implemented as a perfect matching-alleles type of infection.
As our model allowed no mutation, no re-introduction of genotypes and no migration of hosts between subpopulations, any extinction of a genotype was final, and the number of genotypes in both the meta- and the subpopulation could only decrease.

Parasites were simulated to have a virulence of 0.4, meaning that a successful parasite infection decreased the hosts resource availability by 40 %. This led to strongly reduced fecundity of the host and, when resources were scarce, to increased mortality.
To simulate the behaviour of host genotype frequencies without a parasite (drift baseline), we simulated parasite agents with a virulence of 0 %. Parasites of 0 % virulence have no effect on the hosts, and could thus be used to simulate host populations under drift.

The simulation was thus run for 2500 time-steps for the smaller 100 host individual subpopulations and 20000 time-steps for the larger 800 host individual subpopulations. Number of individuals per host and parasite genotype was recorded at 50 evenly spaced time points throughout each run of the simulation. We ran 40 replicates per parameter combination.

The simulations were run on the "Euler" high performance cluster computer at ETH Zurich on R version 3.6.0 linked again OpenBLAS available in module gcc/4.8.2. Euler is a HPC cluster running on Linux and uses the IBM load sharing facility (LSF).

Analysis

All analyses were conducted in R (Team 2020) using RStudio as IDE (Team 2019), the *data.table* package for data handling (Dowle and Srinivasan 2019) and *ggplot2* (Wickham 2016) and *RColorBrewer* (Neuwirth 2014) for visualisations.
We used genetic richness (number of genotypes) as a measure of genetic diversity. We calculated the average genetic richness within subpopulations and within the metapopulation across 40 replicates and calculated a 95% confidence interval assuming a normal distribution.
To measure the extent of NFDS by the parasite, we used Pearson's product-moment correlation as well as the slope of the linear relationships between host and parasite genotype frequencies. Under NFDS there should be a positive correlation of host and parasite genotype frequencies. We calculated Pearson's product-moment correlation using the *cor.test* function in base R. To estimate the linear relationship we constructed four normal distributed linear models using the *lm* function (base R). One model each was calculated for the linear relationship at the subpopulation level, with or without parasite population connectedness, and for the linear relationship at the metapopulation level, with or without parasite population connectedness. We used the parasite genotype frequency as the dependent variable and the host genotype frequency as the explanatory variable. We added population size as an explanatory variable as well, and used the joint data for both population sizes in the same model. See appendix for the full model outputs.
To measure subpopulation divergence we calculated Josts D at the endpoint of the simulation. Josts D can be interpreted analogously to G_{ST} (Jost 2008). Josts D is robust to the amount of polymorphism in subpopulations and detects divergence also for highly polymorphic subpopulations. We used the *basic.stats* function from the *hierfstat* package to calculate Josts D, with the settings for haploid data (Goudet and Jombart 2015) and calculated a 95% confidence interval around the mean assuming a normal distribution.

Results

Genetic divergence between host subpopulations

A connected global parasite metapopulation (Figure 1) led to stronger genetic divergence of host subpopulations than either a situation with unconnected local parasite subpopulations (Figure 2) or a situation without a parasite (drift baseline).
Analysis of genetic divergence using Josts D showed complete divergence between host subpopulations coevolving with a connected parasite metapopulation for both simulated levels of subpopulation size (Table 1). Josts D for host subpopulations coevolving with unconnected, local parasite subpopulations was much lower for both levels of subpopulation size.
Genetic divergence this strong can only be reached when all host subpopulations have fixed for one single exclusive genotype. This can also be seen in Figure 4.

Table 1: Genetic divergence between host subpopulations at the endpoint of the simulation, measured as Josts D. A value of 0 means no divergence, a value of 1 means complete divergence.

Subpopulation size	Parasite population connection scenario	Josts D	Confidence interval
100	No parasite	0.81	[0.78, 0.85]
100	Local parasite	0.77	[0.72, 0.83]
100	Global parasite	1	[1, 1]
800	No parasite	0.78	[0.74, 0.81]
800	Local parasite	0.40	[0.33, 0.47]
800	Global parasite	1	[1, 1]

Loss of NFDS within subpopulations

We found that when isolated host subpopulations coevolve with a connected parasite metapopulation instead of with isolated parasite subpopulations, the focus of NFDS by the parasite shifts from the within host subpopulation genotype frequency to the overall genotype frequency across all host subpopulations (the host metapopulation).
The shift of coevolutionary focus by the parasite leads to a loss of NFDS within the host subpopulations. This is visible when comparing the temporal dynamics of host genotype frequencies within subpopulations in Figure 3 and 4. Figure 3 shows one replicate of five host subpopulations that each coevolve with a local, isolated parasite subpopulation (local scenario, Figure 2). Figure 4 shows one replicate of five host subpopulations that each coevolve with the same global connected parasite metapopulation (global scenario, Figure 1). In Figure 3, with a local parasite, the host genotype frequencies show a clear signal of strong NFDS. No host genotype reaches a frequency that differs much from the average host genotype frequency of 0.2 and no genotypes go extinct. Their behaviour also clearly differs from the drift baseline shown in grey. In Figure 4, with a global parasite, the signal of NFDS is lost. The host genotype frequencies quickly start diverging from the average genotype frequency and four out of five genotypes go extinct in all subpopulations. The behaviour of the host genotype frequencies looks very similar to the drift baseline shown in grey.
Taken together this behaviour strongly suggests that NFDS has been lost when isolated host subpopulations coevolve with a connected parasite metapopulation. In the appendix of chapter 3 are figures of more replicates. Figures 1-3 in the appendix correspond to Figure 3 here, and Figures 4-6 in the appendix correspond to Figure 4 here.

The finding that NFDS within host subpopulations is lost with a connected parasite population is strengthened by the loss of correlation and the decrease in linear slope between host genotype frequency and parasite genotype frequency. Both the correlation and the linear slope are strongly decreased with a connected parasite population compared to an unconnected parasite population (Table 2).

Table 2: Slope of linear relationship ± standard error and Pearson's product-moment correlation incl. 95% confidence interval between host genotype frequencies and parasite genotype frequencies. Stars depict levels of significance.

Situation	Slope of linear relationship	Pearson's product-moment correlation
Subpopulation level, local parasite	1.03 ± 0.04 ***	0.96 (0.94, 0.98) ***
Subpopulation level, global parasite	0.14 ± 0.03 ***	0.59 (0.38, 0.74) ***
Metapopulation level, local parasite	- 2.22 ± 1.33	- 0.16 (- 0.40, 0.11)
Metapopulation level, global parasite	0.95 ± 0.19 ***	0.56 (0.35, 0.72) ***

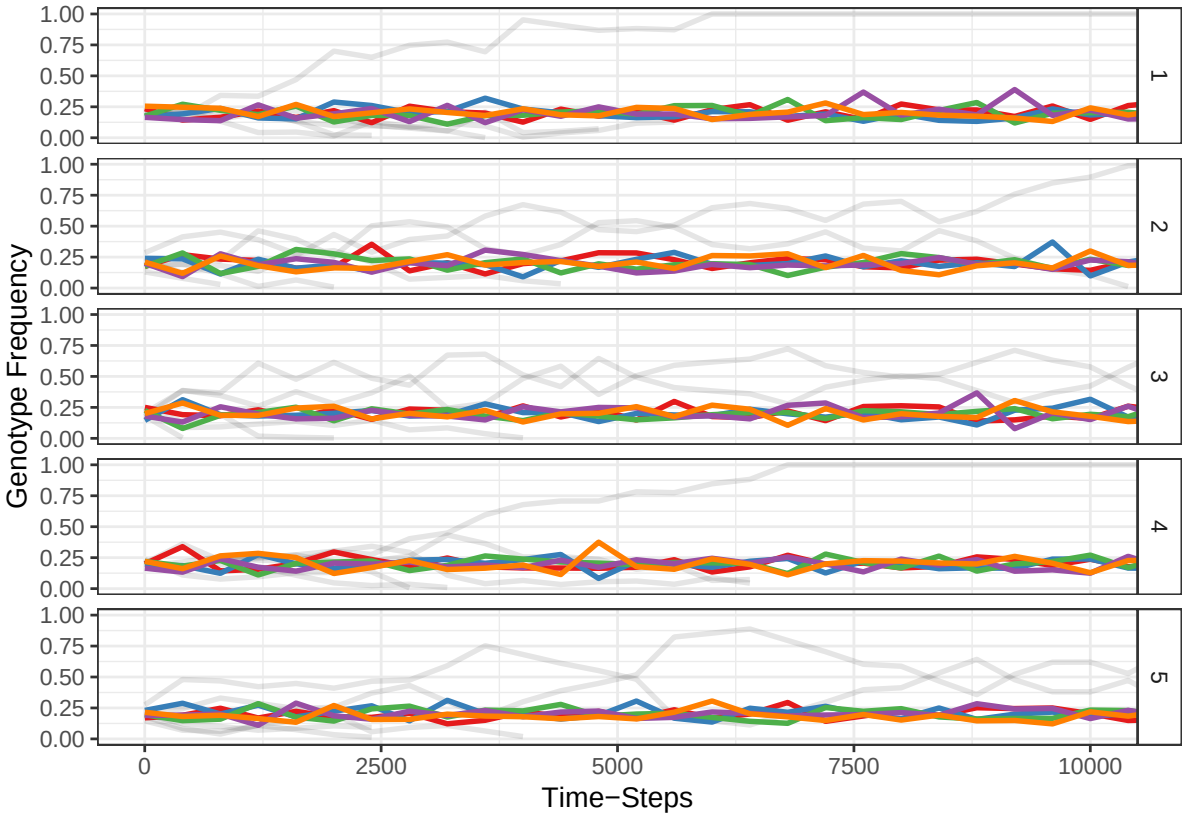

Figure 3: Host genotype frequencies within subpopulations over time. One replicate of five host subpopulations of population size 800 that each have a locally coevolving parasite subpopulation are shown. All subpopulation show the same NFDS dominated dynamics. No host genotype can increase or decrease in frequency as the parasite excerts control. The baseline under drift is in grey in the background, without resolving genotypes. Each subpopulation has about 800 individuals. Different colours depict different host genotypes.

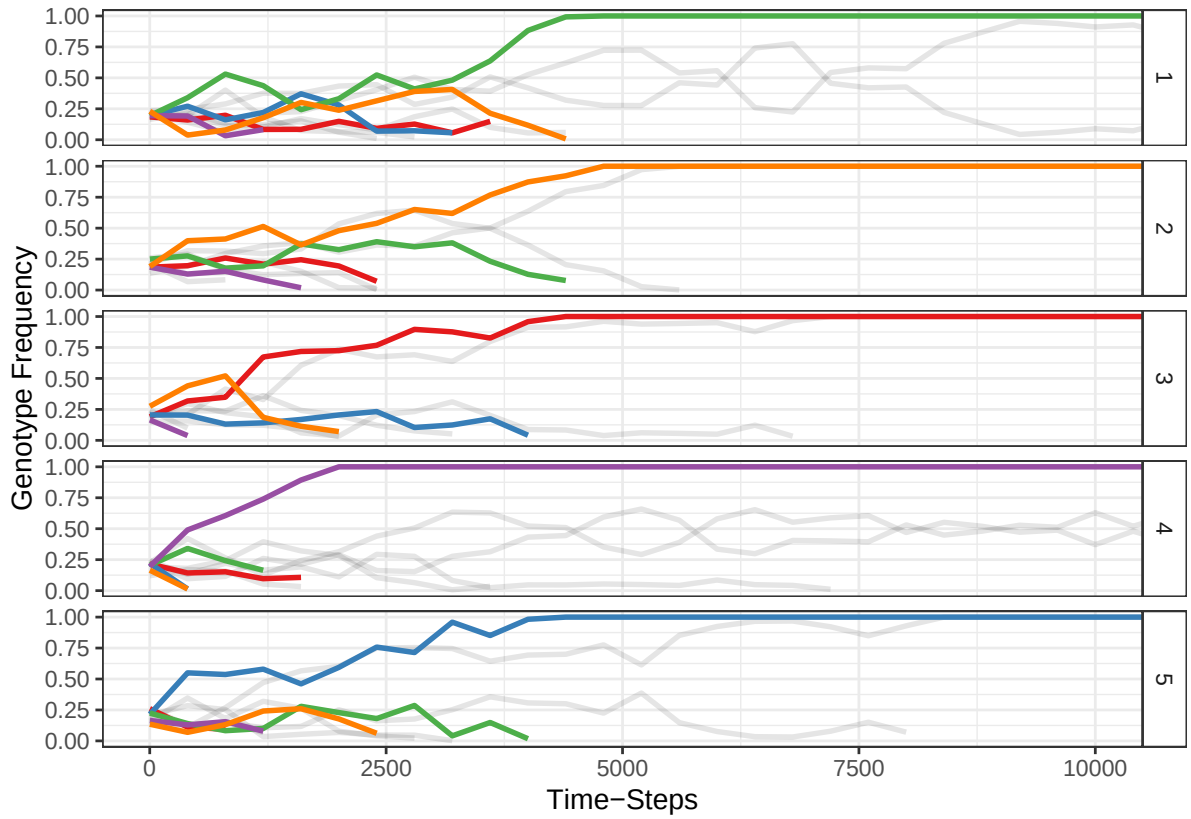

Figure 4: Host genotype frequencies within subpopulations over time. One replicate of five host subpopulations of population size 800 that are all coevolving with the same global parasite metapopulation are shown. There is no evidence of negative frequency-dependent selection. All host subpopulations quickly fix for a single host genotype. Note how each subpopualtion fixes for a different host genotype. The baseline under drift is in grey in the background, without resolving genotypes. Each subpopulation has about 800 individuals. Different colours depict different host genotypes.

Distribution of host genetic diversity

Even though NFDS was lost within host subpopulations coevolving with a connected parasite metapopulation, it was maintained at the level of the host metapopulation. This had consequences for the maintenance and distribution of host genetic diversity. Genetic diversity of the host was being maintained at the level of the metapopulation, but was degraded at the level of the subpopulations (Figure 5). In Figure 5, the dashed lines that show host genetic diversity in the metapopulation stay at the maximum of five host genotypes when the parasite is connected (blue dashed line) for both population sizes, while the metapopulation host genetic diversity with unconnected parasites (red dashed line) is reduced over time in smaller populations. The solid line in Figure 5 shows that the within subpopulation host genetic diversity is strongly degraded over time when the parasite population is connected (solid blue line) compared to the case of unconnected parasite populations (solid red line) as well as drift (solid grey line).
Maintaining full genetic diversity in the metapopulation while there is complete degradation of genetic diversity in the subpopulations is only possible, if each of the subpopulations has fixed for one single exclusive genotype. This means genetic diversity has been compartmentalized into host subpopulations.

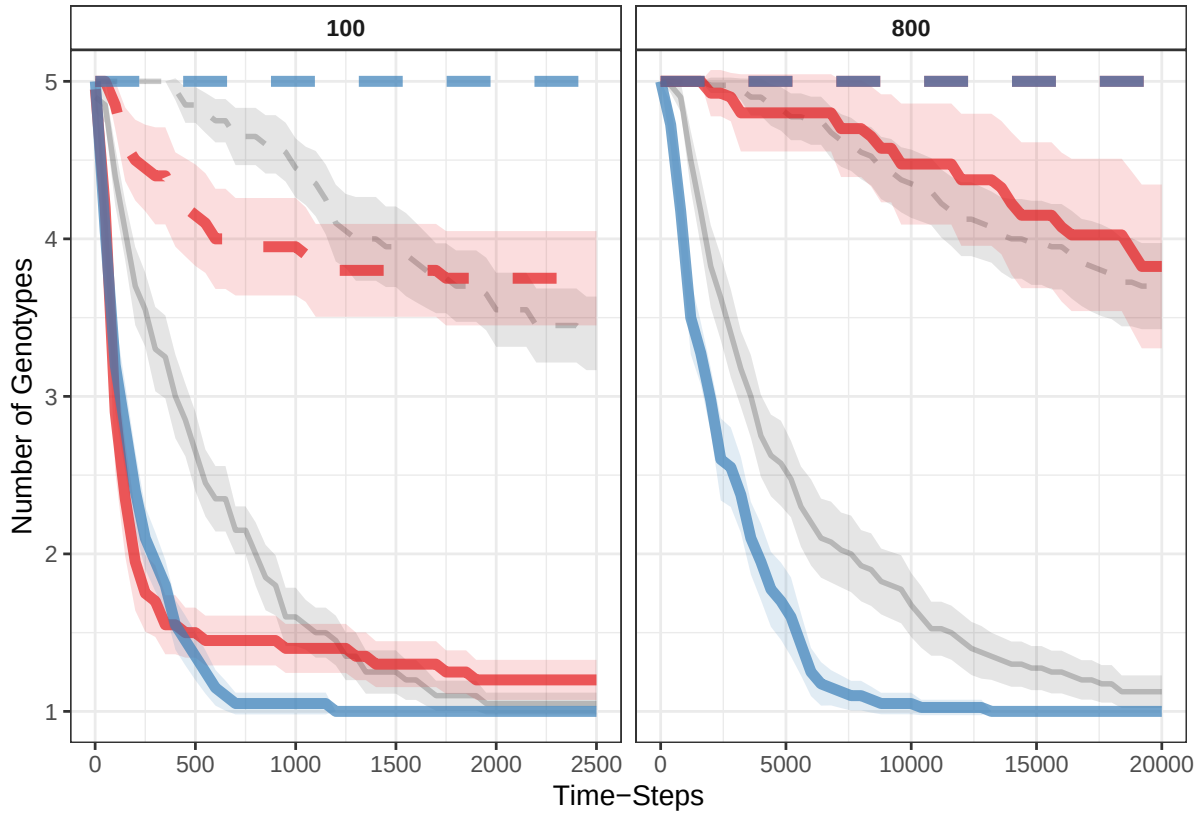

Figure 5: Number of host genotypes over time. Solid lines are number of genotypes within subpopulations, dashed lines are number of genotypes within the metapopulation. Blue coloured lines are with a migrating parasite, red lines with a local parasite. The two panels are the simulated population sizes. The panel on the left are subpopulations of about 100 individuals, the panel on the rigth of about 800 individuals. The red and the blue dashed line in the right panel are overlapping. The grey lines are the drift baseline.

Discussion

The observation that two host entities which coevolve with the same parasite entity will start to diverge genetically has already been noted by Clarke (1979). He proposed that when two separate host species share a common parasite species, those host species could start to diverge, which is an exact analog to what we found for two host populations that share one parasite population.

> "It is worth drawing attention to the fact that frequency-dependent selection will not only maintain diversity within species, but also promote divergence between them. If two closely-related species of hosts share one or more species of parasite, it will be advantageous for the hosts to diverge." (Clarke 1979).

To our knowledge the finding that negative frequency-dependent selection (NFDS) can be switched off within host subpopulations by connecting the parasite subpopulations has not been previously reported. A plausible explanation could be that allowing the parasite to migrate randomly between host populations fuses the separate host subpopulations into one host metapopulation from the perspective of the parasite. This shifts the coevolutionary focus of the parasite and hence also the population structure level at which NFDS may be expected.
A host genotype in a subpopulation can drift freely to any genotype frequency within this subpopulation without triggering a response by the parasite as long as the net change of its frequency in the host metapopulation is zero. In other words, a host genotype can drift in any direction within a certain subpopulation, as long as that drift is compensated by the same corresponding genotype in another subpopulation drifting in the opposite direction.

While NFDS within subpopulations is lost, the parasite still exerts NFDS on the overall metapopulation. All host genotypes, independent of their within host subpopulation genotype frequency, experience the same selection pressure by the parasite. This means that host genotypes that are rare in the overall metapopulation will still experience a rare advantage and have a low likelihood of going extinct in all host subpopulations. On the level of the metapopulation NFDS by the parasite still maintains host genetic diversity.

If we expect genetic diversity to be maintained on the level of the metapopulation but not at the level of the subpopulation, there is only one stable outcome. All subpopulations will diverge as much as possible in terms of genotype frequencies and consequentially genetic diversity. In our case with five subpopulations and five genotypes maximum divergence is observed when each subpopulation has fixed for one exclusive host genotype, as we have seen in Figures 4 and 5. This means that genetic diversity has been degraded within host subpopulations, while it has been maintained in the metapopulation. Allowing the parasite to move around has compartmentalized the genetic diversity into subpopulations, leading to a strong gene storage effect (Hamilton 1993, Judson 1995). That combining highly connected parasite subpopulations with isolated host subpopulations has a profound effect on maintenance and distribution of genetic diversity has been shown by Judson (1995) and Judson (1997) using an agent-based model that has been modified from Hamilton et al. (1990). Our findings confirm those results and advance them in providing an explanation based on the loss of NFDS within host subpopulations.

Matching numbers of host subpopulations and genotypes allowed each subpopulation to fix for one single exclusive host genotype. In reality there might be far more than five host and parasite genotypes present in a metapopulation, and also the number of subpopulations might be greater. In such more general cases, a connected parasite metapopulation like we described it will still shift the coevolutionary focus of the parasite to the metapopulation. This would still lead to the loss of NFDS within and the genetic divergence between host subpopulations. Instead of fixing for one exclusive genotype, host subpopulations might stabilize at a subset of exclusive host genotypes or at a mixture of exclusive genotypes and few shared genotypes at low frequency. The number of genotypes to be expected per subpopulation would simply be the total metapopulation number of genotypes divided by the number of metapopulations.

Several host subpopulations that each consist of exclusive genotypes create reproductive isolation between those host genotypes, increasing the potential for speciation. Such a process kick-starts allopatric speciation by already pre-sorting the available genetic diversity into monomorphic subpopulations. Wright (1931) and Wright (1948) noted that a metapopulation consisting of fairly isolated host subpopulations is going to evolve faster than a well mixed single population of similar size. This effect is only going to be strengthened with a parasite that is allowed to move freely between host subpopulations as we have simulated.

Complete compartmentalisation of genetic diversity through host-parasite coevolution might be relevant to account for in ecological studies of local adaptation. We have seen that subpopulations quickly degrade their genetic diversity to fix for a subset of available host genotypes, creating a strong pattern of subpopulation divergence. If this is happening in a landscape and one would be measuring environmental parameters alongside population genetic information, there is the danger of wrongly attributing genetic differences between subpopulations to differences in environmental parameters. This could potentially mislead conclusions about local adaptation of host subpopulations to environmental parameters, even though the population divergence has been entirely caused by host-parasite coevolution. Such a pattern of genetic diversity that would match ecological parameters but is instead caused by interactions of host subpopulations with a shared parasite could be called "apparent local adaptation", similar to the concept of apparent competition (Holt and Lawton 1994).

Conclusion

We have simulated several unconnected host subpopulations that either interact with a parasite that is local to a host subpopulation or with a global parasite that can migrate between the subpopulations. This structure leads to an asymmetrical shift of the population structure level at which the host and the parasite coevolve. If the parasite is allowed to migrate freely, the metapopulation becomes the relevant population structure for the parasite. This disentanglement of organisational levels of coevolutionary feedback between host and parasite leads to strong subpopulation divergence, the loss of NFDS within host subpopulations and to the compartmentalisation of genetic diversity.
Increased host subpopulation divergence might increase the speed of speciation by pre-sorting existing genetic diversity into host subpopulations. It might also create problems for empirical studies on adaptation to environmental parameters. The potential loss of NFDS within host subpopulations has to be taken into account when testing theoretical expectations of host-parasite coevolution in the wild.
Taken together this means that accounting for the metapopulation structure of both host and parasite is important to fully understand host-parasite co-evolution.

References

Boëte, C., M. Seston, and M. Legros. 2019. Strategies of host resistance to pathogens in spatially structured populations: An agent-based evaluation. Theoretical Population Biology 130:170–181.

Clarke, B. 1976. The ecological genetics of host-parasite relationships. *in* Genetic aspects of host-parasite relationships. Blackwell, London.

Clarke, B. 1979. The Evolution of Genetic Diversity. Proceedings of the Royal Society B: Biological Sciences 205:453–474.

Clarke, B., and P. O'Donald. 1964. Frequency-dependent selection. Heredity:6.

Dowle, M., and A. Srinivasan. 2019. Data.Table: Extension of 'data.Frame'.

Dünner, R. P. 2020. Digital_Coevolution: An agent-based simulation for host-parasite coevolution.

Dybdahl, M. F., and C. M. Lively. 1996. The Geography of Coevolution: Comparative Population Structures for a Snail and Its Trematode Parasite. Evolution 50:2264.

Gandon, S. 1996. Local adaptation and gene-for-gene coevolution in a metapopulation model. Proceedings of the Royal Society of London. Series B: Biological Sciences 263:1003–1009.

Gillespie, J. H. 1975. Natural Selection for Resistance to Epidemics. Ecology 56:493–495.

Goudet, J., and T. Jombart. 2015. Hierfstat: Estimation and Tests of Hierarchical F-Statistics.

Haldane, J. B., S. 1949. Disease and evolution. La Ricerca Scientifica Supplemento.

Haldane, J. B., S., and S. D. Jayakar. 1963. Polymorphism due to selection depending on the composition of a population. Journal of Genetics 3:318–323.

Hamilton, W. D. 1993. Haploid Dynamic Polymorphism in a Host with Matching Parasites: Effects of Mutation/Subdivision, Linkage, and Patterns of Selection. Journal of Heredity 84:328–338.

Hamilton, W. D., R. Axelrod, and R. Tanese. 1990. Sexual reproduction as an adaptation to resist parasites (a review). Proceedings of the National Academy of Sciences 87:3566–3573.

Holt, R. D., and J. H. Lawton. 1994. The Ecological Consequences of Shared Natural Enemies. Annual Review of Ecology and Systematics 2:495–520.

Jost, L. 2008. GST and its relatives do not measure differentiation. Molecular Ecology 17:4015–4026.

Judson, O. P. 1995. Preserving genes; a model of the maintenance of genetic variation in a metapopulation under frequency-dependent selection. Genet. Res.:175–191.

Judson, O. P. 1997. A Model of Asexuality and Clonal Diversity: Cloning the Red Queen. Journal of Theoretical Biology 186:33–40.

Ladle, R. J., R. A. Johnstone, and O. P. Judson. 1993. Coevolutionary dynamics of sex in a metapopulation: Escaping the Red Queen. Proceedings of the Royal Society of London. Series B: Biological Sciences 253:155–160.

Levin, B. R. 1988. Frequency-dependent selection in bacterial populations. Philosophical Transactions of the Royal Society B: Biological Sciences:459–472.

Louhi, K.-R., A. Karvonen, C. Rellstab, and J. Jokela. 2010. Is the population genetic structure of complex life cycle parasites determined by the geographic range of the most motile host? Infection, Genetics and Evolution 10:1271–1277.

Neuwirth, E. 2014. RColorBrewer: ColorBrewer Palettes.

Team, R. 2019. RStudio: Integrated Development Environment for R. RStudio, Inc., Boston, MA.

Team, R. C. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austia.

Thrall, P. H., A.-L. Laine, M. Ravensdale, A. Nemri, P. N. Dodds, L. G. Barrett, and J. J. Burdon. 2012. Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation: Coevolution in a wild host-pathogen system. Ecology Letters 15:425–435.

Wickham, H. 2016. Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wolinska, J., P. Spaak, H. Koerner, A. Petrusek, J. Seda, and S. Giessler. 2011. Transmission mode affects the population genetic structure of Daphnia parasites: Parasite population structure. Journal of Evolutionary Biology 24:265–273.

Wright, S. 1931. Evolution in mendelian populations. Genetics 2.

Wright, S. 1948. On the roles of directed and random changes in gene frequency in the genetics of populations. Evolution 2:279–294.

Chapter 4

Temperature induced demographic shifts in host populations of *Daphnia magna* are altered by infection with *Ordospora colligata* parasites

Authors: Robert P. Dünner[1,2], Aijuan Liao[1,2], Jukka Jokela[1,2].

Author affiliations: ETH Zurich, Institute for Integrative Biology[1], Eawag, Department of Aquatic Ecology[2].

Corresponding author: Robert P. Dünner; robert.duenner@eawag.ch, Eawag ECO BU-G-12, Uberlandstrasse 133, 8600 Dubendorf, Switzerland.

Keywords: GxGxE interactions, context dependent fitness, genetic polymorphism, demography, coevolution, *Daphnia magna, Ordospora colligata.*

Author contributions: RD and JJ conceived the ideas and designed the work, RD and AL collected the data, RD analysed the data, RD and JJ led the writing. All authors approved the final version for publication.

Acknowledgements: We would like to thank Dieter Ebert for providing the *Daphnia* and the *Ordospora* parasites and for helpful comments on the manuscript. We would also like to thank Moritz Lürig, Hanna Hartikainen, Blake Matthews, Linda Haltiner and Amruta Rajarajan for helpful comments on the manuscript. RD was supported by the adaptation to a changing environment center (ACE) at ETH.

Data accessibility: Data will be made available through the Dryad Digital Repository upon publication.

Abstract

Infection success of parasites can be determined by genetically specific interactions with their hosts (GxG interactions). This may lead to negative frequency-dependent selection between the host and the parasite genotypes and maintenance of genetic polymorphism. Host and parasite genotypes may also interact with the environment (GxE interaction). GxE interactions either reduce genetic polymorphism by selective sweeps in stable environments or maintain polymorphism in fluctuating or spatially variable environments. Together, GxG and GxE interactions may interfere and generate complex GxGxE interactions that are likely to be important for maintenance of genetic polymorphism, but the outcome of which will be difficult to predict.

Natural *Daphnia* populations host a diversity of parasites and are a model system for laboratory and field studies of host-parasite coevolution and local adaptation. Here, we examined the magnitude and direction of GxGxE interactions in *D. magna* infected with a low-virulence parasite *Ordospora colligata* in different environmental conditions.

We exposed three different monoclonal *D. magna* host populations to allopatric or sympatric *O. colligata* parasite genotypes in a partially-crossed design. We changed the ambient temperature from 20 °C to 16 °C half-way through the experiment. We recorded juvenile and adult population sizes of *D. magna* over 10 weeks to evaluate the population-level demographic response to parasite infection and environmental change in temperature.

We found that the proportion of juveniles in *D. magna* populations was lower at 16 °C relative to the proportion of juveniles at 20 °C. The magnitude of this demographic shift was strongly reduced by infection with either allopatric or sympatric *O. colligata* parasites, depending on the host genotype (GxGxE interaction).

Our results suggest that GxGxE interactions may generate selection that fluctuate over time and influence host demography and fitness. Such variable and context dependent fitness of both the host and the parasite may be important for maintenance of genetic polymorphism.

Introduction

Outcome of host-parasite interactions often depends on the interacting genotypes of both the host and the parasite (genotype-by-genotype interaction; GxG), with no single host or parasite genotype being universally "resistant" or "virulent" (Carius et al. 2001, Lambrechts et al. 2006). Genotype-specific infectivity/susceptibility may generate reciprocal selection and coevolutionary dynamics between the host and the parasite populations. When rare genotypes have a selective advantage and common genotypes are repeatedly replaced (negative frequency-dependent selection), coevolution can drive cyclical dynamics in genotype frequencies of both host and parasite. This process, often referred to as Red Queen Dynamics (RQD), is thought to be driven by fluctuating selection that may be relevant for maintenance of sexual reproduction and genetic polymorphism (Jaenike 1978, Hamilton 1980, King et al. 2011). Support for the RQD is widespread, as is criticism (see (Lively 2010) for a review), but empirical demonstrations remain rare (Decaestecker et al. 2007, Jokela et al. 2009).

Environmental conditions like temperature can have a major influence on the strength and direction of interactions between hosts and parasites, for example by altering host condition or changing parasite virulence (Mitchell et al. 2005, Tadiri et al. 2013). A genotype can express different phenotypes depending on the environment and according to its reaction norm. Differing reaction norms between genotypes indicate genotype by environment interactions (GxE interaction), where the fitness rank of a genotype becomes environment specific. It is reasonable to expect that genotype-specific environmental variation also influences coevolving host-parasite interactions.

Genotype-specific response to environment and genotype-specific infection dynamics may combine to three-way genotype by genotype by environment interactions (GxGxE interaction). It is not clear what consequences GxGxE interactions have for coevolutionary dynamics between the host and the parasite. Potentially, environmental variation acts as a disturbance that could de-stabilize the coevolutionary dynamics predicted by GxG interactions (Wolinska and King 2009).

Here we assess the direction and magnitude of GxGxE interactions in a temporally changing environment. We conducted a cross-infection experiment using the cladoceran *Daphnia magna* as host and the microsporidian *Ordospora colligata* as parasite. To mimic onset of autumn, we changed temperature and day length half-way during the experiment. We used the juvenile population size (corrected for adult population size) as a response variable to follow the change in demographic structure of the host populations in response to GxGxE interactions. The size of the juvenile population is a direct proxy of the key demographic parameters (age structure, growth potential, fecundity) of the total population. We discuss the potential consequences of widespread GxGxE interactions for host-parasite coevolution.

Methods

Experimental system

Daphnia magna Strauss 1820 (Cladocera, Crustacea) is a small freshwater zooplankter that is abundant in laboratories and natural waterbodies around the world. *Daphnia* are cyclical parthenogens that reproduce clonally for the large part of the favourable growing season. Sexual reproduction occurs at the end of the growing season in autumn or under stressful conditions, and results in the production of durable resting eggs, called ephippia. As clonal reproduction allows maintenance of monoclonal lines in the laboratory, *Daphnia* is a valuable model system that is widely used in biological research. Many parasite species use *D. magna* as their natural host, which makes *D. magna* a good model organism for research on coevolution (Ebert 2008).
Ordospora colligata Larsson 1997 is a parasitic microsporidian that so far has only been found parasitizing *D. magna*. It is an intracellular parasite of the gut epithelium and transmits horizontally via faeco-oral transmission (Larsson et al. 1997). *O. colligata* is widespread across the native range of *D. magna* populations and can reach very high prevalence (Decaestecker et al. 2005). Infections with *O. colligata* are fairly avirulent and will slightly reduce lifespan and lifetime reproductive output of infected *D. magna* individuals (Ebert et al. 2000). *O. colligata* strains are highly specific to particular host genotypes and can change the competitive ranking of *D. magna* clones, changing the outcome of clonal competition in experiments (Capaul and Ebert 2003).
The *Daphnia* clones were obtained from Dieter Ebert's *D. magna* diversity panel at the University Basel. The *Ordospora* lines were obtained as sympatric infections on those *Daphnia*. Two *Daphnia* genotypes originated from the Tvärminne rock-pool metapopulation in Finland (genotypes SKW and OER) and one originated from Great Britain (genotype EP).

Experimental design

The experiment was set up as a randomized block design with partially-crossed factorial treatments (see Table 1). We used three clonal lines of *D. magna* and their respective sympatric *O. colligata* parasites. We repeatedly exposed monoclonal populations of each of the three *D. magna* lines to either the sympatric parasite, an allopatric parasite or a control treatment containing no parasite. Each treatment × clone combination consisted of 12 replicates. Data were recorded weekly as repeated measures per experimental unit. We recorded population size of juveniles and adults separately and counted the number of produced ephippia. Identity of each experimental unit was blinded and randomised to avoid researcher bias.

Table 1: Experimental design, columns are host genotype, rows are parasite source, cells are treatment names

		***Daphnia* Host Genotype**		
		EP	**SKW**	**OER**
Parasite Source (Original Sympatric Daphnia Host Genotype)	**EP**	Sympatric	Allopatric	Allopatric
	SKW	Allopatric	Sympatric	
	OER			Sympatric
	No parasite	Control	Control	Control

Experimental protocol

Clonal *D. magna* lines were grown in 720 mL glass jars filled with approx. 500 mL UV-sterilized ADaM in a climate chamber at 20 °C constant temperature and a 16/8 h light/dark cycle for the first five weeks and at 16 °C constant temperature and a 12/12 h light/dark cycle for the next five weeks (Klüttgen et al. 1994). This change in environmental conditions is similar in magnitude, but not speed, to what would be expected in natural populations between summer and autumn. *Daphnia* were fed 3 times a week with a mixture of live *Scenedesmus subspicatus* and *Chlorella vulgaris*. All treatments were re-exposed to parasites according to their treatment weekly. Spore suspensions for exposure were obtained by grinding and re-suspending monoclonal populations of *D. magna* of known infection status. These suspensions were then distributed to the experimental units (i.e. jars). We measured spore dose by counting the spores in the suspension using a Neubauer cell counting chamber and a phase-contrast microscope. This is a standard method to infect *Daphnia* with microsporidian parasite and was successfully used in prior experiments in our laboratory (unpublished data) (Haag and Ebert 2004). The control treatments received

a suspension of uninfected ground-up *Daphnia*. The experiment was started by creating 12 replicated populations per treatment, each consisting of 19 individuals per clone. Starting individuals were taken from pooled breeder populations and individually placed in a replicate experimental population. Largest individuals from the pool were chosen first, then consecutively smaller, until the breeder pool was depleted. This way the age distribution of the starting populations was standardized across all experimental units. Experimental populations were counted weekly by transferring *Daphnia* individuals to a new jar with a glass pipette, distinguishing between adult and juvenile individuals. During this process ephippia were counted but not transferred. Distinction between adults and juveniles was made mainly by size as well as by presence/absence of a brood pouch or eggs. This type of classification might not yield the exact value of the juvenile-to-adult ratio, but by keeping observer bias constant should reflect shifts in juvenile-to-adult ratio across treatments. Media was changed weekly.

Statistical analysis

Data were analysed using R and RStudio as IDE (Team 2019, 2020). In order to understand how the demography of the host populations changed, we used the numbers of individuals in the juvenile and adult age classes as a measure of demographic structure of the host population. We chose to use the juvenile population size as response variable and the adult population size as a co-variate. This conveniently allows for the use of a negative binomial distributed model and circumvents the caveat that using juvenile-to-adult ratio directly would result in the loss of information of overall population size. Corrected juvenile population size is a direct proxy for the demography of the population.
We used a generalized linear mixed model (glmm) with a log-link, a negative-binomial error distribution $V[\lambda] = \mu * (1 + \phi)$, and a dummy coding scheme, fitted with the *glmmTMB* package (Bolker et al. 2009, Brooks et al. 2017). The log-link in this type of model links the logarithm of the mean of the response variable with the mean of the error distribution. We calculated the exponential of the coefficients of this model to get a multiplicative factor, which shows how much the mean of the response, in percent, will change when the predictor changes by one unit.
The dummy coding scheme we used compares each level of an explanatory variable to the reference level, holding all other explanatory variables fixed. Interaction effects are then the difference between the estimated cell mean using additive lower level effects and the realised cell mean. This type of coding allows for the careful interpretation of the coefficients of lower level effects for certain factor levels in the presence of significant higher order interactions.
We included temperature, treatment, host clone, adult population size, spore dose, ephippia production and block as fixed effects in the model. Putatively biologically relevant two- and three-way interactions were included between temperature, treatment and host clone, as well as temperature, treatment and adult population size, and temperature and ephippia. We captured the within-subject variance caused by the repeated measures using the random intercept structure with jar as grouping variable, accounting for systematic differences in mean juvenile population sizes between jars. We added a random intercept for date of sampling, capturing synchronized fluctuations in juvenile population sizes, as for example caused by population size overshoots after initial setup of experiment. With these two crossed random intercepts we capture both variance in time caused by population fluctuations and variance between experimental units, effectively controlling for repeated measures. We used the control treatment of the EP host genotype at 20°C as the reference level. We used the *sjstats* package to estimate explained variance in the model, which uses a pseudo-R-squared implementation for glmm suggested by Nakagawa and Schielzeth (2013) and Nakagawa et al. (2017, Lüdecke 2020).

Results

We found strong GxGxE interactions in our experiment (Figure 1). In the SKW clone, the juvenile population size in the sympatric treatment at 16 °C was 74 % larger with the three-way interaction effect than predicted by the model without the interaction term (0.555 ± 0.200 log response units (lru), "Temperature: Sympatric: Clone SKW" in Table 1). In comparison, the magnitude of the interaction effect on the juvenile population sizes in the allopatric treatment in the SKW clone was only 15 % and not significant (0.135 ± 0.203 lru, "Temperature: Allopatric: Clone SKW" in Table 2). In the OER clone on the other hand, the juvenile population size in the allopatric treatment at 16 °C was 65 % larger than predicted by the model without the three-way interaction term (0.502 ± 0.191 lru , "Temperature: Allopatric: Clone OER" in Table 2), while the interaction effect on the juvenile population sizes in the sympatric treatment was only 6% and not significant (0.060 ± 0.188 lru, "Temperature: Sympatric: Clone OER" in Table 2). These three-way interaction effects were amongst the largest effects detected in this model (See Table A3 in Appendix A). To put this into perspective, changing between the EP clone and the OER clone in the control treatment at 20°C changed the juvenile population size only by 32 % (0.278 ± 0.091 log response units). This demonstrates the importance of the three-way interaction in determining the demographic response of the host clones in this experiment. In our results, influence of the GxGxE interaction on the host demography was of the similar magnitude, but of opposite sign, than simple GxE interactions and far exceeded the influence GxG interactions alone (Appendix A Table A3). Overall, the model explained 45% (fixed effect only; marginal R2= 0.451) and 62% (including random effects; conditional R2= 0.618) of the variance in the juvenile population size (see Appendix A Table A1 & A2 for model fit).

Table 2: Coefficients for the three-way (GxGxE) interactions from the negative binomial generalized linear mixed model. Exponentiating these coefficients gives percent change in response upon presence of these interactions. Coefficients for the three-way interactions only. See Appendix A Table A3 for the full model output.

Term	Estimate	Std. Error	z value	Pr(>
Temperature : Sympatric : Clone SKW	**0.555**	**0.200**	**2.770**	**0.006**
Temperature : Allopatric : Clone SKW	0.135	0.203	0.666	0.506
Temperature : Sympatric : Clone OER	0.060	0.188	0.318	0.751
Temperature : Allopatric : Clone OER	**0.502**	**0.191**	**2.625**	**0.009**

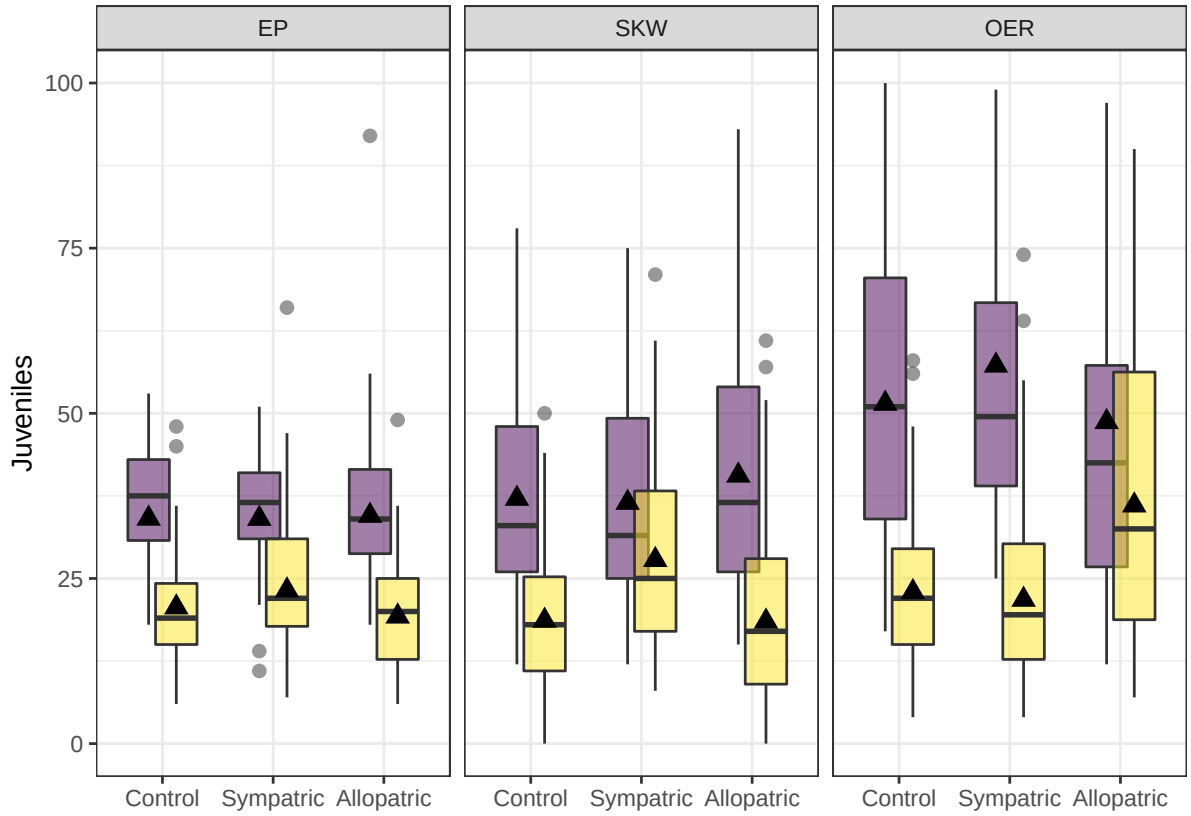

Figure 1: We exposed three host clones of D. magna (EP, SKW, OER) to three different parasite treatments (Control= no parasite, Sympatric parasite, Allopatric parasite). Plotted are the raw data for juvenile population sizes, showing all sampling dates simultaneously. The data for 20°C are in lilac, for 16°C in yellow. The black triangles are medians of the model fitted values. See Appendix B Figure B1 for a full plot of the model fitted values and Figure B2 for a plot of the adult population sizes.

Discussion

Our results show that the demography of the host population, represented by the corrected juvenile population size, was affected by a three-way interaction between temperature and host-parasite interactions (Figure 2). The change in demographic structure under new environmental conditions that we observed in most host-parasite genotype pairs was reduced in the allopatric exposed OER host and in the sympatric exposed SKW host. In other words, we found that the OER host genotype produced more juveniles at 16°C when exposed to an allopatric parasite than predicted by the statistical model by the combined lower level effects. The SKW genotype on the other hand produced more juveniles than predicted at 16°C when exposed to a sympatric parasite rather than an allopatric parasite. Both interaction effects were among the strongest in the statistical model and were more important in determining juvenile population size than for example the main effects for the clonal identity calculated within the control treatment. This shows that certain host and parasite genotype combinations reacted differently to a change in temperature than other combinations, and that this interaction effect has a magnitude that is potentially of biological relevance. The demography of a population, here represented as the corrected juvenile population size, reflects both the current reproductive state and the future reproductive potential of the *Daphnia* population. Life-history traits that govern demographic structure are fitness correlated and likely under optimizing selection given the local conditions. Shifting the demography of a host population as a result of GxGxE interactions reflects a change in selection on key life-history traits and is likely to have consequences on performance of the host clones in competition with other *Daphnia* genotypes.

Another notable result is the absence of a significant main effect for temperature (Appendix A Table A3). Instead of in the main, the effect of temperature on juvenile population size was instead mainly detectable within two-way interactions of host clone identity with temperature (GxE interactions) and in the three-way interactions (GxGxE interactions). This emphasizes the important point that the effect of the environment is context-dependent on both the host and the parasite genotypes.

Finding GxGxE interaction in such a limited subset of host and parasite genotypes suggests that GxGxE interactions may be widespread in the *D. magna* - *O. colligata* system and adds directly to the already published cases of GxGxE interactions in other systems (Tétard-Jones et al. 2007, Bryner and Rigling 2011, Seppälä et al. 2012). This finding also suggests that fitness variation due to GxGxE interactions may be a general characteristic of host-parasite interactions or at least more common than presently considered.

Fluctuating selection and the resulting maintenance of genetic polymorphism are the hallmark features of coevolutionary dynamics driven by negative frequency-dependent selection (the Red Queen Hypothesis) (Clarke 1976, King et al. 2011). Two conditions are necessary for the coevolutionary dynamics under this model. The first is that parasites have a strong negative fitness effect on the host upon infection. The second is that infection has a strong genotype-specific component (a GxG interaction) (Salathe et al. 2008). Both conditions are potentially violated with GxGxE interactions, where the realized virulence of an infection become environmentally context-dependent. The first condition is potentially violated if an infection with a parasite has reduced virulence under certain environmental conditions. The second condition is potentially violated if genotype specificity of infection breaks down when both host resistance and parasite infectivity are environmentally context-dependent (Guay et al. 2009, Studer et al. 2010, Okamura et al. 2011, Vale et al. 2011, Paaijmans et al. 2012). But what are the consequences when the necessary preconditions for coevolution are potentially reached only over limited periods of time or only for subsets of the present host and parasite populations? In temporally variable environments, changing environmentally context-dependent selection might interfere with negative frequency-dependent selection. This interference might create stronger and/or more variable fluctuating selection on both host and parasite than either process by itself. Gibson et al. (2018) have shown that fluctuation in virulence of the parasite can interfere with coevolutionary dynamics and facilitate maintenance of sexual reproduction. GxGxE interactions may provide a parsimonious additional mechanism for the creation of fluctuating selection pressure by the parasite. They may be vital for the maintenance of genetic polymorphism and sexual reproduction.

The proposed mechanism would also still be in congruence with results that document higher genetic diversity under higher parasite pressure (King et al. 2011, Turko et al. 2018).

If, as we have shown, new environmental conditions may lead to changes in the selective landscapes of host and parasite interactions, and if selection by parasites is a major force shaping genetic diversity, then GxGxE interactions will be important for evolutionary response to changing environment (Anderson and May 1982, Capaul and Ebert 2003). Under this model, fitness consequences of host-parasite interaction in response to environmental conditions drives the evolutionary response that determines the genetic

makeup of host and parasite populations.
Selection on life-history traits through environmental local adaptation also provides a mechanism for parasites with apparently low virulence to have major consequences on host genetic structure. For example, infection with a parasite may change time of reproduction of the host, leading to demographic change in the host population (Agnew et al. 2000, Chadwick and Little 2005). This would lead to a mismatch between the realised host demography and the host demography that would be selected for under local conditions, essentially creating a pattern of maladaptation that would face strong selective pressure. A parasite that apparently has a low virulence may thus still create strong selection on the host.

References

Agnew, P., J. C. Koella, and Y. Michalakis. 2000. Host life history responses to parasitism. Microbes and Infection 2:891–896.

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

Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White. 2009. Generalized linear mixed models: A practical guide for ecology and evolution. Trends in Ecology & Evolution 24:127–135.

Brooks, M. E., K. Kristensen, K. J. van Benthem, A. Magnusson, C. W. Berg, A. Nielsen, H. J. Skaug, M. Mächler, and B. M. Bolker. 2017. Modeling zero-inflated count data with glmmTMB. Preprint, Ecology.

Bryner, S. F., and D. Rigling. 2011. Temperature-Dependent Genotype-by-Genotype Interaction between a Pathogenic Fungus and Its Hyperparasitic Virus. The American Naturalist 177:65–74.

Capaul, M., and D. Ebert. 2003. Parasite-mediated selection in experimental Daphnia magna populations. Evolution:249–260.

Carius, H. J., T. J. Little, and D. Ebert. 2001. Genetic variation in a host-parasite association: Potential for coevolution and frequency-dependent selection. Evolution 55:1136–1145.

Chadwick, W., and T. J. Little. 2005. A parasite-mediated life-history shift in *Daphnia Magna*. Proceedings of the Royal Society B: Biological Sciences 272:505–509.

Clarke, B. 1976. The ecological genetics of host-parasite relationships. *in* Genetic aspects of host-parasite relationships. Blackwell, London.

Decaestecker, E., S. Declerck, L. De Meester, and D. Ebert. 2005. Ecological implications of parasites in natural Daphnia populations. Oecologia 144:382–390.

Decaestecker, E., S. Gaba, J. A. M. Raeymaekers, R. Stoks, L. Van Kerckhoven, D. Ebert, and L. De Meester. 2007. HostParasite "Red Queen" dynamics archived in pond sediment. Nature 450:870–873.

Ebert, D. 2008. HostParasite coevolution: Insights from the DaphniaParasite model system. Current Opinion in Microbiology 11:290–301.

Ebert, D., M. Lipsitch, and K. L. Mangin. 2000. The Effect of Parasites on Host Population Density and Extinction: Experimental Epidemiology with Daphnia and Six Microparasites. The American Naturalist 156:19.

Gibson, A. K., K. S. Stoy, and C. M. Lively. 2018. Bloody-minded parasites and sex: The effects of fluctuating virulence. Journal of Evolutionary Biology 31:611–620.

Guay, J.-F., S. Boudreault, D. Michaud, and C. Cloutier. 2009. Impact of environmental stress on aphid clonal resistance to parasitoids: Role of Hamiltonella defensa bacterial symbiosis in association with a new facultative symbiont of the pea aphid. Journal of Insect Physiology 55:919–926.

Haag, C. R., and D. Ebert. 2004. ParasiteMediated selection in experimental metapopulations of *Daphnia Magna*. Proceedings of the Royal Society of London. Series B: Biological Sciences 271:2149–2155.

Hamilton, W. D. 1980. Sex versus Non-Sex versus Parasite. Oikos 35:282.

Jaenike, J. 1978. An Hypothesis to account for the maintenance of sex within populations. Evolutionary Theory:191–194.

Jokela, J., M. F. Dybdahl, and C. M. Lively. 2009. The Maintenance of Sex, Clonal Dynamics, and Host-Parasite Coevolution in a Mixed Population of Sexual and Asexual Snails. The American Naturalist 174:S43–S53.

King, K. C., J. Jokela, and C. M. Lively. 2011. Parasites, sex, and clonal diversity in natural snail populations. Evolution 65:1474–1481.

Klüttgen, B., U. Dülmer, M. Engels, and H. Ratte. 1994. ADaM, an artificial freshwater for the culture of zooplankton. Water Research 28:743–746.

Lambrechts, L., S. Fellous, and J. C. Koella. 2006. Coevolutionary interactions between host and parasite genotypes. Trends in Parasitology 22:12–16.

Larsson, J. R., D. Ebert, and J. Vávra. 1997. Ultrastructural study and description of Ordospora colligata gen. Et sp. Nov. (Microspora, ordosporidae fam. Nov.), a new microsporidian parasite of Daphnia magna (Crustacea, Cladocera). European Journal of Protistology 33:432–443.

Lively, C. M. 2010. A Review of Red Queen Models for the Persistence of Obligate Sexual Reproduction. Journal of Heredity 101:S13–S20.

Lüdecke, D. 2020. Sjstats: Statistical Functions for Regression Models (Version 0.18.0).

Mitchell, S. E., E. S. Rogers, T. J. Little, and A. F. Read. 2005. Host-parasite and genotype-by-environment interactions: Temperature modifies potential for selection by a sterilizing pathogen. Evolution 59:70–80.

Nakagawa, S., P. C. D. Johnson, and H. Schielzeth. 2017. The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. Journal of the Royal Society Interface.

Nakagawa, S., and H. Schielzeth. 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods in Ecology and Evolution 4:133–142.

Okamura, B., H. Hartikainen, H. Schmidt-Posthaus, and T. Wahli. 2011. Life cycle complexity, environmental change and the emerging status of salmonid proliferative kidney disease: PKD as an emerging disease of salmonid fish. Freshwater Biology 56:735–753.

Paaijmans, K. P., S. Blanford, B. H. K. Chan, and M. B. Thomas. 2012. Warmer temperatures reduce the vectorial capacity of malaria mosquitoes. Biology Letters 8:465–468.

Salathe, M., R. Kouyos, and S. Bonhoeffer. 2008. The state of affairs in the kingdom of the Red Queen. Trends in Ecology & Evolution 23:439–445.

Seppälä, O., A. Karvonen, C. Rellstab, K.-R. Louhi, and J. Jokela. 2012. Reciprocal Interaction Matrix Reveals Complex Genetic and Dose-Dependent Specificity among Coinfecting Parasites. The American Naturalist 180:306–315.

Studer, A., D. Thieltges, and R. Poulin. 2010. Parasites and global warming: Net effects of temperature on an intertidal hostParasite system. Marine Ecology Progress Series 415:11–22.

Tadiri, C. P., F. Dargent, and M. E. Scott. 2013. Relative host body condition and food availability influence epidemic dynamics: A *Poecilia Reticulata-Gyrodactylus Turnbulli* host-parasite model. Parasitology 140:343–351.

Team, R. 2019. RStudio: Integrated Development Environment for R. RStudio, Inc., Boston, MA.

Team, R. C. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austia.

Tétard-Jones, C., M. A. Kertesz, P. Gallois, and R. F. Preziosi. 2007. Genotype-by-Genotype Interactions Modified by a Third Species in a Plant-Insect System. The American Naturalist 170:492–499.

Turko, P., C. Tellenbach, E. Keller, N. Tardent, B. Keller, P. Spaak, and J. Wolinska. 2018. Parasites driving host diversity: Incidence of disease correlated with *Daphnia* clonal turnover*. Evolution 72:619–629.

Vale, P. F., A. J. Wilson, A. Best, M. Boots, and T. J. Little. 2011. Epidemiological, Evolutionary, and Coevolutionary Implications of Context-Dependent Parasitism. The American Naturalist 177:510–521.

Wolinska, J., and K. C. King. 2009. Environment can alter selection in hostParasite interactions. Trends in Parasitology 25:236–244.

General Conclusions

In this thesis I studied the context dependency of negative frequency-dependent host-parasite coevolution. Host-parasite coevolution does not happen in isolation from external processes. Host and parasite populations each have different sizes, are embedded in different types of metapopulations and each population faces different environmental conditions. Every interaction between hosts and parasites hence has its own context. I asked if this context influences the coevolutionary dynamics between hosts and parasites.

My results in chapter 2 showed, that in small populations drift and negative frequency-dependent selection (NFDS) by parasites can interact. This interaction boosted the effect of drift and led to the faster-than-drift degradation of host genetic diversity. The population sizes at which this effect was visible in chapter 2 should not be seen as an absolute value. I do not expect that parasites will automatically maintain genetic diversity in populations larger than 2000 individuals, or will always degrade diversity in populations smaller than 200 individuals. Chapter 2 illustrates that in any system, when population size is getting smaller and drift is getting stronger, at one point there is the possibility that NFDS and drift will start interacting in a way that is detrimental to genetic diversity. The exact population size when this point will be reached is dependent on many factors, for example on the virulence of the parasite. This result means that in populations of declining size, there may be a relatively rapid switch from genetic diversity being maintained by parasites counteracting drift to genetic diversity being degraded by parasites boosting the effect of drift. This would lead to an abrupt decline in the genetic diversity, kick starting other processes that small and declining populations face.

In chapter 3 I showed that the outcome of NFDS can be modified greatly by the connectedness of parasite subpopulations in a metapopulation context. I showed that strong gene-flow of parasites between isolated host populations can lead to very strong host population divergence and a complete compartmentalisation of genetic diversity into host subpopulations. I propose that this could lead to a pattern of apparent adaptation of distinct host genotypes on environmental conditions in host subpopulations that is entirely driven by coevolutionary interactions with the parasite. Being aware of the effect that the pattern of parasite gene-flow has on the pattern of genetic diversity in the host should enable better understanding of biodiversity patterns found in the wild.
Studies of local adaptation to environmental conditions should take into account that host-parasite interactions can create a similar pattern to environmental local adaptation. Warning signals of this could be the presence of parasites that are potentially more mobile than the host. Or signatures of selection on known immune loci or pathways in the host.
Investigations of host-parasite coevolution especially should consider the effect of metapopulation connectedness. If for example gene-flow in the parasite is not accounted for, reciprocal transplant experiments might show no pattern of local adaptation in infection rate or virulence, leading to misinterpretations about the coevolutionary process. Or if studies on NFDS are conducted, it is possible that no response in the genotype frequencies of the host will be found, despite finding fluctuations in parasite genotype frequencies. This is possible when the host population level at which the effect is examined does not correspond to the population range of the parasite.

In chapter 4 I found strong GxGxE interactions in an experiment using *Dapnia magna* as the host and *Ordospora colligata* as the parasite. I found that the effect that the parasite had on the host was dependent on both the host and the parasite genotype as well as on the temperature of the environment in which the interaction was taking place. This GxGxE interaction resulted in changes in the hosts demography, shifting the proportion of juveniles to adults. Changes in host demography are at the same time a subtle and a very consequential effect of infection by a parasite. Realizing that the outcome of parasite infection was context-dependent on such a simple parameter as temperature should serve as an example that host-parasite coevolution is a very complex matter. One should be aware that the outcome of a certain host-parasite interaction can change with the environmental conditions. This means that expectations for outcomes of host-parasite interactions that have been built up by experiments or field studies in specific conditions might not be directly transferable to the outcome of the same interaction in different conditions. If the environmental conditions are fluctuating, the outcome of host-parasite coevolution within the same location can fluctuate over time as well, interfering with NFDS and leading to unpredictable outcomes.

Taken together, my findings showed that the outcome of host-parasite coevolution may be highly context-dependent. I showed that interactions of coevolution with several other processes can overlay the signal

that coevolution leaves in host and parasite genotype frequencies. My results put emphasis on the question of how the signal of coevolution can disappear even if coevolution is taking place. Context-dependent interference may alter the signal that one might be expecting from coevolution, so that it may become difficult to detect empirically. This might be especially pronounced if the experimental design does not incorporate, for example, asymmetric population structure of the host and the parasite.
This could be part of the explanation on why there is, despite substantial congruence, still a great deal of theory in host-parasite coevolution that has been difficult to address comprehensively in either natural populations or using controlled laboratory experiments.

Acknowledgements

Foremost I would like to thank Jukka Jokela. I really appreciated having him as my supervisor. Without him this thesis would not have been possible. He always encouraged me to dream big and pursue the important questions, and he never failed in supporting me.

I would like to thank Kyra Gähwiler for all the support she has given me from the start of this thesis to its end and for going with me through this adventure and all the others.

My heartfelt thanks goes to my fellow PhD students at Eawag; Moritz Lürig, Natalia Zajac, Natalie Sieber, Teo Cereghetti, Heidi Käch and all the others. They made the hard parts of the thesis bearable and the fun parts so much more fun.

I thank Jessica Stephenson, Hugo Mathé-Hubert, Anja Felmy, Lukas DeVentura, Hanna Hartikainen, Kirsten Klappert, Emanuel Fronhofer and Stuart Dennis for inspiring and challenging discussions. I thank Arianne Maniglia and Gioia Matheson for taking care of the administrative parts effectively and with fun. The whole Eco department made this thesis unique, they are a great bunch of people.

I would like to thank Dieter Ebert, who provided guiding as well as the organisms that I used for my experiments, most of which did not even make it into this thesis. And Pepijn Luijckx for helpful discussions.

All the helpers that assisted me with counting, feeding and pampering thousands and thousands of *Daphnia* in the climate chambers, especially Aijuan Liao and Alena Streiff as well as Anja Taddei and Nora Weissert.

The team at GDC, notably Aria Minder and Silvia Kobel, who helped me with all the countless hours of molecular lab work.

The adaptation to a changing environment center (ACE) at ETH for its generous grant that started my PhD, for all the cool seminars and lunches and especially for all its people.

To all those people that crossed my path during the years of my PhD and I forgot to mention, goes my thanks as well.

Appendix chapter 3

Full linear model outputs of the models used in Table 2 of chapter 3

Full model output of the linear model between the parasite genotype frequency and the host genotype frequency at the subpopulation level with a local parasite.

```
##
## Call:
## lm(formula = c(Total.Parasite.Number.Individuals/Total.Parasite.Population.Size) ~
##     c(Host.Number.Individuals/Host.Population.Size) + Popsize)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.145798 -0.028306 -0.000221  0.033834  0.164758
##
## Coefficients:
##                                                 Estimate Std. Error t value
## (Intercept)                                    -0.007858   0.019772  -0.397
## c(Host.Number.Individuals/Host.Population.Size)  1.028651   0.041985  24.500
## Popsize400                                      0.003468   0.016715   0.207
##                                                Pr(>|t|)
## (Intercept)                                       0.693
## c(Host.Number.Individuals/Host.Population.Size)   <2e-16 ***
## Popsize400                                        0.836
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0539 on 55 degrees of freedom
## Multiple R-squared:  0.9283, Adjusted R-squared:  0.9257
## F-statistic: 355.9 on 2 and 55 DF,  p-value: < 2.2e-16
```

Full model output of the linear model between the parasite genotype frequency and the host genotype frequency at the subpopulation level with a global parasite.

```
##
## Call:
## lm(formula = c(Total.Parasite.Number.Individuals/Total.Parasite.Population.Size) ~
##     c(Host.Number.Individuals/Host.Population.Size) + Popsize)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.108546 -0.018765  0.001112  0.021765  0.118670
##
## Coefficients:
##                                                Estimate Std. Error t value
## (Intercept)                                     0.15326    0.01349  11.362
## c(Host.Number.Individuals/Host.Population.Size)  0.13786    0.02487   5.543
## Popsize400                                      0.01800    0.01279   1.408
##                                                Pr(>|t|)
## (Intercept)                                    1.05e-15 ***
## c(Host.Number.Individuals/Host.Population.Size) 1.01e-06 ***
## Popsize400                                        0.165
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04378 on 52 degrees of freedom
## Multiple R-squared:  0.3716, Adjusted R-squared:  0.3474
## F-statistic: 15.37 on 2 and 52 DF,  p-value: 5.686e-06
```

Full model output of the linear model between the parasite genotype frequency and the host genotype frequency at the metapopulation level with a local parasite.

```
##
## Call:
## lm(formula =
c(Total.Parasite.Number.Individuals.Between/Total.Parasite.Population.Size.Between) ~
## c(Host.Number.Individuals.Between/Host.Population.Size.Between) +
## Popsize)
##
## Residuals:
## Min 1Q Median 3Q Max
## -1.4919 -0.2317 -0.0720 0.0063 4.8555
##
## Coefficients:
## Estimate
## (Intercept) 1.5779
## c(Host.Number.Individuals.Between/Host.Population.Size.Between) -2.2214
## Popsize400 -0.8714
## Std. Error
## (Intercept) 0.3756
## c(Host.Number.Individuals.Between/Host.Population.Size.Between) 1.3339
## Popsize400 0.2628
## t value
## (Intercept) 4.201
## c(Host.Number.Individuals.Between/Host.Population.Size.Between) -1.665
## Popsize400 -3.316
## Pr(>|t|)
## (Intercept) 9.81e-05 ***
## c(Host.Number.Individuals.Between/Host.Population.Size.Between) 0.10154
## Popsize400 0.00162 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9192 on 55 degrees of freedom
## Multiple R-squared: 0.1866, Adjusted R-squared: 0.1571
## F-statistic: 6.31 on 2 and 55 DF, p-value: 0.00341
```

Full model output of the linear model between the parasite genotype frequency and the host genotype frequency at the metapopulation level with a global parasite.

```
##
## Call:
## lm(formula =
c(Total.Parasite.Number.Individuals.Between/Total.Parasite.Population.Size.Between) ~
## c(Host.Number.Individuals.Between/Host.Population.Size.Between) +
## Popsize)
##
## Residuals:
## Min 1Q Median 3Q Max
## -0.144433 -0.027063 -0.004962 0.025427 0.124140
##
## Coefficients:
## Estimate
## (Intercept) 0.022279
## c(Host.Number.Individuals.Between/Host.Population.Size.Between) 0.948886
## Popsize400 -0.009785
## Std. Error
## (Intercept) 0.038638
## c(Host.Number.Individuals.Between/Host.Population.Size.Between) 0.191265
## Popsize400 0.013364
## t value
## (Intercept) 0.577
## c(Host.Number.Individuals.Between/Host.Population.Size.Between) 4.961
## Popsize400 -0.732
## Pr(>|t|)
## (Intercept) 0.567
## c(Host.Number.Individuals.Between/Host.Population.Size.Between) 7.89e-06 ***
## Popsize400 0.467
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04747 on 52 degrees of freedom
## Multiple R-squared: 0.3219, Adjusted R-squared: 0.2959
## F-statistic: 12.34 on 2 and 52 DF, p-value: 4.101e-05
```

Full outputs of the Pearson‘s product-moment correlation

Pearson‘s product-moment correlation between the parasite genotype frequency and the host genotype frequency at the subpopulation level with a local parasite.

```
##
## Pearson's product-moment correlation
##
## data: c(Total.Parasite.Number.Individuals/Total.Parasite.Population.Size) and
c(Host.Number.Individuals/Host.Population.Size)
## t = 26.908, df = 56, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.9387509 0.9782849
## sample estimates:
## cor
## 0.9634369
```

Pearson's product-moment correlation between the parasite genotype frequency and the host genotype frequency at the subpopulation level with a global parasite.

```
##
## Pearson's product-moment correlation
##
## data: c(Total.Parasite.Number.Individuals/Total.Parasite.Population.Size) and
c(Host.Number.Individuals/Host.Population.Size)
## t = 5.3141, df = 53, p-value = 2.177e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.3844156 0.7392515
## sample estimates:
## cor
## 0.5895844
```

Pearson's product-moment correlation between the parasite genotype frequency and the host genotype frequency at the metapopulation level with a local parasite.

```
##
## Pearson's product-moment correlation
##
## data:
c(Total.Parasite.Number.Individuals.Between/Total.Parasite.Population.Size.Between) and
c(Host.Number.Individuals.Between/Host.Population.Size.Between)
## t = -1.1748, df = 56, p-value = 0.245
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.3974641 0.1075122
## sample estimates:
## cor
## -0.1550905
```

Pearson's product-moment correlation between the parasite genotype frequency and the host genotype frequency at the metapopulation level with a global parasite.

```
##
## Pearson's product-moment correlation
##
## data:
c(Total.Parasite.Number.Individuals.Between/Total.Parasite.Population.Size.Between) and
c(Host.Number.Individuals.Between/Host.Population.Size.Between)
## t = 4.9363, df = 53, p-value = 8.308e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.3476709 0.7193941
## sample estimates:
## cor
## 0.5612047
```

Additional figures of replicates corresponding to Figure 3 in chapter 3

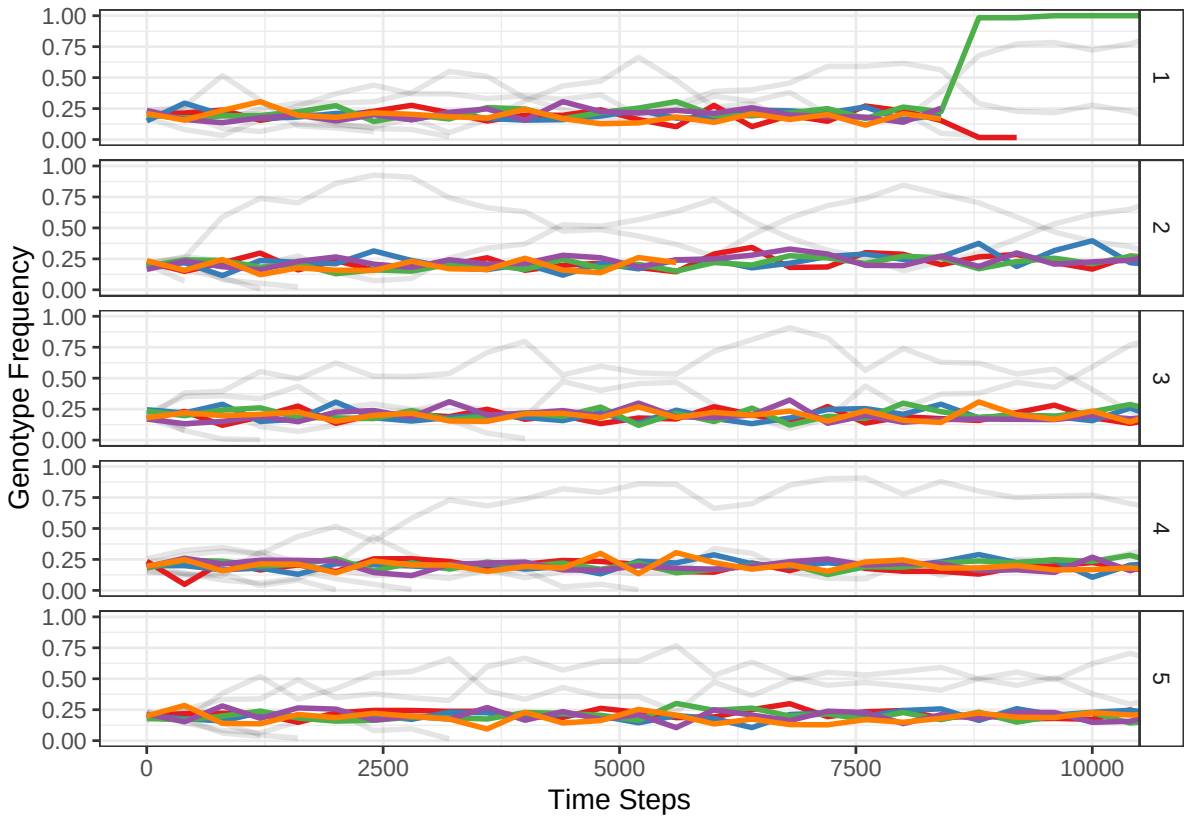

Figure 1: Host genotype frequencies within subpopulations over time. Five host subpopulations that each have a locally coevolving parasite subpopulation are shown. All subpopulation show the same NFDS dominated dynamics. No host genotype can increase or decrease in frequency as the parasite excerts control. The baseline under drift is in grey in the background, without resolving genotypes. Each subpopulation has about 800 individuals. Different colours depict different host genotypes.

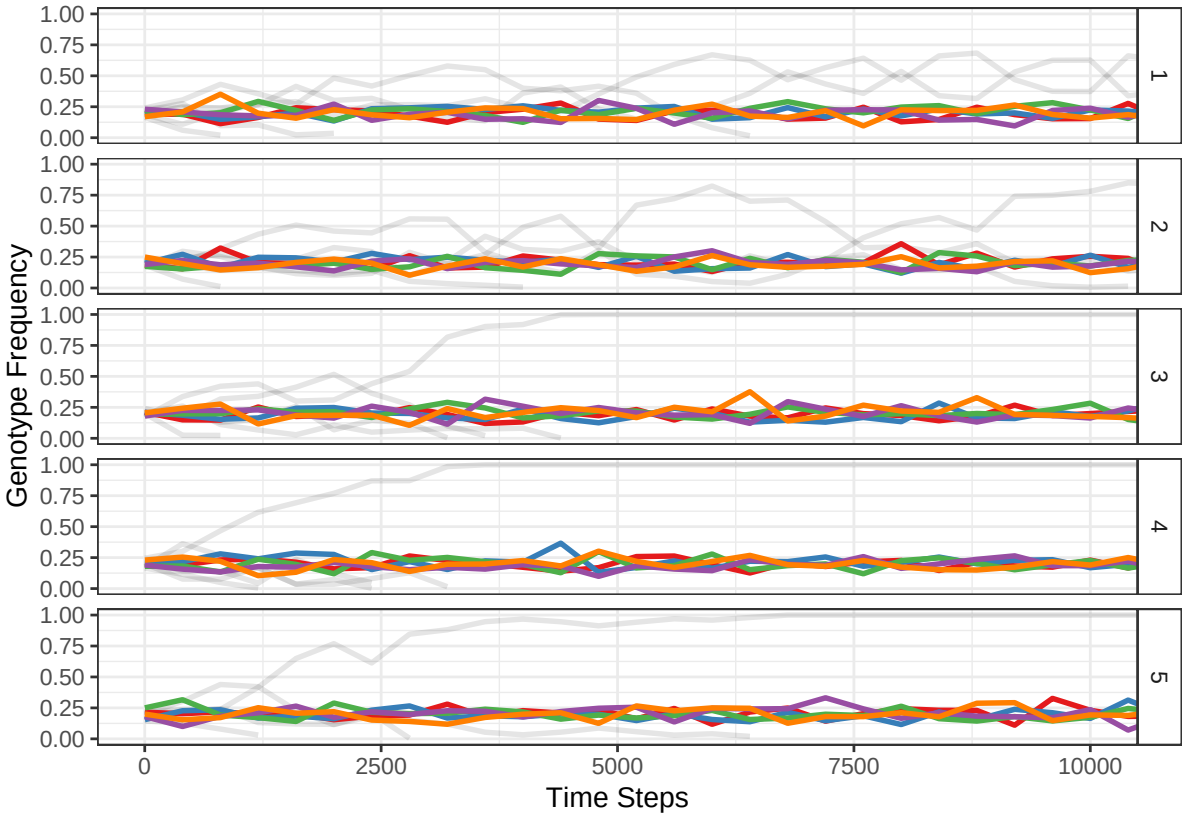

Figure 2: Host genotype frequencies within subpopulations over time. Five host subpopulations that each have a locally coevolving parasite subpopulation are shown. All subpopulation show the same NFDS dominated dynamics. No host genotype can increase or decrease in frequency as the parasite excerts control. The baseline under drift is in grey in the background, without resolving genotypes. Each subpopulation has about 800 individuals. Different colours depict different host genotypes.

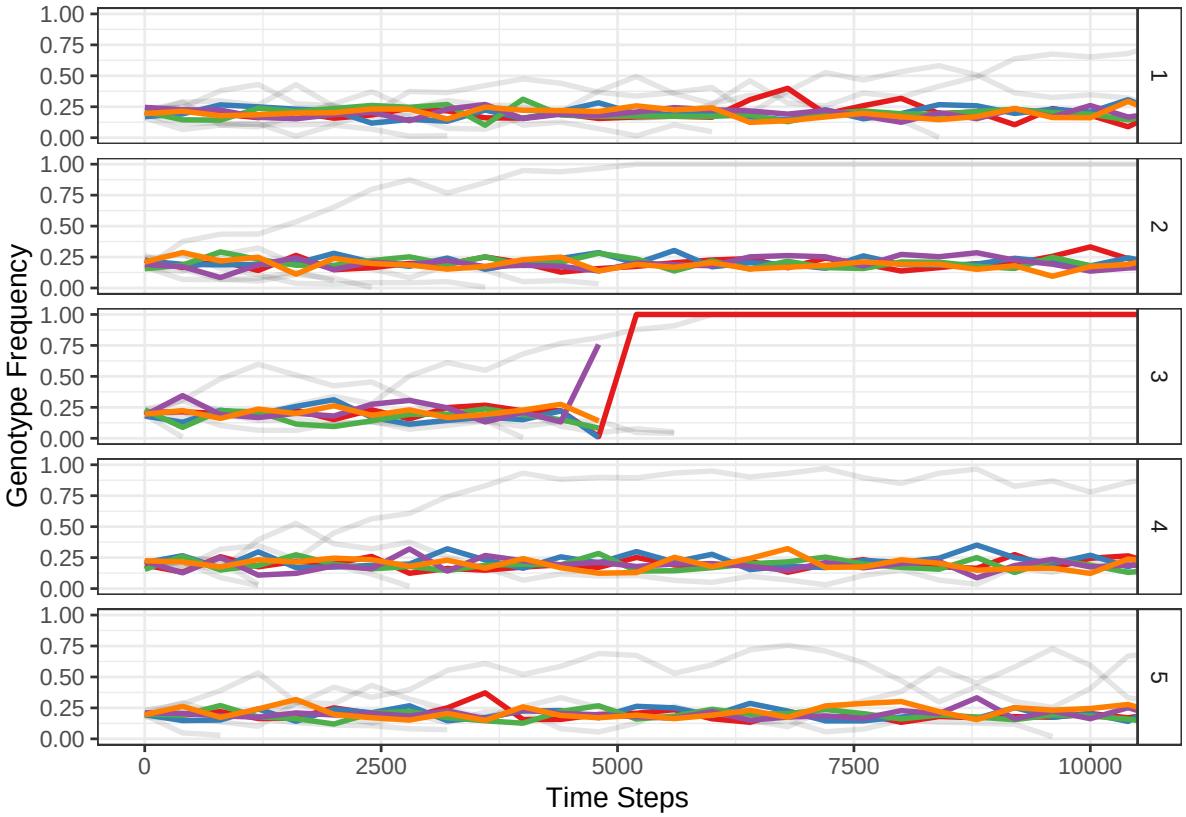

Figure 3: Host genotype frequencies within subpopulations over time. Five host subpopulations that each have a locally coevolving parasite subpopulation are shown. All subpopulation show the same NFDS dominated dynamics. No host genotype can increase or decrease in frequency as the parasite excerts control. The baseline under drift is in grey in the background, without resolving genotypes. Each subpopulation has about 800 individuals. Different colours depict different host genotypes.

Additional figures of replicates corresponding to Figure 4 in chapter 3

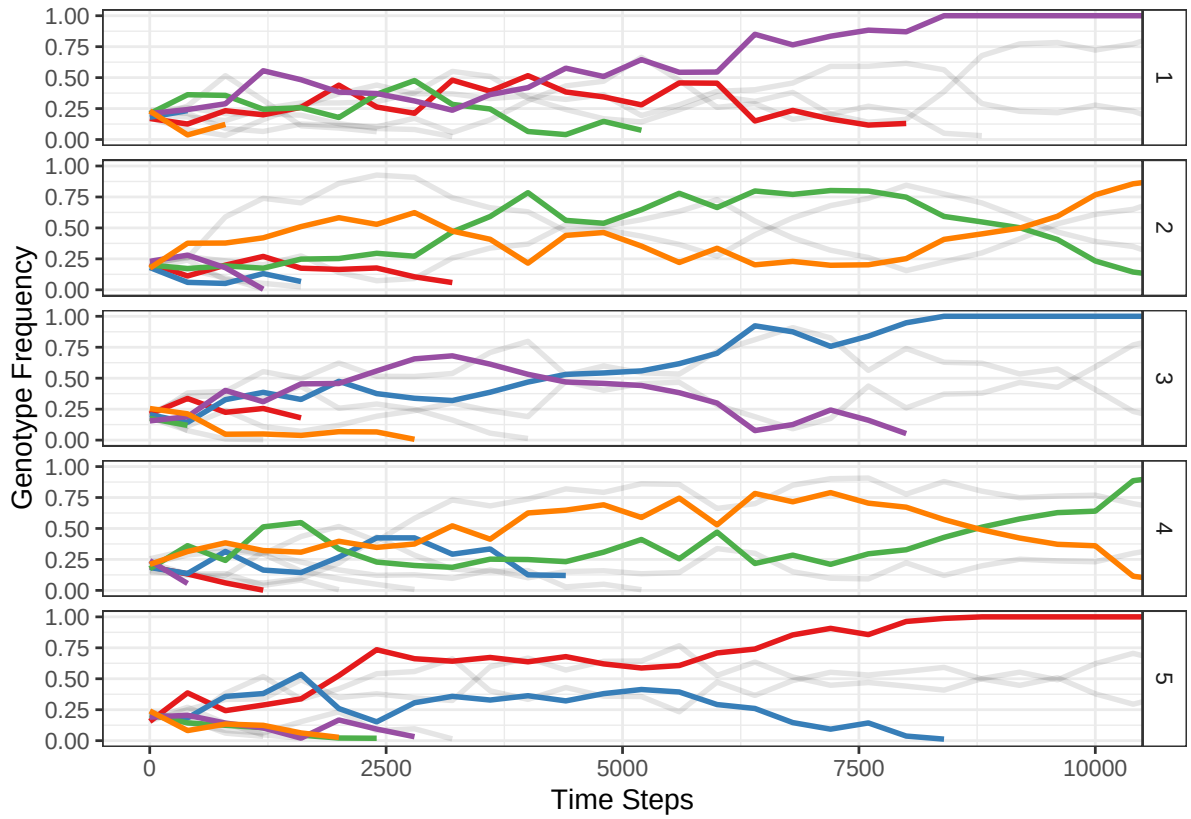

Figure 4: Host genotype frequencies within subpopulations over time. Five host subpopulations that are all coevolving with the same global parasite metapopulation are shown. There is no evidence of negative frequency-dependent selection. All host subpopulations quickly fix for a single host genotype. Note how each subpopualtion fixes for a different host genotype. The baseline under drift is in grey in the background, without resolving genotypes. Each subpopulation has about 800 individuals. Different colours depict different host genotypes.

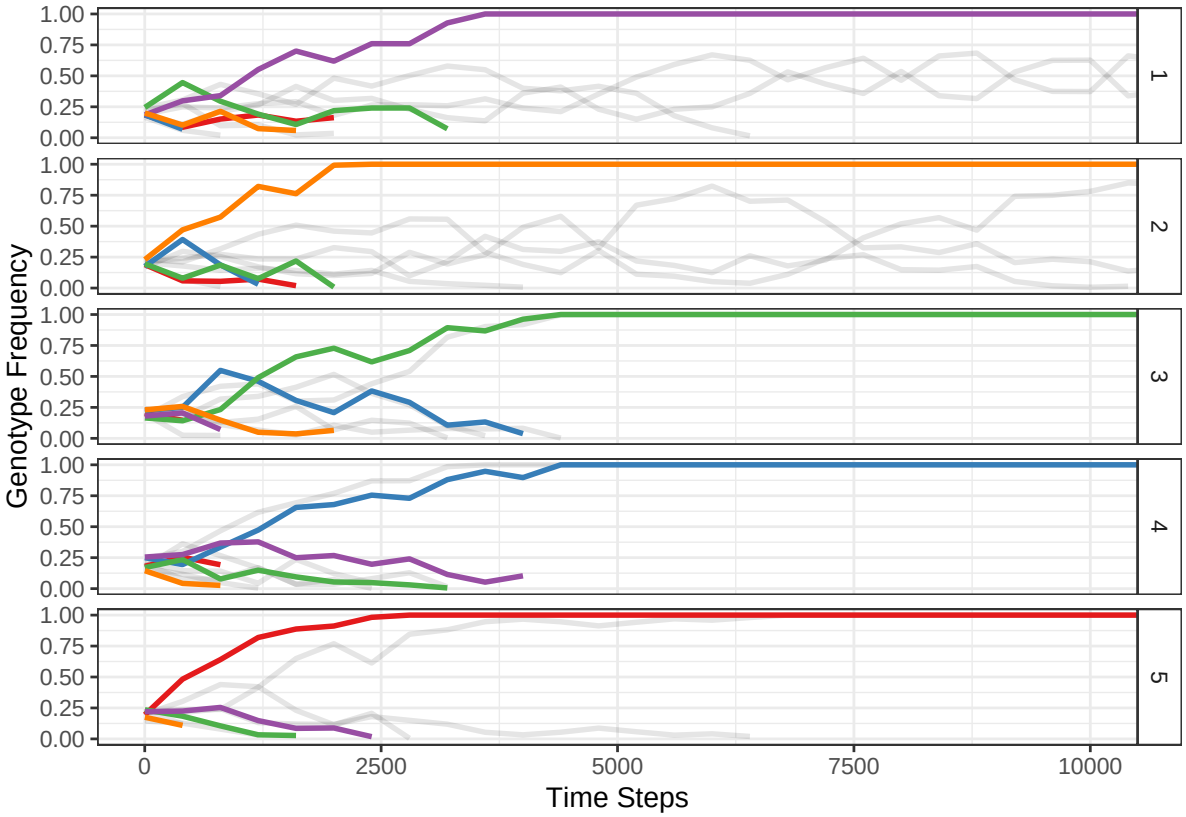

Figure 5: Host genotype frequencies within subpopulations over time. Five host subpopulations that are all coevolving with the same global parasite metapopulation are shown. There is no evidence of negative frequency-dependent selection. All host subpopulations quickly fix for a single host genotype. Note how each subpopualtion fixes for a different host genotype. The baseline under drift is in grey in the background, without resolving genotypes. Each subpopulation has about 800 individuals. Different colours depict different host genotypes.

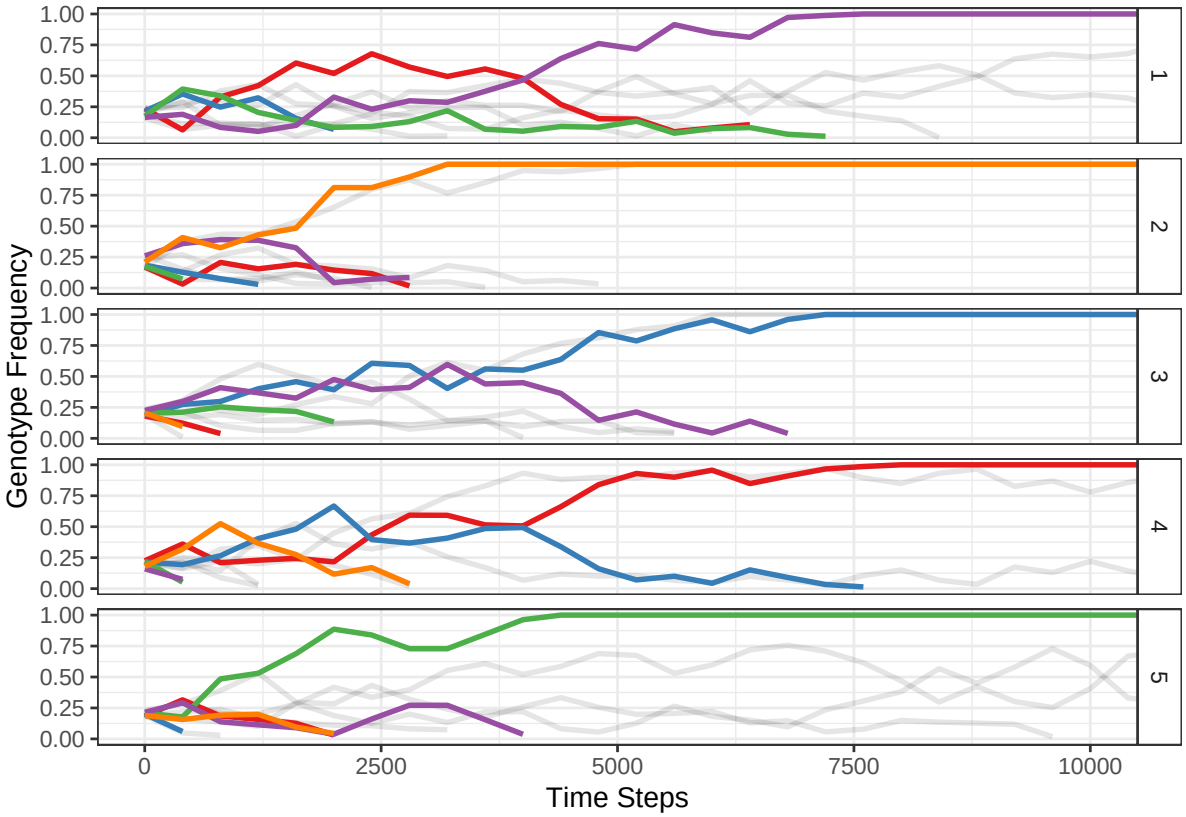

Figure 6: Host genotype frequencies within subpopulations over time. Five host subpopulations that are all coevolving with the same global parasite metapopulation are shown. There is no evidence of negative frequency-dependent selection. All host subpopulations quickly fix for a single host genotype. Note how each subpopualtion fixes for a different host genotype. The baseline under drift is in grey in the background, without resolving genotypes. Each subpopulation has about 800 individuals. Different colours depict different host genotypes.

Appendix chapter 4

Table A1: Model fit of the glmm

AIC	BIC	Log Likelihood	Deviance	Df. residuals
7548.5	7704.6	-3742.2	7484.5	940

Table A2: Random intercepts of glmm

Group	Variance	Std. Deviation
Jar	<0.001	0.010
Date	0.061	0.248

Table A3: Negative binomial glmm, log link, standard R dummy coding, uses the control treatment in the EP genotype at 20°C as reference class.

Term	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	**3.783**	**0.210**	**17.987**	**<0.001**
Temperature	0.352	0.279	1.263	0.207
Sympatric	-0.055	0.222	-0.246	0.805
Allopatric	0.062	0.213	0.292	0.771
Clone SKW	-0.095	0.105	-0.911	0.362
Clone OER	**0.278**	**0.091**	**3.047**	**0.002**
Adult population	-0.020	0.013	-1.528	0.127
Ephippia	0.061	0.236	0.259	0.796
Spore dose	0.003	0.006	0.475	0.635
Block 2	0.050	0.029	1.715	0.086
Block 3	**0.081**	**0.031**	**2.606**	**0.009**
Temperature : Sympatric	-0.491	0.290	-1.689	0.091
Temperature : Allopatric	-0.433	0.284	-1.526	0.127
Temperature : Clone SKW	**-0.397**	**0.147**	**-2.698**	**0.007**
Temperature : Clone OER	**-0.520**	**0.135**	**-3.867**	**<0.001**
Sympatric : Clone SKW	0.036	0.140	0.257	0.797
Allopatric : Clone SKW	0.049	0.135	0.363	0.717
Sympatric : Clone OER	0.118	0.123	0.961	0.337
Allopatric : Clone OER	-0.096	0.119	-0.803	0.422
Temperature : Adult population	-0.025	0.014	-1.783	0.075
Sympatric : Adult population	-0.002	0.017	-0.097	0.923
Allopatric : Adult population	-0.010	0.016	-0.610	0.542
Temperature : Ephippia	-0.123	0.236	-0.520	0.603
Temperature : Sympatric : Clone SKW	**0.555**	**0.200**	**2.770**	**0.006**
Temperature : Allopatric : Clone SKW	0.135	0.203	0.666	0.506
Temperature : Sympatric : Clone OER	0.060	0.188	0.318	0.751
Temperature : Allopatric : Clone OER	**0.502**	**0.191**	**2.625**	**0.009**
Temperature : Sympatric : Adult population	0.024	0.019	1.267	0.205
Temperature : Allopatric : Adult population	0.020	0.017	1.126	0.260

Table A4: Results of the ANOVA

Term	Chisq	Df	Pr(>Chisq)
Intercept	**323.514**	**1**	**<0.001**
Temperature	1.594	1	0.207
Treatment	0.370	2	0.831
Clone	**36.914**	**2**	**<0.001**
Adult population	2.334	1	0.127
Ephippia	0.067	1	0.796
Spore dose	0.226	1	0.635
Block	**7.075**	**2**	**0.029**
Temperature : Treatment	3.360	2	0.186
Temperature : Clone	**14.965**	**2**	**<0.001**
Treatment : Clone	8.075	4	0.089
Temperature : Adult population	3.181	1	0.075
Treatment : Adult population	0.465	2	0.793
Temperature : Ephippia	0.270	1	0.603
Temperature : Treatment : Clone	**39.710**	**4**	**<0.001**
Temperature : Treatment : Adult population	1.860	2	0.395

Appendix B

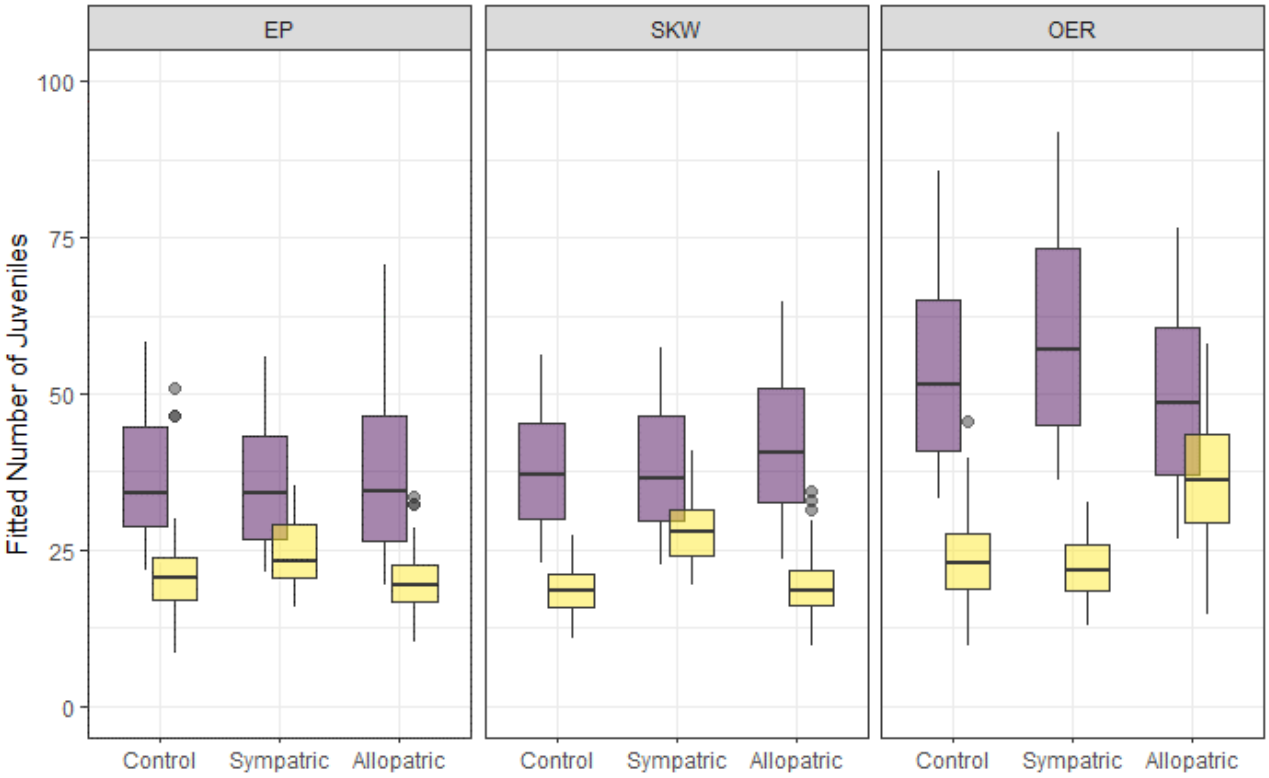

Figure B1: Boxplots of model fitted values of juvenile population sizes, split by treatment, temperature and clone. Lilac is 20°C and yellow is 16°C.

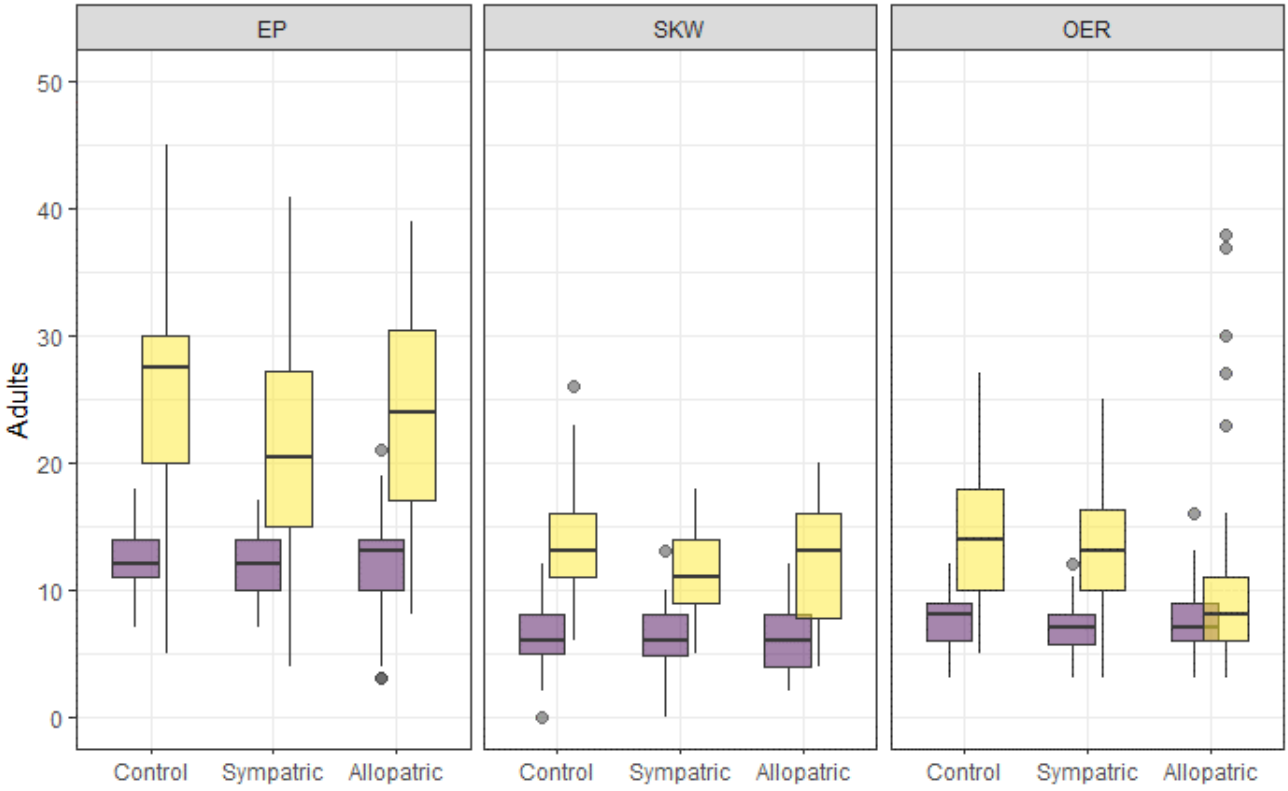

Figure B2: Boxplots of adult population sizes, raw data, split by treatment, temperature and clone. Lilac is 20°C and yellow is 16°C.

Robert Pascal Dünner

PhD Candidate

Lavaterstrasse 73, 8002 Zürich

+41 79 582 58 76 | robert.duenner@gmail.com | RobertDuenner

Education

PhD Candidate

ETH Zurich, Department of Environmental Systems Science and EAWAG, Department of Aquatic Ecology, supervisor: Jukka Jokela — *2015 - now*

Diploma of Advanced Studies ETH in Applied Statistics

ETH Zurich, Department of Mathematics — *2015 - 2017*

Master of Science ETH in Biology: Major in Ecology & Evolution

ETH Zurich, Department of Biology — *2011 - 2013*

Bachelor of Science ETH in Biology

ETH Zurich, Department of Biology — *2007 - 2011*

Key Skills

- **Planning**: Identification of relevant questions, perception of necessary experimental design incl. blocking and power analysis, resource planning, critical path analysis.
- **Research**: Fast and thorough aquisation of knowledge in new field through screening of relevant scientific literature.
- **Experiments**: Execution of experiments at different time scales, establishment of micro- and mesocosms, field experiments.
- **Field studies**: Collection of data from wild systems via appropriate sampling schemes and methods, expedition planning.
- **Molecular laboratory**: DNA extraction with different methods, gel electrophoresis, multiplex PCR, genotyping with microsatellites.
- **Statistics**: Applied statistics, especially; Standard hypothesis testing, regression, analysis of variance, time series analysis, multivariate statistics, repeated measures, generalized linear mixed models, hidden markov chains, descriptive statistics.
- **Programming**: Proficient in R, especially; base, apply family, data.table, ggplot, hierfstat, adegenet, rmarkdown, vitae, msm, parallelisation via parallel, construction of stochastic agent based time-forward simulations from scratch. Use of high performance computing clusters via LSF task arrays.
- **Reporting** Dissemination of gained knowledge and results in scientific journals, presentation of results for lay or specialist audiences, automated dynamic report generation with Rmarkdown.

Languages

- German: C2
- English: C2
- French: B1 - B2

Publications

1. Leventhal, GE, RP Dünner, and SM Barribeau (Apr. 2014). Delayed Virulence and Limited Costs Promote Fecundity Compensation upon Infection. en. *The American Naturalist* **183**(4), 480–493.

Experience

PhD Candidate: Perceiving, planning, conducting, analysing and reporting of various scientific projects

Institute of Integrative Biology ETH and Aquatic Ecology EAWAG — *2015 - now*

Field Assistant: Large scale sampling for long term dataset and field experiments in the New Zealand mud snail system

With Jukka Jokela, New Zealand — *2019*

Member of organising commitee for the 7th and 8th ECO PhD Symposium

Aquatic Ecology EAWAG — *2018 - 2019*

Member of institute council

Institute of Integrative Biology ETH — *2016-2020*

Member of organising commitee for the Monte Verita Conference of the Adaptation to a Changing Environment Center

Adaptation to a Changing Environment Center ETH — *2016*

Scientific Assistant: Organisation of the Biology 2015 Conference, assisting with various experimental procedures

Aquatic Ecology EAWAG — *2014 - 2015*

Sales of technical outdoor equipment

Transa Travel & Outdoor — *2014*

Longboard world trip

Travels by Longboard in Latvia, Lithuania, Estonia, Russia, China, Vietnam, Lao, Cambodia, Thailand, Malaysia, Singapore, New Zealand — *2013 - 2014*

Scientific Employee: Assisting with various experimental procedures

Aquatic Ecology EAWAG — *2013*

Research Assistant: Assisting with various experimental procedures

Aquatic Ecology EAWAG — *2012*

Research Assistant: Breeding of Zooplankton for Laboratory Courses

Teaching Aquatic Ecology EAWAG — *2010*

Research Assistant: Assisting with various experimental procedures, maintenance of Arabidopsis culture in greenhouse

Plant Pathology ETH — *2008 - 2009*

Warehousing and logistics

Peek & Cloppenburg Zurich — *2007*

Teaching

Evolutionary Biology: Field Course

Mentoring and assisting with experimental design and statistical analysis — *2019*

Seminar Environmental Systems

Mentoring and assisting with reports and presentations — *2019*

Biodiversity Excursions

Mentoring and assisting with identification of aquatic macroinvertebrates — *2018 - 2019*

Fundamental Questions in Environmental Science

Mentoring and assisting with reports and presentations — *2017*

Bachelor Thesis Nicolas Meier

Co-supervision of Bachelor Thesis — *2017*

Posters and Presentations

Congress of the European Society for Evolutionary Biology, Turku, Finnland

Poster: Coevolution in a hierarchically structured host-parasite metapopulation *2019*

Modelling Ecology & Evolution Zurich, Zurich, Switzerland

Presentation: Sharing parasites *2019*

Closing Conference of the Adaptation to a Changing Environment Center, Borgen/Hocken, Switzerland

Presentation: Moving parasites accross host ranges *2018*

7th annual ECO PhD Symposium, Dubendorf, Switzerland

Presentation: Parasites, population structure and the role of demography *2018*

Congress of the European Society for Evolutionary Biology, Groningen, Netherlands

Presentation: The strange behaviour of genotype frequencies in a hierarchically structured host-parasite metapopulation *2017*

Joint Meeting of the Society for the Study of Evolution, the American Society of Naturalists and the Society of Systematic Biologists, Portland OR, USA

Presentation: The strange behaviour of genotype frequencies in a hierarchically structured host-parasite metapopulation *2017*

Monte Verita Conference of the Adaptation to a Changing Environment Center, Ascona, Switzerland

Poster: Parasite Ping-Pong; Or how sharing a parasite population can make host populations diverge faster *2016*

5th annual ECO PhD Symposium, Dubendorf, Switzerland

Presentation: Parasite Ping-Pong *2016*

Symposium of the Adaptation to a Changing Environment Center, Zurich, Switzerland

Presentation: Parasite Ping-Pong *2015*

Zurich Interaction Seminar, Zurich, Switzerland

Presentation: Parasite Ping-Pong *2015*

Additional Qualifications

Karch introductory course amphibians *2016*

Karch introductory course reptilians *2015*

PADI open water diver and nitrox *2012*

ETH introductory course dragonflies *2010*

Sportsfishing certificate of competence *2010*

Grants

PhD Fellowship: Adaptation to a Changing Environment Center (ACE), ETH Zürich (150 000 CHF) *2015*