Persisting in an ever-changing world

Integrating plastic and genetic responses across the life cycle

Marjolein Bruijning

Persisting in an ever-changing world

Integrating plastic and genetic responses across the life cycle

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen

in het openbaar te verdedigen op dinsdag 5 februari 2019
om 14.30 uur precies

door
Marjolein Bruijning
geboren op 17 mei 1988
te Leiden

Promotor:
Prof. dr. H. Siepel

Copromotor:
Dr. ir. E. Jongejans

Manuscriptcommissie:
Prof. dr. J.C.J.M de Kroon
Prof. dr. J. Ellers (Vrije Universiteit Amsterdam)
Dr. S. DeClerck (Nederlands Instituut voor Ecologie, Wageningen)

Persisting in an ever-changing world

Integrating plastic and genetic responses across the life cycle

Doctoral Thesis

**to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. J.H.J.M. van Krieken,
according to the decision of the Council of Deans**

**to be defended in public on Tuesday, February 5, 2019
at 14:30 hours**

**by
Marjolein Bruijning
born on May 17, 1988
in Leiden (the Netherlands)**

Contents

General introduction

Hence, as more individuals are produced than can possibly survive, there must in every case be a struggle for existence, either one individual with another of the same species, or with the individuals of distinct species, or with the physical conditions of life.

(Darwin 1859)

We live in a constantly changing world. All populations, whether they are plants, animals or microbes, have to cope with these changing conditions, resulting in a never-ceasing selection on phenotypes[1] that are most successful[2] in the experienced environment[3]. Phenotypes are the product of gene expression, which is influenced by both genetic and non-genetic factors. Ultimately, natural selection will favour those strategies[4] that maximize long-term fitness, via selection on individual

---

[1]A phenotype consists of all individual properties, such as morphological, chemical and behavioural traits. Phenotypic traits have the potential to affect fitness of the focal individual, and/or of relatives.

[2]Throughout this chapter, *being most successful* is used as a synonym for being fittest, or having the highest fitness, and these terms are used interchangeably. Fitness measures the rate 'at which a genotype is able to propagate itself into future generations' (Caswell 2001).

[3]The environment is defined in a broad sense, including both biotic factors, such as effects of predation, symbiosis and competition from conspecifics, and abiotic factors such as temperature and available nutrients.

[4]The terms *population*, *genotype*, and *strategy*, although related, are not identical, and the nuances are important in this introduction. A population is a group of individuals of the same species, inhabiting the same place and competing for space and resources. It generally consists of multiple genotypes. The term strategy is used more abstract, and refers to the way a genotype realizes its fitness, via phenotypic traits. Different genotypes could in principle lead to the same strategy, and the same strategy can lead to different phenotypes (by phenotypic plasticity or intra-genotypic variability, as introduced in this chapter).

phenotypes. Which strategies are beneficial, and under which conditions?

1.1   Phenotypic plasticity and rapid evolution

Individuals change, to some degree, their phenotype in response to the environment, and this is called phenotypic plasticity. For instance, many species have shifted their phenology, such as birds adjusting their egg laying date in response to global warming (Visser and Both 2005). In a hypothetical world where there would be no limits and costs associated with a plastic response, plasticity is an optimal strategy. In reality, however, there are clearly developmental, genetic and physical constraints on plastic responses (a pig will never plastically grow wings, no matter how beneficial it might be). A plastic strategy may also be costly, and finally, environmental cues on the optimal phenotype, are not always reliable (DeWitt et al. 1998; Auld et al. 2009).

Phenotypic plasticity is not the only mechanism by which phenotypes in a population respond to continuously changing environments. When some individuals are more successful than others because of their genetic makeup, natural selection will favour these genotypes. The result is an evolutionary response, in which the population-level mean phenotype will evolve towards the optimal phenotype (Arnold et al. 2001), provided that there are no genetic constraints (Lande and Arnold 1983; Arnold 1992). Over long time scales, evolution is the process that has led to the overwhelming phenotypic variation and enormous species diversity that surrounds us.

Over time, both evolution and plasticity thus change the distribution of individual phenotypes within a population. Evolution, however, is traditionally thought of as a process that acts slowly, on long time scales, as it requires recombination, mutation and/or gene flow in order to create new phenotypes, while plasticity changes phenotypes within the lifetime of an individual. Evolution was (and still is) therefore generally not considered in ecological studies on how species interact with each other and with the (changing) environment. In the past decades, however, biologists have come to the fascinating realization that evolutionary changes can occur on very short time scales. Rapid evolution, defined as genetic changes that are fast enough to have an impact on ecological dynamics (Hairston et al. 2005), has been observed in a wide range of organisms (reviewed in Thompson 1998; Reznick and Ghalambor 2001), with short-term evolutionary changes in beak morphology in Darwin's finches being one of the most illustrative examples (Grant and Grant 2002).

What is the relative importance of plasticity and evolution in changing phenotypes in a population? How adaptive are these processes on both short and long time scales? To what extent can we predict phenotypic changes in natural populations? These are all fundamental and interesting questions. One major stumbling block to answering these is that there are methodological challenges: different approaches to decompose changes in phenotype dynamics have been developed largely independent from each other and differ in what exactly is estimated. The age-structured Price equation (Coulson and Tuljapurkar 2008), Integral Projection Models (Coulson et al. 2010), the Animal Model (Kruuk 2004) and the 'Geber' method (Hairston et al. 2005; Ellner et al. 2011) are four such frameworks that have been used to quantify the importance of different processes underlying changes in trait distributions. Which framework is most suited to answer the fundamental questions raised earlier? In **Chapter 2**, we compare similarities and differences between these four frameworks by applying them to simulated data, and provide some general guidelines on which framework is most suitable with regards to the research question.

1.2 Intra-genotypic variability to induce phenotypic variance

In addition to phenotypic plasticity and genetic variation, intra-genotypic variability is a third mechanism that can induce within-population phenotypic variance (Ayroles et al. 2015), one which has received much less attention. Intra-genotypic variability leads to phenotypic variation among genetically identical individuals even when reared in the same environment, due to stochastic micro-environmental or developmental fluctuations (Morgante et al. 2015). In contrast to phenotypic plasticity, intra-genotypic variability is thus not characterized by a correlation between phenotypes and the environment. Instead, some phenotypic variation is produced irrespective of the environment, and interestingly, the degree of this variation can be under genetic control (Rönnegård and Valdar 2012). In artificial breeding, for example, it is well known that lineages can differ in their 'uniformity', where some families tend to show more variation in certain phenotypic traits (such as birth weight) compared to other lineages (Formoso-Rafferty et al. 2016).

In natural populations, intra-genotypic variability has also been observed (e.g. Mulder et al. 2016), but its fitness consequences are to a large extent unknown (Viney and Reece 2013). One possible advantage of intra-genotypic variability is bet hedging, a topic which is frequently studied in relation to seed germination (Childs et al. 2010; Gremer and Venable 2014). A diversifying bet hedging strategy follows the idea of 'not putting all your eggs in one basket' (Philippi and Seger 1989; Olofsson et al. 2009), and it is proposed to be beneficial in unpredictable environments (Philippi and Seger 1989; Botero et al. 2015). Yet, despite some empirical evidence, it remains unknown how widespread bet hedging is (Simons 2011). In **Chapter 6**, we review empirical evidence for the presence of intra-genotypic variability, identify conditions under which intra-genotypic variability results in a bet hedging strategy, and outline theoretical expectations on when intra-genotypic variability might evolve.

1.3 Density as an ecological factor and its role in eco-evolutionary feedback loops

Evolution is based on the idea that genotypes differ in fitness, and that these differences are moderated by the environment. Indeed, rapid evolutionary responses have been documented in response to various changing abiotic factors, such as changes in pollution levels (Ponti et al. 2010), and changing biotic factors, such as food availability (Grant and Grant 1995) or host-parasite coevolution (Decaestecker et al. 2007). In this dissertation, the effects of several (a)biotic factors on the fitness of different genotypes will be presented. These include the effects of temperature (**Chapters 3** and **4**), parasite infection (**Chapter 4**), and the presence of predators and competitors (**Chapter 5**).

One environmental factor that has often been ignored in studies on how genotypes respond to changing environments, especially when studying individuals in isolation (e.g. Pietrzak 2011; Sommer et al. 2016), is the conspecific density. Yet, conspecific density-dependence is a core concept in the process of natural selection (see the quote at the beginning of this introduction). Density-dependence is caused by for instance competition for resources or space, and it is observed across taxonomic groups (Fowler 1981). Predicting how strategies will respond to environmental changes and which strategies will be most successful, calls for the explicit incorporation of conspecific densities, as the density is inextricably linked to the experienced environment. In all chapters describing empirical studies (**Chapters 3-5**), we therefore treat density-dependence as the core concept it is. We asses its role in individual performance, quantify the relation between density and fitness and quantify how this differs between genotypes. Moreover, we assess how the effects of conspecific density as an environmental factor, interact with effects of other (a)biotic factors.

In addition to ecological conditions (e.g. food availability, temperature or conspecific densities) influencing the rate and direction of evolution, the opposite can also happen: rapid evolution has the potential to influence the ecological conditions (Pimentel 1961). For example, rapid evolution of algae cell clumping has subsequent effects on the dynamics of rotifer populations (Yoshida et al. 2003; Becks et al. 2012), and in both guppies (Bassar et al. 2010b) and sticklebacks (Matthews et al. 2016), the genetic composition affects various abiotic (e.g. nutrient concentrations) and biotic (e.g. zooplankton and phytoplankton biomass) factors. This interplay between ecological and evolutionary processes is called eco-evolutionary dynamics (Schoener 2011; Reznick 2013), and eco-evolutionary dynamics may be important in shaping natural populations, communities and entire ecosystems (Fussmann et al. 2007; Strauss 2014). A two-way relationship between an ecological and evolutionary component, i.e. when the evolutionary response changes the ecological condition which then, in turn, again changes the evolutionary response, results in a so-called 'eco-evolutionary feedback loop' (Schoener 2011).

Conspecific density-dependence plays an essential role in eco-evolutionary feedback loops (Lion 2018): when genotypes respond differently to density (as shown in e.g. Agrawal et al. 2004), density (as an ecological factor) will influence the evolutionary response. This change in the genetic composition is coupled with a change in average fitness (i.e. average population growth rate[5]), which will then affect the density. This new density alters the experienced environment of the genotypes, thereby changing the selection, leading to an evolutionary response. One of the first identified eco-evolutionary feedback loops under field conditions was via conspecific density effects, in the green peach aphid (*Myzus persicae*) (Turcotte et al. 2011b; Turcotte et al. 2013). Populations that consisted of two aphid genotypes (and thus potentially evolving through means of clonal selection), were shown to have higher population growth rates compared to the non-evolving expectations based on the performance of single genotypes. Evolution was thus shown to affect population densities (Turcotte et al. 2011b). Moreover, as was shown in a follow-up experiment, population density had differential effects among genotypes (Turcotte et al. 2013), nicely demonstrating how density can affect evolution and complete the eco-evolutionary feedback loop.

Eco-evolutionary interactions can thus substantially affect the dynamics of a population. Adequately predicting population dynamics therefore requires taking into account the reciprocal interaction between ecology and evolution (Reznick 2013). Turcotte et al. (2011b), as well as many of the studies on eco-evolutionary feedback loops, focus on changes in the density (controlled by population growth rates) as the ecological response variable of interest (Yoshida et al. 2003; Hairston et al. 2005; Ellner et al. 2011; Becks et al. 2012). However, population growth rates, measuring fitness, are directly determined by patterns of individual reproduction and survival. Knowing these underlying individual rates may improve the predictability of eco-evolutionary responses (outlined in more detail below). The aim of **Chapter 5** is therefore to estimate the individual demographic rates underlying the observed eco-evolutionary feedback loop in the aphid populations.

---

[5]Population growth rate describes the change in the number of individuals in a population per time step, and it is a commonly extracted property of population models (λ is often used to denote the asymptotic population growth rate; Caswell 2001). This measure can for instance be used to assess the viability of natural populations (Doak et al. 1994). Likewise, it describes the rate at which a genotype or strategy increases or decreases in time, and throughout my dissertation, I use λ to quantify a genotype's or strategy's fitness (log λ equals Malthusian fitness; Fisher 1930; Caswell 2001)

1.4 **The importance of a full life-cycle approach**

It is the sum of the appearance (birth) and disappearance (death) of all individuals expressing a strategy that determines the strategy's fitness. Birth and death rates depend on multiple fitness components, for example maturation rate, reproduction frequency and lifespan, which are all shaped by individual phenotypes and by the environment. Therefore, to be able to understand, and ultimately, predict the success is of different strategies given the environmental conditions, it is critical to understand: i) which phenotypes are produced by the existing strategies, ii) how phenotypes affect single fitness components, iii) how these components altogether determine fitness, and iv) how the environment (including conspecific densities) interacts with the produced phenotypes and fitness components. A recurrent theme throughout this dissertation (specifically in **Chapters 3-6**), is the explicit link between individual phenotypes, individual performance and the long-term fitness of different strategies, in interaction with the environment.

Single life history traits, such as age at first reproduction, body size, clutch size or number of seeds per flower, are commonly used as a proxy for fitness (Kingsolver et al. 2001). Although those are indeed components of fitness, single life history traits are not necessarily a good proxy for fitness (Metcalf and Pavard 2007). First, this is because there can be correlations between fitness components, for instance through trade-offs (Stearns 1989). Often studied trade-offs, for example, include the negative correlation between offspring number and offspring size (Smith and Fretwell 1974; Einum and Fleming 2000) and the negative correlation between current reproduction, and future reproduction and survival (Snell and King 1977; Stearns 1989). Second, population growth rate is not equally sensitive to each fitness component (Caswell 1978; Kroon et al. 1986), and this sensitivity depends on the life history of the species (Saether and Bakke 2000). Measures of fitness should ideally be based on an integration over all (st)age-dependent growth, survival and reproduction, from here on called 'vital rates' (Metcalf and Pavard 2007). Population models, such as matrix population models (used in **Chapter 5**) or Integral Projection Models (used in **Chapters 3, 4** and **6**; Box 1.1), are useful and widely used tools to integrate over multiple vital rates.

> **Box 1.1 — Integral Projection Models.** Integral Projection Models (IPMs), not to be confused with Integrated Population Models (Schaub and Abadi 2011), are structured, data-driven population models, proposed as an extension of classical matrix population models (Easterling et al. 2000; Ellner and Rees 2006). IPMs describe changes in the distribution of a continuous state variable, due to state-dependent growth, reproduction and survival (i.e. all vital rates). Asymptotic population growth rate λ, measuring fitness, can be extracted from the constructed IPM, as well as other properties (Merow et al. 2014). By using a phenotypic trait as the state variable, often body size, there is an explicit link between phenotypes, vital rates and fitness.
>
> The underlying vital rate models are generally constructed from regressions between the phenotypic state variable and each of the vital rates, using individual-level measurements on these variables (Merow et al. 2014). Using standard methods, these regressions can be extended to include for instance multiple covariates, random effects and different link functions. This makes it straightforward to create IPMs which are not only a function of the state variable, but also of other factors (Bruijning et al. 2017a), such as temperature, population density, genotype and parasite load (**Chapters 3, 4, 6**). ∎

1.5 **Climate change: an ongoing selection pressure**

An environmental change that our planet is currently facing is climate change. It is predicted that surface temperatures will continue to increase during the 21$^{\text{st}}$ century, with predicted average

increases between 1.5 °C and 4 °C, and that extreme climatic events (e.g. droughts, floods) will become more frequent (IPCC 2014). These changes can have large effects on natural populations. Climate change is expected to impose a strong selection pressure on particularly cold-blooded animals (ectotherms), as their physiological rates depend directly on the external temperature (Huey and Berrigan 2001). Changes in the abundance and/or functioning of local populations can have ramifications for whole ecosystems, especially when key stone species are involved (Hoffmann and Sgrò 2011). Therefore, predicting whether and how populations can cope with coming environmental changes is thus not only a fundamentally interesting and challenging concern, but also a highly relevant one (Hoffmann and Sgrò 2011).

Using the small aquatic crustacean *Daphnia magna* (water fleas) as a study species (Box 1.2) in two laboratory experiments, we focus on how different strategies affect fitness, and the interaction with temperature. We study plastic and genetic effects on body size and vital rates, and obtain measures of fitness by integrating over all vital rates. We concentrate both on within-population variance in thermal responses (**Chapter 3**) and on between-population variance, by comparing genotypes originating from populations in Belgium with genotypes from Norway (**Chapter 4**). These studies will contribute to our understanding in how much genetic variation for thermal tolerance there is in natural populations, to what degree individuals are (genetically) adapted to the temperature they are usually exposed to, and to what degree individuals can respond plastically to a range of temperatures.

Natural populations exposed to climate change do not only have to cope with changes in temperature, but also with associated changes in for instance food and oxygen availability (Ficke et al. 2007). Furthermore, local populations will be increasingly likely to experience competition from conspecific immigrants that are shifting their habitat range (Thomas and Lennon 1999; Hickling et al. 2006; Ficke et al. 2007). These conspecific immigrants may be better adapted to increasing temperatures than local populations (Van Doorslaer et al. 2009b), giving immigrants a competitive advantage. Moreover, immigrants may carry along pathogens, to which they may be better adapted than local individuals (Carius et al. 2001; Decaestecker et al. 2004; Ficke et al. 2007). In order to predict the future of local populations, it will be critical to understand these complex interactions between temperature and associated (a)biotic changes. In **Chapter 4**, we investigate the interactive effects of temperature and an ectoparasite infection, on the success of water fleas originating from Belgian and Norwegian populations. By exposing genotypes to a competition treatment, we specifically look at the effects of competition between genotypes originating from different populations.

> **Box 1.2 — Study species *Daphnia magna*.** In **Chapters 3, 4** we have used water fleas as study species. *Daphnia* are primary consumers and considered to be keystone species in many freshwater ecosystems. *Daphnia* is a widely used model system in ecology, genetics, evolutionary biology and toxicology (Stollewerk 2010; Harris et al. 2012; Miner et al. 2012), and *Daphnia pulex* was the first crustacean to have its genome sequenced (Colbourne et al. 2011).
>
> *Daphnia* individuals reproduce by cyclical parthenogenesis (amphitoky): individuals reproduce asexually (apomictic thelytoky) when environmental conditions are favourable, but switch to sexual reproduction (partial mixis: arrhenotoky) when conditions worsen (Kleiven et al. 1992), resulting in the production of resting eggs which can remain viable in the sediment for decades (Hairston et al. 1999). By collecting resting eggs in the field, genetically different lineages can be established. As asexual reproduction can be assured in the lab by keeping conditions favourable and removing any resting eggs, the same combinations of genotypes can be used across treatments.

1.6 Methodological challenges associated with collecting individual and population-level data

Using short-lived species (such as water fleas and aphids) to study how different strategies respond to varying environmental conditions, as presented in my dissertation, has some clear advantages. It is easy to perform multi-generation experiments, culture hundreds of individuals simultaneously and apply different treatments. These aspects are considerably more challenging in larger species, such as mammals. One big advantage, however, of using species with larger body sizes, is that they are generally individually recognizable. This can be achieved by marking individuals, which is commonly done in natural mammal (e.g. Ozgul et al. 2010), bird (e.g. Grant and Grant 2002) and plant (e.g. Bruijning et al. 2017a) populations. This makes it straightforward to perform repeated measurements on individuals to attain information on phenotypic traits and vital rates, and to construct pedigrees (see e.g. Réale et al. 2003).

Being unable to follow individuals within populations, is a drawback of species like water fleas and aphids as study species, and it is likely to, at least partly, explain the low number of studies that link individual vital rates to population-level performance in these kind of species. One solution is to remove individuals from the population and perform measurements on isolated individuals (Ozgul et al. 2012; Cameron et al. 2013; Sommer et al. 2016) but as the isolated individuals do no longer interact with other individuals nor compete for food, density-dependent processes may be overlooked. For the water fleas, we have addressed this by isolating randomly picked individuals within transparent permeable tubes, placed in between the other individuals, and following these isolated individuals for a short time interval (**Chapters 3, 4**). Alternatively, a promising method, and the approach of **Chapter 5**, is to use inverse modelling techniques to estimate individual vital rates. This can be an option when there is population-level data available, but no individual-level data. Temporal changes in stage-frequency data are the direct result of individual performance, and these data can in principle be used to infer these rates (Wood 1994). One major challenge, however, is that multiple combinations of individual rates can result in the same observed dynamics, which becomes clear by the example given by Wood (1994): 'There were two bears yesterday and there are three bears today. Does this mean that one bear has been born, or that 101 have been born

and 100 have died?' In **Chapter 5**, we discuss possibilities of this inverse modelling, as well as limitations and vital assumptions.

Although it is practically possible to bring hundreds, even thousands, of *Daphnia* individuals into the lab, there are other aspects that limit the scope of experiments. Manually counting and measuring the number of individuals is very time-consuming. A solution is presented in **Chapter 7** where we provide a general method to obtain estimates of population densities, individual body sizes and behavioural metrics from video material of moving organisms. The methods are supplied as a new R-package *trackdem* (Bruijning et al. 2017b). We have used this method for counting water fleas for the experiments described in **Chapters 3, 4**.

1.7   Outline and objectives of this dissertation

As outlined in the previous sections, predicting the fate of natural populations is a major challenge we are currently facing, and this is urgent both from a fundamental and from a conservation point of view. The aim of my dissertation is to contribute to the understanding of the importance of phenotypic plasticity, genetic variation and intra-genotypic variability in creating phenotypic variation, given that it is the constant selection on the fittest phenotypes that results in selection on the most successful strategies. In particular, I focus on the role of eco-evolutionary dynamics in the response of ectothermic species to environmental change. To do so I present a combination of theoretical and empirical work. In **Chapter 2**, four theoretical frameworks that can be used to quantify the importance of different processes in shaping phenotypic trait dynamics are discussed. In **Chapters 3, 4**, I present empirical results on the importance of plasticity and genetic variation in response to temperature in the water flea *Daphnia magna*, for genotypes originating from the same pond (**Chapter 3**), and from different latitudes, in interaction with a parasite infection (**Chapter 4**). In **Chapter 5**, the aim is to estimate individual vital rates underlying the eco-evolutionary feedback loop observed in the aphid *Myzus persicae*. In **Chapter 6**, theoretical fitness consequences of intra-genotypic variability are explored and empirical evidence for its presence is reviewed. In **Chapter 7**, I present a general method to obtain population counts and measurements using automated video analysis. Finally, in **Chapter 8**, I synthesize the results of all presented chapters. Here, I present the most important implications of these results with regards to the fate of natural populations facing global changes and how these might be relevant in conservation biology. In addition, I provide opportunities to extend available frameworks and techniques to study the links between genotypes, phenotypes and long-term fitness, in interaction with the environment.

*Koen J. van Benthem[1†], **Marjolein Bruijning**[2†], Timothée Bonnet[1†], Eelke Jongejans[2], Erik Postma[1] and Arpat Ozgul[1]*

1. Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.
2. Department of Animal Ecology and Physiology, Radboud University, 6500 GL Nijmegen, The Netherlands.

† Contributed equally as first authors.

$$\Delta \bar{z} = h^2 \cdot S \quad n(z', t+1) = \int S(z) \cdot G(z', z) + R(z) \cdot D(z', z) n(t) dz$$

$$V_P + V_A + V_E \qquad z = \mu + g + e$$

$$S = \sigma(w, z) \qquad R_z = \sigma(w_i, z_i) + E(w_i \delta_i)$$

$$S = \mu_S - \mu$$

$$h^2 = V_A / V_P \qquad \frac{d\bar{z}}{dt} = \frac{\delta \bar{z}}{\delta \bar{k}} \frac{d\bar{k}}{dt} + \frac{\delta \bar{z}}{\delta \bar{a}} \frac{d\bar{a}}{dt} + \frac{\delta \bar{z}}{\delta \bar{p}} \frac{d\bar{p}}{dt}$$

Disentangling evolutionary, plastic and demographic processes underlying trait dynamics: a review of four frameworks

2.1 Abstract

1. Biologists are increasingly interested in decomposing trait dynamics into underlying processes, such as evolution, plasticity and demography. Four important frameworks that allow for such a decomposition are the quantitative genetic animal model (AM), the 'Geber' method (GM), the age-structured Price equation (APE), and the integral projection model (IPM). However, as these frameworks have largely been developed independently, they differ in the assumptions they make, the data they require, as well as their outcomes and interpretation.

2. Here we evaluate how each framework decomposes trait dynamics into underlying processes. To do so, we apply them to simulated data for a hypothetical animal population. Individual body size was affected by, among others, genes, maternal effects and food intake. We simulated scenarios with and without selection on body size, and with high and low heritability.

3. The APE and IPM provided similar results, as did the AM and GM, with important differences between the former and the latter. All frameworks detected positive contributions of selection in the high but not in the low selection scenarios. However, only the AM and GM distinguished between the high and low heritability scenarios. Furthermore, the AM and GM revealed a high contribution of plasticity. The APE and IPM attributed most of the change in body size to ontogenetic growth and inheritance, where the latter captures the combined effects of plasticity, maternal effects and heritability. We show how these apparent discrepancies are mostly due to differences in aims and definitions. For example, the APE and IPM capture selection, whereas the AM and GM focus on the response to selection. Furthermore, the frameworks differ in the processes that are ascribed to plasticity and in how they take into account demography.

4. We conclude that no single framework provides the 'true' contributions of evolution, plasticity and demography. Instead, different research questions require different frameworks. A thorough understanding of the different definitions of their components is necessary for selecting the most appropriate framework for the question at hand, and for making biologically meaningful inferences. This work thus supports both future analysis as well as the careful interpretation of existing work.

2.2 Introduction

Understanding trait and population dynamics and how the two are intertwined is crucial for predicting population resilience and viability (e.g. Merilä and Hendry 2014). Hence, which processes shape population-level trait dynamics (i.e. changes in trait distributions over time) is a fundamental question in ecology and evolution, and one which is gaining in urgency given mounting concern regarding the consequences of anthropogenic environmental change for natural populations (e.g. Parmesan 2006).

Phenotypic trait distributions may be altered across generations by genetic (i.e. evolutionary) processes, as well as by non-genetic processes, such as phenotypic plasticity. Since the realisation that evolutionary and ecological processes may act on the same time scale, distinguishing between the role of evolution and plasticity has been the subject of a substantial body of research (Hairston et al. 2005; Gienapp et al. 2008; Post and Palkovacs 2009). To complicate matters further, changes in the demographic structure of a population may additionally shape trait distributions (Coulson and Tuljapurkar 2008). Hence, understanding and predicting trait dynamics ideally requires simultaneously taking into account all three processes (Pelletier et al. 2007; Schoener 2011).

To date, four major frameworks aiming at distinguishing between the role of evolution, phenotypic plasticity and demography have been developed: 1) The quantitative genetic framework, particularly the animal model (AM; e.g. Henderson 1950), 2) the 'Geber' method (GM; Hairston et al. 2005), 3) the age-structured Price equation (APE; Coulson and Tuljapurkar 2008), and 4) the application of the APE in conjunction with an integral projection model (IPM; Easterling et al. 2000; Ellner and Rees 2006; Coulson et al. 2010). Several studies have tried to explicitly estimate the relative importance of evolution, plasticity and/or demography using one of these approaches (e.g. Morrissey et al. 2012a; Réale et al. 2003; Ezard et al. 2009; Ozgul et al. 2009; Rebke et al. 2010; Becks et al. 2012). Nevertheless, fully disentangling and quantifying evolutionary, ecological and demographic processes and ultimately predicting the consequential trait dynamics has proven to be problematic (Gienapp et al. 2008; Schoener 2011; Merilä and Hendry 2014). At least some of these difficulties can be attributed to the large amounts of (individual-based) long-term data required, which are often unavailable for natural populations (Clutton-Brock and Sheldon 2010). However, even if sufficient data are available, synthesis of the results from the four frameworks is hampered by the fact that they have been developed largely independently of each other. As a consequence, they differ in their focus and aims, and as we show here, they define biological processes in non-equivalent ways.

Here we provide an overview of the differences, similarities and complementarity of each of these four decomposition frameworks by applying them to the same simulated datasets and comparing their outcomes. Thereby, we evaluate how they quantify the role of different ecological and evolutionary mechanisms in shaping trait dynamics under a range of biological scenarios. Together with a critical review of the theory underlying each of the frameworks, we provide comprehensive insight into their underlying assumptions, as well as the conceptual differences and similarities. This provides a much needed overview of the suitability of each framework with respect to research questions and data availability.

2.3 Applying the four frameworks

Data simulation

Although it comes with the loss of some biological realism, using simulated rather than empirical data enables us to evaluate the frameworks under different scenarios and allows for replication. Furthermore, simulated data do not suffer from the complications introduced by missing data. Finally, it provides a reference that aids the comparison between the results of each framework. Importantly, it is not possible to calculate "true" contributions of for example evolution without first adopting one of the frameworks and their corresponding definitions, therefore, our simulations allow only for a qualitative assessment.

Data were simulated using a two-sex individual-based model of a closed population of a hypothetical animal species, implemented in R (R Core Team 2015). Here, we provide a brief overview, while a more complete description can be found in supporting information A.1. We also provide the R code on
`https://github.com/koenvanbenthem/Disentangling_Dynamics_IBM`. We simulated a single trait, body size z. Size at birth is determined by an individual's genotype (10 loci, with 10 alleles each and mendelian inheritance, more details in A.1.1), the body size of its mother (i.e. a maternal effect as in Falconer 1965), and a stochastic component (drawn from a Gaussian distribution; A.1.2). Ontogenetic growth results in an increase of body size with age. Growth rate, the proportional increase in body size, decreases with age, and is further influenced by per-capita food availability (A.1.3). Males were randomly assigned to females, who have a 50% chance of becoming reproductive after one year and whose reproductive probability increases with age. The litter size that a female produces depends on per-capita food availability, a stochastic component, and body size (A.1.4). Survival probability first increases with age, but starts decreasing after year five, reflecting senescence, and is further influenced by per-capita food availability and body size. Maximum age is 30 years. Furthermore, a trade-off exists between female reproduction and survival, i.e. reproducing at time t decreases survival probability to time $t + 1$ (A.1.5).

We simulated fifty time steps (years). After ten years, total food availability started to decline. Every year the available food is divided over all individuals, with some individuals randomly obtaining more than others. Individual food intake affects survival, growth and (female) reproductive success (A.1.6). The first ten years were discarded from further analyses to allow the age structure to stabilize (Fig. 2.1(f)) The remaining data spanned 40 years (i.e. approximately 13 generations), which is comparable to the length of some of the field studies these frameworks have been applied to (Clutton-Brock and Sheldon 2010).

To evaluate the behaviour of the frameworks under different circumstances, we simulated four different scenarios. First, survival and fertility selection on body size was either present (s_+) or absent (s_0). Under the s_+ scenarios, there was a positive effect of body mass on survival and on litter size for mothers. Second, the relative importance of genetic variation in shaping body size, commonly measured as heritability, was either high (h_+) or low (h_-). This was done by using either of two pre-defined genotype-phenotype maps: one with big and one with small variation in the effects of alleles. Furthermore, to keep the phenotypic variance comparable, we decreased the plastic component in birth size in the h_+ scenarios. The parameter values for each of the four scenarios (s_0h_-, s_0h_+, s_+h_- and s_+h_+) can be found in A.1.7. To evaluate the effect of stochasticity, each scenario was replicated 100 times.

Fig. 2.1 provides an illustration of some of the key characteristics of the datasets simulated under each scenario. Despite a substantial amount of stochastic variation across replicates within each scenario, clear differences in trait and population dynamics are apparent. As expected, the s_+ scenarios show a positive relation between body size and annual fitness, calculated as the sum of survival and litter size to $t + 1$, whereas the s_0 scenarios do not (Fig. 2.1(e)). Furthermore, the

proportion of the phenotypic variance attributable to variance in the simulated genotypic values (i.e. broad-sense heritability) was ca. 0.50 in the h_+ and 0.08 in the h_- scenario.

Although in all scenarios population size first increased (until year 20) and then decreased (Fig. 2.1(a)), the population size averaged across replicates reached up to 322 and 334 individuals in scenarios s_+h_- and s_+h_+, whereas in s_0h_- and s_0h_+ the maximum average population size was 245 and 252 individuals, respectively. Mean body size first increased rapidly, but decreased in all scenarios between the eleventh and fiftieth year (Fig. 2.1(b)): in s_0h_- with (mean $\pm$ SE) -0.47 ± 0.058 $[-1.45; 0.63$ 95% interval], in s_0h_+ with -0.46 ± 0.061 $[-1.59; 0.0.68]$, in s_+h_- with -0.75 ± 0.051 $[-1.87; 0.08]$, and in s_+h_+ with -0.16 ± 0.057 $[-1.12; 0.83]$. Note that the 95% intervals, here and in the rest of the manuscript, are ranges of point estimates across replicates. They reflect the stochasticity of the simulations rather than the precision of the estimates. The standard errors for each average were calculated by dividing the standard deviation of the values of the replicates by 10 (the square root of the number of replicates). A full power analysis of the methods is beyond the scope of this manuscript.

Contrary to average body size, genotypic values for birth size continued to increase only in scenario s_+h_+. Here, the change in average genotypic value (across the entire population) between year 11 and year 50 was 0.62 ± 0.022 $[0.23; 1.04]$ (Fig. 2.1(d)). In s_+h_- a smaller increase was observed 0.08 ± 0.0083 $[-0.074; 0.24]$, whereas s_0h_- and s_0h_+ show on average no change in genotypic values. Correspondingly, average birth size increased only in the s_+h_+ scenario, with 0.58 ± 0.027 $[0.092; 1.11]$, between year 11 and year 50 (Fig. 2.1(c)).
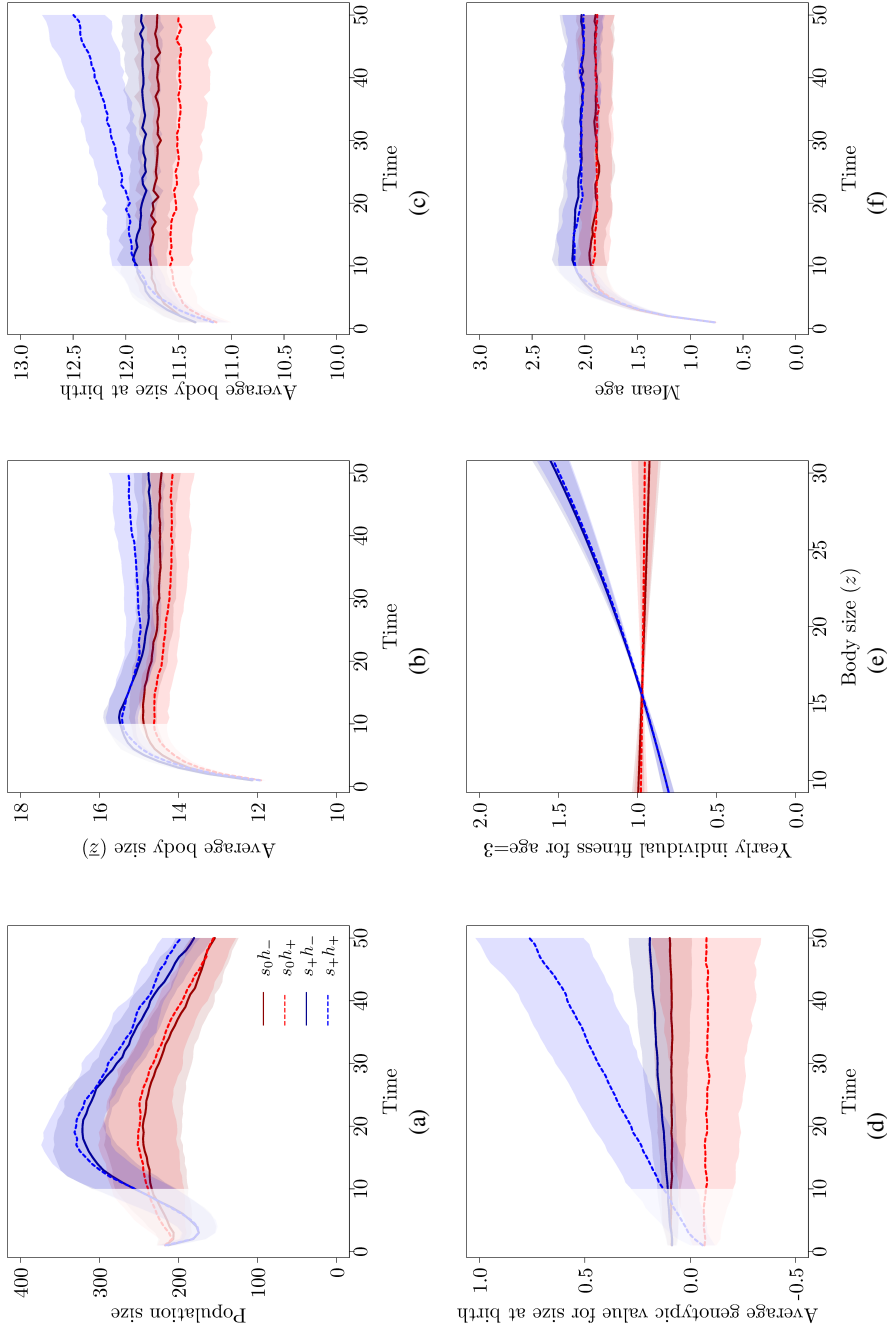
Figure 2.1: Summary of the observed population and trait dynamics of simulated datasets. (a) Trends in population size, (b) changes in mean body size, (c) mean birth size, (d) and genotypic values for body size, (e) relations between body size, (d) and genotypic values for body size, (e) relations between body size and yearly individual fitness (sum of survival and litter size at $t+1$), and (f) changes in mean age. Lines indicate the averages across 100 replicates. Polygons show one standard deviation above and below the average. Red lines indicate s_0 scenarios (no viability and fertility selection), blue lines indicate s_+ scenarios (strong viability and fertility selection). Solid lines indicate h_- scenarios (low heritability), dotted lines indicate h_+ scenarios (high heritability). In a-d and f, the white polygon indicates the first 10 years, which are excluded from further analysis. In e, lines are averaged predictions based on generalized additive models over all replicates.

Decomposing simulated trait dynamics

Rather than providing an exhaustive overview of all methods allowing for the decomposition of trait dynamics, we have chosen to focus on four, commonly-used, frameworks. The four frameworks have different data requirements and do not yield identical results. This is illustrated in the following section, in which we analyse the simulated data using each framework.

Animal Model

The animal model (AM) is a quantitative genetic method that was developed for commercial breeding (Henderson 1950; Henderson 1976), where it has been used successfully for several decades (e.g. Lynch and Walsh 1998). Only recently has it been applied to wild animal (e.g. Réale et al. 2003; Postma 2014) and plant (Stinchcombe et al. 2014) populations. For extensive explanations of the AM as applied to natural populations, see Kruuk (2004) and Wilson et al. (2009).

The AM is a linear mixed effects model that is fitted to individual-level data and assumes a quantitative genetic model, where a phenotypic trait (z) is influenced by a large number of genes with small effects (Roff 2007). The variance in z is partitioned into genetic and non-genetic sources of variation. Under the assumption that this partitioning is additive (i.e. in the absence of genotype-environment correlations and interactions), z can be written as the sum of a population mean (μ), an additive genetic effect (the breeding value, a) and a residual (environmental) value capturing plasticity (e), thus $z = \mu + a + e$. Information on the relatedness between individuals (estimated from a pedigree or genetic markers) is used as a constraint in the fit, allowing for the estimation of a. If the data allow for it, other components contributing to variation in z, such as maternal, common, and permanent environmental effects can be accounted for explicitly. This variance decomposition can be used to estimate genetic change over time—resulting from, for example, selection or genetic drift.

There are several ways to estimate evolution within the AM framework (see *Discussion*), but here we illustrate only one. We fitted a univariate AM and quantified the change in the best linear unbiased predictors (BLUPs) for the breeding values over time (Postma 2006; Hadfield 2010). We used body size as the sole response variable, and intercepts for breeding values, maternal effects, permanent environment, and year were included as random effects. Maternal and permanent environment effects were modelled by fitting maternal and individual identity, respectively. An alternative specification of the maternal effects, more in line with the simulation process, is briefly discussed further below. Age was included as a continuous fixed effect (both as linear and quadratic terms). All fits were performed using the R-package MCMCglmm (Hadfield 2010) using inverse-Wishart priors with variance and degree of belief both set to 1. The posterior distributions were estimated based on 1,000 MCMC samples, from 50,000 iterations with a thinning interval of 40 and a burn-in of 10,000, thus ensuring that the correlation between successive samples of all parameters is below 10%.

We estimated the temporal trend in the BLUPs for all random effects. We accounted for their uncertainty following Hadfield et al. (2010) by performing a regression of the BLUPs on time for each MCMC sample of the model. This provided a posterior distribution of linear slope coefficients, estimating the change in additive genetic, maternal, and permanent environment effects per time step. More details on the fitted models are given in A.2.1.

As depicted in Fig. 2.2(a), in all scenarios the contributions of evolution and individual plasticity were largest, while the contributions of permanent environment and maternal effects were very small. On average, the per year change in breeding values was positive in both scenario s_+h_- (0.0013 ± 0.0003 [$-0.0038; 0.0095$]) and scenario s_+h_+ (0.014 ± 0.0007 [$0.00021; 0.029$]). Note that the large error bars in Fig. 2.2(a)) mostly reflect a substantial amount of variation in the

rate of evolutionary change among replicates due to genetic drift, rather than the uncertainty in the point estimates. Negative contributions of individual plasticity were found, particularly in the scenarios with selection -0.02 ± 0.0013 $[-0.049; 0.0018]$ and -0.019 ± 0.0013 $[-0.045; 0.0029]$ for h_- and h_+, respectively.

Despite substantial drift, we would expect the contribution of evolution averaged over replicates to be 0 in the s_0 scenarios. Instead, our model inferred a genetic decline for h_- and h_+ of -0.0057 ± 0.0005 $[-0.016; 0.0040]$ and -0.0073 ± 0.0009 $[-0.024; 0.0087]$, respectively. The AM therefore estimates evolution with a negative bias. The reason is a mismatch between the model structure and the simulation process. As mean size decreases with time, the maternal contributions to birth size decreases. Because we modelled maternal effects as maternal identity rather than maternal current size, this change is mistaken for evolution. We performed an additional analysis using maternal size instead of maternal identity, which strongly reduced this artefact (details and results in A.2.2).

Geber method

The 'Geber' method (GM) (Hairston et al. 2005) is a very general method that quantifies how temporal changes in various factors influence the response variable of interest. Because of this generality, the biological assumptions depend on the specific implementation. The GM may for example estimate how temporal changes in mean breeding value $\bar{a}$ and in an environmental factor k such as food availability propagate to a population-level response variable X, such as mean trait value. Examples of its application can be found in Ellner et al. (2011) and Becks et al. (2012).

Our implementation of the GM follows the analysis of fledgling mass in Ellner et al. (2011). We took the average body size ($\bar{z}$) as the population-level response variable, and decomposed the change in $\bar{z}$ into a contribution of the environment ($\bar{k}$) and a contribution of a phenotypic change in size at birth. The latter was decomposed further into an evolutionary ($\bar{a}$) and a plastic component ($\bar{p}$):

$$\frac{d\bar{z}}{dt} = \frac{\partial \bar{z}}{\partial \bar{k}}\frac{d\bar{k}}{dt} + \frac{\partial \bar{z}}{\partial \bar{a}}\frac{d\bar{a}}{dt} + \frac{\partial \bar{z}}{\partial \bar{p}}\frac{d\bar{p}}{dt} \qquad (2.1)$$

For each year between years 11 and 50, we calculated the mean body size ($\bar{z}$), mean size at birth of newborns, the average food availability that alive individuals had access to during their life up to that moment ($\bar{k}$), and the mean breeding value as estimated by the AM ($\bar{a}$) (see above). As breeding values can not be observed directly, the application of the GM to empirical data relies on other methods such as the AM for their estimation. Finally, we calculated a plasticity term ($\bar{p}$), equal to the difference between the average size at birth and the average breeding value for size at birth. Thereby this term only captured plasticity in mass at birth. We fitted a linear model to estimate the effects of $\bar{a}$, $\bar{p}$ and $\bar{k}$ on $\bar{z}$. Using this model, together with separate linear models that describe how each of the three underlying factors changes over time, we evaluated their respective influence on $\bar{z}$. This procedure is described in more detail in A.3.1.

The results of the GM are shown in Fig. 2.2(b). The results for the evolutionary component are, as expected, nearly identical to the results of the AM. This evolutionary component is counter-acted by a decrease in food availability, as is shown by the negative 'environmental' contributions. The latter is largest for the s_+ scenarios, under which population size is higher (Fig. 2.1(a)) and per capita food availability therefore lower.

The average contributions of plasticity are more equivocal. Whereas we expected the slight reduction in maternal body size, and hence in the maternal effect, to result in a minor negative contribution of plasticity, we instead see mainly positive contributions. This is the result of the

downwardly biased trend in the breeding values (as discussed above). When the analysis was repeated with the 'true' genotypic values from the simulations instead of the estimated breeding values, all scenarios showed negative contributions of plasticity (A.3.2).

Age-structured Price Equation

The age-structured Price equation (APE) (Coulson and Tuljapurkar 2008) is an extension of the Price equation (Price 1970). The APE does not explicitly consider genetic variation. It decomposes the change in mean trait value into seven additive components. All these contributions are either averages of, or covariances between, observable individual properties (e.g. individual survival and body size).

The two selection terms describe how selective disappearance (viability selection, VS) and selective reproduction (fertility selection, FS) alter the mean trait value. Here, VS is the covariance between z and survival, which scales with the difference in the average trait value of the whole population and the part of the population that survives to the next time step (e.g. Rebke 2012). This is referred to as the selection differential in the evolutionary literature (Robertson 1966; Lande and Arnold 1983). The contribution to the change in mean trait value due to ontogenetic development of surviving individuals is captured by the growth term. The two inheritance-related contributions were combined into one (A.4.3). This combined term measures the contribution to changes in average body size due to the difference between the mother's body size (at time of giving birth) and her offspring's body size at birth (i.e. between generations). Because offspring are generally smaller than mothers, the inheritance contribution will typically be negative. This stresses that the inheritance term should not be confused with heritability, which can not be negative. Finally, the two demography contributions, here also combined into one, describe change resulting from the age structure (A.4.2). The demography term arises because the other contributions are calculated per age class. This takes into account that their values depend not only on the trait value of an individual, but also on its age. The total contribution is obtained by a weighted sum of the age specific contributions.

The APE thus allows for an exact decomposition of $\Delta\bar{z}$ in discrete time into components of viability selection, fertility selection, ontogenetic growth, inheritance, and demography in populations with overlapping generations. It has been applied to a range of mammals species (Coulson and Tuljapurkar 2008; Ozgul et al. 2009; Ozgul et al. 2010; Canale et al. 2016). See A.4.1 for the full equation and an explanation of the terms. Note that a stage-structured version of the Price equation has also been developed (Barfield et al. 2011).

As is commonly done in demographic analyses, we applied the APE to the female part of the population only. Under the s_0 scenarios, we find that the average VS and FS are both indistinguishable from zero (Fig. 2.2(c)). For the s_+ scenarios, the contribution of selection is positive, and there is no difference between the s_+h_+ and s_+h_- scenarios (VS: 0.081 ± 0.0012 $[0.060; 0.10]$ and 0.090 ± 0.0013 $[0.063; 0.11]$ respectively, FS: 0.054 ± 0.0015 $[0.027; 0.079]$ and 0.055 ± 0.0015 $[0.029; 0.082]$ respectively). Finally, the demographic contribution differs between the s_0 and s_+ scenarios, but does not differ between h_+ and h_-. This combined demography term scales with the between-age class covariance between fitness and body size (A.4.2). In agreement with our simulation processes, this covariance is strong and positive, as older age classes have larger average body size, and larger individuals have higher fitness in the s_+ scenarios. The negative contribution in the s_0 scenarios is the result of a negative effect of age on survival, which in the absence of positive selection will dominate the between-age class covariance. The biggest contribution to changes in average body size comes from ontogenetic growth. This component is slightly lower in the s_+ scenarios, due to smaller per capita food availability.

The inheritance term is more negative in the s_+ than in the s_0 scenarios. This is because in the

s_+ scenarios larger mothers produce more offspring, which on average results in a larger difference between mother and offspring size: although the maternal trait value when giving birth is higher, their offspring's trait value at birth does not increase by the same amount. This leads to the average contribution of inheritance becoming more negative. Furthermore, we see that contributions from inheritance are slightly smaller (less negative) under the h_+ scenarios than under the h_- scenarios. This is because with increasing heritability, the mother-offspring difference decreases, leading to a less negative inheritance term.

Integral Projection Model

The integral projection model (IPM) is a general model for projecting continuous distributions in discrete time. When describing a population, it often considers four life history processes: survival, reproduction, growth and inheritance (Ellner and Rees 2006). The dependencies of these processes on a continuous phenotypic trait z are estimated using regression models. No assumptions concerning the underlying genetics are made. Based on these regressions, the trait distribution at time $t+1$ can be predicted from the trait distribution at time t (as well as demographic properties, such as population growth rates, e.g. Adler et al. 2010; Merow et al. 2014). Over the past years, IPMs have been used to address a range of eco-evolutionary questions (e.g. Metcalf et al. 2008; Smallegange and Coulson 2013; Traill et al. 2014). While the specific decomposition we use involves applying the APE to a fitted IPM, as proposed by Coulson et al. (2010), approaches using a sensitivity analysis also exist (e.g. Coulson et al. 2011; Traill et al. 2014).

An IPM was parametrized for each simulated dataset, and as we did for the APE, we only considered females. Models describing individual growth, survival and reproduction (both the probability of reproducing and the number of offspring) were fitted using generalized linear mixed models with appropriate link functions (logit for survival and reproduction probability, log for number of offspring). The contribution of inheritance was estimated as a linear regression of offspring size at birth on the size of the mother at the time of giving birth, as done in Traill et al. (2014). This differs fundamentally from heritability (h^2), where offspring size is related to the mother's size, both at the same fixed developmental stage (e.g. birth) (Chevin 2015). For all life history processes, we tested five different models: a full model containing age, size and their interaction, as well as all models nested within this full model. Furthermore, each model included a random effect for year. The model with the lowest AIC was selected and used for the IPM.

Using the selected models, a 3100×3100 matrix was parametrized (i.e. 31 age classes, 100 size classes per age class, ranging between 1 and 50) for each replicate. See A.5 for more details on model fitting and the construction of the IPMs. For each IPM, we used the observed population vector at each time step (excluding the first ten years) to project the population vector to the next time step ($t+1$). Changes in population structure, and thereby changes in $\bar{z}$, are decomposed into contributions from different life history processes.

We found very similar patterns as in the APE (Fig. 2.2(d)). Both viability and fertility selection were detected in the s_+ scenarios (VS was 0.045 ± 0.00096 $[0.026; 0.063]$ and 0.041 ± 0.00098 $[0.024; 0.060]$; FS was 0.012 ± 0.00074 $[0.00; 0.026]$ and 0.012 ± 0.00076 $[-0.0044; 0.028]$, for h_- and h_+). In contrast, in the $s_0 h_-$ and $s_0 h_+$ scenarios, average viability selection was -0.024 ± 0.0011 $[-0.045; -0.0024]$ and -0.019 ± 0.0010 $[-0.039; -0.00024]$, respectively, and fertility selection was -0.00069 ± 0.00069 $[-0.014; 0.012]$ and 0.00068 ± 0.00059 $[-0.011; 0.014]$. As in the APE, the contribution of inheritance to $\Delta \bar{z}$ was large and negative in all scenarios, and was more negative in the s_+ scenarios. Furthermore, there was a consistently positive contribution of ontogenetic growth, with weaker effects in the s_+ scenarios, again due to lower per capita food availability. As in the APE, we considered both demographic terms together. This term showed positive contributions in all

scenarios.

To allow for a better comparison with the other three frameworks, here we focus on the average value of $\Delta\bar{z}$, and how much various processes contribute to this. When quantifying how much of the *year-to-year variation* in $\Delta\bar{z}$ is explained by each process (as for example in Ozgul et al. 2009), the IPM and APE provide more divergent results (A.6).

2.4   Discussion

We have decomposed changes in mean body size into underlying processes by applying four major frameworks to simulated data. Thereby we have shown that these frameworks differ substantially in their data requirements, which processes they consider, how these are defined, and how changes in the mean trait value are assigned to them. In the following sections we will discuss and compare the theory underlying the four frameworks, illustrated by our simulations. We will discuss the inherent differences among frameworks regarding evolution, plasticity, demography, and measures of uncertainty. These are summarised in Table 2.1. We finish by discussing each framework with respect to data availability and the research question at hand.

We have simulated scenarios with and without selection on body size, and with low and high heritability. As multiple processes influence and interact with body size, these scenarios resulted in divergent and relatively complex population and trait dynamics (Fig. 2.1). For example, in addition to genetic effects, size at birth was influenced by maternal effects and stochasticity. Moreover, ontogenetic growth was subject to both stochastic variation and a decrease in per-capita food availability. We also included a trade-off between viability and fertility. It is exactly this complexity that highlights the need for a robust framework that allows disentangling the underlying processes and quantifying their importance.

Table 2.1: A selection of research questions and to what extent frameworks may be used to answer them, ranging from impossible without major modifications ($--$) to being answered by the standard formulation of the framework already ($++$). AM = animal model, GM = Geber method, APE = age-structured Price equation and IPM = integral projection model. Note that scores are based on the specific application of the frameworks as we reviewed here; this involves the univariate AM, and the application of the APE to the IPM, in case of the IPM. Alternative approaches of the frameworks are mentioned in the discussion.

Question	AM	GM	APE	IPM
Does the change in trait value have a genetic basis?	++	+	--	--
Is selection acting on the trait?	+	+	++	++
Is the trait heritable?	++	±	-	-
Is the age structure responsible for the change in mean trait value?	+	±	++	++
How does individual heterogeneity affect trait value z?	+	±	--	-
How do trait dynamics affect population dynamics?	-	+	-	++
Is an environmental change responsible for the change in mean trait value?	+	++	--	-

Selection and evolution

All four frameworks infer positive selection on body size in the s_+ scenarios, but not in the s_0 scenarios (Fig. 2.2). The APE and IPM detect positive viability and fertility selection in both the s_+h_+ and the s_+h_- scenarios. The AM and GM detect a strong increase in mean breeding values in the s_+h_+ scenario and a small yet positive contribution in the s_+h_- scenario. Importantly, the
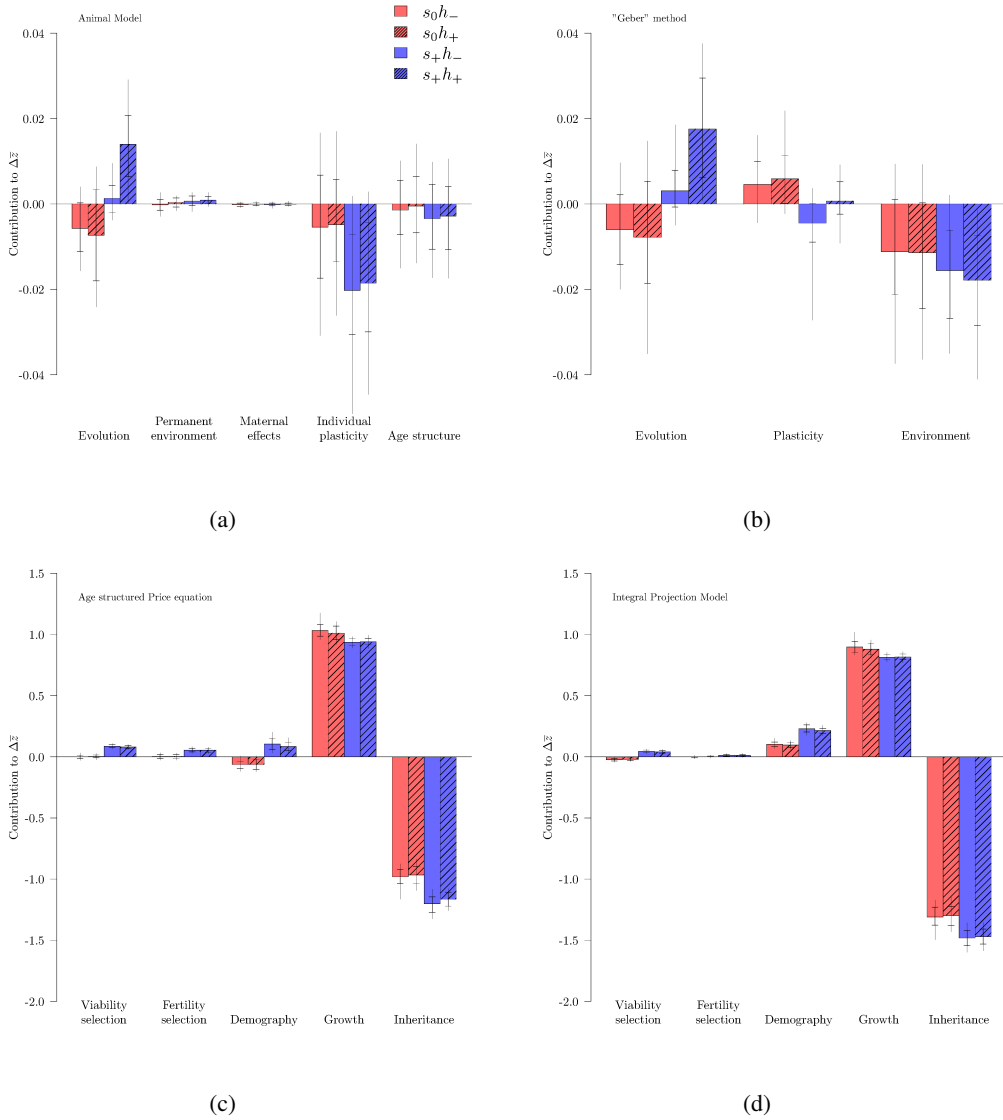
Figure 2.2: Results of the different frameworks when applied to the simulated scenarios. (a) Animal model. (b) 'Geber' method. (c) Age-structured Price equation and (d) Integral projection model. In (c) and (d), demography includes changes in average body size due to the age structure, inheritance is the sum of offspring mother difference and offspring difference covariance. In a-d, red bars indicate s_0 scenarios, blue bars indicate s_+ scenarios. Solid bars indicate h_- scenarios, and shaded bars indicate h_+ scenarios. Error bars represent the range in which 68% (error bars until horizontal lines) and 95% (entire error bars) of the contributions lie when applied to 100 replicates. The y-axis is always average contribution to mean trait change per year, although the scaling is different in a, b versus c, d.

AM and GM estimate a genetic change (due to selection and/or drift) whereas the IPM and GPE estimate selection. This is highlighted by the fact that the AM and GM estimate a much larger contribution of evolution in the s_+h_+ compared to the s_+h_- scenario. This contrasts with the IPM and APE, where the contribution of selection is independent of the heritability.

Due to a misspecification of the maternal effects in the AM, we find a negative contribution of evolution in the s_0 scenarios. This mismatch highlights the need to adapt the model structure to the study system. Only then reliable conclusions can be drawn from the AM (see also Hadfield et al. 2011). Indeed, we show that contributions are closer to the simulation process when we use a more appropriate specification of the maternal effects (A.2.2).

Here we have chosen to quantify the contribution of evolutionary change to trait dynamics by measuring the temporal change in BLUPs for breeding value in a univariate animal model. Within a quantitative genetic framework, we could also have used the heritability estimated by the AM to apply the breeder's equation and estimate the expected response to selection. This approach has proven its effectiveness under breeding conditions, although nonlinearities in the parent-offspring regression or the trait value-fitness relationship may bias predictions (Heywood 2005). More serious difficulties arise in natural populations, where the prediction of evolution can be biased when selection acts on genetically correlated traits or when an environmental variable dominates the covariation between traits and fitness (Rausher 1992; Morrissey et al. 2010).

A third approach relies on a bivariate AM that estimates genetic and environmental (co)variances between a trait and a proxy for relative fitness (Lande 1979; Lynch and Walsh 2014). The additive genetic covariance is of particular interest, as following the Robertson-Price identity it provides a direct estimate of the evolutionary change (Robertson 1966; Price 1970; Lynch and Walsh 2014). Although more data demanding, this approach does not require the assumptions of the breeder's equation to be fulfilled (Morrissey et al. 2012b), and avoids potentially biased trends in breeding values (Postma 2006).

Unlike the AM and GM, which quantify the change in breeding values, the APE and IPM estimate the contribution of selection, irrespective of whether this yields a genetic response. The overall contribution of selection is obtained by summing over all age-specific selection contributions. This is an attempt to remove the between-age covariation between traits and fitness (Engen et al. 2014), which is instead captured by the demography term. However, the age correction is not continuous, and therefore the choice of age classes determines how this total contribution of demography and selection is partitioned (see A.4.4 for an example).

Most studies that have applied the APE or IPM framework to natural vertebrate populations have found a relatively small role for selection in shaping trait dynamics (e.g. Ozgul et al. 2009; Traill et al. 2014). This is in line with our application, as even in the s_+ scenarios, the contribution of the other processes was estimated to be many times larger. In the IPM, the interpretation of selection in terms of evolutionary potential critically depends on the heritability. Heritability is, however, not assessed by the IPM. Indeed, the inheritance function relates juvenile to adult (maternal) trait values, and ignores the fact that individual growth trajectories may be heritable (Chevin 2015). Alternatively, trait inheritance can be incorporated in the IPM by implementing size at birth as a fixed trait influencing offspring size (Vindenes and Langangen 2015), or by explicitly modelling the transmission of additive genetic effects within the IPM (Coulson et al. 2015; Childs et al. 2016).

Plasticity

Plasticity includes all individual-level phenotypic changes that are not attributable to genetic changes. While all four frameworks estimate a large contribution of plasticity in all scenarios, they attribute them to different biological processes. This makes it difficult to directly compare the

importance of plasticity across frameworks and may potentially lead to confusion. In this section we will focus on plasticity in birth size.

We used the AM to separately estimate plasticity due to maternal and permanent environment effects (Fig. 2.2(b)). The contribution of maternal effects was very small. This may seem at odds with the effect of maternal adult size on offspring size at birth in our simulations, but as explained above, this was due to a mismatch between the model structure (which included a random effect of maternal identity) and the data generating process (which included an effect of maternal body size). The contribution of permanent environment was low, which is in line with the lack of a trend in the stochastic component of birth size in our simulations.

The GM captures plasticity in size at birth due to both maternal effects and stochasticity in one single term (Fig. 2.2(a)). Because plasticity at birth is here defined as the difference between actual birth weight and the breeding value for birth weight of an individual, by construction, the plasticity term has to compensate for the bias in estimated breeding values.

In the APE and IPM frameworks, plasticity at birth and growth are intrinsically entangled. Whereas ontogenetic growth forms the main plastic contribution to $\Delta\bar{z}$ (Figs. 2.2(c) and 2.2(d)), the body size that is attained through ontogenetic growth is only partially (through maternal effects) transmitted to the offspring. Most of the ontogenetic growth will thus be reset in the offspring: this is reflected in the strong negative contribution from inheritance (for a more detailed explanation of the inheritance terms, see A.4.3.1). Also, because we applied the APE only on the female part of the population, changes in offspring body size due to selection on males (and thus fathers) will be attributed to the inheritance term.

The role of the environment

Whereas the GM defines an explicit environmental factor, in the other frameworks, the environment influences trait dynamics only indirectly through selection, plasticity and/or demography. For example, high food availability may lead to an increase in average body size through plasticity. At the same time, increased food availability may decrease competition, and thereby affect selection.

In our implementation of the GM, we defined the environment as the total food intake of an individual. Hence, the environment mainly acts through within-individual plasticity through its effect on ontogenetic growth. Importantly, the outcome of the GM depends fully on how evolution, plasticity and environment are defined. When applying the GM to field data, where not all processes are known, it is thus crucial to first identify the main drivers and attribute them to evolutionary, plastic or demographic processes.

Although in the APE and IPM effects of the environment are implicitly present in all terms, in our implementation there is no explicit quantification of this environmental effect. Although an IPM can include an environmental variable, its contribution will not be quantified by the APE when applied to that IPM. However, alternative applications of the IPM that allow exploring the effects of such an environmental variable do exist (e.g. Vindenes et al. 2011). Alternatively, one can parametrize different IPMs for different environments (e.g. Ozgul et al. 2010) and use comparison methods such as life table response experiments to see how population and trait dynamics differ between these environments (Rees and Ellner 2009).

In our version of the AM, all contributions of changes in the environment, such as decreasing food availability, are captured within the residual individual plasticity term. Although not commonly done, environmental contributions can be estimated more explicitly by including additional fixed or random effects (Charmentier et al. 2014). One possibility is the inclusion of a fixed effect of food availability. Furthermore, it is possible to model interactions between the environmental variable and the additive genetic effects.

Demography

We showed how the combined demography terms in the APE scale with the covariance of age class-specific fitness and age class-specific average body size. The demography terms hence do not reflect the effect of changes in the age structure between time t and $t + 1$, but rather differences due to the existing age structure at time t. As such it provides a demographic correction of estimates of selection, similar to the one proposed by Engen et al. (2014).

In the AM we have quantified the demographic contribution by multiplying the slope of body size with respect to age with the predicted change in average age. This contribution is most negative in the s_+ scenarios, meaning that here a change (decrease) in the average age in the populations over time led to a decrease in the average body size in these scenarios, in agreement with the observed slight decrease in average age as shown in Fig. 2.1(f).

Unexplained variation and uncertainty

Making conclusive statements regarding which factor has the largest influence on $\Delta \bar{z}$ requires a measure of the uncertainty in the estimates of each contribution. So far we have only considered the range of point estimates over the replicates, generally showing smaller ranges for APE and IPM. However, APE and IPM were estimating processes that were constant throughout replicates (e.g. selection), whereas the AM and GM were estimating quantities subject to stochasticity (e.g. genetic drift). Differences in range are thus due to the stochasticity in the simulations rather than the uncertainty in the point estimates.

While the AM allows the estimation of confidence intervals for each estimated contribution, in our implementation of the IPM, APE and GM there is no direct measure of uncertainty. For the GM, confidence intervals can be obtained using bootstrapping methods (as in Ellner et al. 2011). As of yet, the lack of uncertainty quantification is a major drawback of the application of the IPM and APE. However, measures of uncertainty accompanying parameter estimates could be propagated to the decomposition, by using bootstrapping, and in the case of the IPM also by MCMC sampling.

Residual variance is explicitly quantified in the AM. The GM does evaluate the residuals of the underlying regressions, but does not include these in the final results (Ellner et al. 2011). In contrast, the APE is an exact framework and hence the residual variance is zero. However, it is still subject to sampling variance. Although the IPM uses the APE, it is constructed by fitting statistical models to the data, each with their own residual term.

The AM can also account explicitly for additional sources of variation, by including the corresponding random effects (for example, we incorporated individual identity as a random effect to account for individual heterogeneity that could not be explained by additive genetic variation). IPMs can also include a random individual effect in the underlying fitted functions. This inclusion accounts for individual heterogeneity when estimating vital rates. However, although this individual heterogeneity should explicitly be propagated to the actual IPM (Vindenes and Langangen 2015), the IPM is often parametrized with the random effect set to zero. Thereby not all individual heterogeneity is accounted for. Setting the random effect to zero might also bias the prediction because of Jensen's inequality (e.g. Fox and Kendall 2002). Individual heterogeneity can be incorporated by defining a "static trait", in addition to the continuous state variable. This static trait does not change during development, and reflects fixed individual heterogeneity caused by e.g. differences in size at birth, genetics or experienced environment (e.g. Ellner and Rees 2006; Vindenes and Langangen 2015). The role of individual heterogeneity is not captured in the GM and APE. In case of the GM, the effects of individual heterogeneity, as estimated by the AM, can be propagated to the response variable.

Conclusions and future directions

The urge for a better understanding of eco-evolutionary dynamics is reflected in the range of frameworks that have been developed over the last few years aiming at quantifying the underlying processes (Pelletier et al. 2009; Schoener 2011), especially within the light of the consequences of climate change (Gienapp et al. 2008; Lavergne et al. 2010). Yet, a general, predictive framework is lacking, and applications to field data remain scarce. We have shown that the animal model (AM), 'Geber' method (GM), age-structured Price equation (APE) and integral projection model (IPM) frameworks differ in generality and data requirements. Importantly, key processes are defined and interpreted differently in the different approaches. We emphasize that one should be careful when applying one of the frameworks and interpreting the outcomes as being the "true" contributions of different processes. Indeed, we have shown that each framework has its own set of components and definitions.

All four frameworks have only recently been proposed in their current form, and are only starting to be applied to conservation-related questions. In this review we have explored the frameworks and their assumptions and limitations. Our findings are summarized in Table 2.1, where we provide an overview of which framework seems most suitable for which research question. The AM enables estimation of quantitative genetic parameters, and genetic change in particular, that cannot be estimated by the other frameworks. However, the AM, and the estimation on quantitative genetic parameters in general, is data demanding and it can be difficult to isolate confounding sources of variation when data sets are small. When individual data on reproduction, survival and growth are available, and one is interested in explicitly quantifying the contribution of within-age class selection, IPM and APE are logical choices. The AM can explicitly evaluate the effect of individual heterogeneity. Although the IPM can take this information into account as well by fitting mixed effects models, it does not evaluate its effect on trait dynamics. In contrast to the other frameworks, only the GM focuses on population-level parameters, but knowledge (or assumptions) on processes is required beforehand, i.e. it must be known what processes are shaped by evolution (or plasticity) and which by the environment.

We conclude that in isolation none of the frameworks provides a full picture. Instead, each framework answers different questions and has different data requirements. By highlighting both the similarities and the differences, we hope to have aided in the interpretation of existing work. Furthermore, we hope this work will help researchers interested in eco-evolutionary questions in making an informed choice regarding the most suitable framework for their particular question.

Acknowledgements

Marjolein Bruijning[1], *Anne C. M. ten Berge*[1] *and Eelke Jongejans*[1]

1. Department of Animal Ecology and Physiology, Radboud University, 6500 GL Nijmegen, The Netherlands.
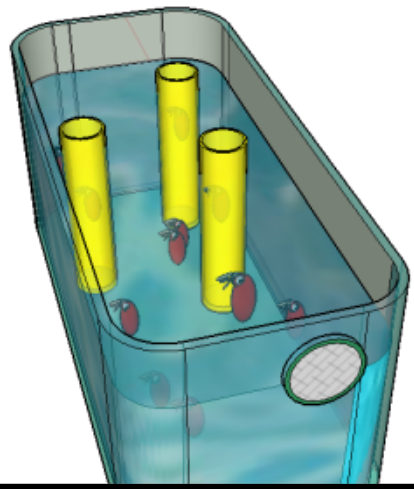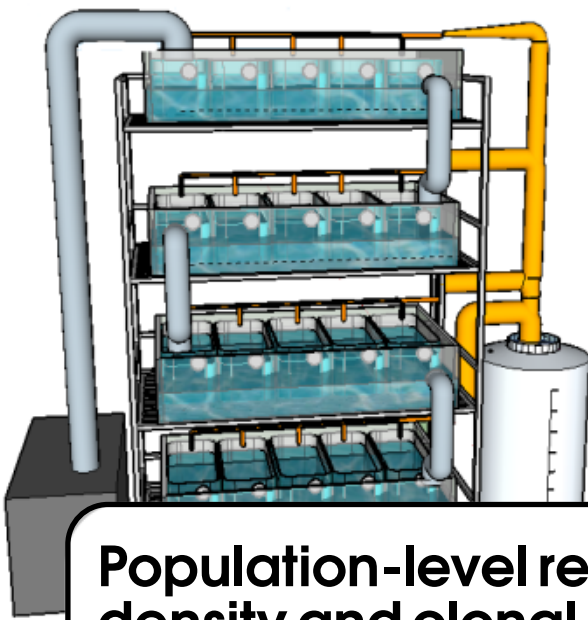
Population-level responses to temperature, density and clonal differences in *Daphnia magna* as revealed by Integral Projection Modelling

3.1 Abstract

1. Raising global temperatures are predicted to have strong consequences for ectotherms, as metabolic rates depend directly on external temperatures. To understand consequences for population fitness, a full life-cycle approach is important because i) temperature can have opposite effects on different vital rates (growth, survival, reproduction), and ii) sensitivities of population growth rate to changes in vital rates can vary in magnitude. Since vital rates are concurrently influenced by other factors, adequately predicting temperature effects requires factors like body size, population density and genetics to be taken into account.

2. The aim of this study was to quantify the role of temperature on all vital rates of *Daphnia magna* individuals, and their integrated effects on population dynamics. Additionally, we evaluated how clonal lineages differed in their temperature response, both on the vital rate and population-level.

3. We performed a laboratory experiment, in which we followed 40 populations (five clonal lineages × eight temperatures) during 80 days. Due to our novel setup, we were able to quantify vital rates of individuals within those populations. We identified relations between vital rates and body size, lineage, temperature and population density and used a size-structured Integral Projection Model to integrate the experimental effects over all vital rates.

4. We found negative density-dependence in growth and reproduction, resulting in lineage-specific carrying capacities. Population fitness showed a thermal optimum that differed among genotypes. Interestingly, we found that clones had different life history strategies, optimizing population fitness via different routes. As no lineage outperformed the others in all vital rates, we identified trade-offs between vital rates, which had strong effects on the dynamics of the population. Moreover, simulations suggest that the genetic composition of

mixed populations is temperature-dependent.

5. Our results underscore the importance of studying individuals within their population when predicting responses to environmental change. The observed density effects, which were as strong as temperature effects but explained considerably more variation in population growth, would have been overlooked in life table experiments. Furthermore, differential temperature responses emphasize the importance of genetic variation in the ability of ectotherm species like *Daphnia magna* to respond to climate change.

3.2 Introduction

Raising global temperatures are predicted to have strong consequences for ectotherms, as their metabolic rates directly depend on external temperatures (Huey and Berrigan 2001). To avoid local extinction populations must appropriately respond to these increasing temperatures, for instance by phenotypic plasticity or by evolution (Gienapp et al. 2008; Hoffmann and Sgrò 2011). However, disentangling plastic and evolutionary processes is not straightforward, and predicting their relative importance in natural populations is a major challenge (Pelletier et al. 2009; Lavergne et al. 2010; Schoener 2011; Chevin et al. 2013; Benthem et al. 2017).

To understand short-term environmentally induced changes in population dynamics, it is important to know how the performance of individuals within the population is affected, as it is the sum of the number of surviving individuals and number of newborns that determines the success of a population. Temperature effects on various life history traits of ectotherms have been studied extensively, using for instance life table experiments (e.g. MacArthur and Baillie 1929; Carvalho 1987; Van Doorslaer et al. 2009a). However, in addition to temperature, vital rates (growth, survival and reproduction) are shaped by many other factors, such as body size (Ozgul et al. 2012; Brooks et al. 2016), genotype (Dudycha and Tessier 1999; Geerts et al. 2015) and population density (Guisande 1993; Ozgul et al. 2012). The complex interplay between all these factors influences how individual plasticity and evolution will alter vital rates. Therefore, adequately predicting climate-driven changes in vital rates requires taking into account all these factors.

Integrating over all vital rates is a key element when studying eco-evolutionary dynamics (Smallegange and Coulson 2013). Without doing so, the combined effect for the population remains unknown (McLean et al. 2016). This is because population-level effects do not only depend on the observed effect sizes of changes in vital rates, but also on the sensitivity of population growth rate to these vital rate changes (Kroon et al. 2000). Moreover, changes can have opposite effects in different life stages. Positive and negative (including trade-offs) correlations between vital rates exist (Stearns 1989), and their net effects will be overlooked without integrating the effects over all life stages (Villellas et al. 2015). For example, a widespread phenomenon among ectotherms is that at higher temperatures, individual development rates increase, but individuals tend to mature at a smaller body size (Atkinson 1994; Atkinson 1995; Kingsolver and Huey 2008). Given that fecundity is often related to body size, temperature may result in life history changes that have opposite effects on population growth. Hence, the estimation of the net population-level effects of this 'temperature-size rule' requires integration over all life stages and multiple vital rates. Integral Projection Models (IPMs) are a powerful tool to integrate vital rates fitted to individual level data (Ellner and Rees 2006; Ellner et al. 2016), and have been used to study eco-evolutionary dynamics (e.g. Ozgul et al. 2010; Coulson et al. 2011; Traill et al. 2014; Chevin 2015).

In this study, we use the water flea *Daphnia magna* as a study system. This species has been widely used in studies on genetics (Colbourne et al. 2011), toxicology (e.g. Gust et al. 2016), as well as in studies on rapid evolution and eco-evolutionary dynamics (Hairston et al. 1999;

Van Doorslaer et al. 2009a; Van Doorslaer et al. 2010; De Meester et al. 2011). In addition, various factors have been shown to affect specific vital rates, such as effects of temperature (Henning-Lucass et al. 2016), genetic background (Pietrzak 2011; Henning-Lucass et al. 2016), food concentration (Gabsi et al. 2014) and population density (Guisande 1993). *Daphnia magna* individuals reproduce parthenogenetically when environmental conditions are favourable, and switch to sexual reproduction when conditions worsen (Kleiven et al. 1992), which results in the production of long-lived dormant eggs. This has the advantage that asexual reproduction can be assured in the lab by keeping conditions favourable. Therefore, genetic variation can be controlled, and the same genotypes can be used across treatments.

Although there are multiple studies on eco-evolutionary dynamics in *D. magna*, we are not aware of any study quantifying the role of ecological and genetic factors on the success of a population of interacting individuals, via their integrated effects on reproduction and survival of all life stages (see Duchet et al. 2010; Sommer et al. 2016, for parameterizations of a matrix population model based on isolated *Daphnia* individuals). Due to our novel setup, we are able to follow individuals within their population, to explicitly quantify how population density affects vital rates, in addition to the effects of temperature and genotype.

The aim of this study is to quantify how temperature, genetic background and population density affect the dynamics of *Daphnia magna* populations, and through which vital rates. We do so by performing a lab experiment, exposing five clonal lineages to a temperature gradient (10-26 °C), and following the populations for 80 days, while collecting both population-level and individual level data. By identifying relations between body size and demographic processes, and combining them into an Integral Projection Model, we aim to answer the following questions: 1) How does temperature affect vital rates of *D. magna* individuals within dynamic populations? 2) How do these effects on the individual level propagate to the population-level? 3) How do clonal lineages differ in their vital rates and population-level responses, and can we identify trade-offs between vital rates? 4) What is the relative importance of temperature, compared to genetic background and population density, in shaping populations dynamics?

3.3 Material and methods

Clonal lineages

In June 2014, we extracted *Daphnia magna* dormant eggs from mud collected in a small lake in Hilversum (Laapersveld), the Netherlands. These eggs were stored in the dark at 4 °C. From September 2014, we stimulated hatching by exposing the eggs to light and 20 °C. The eggs were checked daily and neonates were placed in individual 100 mL tubes, held in Dutch Standard Water (DSW; Hoefnagel et al. 2018) (200 mg l^{-1} $CaCl_2 \cdot 2H_2O$, 180 mg l^{-1} $MgSO_4 \cdot 7H_2O$, 100 mg l^{-1} $NaHCO_3$, 20 mg l^{-1} $KHCO_3$; NEN6503 1980) and fed with instant algae ($1.6 \cdot 10^5$ cells/mL; Shellfish 1800, Reed Mariculture). Medium, containing new food, was refreshed three times per week. In total, out of the 50 dormant eggs, we established 22 lineages. Of these lineages, 12 were successfully kept alive until the beginning of the experiment, in December 2015. Four lineages were randomly chosen for the experiment. Additionally, we included a *D. magna* lineage that had been successfully held in the lab for more than ten years, originally also extracted from a lake in The Netherlands (Lürling and Tolman 2010). This enabled us to specifically look at within population variation as well as comparing dynamics to a lineage that has been known to perform well under lab conditions. In total, this resulted in five clonal lineages (A1-A4 indicating the newly collected lineages, L indicating the existing laboratory lineage). Note that lineages L, A3 and A4 are referred to as C, D and E, respectively, by Hoefnagel et al. (2018).

Experimental setup

The experiment was conducted in a climate chamber, with a 16:8 hour light:dark regime and temperature set at 18 °C. We established a gradient of eight temperatures ranging between 10 °C and 26 °C. We achieved the cooler temperatures by placing four basins containing demineralized water above each other and cooling the upper basin to 10 °C using a water bath. Demineralized water was continuously pumped from the lowest basin to the upper basin. Using an overflow system, cold water continuously flowed from a basin to the one below it, slowly reaching chamber temperature (18 °C), resulting in a gradient from 10 to 18 °C. The same was done for the four warmer temperatures: the upper basin was heated to 26 °C, and by using a continuous overflow, a gradient was obtained. In this way, we generated a stable temperature gradient of eight temperatures: 10.5, 14.3, 15.5, 17.0, 20.0, 22.3, 23.5 and 25.9 °C. Temperatures were constantly measured by temperature loggers. In each basin, we placed five 2-liter aquaria, each aquarium containing one population of a different lineage (see Appendix B.1 for a schematic drawing).

The experiment ran between December 1st, 2015 and February 18th, 2016, lasting 80 days. Prior to the experiment, individuals were placed in the experimental setup for more than three weeks to acclimatize to their respective temperature treatments and to reduce differences due to maternal effects. Using eight temperatures, and five lineages, this resulted in 40 experimental units. On day 1, we arbitrary chose 20 individuals from each aquarium, reflecting the full range of body sizes (i.e. from small juveniles to adults), to start the experiment with. Every eight hours, an automatic pump system added 200 mL fresh DSW medium including instant algae ($8 \cdot 10^6$ cells/mL) to each aquarium. Volume in each aquarium was held constant by an overflow system, as each aquarium contained a sieve (0.3 mm sieve size). For comparison, the smallest measured neonate was 0.67 mm, and 95% of the measured neonates were larger than 0.83 mm. Of all measured individuals, 95% was > 0.97 mm.

We placed three transparent PVC tubes (4 cm diameter) in each aquarium, containing 12 holes (1 cm diameter) covered with permeable filters (0.125 mm sieve size; preventing neonates to escape or enter the tubes), allowing food and other cues to pass. These were used to isolate individuals for either three or four days, to obtain individual measurements, while ensuring that the individuals experienced the same environment as the rest of the population.

Population counts

For each aquarium, measurements were done twice per week; for half of the aquaria this was on Monday and Thursday, for the other half on Tuesday and Friday. This resulted in a time interval of either three or four days. On these days, each population was transferred to a petri dish, which was placed in a fixed camera setup. A movie of approximately 4 seconds was made with a digital camera (Sony Handycam, HDR-CX115). We used newly developed R-package *trackdem* to obtain estimates of population counts (Bruijning et al. 2018c). In short, movies were converted to an image sequence and loaded in the R-environment (R Core Team 2016). As only individuals move, a background image was created, containing all motionless objects. By subtracting all images from this background, moving particles were detected. Identification was optimized using machine learning. Individual trajectories were subsequently reconstructed (Jaqaman et al. 2008). We obtained 20 counts for each of the 40 populations.

Individual measurements

At the same time as the population counts, we collected data on individuals in the tubes, using a stereo microscope. We noted whether the individual was alive and measured its body size, as measured from the base of the spine until the middle of the eye. We counted the number of eggs in the brood pouch, and noted the stage of the eggs (1: round, no eyes, 2: oval shape, 3:

development of eyes and limbs) for 317 individuals, across all temperatures and lineages. If present, we counted the number of released (alive) offspring and measured the size of one of the neonates. All these individuals were joined with their respective population. The aquarium and tubes were rinsed with hot water, and we arbitrarily selected a new individual for each tube. These newly chosen individuals were also measured and their number of eggs counted. Finally, all other individuals were placed back in the aquarium, and the aquarium was randomly positioned in the appropriate basin (using lists of random placements, made in R). This resulted in 2293 observations of demographic rates (3 individuals $\times$ 5 clones $\times$ 8 temperatures $\times$ twice a week $\times$ 11 weeks of measuring; three populations went extinct early during the experiment, see *Results* section, explaining the discrepancy) over either a three or four days interval, providing information on size-dependent survival, growth, probability of carrying eggs, probability of reproduction, clutch size and neonate size (from now on called 'vital rates') (Bruijning et al. 2018a).

Explanatory variables

We explored population density effects on vital rates (Appendix B.2) and found negative density dependence in growth and reproduction, and positive density dependence in survival. We therefore included density in regressions related to growth and reproduction. The apparent positive density dependence found for survival is addressed in the discussion. Prior to all analysis, we standardized body size, temperature and population size to enable a comparison of effect sizes. For regressions including lineage, we used lineage A1 as a reference category, and all other lineages as contrasts to A1.

Model framework

Integral Projection Models (IPMs) describe the dynamics, of a population in which individuals are characterized by a continuous state variable, in discrete time (Ellner and Rees 2006; Ellner et al. 2016). We used (standardized) body size z as the continuous state variable. The IPM consisted of four kernels, describing how z influences all vital rates.

   We have constructed IPMs following four different procedures, with increasing complexity. IPM_1 was constructed using the collected data directly, describing transitions on an approximately 3.5-day basis. IPM_{2-4} described daily transitions, requiring a translation of collected data (twice per week) to daily rates, as explained below. For notation, daily rates have subscript d. IPM_{2-4} differed from each other in the reproduction kernel. For IPM_2 we divided estimated reproduction by 3.5, the average time interval. For IPM_3 and IPM_4, we used data on the egg stages to estimate temperature-dependent daily development rates and the average number of days it takes early-stage eggs to develop into neonates (Appendix B.3 for more details). Finally, IPM_4 was a size- and stage-structured model in which individuals were, in addition to body size, characterized by a discrete developmental stage of the eggs they carried. Here, we defined four discrete stages (1: round, no eyes, 2: oval shape, 3: development of eyes and limbs, 4: released neonates); individuals had to move through all stages before offspring was born and added to the population.

   We here provide details and results for IPM_2. Because we preferred an IPM structure with a daily time step, we did not use IPM_1. IPM_{2-4} used the same vital rates except for details on reproduction, but IPM_2 was the least complex. See Appendices B.4-5 for details and results of the other approaches, which were to a large degree similar to those of IPM_2.

   The constructed IPM predicts the body size distribution at day $t+1$ ($n(t+1,z')$), given the body size distribution at day t ($n(t,z)$). The four kernels describing all daily transition probabilities were 1) survival $S_d(z)$, 2) growth $G_d(z'|z)$, describing probabilities for surviving individuals of size z at time t to obtain size z' at day $t+1$, 3) reproduction $R_d(z)$, and 4) an offspring size distribution $D_d(z'|z)$ describing probabilities of obtaining offspring with size z' at $t+1$ given a maternal size z

at day t. We created a composite IPM, whereby the four kernels were functions of, in addition to z, temperature T, observed population density N as estimated using *trackdem* (Bruijning et al. 2018c) (except for survival) and lineage C:

$$n(t+1,z') = \int [S_d(z,T,C) \cdot G_d(z'|z,T,N,C) + R_d(z,T,N,C) \cdot D_d(z'|z,T,N,C)]n(t,z)dz \quad (3.1)$$

Reproduction ($R_d(z,T,N,C)$) was defined as the product of the probability of carrying eggs (p), the probability of having live offspring at the end of a half-week interval conditional on carrying eggs (f), and clutch size at birth (L_0), divided by 3.5 to translate to daily estimates:

$$R_d(z,T,N,C) = p(z,T,N,C) \cdot f(z,T,N,C) \cdot L_0(z,T,N,C) \cdot \frac{1}{3.5} \quad (3.2)$$

Note that $p(z,T,N,C)$ and $f(z,T,N,C)$ do not have subscript d, since data did not need to be translated to obtain daily estimates. We translated observations on clutch size to predict clutch size at birth $L_0(z,T,N,C)$, as explained below.

New size distribution and offspring size distribution were functions of the size-dependent expected growth ($\hat{g}_d(z)$) and expected offspring size at birth ($\hat{\phi}_0(z)$), respectively, and the estimated variation around these means (σ_g and σ_ϕ):

$$G_d(z'|z,T,N,C) = Normal(z'|\hat{g}_d(z,T,N,C), \sigma_g) \quad (3.3)$$

$$D_d(z'|z,T,N,C) = Normal(z'|\hat{\phi}_0(z,T,N,C), \sigma_\phi) \quad (3.4)$$

The IPM was discretized into a 100×100 matrix, with (standardized) z ranging between -3 and 3, corresponding to 0.07 and 4.1 mm, respectively.

Estimation of vital rates

The collected data on individuals were used to estimate all vital rates needed to parameterize the IPM kernels. For each vital rate, we tested all models including additive effects of body size z, temperature T, population density N and clonal lineage C. We also included a quadratic effect of z and T because visual inspection revealed non-linear effects. Finally, we included two-way interactions between z and T, z and C, C and T and C and N. The most complex model was thus:

$$y(z,T,N,C) = \beta_0 + \beta_1 z + \beta_2 T + \beta_3 N + \beta_4 z^2 + \beta_5 T^2 + \beta_6 zT +$$
$$\sum_{i=1}^{4} (\beta_{6+i}C_i + \beta_{10+i}zC_i + \beta_{14+i}TC_i + \beta_{18+i}NC_i) + \varepsilon_{res} \quad (3.5)$$

Here, C_i is clone i (A2, A3, A4 or L); effects of clone i are compared to those for clone A1. We fitted all 196 models nested within this model. Instead of choosing the best model based on AIC, we applied model averaging (based on AIC weights) over all models to obtain averaged parameters and standard errors using the R-package *MuMIn* (Burnham and Anderson 2002; Bartoń 2016), after ensuring that explanatory variables were only weakly correlated (r^2_{T-N}=0.039, r^2_{z-N}=0.00029, r^2_{T-z}=0.028). We chose to perform model averaging because this results in more robust models, where there is model uncertainty (reflected by similar AIC values across different models). Model averaging has been shown to improve prediction accuracy and reduce the risk of finding spurious effects (e.g. Madigan and Raftery 1994; Raftery et al. 1997; Hoeting et al. 1999; Burnham and Anderson 2004; Yang 2007; Lukacs et al. 2010). We used the conservative zero method for averaging coefficients, in which parameters are assigned a zero if not present in a model (Grueber et al. 2011). In the case of a log or logit link function, averaging coefficients may yield different results than averaging predictions, but differences were negligible in our case (see Appendix B.6). The procedure for each vital rate will now be explained.

Survival probability

As time intervals between measurements varied, we estimated survival as a function of time between measurements (in days; Δt):

$$\left[\frac{1}{1+\exp(-k(z,T,C))}\right]^{(1/\Delta t)} = S_d(z,T,C)^{(1/\Delta t)} \tag{3.6}$$

Here, $k(z,T,C)$ is the linear relation between the explanatory variables (Eq. 3.5). To calculate the exact time interval (Δt), we used the time at which the population was filmed. Intervals ranged between 2.8-3.2 days, and between 3.8-4.2 days. We optimized a likelihood function to fit $S_d(z,T,C)$. In accordance with the other vital rates, we tested all different models and performed model averaging based on AIC.

Growth

To fit daily growth $\hat{g}_d(z,T,N,C)$, individual growth was first calculated by dividing observed size increment by the time interval (Δt), assuming that growth rates were constant within these days. Daily growth was then fit as a linear function of size, temperature, density and lineage. Growth variation σ_g was calculated as the standard deviation of the residuals.

Probability of carrying eggs

We only included individual measurements when they were placed in the tube (i.e. not using individual measurements 3 or 4 days later) to avoid pseudoreplication caused by repeated measurements. We dichotomized the number of eggs in the brood pouch into zeros (0 eggs) and ones (>0 eggs). By performing a logistic regression and model averaging, we fitted $p(z,T,N,C)$.

Probability of producing offspring

We included all individuals carrying eggs at day 1 (i.e. the day when they were placed into the tubes). Individuals that produced neonates when remeasured were assigned ones, other individuals were assigned zeros. These binomial data were used to fit probability of producing offspring after, on average, 3.5 days, $f(z,T,N,C)$, conditional on carrying eggs, using logistic regression.

Clutch size

To estimate clutch size at birth $L_0(z,T,N,C)$, we took into account that born offspring (observed when remeasuring the parent) could have been born 0-4 days earlier. To do so, we first used $\hat{g}_d(z,T,N,C)$ to predict neonate body sizes 0, 1, 2, 3, 4 days earlier, based on neonate size when measured (and relevant temperature, density and lineage). Second, we predicted survival probabilities for each day, given predicted body sizes, using $S_d(z,T,C)$. For all days, the probability of surviving until the measurement was calculated, by multiplying probabilities with probabilities of consecutive days. Subsequently, observed clutch size was divided by each of these probabilities. This gave the expected clutch size, if offspring would have been born on that day, given the observed clutch size and given expected survival probabilities. We assumed equal birth probabilities for each day, and averaged these predicted clutch sizes. We log-transformed these estimates and fitted clutch size at birth $L_0(z,T,N,C)$, using linear regression.

Neonate body size at birth

A similar procedure was followed to estimate neonate body size at birth $\hat{\phi}_0(z,T,N,C)$. We used the fitted growth function ($\hat{g}_d(z,T,N,C)$) to back-calculate the sizes 0-4 days earlier. Assuming

equal birth probabilities for each day, we took the average of these numbers as an estimate for offspring body size at birth. Estimates were fit as a function of maternal size, temperature, density and lineage. We calculated the standard deviation of residuals to estimate variation in estimated offspring body size σ_ϕ.

Integrated effects of temperature and density

The fitted vital rates were used to quantify the population-level effects of the temperature and density, using an IPM (Eq. 3.1). To evaluate how temperature and density effects in single vital rates propagated to the population, we performed the following analyses: first, we calculated temperature effects on projected (asymptotic) population growth rate (λ), which is the dominant eigenvalue of the discretized IPM. To do so, we constructed an IPM for temperatures ranging between 9 °C and 28 °C, for each lineage, and calculated λ. Here, we set N at 0, i.e. average density (corresponding to 96 individuals). Second, density effects (at average temperature, 18.6 °C) were calculated by varying density between 1 individual and the maximum lineage-specific density across all temperatures. We calculated λ for each density and lineage. Third, we analyzed how much of the variation in projected population growth rate was explained by variation in temperature, lineage and density. We used each of the observed population sizes during the experiment, and projected λ based on an IPM with the corresponding lineage, temperature and density, resulting in 915 growth rates. A linear regression was performed between λ and T (and T^2), C or N, and we calculated coefficients of determination (R^2) for each of these three regressions.

Population-level consequences of lineage differences

We quantified the population-level consequences of lineage differences in single vital rates. We started by calculating the average IPM among all lineages, by taking the average matrix for a specific temperature (λ_T). Subsequently, we systematically replaced one of the vital rates (averaged across lineages) by the vital rates of each lineage and recalculated λ. The difference between λ_T and λ when one of the vital rates is replaced combines the effect size and the sensitivity of λ to that vital rate. This was done for temperatures ranging between 9 °C and 28 °C. To evaluate within-population differences, we performed the same analysis excluding lineage L, which originated from a different population.

Quantifying uncertainty in population growth rates

We used bootstrapping to obtain measures of uncertainty in population growth rates. Per population, we resampled individual observations from the dataset, with replacement. This way we created 500 bootstrapped datasets. Vital rate models were fitted and IPMs were constructed as described above for each of the 500 datasets. We performed all above described analysis with the constructed IPMs, i.e. lineage specific temperature and density effects, and how differences in vital rates between lineages propagated to the population-level.

Simulating dynamics of a hypothetical mixed population

Finally, we used IPMs to simulate a scenario in which lineages were mixed in one population, to explore whether the genetic composition is predicted to be temperature-dependent. We simulated density-dependent dynamics with all lineages starting with four 1.7 mm individuals, which is the average observed body size (i.e. standardized body size set at 0). For each time step, we constructed a density-and temperature-dependent IPM per lineage, using the total density across all lineages. We projected population size, per lineage, at $t + 1$ by matrix multiplying the appropriate IPM with the lineage-specific population structure at time t. Populations were projected for 100

days, and this analysis was done for temperatures between 9 °C and 28 °C. At day 100, for each temperature, the lineage composition was evaluated. To focus specifically on within-population clones, this analysis was also performed without lineage L.

3.4 Results

Population trends

Thirty-seven out of the 40 populations remained viable during the 80 days of the experiment. Three populations of lineage A4 (at 17 °C, 22.3 °C and 25.9 °C) went extinct after 10 days. After acclimatization, we restarted these populations with 20 individuals on day 49, from which they remained viable until day 80. Most populations started with an initial increase, although the rate and timing of that increase differed between temperatures and lineages (Appendix B.7). The populations fluctuated considerably in size over time, with a maximum number of 500 individuals (for lineage L at the lowest temperature, after 40 days).

Vital rates

For all vital rates, estimated rates are shown for lineage A1, average density, and for both the highest (25.9 °C), and the lowest (10.5 °C) temperature (Fig. 3.1). See Appendix B.8 for estimated vital rates per temperature.

Survival

The best model (having 67% of the weight) describing daily survival included an effect of size, squared size, temperature, lineage and an interaction between size and temperature (Tables S2, S8-9 in Appendix B.9). Survival probabilities showed an optimum for medium sizes, and lower survival at higher temperatures (β_4, β_2 in Table 3.1; Fig. 3.1a), as well as a negative interaction between size and temperature (β_6), such that survival was further reduced at higher temperatures for larger animals. Parameters β_2, β_4, β_6 were all significant in both the averaged model and in the best model (Tables 3.1, S8). Lineage A2 had highest and lineage A3 the lowest daily survival probabilities (β_{7-10}).

Growth

Daily growth $\hat{g}_d(z, T, N, C)$ was on average 0.058 mm (for lineage A1, and body size, density, temperature all set at their mean) and decreased with body size (β_1; Fig. 3.1b). The best model (17% of the weight) included effects of size, squared size, density, lineage, temperature, squared temperature, and interactions between size and lineage, and size and temperature (Tables S3, S9). A negative interaction between temperature and size was found (β_6; Fig. 3.1b), suggesting that at higher temperatures, individuals initially grew faster, but stop growing at a smaller size. The cumulative sum of the AIC weights for main effects of size, temperature, density and genetic lineage was 1.00 for each of these variables, as well as a size × temperature interaction. All averaged coefficients that differed significantly from zero, were also significant according to the best model (Tables 3.1, S10-11). Variation in growth (σ_g) equaled 0.048.

Carrying eggs

Probability of carrying eggs $p(z, T, N, C)$ equaled practically zero for individuals of up to 1.5 mm and increased with size (β_1; Fig. 3.1c). Temperature had a positive effect on estimated probabilities (β_2), and this effect was significant in both the averaged and best model (Tables 3.1, S12). For the largest individuals, probabilities started decreasing as reflected by the significant negative squared size effect (β_4), in both the averaged and best model. Significant lineage effects were

found (β_{7-10}). Finally, density had a negative effect on the probability, although insignificant in the averaged and best model. The best model (with 32% of the weight) was the full model, including all additive effects and interactions (Tables S4, S12-13). Cumulative AIC weights for main effects of size, temperature, density and genetic lineage, and interactions between size and temperature, equaled 1.00.

Probability of producing offspring

The best model (17% of the weight) describing included all additive effects, as well as the interactions between lineage and density (Tables S5, S14-15). The cumulative sum of the AIC weights was 1.00 for main effects of size, temperature and density, and 0.97 for genetic lineage. Probabilities $f(z,T,N,C)$ increased with temperature and body size (β_1, β_2 in Table 3.1; Fig. 3.1d), and lineage A4 showed significant higher probabilities compared to lineage A1 (β_9). A significant negative effect of density was found (β_3). All significant coefficients were present, and significant, in the best model (Tables 3.1, S14).

Clutch size at birth

Number of offspring $L_0(z,T,N,C)$ increased with size (β_1; Fig. 3.1e; significant in both the averaged and best model, see Tables 3.1, S16-S17). At higher temperatures, slightly higher clutch sizes were reached, but temperature effects were insignificant and the cumulative sum of the AIC weights of models including main effects of temperature was 0.57. Differences in AIC values were small, indicating that there was support for a wide range of models (Tables S6).

Neonate body size distribution at birth

Predicted neonate body size at birth averaged 0.80 mm for a mother of 3 mm (for lineage A1 and all other variables set at 0) and increased with maternal squared size (β_4; Fig. 3.1f; significant in both the averaged and best model, see Tables 3.1, S18-19). Temperature had practically no effect and lineage A2 produced significant (in both the averaged and full model) larger offspring compared to lineage A1 (β_7). Again, differences in AIC were small, indicating support for a wide range of models (Table S7), although the cumulative sum of AIC weights including main effects of temperature, genetic lineage and squared body was, equaled 1.00 for each of these variables. Variation in offspring size σ_ϕ equaled 0.25.
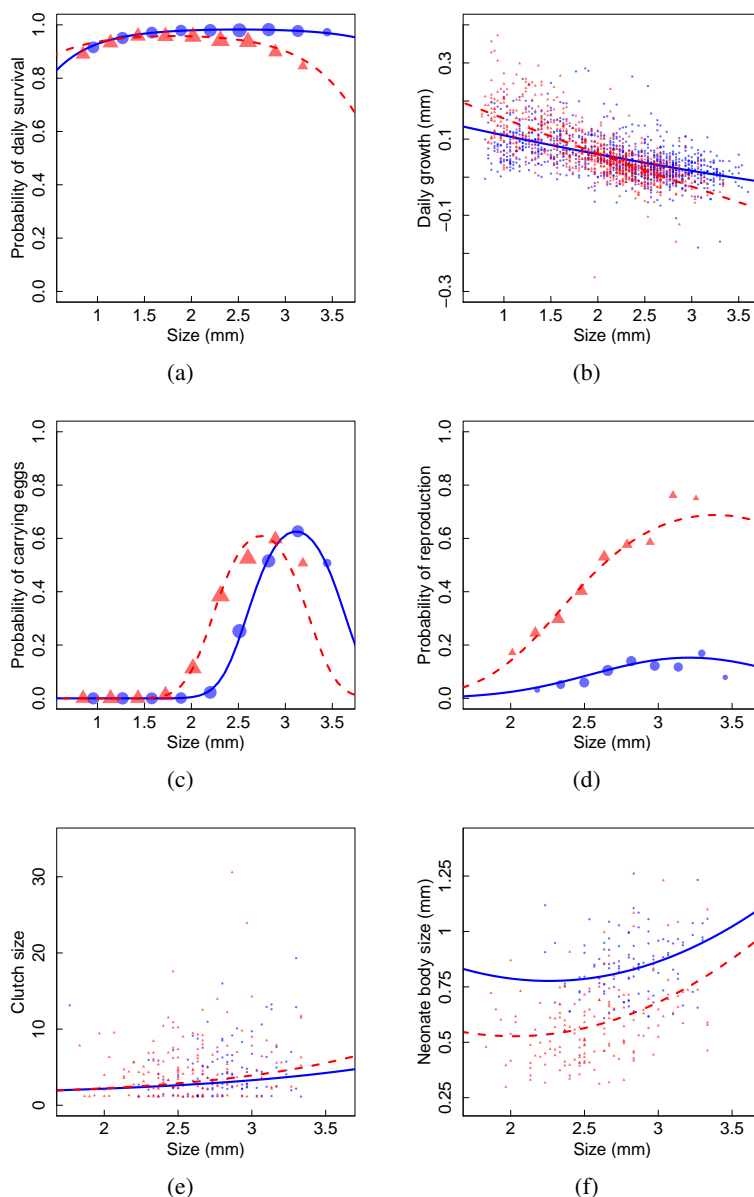
Figure 3.1: Estimated vital rates, shown for lineage A1 at high temperatures (25.9 °C; red dotted lines) and low temperatures (10.5 °C; blue solid lines), according to the full averaged model (fitted across all temperatures and linages) (coefficients in Table 3.1). Note that the six intermediate temperatures fall between the shown model predictions for the lowest and highest temperatures. In a) daily survival probability $S_d(z,T,C)$, b) daily growth $g_d(z,T,N,C)$, c) probability of carrying eggs $p(z,T,N,C)$, d) probability of reproduction $f(z,T,N,C)$, e) clutch size $L_0(z,T,N,C)$ and f) neonate body size $\phi_0(z,T,N,C)$. All predictions are shown for lineage A1, and for average population density. Dots represent partial residuals, which plot the residuals 'around' the two plotted lines of size-dependent model predictions. We separately did so for the four highest temperatures (red triangles around the red lines), and for the four lowest temperatures (blue circles around the blue lines). Those show the density of observations over all body sizes, irrespective of density, lineage or temperature. In a), c) and d), partial residuals are averaged per size class, and dots are scaled to the number of data points.

		$S_d(z,T,C)$	$G_d(z,T,N,C)$	$p(z,T,N,C)$	$f(z,T,N,C)$	$L_0(z,T,N,C)$	$\phi_0(z,T,N,C)$
	β_0	**3.45e+00 [1.02e-01]**	**5.82e-02 [3.27e-03]**	**-2.72e+00 [3.27e-01]**	**-2.24e+00 [5.26e-01]**	2.57e-01 [2.35e-01]	**-2.15e+00 [8.32e-02]**
z	β_1	4.55e-02 [4.04e-02]	**-4.59e-02 [2.51e-03]**	**5.31e+00 [5.58e-01]**	**2.20e+00 [9.52e-01]**	**5.01e-01 [1.80e-01]**	-1.52e-02 [8.45e-02]
T	β_2	**-2.65e-01 [4.91e-02]**	-6.90e-04 [1.97e-03]	**9.32e-01 [2.14e-01]**	**6.10e-01 [2.24e-01]**	1.96e-02 [1.33e-01]	**-1.20e-01 [6.10e-02]**
N	β_3		**-6.45e-03 [3.07e-03]**	-2.50e-01 [1.77e-01]	**-7.67e-01 [3.84e-01]**	-8.42e-02 [1.25e-01]	1.20e-02 [2.36e-02]
z^2	β_4	**-2.92e-01 [4.11e-02]**	1.26e-03 [1.42e-03]	**-2.14e+00 [2.58e-01]**	-6.04e-01 [4.83e-01]	9.99e-03 [6.84e-02]	**1.07e-01 [2.70e-02]**
T^2	β_5	1.04e-02 [1.10e-02]	-1.32e-03 [1.45e-03]	-1.19e-01 [9.10e-02]	-7.48e-02 [1.11e-01]	-2.49e-03 [5.57e-02]	-8.33e-05 [1.14e-02]
zT	β_6	**-1.88e-01 [4.00e-02]**	**-9.02e-03 [1.31e-03]**	**-7.60e-01 [1.69e-01]**	9.96e-02 [2.22e-01]	4.16e-02 [1.11e-01]	2.41e-02 [3.57e-02]
C	β_7	**5.67e-01 [1.50e-01]**	**-1.24e-02 [3.79e-03]**	-7.81e-02 [3.71e-01]	3.27e-01 [5.03e-01]	-3.93e-01 [3.76e-01]	**2.68e-01 [1.01e-01]**
	β_8	**-2.40e-01 [1.21e-01]**	-4.30e-03 [3.71e-03]	**1.03e+00 [3.89e-01]**	-7.48e-02 [4.84e-01]	1.47e-01 [2.47e-01]	-7.55e-03 [8.76e-02]
	β_9	-9.24e-02 [1.32e-01]	4.64e-03 [3.89e-03]	5.51e-02 [4.16e-01]	**1.14e+00 [4.61e-01]**	9.94e-02 [2.51e-01]	5.27e-02 [1.45e-01]
	β_{10}	5.45e-03 [1.29e-01]	-6.85e-03 [3.55e-03]	**1.83e+00 [3.99e-01]**	6.94e-01 [3.76e-01]	1.05e-01 [2.32e-01]	1.65e-01 [1.10e-01]
zC	β_{11}	-8.41e-04 [3.05e-03]	7.77e-03 [4.23e-03]	-1.40e-01 [3.94e-01]	8.13e-02 [3.59e-01]	5.33e-03 [7.93e-02]	-8.37e-02 [9.56e-02]
	β_{12}	-3.37e-03 [3.88e-03]	1.69e-03 [3.49e-03]	-3.42e-01 [4.12e-01]	-4.65e-03 [1.83e-01]	-6.75e-03 [8.33e-02]	-5.19e-02 [7.66e-02]
	β_{13}	-2.78e-03 [3.54e-03]	6.40e-04 [3.58e-03]	2.98e-01 [4.17e-01]	1.62e-02 [2.07e-01]	-3.48e-03 [1.02e-01]	-1.39e-01 [1.41e-01]
	β_{14}	6.35e-05 [2.55e-03]	-4.44e-03 [3.85e-03]	-5.45e-01 [5.37e-01]	-1.94e-02 [1.94e-01]	1.06e-02 [9.46e-02]	-1.29e-01 [1.22e-01]
TC	β_{15}	2.02e-02 [3.92e-02]	-1.97e-05 [1.76e-03]	**-6.15e-01 [2.55e-01]**	-3.81e-03 [6.49e-02]	1.63e-03 [7.13e-02]	5.00e-02 [7.72e-02]
	β_{16}	-3.97e-02 [4.55e-02]	1.87e-03 [3.65e-03]	-2.72e-01 [2.29e-01]	1.54e-02 [5.32e-02]	1.41e-02 [9.22e-02]	2.67e-02 [5.13e-02]
	β_{17}	-5.48e-02 [5.75e-02]	6.64e-04 [2.21e-03]	-1.06e-01 [2.08e-01]	-2.05e-03 [5.63e-02]	1.05e-02 [7.97e-02]	3.49e-03 [3.95e-02]
	β_{18}	2.66e-02 [3.83e-02]	1.06e-03 [2.61e-03]	**-5.34e-01 [2.56e-01]**	3.53e-03 [5.61e-02]	-4.97e-03 [6.82e-02]	3.50e-02 [6.14e-02]
NC	β_{19}		2.74e-03 [4.38e-03]	-1.85e-01 [2.35e-01]	3.26e-01 [4.52e-01]	-3.53e-03 [5.83e-02]	-7.24e-04 [1.30e-02]
	β_{20}		-3.44e-04 [3.15e-03]	-4.90e-01 [4.32e-01]	-4.46e-01 [5.84e-01]	-4.63e-03 [7.41e-02]	-9.87e-04 [1.62e-02]
	β_{21}		3.50e-04 [2.95e-03]	-8.58e-02 [2.38e-01]	4.28e-01 [5.04e-01]	-2.24e-04 [5.44e-02]	-5.70e-04 [1.31e-02]
	β_{22}		2.04e-03 [3.61e-03]	-1.94e-02 [1.85e-01]	5.70e-01 [5.15e-01]	2.55e-03 [5.69e-02]	-1.91e-03 [1.79e-02]

Table 3.1: Weighted coefficients and standard errors for each of the vital rates. Daily survival probability $S_d(z,T,C)$, daily growth $G_d(z,T,N,C)$, probability of carrying eggs $p(z,T,N,C)$, probability of reproduction $f(z,T,N,C)$, clutch size at birth $L_0(z,T,N,C)$ and neonate body size at birth $\phi_0(z,T,N,C)$. Bold numbers indicate significant effects ($p \leq 0.05$). The following linear regression was used for each vital rate (but with different link functions; see main text): $y(z,T,N,C) = \beta_0 + \beta_1 z + \beta_2 T + \beta_3 N + \beta_4 z^2 + \beta_5 T^2 + \beta_6 zT + \sum_{i=1}^{4}(\beta_{6+i}C_i + \beta_{10+i}zC_i + \beta_{14+i}TC_i + \beta_{18+i}NC_i) + \varepsilon_{res}$.

Integration over all vital rates

We here report the results using estimates of daily vital rates to construct IPM_2. Results from daily based IPM_3 and IPM_4 are similar and do not change our conclusions. IPM_1, describing transitions after one measurement interval (on average 3.5 days), gave more divergent results (more details and results in Appendix B.4-5).
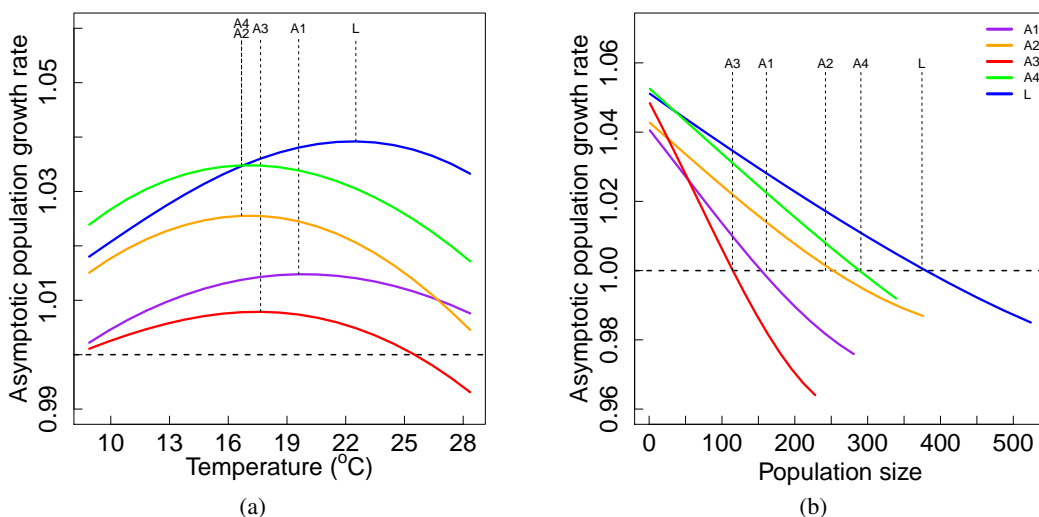

Figure 3.2: Population growth rate (λ) as a function of a) temperature and b) population size. In a) population size is set at average, in b) temperature is set at average. Each colour indicates a different lineage. See Appendix B.11 for results taking into account uncertainty in vital rates estimates.

Effects of lineage, temperature and density

We found that lineage, temperature and density all affected asymptotic population growth rate (λ) (Fig. 3.2; Appendix B.10). In general, lineage L showed the highest λ across all temperatures, with rates above 1 for up to almost 400 individuals. At average density, it is only at the lowest temperatures that lineage A4 showed slightly higher growth rates. Although all lineages showed a wide thermal optimum within the range of tested temperatures, this optimum differed between lineages (Fig. 3.2a). Population growth rate of lineage A2 started decreasing more rapidly with increasingly high temperatures, whereas lineage L outperforms the other lineages in particular at the highest temperatures. Lineage A1 and A3 showed lowest and similar λ, and a relatively weak response to temperature. Bootstrapped datasets were used to obtain measures of uncertainty in population growth rates. Despite the considerable amount of variation in asymptotic growth rates (Fig. S16 in Appendix B.11), lineage L has the highest thermal tolerance in the majority of the cases. In 96%, 98%, 95% and 83% of the cases, lineage L showed a higher population growth rate at the highest temperature, compared to lineage A1, A2, A3 and A4, respectively. A thermal optimum was within the range of tested temperatures in 62% (lineage L) to 86% (lineage A2) of the bootstraps, which suggests the presence of thermal optima in population growth rates. When looking at all pairwise combinations of lineage L and each other lineage, in 55% of the comparisons both lineages showed a temperature optimum within the tested temperature range. Within this subset, the temperature optimum of lineage L was higher than that of A1, A2, A3 and A4, in 78%, 89%, 79% and 90% of the cases.
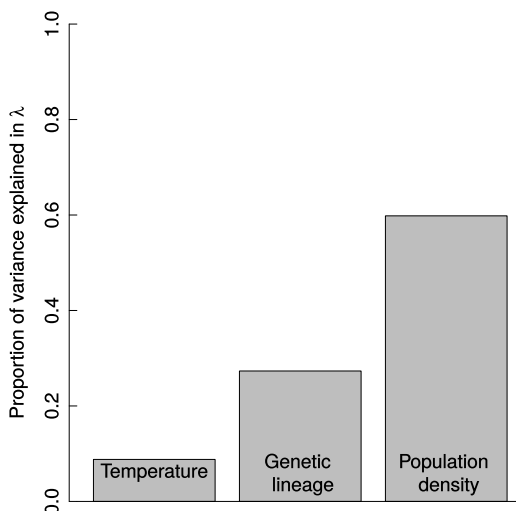
Figure 3.3: Variance explained in λ across all observed lineages, temperatures and densities. Results are based on linear regressions with one explanatory variable at a time.

All lineages show a decline in λ with increasing density, with lineage A3 suffering the most from increasing densities (Fig. 3.2b). Ordering the lineages by the density at which λ dropped below 1, matched to a large extent with the order of observed maximum population sizes (Fig. 3.2b). This resulted in the lowest equilibrium population sizes for lineage A3 and A1. Using the bootstrapped datasets, we evaluated uncertainty in population-level density effects. Population growth rates decreased with increasing density in 99% of the bootstraps, providing strong evidence for negative density dependence (Fig. S17). In 91% of the bootstraps, the carrying capacity was within the range of observed population sizes. Evaluating all pairwise combinations, linage A3 showed the lowest carrying capacity in 96% of all pairwise comparisons, while lineage L had the highest carrying capacity in 88% of all pairwise comparisons.

Using all observed population sizes during the experiment to project population growth rates, variation in density explained 60% of the variation in λ (Fig. 3.3), followed by lineage (27%). Variation in temperature explained only 9% of the variation in λ, indicating that thermal responses in individual vital rates cancelled each other out to a considerable degree when integrated at the level of the whole population.

Differences between lineages

Starting with an IPM which was averaged over the five lineages, and systematically replacing one of the vital rates with each of the lineage-specific vital rates, we evaluated population-level consequences of lineage differences. Different lineages realized higher population growth rates through different vital rates and none of the lineages had a demographic advantage in all vital rates (Fig. 3.4). Results were similar when performing this analysis without lineage L (thereby changing the reference population growth rate; Appendix B.12).

Most notable, lineage A2 benefitted from above-average survival over the complete range of temperatures (Fig. 3.4a; Fig. S18 in Appendix B.11), but this benefit was offset by a great disadvantage in the probability of carrying eggs, especially at the high temperatures (Fig. 3.4c-e; Figs S20-22). Lineage L, in contrast, benefitted most from the vital rates related to reproduction (Fig. 3.4c-e; Figs S20-22), and had a slight disadvantage in individual growth. Finally, the small

differences in offspring size (Table 3.1) hardly contributed to differences in population growth rates among lineages (Fig. 3.4f; Fig. S23).
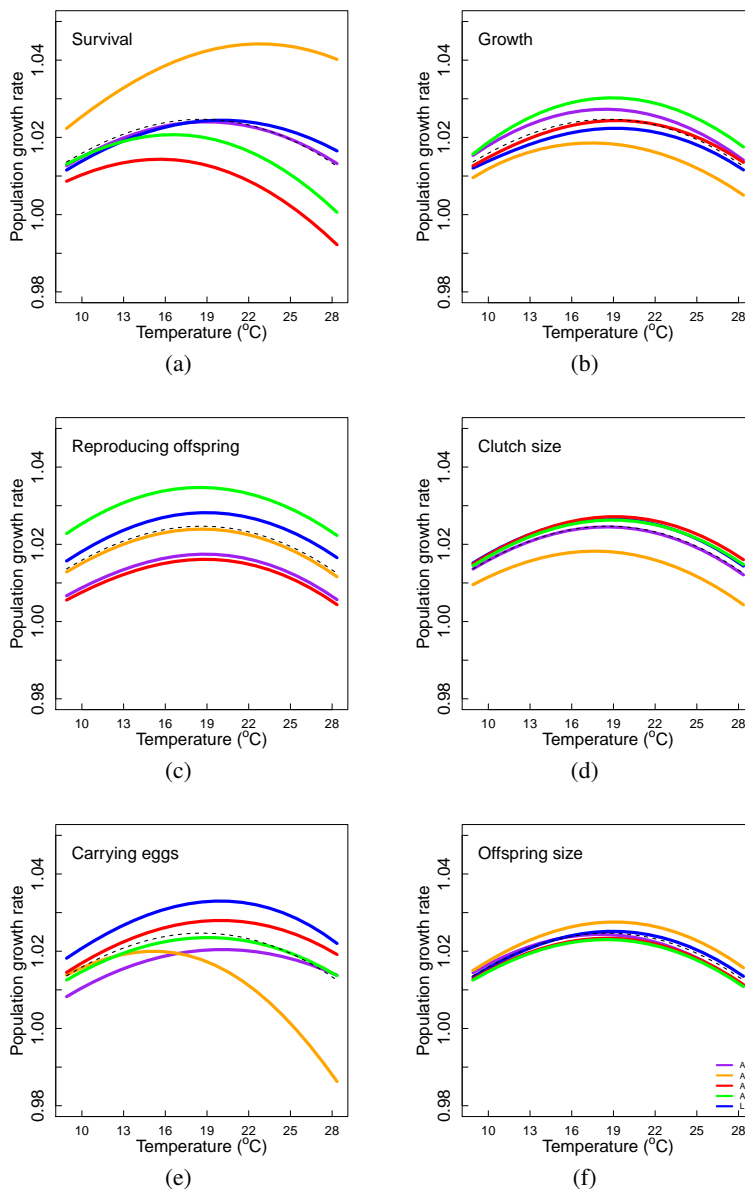
Figure 3.4: Effects on population growth rates (λ) due to differences between genetic lineages in each vital rate, compared to λ of the temperature-specific IPM averaged across lineages (dotted lines). Different colours indicate different genetic lineages. Effects were calculated by replacing one vital rate function by the corresponding vital rate function of a specific lineages, and recalculate λ across the temperature range.
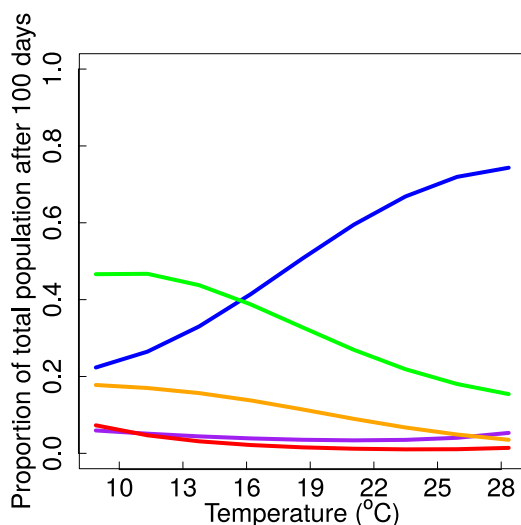
Figure 3.5: 100-day density-dependent simulation of a hypothetical mixed population, containing all five lineages. Each lineage started at day 0 with four individuals. Graph shows proportion of each lineage at day 100 over the range of temperatures (A1: purple, A2: orange, A3: red, A4: green, L: blue).

Density-dependent simulation

A density-dependent simulation revealed that, after 100 days, relative abundance of the lineages changed with temperature. Overall, lineage L became more abundant at higher temperatures (Fig. 3.5), whereas at the coldest temperatures, lineage A4 was most abundant. Lineage A2 and A4 showed the strongest decrease with increasing temperature. The proportions of lineage A1 and A3 were small and did not change much with temperature. When doing this analysis for lineage A1-A4, lineage A4 was most abundant across all temperatures. Lineage A1 and A3 outcompeted lineage A2 only at the highest temperatures (Appendix B.12).

3.5   Discussion

The success of a population is directly determined by the performance of its individuals. To get a more mechanistic insight in the extent to which populations can adapt towards changing environments, it is important to understand the environmental effects on separate vital rates and their integrated effect on population dynamics (Pelletier et al. 2007; Pelletier et al. 2012). In addition to environmental effects, the performance of individuals is simultaneously affected by many other factors, such as genotype and body size (Ozgul et al. 2010; Coulson et al. 2011; Pelletier et al. 2012; Brooks et al. 2016). Moreover, individuals may suffer from negative density-dependent processes, such as competition for food, release of chemical substances or due to physical contact, as has been shown in *Daphnia* (Goser and Ratte 1994), potentially manifesting in all life history traits.

We have studied the importance of all of the above factors, influencing population dynamics of *Daphnia magna*, via their effects on single vital rates. We have shown that *Daphnia magna* individuals embedded in populations were able to respond plastically to higher temperatures, by accelerating their life cycle, reflected by increased growth and earlier maturation. Clonal lineages showed differences in growth, survival and reproduction, and, at a population-level, responded differently to temperature (Fig. 3.2). Results indicate trade-offs between growth, survival and

reproduction, as no lineage performed the best in all vital rates (Fig. 3.4). Our study stresses the importance of studying individuals within a population. First, following only population trends does not give information on how individuals respond and how vital rates contribute to the overall trend. Second, studying individuals without a population setting ignores density-dependent effects, which we found to have large, but variable effects (Figs 3.2, 3.3).

The structure and complexity of a population model can potentially influence results (Salguero-Gómez and Plotkin 2010; Jongejans et al. 2011). To test the robustness of our results and ensure that our results were not driven by particular choices made to parameterize the IPM, we compared four different approaches, differing most notably in how reproduction was incorporated. As shown in Appendix B.5, these choices resulted in similar model outcomes, and hence did not affect our conclusions.

Effects on single vital rates

Given that body size had a significant effect on all vital rates (β_1, β_4 in Table 3.1), understanding thermal responses requires understanding how temperature affects body size and emphasizes the importance of a trait-based approach when investigating population dynamics (Ozgul et al. 2012; Ronget et al. 2017). For all vital rates except clutch size, the best model included a temperature effect, and in 4 out of 6 vital rates, this effect was significant. Individuals reared at higher temperatures became mature at a smaller size. Moreover, individuals initially grew faster, but this effect reversed at larger sizes. These plastic responses on growth and maturation have previously been described for *Daphnia* (Mitchell and Lampert 2000; Van Doorslaer et al. 2009a; Henning-Lucass et al. 2016), and are believed to be important in generating the temperature-size rule, which is followed by the majority of ectotherms (Atkinson 1994; Atkinson 1995). Survival was negatively affected by temperature, in agreement with previous studies (MacArthur and Baillie 1929; but see Henning-Lucass et al. 2016).

We conclude that the found temperature effects on single vital rates are mostly in line with results from life table experiments, in which individuals are followed throughout their life. In our study, instead, we observed transitions over three or four days at a time. The big advantage of our setup is that we were able to simultaneously quantify density effects. Somatic growth and reproductive output decreased with increasing densities, in agreement with previous work (Frank et al. 1957; Guisande 1993; Goser and Ratte 1994), and these density effects were often in the same order of magnitude as the temperature effects (compare β_2 and β_3). Survival probabilities showed a positive correlation with density, which seems surprising at first sight, but this was also found for individuals in different developmental stages of soil mites (Ozgul et al. 2012). We suspect a reverse causality for this correlation: populations in which individuals survive better, reach higher densities. We therefore decided to drop density in the survival models (as was also done in Traill et al. 2014, because of similar findings in bighorn sheep). Future studies could disentangle these relationships between density and survival by manipulating densities to remain constant at different levels (unlike the dynamic populations that were the focus here) or perhaps by using flow-through systems (Gliwicz 1990; Giebelhausen and Lampert 2001).

Integrating vital rates to predict population-level consequences

Integrating over all demographic rates enabled a quantification of the net result of the temperature effects on growth, survival and reproduction. Using asymptotic population growth rate as a proxy for population fitness (Metcalf and Pavard 2007), we have shown that individuals were largely able to compensate for the increased mortality over the range of tested temperatures. Although there were only weak indications of temperature optima per vital rate, all lineages showed a thermal optimum when integrating all vital rates, ranging between 16.6 °C and 22.5 °C.

Density had equally large effects on λ as temperature, clearly resulting in a carrying capacity for all lineages, and explained considerably more of the variation in daily growth rates than temperature. Furthermore, density-dependent effects differed in strength among the lineages. This, together with the stochastic component of population dynamics, would complicate direct usage of performance of isolated individuals for parameterizing population models, as was done for instance by Sommer et al. (2016). Future research will have to show under what circumstances and with which assumptions the wealth of life table experiments (that efficiently study the effects of various environmental factors on performance in isolation) can be properly scaled up to population dynamics.

Population models for zooplankton populations are rare, because it is difficult to collect demographic data on individuals embedded in the population (Jiménez-Melero et al. 2013), as individuals cannot easily be marked or recognized. To overcome this, earlier studies have used 'inverse' methods to estimate vital rates using data on population abundances and structure over time (Jiménez-Melero et al. 2013). This is, however, a complex problem as many combinations of vital rates can result in the same dynamics (Wood 1994). Temporarily caging individuals in tubes enabled us to collect individual data to parameterize population models. Several factors, however, could in theory have led to differences in individual performance inside and outside the tubes. First, isolated individuals experienced a larger volume-per-individual medium, especially at higher densities, which could have led to differences in food access (despite the use of permeable filters). Second, they did not experience physical contact with other individuals, and, finally, they had less freedom to move. To what extent these factors may have influenced the isolated individuals, remains to be investigated. However, individuals were isolated for only three or four days, which is short compared to their lifespan. Moreover, the observed density effects were both convincing and remarkable; vital rates were estimated using observations on the isolated individuals, while density estimates were obtained from completely independent video analysis on the entire population. Finally, ordering the lineages by their predicted carrying capacity, based on the IPMs, matched almost perfectly with the maximum observed population sizes (Fig. 3.2). This gives confidence in our methods and results.

Interclonal differences and life history strategies

Individuals hatched from sexual dormant eggs differ in their genetic makeup, which can lead to differences in performance. We found clear lineage effects on vital rates, propagating to the population-level. In general, lineage L showed highest population fitness over the range of tested temperatures, in particular for the higher temperatures. The overall best performance of lineage L is perhaps not surprising, as this lineage has been successfully held in the lab for more than ten years (Lürling and Tolman 2010), thereby having proved its ability to perform well under laboratory settings. In contrast, the other four lineages (A1-A4) were used only one year after hatching from dormant eggs collected in the field. Model species such as *Daphnia magna* play a central role in climate change research. However, our results warn against extrapolating conclusions based on lineages that have been raised and selected to do well under laboratory conditions, as these may not be representative for natural populations.

Within-population variation in vital rates and population-level responses have been demonstrated in *Daphnia* (Carvalho 1987; Van Doorslaer et al. 2009a; Jansen et al. 2011; Pietrzak 2011; Pantel et al. 2015; Stoks et al. 2016), and is in agreement with our results. When comparing lineage A1-A4, the differential life history strategies suggest that there is substantial within-population variation for natural selection to act upon (see also Van Doorslaer et al. 2009a; Van Doorslaer et al. 2010). Differentiation in thermal tolerance among lineages (Mitchell and Lampert 2000) may indicate the presence of seasonal clones, which are adapted to specific periods of the growing

season (Carvalho 1987). As no lineage outcompeted other lineages in all vital rates, our results suggest trade-offs between survival, growth and reproduction, in accordance with Dudycha and Tessier (1999) and Reznick et al. (2000). Interestingly, we have shown that these trade-offs between multiple vital rates were meaningful in a population context, and that different clonal lineages maximized their fitness via different routes (e.g. De Meester et al. 1995).

Evolutionary potential for thermal adaptation

We used Integral Projection Models to make testable predictions on evolutionary change and predicted that, after only 100 days, the genetic composition of a mixed population can be substantially altered due to interclonal fitness differences, and that these changes are temperature-dependent. These predictions, based on the performance of single lineages, are yet to be tested. This could be done by following populations consisting of two or more lineages, while following the abundance of each lineage, for instance using genetic markers (Turcotte et al. 2011b). Discrepancies between predicted and observed evolutionary changes could indicate that the presence of other genotypes has differential effects on the performance of individuals, compared to density effects of individuals of the same lineage.

The role of evolution in short-term adaptive trait changes, is recently receiving much attention (Pelletier et al. 2009; Ellner et al. 2011; Schoener 2011; Benthem et al. 2017) and rapid evolutionary responses have now been shown in many experimental systems (e.g. Becks et al. 2012; Cameron et al. 2013; Turcotte et al. 2013). In *Daphnia*, thermal evolution experiments have shown rapid adaptive responses towards changing temperatures (Van Doorslaer et al. 2009a; Van Doorslaer et al. 2010; Geerts et al. 2015), and this evolutionary potential for thermal adaptation is supported by our results. Our work differs from earlier studies in that we quantified thermal responses by integrating all fitness components to obtain estimates of population growth rate as a proxy for fitness, which is the metric that is optimized by natural selection. From a population adaptation perspective, this gives a more complete picture compared to studies documenting thermal responses in single fitness components such as somatic growth (Mitchell and Lampert 2000) or maximum thermal tolerance (CT$_{max}$; Geerts et al. 2015). Moreover, when obtaining these fitness components from single individuals, information on the interplay between density, the environment and genetic background, and integrated effects of all vital rates, is lacking. On the other hand, previous studies on eco-evolutionary dynamics that did focus on populations trends (Van Doorslaer et al. 2010; Turcotte et al. 2011b; Turcotte et al. 2013), generally lack information on performance of individuals within the population.

As shown here, using individual demographic rates, as obtained from individuals embedded in a population, and integrating these into a population model, gives a more mechanistic understanding of population-level responses towards climate change. Moreover, density had equally large effects as temperature on vital rates, and explained more of the variation in daily growth rates. We conclude that these meaningful density effects cannot safely be ignored when predicting population responses to environmental change. Future studies applying our approach to different clones or species, e.g. from different locations (Yampolsky et al. 2013) or periods (Geerts et al. 2015), will greatly improve the understanding of evolutionary potential for thermal adaptation. It will also help to identify life stages whose expected demographic responses to future environmental change contribute most to changes in population fitness. These life stages could be the most promising targets for conservation strategies. Natural populations facing climatic changes are not purely affected by the changes in temperature, but also by associated changes in for instance densities, food availability and pathogen dynamics. Scaling our technique of quantifying demography on temporarily partially-isolated individuals up to mesocosm or field studies will provide novel insights into eco-evolutionary responses to climate change in a more natural setting. Together

with other types of 'evo-demo' studies (Ronget et al. 2017), our approach should lead to enhanced understanding of how much resilience we can expect due to phenotypic plasticity and rapid evolution on relevant short time scales, when assessing the vulnerability of animal communities to the effects of global change.

Acknowledgements

Marjolein Bruijning[1], *Erlend Ignacio Fleck Fossen*[2], *Eelke Jongejans*[1], *Héléne Vanvelk*[3], *Joost Raeymaekers*[4], *Lynn Govaert*[3,5,6], *Kristien Brans*[3], *Sigurd Einum*[2] and *Luc De Meester*[3]

1. Department of Animal Ecology and Physiology, Radboud University, 6500 GL Nijmegen, The Netherlands.
2. Centre for Biodiversity Dynamics, Department of Biology, NTNU, Norwegian University of Science and Technology, Trondheim, Norway.
3. Laboratory of Aquatic Ecology, Evolution and Conservation, KU Leuven, Leuven, Belgium.
4. Faculty of Biosciences and Aquaculture, Nord University, N-8049, Bodø, Norway.
5. Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland.
6. Department of Aquatic Ecology, Eawag, Swiss Federal Institute of Aquatic Science and Technology, Uberlandstrasse 133, CH-8600 Dübendorf, Switzerland.
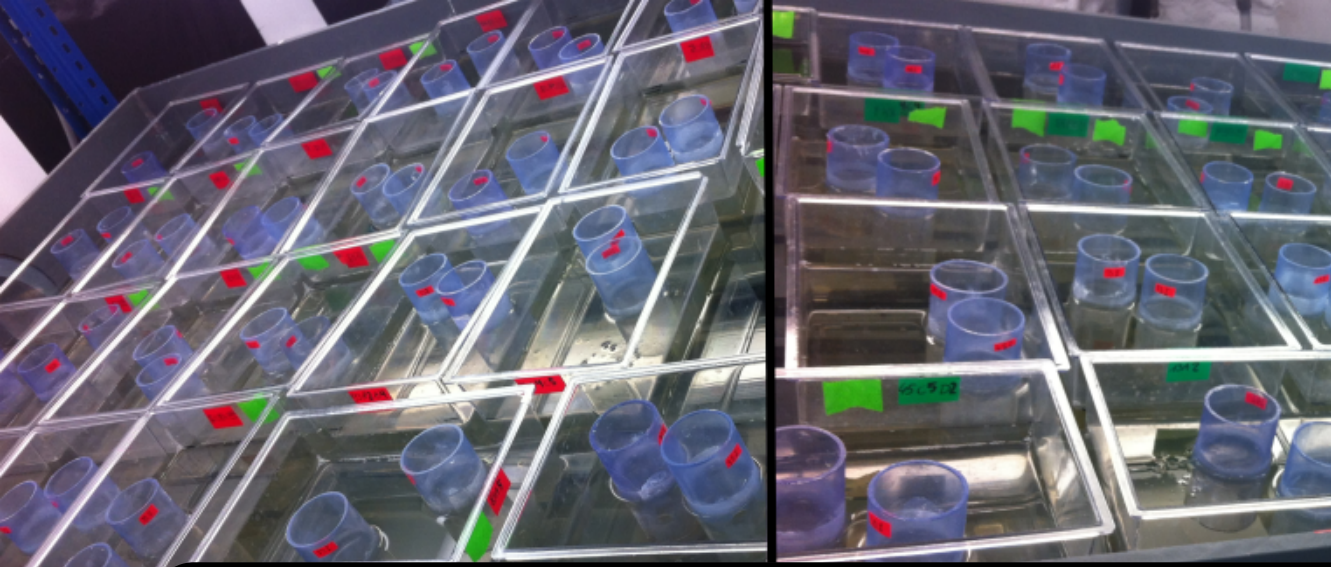
Host-parasite dynamics in *Daphnia magna* populations shaped by temperature and host genotype

4.1 Abstract

1. Global warming is predicted to impose strong selection pressures on natural populations. Moreover, local populations will be increasingly likely to encounter conspecific immigrants that are preadapted to higher temperatures, which can further change the local gene pool. The fate of natural populations additionally depends on the capability of local individuals to cope with temperature-associated changes in the environment, such as changes in parasite dynamics, potentially carried along with immigrants.

2. To look at these interactive effects of temperature and parasite dynamics, we exposed *Daphnia magna* genotypes originating from two Norwegian and two Belgian ponds to a laboratory experiment in Belgium. We followed populations consisting of single genotypes and of combinations of two genotypes, and exposed populations to temperature treatments ranging between 14 °C and 26 °C. During 81 days, we measured survival, reproduction and somatic growth on individuals within the population. Halfway the experiment, all populations became infected with the ectoparasite *Amoebidium parasiticum* leading to large reductions in population densities.

3. We show that parasite load had strong negative effects on individual survival, reproduction and growth. Temperature had negative effects on survival, but positive effects on reproduction components. We integrated these fitness components using an Integral Projection Model (IPM). Results of this IPM suggest large differences in equilibrium densities between genotypes originating from the four ponds, with the highest densities reached by the Belgian genotypes.

4. Parasite infection probabilities and temperature-dependent probabilities of transitioning between parasite load classes were estimated and used to create a 2-state host-parasite IPM. Using simulations, we show that infected *Daphnia* host populations perform better at higher

temperatures, due to higher host resistance and tolerance against infection. Moreover, we show that the Belgian genotypes have a higher parasite tolerance, and that this is mainly due to benefits related to survival. These results are supported by the results of the competition experiments, where we showed, using genetic analyses, that the Belgian genotypes largely outcompeted the Norwegian genotypes at all temperature treatments.

5. Our results show that the dynamics of *Daphnia* populations were severely affected by a parasite infection, through effects in each fitness component. Moreover, responses differed greatly among genotypes, and we showed clear between-population variation in parasite tolerance. Finally, as both host and parasite dynamics were influenced by temperature, we conclude that climate change can have a large impact on local gene pools, not only through changes in temperature, but also through associated changes in parasite dynamics.

4.2 Introduction

Climate change is predicted to impose strong selection pressures on natural populations. Changes in temperature can strongly affect ectotherms, as their physiological functions depend directly on temperature (Huey and Berrigan 2001). The fate of local populations will depend on the degree to which they are able to adapt to increasing temperatures through standing genetic variation (resulting in an evolutionary response), or to acclimate by means of individual phenotypic plasticity (Gienapp et al. 2008; Hoffmann and Sgrò 2011). In addition, local populations will be increasingly likely to encounter conspecific immigrants that are migrating northwards, following shifts in thermal isolines (Thomas and Lennon 1999; Hickling et al. 2006). These conspecific immigrants have the potential to outcompete local populations when they are preadapted to the new local conditions, leading to a change in the local genetic pool (Van Doorslaer et al. 2009b). Finally, local populations facing climate change do not only have to cope with changes in temperature and competition from immigrants, but also with other temperature-associated changes in the abiotic and biotic environment. These include for instance changes in dissolved oxygen and acidity (in aquatic systems), changes in community structure, and altered pathogen dynamics (Ficke et al. 2007; Hoffmann and Sgrò 2011). In this study, we focus on the latter.

The water flea *Daphnia magna*, a small crustacean, is a keystone species in freshwater systems and an excellent model species to study the joint effects of temperature, competition and pathogen exposure. Both within-population genetic variation in thermal tolerance has been shown (Van Doorslaer et al. 2009a; Bruijning et al. 2018b) enabling rapid evolutionary responses to temperature, as well as between-population variation, resulting in local adaptation to temperature (Van Doorslaer et al. 2009b; Yampolsky et al. 2013; Geerts et al. 2015). In addition, natural *Daphnia* populations often show heavy parasite infections (Decaestecker et al. 2004) and responses to parasites can differ greatly among genotypes (Carius et al. 2001). Coevolution of the host *Daphnia* and pathogens (Ebert 1994; Decaestecker et al. 2007) can lead to differences in pathogen responses among host populations, which may be relevant when climate migrants carry along and introduce pathogens to local populations.

To assess the performance of local and immigrant populations and to predict the consequences when exposed to each other, it is important to know how individual life history traits (i.e. stage-specific growth, survival and reproduction) respond to changing temperature and pathogen conditions, and how life history traits together determine population fitness. Not all fitness components have equally large effects on population fitness (Caswell 1978), and moreover, different fitness components can co-vary positively or negatively (for instance through trade-offs; Stearns 1989). Therefore, one life history trait is not necessarily a good proxy for fitness, making it important to integrate over all fitness components (or: vital rates) in order to obtain a measure for

population fitness (Metcalf and Pavard 2007). Integral Projection Models (IPMs) are a useful tool for integrating over multiple vital rates, and describe changes in a continuous trait distribution (often body size), due to trait-dependent survival, growth and reproduction (Ellner and Rees 2006; Merow et al. 2014). Recently, an extension of IPMs has been proposed to describe host-parasite dynamics, by combining IPMs with SIR (or SIS) models (Metcalf et al. 2016; Wilber et al. 2016). SIS models consist of a set of differential equations to describe the number of Susceptible and Infected hosts in a population through time, based on transition probabilities between these compartments (i.e. transmission, recovery and mortality rates). SIS models differ from SIR models in that there is no Recovery state (i.e. individuals become susceptible again directly after recovering). Such epidemiological models can be combined with IPMs in a quite straightforward manner, by including a second (discrete) state variable describing parasite load (Metcalf et al. 2016; Wilber et al. 2016). A combined SIS-IPM allows for projecting parasite as well as host dynamics through time, and a quantification of the contribution of processes leading to host variation in both resistance and tolerance. Resistance is defined as the ability of the host to limit parasite burden (Råberg et al. 2007), and on a population-level, higher resistance leads to a reduction in average parasite loads. Tolerance is the ability of the host to limit the negative effects of the parasite burden (Råberg et al. 2007). Thus, given a parasite load, the host suffers less from it.

In this study, we focus on between-population differences in performance in response to temperature and parasite infection. We used *Daphnia magna* as a study species and compare clones originating from four ponds, located at two different latitudes, to look at the importance of local adaptation to temperature and infection. In a laboratory experiment, we followed populations of *D. magna* during 81 days, while collecting individual life history data from individuals embedded in the population, enabling the estimation of density-dependency in these rates (Bruijning et al. 2018b). Because we were specifically interested in the performance when clones from different locations were put together, reflecting a scenario in which local individuals face immigrants, we exposed clones to a competition treatment and used genetic markers to follow genotype frequencies through time. Infection with the ectoparasite *Amoebidium parasiticum* in all experimental units halfway the experiment (see *Methods*) gave us the opportunity to evaluate clone and temperature-specific tolerance and resistance against infection. Ectoparasites, such as *A. parasiticum*, attach to the exoskeleton of *Daphnia* and can have negative effects on performance when abundant (Chiavelli et al. 1993; Threlkeld and Willey 1993).We specifically aim to answer the following questions: 1) What is the effect of *A. parasiticum* infection on different vital rates of the host *D. magna*, and how does this host tolerance vary between clones originating from different locations? 2) How is host resistance, i.e. the ability to limit parasite burden, affected by temperature? 3) What is the combined effect of infection and temperature on host fitness, and finally, 4) what is causing differences in host-parasite dynamics between clones from the four locations, and between temperatures?

4.3 Materials and Methods

Data collection

Selection of *D. magna* clones

We used clones originating from four fishless ponds. First, these were the Asklund and Ringve pond, both located nearby Trondheim (Norway), which we refer to as location of origin N1 and N2. The distance between these ponds is ∼20 km and they are separated by the Trondheim fjord. Second, these were ponds in Zonhoven and Lapscheure (Belgium), which we refer to as locations B1 and B2 (distance: ∼145 km). From each of these four ponds, resting eggs were collected, hatched and raised in the lab to establish different clonal lineages. From the successfully

established lineages, three were randomly selected from each location of origin, resulting in a total of twelve clones ($N1_{A-C}$, $N2_{A-C}$, $B1_{A-C}$, $B2_{A-C}$) that were used for the experiment.

Experimental design

The experiment was performed between August and November 2017, at the KU Leuven (Belgium). Four experimental temperatures were used: 14, 18, 22 and 26 °C, which were controlled using a water bath system. We followed dynamics of populations that consisted of single clones and populations that consisted of two clones. Every clone was exposed to each of the four temperatures, resulting in $12 \times 4 = 48$ single clone populations. For the mixed populations, we created each pairwise combination once, which results in 66 unique combinations. These 66 combinations were randomly distributed over the three highest temperatures (18, 22 and 26 °C), ensuring that at each of these temperatures, we obtained 4 within-population combinations (e.g. $N1_A$-$N1_B$), 6 within-latitude combinations (e.g. $N1_A$-$N2_A$), and 12 between-latitude combinations (e.g. $N1_A$-$B1_A$). Moreover, we ensured that each clone was equally represented at each temperature. In total, we followed 48+66=114 populations.

Prior to the experiment, individuals of each clone were reared in five 500 mL jars at 20 °C for four generations, whereby first clutches were discarded. At the start of the experiment, we started with 12 neonates (again excluding first clutches) that were 0-24 hours old, for each population (either 12 neonates from the same clone, or 6 neonates from each of the two clones in case of the mixed populations). We randomly picked neonates from different jars, to minimize potential effects of clutch and jar. For logistic reasons, the experiment was started over the course of four days (58%, 29%, 11% and 3% of the populations were started at day 0-3, respectively). See Appendix C.1 for a timeline of the experiment.

During the first week of the experiment, individuals were kept in 500 mL jars. After that, they were transferred to 1.5 L aquaria. In each aquarium, we placed two transparent tubes (4 cm diameter), containing 12 holes (1 cm diameter) covered with a permeable filter allowing medium to pass, but preventing individuals (including neonates) to pass. These tubes enabled us to isolate individuals for a few days to measure the performance of individuals, while they were still part of the population (see below, and Bruijning et al. 2018b). All aquaria (or jars, in the first week) were placed at the relevant temperature, at a randomized location in the water bath. During the experiment, we randomly reorganized all aquaria three times. Populations were kept in dechlorinated tap water and fed daily (0.1 mL Shellfish Diet 1800 L^{-1} medium). Twice per week, when performing the measurements (see Section *Population and individual measurements*), the medium was refreshed. Aquaria and tubes were rinsed with hot tap water, while keeping 0.5 L of the old medium. The experiment lasted 81 days (ranging between 77 and 81 days per population, depending on the start day).

Population and individual measurements

All aquaria were randomly divided into two groups. For group 1, measurements were performed on Monday and Thursday; for group 2, measurements were performed on Tuesday and Friday. Each time individuals were transferred to a glass oven dish using a sieve, and dead individuals and resting eggs were removed from the population using a pipet. The oven dish was placed in a fixed camera set up (details on the setup can be found in Czypionka et al. 2016) to make a short movie of the population. These movies were used to obtain population size estimates, using the R-package *trackdem* (Bruijning et al. 2018c). Individuals were then transferred back to the (rinsed) aquarium, containing fresh medium (see above).

Concurrently, the isolated individuals were taken out of the tubes. We noted whether the individual was alive, its size (measured using a stereomicroscope), the number of carried eggs and the number of produced neonates (if any). If neonates were produced, we additionally measured

the size of one of the neonates. The individuals and neonates that were in the tube were either put back in the population or sampled (see Section *Genetic analyses*). Two new individuals were arbitrarily picked from the population (generally one juvenile and one adult). We measured their size, counted the number of eggs and placed those in the tubes. Finally, the aquarium was placed back in the water bath. Per measuring day, 57 populations were handled in this way. After 81 days, this resulted in 4651 individual measurements on survival, growth and reproduction rates and 2389 population size estimates.

Parasite load

After an initial increase in population size in all aquaria, all populations strongly declined in size during the second half of the experiment (See *Results*). On day 60, we realized that all cultures were infected by the ectoparasite *Amoebidium parasiticum*. Infection with this parasite had occurred before in the laboratory facilities, and is hard to prevent under standard rearing conditions. We therefore anticipated that infection could occur and spread rapidly among the experimental clones. From day 60 onwards, we scored the parasite load for each measured individual, on a 0-3 scale (0: no signs of infection, 3: heavily infected). This was done both when first measuring the individual (putting it in the tube) and when remeasuring it after three or four days. This resulted in 1167 individual measurements on parasite load, as well as transitions between parasite load classes after three or four days. To reconstruct the dynamics of the infection before day 60, we scored parasite loads afterwards for as subset of individuals that were sampled before day 60 (see *Genetic analyses*). In this procedure, we scored individuals going back from day 60 until day 38, which we estimated to be just before the infection became widespread. For those populations in which at least one of the sampled individuals was infected at day 38 (which was the case in 10% of the populations), we continued to score individuals from earlier days, until only uninfected individuals were observed. The first infected individual was observed at day 29. This resulted in an additional 654 measurement on individual parasite loads. Finally, we set parasite load for all other (i.e. unchecked) individuals from before day 38 at zero.

We acknowledge that there are some shortcomings of the collected data on parasite dynamics. To begin with, we could not control the onset of infection, nor differences in exposure among aquariums. However, *A. parasiticum* certainly infected each population, and the observed declines in population sizes suggest that this occurred around the same time. We started recording individual infection levels late in the experiment, when populations were already heavily infected, resulting in limited individual data on transition probabilities between parasite load classes. However, because we were able to reconstruct the start of infection based on measurements on sampled individuals and due to the clear patterns, we were still able to model temperature-dependent parasite dynamics (as explained below, see also Fig. 4.3).

Genetic analyses

To identify genotypes in the competition treatments, we sampled individuals during the experiment of which we obtained demographic data on (i.e. those individuals that were temporarily placed in the tubes). This sampling was started at day 17, when the initial cohort had produced their first clutches across all temperature treatments. After remeasuring an individual, it was put in an Eppendorf tube containing ethanol and stored at 4 °C. Note that this was also done for the single clone populations (even though these did not need to be identified). Whenever a neonate was present, the neonate was sampled instead of the mother, to reduce the impact on the population. In case of a dead individual, it was sampled when not too much decayed. Whenever population size was smaller than ~50 individuals, we did not sample living individuals, except in cases when the measured individual released offspring in the tube. This resulted in 1676 sampled individuals from those competition treatments that contained between-population combinations (e.g. $N1_A$-$N2_A$),

from which 108 individuals were dead and 1568 individuals were alive when sampled. Additionally, it resulted in 1587 sampled individuals from either the competition treatments containing within-population comparisons (e.g. $N1_A$-$N1_B$) or single clone populations. We here only focus on the between-population comparisons, and we used a total of nine microsatellite markers to identify clones from these populations. From these 1676 individuals, DNA was extracted NucleoSpin 96 Tissue Core Kit. Results were manually scored using software GeneMarker. We were able to successfully identify 57 out of 108 individuals that were dead when sampled, and 1420 out of the 1568 individuals that were alive when sampled.

Modelling framework
Vital rate fitting
The individual measurements were used to fit vital rate models. We defined six vital rates: survival, somatic growth, probability of carrying eggs, probability of producing offspring (conditional on carrying eggs), clutch size and offspring size. More details on each of the vital rates are given below. Per vital rate, we fitted generalized linear models including additive effects of body size (z), squared body size (z^2), temperature (T), population density (N), parasite load (P), location of origin (L), and experienced between-location competition (C). Variables z, T and N were included as continuous variables and were Z-score transformed prior to the analyses. Parasite load was included as a continuous variable; L was included as a factor with four levels ($N1$, $N2$, $B1$, $B2$), whereby we specified contrasts comparing location $N1$ and $N2$ with location $B1$ an $B2$, and comparing $N1$ with $N2$, and $B1$ with $B2$, to test the hypothesis that rates were affected by latitude, in addition to estimating differences between locations. Between-location competition was included as a continuous variable, ranging between 0 and 1, indicating the (time-dependent) proportion of the population consisting of another genotype as a measure for the experienced competition against different genotypes. For the single clone treatments and for the within-population mixes (e.g. $N1_A$-$N1_B$), C equaled 0. For the other mixed populations, we predicted genotype frequencies based on a logistic regression fitted per population, as a function of day (Appendix C.2). Experienced competition C at time t for genotype G was then calculated as: $C(t, G) = 1 - p(t, G)$, where $1 - p(t, G)$ is the predicted proportion of genotype G at day t. Note that we assume that genotypes experienced no extra competition (in addition to potential effects of density) from genotypes originating from the same location, and equal competition strength by different genotypes from a different location from the same latitude as from genotypes from locations from a different latitude.

In addition, to test for differential responses across locations, we included all two-way interactions between location and z, T, N, P, and C. Finally, we included an interaction between temperature and parasite load, to allow for different responses to infection at different temperatures. The most complex model contained 26 parameters. We used model selection to select the most parsimonious model per vital rate. Exploratory analyses indicated that additive effects of z, P, L and T resulted in model improvement across all vital rates, so we used models including these four predictors as a basis. We then tested all models nested within the most complex model, resulting in 288 models. These were fitted to each of ten imputed data sets (see Appendix C.3). We continued with the model that resulted in the lowest ΔAIC on average, across the ten imputed data sets (Appendix C.4). Finally, we confirmed that residuals were unbiased across all body sizes, temperatures and parasite loads, in particular to check that both temperature and parasite load could appropriately be included as linear continuous variables, despite only four levels for both variables (results not shown).

Observations were on a three or four-day basis. In order to obtain daily vital rates, duration had to be taken into account when fitting the vital rate models. The appropriate way of doing this depended on the vital rate and is explained below. We use θ to refer to the parameter set including

all coefficients describing effects of z, z^2, T, N, P, C, L and relevant interactions, for a certain vital rate. Each vital rate could be a function of z, T, N, P, C, L (depending on the selected model, see above), but to improve readability, we omit those in notations, except for body-size dependence (see below). The procedure per vital rate, using generalized linear models, was as follows:

1. Survival probabilities $S(z, \theta)$. We optimized the log-likelihood function comparing observed survival (n=4594) with predicted survival probabilities. To do so, we converted daily predictions (based on the generalized linear model) to predictions after one time interval (Δt, three or four days), by raising those to the power of Δt. A logit link function was used to relate the linear predictor to the response scale. Exploratory analyses showed a positive relation between population size N and survival probability. We decided to leave out density in the survival models, and address this relation between density and survival in the discussion.

2. Growth was calculated as the size difference between the two measurements (n=4016). To obtain daily growth, we divided the size difference with Δt, and used an identity link function to fit daily growth $g(z, \theta)$. Variation in growth σ_g was calculated as the residual standard deviation.

3. To fit the probability of carrying eggs $p_e(z, \theta)$, we used a logistic regression (0: no eggs, 1: at least 1 egg; n=4334). Here we only used observations when first measuring individuals (i.e. when placing them in the tube).

4. Probability of having produced offspring $p_r(z, \theta)$ by the time of the second measurement was fitted conditional on carrying eggs (i.e. only including those individuals that carried at least 1 egg when first measuring; n=1000). We took into account that observed offspring could have been born up to Δt days earlier than the second measurement. This was done by calculating the cumulative probability of producing offspring after days, given a predicted daily probability. We optimized the log-likelihood function comparing these cumulative probabilities with the observations, using a logit link function.

5. Clutch size $F(z, \theta)$ was fitted conditional on producing offspring, using a log link function (n=879).

6. Offspring size $\phi(z, \theta)$ was fitted using an identity link function (n=879). Variation in offspring size σ_ϕ was calculated as the residual standard deviation.

Construction of Integral Projection Models

We used Integral Projection Models (IPMs) to integrate over the six vital rates. IPMs describe the dynamics of a population in which individuals are characterized by a continuous state variable (body size z in our case), in discrete time, and typically consists of two kernels (Ellner and Rees 2006; Ellner et al. 2016). In addition to body size, we created an IPM which was a function of temperature, population size, temperature, parasite load, and population origin, but have omitted these in the equations for convenience. The IPM is described as:

$$n(t+1, z') = \int [S(z, \theta) \cdot G(z', z, \theta) + R(z, \theta) \cdot D(z', z, \theta)] n(t, z) dz \qquad (4.1)$$

Eq. 4.1 describes the body size distribution at day $t+1$ ($n(t+1, z')$), given the body size distribution at day t ($n(t, z)$), for parameter set θ. Daily survival probability is given by $S(z, \theta)$. The reproduction function $R(z, \theta)$ is defined as the product of three vital rates (probability of carrying eggs, probability of producing offspring and clutch size):

$$R(z, \theta) = p_e(z, \theta) \cdot p_r(z, \theta) \cdot F(z, \theta) \qquad (4.2)$$

Finally, $G(z',z,\theta)$ and $D(z',z,\theta)$ are probability density functions, describing size distributions at $t+1$ of growing survivors and of neonates, respectively. These were functions of the size-dependent expected growth $(g(z,\theta))$ and expected offspring size at birth $(\phi(z,\theta))$, respectively, and the estimated variation around these means $(\sigma_g$ and $\sigma_\phi)$.

$$G(z',z,\theta) = Normal(z'|g(z,\theta),\sigma_g) \tag{4.3}$$

$$\phi(z',z,\theta) = Normal(z'|\phi(z,\theta),\sigma_\phi) \tag{4.4}$$

The constructed IPM was discretized into a 100×100 matrix, with (standardized) z ranging between -3 and 3 (corresponding to 0.4-3.8 mm), and subsequent analyses were performed on the obtained matrix (described in *IPM analyses*). To avoid eviction (Williams et al. 2012), predicted values outside the implemented size range were added to the first or last size class.

Modelling infection dynamics of the host

In addition to the one-state IPM shown in Eq. 4.1, we constructed a two-state IPM. Here, individuals were characterized by two state variables: body size z, and parasite load P. Changes in the distribution of z (due to z-specific survival, growth and reproduction) are described by Eq. 4.1. Changes in the distribution of P are due to all transition probabilities between parasite classes (i.e. describing infection and recovery dynamics) and due to P-specific mortality patterns. To estimate the transition probabilities between parasite classes, we defined a 5×5 transition matrix **A** describing daily probabilities for moving from parasite load P_t at time t (columns), to P_{t+1} at time $t+1$ (rows). The first class represents uninfected individuals, the 2nd-4th classes represent parasite load class 1-3, and the last class represents dead individuals:

$$\mathbf{A} = \begin{bmatrix} S_{P=0}(1-\alpha) & S_{P=1}\rho & 0 & 0 & 0 \\ S_{P=0}\alpha & S_{P=1}(1-\rho)(1-\gamma) & S_{P=2}\omega & 0 & 0 \\ 0 & S_{P=1}(1-\rho)\gamma & S_{P=2}(1-\omega)(1-\gamma) & S_{P=3}\omega & 0 \\ 0 & 0 & S_{P=2}(1-\omega)\gamma & S_{P=3}(1-\omega) & 0 \\ 1-S_{P=0} & 1-S_{P=1} & 1-S_{P=2} & 1-S_{P=3} & 1 \end{bmatrix} \tag{4.5}$$

Here, α is the probability of becoming infected, ρ is the probability of recovering, ω is the probability of transitioning to an earlier parasite load stage, and γ is the probability of moving to the next parasite load stage. The bottom row contains the estimated survival probabilities for different parasite load stages as predicted using the fitted vital rate model.

Infection probabilities were modeled as a function of parameter β and the (time-dependent) average parasite load $(\overline{P_t})$: $\alpha_t = 1 - \exp(-\beta\overline{P_t})$. We thus assumed that infection probabilities only depended on the average parasite load in the population, and not on population size. We made this assumption because exploratory analyses suggested no positive relation between parasite load and population density; in contrast, a weak negative relation was observed, which we address in the discussion. Transition probabilities γ, ω and ρ were fitted as a function of temperature T, as observations show temperature-dependence in parasite dynamics (see Fig. 4.3). A logit link function was used to obtain probabilities. Fitting models describing the transition probabilities α, γ, ρ, ω involved a total of 7 parameters, and these were estimated simultaneously, as explained below. An additional two parameters were estimated in this procedure: first, we included a parameter π to estimate the 'start' proportion of infected individuals (with a parasite load of 1) at the day

the infection started. We set this day at day 29, which is the day the first infected individual was observed (see *Parasite load*). Second, we estimated the standard deviation of the normal distribution comparing predicted with observed parasite loads, as explained below.

To estimate the required 9 parameters, we included two sources of data. First, we included 1402 individual observations for which we had data on parasite load both when first measuring and when remeasuring after one time interval (including dead individuals when remeasuring). To compare individual transitions with predictions, for each individual i with parasite load P_i we obtained the expected distribution at $t + \Delta t$ by projecting the constructed matrix $\mathbf{A}$ for Δt time steps. The resulting distribution was subsequently compared to the individual's parasite load at $t + \Delta t$ using a multinomial distribution. These measurements were obtained during a limited time period (all collected between days 60 and 81, as described in subsection *Parasite load*). Second, we additionally used population-averaged parasite loads over time, calculated per temperature, for which we had data during the entire period (days 0-81). To compare population averages with observations, we simulated the spread of infection in a host population using matrix $\mathbf{A}$, and compared average parasite loads with observations. While constructing temperature- and day-specific 5×5 transition matrices, we projected parasite load dynamics through time and calculated average parasite loads, which were then compared to observations using a normal distribution. Here, we made the simplifying choice to implement survival probabilities for average body size, and averaged predictions from each location of origin. We scaled the log-likelihoods from both data sources in order to give them similar weights, and optimized the total log-likelihood.

Several assumptions were made for this approach. We assumed that individuals could transition only one parasite class per day. Moreover, we assumed that individuals became susceptible again directly after recovering (i.e. we assumed that individuals did not attain immunity). We further assumed that newborns were proportionally distributed across all classes, and that each aquarium experienced initial parasite exposure at the same time. Finally, we assumed that infection from external sources occurred only at day 29. In other words, after day 29, individuals could only become infected by infected individuals already inside the population and no longer from external sources.

The discretized body-size structured IPM shown in Eq. 4.1, and the parasite transition matrix $\mathbf{A}$ (excluding class 5, reflecting the dead individuals) were combined into a matrix describing all transitions between body sizes, as well as all transitions between parasite load classes. We here assumed that neonates were born with the same parasite load as their mother. The resulting 400×400 matrix was used for subsequent analysis, as described next.

IPM analyses

We performed three analyses to assess the interactive effects of temperature and infection on clones originating from the four locations. First, the 1-state IPM was used to project asymptotic population growth rates for different densities, temperatures, parasite loads and locations. We here let temperature range between 14 °C and 26 °C, densities between 1 and 1000 individuals and parasite load between 0 and 3. The constructed IPMs were used to obtain equilibrium densities (i.e. the density at which population growth rates are equal to one).

Second, we evaluated which vital rates caused the differences in performance among locations when exposed to infection. We used the 2-state IPM to simulate host-parasite dynamics, with an infection starting at day 1, at the estimated proportion π, in order to simultaneously project changes in body size structure and changes in parasite load structure. Per temperature, we started by simulating dynamics of the 'average' location. This was done by averaging each vital rate function over the four locations ($N1$, $N2$, $B1$, $B2$), per time step, and combining those into an IPM. The resulting averaged IPM was multiplied with the time-specific body size and parasite

load distribution, to project the distribution at $t + 1$. In this way, we simulated 150 time steps until a stable host density and parasite load structure was reached, and obtained the equilibrium density of this average location. We then, one by one, replaced each of the averaged vital rate functions by a location-specific vital rate function, reran the whole simulation, and again obtained the equilibrium host density. The difference between these two equilibrium densities measures the population-level effect of vital rate differences between locations when exposed to infection, compared to the average. In other words, it quantifies which vital rates contribute to differences in host tolerance among locations. This analysis was repeated for each temperature.

Finally, we used the two-state IPM to quantify how different factors contributed to differences in host-parasite dynamics at 14 °C compared to 26 °C. We first projected the stable parasite load structure and host equilibrium density at 14 °C, by projecting dynamics for 200 time steps, for one of the four locations. We then replaced different functions by the function when temperature was set at 26 °C, and projected parasite load structure and equilibrium density again. These rates were the temperature-dependent transition probabilities between parasite classes (ρ, γ, ω; describing host resistance) and the host's demographic responses to temperature, in interaction with parasite infection (describing host tolerance). We here separated the effects of survival and growth (i.e. effects on survivors, the first term on the right-hand of Eq. 4.1), and the effects of reproduction and neonate body size (i.e. effects on newborns, the second term on the right-hand of Eq. 4.1). This analysis was repeated for each location.

4.4   Results

Population trends and genotype frequencies

After a strong increase in densities in each of the 114 populations, all *Daphnia* populations showed strong reductions in size during the second half of the experiment, after arrival of the parasite. In total, 37 populations went extinct prior to the end of the experiment (Appendix C.5): 62% of the Norwegian populations (either single clone populations, or Norwegian mixes) went extinct, 21% of the Belgian populations (single clone and mixed), and 14% of the Norwegian-Belgian mixes went extinct. The largest difference between Norwegian and Belgian population performance was found at the highest temperature, at which none of the Belgian populations went extinct, but 63% of the Norwegian populations.

Genetic analyses of the sampled individuals, show a clear hierarchy in the competitive strength of clones from different locations of origin (Fig. 4.1), and are in agreement with the extinction patterns. Clones from location $B2$ practically outcompeted each of the other clones at each temperature. Clones from location $B1$ outcompeted clones from $N1$ and $N2$, but not from $B2$, whereas the Norwegian clones ($N1$ and $N2$) show mixed outcomes in competition amongst themselves. Looking at the temporal changes in genotype frequencies in each population (Appendix C.2 for more details and results), most populations show a clear change in the ratio of the two genotypes. In only 4 out of the 54 mixed populations, the best model based on AIC values included no temporal trends. In 14 mixed populations, the best model included a directional temporal change in proportions. Finally, in 36 mixed populations, the best model included both an effect of day and a quadratic effect of day, with opposite signs, suggesting a change in the direction during the experiment. In the 50 populations which show a trend, the ratio between the two genotypes is predicted to reach at least 25%-75% within the course of the experiment. Moreover, in half of these populations, this is predicted to occur before day 24.
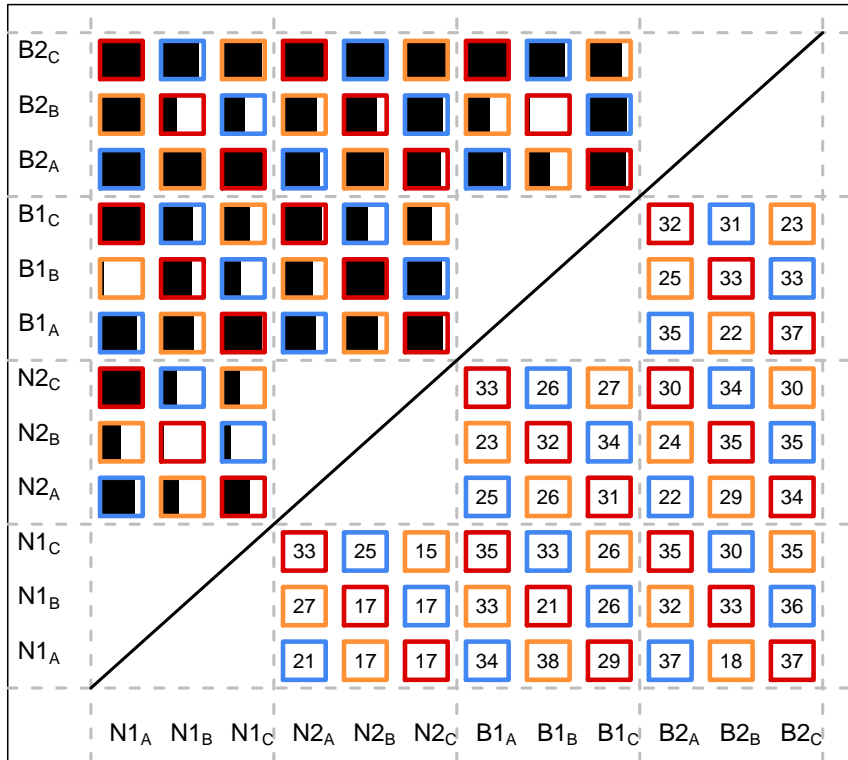
Figure 4.1: Average clonal frequencies of different combinations of clones, based on sampled individuals during the experiment, as identified using genetic markers. Rows (grey) are compard to columns (white). Coloured borders indicate the temperature treatment (blue: 18 °C, orange: 22 °C, red: 26 °C). For example, the first row shows in black the proportion of Belgian clone $B2_C$, when compared to each of the other genotypes. Below the diagonal, sample sizes for each comparison are given.

	Survival ($S(z, \theta)$)	Growth $g(z, \theta)$	Carrying eggs ($p_e(z, \theta)$)	Reproduction ($p_r(z, \theta)$)	Clutch size ($F(z, \theta)$)	Neonate size ($\phi(z, \theta)$)
β_0	**4.23e+00 [1.16e-01]**	**7.65e-02 [1.60e-03]**	**-2.55e+00 [1.39e-01]**	**-4.25e-01 [1.40e-01]**	**5.16e-01 [5.88e-02]**	**-1.98e+00 [3.04e-02]**
z	-2.92e-02 [4.92e-02]	**-4.11e-02 [8.41e-04]**	**4.14e+00 [2.75e-01]**	1.75e-01 [1.06e-01]	**1.16e+00 [3.47e-02]**	**2.47e-01 [4.58e-02]**
z^2		**6.77e-03 [6.48e-04]**	**-1.03e+00 [1.24e-01]**			**-8.90e-02 [2.07e-02]**
C	-4.83e-01 [3.95e-01]	**1.20e-02 [4.41e-03]**		**2.42e-01 [1.04e-01]**		-4.10e-02 [8.28e-02]
T	**-2.66e-01 [8.68e-02]**	3.12e-04 [1.12e-03]	**2.02e-01 [7.93e-02]**	**5.12e-01 [8.32e-02]**	6.73e-03 [1.45e-02]	**1.50e-01 [1.86e-02]**
N		**-1.79e-02 [1.20e-03]**	**-4.86e-01 [7.21e-02]**	**-2.31e-01 [6.29e-02]**	**-6.14e-01 [2.36e-02]**	**5.49e-02 [1.22e-02]**
P	**-7.74e-01 [6.23e-02]**	**-1.83e-02 [1.32e-03]**	**-5.33e-01 [5.95e-02]**	-7.82e-03 [9.32e-02]	**-4.88e-01 [3.62e-02]**	**-4.53e-02 [1.55e-02]**
L_1	**-8.16e-01 [1.48e-01]**	**-1.16e-02 [2.08e-03]**	**-3.84e-01 [1.96e-01]**	**6.98e-01 [2.51e-01]**	1.04e-01 [9.86e-02]	**-1.16e-01 [2.88e-02]**
L_2	-8.58e-03 [1.03e-01]	2.52e-03 [1.61e-03]	6.81e-02 [1.60e-01]	**5.33e-01 [2.50e-01]**	**2.01e-01 [8.44e-02]**	**-5.10e-02 [2.16e-02]**
L_3	1.04e-01 [1.16e-01]	**7.69e-03 [1.43e-03]**	**-9.47e-01 [1.23e-01]**	-1.18e-02 [9.20e-02]	2.88e-02 [7.52e-02]	2.37e-02 [1.90e-02]
$P{:}L_1$	4.05e-01 [9.26e-02]	9.41e-04 [2.69e-03]		**-4.39e-01 [2.21e-01]**	1.38e-01 [7.34e-02]	
$P{:}L_2$	**-2.06e-01 [7.01e-02]**	**-5.39e-03 [2.31e-03]**		**-5.20e-01 [2.17e-01]**	**-2.85e-01 [7.47e-02]**	
$P{:}L_3$	4.92e-02 [6.08e-02]	4.09e-04 [1.30e-03]		1.74e-01 [9.30e-02]	2.21e-02 [3.77e-02]	
$T{:}L_1$	1.10e-01 [9.57e-02]	-4.34e-04 [1.86e-03]	1.22e-01 [1.10e-01]	3.44e-01 [2.05e-01]		-2.33e-03 [2.61e-02]
$T{:}L_2$	**-2.13e-01 [7.10e-02]**	**-3.10e-03 [1.41e-03]**	-7.44e-02 [8.30e-02]	**4.35e-01 [1.87e-01]**		3.03e-02 [1.91e-02]
$T{:}L_3$	1.02e-01 [6.61e-02]	**3.82e-03 [1.12e-03]**	**-3.05e-01 [7.04e-02]**	**-1.61e-01 [7.73e-02]**		**-4.32e-02 [1.69e-02]**
$N{:}L_1$		**-4.90e-03 [1.85e-03]**	**-3.21e-01 [1.14e-01]**		**1.15e-01 [4.21e-02]**	
$N{:}L_2$		-5.30e-04 [1.45e-03]	7.27e-02 [9.17e-02]		3.97e-02 [2.79e-02]	
$N{:}L_3$		**2.59e-03 [1.27e-03]**	-1.25e-01 [6.93e-02]		4.10e-02 [2.49e-02]	
$z{:}L_1$			1.57e-01 [1.96e-01]		**-1.38e-01 [5.58e-02]**	
$z{:}L_2$			-1.12e-02 [1.54e-01]		1.66e-02 [5.49e-02]	
$z{:}L_3$			**4.52e-01 [1.36e-01]**		**-1.31e-01 [4.68e-02]**	
$C{:}L_1$	**1.28e+00 [5.57e-01]**				2.53e-02 [1.53e-01]	**2.47e-01 [1.24e-01]**
$C{:}L_2$	2.12e-01 [5.25e-01]				-2.70e-01 [1.49e-01]	6.50e-02 [8.78e-02]
$C{:}L_3$	-3.85e-01 [4.46e-01]				1.53e-02 [1.20e-01]	4.32e-02 [8.46e-02]
$T{:}P$	7.57e-02 [4.22e-02]		**1.31e-01 [5.97e-02]**			-2.44e-02 [1.62e-02]

Table 4.1: Coefficients and standard errors for each of the vital rates. Bold numbers indicate significant effects ($p \leq 0.05$). Coefficient β_0 indicates the intercept, z, z^2: effects of body size and squared body size, C: effects of competition against clones from different locations, T: effects of (standardized) temperature, N: effects of (standardized) population density, P: effects of parasite load. L_1 refers to the effects of Norwegian populations N1 and N2, compared to Belgian populations B1 and B2, L_2 refers to effects of $N2$ compared to $N1$, and L_3 refer to $B2$ compared to $B1$.

Effects on single vital rates and location differences

Estimated coefficients for each of the vital rates can be found in Table 4.1; we will here summarize the most notable, and significant, effects. Temperature had positive effects on vital rates associated with reproduction, but negative effects on survival. Parasite load had negative effects on all vital rates, except on the probability of reproducing offspring conditional on having eggs (i.e. neonate development). A positive interaction between parasite load and temperature was estimated for probability of carrying eggs, implying that the negative effect of infection diminishes with temperature. Population density negatively affected each of the vital rates, except survival and neonate size.

Coefficients comparing locations $N1$ and $N2$ to locations $B1$ and $B2$ (L_1) are negative for survival, somatic growth, carrying eggs probability and neonate size, while positive for neonate development, indicating that for most vital rates, the Norwegian lineages ($N1$ and $N2$) performed worse compared to the Belgian lineages ($B1$ and $B2$). In four out of six vital rates, clones from population $N2$ responded stronger to the parasite load than $N1$ (interaction coefficient $P \times L_2$), indicating a lower parasite tolerance for population $N2$.

Interactive effects of temperature and parasite load

Integrating over all vital rates using the one-state IPM revealed a strong negative effect of parasite load on projected equilibrium densities (Fig. 4.2). Recall that tested temperatures ranged between 14 °C and 26 °C, and individual parasite loads ranged between 0 (uninfected) and 3 (heavily infected). Effects of temperature were weak compared to the effects of infection. Clones from Belgian locations $B1$ and $B2$ were projected to achieve the highest densities at all temperatures, and this pattern was the same across parasite loads. Clones from location $N2$ responded strongest to parasite infection (i.e. was the least tolerant), and could only persist when parasite load was smaller than ~1.5. Clones from location $B2$ were able to persist under highest parasite loads, indicating the highest tolerance.

Host-parasite dynamics at different temperatures

The proportion of infected individuals at day 29 (π) is estimated to be 3.4%. Estimated daily infection probabilities α increased with increasing average parasite load up to ~40% when the average parasite load is at its maximum of 3 (Fig. 4.3). Transition probabilities between parasite load classes ρ, ω and γ all increase with temperature, with the probability of transitioning to a previous class (ω) showing the strongest temperature response (Fig. 4.3). This results in a decreased host resistance at low temperatures, reflected by a higher population-average parasite load (Fig. 4.3). Population-level predictions on parasite loads through time, as well as on the parasite load distributions at the end of the experiment, capture the observed dynamics reasonably well (Fig. 4.3 and Appendix C.6, for average parasite loads and stage distributions, respectively).

The 2-state IPM was used to quantify the contribution of different vital rates to the differences in performance of infected populations, for different temperatures (Fig. 4.4). Overall, vital rates of clones from location $B2$ had the largest positive effects on equilibrium densities, with the largest benefits caused by survival and somatic growth (Fig. 4.4). For example, equilibrium density (N_{eq}) at 14 °C increases with ~50% when $B2$-specific survival is implemented instead of the average survival function. This increases N_{eq} from 124 to 184 individuals. In contrast, implementing survival of $N1$ decreases N_{eq} to only 71 individuals. At 26 °C, location $B1$ benefits from higher probabilities of carrying eggs, increasing N_{eq} from 461 to 540 individuals. In general and across temperatures, almost all vital rates of locations $N1$ and $N2$ decrease equilibrium densities.

Finally, per location, we assessed which processes caused temperature-related differences in host-parasite dynamics. The 2-state IPM predicted higher equilibrium host densities and
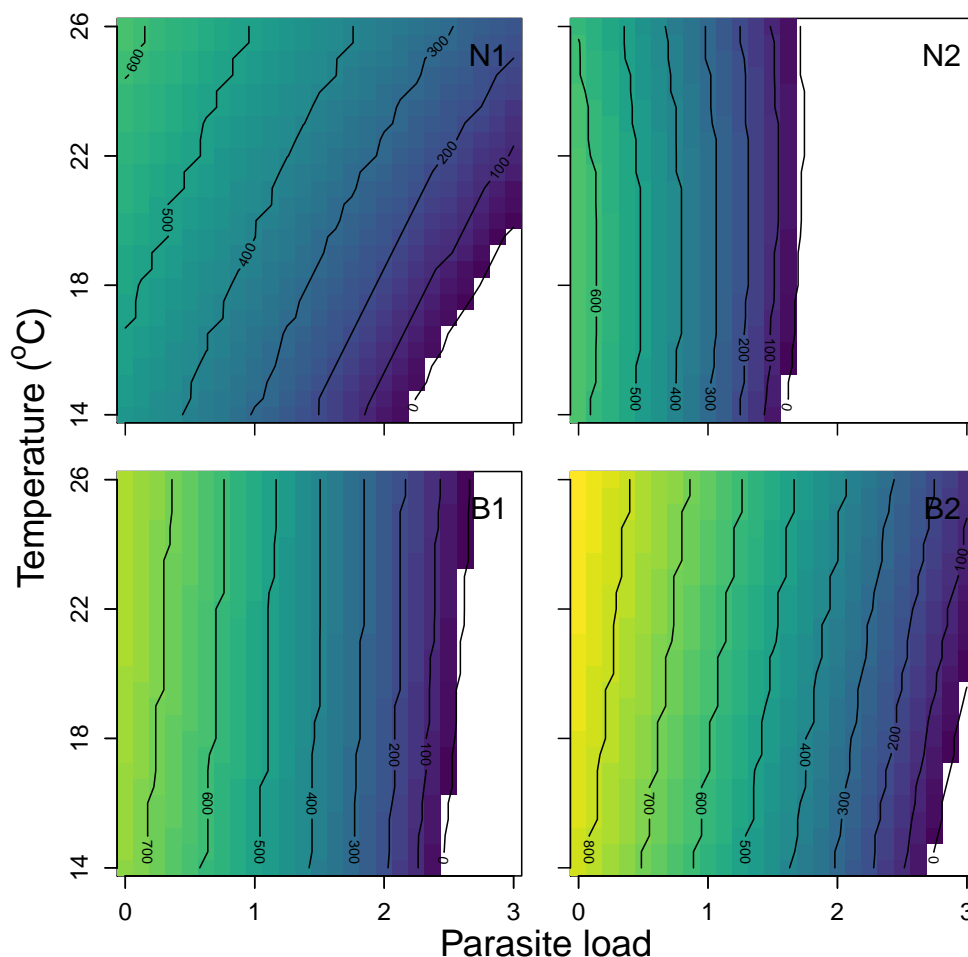
Figure 4.2: Projected equilibrium densities for different combinations of temperature and parasite load, based on an integration of all vital rates using Integral Projection Models. Different panels show results for different locations of origin. Values smaller than zero are not shown.

lower infection loads at higher temperatures (Fig. 4.5, see also Fig. 4.4). This is mostly due to temperature-related differences in resistance. More specifically, this is due to the positive temperature effects on daily transition probabilities ω, describing probabilities of moving from stage 3 to 2, and from 2 to 1 (Fig. 4.5). Running a simulation for 14 °C when replacing ω by the 26 °C specific probability, increases N_{eq} from 238 to 493 individuals, for location $B2$ (Fig. 4.5; see Appendix C.7 for results for the other locations which were qualitatively similar). Moreover, it decreases the average parasite load (Fig. 4.5). When doing the same analysis, but now replacing recovery probabilities (i.e. moving from parasite load 1 to 0) by the 26 °C specific probability, increases N_{eq} to 344 individuals. Probabilities of transitioning to the next parasite class (γ) have the opposite effect and reduce the equilibrium densities (caused by the increasing effect of temperature on these probabilities; Fig. 4.3). Finally, temperature-related host tolerance to infection also
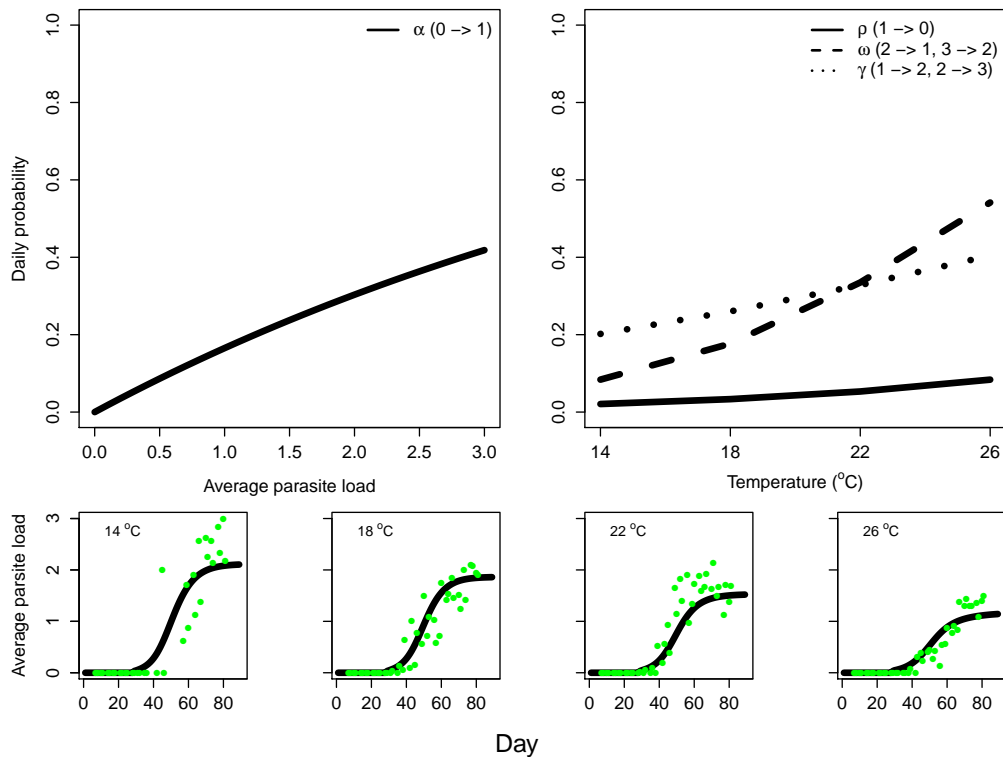
Figure 4.3: Estimated probabilities of moving between parasite classes. Top left graph shows daily infection probabilities α (i.e. probabilities of moving from parasite load class 0 to 1, as indicated in the graph) as a function of the population-level average parasite load. Top right graph shows recovery probabilities ρ, probabilities of moving to the previous parasite load class ω and probabilities of moving to the next parasite class γ, as a function of temperature. Bottom graphs show, per temperature, predicted parasite loads through time, and green dots show observations averaged over all clonal treatments.

contributes to the higher equilibrium densities at higher temperatures, increasing N_{eq} to 305 and 295, due to dynamics of the newborns and the survivors, respectively (Fig. 4.5).

4.5 Discussion

Climate change is predicted to have profound and complex effects on local populations (Hoffmann and Sgrò 2011). In addition to direct temperature effects, local gene pools may be altered by preadapted immigrating conspecifics and by temperature-related changes in disease dynamics. In this study, results suggest strong negative impacts of an ectoparasite on the performance of the host *Daphnia magna*. Moreover, the effects of parasite infection were much stronger than the effects of temperature, for both water fleas originating from northern (Norwegian) and southern (Belgian) ponds. *Daphnia* individuals showed clear between-population differences in tolerance against infection, with clones from both Belgian ponds being able to realize a higher fitness than the Norwegian clones (Figs 4.1, 4.2). Moreover, when taking into account temperature-dependence in recovery and transition rates between parasite loads (Fig. 4.3), we showed that host-parasite dynamics differed among host clones from different locations (Fig. 4.4) and a strong
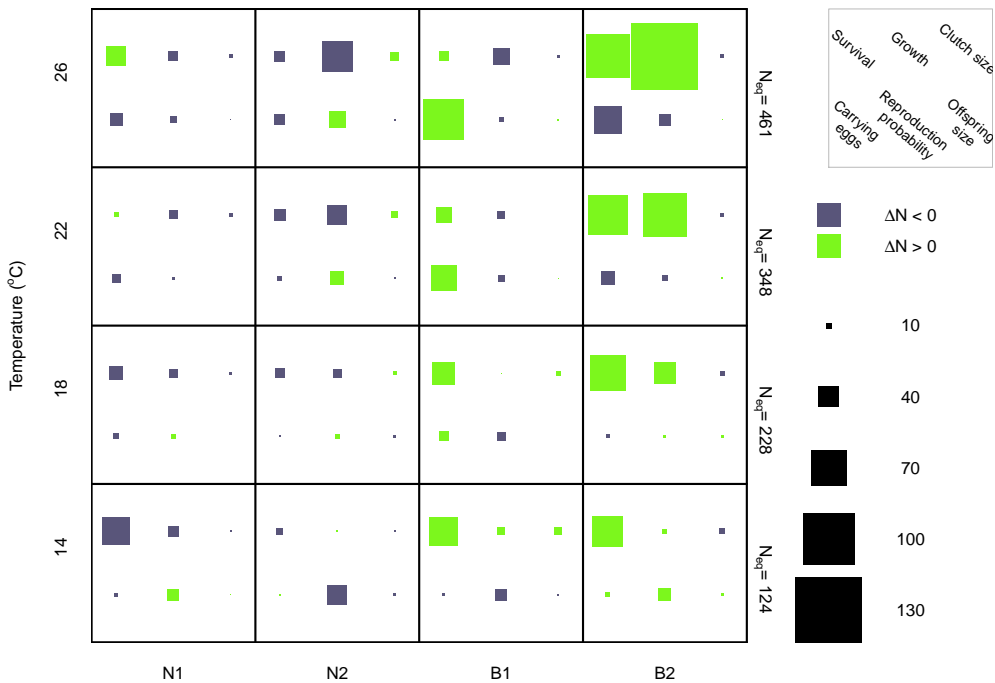
Figure 4.4: Differences in equilibrium densities N_{eq} between the four locations due to the contribution of different vital rates. For each temperature, the two-state IPM was used to simulate host-parasite dynamics during 150 time steps. This was first done for the 'average' location, by averaging vital rate functions at each time step. Equilibrium densities for the average locations are given in the graph, per temperature. Next, we ran simulations in which we, one by one, replaced each averaged vital rate function by a location-specific vital rate, and obtained equilibrium densities. The difference between the equilibrium density of the average location and this density, measures the contribution of that specific vital rate to the projected host density when infected, for each location. These values are depicted in the graph. Green colours indicate positive effects on equilibrium density (i.e. a demographic benefit related to the vital rate), and blue colours indicate negative effects. Squares are scaled to their effect size (see legend). See the legend for the order in which vital rates are depicted.

temperature-dependence in these dynamics (Fig. 4.5). Altogether, these results underscore that the consequences of climate change will not only depend on the capability of local populations to respond plastically or genetically to changes in temperature, but also largely on their capability to respond to associated parasite dynamics.

Host performance when uninfected

All 114 populations, across all temperature treatments, showed strong increases in population size during the first weeks, growing up to a few hundred individuals per population. All clones could thus respond plastically to the tested temperature range and maintain positive population growth rates. Our unique experimental setup enabled measuring of individuals from within the populations,
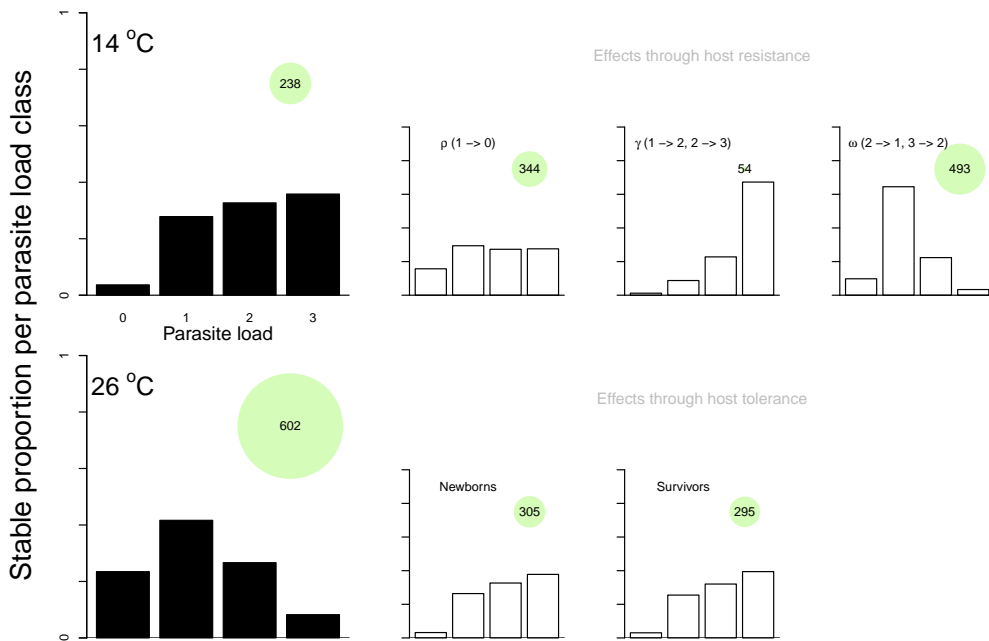
Figure 4.5: Host-parasite dynamics at 14 °C and 26 °C, based on a host-parasite IPM. Barplots on the left show the projected proportion of individuals in each parasite class after 200 time steps. Green circles show stabilized densities. Starting with the host-parasite dynamics at 14 °C, white barplots show proportions and densities when one function is replaced by the relevant function at 26 °C. Upper row shows results when changing functions associated with the host resistance: recovery probability, probability of moving to the next parasite class, and probability of moving to the previous parasite class (see Eq. 4.5). Bottom row shows the temperature effects on host tolerance, through the dynamics of newborns (reproduction and neonate body size) and survivors (survival and growth). Results are shown for location *B*2, but were largely similar for each location (see Appendix C.7).

by isolating them for a few days (which is short compared to their total life span) in permeable tubes. This allowed us to quantify how different vital rates, together determining population growth rates, were affected by body size, environmental factors and genetic background. We found that, at higher temperatures, components related to development were increased (i.e. size-dependent probability of carrying eggs and neonate development), while survival was decreased (Table 4.1). These findings are mostly in line with previous life table studies on *Daphnia* (MacArthur and Baillie 1929; Van Doorslaer et al. 2009a; Henning-Lucass et al. 2016). Vital rates differed between clones originating from different locations, whereby most notable are the negative effects in four out of six vital rates for the Norwegian populations compared to the Belgian populations (coefficients *L*1 in Table 4.1). This results in a predicted lower fitness (equilibrium densities; Fig. 4.2, as well as population growth rates; Appendix C.8) for the Norwegian clones. This is in agreement with data on genotype frequencies, where we showed that both Belgian populations had a clear competitive advantage compared to the Norwegian clones, and that for many populations, these effects already occurred during the first half of the experiment (i.e. before the infection became prevalent) (Appendix C.2). If clones would be thermally adapted to the temperature they usually experience, we would expect the Belgian and Norwegian clones to have differerent thermal

optima. Moreover, we would expect the Belgian clones to outcompete the Norwegian clones at high temperatures, while the opposite to be true at low temperatures. However, we found that the Belgian clones showed higher fitness across all temperatures and thus we did not find any evidence for local thermal adaptation, in contrast to other studies, for example Yampolsky et al. (2013).

Another important environmental factor that influences vital rates is the conspecific density. Individual performance generally decreases with increasing density (Guisande 1993; Goser and Ratte 1994), for instance through effects on per capita food availability (Lampert 1978). Moreover, responses to density may differ between genotypes (Agrawal et al. 2004; Fitzsimmons and Innes 2006), which can lead to eco-evolutionary feedback loops (Turcotte et al. 2013). Therefore, assessing the competitive strength of local versus immigrating genotypes and predicting the evolutionary responses, requires explicit knowledge of density-dependence in the performance of different genotypes. We observed significant negative density effects in four out of six vital rates, on a population level resulting in a negative relation between density and population growth rate. We also found evidence for differential density responses among clones from different locations and location-specific equilibrium densities (Fig. 4.2). Norwegian populations responded stronger (more negative) to density in somatic growth and carrying eggs probability, while weaker effects were found in clutch size. Water fleas from location $B2$ showed a weaker response to density in both somatic growth compared to $B1$ (Table 4.1). For survival, an opposite result was observed than for other vital rates, since exploratory analyses revealed a positive relation between density and survival. This relation between density and survival has been documented earlier, when also following the dynamics of populations (Bruijning et al. 2018b). As discussed in Bruijning et al. (2018b), we expect a reverse causality here: in populations in which individuals survive better, higher densities are reached. Future studies experimentally manipulating densities and measure survival rates, could help to determine the actual density effects on individual survival.

Host tolerance and resistance

Natural zooplankton populations generally show high parasite infection rates, both by endoparasites and by epibionts (also called ectoparasites) (Chiavelli et al. 1993; Stirnadel and Ebert 1997; Decaestecker et al. 2004). These host-parasite interactions have the potential to influence the ecological and evolutionary dynamics of host populations (Ebert 2005; Decaestecker et al. 2007). How host population dynamics are affected by a parasite depends on both the host's resistance and tolerance. Recall that resistance is the ability of the host to limit parasite burden, while tolerance is the ability of the host to limit the negative effects of the parasite burden (Råberg et al. 2007). In this study, both these factors were considered, as discussed next.

Ectoparasites, such as *Amoebidium parasiticum*, generally have small effects on host performance compared to endoparasite infections (Ebert 2005; Decaestecker et al. 2004), implying a high host tolerance to ectoparasites. However, when abundant, ectoparasites have the potential to negatively effect host reproduction and survival as well as population growth (Kuno 1973; Brambilla 1983; Stirnadel and Ebert 1997), by impeding movement and reducing buoyancy (Chiavelli et al. 1993). A relation between levels of the stress protein Hsp60 in *Daphnia*, and the abundance of *A. parasiticum*, has been suggested (Pauwels et al. 2007), also indicating that *A. parasiticum* affects its host. This is in agreement with our results, which indicate negative effects of parasite load on basically all vital rates. On a population-level, increasing parasite load resulted in large declines in projected equilibrium densities, showing much stronger effects than temperature (Fig. 4.2). We also found evidence for host differences in tolerance, reflected by significant interactions between parasite load and location (Table 4.1). Most notable, water fleas from location $N2$ responds stronger to infection compared to those from location $N1$ in components related to survival, growth and reproduction. This resulted in the lowest tolerance against infection for clones from location $N2$

(Fig. 4.2).

Host resistance depends on the transmission and recovery probabilities. For *Daphnia*, it has been experimentally shown that parasite transmission probabilities are a function of the host density. Ebert (1995), for instance, has shown that the transmission probability increases with a higher spore density, as a measure for the density of infected hosts, of the microsporidian parasite *Pleistophora intestinalis*. In natural populations, however, density-dependent parasite transmission in *Daphnia* has rarely been described, and it remains unknown how important it is (Ebert 2005). Under density-dependent transmission probabilities, we expect a reduction in infection rates once populations are declining. However, this was not what we observed in this study: average parasite loads continued to increase (with the exception of a few populations, which indeed managed to partly recover), and showed a tendency to stabilize at the end of the experiment (Fig. 4.3). *A. parasiticum* reproduces by the production of spores (Kuno 1973), which can remain viable for a long time (Ebert 1995; Decaestecker et al. 2004). As we kept 1/3 of the old medium when cleaning the aquariums, we suspect a build up of spores in the aquariums (being closed systems) through time. This might explain why we did not observe this reduction in infection rates once hosts became less abundant, and why we observed continuing increases in infection rates. We therefore decided to model transmission probabilities as a function of the population-level average parasite load, as this might be a better proxy for the total amount of present parasite spores. Finally, we show that hosts were able to reduce parasite burden: the daily probabilities of moving to a previous parasite class were less than 10% at 14 °C, increasing to a maximum of ∼50% at 26 °C (Fig. 4.3). Estimated probabilities of full recovery, however, were much lower. This is in agreement with the observation that *Daphnia* parasites generally result in chronic infections with low recovery (Ebert 2005).

Although the *A. parasiticum* infection was initially not part of the study design, we had anticipated it could occur, as it had been observed in the laboratory facilities before. Differential parasite effects on the Belgian compared to Norwegian populations (Appendix C.5) and the interaction with temperature (Fig. 4.3) became immediately apparent during the experiment, and it has given us the exciting opportunity to study the interactive effects of temperature and parasite infection, and the role of genetic background of the host. Unfortunately, we lacked the power to detect potential differences in resistance among locations, while it is known that genotypes of *Daphnia* can differ in their resistance (Carius et al. 2001; Decaestecker et al. 2007; Bourgeois et al. 2017; Krebs et al. 2017). Given the observed differences in tolerance, there might also be resistance differences among the genotypes studied here. It remains to be tested whether this will result in even larger predicted performance differences between Belgian and Norwegian populations, or whether tolerance differences can be partly counterbalanced by differences in resistance.

Interactive effects of temperature and infection and implications for natural populations

In order to predict the fate of local populations in response to climate change, it is important to incorporate how other abiotic and biotic factors that influence local populations, are affected by climate change (Ficke et al. 2007; Tylianakis et al. 2008; Labaude et al. 2017). One of the most common biotic interactions are those between pathogens and their host, and pathogen dynamics themselves are highly affected by temperature (Morley and Lewis 2014). For the host-parasite interaction studied here, increased temperatures reduced the impact of infection on host fitness. We showed that this was mainly due to a temperature-dependence in the host ability of reducing parasite burden (i.e. host resistance), and to a lesser extent, we found that hosts showed increased tolerance against infection at higher temperatures. Although this clearly emphasizes the importance of considering non-additive effects of temperature and pathogens, these benefits of

higher temperatures may be an exception, as it has been suggested that pathogen infections will become more severe and prevalent under climate change, imposing an extra pressure on natural populations (Harvell et al. 2002; Mouritsen et al. 2005).

Climate migrants can introduce new pathogens to local populations (Holdich and Reeve 1991; Prenter et al. 2004), which can further challenge the persistence of local populations. We found clear between-location differences in fitness and in parasite tolerance, which supports the idea that local gene pools can be altered by the invasion of migrating individuals. Both Belgian populations were able to realize a higher fitness, and this was through benefits related to multiple vital rates. It remains to be tested whether the increased parasite tolerance of the Belgian genotypes reflects an evolutionary adaptation in response to increased exposure, as we have no data on the prevalence of *A. parasiticum* in the ponds where we obtained our clones from. We do note that *A. parasiticum* was shown to be the most prevalent ectoparasite in another Belgian pond (Decaestecker et al. 2004; Pauwels et al. 2007), located 50 kilometers (*B*1) and 100 kilometers (*B*2) away from the locations studied here.

To conclude, we have shown that *Daphnia magna* clones from different latitudes differed in their performance. A parasite infection was shown to have stronger fitness effects than temperature, and moreover, dynamics of the parasite themselves were highly influenced by temperature. With temperatures that are predicted to continue to increase during the coming century, natural populations will become more likely to experience competition from immigrants. These climate migrants may be preadapted to temperature and to the pathogens they will bring along, which might even be a stronger selection pressure than direct temperature effects. Altogether, this can have large implications for natural populations that have to deal with all these biotic and abiotic changes concurrently.

Acknowledgements

Marjolein Bruijning[1], Eelke Jongejans[1] and Martin M. Turcotte[2]

1. Department of Animal Ecology and Physiology, Radboud University, 6500 GL Nijmegen, The Netherlands.
2. Department of Biological Sciences, University of Pittsburgh, Pittsburgh, USA.

Demographic responses underlying eco-evolutionary dynamics as revealed with inverse modelling

5.1 Abstract

1. Changes in population dynamics due to interacting evolutionary and ecological processes are the direct result of responses in vital rates, i.e. stage-specific growth, survival and fecundity. Quantifying through which vital rates population fitness is affected, instead of focusing on population trends only, can give a more mechanistic understanding of eco-evolutionary dynamics.

2. The aim of this study was to estimate the underlying demographic rates of aphid (*Myzus persicae*) populations. We analyzed the unpublished stage-structured population dynamics data of a field experiment with caged and uncaged populations in which rapid evolutionary dynamics were observed, as well as unpublished results from an individual life table experiment performed in a greenhouse.

3. Using data on changes in population abundance and stage distributions over time, we estimated transition matrices with inverse modelling techniques, in a Bayesian framework. We validated our modelling framework using simulated data, and show that parameters could be estimated accurately and unbiased when the true underlying model structure is known. Based on cross-validation, we compared models including stage-dependency in vital rates, exposure to predators and competitors, population size, plant size and density as explanatory variables. Density resulted in the largest model improvement, and the final model used to fit across all experimental treatments therefore included density as well as clone-specific caging effects.

4. Results suggest that clones varied considerably in vital rates, and imply trade-offs between reproduction and survival. Responses to densities also varied between clones. Negative density-dependence was found in growth and reproduction, and the presence of predators and competitors further decreased these two vital rates, while survival estimates increased.

Under uncaged conditions, population growth rates of the evolving populations were increased compared to the expectation based on the pure clones.

5. Our inverse modelling approach revealed how much vital rates contributed to the eco-evolutionary dynamics. The decomposition analysis suggested that variation in density-dependent population growth rates in the evolving populations were to a large extent shaped by plant size. Finally, we emphasize that inverse modelling is a complex problem, as multiple combinations of individual rates can result in the same dynamics. We discuss assumptions and limitations, as well as opportunities, of this approach.

5.2 Introduction

Rapid evolution, defined as genetic changes that are fast enough to have an impact on ecological dynamics (Hairston et al. 2005), has been observed in a wide range of organisms (see Schoener 2011). Given that evolutionary and ecological processes can act simultaneously, they also have the potential to interact with each other. These eco-evolutionary dynamics potentially play an important role in shaping populations, communities and ecosystems (Fussmann et al. 2007; Bassar et al. 2010b; Strauss 2014; Matthews et al. 2016). Discriminating between ecological and evolutionary processes and quantifying their relative importance is challenging, especially in natural populations, but different frameworks exist that aim to disentangle different processes (Hairston et al. 2005; Coulson and Tuljapurkar 2008; Ellner et al. 2011; Benthem et al. 2017). Experiments on eco-evolutionary dynamics can be very useful in addition to long-term field observations, as experiments allow for manipulating and tracking ecological and evolutionary processes (Yoshida et al. 2003; Becks et al. 2012; Turcotte et al. 2013). Experiments not only strongly test causality, but can help us understand how ecological and evolutionary processes influence each other.

Various experimental studies have now shown how density-dependent selection can result in an eco-evolutionary feedback loop (Strauss 2014), both within species (Turcotte et al. 2011b; Turcotte et al. 2013) and between species (Yoshida et al. 2003; Becks et al. 2012). These, as well as other studies on eco-evolutionary feedback loops, have often focused on population size as a response variable (Hairston et al. 2005; Ellner et al. 2011; but see Pelletier et al. 2007; Cameron et al. 2013; Bassar et al. 2015). Those changes in population size however are the direct result of changes in vital rates, i.e. age- or stage-specific survival, growth and reproduction. In other words, eco-evolutionary dynamic effects on population growth occur through effects on vital rates.

Eco-evolutionary studies that looked at single vital rates exist, but these have generally not considered their integrated effect on population fitness, or assumed one vital rate to be an appropriate proxy for fitness (e.g. Matthews et al. 2016). Population fitness is not equally sensitive to all vital rates (Caswell 1978), and changes in one vital rate can be coupled with (opposite) changes in other vital rates (for instance through trade-offs) (Stearns 1989). Therefore, studies quantifying population fitness should ideally integrate over all vital rates (Metcalf and Pavard 2007). For instance, Cameron et al. (2013) showed that evolution led to higher population growth rates due to increased fecundity, while survival remained unchanged. Estimating these individual vital rates gives a more mechanistic insight in the processes underlying eco-evolutionary dynamics. Moreover, it can help us to better understand whether eco-evolutionary dynamics operate through similar demographic mechanisms across species and systems, and determine to what extent eco-evolutionary dynamics are repeatable.

The lack of information on the vital rates through which eco-evolutionary dynamics operate is, at least in part, because it can be difficult to collect demographic data on individuals embedded within a population. This is especially true for the short-lived species (e.g. zooplankton) that

are typically used in multiple generation studies, as those individuals cannot easily be marked or recognized. One solution is to remove individuals from the population and measure performance on isolated individuals (Cameron et al. 2013). A drawback of this approach is, however, that density-dependency in vital rates is ignored (Fowler 1981; Bassar et al. 2010a). Alternatively, an interesting possibility is to use data on changes in both population size and age/stage structure over time. As those changes are the direct result of the individual vital rates, they contain information on individual survival, growth and reproduction and have been used to infer these rates. Previous studies applying this 'inverse' modelling, have estimated demographic rates for a broad range of different species, such as sea lions (Wielgus et al. 2008), blue rockfish and gopher rockfish (White et al. 2016), tropical palm species (Cropper Jr et al. 2012), the perennial plant *Cryptantha flava* (González et al. 2016), tulip trees (Ghosh et al. 2012) and aphids (Gross et al. 2002). One major difficulty with the inverse estimation of individual vital rates is that many combinations of individual rates can theoretically result in the same population-level observations (Wood 1994). Another complicating factor is that the true underlying demographic model is often unknown, which makes it challenging to decide on the functional form of the underlying vital rates, and on which covariates to include. Therefore, some prior knowledge on the biology of the system is required, for instance some demographic rates must be known beforehand (González et al. 2016).

In this study, we estimate the demographic changes in vital rates and investigate how they contribute to the eco-evolutionary dynamics observed in the green peach aphid (*Myzus persicae*) (Turcotte et al. 2011b; Turcotte et al. 2011a; Turcotte et al. 2013). We analyze the unpublished stage-structure population dynamics data of a field experiment (Turcotte et al. 2011b) as well as unpublished results from an individual life table experiment. In this field experiment, rapid evolution significantly altered concurrent population dynamics (Turcotte et al. 2011b). The dynamics of replicated single-clone populations were compared to potentially evolving populations (consisting of two clones) over the course of a month, approximately 3-5 generations. Rapid evolution was observed and quantified as changes in the frequency of genotypes. Rapid evolution increased exponential population growth rates by 33% to 42%, compared to non-evolving controls, when populations were exposed to herbivores, predators and competitors. Additionally, results suggested that population density had differential fitness effects on competing clones, implying possible two-way eco-evolutionary dynamics between density (an ecological feature) and evolution.

In order to gain a more mechanistic understanding of the eco-evolutionary processes shaping the aphid populations we here aim to answer the following five questions: 1) Which vital rates underlie the differences in intrinsic growth rate among aphid clones? 2) Can we detect trade-offs between clones, for instance in survival and reproduction? 3) What is the impact of the changes in population density on vital rates of the three clones? 4) What demographic mechanisms evolved leading to more rapid growth in evolving populations compared to controls? Finally, 5) to what degree can we understand the evolutionary response in evolving populations, based on the vital rates of single clones?

5.3 Materials and Methods

Experimental design

We used data from two different experiments. In both experiments, three aphid clonal lineages were used (which we refer to as 'A', 'B' and 'C'), which differ in intrinsic growth rate (Turcotte et al. 2011b). First, we used data from a field experiment on the effects of ecological context and evolution on population dynamics. The three clones were tested individually, and in each combination ('AB','BC', 'AC'), allowing for evolution (by clonal selection) to occur. At the start of the experiment, 20 third-instar aphids (i.e. 20 individuals of one clone, or 10 individuals from

each of two clones) were placed on a caged host plant (mustard; *Hirschfeldia incana*). For half of the populations, the cages were removed at day 13, allowing competitors, predators and pollinators to access the plants, resulting in a strong reduction in plant sizes compared to the caged plants (Appendix D.1). In total, this resulted in 12 treatments (6 clonal treatments, fully crossed with the caging treatment), which were replicated 8 times. Populations were followed over 36 days (Appendix D.2). Every three or four days, the number of first/second, third, fourth/fifth instar and winged individuals were counted. Additionally, on these days, the number of leaves, which used as a proxy for plant size, was counted. Plant sizes were not recorded daily; to predict daily plant size, which was implemented in the model, we used smooth functions, fitted per plant separately based on generalized additive models (Appendix D.1). We excluded data from day 36, as aphid populations crashed due to plant senescence. More details on the included clones, experimental design and data collection can be found in Turcotte et al. (2011b).

Second, we used individual aphid life table data, not published previously, which were collected during a greenhouse experiment. All aphids were maintained as clonal colonies on *H. incana* in the same greenhouse. For the experiment, on each host plant of *H. incana*, four clip cages were attached, each containing two adult female aphids. In each cage, once an offspring was born the adults were discarded. This individual was followed during its complete life and moved to a fresh leaf when leaves turned yellow. Any offspring produced were counted and removed from the cage approximately every two days. An average of 15.5 aphids was tested in this manner for each clone. These individual-level data on life span, development and reproductive output were, in combination with the field experimental data, used to estimate daily survival, growth and reproduction, as explained below.

Modelling framework

Changes over time in the number of individuals in each stage were used to estimate demographic rates (survival, growth and reproduction). To do so, we defined three stages: 1) first/second instar aphids, 2) third instar aphids and 3) fourth/fifth instar and winged aphids combined. Daily changes in population structure from time t to time $t+1$, were described by a 3×3 transition matrix $\mathbf{A}$.

$$\mathbf{A} = \begin{bmatrix} \sigma_1(1-\gamma_1) & 0 & \phi \\ \sigma_1\gamma_1 & \sigma_2(1-\gamma_2) & 0 \\ 0 & \sigma_2\gamma_2 & \sigma_3 \end{bmatrix} \tag{5.1}$$

Matrix $\mathbf{A}$ describes all daily probabilities of moving from stage i at time t, to stage j at time $t+1$, and depends on six vital rates: stage-specific survival rates σ_{1-3}, probabilities of moving to the next stage γ_{1-2}, and daily reproduction ϕ. These six vital rates were described by underlying functions, which could include various predictors. We made the assumption that each vital rate is influenced by the same predictors (except for stage-dependent effects, see below), but that the strength and direction of effects can differ. Each vital rate was determined by a linear function including one or more coefficients and predictors, resulting in the linear predictor $\hat{y}$, which was related to the response variable by an appropriate link function. A log link function was used for reproduction ($\phi = \exp(\hat{y})$) and a logit link function was used for survival and growth (e.g. $\sigma = 1/(1+\exp(-\hat{y}))$).

Model fitting and selection

We first tested a set of seven 'basic' models that included various covariates to explore which single covariate resulted in the best model fit, based on cross-validation. These covariates were:

stage-dependency in growth or survival, caging, population size, plant size and population density. See Appendix D.3 for details; see Table 5.1 for an overview of the tested models. Next, we used the structure of the best (i.e. highest cross-validation score) model to construct a final model, which was the model including aphid density (number of individuals leaf^{-1}). In this model, each vital rate was a function of density, aphid treatment and caging. We included caging, because visual inspection of density trends through time (Appendix D.4) revealed effects of caging which remained after correcting for plant size. Caging effects were estimated per aphid treatment, to allow different responses to cage removal among treatments. For each vital rate, this resulted in:

$$\hat{y} = \beta_0 + \beta_1 D + \beta_{1+i} T + \beta_{7+i} CT_i \qquad (5.2)$$

where D indicates density and T_i indicates aphid treatment i. C is a dummy variable that is either 0 (caged conditions) or 1 (uncaged conditions). From here, we refer to this model as the 'full model'. The estimates from the full model were used to perform various analyses to quantify the demographic differences between the different experimental treatments, as described in *Population level effects of clonal identity and evolution*. All models were fitted in a Bayesian framework, implemented in JAGS software using the R-package *rjags* (Plummer 2016) (details in Appendix D.5).

		C	A	B	AB	AC	BC	Mean
Model 0	$\hat{y} = \beta_0$	1104	2273	1633	1899	1276	987	1529
Model 1	$\hat{y}_s = \beta_{0_s} + \beta_{1_s} x$	1102	2276	1634	1873	1252	974	1518
Model 2	$\hat{y}_g = \beta_{0_g} + \beta_{1_g} x$	1104	2283	1640	1886	1256	971	1523
Model 3	$\hat{y} = \beta_0 + \beta_1 C$	1136	2363	1678	1957	1271	1034	1573
Model 4	$\hat{y} = \beta_0 + \beta_1 N$	1120	1901	1634	1695	1357	960	1444
Model 5	$\hat{y} = \beta_0 + \beta_1 P$	930	2736	5771	1708	1327	894	2228
Model 6	$\hat{y} = \beta_0 + \beta_1 D$	982	1871	1111	1551	1252	864	**1272**

Table 5.1: Root of the mean squared error of out-of-sample prediction for different basic models, for each aphid treatment. Models include different predictors, and thereby differ in how the vital rates were calculated. Variable x is developmental stage, C is caging, N is population size, P is plant size and D is density. Predicted $\hat{y}$ is translated to daily survival (σ), growth (γ) and reproduction (ϕ) using the relevant link function. See Appendix D.3 for more details. Model 6 resulted in the best model overall and was used as a basis for the full model.

Prior distributions and likelihood

We used vague priors for all coefficients for both the basic models (Table 5.1) and the full model (Eq. 5.2). To compare population-level observations with predictions, the likelihood was calculated in accordance with González et al. (2016); to optimize the stage distribution, we used a multinomial distribution:

$$\mathbf{p}(t) \sim Multinom[\hat{N}(t), \hat{\mathbf{p}}(t)] \qquad (5.3)$$

Here, $\mathbf{p}(t)$ is a vector containing the observed proportions of individuals stage 1-3 at day t and $\hat{\mathbf{p}}(t)$ are the predicted proportions. Total predicted population size is given by $\hat{N}(t)$. To compare the total estimated and observed population size, we used a Poisson distribution:

$$N(t) \sim Pois[\hat{N}(t)] \qquad (5.4)$$

where $N(t)$ is the total observed population size.

Both $\hat{\mathbf{p}}(t)$ and $\hat{N}(t)$ were predicted by the following procedure: we started with the observed population structure at the previous measurement day. Given matrix $\mathbf{A}$, calculated with parameters θ and using the functions described in Eq. 5.2 and Table 5.1 (with the relevant link function), we projected population structure one day later by multiplying the observed population structure with $\mathbf{A}_\theta$:

$$\mathbf{n}(t+1) = \begin{bmatrix} n_1(t+1) \\ n_2(t+1) \\ n_3(t+1) \end{bmatrix} = \mathbf{A}_\theta \cdot \begin{bmatrix} n_1(t) \\ n_2(t) \\ n_3(t) \end{bmatrix} \tag{5.5}$$

The resulting population structure was multiplied with $\mathbf{A}_\theta$ again, for a total of Δt times, where Δt indicates the time interval between measurements (either three or four days). Finally, $\mathbf{n}(t+\Delta t)$ was divided by its sum ($\hat{N}(t)$), obtaining $\hat{\mathbf{p}}(t)$. Relevant in the case of the population, density and plant size-dependent vital rates (basic models 4-6 and full model; Table 5.1), we recalculated $\mathbf{A}_\theta$ every time step, taking into account the population size and/or plant size on each day. This approach enabled us to estimate the daily transition matrix $\mathbf{A}$, even though observations were on a three- or four-day interval.

For the full model, we additionally compared life table data on individual performance, to the predicted individual survival, growth and reproduction rates. The life span of a total of 46 individuals (15 or 16 individuals for each clone A-C) was recorded during a greenhouse experiment and was on average 24 days. We calculated the predicted survival probability when density was set at 1 individual per leaf, as this is in agreement with the life table experimental conditions. Each observed life span was then compared to the predicted daily mortality probability (calculated as one minus the daily survival probability) using an exponential distribution. For reproduction, we included daily reproduction rates for individuals from the day they started reproducing and onwards. On average, daily reproductive output of adult individuals equalled 2.2 and ranged between 0 and 7. These 635 observations on numbers of offspring were compared to the predicted reproduction when density set at 1 using a Poisson distribution. Finally, we used 45 observations on the day of maturation; on average individuals first reproduced when they were 11.4 days old. Translating this to the population matrix shown in Eq. 5.1, this implies that individuals reach stage 3 after on average 11.4 days. The predicted growth when density set at 1 ($\gamma(D=1)$) was used to calculate the expected time before first reaching stage 3 (i.e. the mean first passage time), conditional on survival, as: $1+2/\gamma(D=1)$. We compared this expected time to the observed individual maturation times using a gamma distribution, in which we estimated both the shape and rate parameter.

Note that, although we used the individual life table data to inform the model on survival, growth and reproduction rates, we purposely did not use clone-specific life table data to estimate effects of clonal treatment, but instead combined all data. This was done in order to estimate the clonal treatment effects based on only the population-level data.

Population level effects of clonal identity and evolution

Using the median of the posterior distributions for each estimated parameter, we projected transition matrices for each treatment, for densities ranging between 0 and the 95% quantile per caging treatment (4274 and 2100 individuals leaf^{-1} for the caged and uncaged conditions, respectively). Average density was 1024 and 416 individuals leaf^{-1} for the caged and uncaged conditions, respectively. For each matrix, asymptotic population growth rate was computed, which is the dominant eigenvalue. The matrices were used for subsequent analyses. Note that, as all vital rates

were density-dependent, the resulting matrices and thus the obtained population growth rates were density-dependent.

First, we compared the three pure clones to evaluate how clonal differences in vital rates led to differences in population growth rate. To do so, we used Life Table Response Experiments (LTREs; Caswell 1989). An LTRE decomposes differences in population growth rate into the contribution of differences in each underlying matrix element or vital rate. As we were interested in vital rate differences between treatments, we quantified the effects of vital rate differences, and not matrix elements, on population growth rates. We created a matrix for the 'average' pure clone using the average of each of the estimated clone-specific parameters, from which we obtained asymptotic 'reference' density-dependent population growth rate. Here, we first applied the relevant link function for the parameters describing survival, growth and reproduction to get averages on the response scale. For each clone, we then replaced one of the averaged vital rates by the clone-specific vital rate, and recalculated population growth rate. The difference in growth rate between the reference growth rate and the growth rate in which one of the vital rates is replaced by a clone-specific vital rate, quantifies the population-level effects of clonal differences in each of the vital rates. This analysis was repeated for each density, and both for the caged and uncaged treatment.

Second, we quantified the effects of evolution across densities, following a similar procedure. As a reference matrix, we calculated the average matrix over each combination of two clones, by using averaged vital rates. This reflects the 'expected' transition matrix, when both populations occur at a constant frequency of 50%, which represents a non-evolving population. We then replaced one of the vital rates by the vital rate of the corresponding mixed population, and calculated the difference between the reference population growth rate and the population growth rate in which the vital rate is replaced. This was done for each of the three mixed populations, for all densities, and both for the caged and uncaged treatment. To quantify uncertainty in the population-level effects of clonal differences and of evolution, the above analyses were repeated 1000 times with coefficients randomly obtained from the posterior distributions of each parameter.

Predicting population dynamics in evolving populations based on pure clones

The above analyses were based on asymptotic measures of (density-dependent) fitness, i.e. assuming a stabilized stage structure. Here, we were interested in quantifying the importance of various processes leading to differences in transient daily population growth rates of the evolving populations compared to the pure clone populations, using population structures observed during the experiment.

The following steps were repeated for each observed population structure of the evolving populations. We projected population size one time interval (three days) later based on the estimated vital rates and the observed plant size for the corresponding evolution treatment, and considered this to be the 'true' reference population size prediction at $t + 3$. We then quantified to what extent we could predict these true population sizes based on 1) the dynamics of the pure clones, 2) the observed plant sizes, 3) changing clone frequencies and 4) vital rate type-specific changes in the evolving populations. We did so in a cumulative fashion, stepwise adding each of these factors. For 1) we averaged vital rates and day-specific plant sizes of the relevant pure clones and projected population size at $t + 3$. We started from the same population structure and size, but implemented the average plant size, resulting in a different density. This reflects the expected dynamics of a non-evolving population (in which both clones occur at a constant frequency of 50%), the same plant size in the evolving and non-evolving populations, and no interactions in vital rates among clones. These three factors could all contribute to differences between the predicted true population size and the predictions of a non-evolving population. In

2), we tested the importance of plant size. Population dynamics were projected based on the non-evolving averaged vital rates of the pure clones, but instead of using mean plant size from the pure clones, the actual observed day-specific plant size from the evolving population was included to calculate the density at time t. In 3) we added changing frequencies (evolution). We no longer assumed a constant frequency, but implemented the observed genotype frequencies of both clones, for a given day (Appendix D.6). We calculated average vital rates weighted by the frequency of each of the clones and used these to predict population size at $t + 3$. Finally in 4), we tested for the presence of interactions among clones resulting in changed vital rates. Survival, growth and reproduction (weighted averages from the pure clones) were one by one replaced by the estimated vital rate of the evolution treatment, and again population dynamics were projected.

For all scenarios, we calculated growth rates by dividing population sizes at day $t + 3$ by population size at day t and translated these values to daily population growth rates. We calculated the proportion of variance explained by each of the scenarios, to assess the predictability in transient population dynamics of the evolving populations. This analysis was done separately for the caged and uncaged treatments. For the uncaged treatments, we excluded projections from day 0-13, as cage removal took place on day 13.

5.4   Results

Simulated data

We tested our inverse modelling approach with simulated data, for which the true relationships were known. We simulated data according to three different basic modelling structures (model 0, 1 and 4; see Table 5.1). In general, the correct model resulted in the lowest mean error, in particular for the density-dependent model (Table 5.2), and parameter estimates were accurate and unbiased overall (Fig. 5.1). See Appendix D.7 for details.

	Model 0	Model 1	Model 4
Model 0	51	49	62
Model 1	100	104	117
Model 4	729	690	77

Table 5.2: Model selection results for simulated data sets. Three modelling structures (rows) were used to simulate ten independent data sets per model, each with randomly drawn parameters (see Appendix D.7). Noise was added to the parameters, and three replicates were simulated. Each of the three models (columns) were then fitted to data on two replicates. Root mean squared errors were calculated as the difference between predictions and observations for the third replicate, which was not used for fitting. Values show the root of th mean squared error across the ten simulations. Values on diagonal indicate model performance when the true model is fitted to the data. Modelling structures correspond to models shown in Table 5.1. Model 0 includes only intercepts, model 1 includes stage effects for survival, and model 4 includes population size.

Model selection and estimated coefficients

When fitting the seven basic models for each treatment separately, the density-dependent model (model 6) resulted in the highest predictive ability (lowest root mean squared error; Table 5.1). We thus continued with density to construct the full model which was fitted using the full dataset and included an additional effect of caging, estimated per aphid treatment. This final model resulted in accurate predictions of numbers of individuals in each stage (r^2=0.89). Throughout the rest of the *Results*, we present analyses using the estimates of this final model, assuming that these are
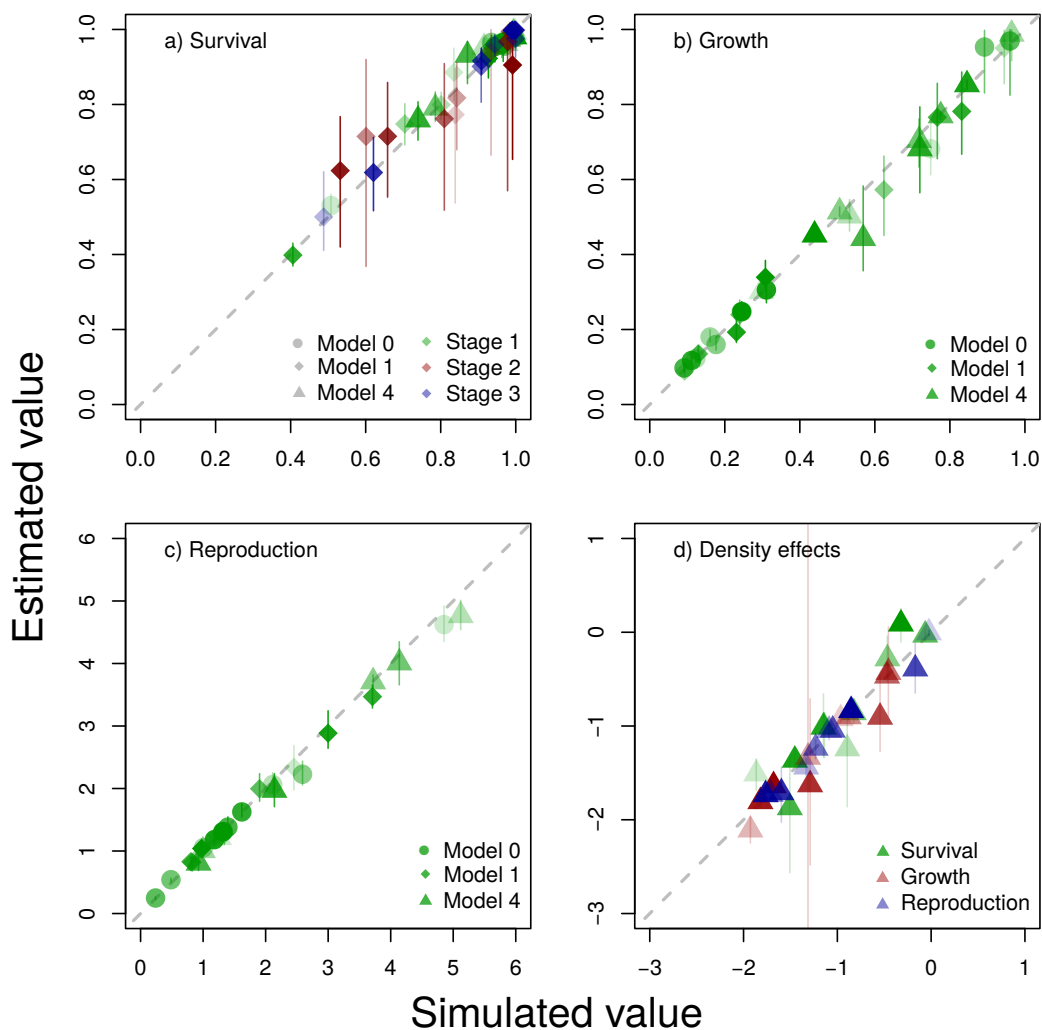
Figure 5.1: Simulated parameters against estimated parameters for a) survival, b) growth, c) reproduction and d) density dependence. Different symbols indicate different simulated modelling structures (Table 5.1), and levels of transparency indicate different simulations. Error bars show 95% credible intervals. Model 0 includes only intercepts, model 1 includes stage effects for survival, and model 4 includes population size.

indeed correct estimates of the individual rates. In the discussion, we address some limitations and uncertainties that are associated with this inverse modelling approach.

Predicted daily survival probability in the caged populations, at average density (across all observations; 815 aphids leaf^{-1}), ranged between 0.87 and 0.97, and increased with density (Fig. 5.2a). Survival estimates were significantly higher for clone B compared to the other aphid treatments. Average daily probabilities of moving to the next stage (growth) for caged populations ranged between 0.39 and 0.65 and decreased with density (Fig. 5.2b). Finally, daily reproduction when caged ranged between 0.89 and 1.70 and decreased with density (Fig. 5.2c). Results suggest that both growth and reproduction were strongly decreased in the uncaged populations, in all aphid treatments (open dots in Fig. 5.2). In contrast, survival was increased, implying a survival
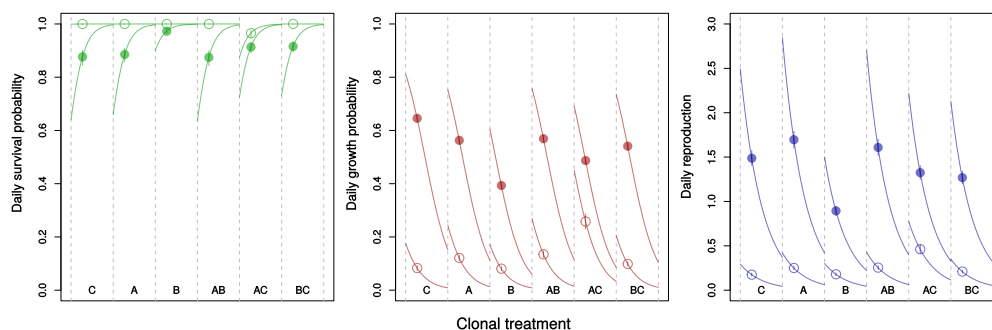
Figure 5.2: Vital rate estimates for the six aphid treatments based on the final full model. Error bars show 95% credible intervals of the estimates due to uncertainty in aphid treatment effects. Colours indicate different vital rates (green: survival, red: growth and blue: reproduction). Dots show the estimates at average density (calculated across all observations) for the caged (closed dots) and uncaged treatments (open dots). Lines show the effects of density (aphids per leaf), ranging between zero (left) to one standard deviation (right) above the average.

probability of practically 1 under uncaged conditions (Fig. 5.2a). These findings are supported by a visual comparison of the observed population structures for both the caged and uncaged populations (Appendix D.8): the proportion of individuals in stage 1 is lower in the uncaged populations, while the proportion of individuals in stage 2 and 3 is higher. This proportional decrease of stage 1 individuals is in agreement with a reduced reproduction, and the higher proportion of individuals in later stages is in agreement with higher survival. The estimated parameters were used for the subsequent analyses, in which we combined the vital rates to construct transition matrices (according to Eq. 5.1).

Vital rates underlying population-level differences among pure clones

In this section, we present the results for the caged treatments; see Appendix D.9 for the results of the uncaged treatments. Projected density-dependent population growth rates decreased with density, after an initial increase for clone A and C (Fig. 5.3a). Clone B had the highest population growth rate only at the lowest densities (Fig. 5.3a), but shows the strongest negative effect of density. This results in the lowest growth rates overall. Clone A, in contrast, generally shows the highest growth rates, although the difference with clone C diminishes at higher densities. This is mostly in line with the observed trends: although clone B has higher population sizes for most of the time compared to clone C (Appendix D.2), when correcting for plant size, clone B reaches lower densities (Appendix D.4). Finally, population growth rates of the uncaged populations were greatly reduced (Appendix D.9), despite a higher survival probability, which is in agreement with the observed trends (Appendix D.4).

Results of the LTRE suggest that the lower population growth rate of clone B (for a given density) is caused by the lower reproduction and slower development (Fig. 5.3b-d). This effect is partly counterbalanced by increased survival. Clone A has a slightly higher population growth rate due to a significant benefit related to reproduction (Fig. 5.3d). These negatively correlated findings suggest trade-offs between growth, survival and reproduction, as no clone benefits from increases in each vital rate.
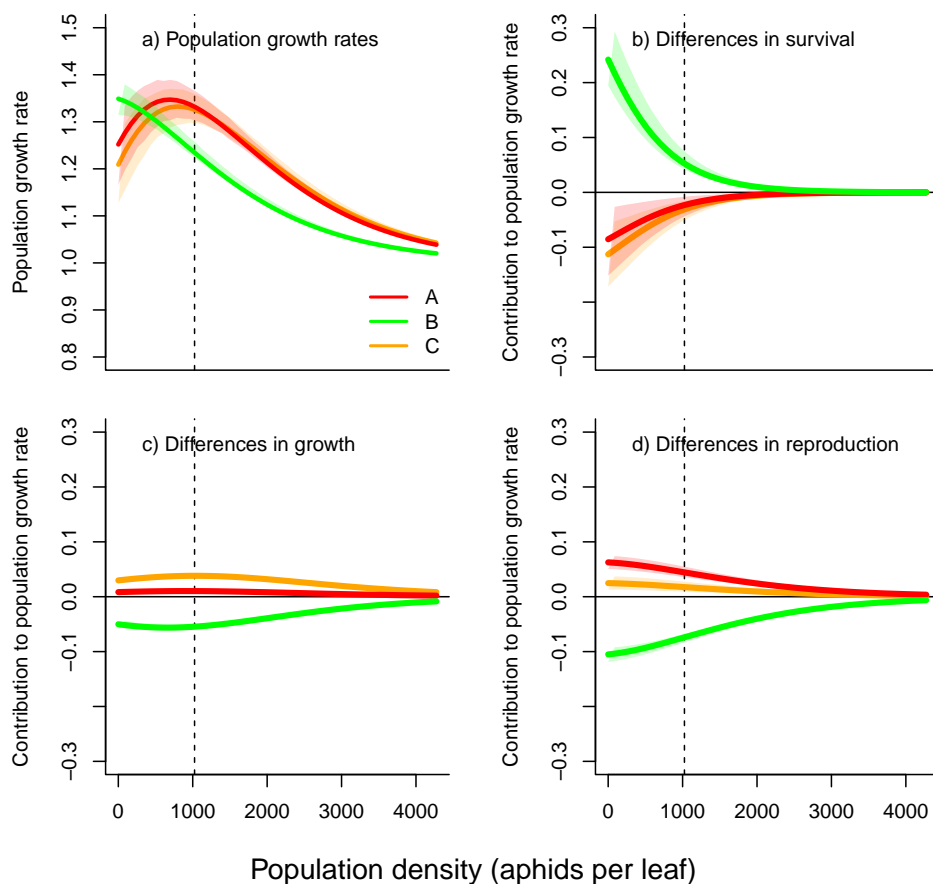
Figure 5.3: a) Projected population growth rate as a function of density for the three pure clones, under caged conditions. b-d) Life Table Response Experiment comparing different clones, as a function of density (individuals per leaf). Clone A-C were compared to the average matrix across the three clones. Densities range between 0 and the 95% quantile of observed densities, under caged conditions. Different colours indicate different clones and vertical lines indicate the average density under caged conditions. Shaded polygons show 95% confidence intervals in the predictions, obtained by simulating 1000 transition matrices by drawing coefficients from the posterior distributions of the clonal effects.

Vital rates underlying population-level evolutionary effects

Comparing population growth rates of the evolving populations with the expected population growth rate when both clones occur at a frequency of 0.5, complex interactions with density are found for the caged treatments (Fig. 5.4a). At uncaged conditions, population growth rate was higher in all evolving populations across all densities (Fig. 5.4b). At both caged and uncaged conditions, population growth rate of treatment BC is higher than the mean growth rate of B and C. The same applies for treatment AB under uncaged conditions, and at higher densities when caged.
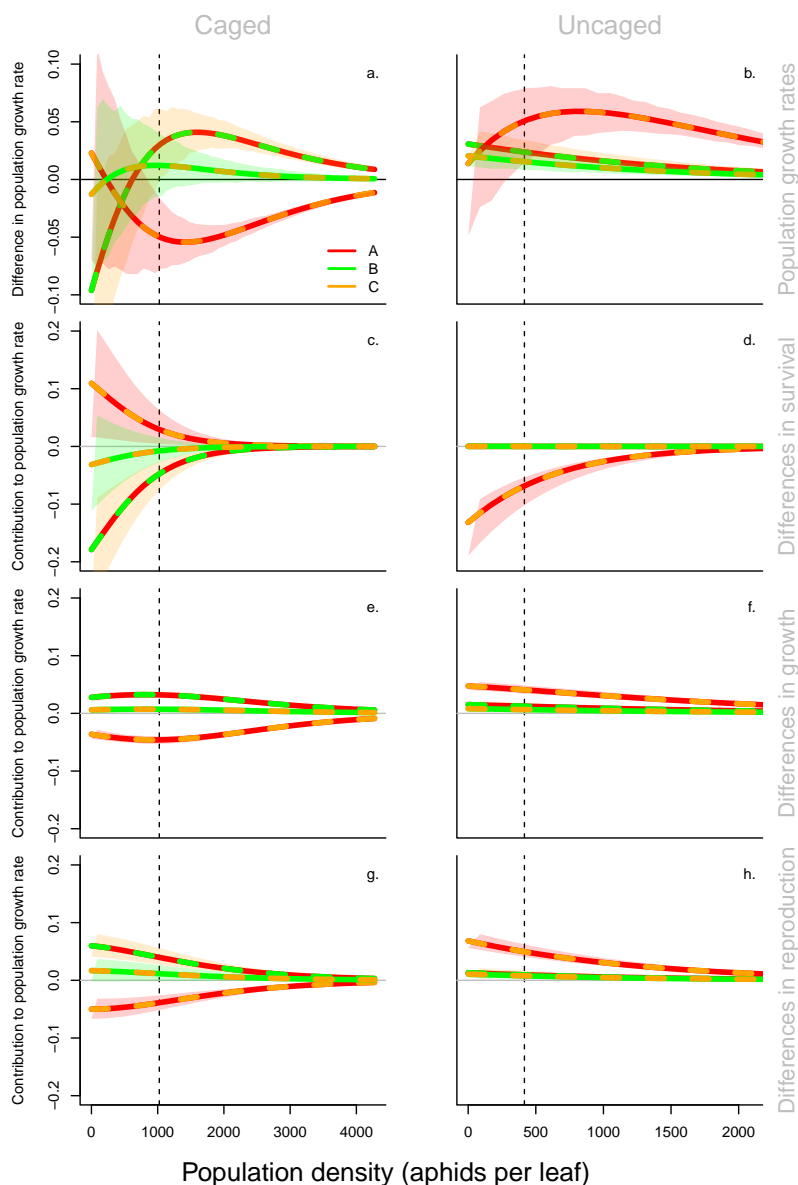
Figure 5.4: a-b) Differences in population growth rate between the evolving population and the expected population growth rate calculated as the average of the two pure clones (i.e. at a 1:1 ratio), under caged (left) and uncaged (right) conditions. Values above zero imply that population growth rates of the evolving population is higher than expected based on the pure clones. c-h) Life Table Response Experiment, comparing population matrices of each mixed population to the average matrix of the two pure clones (i.e. at a 1:1 ratio): contribution of difference in c-d) survival, d-f) growth and g-h) reproduction. Positive values indicate a higher population growth rate in the evolving population due to differences in either survival, growth or reproduction. Different colours represent different combinations of pure clones. Densities range between 0 and the 95% quantile of observed densities, under caged (left) or uncaged (right) conditions. Vertical lines indicate the average density, under either caged or uncaged conditions conditions. Shaded polygons indicate 95% confidence intervals in the predictions, obtained by simulating 1000 transition matrices by drawing coefficients from the posterior distributions of the clonal effects.
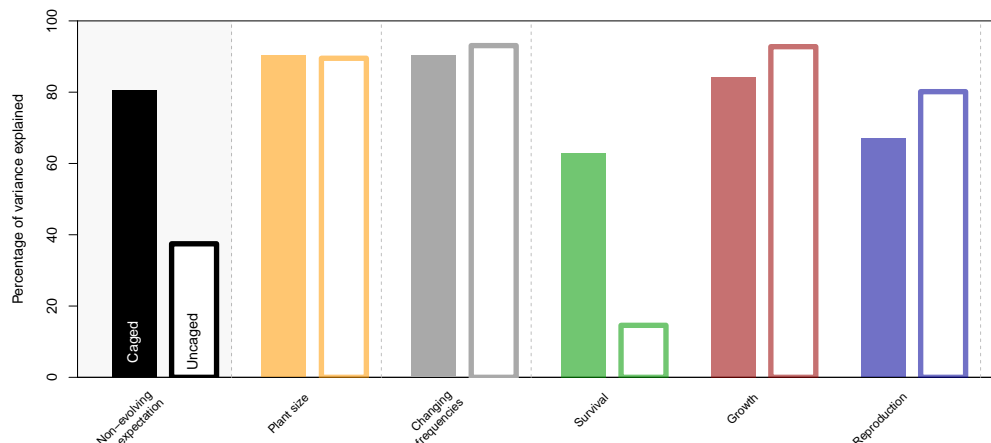
Figure 5.5: Variance explained in the transient population growth rates (over three days) of the evolving populations. Black bars show the explained variance when assuming a non-evolving population, in which the vital rates equaled the average vital rates of the two pure clones. Orange bars show the explained variance when plant size is replaced by the day-specific plant size of the mixed population. Using this as a starting point, we quantified the effect of changing frequencies, instead of assuming a constant frequency of 0.5 (grey bars). Green, red and blue bars show the proportion of variance explained when each of the averaged (weighted by the frequency) vital rates is replaced by the vital rate of the mixed population. Replacing all three vital rates at the same time results in the reference model and hence a 100% of the variance explained. For the caged populations, we used each observed population structure from day 0 until day 31; for the uncaged conditions we used each observed population from day 17 until day 31.

Treatment AC results in higher population growth rates than expected only at uncaged conditions. Results of the LTRE suggest that the higher growth rates are mostly due to benefits related to faster development (growth and reproduction), at both caged and uncaged conditions (Fig. 5.4c-h).

Predicting the dynamics in the evolving populations

When assuming a non-evolving population, in which vital rates and plant dynamics equal the 50-50 average of the two pure clones, 80% and 37% of the variance in daily population growth rates in the evolving treatments can be explained, for the caged and uncaged conditions, respectively (Fig. 5.5). For the uncaged conditions, this proportion greatly increases when including observed treatment-specific plant sizes instead of the averaged plant size at a certain point in time (orange bars). When allowing clonal frequencies to change through time, R^2 increases from 89.5% to 93% in the uncaged conditions, but not for the caged populations (grey bars). Finally, separately replacing each of the averaged vital rates by the treatment-specific vital rates, did not improve the predictability, suggesting that the entire life history is evolving in the mixed populations instead of isolated vital rates (Fig. 5.5).

5.5 Discussion

The main goal of this study was to gain a more mechanistic understanding of the eco-evolutionary processes shaping aphid populations, by quantifying how clones differ in individual growth,

survival and reproduction and how these differences contribute to responses of evolving populations. Our density-dependent population models show clear intraspecific variation in the degree of density-dependence (Fig. 5.3a), which is in agreement with the aphid study by Agrawal et al. (2004). According to our results, clone B showed the strongest negative response to density, resulting in the competitive strength of clone B being highest only at very low densities. Clone A had the highest fitness at intermediate densities, and clones A and C are equally fit at high densities. The novelty of our study is that we additionally aimed to assess which vital rates caused the variation in density-dependence and which vital rates were altered in the evolving populations. Modelling results suggest that density negatively affected growth and reproduction (Fig. 5.2), but, on a population-level, that interactions with density were mainly driven by survival differences between clones, with survival differences becoming smaller at higher densities (Figs 5.2, 5.3). These genotype-specific density effects are in accordance with experimental evidence for the presence of a full eco-evolutionary feedback loop, as was found in a follow-up experiment by Turcotte et al. (2013), by showing that initial density affected the rate and direction of evolution, and that evolution altered population growth.

Comparing density-dependent population growth rate of each of the potentially evolving populations with the expected growth rate (for that given density) of a non-evolving mixed populations, imply an effect of evolutionary dynamics on ecological dynamics, in particular under uncaged conditions (Fig. 5.4). For all three combinations, an overall increase in population growth rate was found across all densities. Increases were due to both increased growth and reproduction, despite a decrease in survival for combination AC. Under caged conditions, patterns were less straightforward, although for two out of three combinations, population growth rate of the evolving population was higher across most densities (Fig. 5.4). Studies quantifying the importance of evolutionary versus ecological factors have found varying results, with the importance of evolution practically ranging between 0% and 100% (Ellner et al. 2011; Becks et al. 2012; Govaert et al. 2016), depending on the system and evaluated response variable. For the aphid populations under caged conditions, transient dynamics of the mixed populations were well predicted by assuming a non-evolving population in which both clones occur at a 0.5 frequency, which is in agreement with the finding that evolution did not affect the growth of caged aphid populations (Turcotte et al. 2011b). In contrast, under uncaged conditions, dynamics were relatively poorly predicted when assuming a non-evolving population. Results show that here, daily per capita growth rate was in the first place shaped by changes in plant size, as including plant size increased the explained variation in transient population growth rates from 37% to 90% (Fig. 5.5). It is perhaps not surprising that plant size, as a proxy for available resources, shapes population dynamics to a large extent.

We then quantified the importance of evolution, in the form of changing clonal frequencies, which had no effect under caged conditions and only a 3.5% increase in explained variance under uncaged conditions. This supports earlier findings that evolution influences population dynamics under uncaged conditions (Turcotte et al. 2011b), although the effect is very small compared to the effects of plant size. Evolution is expected to have larger impacts on population growth through time, when clonal frequencies start to deviate further from starting distributions. We would thus predict that the importance of evolution increases with time. However, given the short duration of the experiment, these temporal effects are difficult to assess, in particular for the uncaged conditions as cages were removed only at day 13. Given that the duration of the uncaged conditions was only $\sim$ 2-3 generations, the 3.5% increase in explanatory power due to evolution could suggest that evolutionary dynamics potentially play an important role in shaping ecological dynamics over the longer term. Future experiments will be required to test this further.

Second, we looked at the importance of evolutionary dynamics leading to changed interactions among clones, which can also lead to increased performance, due to for instance resource

partitioning or facilitation. It is widely known that grasslands with higher plant species richness show increased productivity (Tilman et al. 1996); however, also within a species, performance can increase with increasing genetic diversity (reviewed in Hughes et al. 2008). For example, in springtail populations, various life history traits improve with genetic richness (Ellers et al. 2011). We found some evidence for interactive effects on vital rates, as 7% of the variation in transient growth rates remained unexplained after taking into account plant size and evolution. This suggests non-additive effects of combined genotypes, although it could also (partly) reflect uncertainty in the estimates or perhaps changes in the composition in winged versus unwinged adults, which both fall into the same stage in our analysis. We were unable to pinpoint a specific vital rate that explained the remaining variation all by itself (Fig. 5.5). This could suggest that these non-additive effects of genotypic diversity do not necessarily operate through the same demographic rates, even within the same species, but future experiments will have to test this. We also note that differences in vital rates between the pure and mixed populations were generally small (Fig. 5.2), so we might also lack the power to detect these interactive effects, if present at all.

Caging and density

Our findings suggest that density is the foremost important factor determining daily population growth rates (Table 5.1, Figs 5.3, 5.4), making it critical to include plant size in the analysis (Fig. 5.5). Results indicate negative density dependence in population growth, as was already found for these populations by Turcotte et al. (2011b), in other aphid populations (Breton and Addicott 1992; Agrawal et al. 2004), as well as across other taxonomic groups (Fowler 1981). With our approach, we were now able to quantify through which vital rates population growth rate decreased with density. Results suggest that this is due to reduced reproduction and growth, which is mostly in line with earlier studies, on for example *Daphnia* (Guisande 1993; Goser and Ratte 1994) and soil mites (Ozgul et al. 2012). More surprising is the apparent positive relation between survival and density. According to our results, daily survival probability ranges between 0.88 and 0.97 at average density in the caged populations, but becomes >0.99 at one standard deviation above the average density. Positive correlations between population density and survival have also been observed in *Daphnia* (Bruijning et al. 2018b) and some developmental stages of soil mites (Ozgul et al. 2012). It could be that populations reach higher densities, because individuals survive better, leading to a positive correlation between survival and density.

We also noted a positive effect of cage removal on survival (Fig. 5.2). Individual growth and reproduction are strongly decreased, while survival approaches 100%. While this may seem strange at first, we have two likely explanations for this pattern. First, cages were removed only at day 13 and in half of the populations. We therefore have relatively little data, from a short time period, on the uncaged dynamics (compared to the caged dynamics). During the two weeks of uncaged dynamics, it could -in principle- have been that almost all individuals survived. Second, higher survival in uncaged conditions might reflect the same pattern as the found positive effects of density: aphids under uncaged conditions experienced a larger (interspecific) density, due to the presence of competitors. If experienced density indeed reduces reproduction and growth but increases survival, as the estimated density effects suggest, it is perhaps not surprising that interspecific density of competitors has the same effects.

Reduced population growth rates in the uncaged populations, are partly due to the smaller plant sizes. These effects are captured by the inclusion of density (population size corrected for available resources) as a covariate in the analyses. However, even after correcting for density, uncaged populations reach smaller population densities, suggesting additional effects of the cage removal (Appendix D.4). Based on our results, we predict that this is not so much because of predation, as this would lead to a decrease in survival. Instead, it could be due to competition for

resources by other herbivores, such as other aphid species.

Matrix model parameterization with inverse modelling

It has been shown that estimates of individual rates based on stage-frequency data can be very sensitive to the chosen model structure (Manly and Seyb 1989). By first exploring which single covariate resulted in the largest model improvement, and by doing so for each clonal treatment separately, we have attempted to find the model structure that is most likely to represent the true dynamics. We show that including density (number of individuals leaf^{-1}) resulted in a major model improvement in most clones, suggesting a strong support for this covariate. We note, however, that we made the simplifying assumption that each vital rate is affected by the same predictors, which does not necessarily have to be the case. Moreover, we considered only one type of life cycle (Eq. 5.1), which seemed realistic for our study species as was also confirmed by the individual life table data. Finally, we were able to inform the model on the parameters making use of the life table data, as was suggested in David et al. (2010). Including these life table data did not affect the estimated treatment effects, but helped to centre the coefficients and making the model identifiable.

Whether the full model including effects of density, caging and treatment indeed captures the true observed dynamics, is of course unknown. It could be that the model can be further improved by including other (unknown) covariates, interactions, non-linear effects, and/or different structures for different vital rates. However, as the final model explained 89% of the variation in one-time interval changes in stage-specific population numbers, this gives confidence in that the most important factors influencing dynamics of the aphid populations were included. In addition, we used simulated data to test the reliability of our methods. Estimates of the simulations give confidence in the identifiability of the model as parameters can, in principle, be estimated accurate and unbiased (Fig. 5.1). In addition, results of these simulations show that, if density indeed is important, a model ignoring density will lead to a large error, and that density is not likely to be erroneously selected (Table 5.2). However, the simulations contained little noise and a known modelling structure, while the experimental data were noisier, both within and between treatments.

Moreover, although the above points make that we believe that plant size was indeed an important factor shaping the dynamics in this system, differences among clonal treatments (which was the main focus of this study) were much more subtle than the effects of plant size. Indeed, most of the estimated vital rates did not differ significantly between pure clones, nor between evolving versus non-evolving treatments (Fig. 5.2), possibly indicating a lack of power to detect differences, if any. Whether or not the few observed vital rate differences among clonal treatments, indeed reflect biological differences in life history traits, can only be confirmed by collecting the required individual-level data. Finally, whether there are multiple combinations of parameters resulting in a similar fit (i.e. multiple optima), and if so, to which extent the results and conclusions will prove robust, remains to be investigated.

The inverse estimation of transition matrices obviously comes with challenges and measuring the individual rates directly (on individuals embedded in the population) is preferred. However, for small-sized species often used in this type of experimental studies, such as zooplankton (Van Doorslaer et al. 2009a), mites (Cameron et al. 2013) and aphids, it is difficult to follow individuals within their population. This in contrast to studies on e.g. mammals, birds or fish (Grant and Grant 2002; Pelletier et al. 2007; Bassar et al. 2010a; Traill et al. 2014), where it is common practice to mark individuals in order to obtain demographic data. One solution is to measure individual rates on sampled individuals/genotypes, held in isolation (Van Doorslaer et al. 2009a; Cameron et al. 2013). A drawback is that density-dependent effects will be overlooked, while these are known to impact population dynamics, as also underscored by our results. Alternatively, individuals can be isolated within their population to measure individual rates during a short interval (Bruijning et al.

2018b). However, if these individual data are not available, we show that estimating individual rates based on stage-frequency data can provide useful insights in how ecological and evolutionary dynamics shape populations. Moreover, it results in predictions on individual vital rates, which can subsequently be tested by collecting the relevant data. This will further inform us on the reliability, robustness and opportunities of inverse modelling to estimate the individual vital rates underlying changes in population dynamics.
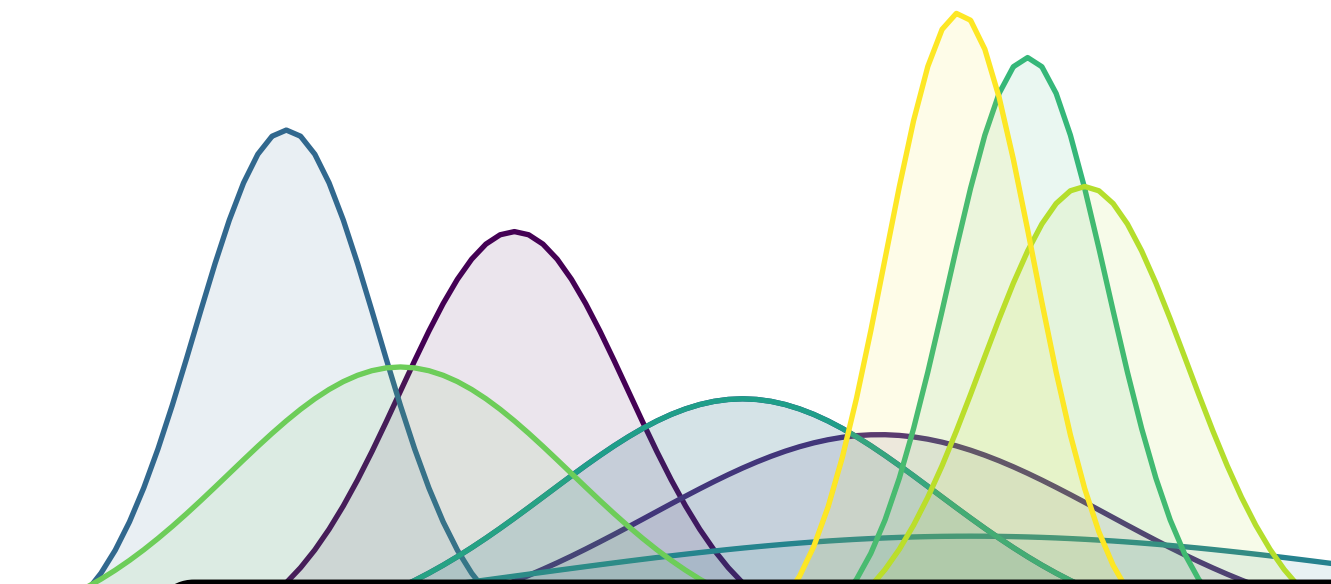
Acknowledgements

Marjolein Bruijning[1], *C. Jessica E. Metcalf*[2], *Eelke Jongejans*[1] *and Julien Ayroles*[2]

1. Department of Animal Ecology and Physiology, Radboud University, 6500 GL Nijmegen, The Netherlands.
2. Princeton University, Department of Ecology and Evolutionary Biology, 08540 Princeton, New Jersey, United States of America.

Fitness-consequences of intra-genotypic variability

6.1 Abstract

Intra-genotypic variability leads to phenotypic variation between genetically identical individuals, even under constant environmental conditions. Although it has long been known that intra-genotypic variability can be adaptive, well-described empirical examples remain very rare. We argue that the time is ripe to develop formal expectations for the distribution of adaptive intra-genotypic variability across species and environments. Doing so, we take advantage of both theoretical advances, as well as increasingly well-resolved descriptions of the genetic underpinnings of this phenomenon. We review empirical evidence for the presence of intra-genotypic variability and its genetic basis, and synthesize theoretical expectations to encompass non-linear averaging and bet hedging. We illustrate that the optimal intra-genotypic variance in one fitness component depends on the environmental variance and the life history of a species. We conclude by discussing ways to bridge the gap between empirical evidence of intra-genotypic variability, studies demonstrating its genetic component, and evidence that it is adaptive.

6.2 Introduction

Few environments on this planet are stable. Most organisms experience short-term fluctuations (such as seasonal variation, or stochastic fluctuations in temperature, rainfall and nutritional resources) as well as longer-term trends (such as those driven by climate change). This environmental variation drives eco-evolutionary dynamics that contribute to maintaining genetic variation between individuals, and shape phenotypic change. In turn, the fate of populations is determined in part by whether they harbour sufficient phenotypic variation to adapt to changing environments (Lande 1976). Understanding the causes and consequences of the propensity of phenotypes to vary between individuals is a central problem in biology.

Perhaps the most frequently discussed source of phenotypic variation is genetic variation between individuals. Where sufficient genetic variation exists, a shift in the genetic makeup of the population will allow populations to persist in new conditions. This process can result in adaptive tracking, where the phenotypic mean of a population tracks a fitness optimum. Alternatively, phenotypic plasticity may allow populations to keep pace with environmental change, as specific genotypes modulate their phenotype in response to experienced environmental conditions (Piersma and Drent 2003). These two components of population persistence in variable environments have been extensively studied both theoretically and empirically. An alternative strategy has received far less attention: changes in the level of phenotypic variability (rather than the mean) may be a key element in allowing populations to cope with changing environments. The evolution of intra-genotypic variability, phenotypic variation observed between genetically identical individuals reared in a constant environment, leads to phenotypic variation even in a constant environment (Bradshaw 1965). This is in contrast to phenotypic plasticity, in which phenotypic variation is correlated with the environment (Fig. 6.1a-b). A life history strategy that maintains intra-genotypic variation can guarantee that at least some individuals are well suited to any environmental conditions (Philippi and Seger 1989). Beyond genetic variation and phenotypic plasticity, intra-genotypic variability can thus be an important third axis by which population growth can be maintained across changing environmental conditions, and is our focus here. We start by reviewing empirical evidence for intra-genotypic variability. We then delineate theoretical expectations for the evolution of intra-genotypic variability across an array of life history and environmental contexts, while considering the genetic effects on both the mean and the variance. This is followed by a discussion relating empirical observations to theoretical expectations. We conclude by defining a research agenda designed to bridge the gap between empirical evidence of intra-genotypic variability and evidence that it is adaptive.

6.3   What is known about intra-genotypic variability?

Although there are surprisingly few studies investigating the contribution of intra-genotypic variability to phenotypic variation between individuals, various lines of evidence indicate that intra-genotypic variability is widespread. It emerges in artificial breeding programs as a nuisance to farmers who seek uniformity in their crops or livestock, e.g. observed in birth weight of rabbits (Garreau et al. 2008), mice (Formoso-Rafferty et al. 2016) and rainbow trout (Sae-Lim et al. 2015). It has also been commonly observed in experimental model systems such as flies, yeast or *C. elegans* (Acar et al. 2008; Diaz and Viney 2014; Ayroles et al. 2015). Finally, it has been observed in natural populations of plants, typically focusing on seed banking (Childs et al. 2010; Gremer et al. 2016), mammals (e.g. in docility behaviour in yellow-bellied marmots; Martin et al. 2017) and birds (in fledgling weight in great tits; Mulder et al. 2016), to list a few. Intra-genotypic variability is believed to be due to differences in the individual's ability to buffer micro-environmental perturbations, which lead to stochastic fluctuations in development. Referring to this plausible mechanism, intra-genotypic variability is also called 'micro-environmental plasticity' in the literature (Morgante et al. 2015).

For intra-genotypic variability to be of relevance for evolutionary ecology, it must be under genetic control. There is now ample evidence that this is the case, and heritabilities for intra-genotypic variability are reviewed by Hill and Mulder (2010). Moreover, recently, it has become possible to map individual loci that regulate the degree of variability of a trait, rather than its mean (Ansel et al. 2008; Rönnegård and Valdar 2012; Rönnegård and Valdar 2011). Such loci have been mapped in a wide range of organisms and for a variety of traits, for example locomotor handedness (Ayroles et al. 2015) and bristle numbers (Mackay and Lyman 2005) in *Drosophila*, flowering
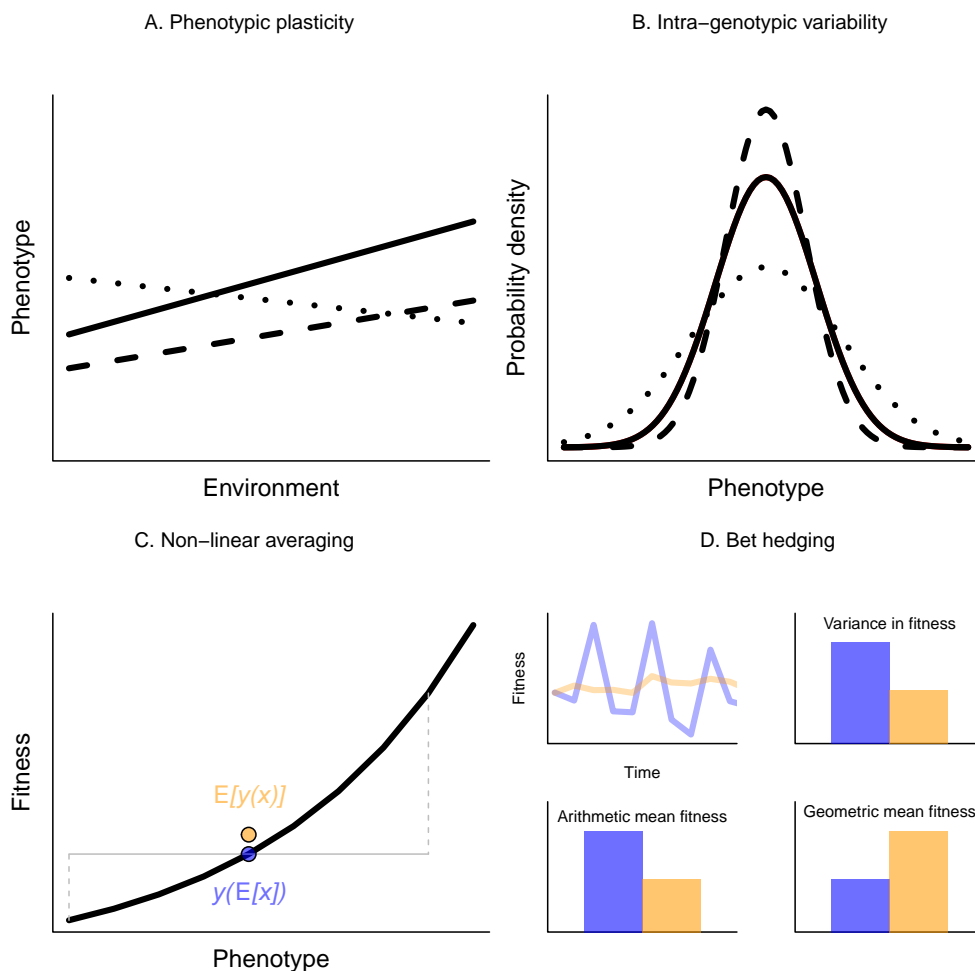
Figure 6.1: Schematic overview of the different concepts related to phenotypic variation. In a) and b), different lines represent three different genotypes, that differ in their a) reaction norm, and b) environment-independent phenotypic variance. a) Phenotypic plasticity results in the expression of different phenotypes across environments, and involves a correlation between phenotype and the environment. b) Intra-genotypic variability results in a distribution of phenotypes within an environment. In c) and d), two conditions that can result in selection favouring intra-genotypic variability. c) Relation between phenotype x and fitness component (e.g. fecundity). If the relation is convex, an increase in the phenotypic variance results in a higher expected fitness (orange) then the fitness of the average phenotype is (blue), due to non-linear averaging. d) Compared to a non bet-hedging strategy (blue), a bet-hedging strategy (orange) reduces the variance in fitness. Despite a decrease in arithmetic mean fitness, geometric mean fitness is increased, and thus expected to be favoured by natural selection.

time in *Arabidopsis* (Jimenez-Gomez et al. 2011; Shen et al. 2012), morphological traits in maize (Ordas et al. 2008), and calcium excretion in rats (Perry et al. 2012). The number of loci associated with variance may be on par with those associated with the mean (Jimenez-Gomez et al. 2011; Morgante et al. 2015) and there can be little overlap between loci affecting the mean of a trait and its variance (Ayroles et al. 2015; Morgante et al. 2015), but see Ros et al. (2004).

The emerging conclusion is that variance control is ubiquitous. However, observing heritable variation for variability tells us nothing about whether it is adaptive or not, just that it can evolve. It is largely unknown whether alleles controlling variance are maintained by neutral or adaptive processes. How selection might drive the evolution of these loci in natural systems, and more generally, how intra-genotypic variability links to fitness remains poorly characterized (Hill 2007; Viney and Reece 2013).

6.4   When do we expect selection for increased intra-genotypic variability?

While it is possible that variability represents maladaptive deviation from a developmental ideal, and the fitness of the genotype is reduced because of individual variability, it is also clear that the magnitude of intra-genotypic variability may be tuned by natural selection. An array of results from basic theory can be used to delineate the range of contexts in which intra-genotypic variability can evolve.

We begin with an important result termed 'Jensen's inequality' (Jensen 1906). The relationship between traits and fitness is often nonlinear. If the shape of the relationship is accelerating (convex), this means that trait values some magnitude larger than the mean will result in large increases in fitness, while trait values equivalently smaller than the mean will result in smaller reductions in fitness. On average, then, if the trait mean is fixed, fitness can nonetheless be increased by increasing trait variance, as a result of non-linear averaging (Fig. 6.1c). The consequences of non-linear averaging have been described with respect to species coexistence (Hart et al. 2016), and it may also be important in the evolution of variance control (Koons et al. 2009). For example, Mulder et al. (2016) found that stabilizing selection acts on variance in fledging weight in great tits due to the sigmoidal curve describing the relationship between fledging weight and recruitment.

Moving from constant environments, to environments that vary (spatially or temporally), a next core theoretical driver of the evolution of intra-genotypic variability is bet hedging (which can act in tandem with non-linear averaging; Rees et al. 2004). Bet hedging is formally defined as a strategy that leads to a reduction in fitness variance, thereby increasing geometric mean fitness, at the cost of arithmetic mean fitness (Philippi and Seger 1989) (Fig. 6.1d). In other words, while year to year, on average, a bet-hedging strategy might produce fewer offspring, over the long term, its fitness will be larger. Germination delays in annual plants provide a classic example of diversifying bet-hedging, which is our focus here. A non bet-hedging strategy, where every year all seeds present germinate, leaves itself vulnerable to years of difficult environmental conditions where no germinants survive, and thus the population is at risk of extinction. A bet-hedging strategy guarantees some seeds will have delayed germination, and thus persist in the seedback during harsh years. Identifying bet hedging requires quantifying the fitness consequences in terms of arithmetic mean and fitness variance, to properly distinguish between bet hedging and non-linear averaging (Olofsson et al. 2009; Rees et al. 2010; Ripa et al. 2010).

Work on bet hedging often evaluated its role for traits related to timing, for example, seed banking (Evans et al. 2007; Gremer and Venable 2014; Gremer et al. 2016) or diapause (Menu et al. 2000; Graham et al. 2014; Rajon et al. 2014; García-Roger et al. 2014; Furness et al. 2015; Tarazona et al. 2017). For these timing-related examples, the process underpinning bet hedging is

often framed as a probability (e.g., probability of germinating). This implicitly defines the core trait underpinning fitness as discrete (e.g., being a seed or a germinant), with variance emerging purely from the binomial variance associated with the probability (implying that variance is maximal for a probability $p = 0.5$, all else being equal). Interestingly, a large majority of both empirical and theoretical studies describing bet hedging in other traits (not related to timing) have also mostly considered discrete traits. For example, whether or not aphids develop wings (Grantham et al. 2016), left or right handedness in *Drosophila* (Ayroles et al. 2015), investment in spores or vegetative cells in slime molds (Martínez-García and Tarnita 2017), stochastic on- and off switching of gene expression in yeast cells (Acar et al. 2008) as well as an array of theoretical models (Botero et al. 2015; Crowley et al. 2016). On consideration, a clear reason emerges for the observation that studies of bet-hedging tend to predominantly focus on discrete traits: where the optimal trait value fluctuates unpredictably due to a changing environment, and there is no 'intermediate' trait value possible, it is beneficial to express both discrete traits every time step (Bull 1987). Bet-hedging has also been observed in continuous traits, for example thermal and phototactic preference in *Drosophila* (Kain et al. 2015), egg size in gypsy moths (Rossiter 1991) and macrophyte offspring size (Charpentier et al. 2012). Moreover, the processes underlying timing traits can also be reframed as continuous (Metcalf et al. 2015). For continuous traits, theory indicates that bet hedging can be beneficial in traits for which the phenotypic optimum varies between years, provided that the magnitude of environmental fluctuation is sufficiently large (Bull 1987).

In spite of striking advances in knowledge of the genetics of intra-genotypic variability, with a growing number of studies providing heritability estimates and mapping loci associated with variation in variance heterogeneity, these studies have rarely related the degree of variance to fitness, or predicted optimal expected variance (Simons 2011, for a review on studies specifically focusing on bet hedging). On the other hand, studies quantifying the fitness consequences of variance heterogeneity (Kain et al. 2015) or aiming to predict the optimal degree of variability for a given trait (Charpentier et al. 2012; Gremer and Venable 2014) have generally not looked at the heritable component of the mean and variance. Studies considering both sides of this issue are needed to understand when we expect intra-genotypic variation to evolve, but these are few and far between (Box 6.1).

Box 6.1 — The impact of genetic covariance between trait mean and trait variance. The analyses presented in Figs 6.3 and 6.4 define the adaptive landscape for the mean and variance of a phenotypic trait affecting reproduction. They, however, neglect the impact of the underlying genetic variances and covariances. Combining these two perspectives will shed light on an array of interesting research questions.
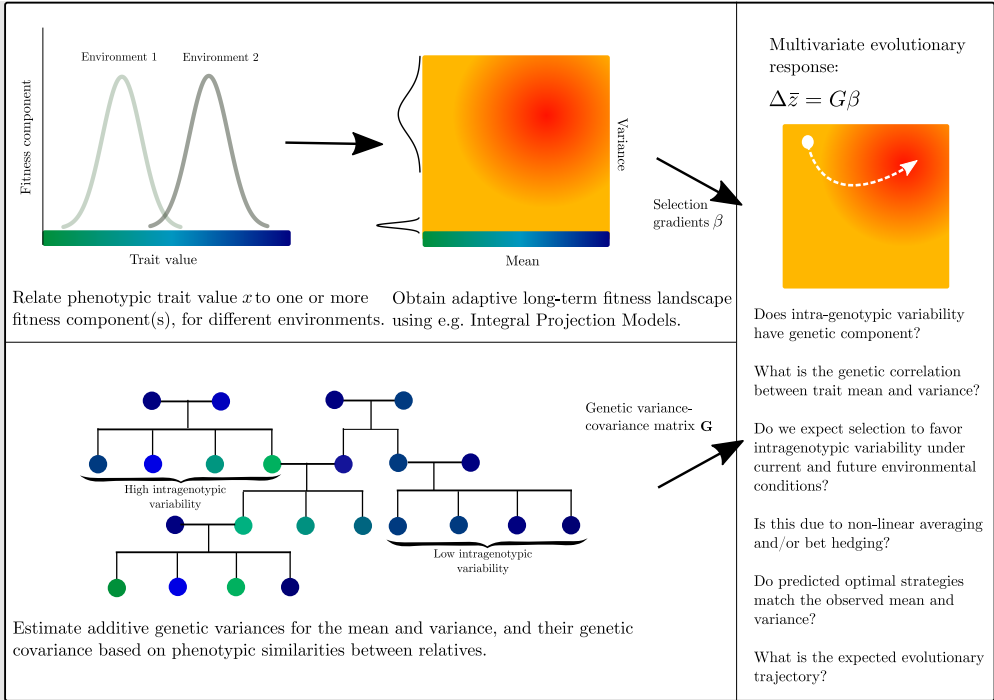
Proposed workflow to study the fitness consequences of intra-genotypic variability.

The evolutionary response for multiple traits can be written as the multivariate breeder's equation (Lande 1979), as the product of the genetic variance-covariance matrix **G** and the partial selection gradients for each trait. By using a demographic model to define the adaptive landscape, partial selection gradients at a given point in the landscape ($\delta\tilde{r}/\delta\overline{m}$ and $\delta\tilde{r}/\delta\overline{s}$) can be obtained. The **G**-matrix can be estimated using linear mixed effects models, based on knowledge on phenotypic similarities between relatives (Hill and Mulder 2010; Sae-Lim et al. 2015; Mulder et al. 2016; Martin et al. 2017). Changes in the **G**-matrix can have large effects on the expected evolutionary trajectory.
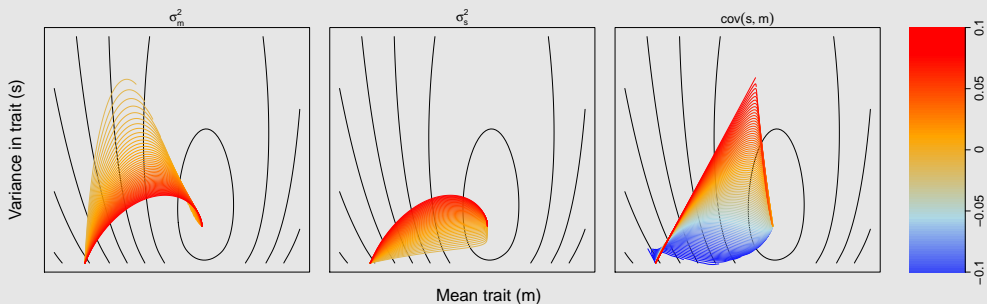
Evolutionary trajectories across a fitness landscape for a varying genetic variances and covariances, starting from a population with a mean trait value of -1.5, and a variance of 0.01. Fitness landscape corresponds to the fitness landscape obtained from Scenario C (the optimum relation between phenotype and reproduction, and survival set at 0.5), under a fluctuating environment (see *Modelling the evolution of variability* and Fig. 6.3). Additive genetic variances for the mean (σ_m^2) and variance (σ_s^2) in phenotypic

trait were both set at 0.1, and their covariance (cov(m,s)) at 0. We then, one by one, varied σ_m^2, σ_s^2, and cov(m,s), while keeping the rest of the **G**-matrix constant. We evaluated how average m and s change over the course of 200 time steps. Graph on the left shows the effect of varying values for σ_m^2, whereby different colours correspond to different values. Middle graph shows the effect of varying σ_s^2, and graph on the right the effect of changes in cov(m,s).   ∎

6.5 Modelling the evolution of variability

To explore the fitness consequences of intra-genotypic variability, we simulated the dynamics of a population consisting of individuals that are characterized by phenotypic trait x, which influences an individual's fitness via an interaction with the environment k. We used a full life-cycle approach where yearly individual fitness is the sum of its survival and reproduction, and use long-term population growth rate as a measure for fitness (Sæther and Engen 2015). An individual's trait x is affected by its genetic background and is determined at birth. This is consistent with observations in e.g. *Drosophila* (Ayroles et al. 2015) and *C. elegans* (Diaz and Viney 2014) where intra-genotypic variability (in locomotor handedness or mean lifetime fecundity, respectively) is observed across individuals, but the phenotype of each individual is stable over its lifespan. Since our focus is on intra-genotypic variability itself as an evolvable trait and the selective pressures it will experience in the context of different values of mean traits and their own selection pressures, we account for genetic effects on both the mean (m) and the variance (s) on trait x. The genetics of an individual determine the probability density distribution from which a phenotype x is drawn from (Appendix E.1 for more details). We present results in which trait x influences reproduction, because it is easier to titrate the impact of non-linearities on fitness outcomes for reproduction than for survival, which is naturally bounded between 0 and 1, and thus inherently non-linear. However, qualitative results are the same when focusing on survival.

We implemented four different functional forms describing relations between x, the environment k and reproduction, which we refer to as scenario A-D, while keeping survival probability constant at 0.5. First, a linear relationship (Fig. 6.2a), which might emerge where for example plant size is allometrically related to seed output, e.g., measured as the relation between vegetative and reproductive weight (Samson and Werk 1986; Klinkhamer et al. 1992). Second, an exponential relationship (Fig. 6.2b), a situation that can emerge under different measures of plant size (Klinkhamer et al. 1992), but might also be reflected in e.g. body size and egg number in mites (Ozgul et al. 2012) or antler size and lifetime breeding success in red deer (Kruuk et al. 2002). Third, an intermediate optimum phenotype, e.g. body size in zooplankton (Lynch 1977) or macrophyte offspring size (Charpentier et al. 2012) (Fig. 6.2c). Finally fourth, a discrete (threshold) relationship between trait x and reproduction (Fig. 6.2d), for instance the development of spores or vegetative cells in slime molds (Martínez-García and Tarnita 2017), coat colour in wolves (Coulson et al. 2011) or germination in plants (Evans et al. 2007). The environment modulates these relationships and this is captured by an interaction term (solid and dotted lines in Fig. 6.2a-d). All scenarios were evaluated for both a stable environment and a fluctuating environment randomly drawn from normal distribution (Fig. 6.2e-f) (Appendix E.1 for more details).

Our analysis required accounting for fitness given a life history shaped by a continuous underlying trait, x. We used an Integral Projection Model (IPM; Box 6.2) to calculate long-term population growth rates for different combinations of m and s (that result in different distributions of x), for each of the four scenarios and environmental conditions (Fig. 6.2). This resulted in long-term fitness landscapes for combinations of m and s values, allowing us to identify optimal strategies (Appendix E.2 for more details). We assumed asexual reproduction and a heritability of 1, and thus did not include for instance maternal or environmental effects on either the mean or variance. Furthermore, note that we purposefully only included relations between phenotype

x and environment *k* as described above. There is thus no interaction between the genotype and the environment, to not obscure the evolution of intra-genotypic variability with effects due to plasticity.
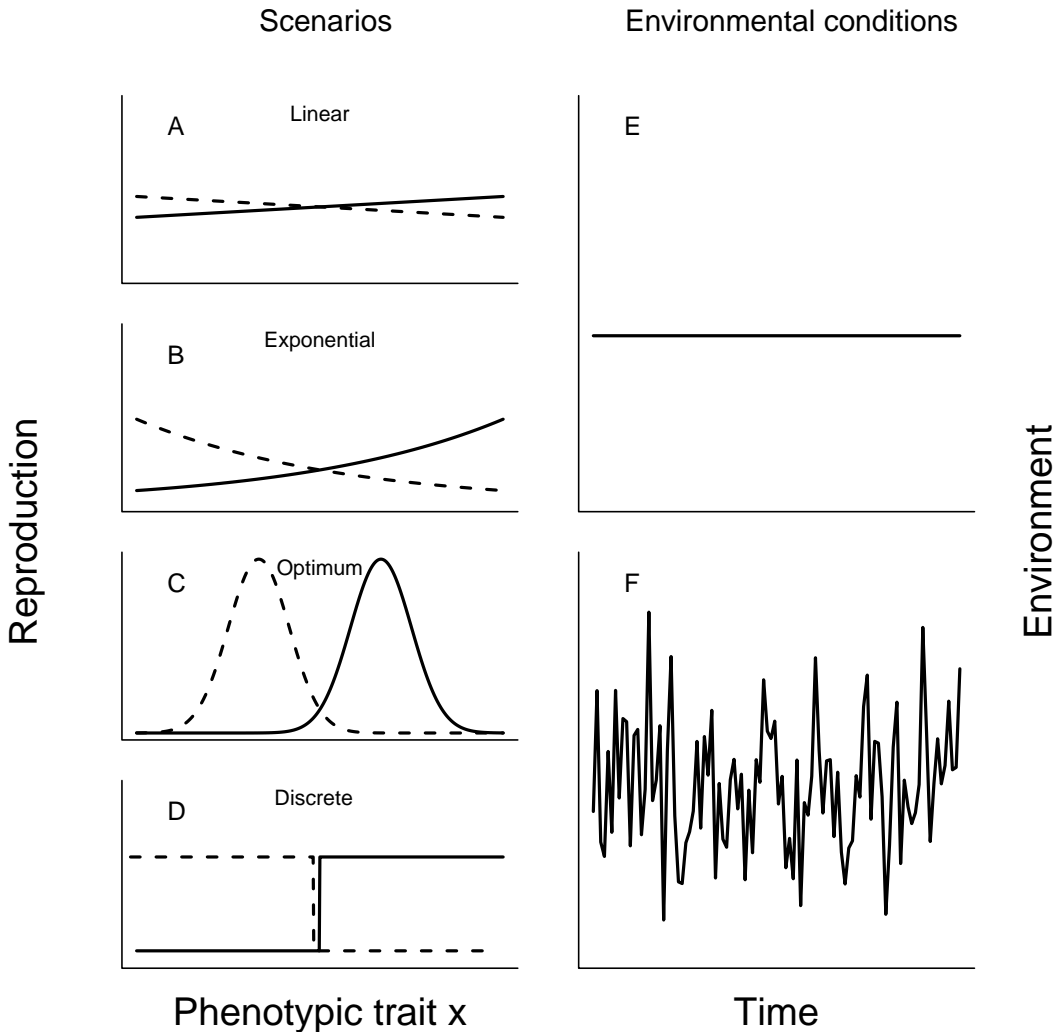

Figure 6.2: Overview of different simulations. In a-d) different relations between phenotypic trait *x* and reproduction are shown for two different environments *k*. Axes are the same in a-d; on the x-axis, trait value *x* is shown, and reproductive output is shown on y-axis. Solid lines show relation for $k = 1$; dotted lines for $k = -1$. Relations shown in a-d are referred to in the text as scenario A-D. In e-f) two different environmental conditions (stable and randomly fluctuating environment). Time is on the x-axis, and the environment *k* is on the y-axis. The average environment is 0.5 in both e) and f), and the variance equals 0 and 1, in e) and f), respectively.

Box 6.2 — Integral Projection Models in Evolutionary Biology. Demographic models have been used to tackle evolutionary questions ranging from why senescence evolves (Hamilton 1966) to the age or size at which reproduction should occur (Metcalf et al. 2003). Typically, such analyses map the relationship between a focal trait (such as survival at any age, or onset of reproduction) and a fitness measure like the population rate of increase, λ, or the net reproduction number, R_0, which emerges from the demographic model. The mapping may be either direct, or via sensitivity analyses.

Within this context, integral projection models (IPMs) are an extension of matrix population models to reflect continuous classes such as body size (Easterling et al. 2000), rather than the discrete stage-class framing of a typical matrix population model. This characteristic allows more realistic reflection of ontogenetic trajectories that characterize many species. IPMs can also be easily modified to capture age as well as stage and size dynamics (Ellner and Rees 2006); and standard measures of fitness can be easily extracted, including extensions to temporally varying environmental contexts (Rees and Ellner 2009), and frequency or density dependent settings (Metcalf et al. 2008).

Typically, evolutionary analyses based around Integral Projection Models assume very simple genetics (Rees and Ellner 2016), in part because of limited knowledge for traits of interest. This is likely to be sufficient for analyses that aim to predict the long-term outcomes of evolution rather than short term transient changes. The latter may require a more nuanced framing of the genetic underpinnings (Chevin 2015; Janeiro et al. 2017), e.g., leveraging knowledge of the existing quantitative genetics context (Childs et al. 2016; Coulson et al. 2017). To date, IPM–based predictions of long-term evolutionary outcomes have been developed for i) size at flowering in many plants species (Rees et al. 2004; Kuss et al. 2008; Miller et al. 2012); ii) predictions of seed dormancy (Rees et al. 2006); and iii) predictions of twinning in sheep (Childs et al. 2011). ∎

6.6 Consequences of intra-genotypic variability shaped by phenotype relationship to fitness

The degree to which intra-genotypic variance is adaptive is affected by the functional form of the relation between phenotypes and fitness, and by the environmental conditions (Fig. 6.3; Table 6.1). There is no selection on the variance in the linear scenario, whether the environment is variable or not. If the trait-fitness relation is convex, variance is favoured due to non-linear averaging (and similarly, if the relation is concave, variance will be selected against), irrespective of the environmental conditions. Selection will favour intermediate phenotypic variance only if there is an optimum trait value (either continuous or discrete) that maximizes fitness and which changes through time. As environmental conditions are rarely stable, and scenarios C and D are common in nature, we expect that intra-genotypic variability can be the outcome of adaptive evolution across many traits and species, provided that variability is under genetic control, for which we find more and more evidence (e.g. Mackay and Lyman 2005; Ayroles et al. 2015; Morgante et al. 2015, see references in introduction). Note that even if selection does not act on variance itself (e.g. as in scenario A), we still expect evolution of variance control if genetic correlations between the mean and variance exists (Box 6.1).

	Scenario A: Linear relation between trait and reproduction	Scenario B: Exponential relation between trait and reproduction	Scenario C: Highest reproduction for intermediate trait values	Scenario D: A threshold relation between trait and reproduction
Constant environment	Directional selection acts on the mean trait value. There is no effect of variance on long-term population fitness, whether the environment is fluctuating or not. A linear relation between continuous traits and fitness alone cannot favour (of disfavour) an increase in variance, unless additional processes play a role, such as frequency dependence (Metcalf et al. 2015).	Selection on the mean is directional. There is also directional selection towards an increase in variance, even in the stable environment, as a result of non-linear averaging (Jensen's inequality).	Selection on the mean is stabilizing, and selection favours low variability.	Selection favours the expression of only one of the two trait values, which is best fit to the environment.
Fluctuating environment			Selection on the mean phenotype is stabilizing. Here we find selection for intermediate variance. Comparison of arithmetic and geometric mean shows that the variance resulting in the highest arithmetic mean is not the variance resulting in the highest long-term fitness (geometric mean), indicating that intermediate variance is favoured as a result of bet hedging.	Selection favours a combination of the expression of both traits when the environment fluctuates. This reflects bet hedging, as it comes at the cost of arithmetic mean fitness.

Table 6.1: Summary of patterns of selection pressures on intra-genotypic variance across a spectrum of environments and trait-reproduction relationships.
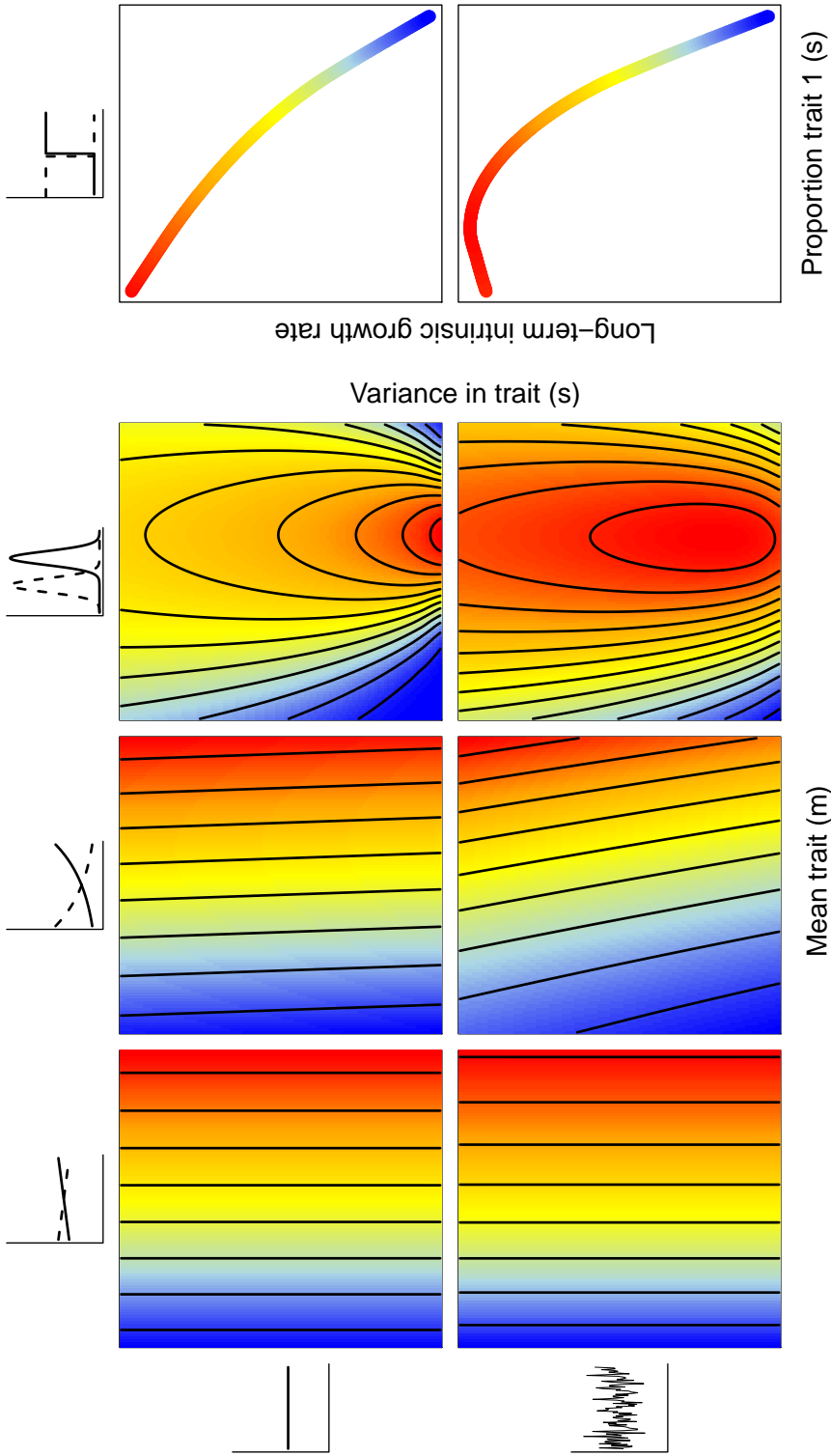
Figure 6.3: Long-term fitness landscapes for different combinations of mean trait (m) and variance in trait (s), that together define the probability density function where phenotypic trait x is drawn from. Upper row shows results for a stable environment, bottom row shows results for a fluctuating environment. Different columns show results for different relations between phenotype x and reproduction (scenarios A-D). From the left to the right: linear, exponential, optima and discrete relations. Red colours indicate higher long-term fitness, blue colours indicate lower values, and values are scaled for each graph separately. In the right column, s indicates the proportion of expressing trait 1 (x-axis), and is plotted against long-term fitness (y-axis).

6.7   Interactive effects of environmental variance and life span on the evolution of intra-genotypic variability

In the case of stabilizing selection on the variance (scenarios C and D; Fig. 6.3), the fitness consequences of variance depend also on the broader demographic context, in addition to relations between fitness, phenotype and the environment, and to the experienced environmental conditions. This is because fitness is the outcome of the integration across the life cycle (Metcalf and Pavard 2007), and other life stages will thus define selective pressures on intra-genotypic variation in traits affecting reproduction. To expand our investigation beyond our core analyses (in which survival was set at 0.5 and environmental variance at 1), we varied both survival and the environmental variance, and obtained the optimal variance for each combination using models proposed (Bull 1987) (see Appendix E.3 for more details).

Higher environmental variance, unsurprisingly, increases the optimal variance in both the discrete and optimum scenario, favouring bet hedging (Fig. 6.4). Interestingly, increasing survival probabilities decreased the optimal variance in both explored scenarios, C and D (Fig. 6.4). This demonstrates that modifying survival can alter the optimum variance in another, uncorrelated, fitness component (reproduction), that together determine population fitness. Furthermore, it implies that intra-genotypic variability, when affecting a trait related to reproduction, may be more adaptive in short-lived species (e.g. annual plants, *Drosophila*), where individuals are at a higher risk of dying without having reproduced. For long-lived life cycles (e.g. trees, blue whales), individuals can afford to have one year of zero reproduction, because a higher yearly fitness is guaranteed due to the high survival (Visser et al. 2011).

Using a different approach, Koons et al. (2009) came to essentially the same conclusion. They concluded that, in fluctuating environments, demographic lability will be favoured in short-lived species, while buffering is favoured in long-lived species. Results obtained by Koons et al. (2009) are due to non-linear averaging: by varying average survival on a logistic curve, it was shown that lability (i.e. variance) is beneficial when the average is low, while disadvantageous when the average survival is almost at its maximum of 1. This is because at low values, the logistic curve is convex, while at high values, the curve is concave. Our results support the idea that lability may be more adaptive in short-lived species. Moreover, we show that these patterns can not only be driven by non-linear averaging (and thus by the shape of the curve), but also by effects via an uncorrelated vital rate, that sets the lower boundary of yearly fitness and co-determines the contribution of reproduction to fitness. We conclude that in order to determine how intra-genotypic variability affects fitness, it is critical to consider the full life cycle, in addition to considering the trait-fitness relation, as was illustrated in the previous section.

6.8   Evolution of intra-genotypic variability in natural populations

Results of the demographic models suggest that, on a phenotypic level, non-zero variance is adaptive under commonly encountered environmental conditions and trait-phenotype relations (Figs 6.3, 6.4). This phenotypic variance may be due to stochastic variation in gene expression ('gene expression noise') (Raser and O'Shea 2005). Work on yeast cells has shown that the average gene expression noise differs between genes, and the degree of variability is suggested to be dependent on the gene function (Newman et al. 2006). For instance, essential genes tend to have low stochastic variance in expression (Fraser et al. 2004; Newman et al. 2006), while genes related to stress and plasma-membrane transporters show elevated noise levels (Bar-Even et al. 2006; Newman et al. 2006; Zhang et al. 2009). Whether increased gene expression noise is adaptive seems under debate, with some studies showing selection against variance-increasing alleles (Lehner 2008; Metzger et al. 2015), while others suggest fitness benefits in response to
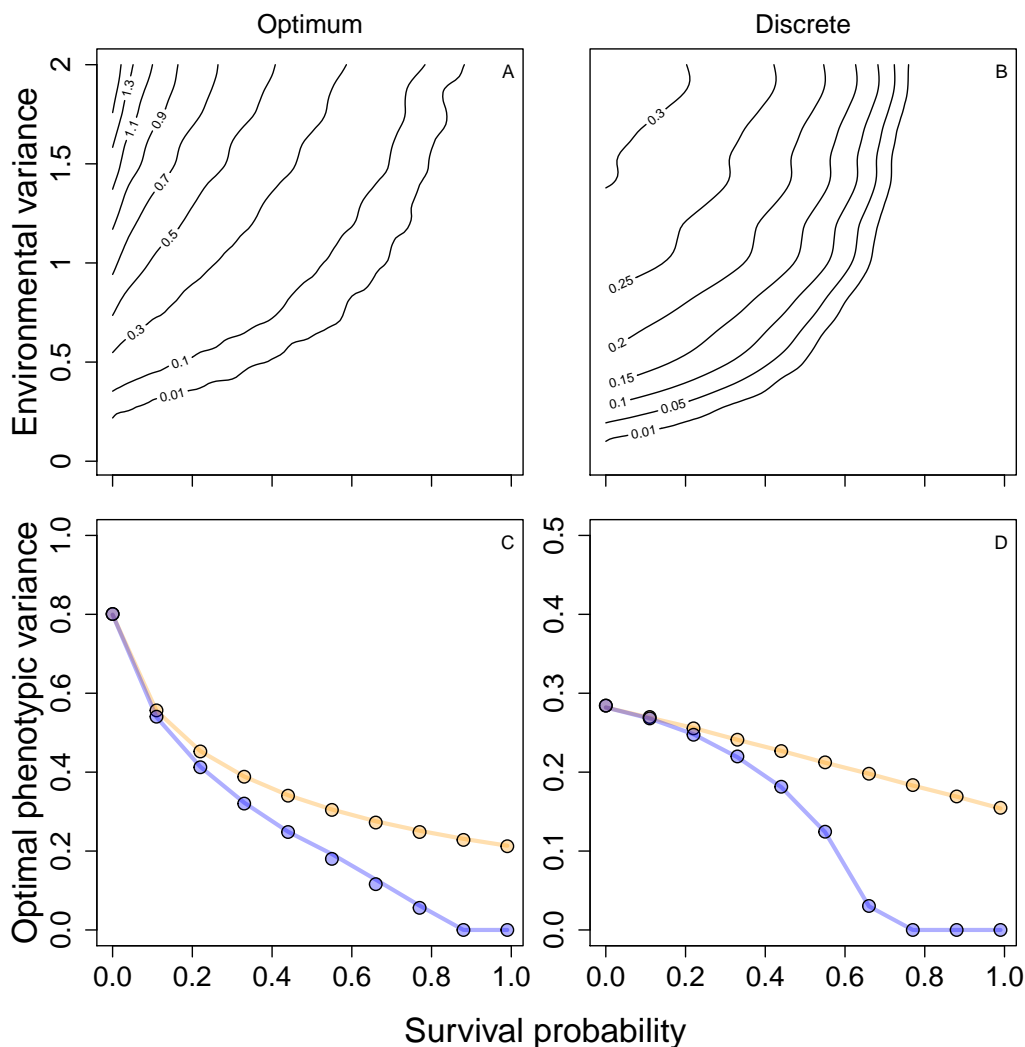
Figure 6.4: In a-b) optimal variance in phenotypic trait *x* for different combinations of survival probability and environmental variance. Here, reproduction is varied together with survival, in order to keep population growth rates constant. In c-d) comparison of a scenario in which reproduction is kept constant (orange lines), and environmental variance is set at 1. Effects were similar, although weaker, compared to a scenario in which reproduction is changed accordingly (blue lines). Dots represent optimal values as predicted by the model proposed by Bull (1987) which is essentially identical to the constructed IPM (See Appendix E.3 for more details). In a,c) results for scenario C, in b,d) results for scenario D.

environmental stress (Blake et al. 2006; Viney and Reece 2013). Optimal levels of expression noise will depend on the relation between gene expression and performance, following the same general principles as shown in our study, where we considered phenotypic traits at the level of the individual instead of gene expression at the cellular level. Determining these relations will help us to predict in which genes we expect stochastic variation in expression to be beneficial, under which conditions, and potentially help explaining why we observe gene expression noise in some genes, but less so in others. For example, for dosage-sensitive genes, in which changing expression levels is harmful, most genes indeed show low expression noise (Lehner 2008). Whenever relations between expression level and performance are convex, we expect variance in gene expression to be beneficial, as was also pointed out by Zhang et al. (2009). Further, whenever an unpredictable environment changes the optimum gene expression level, bet hedging might be a beneficial strategy. Such an optimum relation has for instance been observed between Lac protein expression and fitness of *Escherichia coli*, whereby the optimum changes with lactose level (Dekel and Alon 2005). Finally we note that, even without variation in gene expression noise, non-linearities in the genotype-to-phenotype map can result in differences in the degree of phenotypic variance between different genotypes (Green et al. 2017).

Evidence available to date suggests that heritabilities for intra-genotypic variability are quite low (Hill and Mulder 2010; Sae-Lim et al. 2015; Martin et al. 2017). Since the heritability is a ratio of variances, and the variance of a variance is a difficult thing to measure, measurement error is an important consideration. However, assuming that reasonably unbiased estimates are obtained, we have three possible explanations for a low heritability of intra-genotypic variability. First, it may be a relatively constrained trait, with little variance. This will be modulated in part by the architecture of the gene networks regulating the degree of gene expression noise, and the degree to which it responds to mutations (Chalancon et al. 2012; Shreif and Periwal 2014). Molecular factors that may contribute to the degree of expression noise, for instance, include cell-to-cell variation in available machinery, mRNA and protein degradation and epigenetic histone modifications (Newman et al. 2006; Chalancon et al. 2012).

Second, populations may be close to their (long-term) fitness optima, such that all additional additive genetic variance has been eroded. However, evidence on whether populations are close to their optima, is mixed: theoretical predictions on optimal diapause timing in copepods (Hairston and Munns 1984) and on optimal germination and flowering probabilities in plants (Rees et al. 2006; Gremer and Venable 2014; Gremer et al. 2016), have been shown to closely match observations. In contrast, Philippi (1993), also focusing on germination probabilities, showed that predicted optimal strategies did not match the observations. An apparent weak selection on optimal bet-hedging strategies in specific cases can be due to: i) selection for bet hedging being of recent origin, simply resulting in insufficient time for optimal strategies to have evolved, or ii) selection on bet hedging being infrequent, implying that bet hedging promoting alleles disappear due to drift or short-term selection against them, despite being beneficial on the long-term (King and Masel 2007). Our theoretical predictions assume infinite large populations. However, real-world populations consist of individuals that are subject to demographic stochasticity, resulting in genetic drift. Whether a variance-increasing mutation appearing at one time point will get fixated in the population, depends on the population size, in addition to the selective advantage and the experienced environment (King and Masel 2007). Moreover, responses to selection are additionally shaped by genetic (co)variances, which have the potential to create a wide range of evolutionary trajectories towards the optimum (Box 6.1).

Third, intra-genotypic variability may be highly context-dependent, and thus heritable in the broad sense but not the narrow sense. In other words, there may be standing genetic variation, but these alleles appear neutral until particular environmental circumstances occur, and thus their

contribution to heritability would be cryptic. Our fitness modelling focused on fitness outcomes across environments that fluctuate, but in a stationary manner (i.e. around a consistent mean) over the long term, and we did not allow for context-dependent manifestation of intra-genotypic variability (Bishop et al. 2007; Raj et al. 2010). Transient and directional changes in environmental conditions that co-occurred with expression of intra-genotypic variability could help sustain populations across challenging periods (Bishop et al. 2007). Characterizing these links empirically would require sufficient knowledge of the system as to be able to identify the condition-dependence of expression of intra-genotypic variability, as well as its demographic consequences as outlined above.

6.9 Going forward

Increasing recognition of the ubiquity of intra-genotypic variation across natural systems and characterization of its mechanistic underpinnings remain to date relatively divorced from theoretical expectations for its evolution and spread, with a few exceptions. For instance, Mulder et al. (2016) demonstrated a genetic component in the degree of variation in fledgling weight, in a wild great tit population. Moreover, evidence for a relation between the degree of variation and fitness was found, indicating stabilizing selection on intermediate phenotypic variance (Mulder et al. 2016). A clear next step is to use theoretical expectations to characterize more empirical examples. This remains challenging for a number of reasons. First, some of the best characterized empirical examples are from laboratory or domestic animals, making direct application of evolutionary predictions complicated (as populations may not be exhibiting the outcomes of natural selection). Second, inferring the degree to which intra-genotypic variability is adaptive requires quantifying the impacts of a varying environment on (components of) fitness. Although our taxonomy of patterns of selection on variance, across patterns of environmental variability and trait-demography reveals some generalities (Table 6.1), in general, our predictions are more qualitative than quantitative because of the nuance of other traits, fitness components and environmental context. Quantifying the impact of a varying environment on fitness is intractable unless a life history model can be constructed. This in turn requires measurement of core individual demographic rates (survival, growth, reproduction) as well as measurements of how variation in both the trait and the environment affect these demographic rates. The required information is rarely available (with the notable exception of plants with seed banks; Childs et al. 2010; Gremer et al. 2016). Third, and finally, evolutionary outcomes will be modulated by the degree to which genetic variation is available for selection to act upon, e.g., both additive genetic variance, and covariances with other traits (as well as selection on those traits) (Box 6.1). Estimating genetic variances and covariances (e.g. using mixed effects models; Cleasby et al. 2015) is very data-hungry for mean trait values, and this challenge is amplified when attempting to do the same for variance in trait values. Estimating these components requires data on the relatedness between individuals, which generally calls for genetic analyses. Current developments improving efficiency and quality of genetic analyses (De Barba et al. 2017) will help making these pedigree data available on a larger scale.

   Making robust quantitative predictions on the evolution of variance control in natural populations thus require long-term data on i) temporal changes in one or more environmental variables, ii) individual phenotypes and demographic rates, interacting with the environment, and iii) the relatedness between individuals. Whenever those data are available, a first step will be to estimate the genetic component of the phenotypic variance, and its genetic correlation with the mean trait (Sae-Lim et al. 2015; Mulder et al. 2016; Martin et al. 2017). Second, by relating the focal phenotypic trait to one or more fitness components, a life history model can be constructed. This model can be used for instance to identify optimal combinations of the phenotypic trait mean and

variance under actual or future environmental conditions, compare predicted optimal strategies to observations, and determine the relative roles of non-linear averaging and bet hedging. Combining the adaptive landscapes with the evolutionary potential of intra-genotypic variability, finally, will be essential for making predictions on the evolution of intra-genotypic variability. Future studies applying demographic models, such as the one proposed here, to empirical data, will help us to explain the ubiquity of the genetic control of phenotypic variation and to understand its relevance for natural populations facing environmental changes.
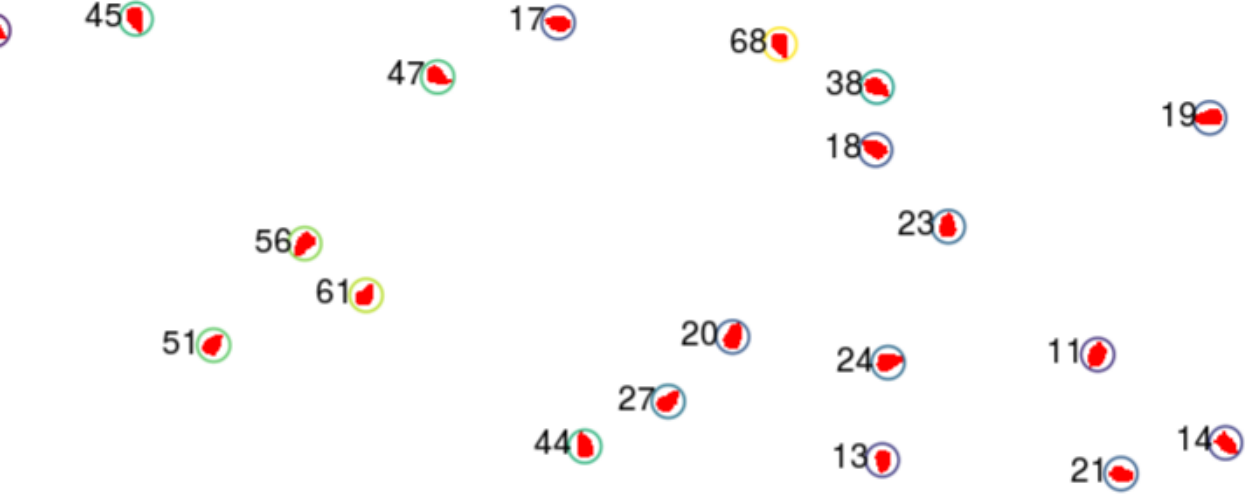
Marjolein Bruijning[1], Marco D. Visser[2], Caspar A. Hallmann[1] and Eelke Jongejans[1]

1. Department of Animal Ecology and Physiology, Radboud University, 6500 GL Nijmegen, The Netherlands.
2. Princeton University, Department of Ecology and Evolutionary Biology, 08540 Princeton, New Jersey, United States of America.

trackdem: automated particle tracking to obtain population counts and size distributions from videos in R

7.1 Abstract

1. The possibilities for image analysis in scientific research are substantial: the costs of digital cameras and data storage are sharply decreasing, and automated image analyses greatly increase the scale, reproducibility and robustness of biological studies. However, automated image-analysis in ecological and evolutionary studies is still in its infancy. There is a clear need for easy to use and accessible tools.

2. Here, we provide a general-purpose method to obtain estimates of population densities, individual body sizes and behavioural metrics from video material of moving organisms. The methods are supplied as a new R-package *trackdem*, which provides a flexible, easy to install and use, generally applicable, and accurate way to analyse ecological video data. The package can detect and track moving particles, count individuals and estimate individual sizes using background detection, particle identification and particle tracking algorithms. Machine learning is implemented to reduce the influence of noise in lower-quality videos, or to distinguish a single species in multi-species systems.

3. We show that *trackdem* provides accurate population counts and body size distributions. Using a series of simulations, we show that our estimates are robust against high levels of noise in videos. When applied to live populations of *Daphnia magna*, our methods obtained accurate and unbiased estimates of population counts, individual sizes, and size distributions, as verified by manual counting and measuring. The package *trackdem* is also directly usable for movement analysis, for instance in behavioural ecology, as illustrated by the tracking of insects, fish, cars and humans.

4. Within 24 hours, we obtained 192 accurate population counts and body sizes of 22,154 individuals. Such results underscore that automated analysis can improve robustness and reproducibility, and greatly increase the scope of studies in ecology and evolution.

7.2   Introduction

Many ecological studies rely on accurate measurements of the number, size distribution or be-
haviour of individuals in populations. Commonly, such metrics are obtained by manual counting
and measuring, which can be time consuming and labour intensive, thereby limiting the scope of
the studies. Moreover, manual measurements may be subjective, and are often not reproducible,
as no records or materials are preserved. Improved measurement accuracy and data management
are especially important when the reproducibility of science is being challenged (Ioannidis 2005;
Collins and Tabak 2014; Errington et al. 2014) and scientists are ever increasing the scale and
scope of data and research questions (Petrovskii and Petrovskaya 2012; Bolker et al. 2013).

   A promising alternative to manual counting is automated image or video analysis, however,
to date, these tools have only been used occasionally in ecological studies (Dell et al. 2014;
Pennekamp et al. 2015; Gerum et al. 2017). Studies that do use video analysis often either 1)
lack the generality and flexibility needed for easy application across systems or set-ups; 2) are not
freely available or platform-independent; 3) are cumbersome to install or use; 4) are not accurate
enough to base demographic studies on, or 5) fail to provide measures of accuracy and robustness
(e.g. Færøvig et al. 2002; Noldus et al. 2002; Hooper et al. 2006; Auclerc et al. 2010; Mallard et al.
2013; Bánszegi et al. 2014; Pérez-Escudero et al. 2014; Pennekamp et al. 2015). These drawbacks
need to be addressed, before ecology can harness the full power and potential of automated image
and video analysis.

   We improve on all the above points in a new R-package *trackdem*, which provides a method
to obtain estimates of the number, individual body sizes as well as behavioural metrics from
recordings of multiple organisms simultaneously. We implemented machine learning to filter out
noise or to distinguish single species in multi-species systems, based on visual characteristics.
Here, we briefly describe the methodology, and evaluate its performance on both simulated and live
populations. We show that applying *trackdem* is flexible and straightforward, and that it provides
accurate and reproducible results in a wide range of study systems without the need for expensive
equipment.

7.3   A brief overview of the methodology of *trackdem*

To apply *trackdem* to live populations, the first step is to record a short movie. Since *trackdem*
is based on movement detection, it is crucial that particles of interest move. In most systems, a
video of approximately four seconds will suffice, and the best results are obtained with a stable
(contrasting) background, and strong (background) illumination (Pennekamp and Schtickzelle
2013). The implemented methodology consists of four steps: 1) image sequence creation, 2)
particle identification, 3) particles tracking, and 4) machine learning for detection optimization
and noise filtering. We explain the main aspects of each step below, while detailed information
can be found in Appendix F.1. All functions are combined in an open-source software package
for "particle *track*ing and *dem*ography", which is freely available as the R-package *trackdem*
on CRAN (Bruijning et al. 2017b, github.com/marjoleinbruijning/trackdem, https://CRAN.R-
project.org/package=trackdem). An overview of the most relevant functions, for version 0.3.1, is
given in Fig. 7.1, and a tutorial can be found in Appendix F.2.

   Since all functions run in R, installation is straightforward, and full transparency and user-
modification of all functions is enabled. A native R implementation is both convenient and flexible:
users have access to the full range of statistical packages in R, and users can skip steps depending
on preference. Moreover, we optimized code to run fast and refactored key functions in C++
(*sensu* Visser et al. 2015), using *Rcpp* and the plugin *RcppArmadillo* (Eddelbuettel and Sanderson
2014), showing that a native R solution is not necessarily slow, as suggested in (Pennekamp and

Schtickzelle 2013).

Creating and loading an image sequence in R

The package uses Python code and the open source software libav (*libav.org*) to automatically convert recorded movies of study organisms into image sequences. These sequences are then stored in R as an array A with four dimensions (*X* and *Y* grid of the pixels, colour layers, and time), making use of R-package *png* (Urbanek 2013).

Particle identification

The particle identification starts with the detection of a 'still background', containing all motionless objects. As *trackdem* is developed to work with short movies, we use static background models, which are not continuously updated, and require a stable background where only focal particles move. Whenever backgrounds change, in e.g. longer videos or more complex field conditions, more advanced dynamic background modelling can be implemented (e.g. Barnich and Van Droogenbroeck 2011; Bouwmans 2014, Appendix F.2.3). Three different methods for background detection are implemented natively in *trackdem* (Appendix F.1.3). Movement is subsequently detected by subtracting each original frame from the still background, where large differences in colour values indicate movement. Here, pixels are defined to be background (0) or part of a moving particle (1) depending on a movement threshold T: its value either automatically determined by minimizing within-class variance in colour values, or provided by the user (Appendix F.1.3). Finally, individual particles are identified and labelled by merging all eight adjacent particle-pixels, using connected component labelling (Chang et al. 2004) and package *SDMTools* by VanDerWal et al. (2014). If the minimum and maximum sizes of the study organism (in pixels) are known, a first filter can now also be applied to remove noise such as small air bubbles or large moving debris.
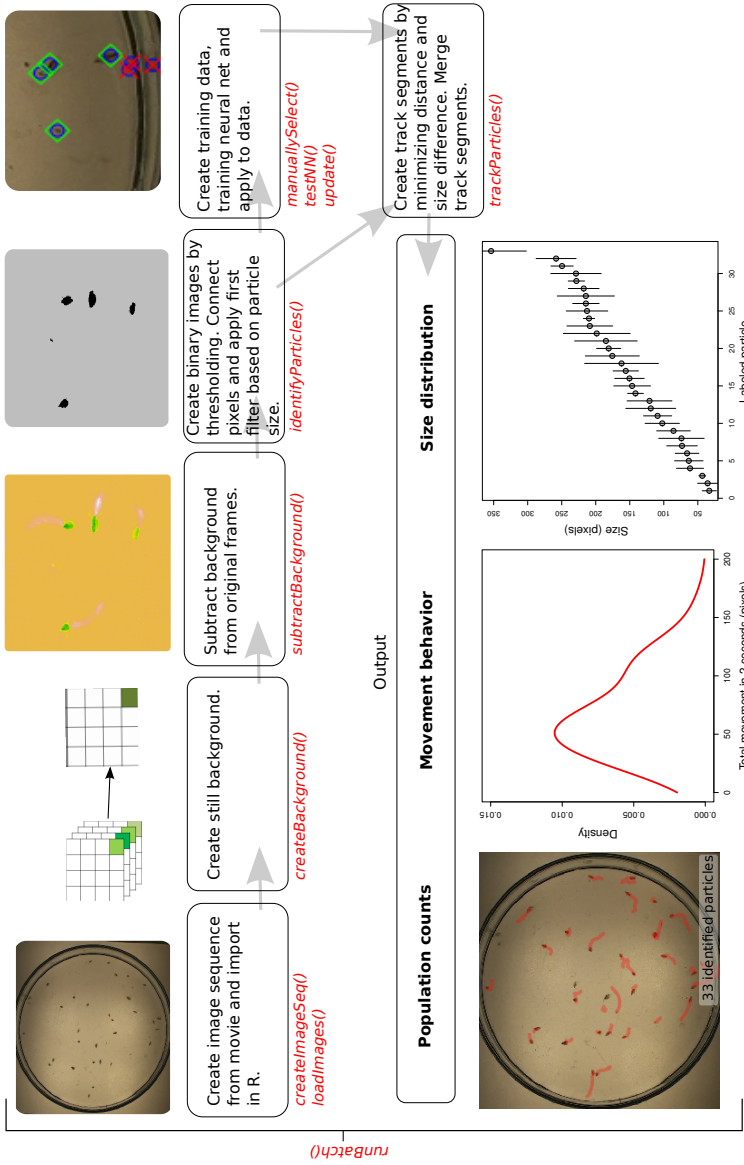
Figure 7.1: Schematic overview of a typical workflow, including most relevant functions and output provided in R-package *trackdem*. After initial setup the function *runBatch* may automate the analysis of large amounts of video material.

Tracking of particles

We implemented a tracking algorithm similar to the one described in Sbalzarini and Koumoutsakos (2005) to identify and reconstruct the movement trajectories of all individual particles over time. In short, particles between frames are assumed to belong to the same individual if they are similar in size, speed, direction and location. Similarity is measured by a cost function, which describes the costs to link a particle i in frame k to a particle j in frame $k+1$, and which is minimized across particles. In some instances, particles may disappear (e.g. due to overlap), or conversely appear (e.g. due to noise), which may result in unequal number of particles between frames. Whenever this occurs, the start or end of an individual trajectory is generated (more details in Appendix F.1.4).

This first step results in tracked segments where particles are linked between consecutive frames. When all existing particles are detected and there is no noise, this gives a complete and correct tracking record. However, in real life applications this may not be the case, as particles are sometimes overlapping, not detected in some frames, or confounded with noise. Accuracy is then further improved by merging incomplete track segments. The same cost function is used to calculate the cost of linking the end of segment i in frame k to the start of a segment j in frame $k+r$. In an iterative process ranging r between 1 and maximum R, incomplete track segments across frames are linked.

Particle tracking then provides a sequence of locations and sizes of identified particles which can be used to calculate statistics of interest. Note that *trackdem* assumes that splitting and merging of particles (as explicitly incorporated in Jaqaman et al. 2008) does not occur, as such processes, akin to cell division, do not generally occur in ecological systems of short time spans of typical video material (seconds to minutes).

Noise reduction and detection optimization

A particle tracking algorithm will overestimate actual numbers whenever movies contain identifiable moving particles which are not of interest (noise). Noise may include false positives such as moving debris, or non-focal particles, such as other species in the case of multi-species systems. When such noise has not been filtered out in previous steps by thresholding and size filtering, it may be removed by implementing supervised machine learning. In *trackdem*, a set of functions allow users to interactively select true and false positives (using the package *shiny*; Chang et al. 2017), and then train an artificial neural network (using the package *neuralnet*; Fritsch and Guenther 2016). Neural networks are algorithms that are used, here, to distinguish between objects of interest and noise, based on visual characteristics (e.g. shape or colour), and are described in detail in various textbooks (e.g. Fielding 1999; Bishop 2006). In *trackdem* a neural network is optimized through cross-validation, and the final trained neural network is then applied to the identified particles. After application of the neural network, noise is removed, and particle tracking is reapplied, now with increased precision of estimates (see *Application and validation*; Fig. 7.2).

Estimates of population count, structure, and behaviour

The number of track segments is an estimate of the number of individuals. In the presence of incomplete tracks, a 'minimum presence' is defined, where a particle is only considered to be of interest when it is present in at least z of the frames. The value of z can be supplied by the user, or can be estimated through k-means clustering. This process ensures that the short segments, caused by noise, are excluded. From the remaining track segments, estimates of body size distributions are obtained as the mean or median number of pixels. A range of behavioural metrics can be obtained from particle coordinates, such as the distributions of speed and directionality, or a preference for

certain regions (Figs 7.1, 7.4, 7.5). Users can apply S3-generic-functions as *summary()* and *plot()* on *trackdem* objects to obtain these results.

7.4   Application and validation

To test the accuracy, and evaluate limitations, we applied our methods both to simulated particles and to video material on *Daphnia magna* populations. Benchmark times depend on resolution, image sequence length, noise, and number of particles. It took approximately 24 seconds to obtain populations counts and body size distributions from 30 frames of a simulated video with a 480×480 pixel resolution, and 78 seconds for a 1000×1000 pixel resolution (both containing 100 particles). It took 3-4 minutes for 30 frames from a 2 second HD movie (1080×1915) on live *Daphnia magna* populations (see below). Benchmarks were taken on a Lenovo z580 (Intel i7-3632QM CPU @ 2.20GHz, 8 GB of RAM, running Ubuntu 16.04).
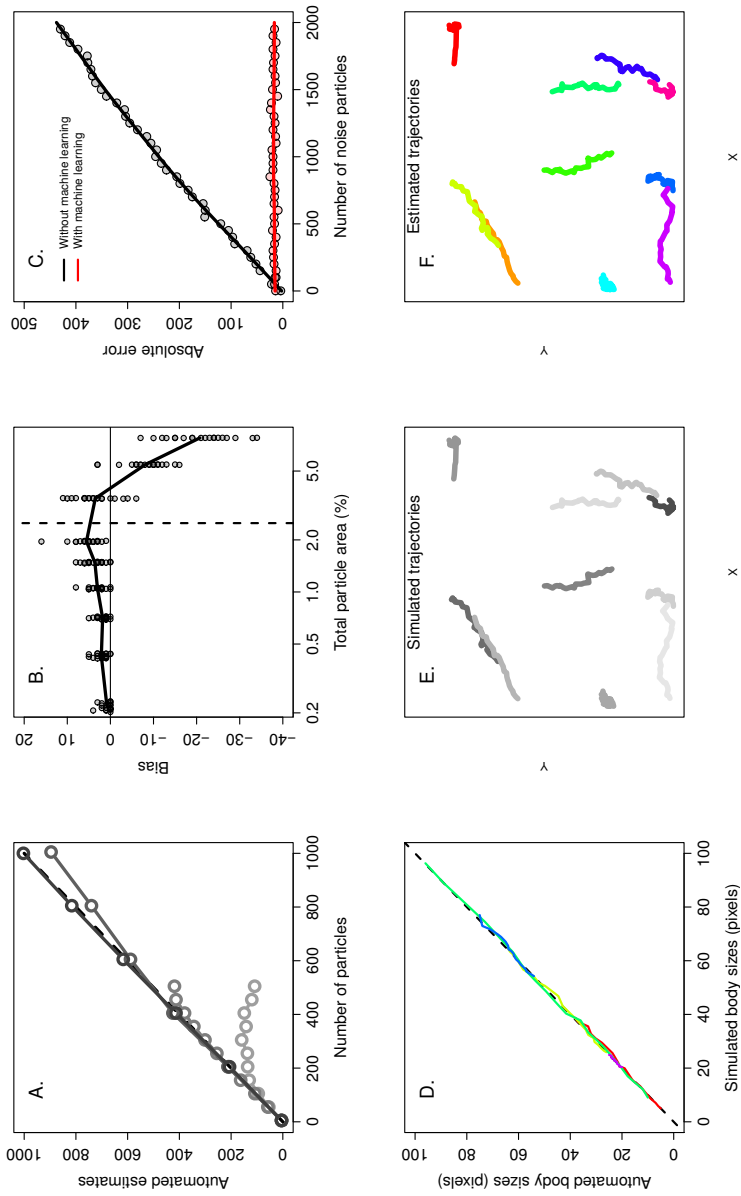
Figure 7.2: Application of *trackdem* to simulated movies. In a) the estimated number of particles is plotted against the simulated number of particles. Different colours show different relative densities per pixel, from low (dark grey) to high (light grey). In b) we show bias as a function of total particle area, simulating 100 particles per image sequence. In c) the accuracy of *trackdem* was tested for different numbers of noise particles. Black points indicate estimates without machine learning, red points with machine learning. Lines show moving averages based on generalized additive models. In d) a Q-Q plot shows estimated body size distributions versus simulated body size distributions. Ten independent image sequences were simulated, with 100 particles each having a randomly drawn size. In e) and f), simulated trajectories are compared with estimated trajectories, for ten simulated particles.

Simulated trajectories

We developed code to simulate movies, as this enables us to find the theoretical limits of our methods by exactly controlling the number of particles, their size, video resolution, and noise. We simulated image sequences by generating movement paths for particles that had both movement speeds and directions drawn from pseudo-random distributions. Simulation functions are available to the user (Appendix F.2.1).

The effect of particle number and sizes on the precision of *trackdem* was tested by modifying the number and average size of particles (Fig. 7.2a-b). When small particles were simulated, estimates were precise for the entire range of tested particle numbers (up to 1000 particles, making up $\pm2.6\%$ of all pixels of a 480×480 image). Results show that the maximum number of detectable particles was asymptotically dependent on the resolution: when the total area occupied by particles increased, overlap increased and estimates were more error prone. When the total area of each image occupied by moving particles (particles/pixels) exceeded $\pm3\%$, the method tended to underestimate the true number of particles by >10%, and the absolute error in population count increased substantially (Fig. 7.2a-b). Systems with higher densities of organism will therefore require higher resolution videos and corresponding equipment, or sub-sampling of the population at lower densities per pixel. Note that computational intensity will increase with resolution.

In the implemented algorithm, particles are linked from frame to frame, not taking into account the full particle history. This means that the individual identities can be maintained only to a certain extent. Whenever there is too much overlap between particles, individual identities may be lost, and reconstructed trajectories may contain errors. In such cases (if reducing density is not an option), pattern recognition methods to identify individuals (Bishop 2006; Szeliski 2011; Pérez-Escudero et al. 2014) or more complex tracking algorithms (Savin and Doyle 2005; Chenouard et al. 2014) may offer solutions.

Robustness to noise was evaluated for a range of different simulated noise levels, keeping the simulated number of focal particles at 100. Noise was simulated as randomly appearing particles each with a randomized colour. Noise particles moved according to a random walk and remained for a random number of frames (ranging between 1 and $\frac{1}{3}$ of the total number of frames). This simulates different types of noise, e.g. the appearance of reflection on water surfaces or air bubbles. Results show that absolute error increased proportionally with noise (grey dots in Fig. 7.2c, $r^2=0.99$). When applying a trained neural network, noise could be successfully removed, and population estimates remained accurate for all levels of noise (red dots in Fig. 2c), with the absolute error becoming independent of noise density ($r^2=0.017$). Here, the neural-net was trained with a small dataset containing only 40 true positives and 22 false positives. Machine learning filter can, in principle, also be applied in multi-species systems to distinguish between species; the neural net can be trained to recognize the species of interest based on shape, size and colour (Soleymani et al. 2015), and species can then be counted in turn.

Additionally, the accuracy of estimated particle size distributions was tested, using five independent simulations of 100 particles each with randomly drawn body sizes. Estimated body size distributions matched perfectly with simulated body size distributions (Q-Q plot in Fig. 7.2d; $r^2=1$). Finally, *trackdem* proved able to reconstruct movement trajectories, as illustrated in Fig. 7.2e-f.

Application to *Daphnia magna* populations

Using video material of live *Daphnia magna* populations, we evaluated accuracy in population counts, individual sizes and body size distributions compared to manual estimates. We used two sources of video material. First, we included 39 movies that were recorded in our lab at Radboud University, with the objective to test *trackdem*. These movies were made using the same HD
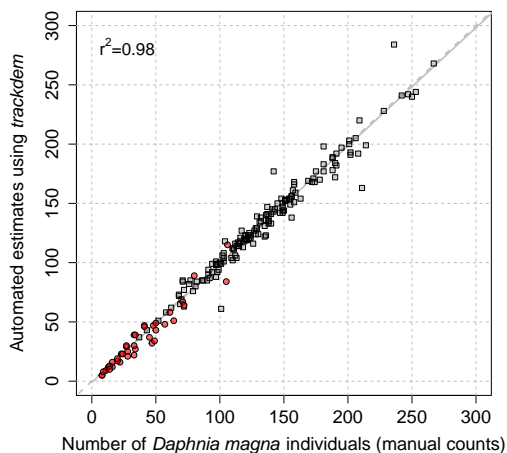
Figure 7.3: Manual counts versus automated counts. Each dot represents one video on *Daphnia magna* populations. Interrupted line shows x=y line, solid line shows relation between x and y based on a linear regression. Red circles indicate movies which were made to develop *trackdem*, whereas grey squares indicate movies which were recorded separately from the development of *trackdem*.

camera (1024×1950 pixels) on different populations in February 2016 (details of the setup are given in Appendix F.3.1). A neural network was trained by randomly selecting one frame from each of 10 movies. Second, we included 153 movies, recorded at the Norwegian University of Science and Technology, which were recorded independently from the development of *trackdem*, between September and November 2016. Here, the neural net was based on one frame from each of five randomly selected movies. Overall, *trackdem* provided highly accurate and unbiased population counts for the complete range of tested population sizes (Fig. 7.3; r^2=0.98).

The reliability of automated estimates of individual sizes was evaluated with 53 movies of single *Daphnia magna* individuals, recorded in our lab. For each video, a set of 30 images was used. Manual estimates of body size in pixels were obtained by tracing polygons by hand on four frames using imageJ (Schneider et al. 2012). Additionally, we compared automated estimates with manual body-length measurements performed under a stereoscope. Individuals ranged from 1.1 mm to 3.5 mm length, as determined by body length measurements (measured from the base of the spine until the middle of the eye). Length (L) measurements were transformed to area in mm^2 assuming *D. magna* are perfect circles ($A = \pi \cdot (L/2)^2$). For 35 successfully detected individuals, we found a strong correlation between automated estimates and manually traced polygons (Fig. 7.4a; r^2=0.80) and manual body length measurements (Fig. 7.4b; r^2=0.68). Note that manual measurements are also subject to uncertainty: variation within individual manually traced polygons was high (see horizontal error bars in Fig. 7.4a), and manual length measures assume a circular shape and use only a single length measurement. In contrast, *trackdem* estimates body size by counting pixels and averaging over all frames, assuming no particular shape.

For 18 out of 53 recorded movies, we did not successfully detect the individual of interest, due to for instance limited movement. This illustrates that *trackdem* estimates are subject to the condition of movement; non-moving particles are seen as background and disappear after background subtraction - which is true for all methods based on background subtraction (e.g.

Bánszegi et al. 2014). Whenever possible, researchers should ensure that study organisms move, or are induced to move, which is highly system-dependent. In the case of *Daphnia magna*, movement may be induced by a sudden light stimulus (De Meester 1993). The detection of fully motionless particles requires object and pattern recognition methods (e.g. SIFT or Haar classifiers; Bishop 2006; Szeliski 2011), however, this is beyond the scope of *trackdem*, where machine learning currently only removes false positives.

Finally, automated and manual measurements of size distributions from seven different *D. magna* populations strongly resembled each other (Q-Q plots shown in Fig. 7.4c, r^2=0.88). Here, size measurements were normalized once over all seven movies for both manual (mm^2) and automated estimates (pixels) to facilitate comparisons. Kolmogorov-Smirnov tests showed no significant differences between manual measures and automated estimates of population size structures (p-values ranging between 0.12 and 0.84). Accurate and unbiased estimates of both population size and structure are required if one wants to estimate demographic parameters in ecological and evolutionary studies. For instance, time series of population size and structure allow for the estimation of growth, survival and reproduction rates, and the construction of population models (González et al. 2016).

Applying *trackdem* to other study systems

The standard analysis in *trackdem* also directly results in information on behavioural metrics, such as distributions of movement speed and direction (Fig. 7.1), as well as locations and trajectories (Fig. 7.2e,f). These metrics can be used for various objectives, as illustrated in Fig. 7.5. We estimated numbers of both fish and honey bees (Ma et al. 2015) when individuals can leave the filmed area, resulting in changes in numbers over time. Furthermore, we used *trackdem* to count the number of cars on a highway (Vezzani and Cucchiara 2010). Finally, we generated movement density maps for insects in the field (Jongejans et al. 2015) and human movement on a sidewalk (Pellegrini et al. 2009). More information and R-code can be found in Appendix F.3.2.

7.5 Conclusions

Automated image analyses are repeatable, not subject to observer bias, and video material can be easily archived. Our results also underscore that automated analysis can improve robustness, reproducibility and greatly increase the potential scope of ecological science. Most notably, it reduces human workload per data unit, which essentially allows experimental units to increase in number. Testing it on live populations, we were able to acquire population abundance for 192 different *Daphnia* populations and body sizes of 22,154 individuals within 24 hours, after initial setup. Obtaining the manual measurements or validation from these populations took an order of magnitude longer. The *trackdem* package is a flexible toolbox for automated image analysis in R, generally applicable to many systems, and straightforward to install and use. It yields accurate and unbiased estimates of population size and structure, suitable for demographic modelling. Dell et al. (2014) compared 16 software packages for automated tracking using image analysis and concluded that no single package was open-source, platform-independent and able to follow trajectories of more than 80 particles (see table S1 in Dell et al. 2014). The *trackdem* package is the only available method that ticks all these boxes, and hence can be seen as a response to Dell et al.'s (2014) 'call to developers'. Our work attempts to bridge the gap between the advanced tracking technology available today and the specific needs of ecologists to track, count and measure individuals. With *trackdem*, we hope to ease the broad application of automated image and video analysis in ecology and evolution.
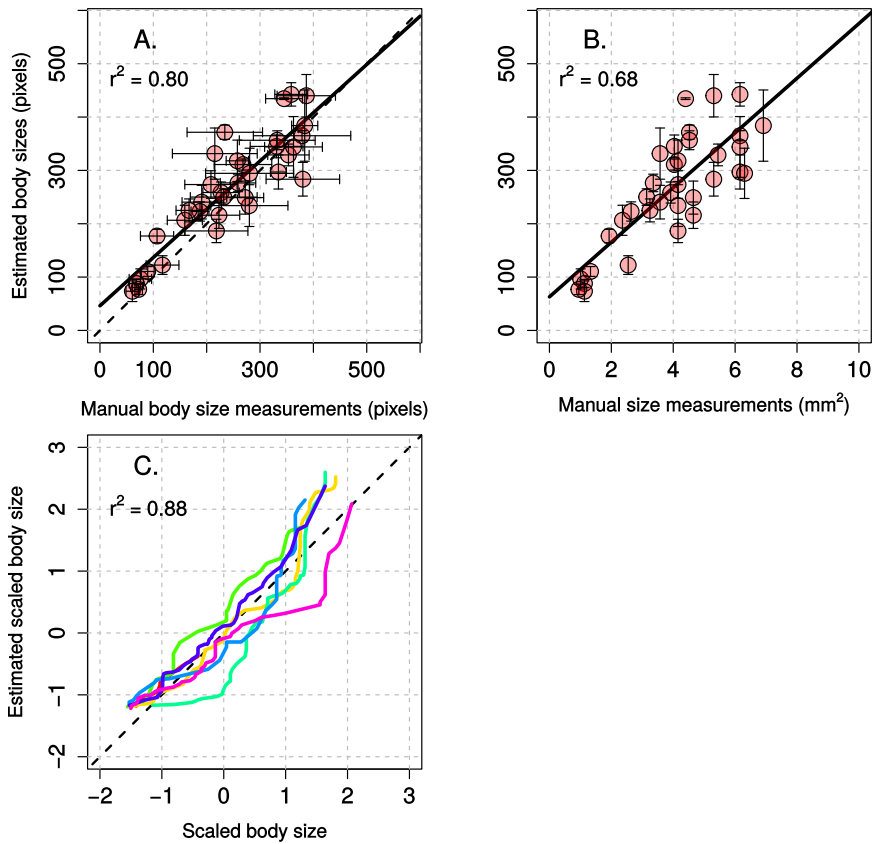
Figure 7.4: Body size estimates of individual *Daphnia magna* individuals using *trackdem* are compared to a) manual measurements performed on the obtained movie, and b) manual body length measurements using a stereo microscope. Body length was used to calculate the area in mm^2, assuming a perfect circle. Error bars show variation within an identified particle over all frames (1 SD). Each dot represents an independent movie. In a) the interrupted line shows the x=y line. In c) body size distributions are compared for seven movies (indicated by different colours). Both estimated and observed sizes were normalized once over all seven movies.
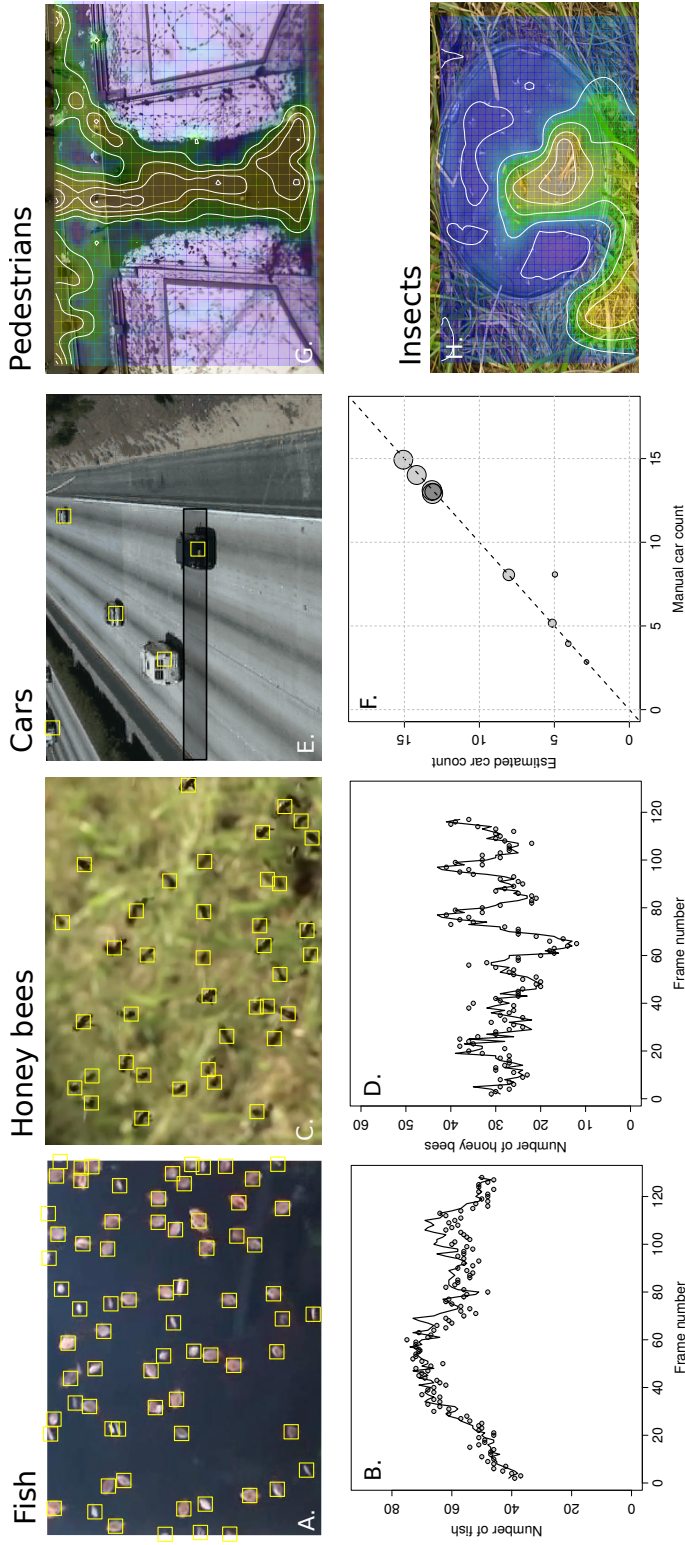
Figure 7.5: Illustration of the application of *trackdem* to other study systems. Upper row (a, c, e) shows the identified particles for one frame, and bottom row (b, d, f) compares estimates to the ground truth. In b) estimated number of fish and d) estimated number of honey bees. Lines show true numbers per frame, dots show automated estimates. Average error was 2.1 and 0.017, average absolute error was 3.6 and 2.6, and coefficients of determination were 0.80 and 0.74, for the fish and honey bees, respectively. In e-f) identified vehicles on a high way. Number of cars passing the drawn polygon were compared to manual counts. This was done for ten subsets of images with different durations; dots are scaled to the length of the image sequence, which ranged between 44 and 190. In 9 out of 10 videos, a perfect match was found (r^2=0.96). For the graph, a little noise was added to the numbers as three datapoints overlapped. In g-h) heat maps show where pedestrians move on a side walk, and where insect movement occurs.

Acknowledgements
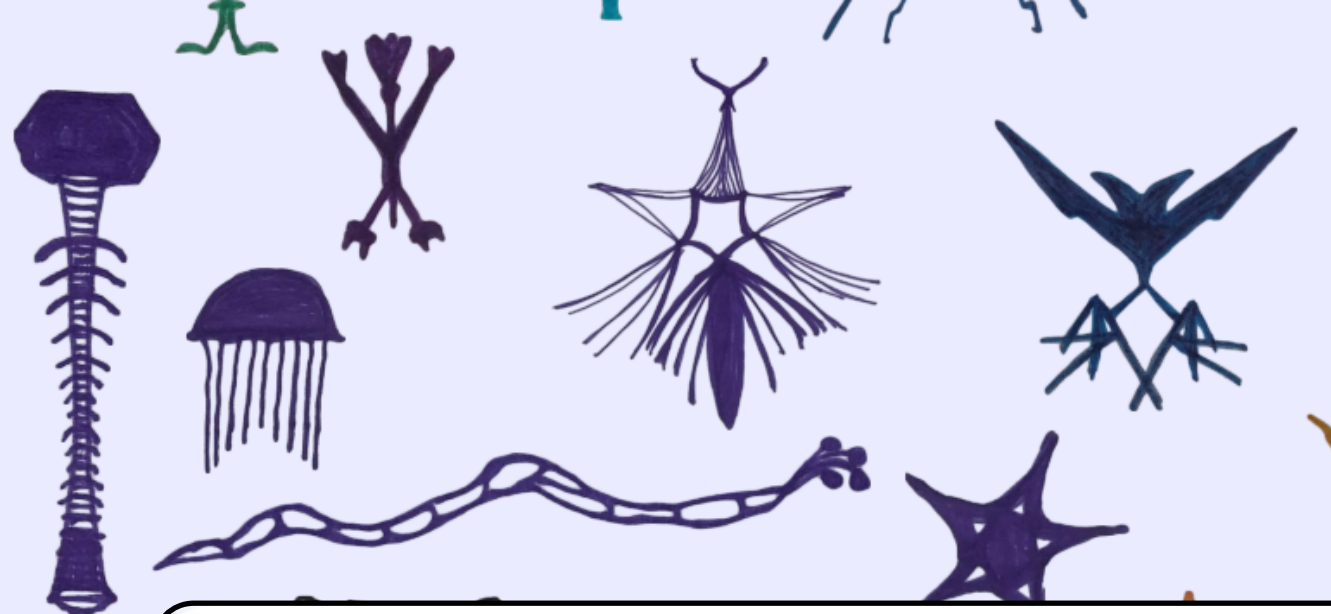
Synthesis: Integrating plastic and genetic responses across the life cycle

All organisms in the world experience environmental changes. These changes include for example seasonal changes, changing exposure to predators or pathogens, and long-term directional changes such as climate change. How individual organism responds to a changed environment is a function of an individual's phenotype, consisting of all individual morphological, chemical and behavioural properties. Some phenotypes perform better than other phenotypes in a given environment, reflected by an increased life-time reproductive success. Although selection acts on phenotypes, it is the individual's genes that are inherited to the next generation. Conceptually, three different strategies exist for how a genotype is translated into a phenotype with regards to its environment. First, a genotype can influence the average phenotypic trait value that is produced. Second, a genotype can affect the ability of an individual to plastically adjust its average phenotypic trait in response to the environment (i.e. a genotype-specific reaction norm). Third, a genotype can influence the phenotypic trait variation that is produced irrespective of the environment: intra-genotypic variability.

All three strategies can change phenotype dynamics in a population, and all three are the focus of this dissertation. **Chapter 2** compares four frameworks that aim to quantify the contribution of different processes to a change in a population-level mean phenotype. **Chapters 3-5** present empirical studies on the role of genetic variation and plasticity in response to changing environmental conditions, using the water flea *Daphnia magna* and the green peach aphid *Myzus persicae* as study species. **Chapter 6** explores theoretical expectations on the evolution of intra-genotypic variability. Finally, in **Chapter 7**, I present a method to obtain automated estimates of population sizes and structure, which enabled us to perform the large-scale experiments presented in **Chapters 3-4**. In this final chapter, I synthesize the presented studies. The chapter consists of four sections:

  (a) I start by providing a conceptual overview of the different factors and processes influencing the long-term fitness of a strategy, and how these processes relate to each other.
  (b) I follow by discussing what the most successful strategies are expected to be, depending

on the environmental conditions and species life history. This section is illustrated by a simulation study, building on the model proposed in **Chapter 6**.

(c) I then discuss a framework to describe eco-evolutionary feedback loops (Lion 2018), discuss opportunities to extend it, and lay out where all components described in a) are incorporated in this framework. In particular, I explore how we can use this framework to identify the presence and the success of different strategies.

(d) I conclude by speculating on the implications of the presented results for the fate of natural populations facing global change.

8.1   From genes to long-term fitness

Natural selection favours the genotypes that maximize long-term fitness, through selection on phenotypes. To understand and predict the long-term fitness of genotypes, it is crucial to understand how genotypes translate into phenotypes, how the produced phenotypes influence vital rates and fitness, and how this interacts with the environment. Figure 8.1 shows these processes and their relations. All chapters presented in this dissertation focus on subsets of these processes. The numbered arrows describe the following processes:

1. The presented chapters focus on three ways in which an individual set of genes (a genotype) can shape phenotypes. First, by influencing the average phenotype. Second, by interacting with the environment (which involves an interaction between arrows 1 and 6; phenotypic plasticity). Third, by influencing the variation in phenotypes independent of the environment (intra-genotypic variability) (**Chapters 2-4, 6**).

2. Phenotypes influence (age- or stage-specific) patterns of reproduction and survival, in interaction with environmental effects (arrow 5). Note that survival and reproduction can consist of multiple fitness components, or vital rates (e.g. maturation rate, clutch size and age-dependent survival) and there exists correlations between these components (for instance caused by trade-offs; Stearns 1989). Selection acts on phenotypic properties due to these relations between phenotypes and fitness components (**Chapters 2-4, 6**).

3. All vital rates together determine fitness, which measures the rate 'at which a genotype is able to propagate itself into future generations' (Caswell 2001) (i.e. exponential growth rate). Long-term fitness is the geometric mean of the population growth rates over time (**Chapters 3-6**).

4. Population growth rates determine changes in the number of individuals and determine their density. The change in the frequency of a genotype (evolution), depends on the relative fitness of the genotype and its current frequency (measured as the ratio of its density over the total density of the population), as explained below (Eqs 8.2, 8.3). The density is a component of the experienced environment. For example, higher densities increase competition for space or light, or reduce the per capita nutrient or food availability (Lampert 1978) (**Chapters 3-5**).

5. Relations between phenotypes and vital rates depend on the environment. The environment consists of all abiotic factors, such as temperature, nutrient availability and oxygen, and biotic factors as density. In addition to the genotype-specific density (arrow 4), biotic environmental factors include the densities of different genotypes from the same species and densities of different species, for instance reflecting predator-prey dynamics. Note that interactive effects of multiple environmental factors, e.g. an interaction between temperature and food concentration (Giebelhausen and Lampert 2001), can shape reproduction and survival (**Chapters 3-6**).

6. Environmental effects, together with the genetic component, determine an individual's
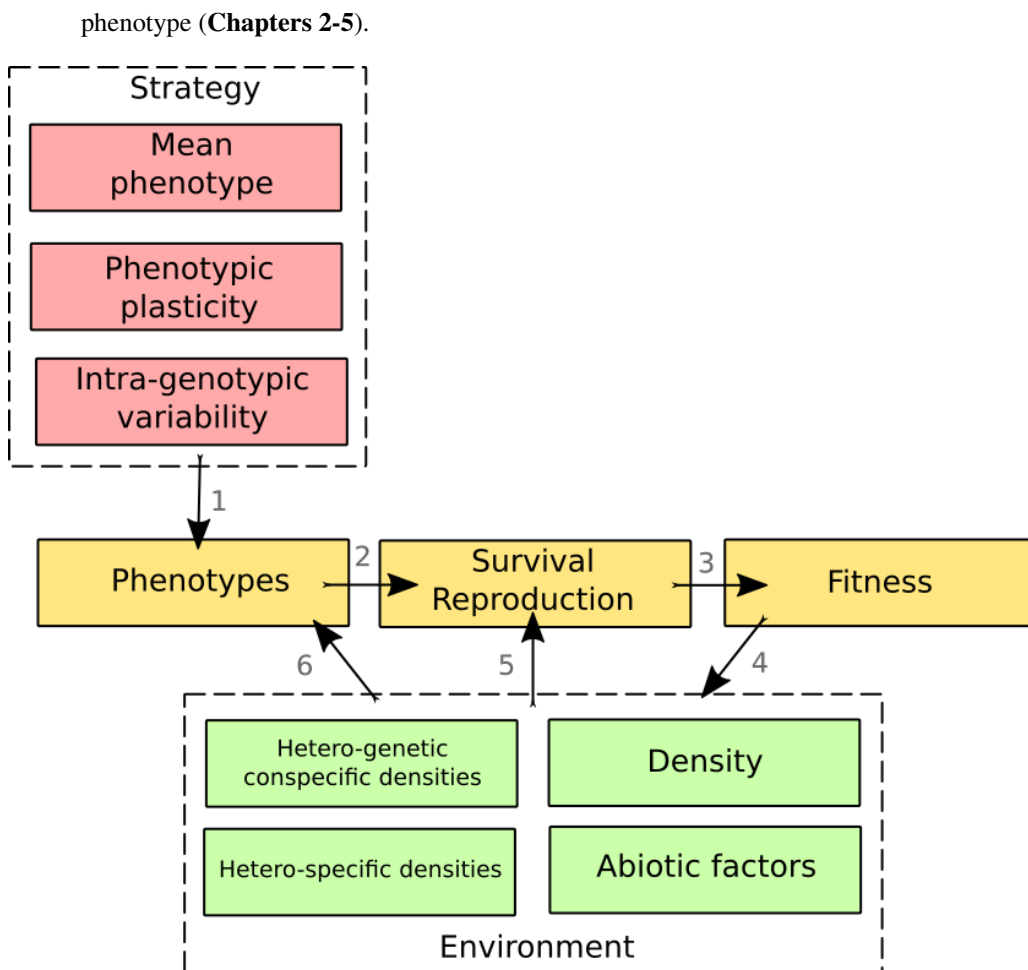
phenotype (**Chapters 2-5**).

Figure 8.1: Flow chart presenting the processes studied in this dissertation, and their relations. Red boxes indicate the genetic dynamics: three different ways in which genotypes produce phenotypes are shown. Yellow boxes indicate the individual-level dynamics, specifying the relations between individual phenotypes and performance. Green boxes indicate ecological dynamics. Numbered arrows indicate different processes and are explained in the text. Note that numerous variations on such an overview exist (see e.g. Scheiner 1993; Coulson et al. 2017). I chose for this simple and very general scheme, while including all processes studied in the presented chapters, and moreover, all relevant processes to describe the interactions between ecological and evolutionary dynamics (see also Fig. 8.3).

8.2 Adaptive tracking, phenotypic plasticity and intra-genotypic variability

Throughout this dissertation, three different mechanisms that induce phenotypic variation stand central: genetic variation, phenotypic plasticity and intra-genotypic variability. When genetic variation leads to variation in the average phenotype and some phenotypes have higher fitness than others, this will result in adaptive tracking, where the mean phenotype evolves towards the optimal phenotype (Arnold et al. 2001). Plasticity refers to the ability of an individual to change

its phenotype in response to the environment, in which the norm of reaction can differ between genotypes (Gotthard et al. 1995). Intra-genotypic variability, finally, creates phenotypic variation even in the absence of environmental variation, whereby the degree of variance can evolve (Hill and Mulder 2010). These three strategies differ in the relation between a genotype and the produced phenotypes, in interaction with the environment (arrows 1, 6 in Fig. 8.1). Their success depends on how these phenotypes subsequently influence vital rates (arrow 2) and fitness (arrows 3), given the biotic and abiotic environmental factors (arrows 5, 6). Is it most beneficial to express the phenotype that is optimal on average, to respond plastically to the experienced environment, or to always ensure some phenotypic variation, irrespective of the environment?

In **Chapter 6**, we explored the conditions that will select for intra-genotypic variability, compared to adaptive tracking. I now extend the results presented in **Chapter 6** and include phenotypic plasticity as a third trait that can evolve, by using a simulation (Box 8.1; see Fig. 8.2 for the processes implemented in this simulation). When there are no costs or constraints associated with plasticity, it is the optimal strategy: producing individuals that can transform their phenotype in an unlimited way to match the environment is obviously superior. However, in nature, the benefit of plasticity depends on the reliability of the environmental cues (Reed et al. 2010; Botero et al. 2015; Donelson et al. 2018) and individuals are limited by genetic, physical and developmental constraints and costs (DeWitt et al. 1998; Auld et al. 2009). When limits to plastic responses exist, selection will favour adaptive tracking and/or intra-genotypic variability in addition to plasticity. Here, in stable environments, it is the most beneficial to only express the optimal phenotype for that environment, and this phenotype is expected to evolve (Table 8.1). When the environment is stochastic, intra-genotypic variability can be selected for (Table 8.1). Here, the benefits of increasing phenotypic variation outweigh the costs only under sufficiently strong stabilizing selection on the phenotypic trait (Bull 1987). Moreover, the broader demographic context (i.e. life history of a species) affects the fitness consequences of the three strategies. As long-term fitness is a multiplicative process, it is proportionally more affected by low values. Therefore, for life histories that have a low 'baseline' fitness (e.g. annual plants), intra-genotypic variability in a reproduction-related trait is expected to be more important, to ensure at least some reproductive output every time step (discussed in more detail in **Chapter 6**).

Concluding, combinations of each of the three strategies can be favoured by selection, under scenarios of varying costs of plasticity, demographic context, environment and strength of selection. All these factors are likely to vary substantially in natural populations, and hence we may expect to find all three strategies in nature, which is in agreement with observations (Reznick and Ghalambor 2001; Grant and Grant 2002; Visser and Both 2005; Gremer et al. 2016; Mulder et al. 2016). Adequately predicting optimal strategies for particular cases will be challenging, and the devil is in the details. Critical questions include: what is the functional form of the relation between phenotype and fitness components; is there an optimum, is it accelerating, or is it linear (see **Chapter 6** for how this affects the expectations)? How does the focal phenotypic trait covary with other phenotypic traits and fitness components (e.g. are there trade-offs between vital rates)? What is the effect of different environmental factors on all these traits and fitness components? These are challenging questions to answer for natural populations, and data will often be limiting. I think that theoretical models or simulation studies, considering the full life-cycle, as for example presented in this dissertation, will be essential. Such theoretical studies will help us to identify the most important aspects, inform us on which data are most essential to collect, extract general patterns, and come with predictions that can subsequently be tested in natural populations.

Finally, I emphasize that whether a trait will evolve, additionally depends on the genetic architecture, including the genetic component of the trait as well as its genetic correlations with other traits (Lande 1979; Lande and Arnold 1983; Merilä and Sheldon 1999). This has
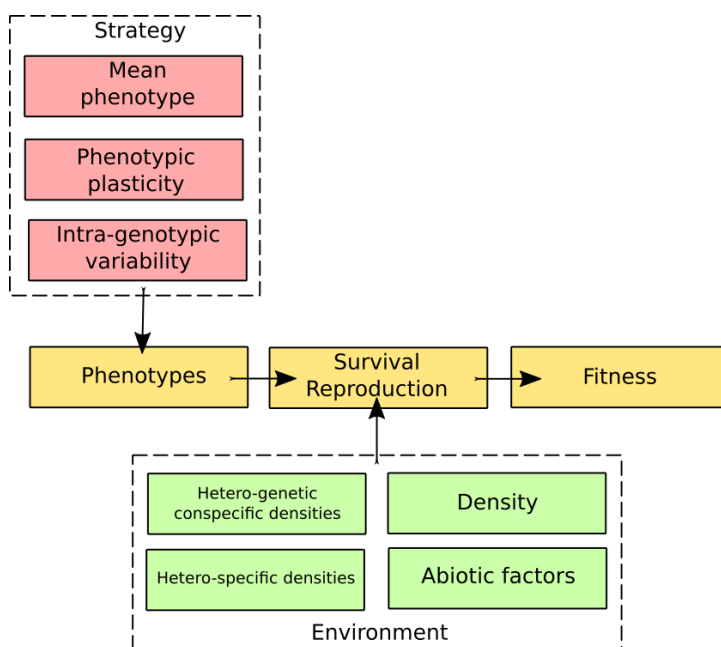
Figure 8.2: Flow chart similar to the one shown in Figure 8.1, but including only those arrows that indicate the processes which were considered for simulating the evolution of the three strategies (Box 8.1). I defined relations between phenotypes and fitness components, whereby the environment determined the optimal phenotype. The produced phenotypes, however, were not influenced by the environment (therefore lacking arrow 6, see Figure 8.1). I did not consider density-dependence, ignoring how fitness influences the density and how this feedbacks to fitness components (lacking arrow 4, see Figure 8.1).

four implications: first, the most beneficial strategy does not necessarily evolve, due to low genetic variation (Oostra et al. 2018) or correlations with other traits. Second, a strategy that is disadvantageous might still evolve due to genetic correlations with other traits. Third, how often a new strategy will appear due to a mutation, depends on the genetic architecture, with some traits showing higher mutational variability than others (Merilä and Sheldon 1999). Finally, once appeared, whether this new strategy will get fixated in population, is a stochastic process which also depends on the population size (i.e. genetic drift) (King and Masel 2007).

Box 8.1 — Simulating the evolution of different strategies. I simulated stabilizing selection on a trait that affects reproduction, with an environment-dependent optimal trait value. The mean expressed trait value can evolve. Plasticity acts, per time step, by shifting trait values towards their optimum. The degree of plasticity can evolve, and ranges between 0-100%, which indicates the proportional decrease in the mismatch (i.e. 0% is no plasticity, 100% results in always expressing the optimal phenotype). Note that I implemented only within-individual plasticity, and no transgenerational plasticity (Donelson et al. 2018). A cost of plasticity is implemented as a proportional decrease in either survival, reproduction, or both, and depends linearly on the degree of plasticity. Intra-genotypic variability is included as normally distributed variation in the phenotypic distribution, whereby the variance of this

Species and trait characteristics	Stable environment	Stochastic environment
No costs of plasticity	Plasticity / adaptive tracking	Plasticity
High costs of plasticity Weak stabilizing selection	Adaptive tracking	Adaptive tracking
High costs of plasticity Strong stabilizing selection High baseline fitness (long-lived species)	Adaptive tracking	Adaptive tracking
High costs of plasticity Strong stabilizing selection Low baseline fitness (short-lived species)	Adaptive tracking	Intra-genotypic variability

Table 8.1: Summary of the most successful strategies under different environmental conditions and life histories, as obtained by running a simulation (Box 8.1). Note that more moderate settings can result in combinations of intermediate outcomes.

distribution can evolve. Long-term fitness is obtained by using the framework presented in **Chapter 6**, based on Bull (1987). Running this simulation for different settings, the evolution of the average phenotype, plasticity and intra-genotypic variability can be simulated (summarized in Table 8.1). An interactive tool for running the simulation can be found at: https://marjoleinbruijning.shinyapps.io/popsimul/.
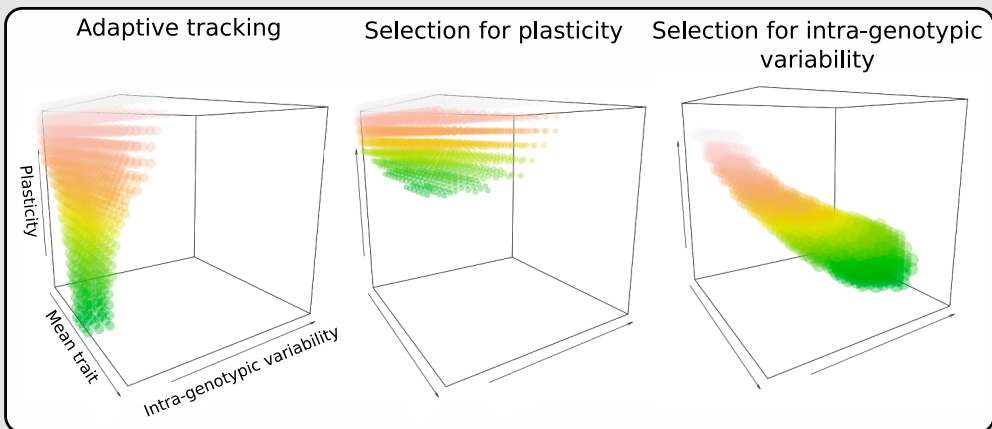
Three-dimensional fitness landscapes for combinations of the mean phenotypic trait, the degree of plasticity and the intra-genotypic variability. Dots show the 25% combinations that result in the highest long-term fitness, and are scaled to their values (the highest densities of dots indicate the combinations resulting in the highest fitness). For visibility, colours show variation in the z-axis, ranging from green (low plasticity) to grey (high plasticity). a) A constant environment, with a cost of plasticity. This results in adaptive tracking. b) A fluctuating environment with unlimited plasticity, resulting in the evolution of plasticity. c) A fluctuating environment with costs of plasticity, resulting in the evolution of intra-genotypic variability.   ∎

8.3 A general framework to describe eco-evolutionary dynamics

Evolutionary processes cannot be considered separate from ecological processes (Schoener 2011; Reznick 2013). This is because changes in allele frequencies through time (i.e. evolution) are due to genotype differences in life-time reproduction, which is under control of ecological processes. In turn, evolution changes population growth rates and thereby densities, which may affect a whole range of ecological factors, such as predator-prey dynamics. This results in a permanent feedback loop between ecology and evolution (Schoener 2011; Lion 2018). In the previous section, I explored the conditions that will select for different strategies, via relations between produced phenotypes and fitness components. I simulated an environmental effect on fitness and thereby on the evolutionary response, but I overlooked potential subsequent effects of evolution on the environment (Fig. 8.2). In this section, I will focus on two-way interactions between ecology and evolution, whereby I start by ignoring phenotypes and vital rates (Fig. 8.3). To formalize this, I will make use of equations presented by Lion (2018). I will then discuss how to extend these equations to incorporate relations between phenotypes and vital rates, considering all processes shown in Fig. 8.1. Finally, I will propose some ideas on how this framework can be used to quantify genotypic differences in mean phenotype, plasticity and intra-genotypic variability, and, finally, to predict the evolution of those strategies in real-world populations.

The total environment a population experiences at a given point in time can be described by an environmental vector $\mathbf{E}$ (green boxes in Fig. 8.3). This vector consists of the densities of each genotype ($\mathbf{n}$) and a vector of all other environmental variables (both biotic and abiotic) ($\mathbf{e}$) (Lion 2018):

$$\mathbf{E} = \begin{pmatrix} \mathbf{n} \\ \mathbf{e} \end{pmatrix} \tag{8.1}$$

The change in the total population density n (which is the sum of all values in $\mathbf{n}$), can be described by:

$$\frac{dn}{dt} = \bar{r}(\mathbf{E})n \tag{8.2}$$

Here, $\bar{r}$ is the average per-capita population growth rate (fitness) of the entire population (including all genotypes), as a function of the environment $\mathbf{E}$. Evolution is a change in the frequency of a genotype i, whereby the frequency is calculated as the proportion of the density of individuals with genotype i compared to the total density ($f_i = n_i/n$). The change in f_i depends on f_i, and on the relative fitness of genotype i ($r_i(\mathbf{E})$ compared to $\bar{r}(\mathbf{E})$):

$$\frac{df_i}{dt} = f_i(r_i(\mathbf{E}) - \bar{r}(\mathbf{E})) \tag{8.3}$$

Equations 8.2 and 8.3 describe the link between changes in population density (an ecological factor) and changes in genotype frequencies (evolution), and can describe full eco-evolutionary feedback loops via both conspecific densities and other environmental factors (Box 8.2).
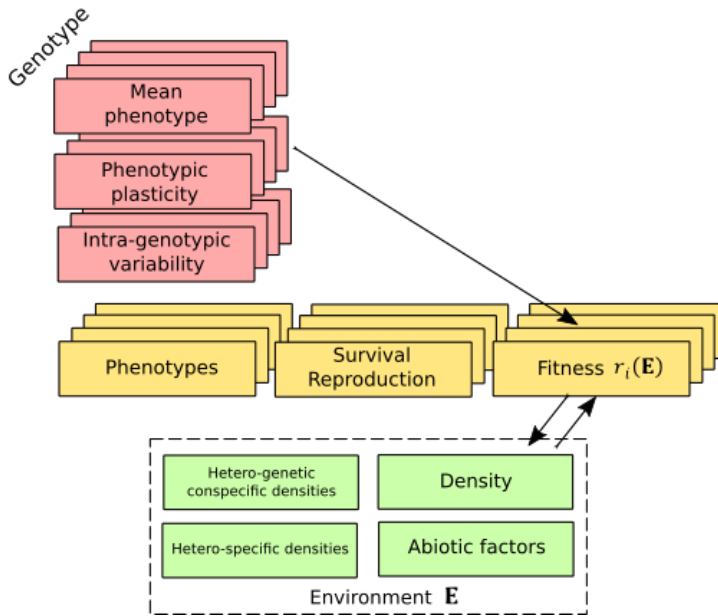
Figure 8.3: Flow chart similar to Figure 8.1, here showing the dynamics of multiple genotypes. Arrows indicate the processes as described by Equations 8.1-8.3. Fitness of genotype i ($r_i(\mathbf{E})$) is a function of the environment $\mathbf{E}$. Fitness differences among genotypes will result in a change in genotype frequencies (evolution) (Eq. 8.3). The resulting change in the genetic composition and densities, will alter $\mathbf{E}$. The changed environment feedbacks on each genotype's fitness, completing an eco-evolutionary feedback loop (Box 8.2). These simplified relations ignore that each genotype realizes fitness via phenotype dynamics and vital rates (lacking arrows 2, 3, 5, 6 in Figure 8.1).

Box 8.2 — Eco-evolutionary feedback loops via conspecific densities and other environmental factors. Equations 8.1-8.3 can describe a full eco-evolutionary feedback loop. Whenever $r_i(\mathbf{E}) \neq \bar{r}(\mathbf{E})$ for at least one genotype, ignoring stochastic events, genotype frequencies will change (evolution) (Eq. 8.3). Additionally, whenever the term $r_i(\mathbf{E}) - \bar{r}(\mathbf{E})$ for at least one genotype changes with n (as part of $\mathbf{E}$), evolution will be density-dependent. This shift in genotype frequencies may then alter $\bar{r}(\mathbf{E})$, which affects n (Eq. 8.2), completing the feedback loop. This makes density-dependence an important ingredient for eco-evolutionary feedback loops, and it is the type of feedback loop that was observed in the aphid populations, described in Turcotte et al. (2011b) and **Chapter 5**. Here, $r_i(\mathbf{E})$ of different aphid genotypes changed with density (Turcotte et al. 2013), resulting in a full eco-evolutionary feedback loop. Furthermore, it was shown that growth rates of populations consisting of two genotypes were higher compared to the non-evolving expectation, thus: $\bar{r}(\mathbf{E}) > (0.5 \cdot r_1(\mathbf{E}) + 0.5 \cdot r_2(\mathbf{E}))$, illustrating the importance of taking into account evolutionary dynamics when predicting population growth. Finally, the rate of evolution depended on whether populations were exposed to competitors and predators (captured in $\mathbf{e}$, Eq. 8.1). This emphasizes that $r_i(\mathbf{E})$, and thereby the evolutionary response, is a function of all components that the environment $\mathbf{E}$ consists of.

Although not explicitly tested, results from the empirical studies on water fleas (**Chapters**

3, 4), also suggest the presence of eco-evolutionary feedback loops. This is because genotypes responded differently to density (e.g. Fig. 3.2), resulting in changes in $r_i(\mathbf{E}) - \bar{r}(\mathbf{E})$ as densities change. For example, results suggest that, at low densities, fitness of lineage A3 is close to the fitness average among all lineages, but that it shows the strongest negative response to density. As a result, at densities higher than ~ 60 individuals, A3 shows the lowest fitness. In other words, this indicates that the evolutionary response (i.e. the change in the frequency of lineage A3) will depend on the density (indicating an eco-to-evo effect). Changing frequencies of the lineages will then affect the overall growth rate (indicating evo-to-eco effect), and thus completing a full feedback loop. Moreover, we show that for different water flea genotypes, relative fitness ($r_i(\mathbf{E}) - \bar{r}(\mathbf{E})$) does not only change with density, but also with temperature. This is illustrated in Fig. 3.5, where we predict that the evolutionary response is temperature-dependent. Although temperature is not part of the feedback loop (i.e. changing genotype frequencies do not affect temperature), it again emphasizes the importance of including all (relevant) environmental factors in predicting eco-evolutionary changes.

As vector $\mathbf{n}$ includes the densities of other genotypes, it can be used to describe interactions among genotypes and frequency-dependence. In principle, $r_i(\mathbf{E})$ can depend not only on the total density n, but also specifically on its own density n_i and/or the density of genotype j (n_j). When $r_i(\mathbf{E})$ is differently affected by n_j than by n_i, this implies genotype × genotype interactions. Genetic diversity can lead to an increased performance due to niche partitioning and facilitation (Hughes et al. 2008; Ellers et al. 2011). Moreover, fitness of genotype i can depend negatively on its own frequency f_i, resulting in frequency-dependent selection. When individuals of genotype j are phenotypically more divergent, they might compete less strongly with i, for instance due to specialization to different food types (Bolnick et al. 2011). This will be reflected by lower negative effects of j on i (and/or i on j) compared to i on i (and/or j on j). I here note the resemblance to Lotka-Volterra models, in which coefficients describe competition effects among species/genotypes.

In addition to effects through conspecific density, an eco-evolutionary feedback loop can be realized when genotypes differently affect the environmental vector $\mathbf{e}$, or in other words, when the genetic composition has an effect on the environment (which can be both through abiotic and biotic components; Fig. 8.3). Guppy genotypes, for instance, have been shown to have differential effects on ecosystem properties such as nutrient fluxes, invertebrate communities and leaf decomposition rates (Bassar et al. 2010b). Whenever changes in $\mathbf{e}$ subsequently alter fitness of different genotypes (for instance, if some genotypes can use food more efficiently; Jeyasingh et al. 2009), this will lead to a full feedback loop. This has been demonstrated in a rotifer-algae system: algal cell clumping (a defense-related trait) was shown to evolve when rotifer abundances increased, showing how ecological conditions affect evolution (Becks et al. 2012). Rotifer population growth rates were subsequently affected by prey evolution (Becks et al. 2012), completing an eco-evolutionary feedback loop through effects on $\mathbf{e}$. ∎

Fitness as the sum of phenotype-dependent survival and reproduction

Instead of directly relating genes to fitness (Fig. 8.3), thereby overlooking the fact that selection acts on phenotypes, I will now formally incorporate all processes shown in Fig. 8.1. The temporal change in abundance of a genotype (i.e. its fitness) is the sum of the average genotype-specific survival and reproduction rate (arrow 3 in Fig. 8.1). As natural selection acts on phenotypes (i.e. differential fitness among genotypes can arise *only because* different phenotypes are produced), survival and reproduction rates are a function of phenotypic traits (arrow 2 in Fig. 8.1). For simplicity, we assume that a phenotype can be summarized by a single variable z. We can rewrite fitness of genotype i to include phenotype-dependent survival and reproduction, as well as within-

individual phenotype changes (growth) and between-individual phenotype changes (inheritance). Doing this, we obtain an Integral Projection Model (IPM) (see for more information on IPMs e.g. Ellner and Rees 2006; Smallegange and Coulson 2013; Merow et al. 2014; Childs et al. 2016; Rees and Ellner 2016). An IPM describes, in discrete time, how a phenotypic distribution at time t changes into a phenotypic distribution at time $t + 1$:

$$w_i(t+1,z') = \int [S_i(z,\mathbf{E}) \cdot G_i(z',z,\mathbf{E}) + R_i(z,\mathbf{E}) \cdot D_i(z',z,\mathbf{E})]w_i(t,z)dz \qquad (8.4)$$

Here, $S_i(z,\mathbf{E})$ and $R_i(z,\mathbf{E})$ indicate genotype-specific reproduction and survival, for individuals with phenotype z. The terms $D_i(z',z,\mathbf{E})$ and $G_i(z',z,\mathbf{E})$ indicate how phenotypes are produced and develop due to inheritance and changes during a surviving individual's life, respectively. Fitness λ_i can now be calculated by summing over all trait values z' and taking the ratio between the density at time $t + 1$ and at time t:

$$\lambda_i = \frac{\int w_i(t+1,z')dz}{\int w_i(t,z)dz} \qquad (8.5)$$

Calculating fitness using Eqs 8.4-8.5 prior to implementing it into Eqs 8.2-8.3, and combining this with sensitivity analyses (Caswell 1978; Caswell 1989; Merow et al. 2014) enables the identification of the vital rates that contribute most to fitness differences among genotypes and environments (**Chapter 3-5**).

Phenotype dynamics with Integral Projection Models

Eqs 8.2 and 8.3 describe how evolutionary dynamics shape changes in population densities. Evidently, evolution does not only change population dynamics, but also phenotype dynamics (Fig. 8.1, see also **Chapter 2**), and this becomes clear by making fitness phenotype-dependent, using an IPM (Eqs 8.4, 8.5). These functions contain the required components to quantify all processes shown in Fig. 8.1 for a given genotype, as outlined in the next paragraphs.

The functions $D_i(z',z,\mathbf{E})$ and $G_i(z',z,\mathbf{E})$ contain information on the average expressed phenotype, the plastic response and the degree of intra-genotypic variability (i.e. arrow 1 in Fig. 8.1). First, the function $D_i(z',z,\mathbf{E})$ describes the average produced phenotype z' at birth, and $G_i(z',z,\mathbf{E})$ describes how a phenotype is expected to develop during an individual's life from phenotype z at time t, to phenotype z' at $t + 1$. When these functions (for a given $\mathbf{E}$) differ among genotypes, this implies a genetic component in average phenotype production. We found indications for such differences among water flea genotypes (specifically, in birth size and somatic growth; **Chapters 3, 4**). Second, phenotypic plasticity is described by effects of the environment ($\mathbf{E}$) on $D_i(z',z,\mathbf{E})$ and $G_i(z',z,\mathbf{E})$, as these effects describe the average change in the expected phenotypes in response to the environment (arrow 6). Differential effects of $\mathbf{E}$ on these functions among genotypes, indicate genotype-specific reaction norms (Gotthard et al. 1995). Third, variation in expected offspring phenotype and in expected within-individual phenotypic trajectory, are implemented in functions $D_i(z',z,\mathbf{E})$ and $G_i(z',z,\mathbf{E})$. In **Chapters 3, 4**, this is calculated as the standard deviation of the residuals, quantifying the amount of variation irrespective of the environment: intra-genotypic variability. In **Chapters 3, 4**, we calculated this variance across all genotypes. This could, however, also be calculated per genotype. Differences in the deviation among genotypes would indicate a genetic component in the control of variance, and this is precisely how intra-genotypic variability is incorporated in the model presented in **Chapter 6**.

The survival and reproduction functions $S_i(z,\mathbf{E})$ and $R_i(z,\mathbf{E})$ describe the relation between phenotypes and fitness components (arrow 2 in Fig. 8.1), and thus indicate the selection on

phenotype z. Phenotypes provide the raw material for natural selection to act upon. Therefore, differences in survival between genotype i and j (experiencing the same environment), *must* be caused by differences in the produced phenotypes. Finding an additional genotype effect on either survival or reproduction, implies that z does not describe all relevant phenotypic properties. Of course, in practice, it will be challenging (even impossible) to measure all relevant phenotypic properties. It is thus not surprising to find remaining genotype-effects on survival and reproduction after correcting for z (especially when considering only one phenotypic trait such as body size; **Chapters 3, 4**). Interactive effects of z and $\mathbf{E}$ on $S_i(z, \mathbf{E})$ and $R_i(z, \mathbf{E})$ (arrow 5 in Fig. 8.1) can be interpreted as changing the selection. In water fleas, we showed that temperature changed the selection, reflected by significant interactions between body size and temperature. Finally, integrating over all fitness components enables the calculation of environment-dependent growth rates as a measure for a genotype's fitness (Eqs 8.4-8.5; arrow 3 in Fig. 8.1), and subsequent changes in densities (arrow 4 in Fig. 8.1).

In theory, parameterizing such an IPM for each genotype and projecting the dynamics of each genotype, while taking into account changes in densities and other environmental variables (Figs 8.1, 8.3), will result in predictions on eco-evolutionary dynamics. However, in practice, fitting the required functions for each genotype will be nearly unfeasible, especially in sexual reproducing species in which genotypes are not entities that remain unchanged over time. I will conclude this section by providing some ideas on how to construct an IPM that we can use to predict eco-evolutionary dynamics without the need to parameterize functions for each genotype specifically.

Predicting evolutionary responses with IPMs

In **Chapter 2**, we constructed one IPM for all genotypes, and used it to quantify how selection, amongst others, contributed to changes in the population-average phenotype. We did so by combining an IPM with an age-structured version of the Price equation (Price 1970; Coulson and Tuljapurkar 2008; Frank 2012), proposed by Coulson et al. (2010). A limitation of this approach, as was shown, is that it only captures effects of phenotypic selection on trait dynamics, ignoring the genetic basis of these phenotypes. This approach can therefore not distinguish between trait dynamics generated in low heritability scenarios or in high heritability scenarios. To quantify how evolution contributes to phenotypic dynamics, a quantitative genetic framework is more appropriate (see **Chapter 2**), as this framework estimates how the genetic component of a phenotype changes through time (i.e. explicitly distinguishing between effects through arrows 1 and 6, see Fig. 8.1). In order to quantify how the evolutionary change subsequently alters population dynamics (arrows 2-4), this quantitative genetic framework can be combined with IPMs, as proposed by Coulson et al. (2017). Instead of defining an IPM for each genotype i (Eq. 8.4), this IPM would thus still focus on phenotypic dynamics while explicitly incorporating the genetic component of these dynamics.

In quantitative genetics, it is assumed that an individual's phenotype is determined by many unlinked loci with small additive effects (Fisher 1930). Assuming no genotype-environment interactions, a phenotype can be described by: $z = A + \varepsilon$, where A is the additive genetic component or breeding value (i.e. the portion of an individual's phenotype that is passed on to its offspring), and ε is a residual term (Falconer 1960). Assuming no epistasis or dominance, the residual term ε can be interpreted as the non-genetic, or environmental, contribution to z. Components A and ε can be estimated using a so-called 'animal model', which is a mixed-effects model (Kruuk 2004; Wilson et al. 2009). Moreover, the residual component ε can be further split into contributions of, for instance, sex, age, environmental drivers and maternal effects (Kruuk 2004; Wilson et al. 2009) (see also **Chapter 2**). This quantitative genetic model, separating the dynamics of a phenotypic trait z into the contribution of genetic and non-genetic factors, can be implemented in

all trait-dependent functions underlying the IPM.

By doing so, changes in phenotype dynamics and subsequent population dynamics, through different processes, can be quantified. Selection still acts on phenotypes, through phenotype-dependent survival and reproduction (arrow 3 in Fig. 8.1). However, mean offspring phenotype is no longer a function of the parental phenotype (as criticized by Chevin 2015). Instead, what is inherited, are the breeding values for the average trait. This makes it possible to track evolutionary changes in mean offspring phenotype, as this is a change in these breeding values over time. Further, the effect of an environmental variable on z (i.e. plasticity) can be estimated and implemented in the IPM. This will enable the quantification of environmental changes on phenotype and population dynamics. A deteriorating environment, for instance, can shift a phenotype distribution to one direction despite an opposing selection pressure. In sheep populations on the Scottish island of St. Kilda, it has been shown that average birth weight (plastically) decreases over time, despite positive selection on birth weight (Ozgul et al. 2009). Genetic effects on the plastic response can be incorporated by estimating the genetic component of this environmental effect (Wilson et al. 2009), and this will make it possible to make predictions on the evolution of plasticity. Intra-genotypic variability can be interpreted as the genetic component of the residual term (Hill and Mulder 2010), and implementing this will allow one to make predictions on the evolution of intra-genotypic variability. Cryptic variation, whereby variance can be released under certain environmental conditions, could be implemented as a correlation between an environmental factor and variation in breeding values (Gibson and Dworkin 2004; Ashander et al. 2016). Finally, covariances between the genetic components are estimated in the animal model, which will make it possible to project evolutionary changes while taking into account genetic correlations, for instance between the mean phenotypic trait and the degree of variance (Morgante et al. 2015).

Concluding, to predict how evolution affects phenotype dynamics, a quantitative genetic approach (separating phenotypes into genetic and non-genetic components) is more appropriate than a purely phenotype-based model. A combined QG-IPM model (Coulson et al. 2017) seems a promising and exciting future avenue of research for jointly tracking population and phenotype dynamics through time, via both evolutionary and environmental processes. The drawback is, that it will be very data-hungry to parameterize such a model, as it requires long-term individual data on phenotypes and vital rates, data on environmental variables, and a pedigree to estimate breeding values (Kruuk 2004; Wilson et al. 2009; Clutton-Brock and Sheldon 2010). However, as the costs required for (genetic) analysis decrease over time (De Barba et al. 2017), while techniques to automate data collection (**Chapter 7**) will become increasingly available due to higher computational power and skills, I am hopeful that we can start parameterizing such models in the near future. In the end, these developed frameworks can only prove their worth once we can apply them to real data.

8.4 Climate change: what can we expect for natural populations?

Climate change represents a selection pressure that natural populations are currently facing (Hoffmann and Sgrò 2011). In this dissertation, we have used water fleas (*Daphnia magna*) as a study species to improve our understanding in the potential consequences of climate change. Water fleas are key stone species in fresh water systems, as changes in their prevalence and functioning can have cascading effects on whole communities and ecosystems (Walther et al. 2002). Following Gienapp et al. (2008), natural populations facing climate change can respond in three ways: i) they can adapt by means of evolution, ii) individuals can acclimate through phenotypic plasticity, and iii) individuals can disperse. What do the results presented in this dissertation suggest for the persistence of natural (water flea) populations and for the importance of each of these three

responses?

My results indicate that water fleas show substantial within-population variation in fitness due to genotypic variation (**Chapter 3**). I showed that different genotypes can optimize fitness via different vital rates, with some genotypes profiting from increased survival, while other genotypes profiting from components related to reproduction. Furthermore, water fleas originating from the same population responded differently to density. As I outline in this synthesis, these differential responses can affect the direction and rate of evolution (Eqs. 8.2-8.3, Box 8.2). To be able to predict evolutionary responses, the above points emphasize the importance of considering the full life-cycle and incorporate density effects, instead of focusing on single life history traits, measured on isolated individuals. However, in commonly performed life-table studies, these important factors are generally overlooked (e.g. Geerts et al. 2015; Henning-Lucass et al. 2016; Sommer et al. 2016). The found genotypic differences in performance depended on the temperature, showing within-population variation in thermal tolerance. We predicted that the genetic composition of a water flea population can be substantially altered within a few months, and that the rate and direction of evolution depends on the temperature (**Chapter 3**). This variation in thermal tolerance suggests that natural populations of water fleas have, at least some, capacity for thermal adaptation, in agreement with earlier studies (Van Doorslaer et al. 2009a; Geerts et al. 2015).

At the same time, however, results indicate substantial between-population variation in performance. Genotypes originating from two Belgian ponds were better adapted to the tested conditions than genotypes from two Norwegian ponds. Within 80 days, most Belgian genotypes were able to outcompete the Norwegian clones to a large extent, and this was observed across all temperature treatments (**Chapter 4**). Moreover, we did not find any differences in thermal optima among genotypes originating from the two latitudes. Rather, in this experiment, individuals were able to adjust plastically to a wide range of temperatures, with all genotypes performing better at the higher tested temperatures (**Chapter 4**). Contrary to the expectation, we thus did not find indications for thermal adaptation, but instead, showed that the Belgian genotypes realize a higher fitness at all temperatures. These fitness differences between latitudes became even more pronounced when populations were infected with an ectoparasite (*Amoebidium parasiticum*). Temperature-associated environmental changes such as changes in parasite dynamics, can thus greatly impact fitness and these effects might even impose a stronger selection pressure on local populations than direct effects of temperature. Moreover, temperature highly influenced the dynamics of *A. parasiticum* themselves. This complicates predictions of evolutionary responses to climate change based on only responses to temperature, and underscores the importance of considering the dynamics of other abiotic and biotic factors altered by climate change (Labaude et al. 2017).

Migration in response to climate change is widely observed and can result in range shifts (Thomas and Lennon 1999; Hickling et al. 2006). Local populations risk being outcompeted by conspecific immigrants that are preadapted to the changing conditions (Van Doorslaer et al. 2009b), and this is in agreement with results of our competition experiment (**Chapter 4**). Especially when these immigrants bring pathogens to which they are better resistant or tolerant (Prenter et al. 2004), results presented in **Chapter 4** suggest that this can have an enormous impact on local gene pools. To conclude, the observed genetic basis in both the response to temperature and to associated (pathogen) changes, suggests that we might expect drastic changes in the genetic composition of populations as a result of climate change, both due to within-population selection and due to gene flow from climate migrants. As a final note, it is important to realize that the presented empirical studies were performed under controlled conditions, using (asexually reproducing) model species that experienced limited competition and no predation, and were exposed to constant abiotic factors. Obviously, natural populations have to cope with many more challenges at the same time which will complicate making predictions even more.

In addition to the three responses proposed by Gienapp et al. (2008) (i.e. adaptive tracking, acclimation and migration), I propose that intra-genotypic variability (for instance leading to a bet-hedging strategy) can act as a fourth mechanism by which populations can show an adaptive response towards global change. Extreme climate events are becoming more common, in addition to the predicted changes in mean temperature (IPCC 2014). This higher environmental stochasticity is predicted to lead to an increased production of phenotypic mismatches due to plasticity (Oostra et al. 2018). In this case, intra-genotypic variability may become more important, as intra-genotypic variability is expected to be favoured under increasing environmental variances (**Chapter 6**). Indeed, Donelson et al. (2018) recently suggested that the role of bet hedging in relation to climate change is underestimated.

It is clear that predicting the consequences of climate change for natural populations is a major challenge (Gienapp et al. 2008; Hoffmann and Sgrò 2011), due to the complexity of the relations between all involved biotic and abiotic factors. Even in cases where it is possible to predict phenotypic changes to climate change, without taking into account the genetic architecture, it will be impossible to predict the genetic changes. This is because selection, acting on phenotypes, is oblivious for the genetics underlying the realized phenotypes. Indeed, changes in multiple genetic pathways can result in the same phenotypic response (Becks et al. 2012; Ellner 2013). Future research should continue to explore modelling approaches to quantify and predict changes in phenotype and population dynamics, while disentangling plastic and genetic factors. In addition, the methodology presented in this dissertation, which is the linking of individual phenotypes embedded in a population, to fitness components and then to fitness, can be applied to natural populations. For this, long-term individual-based field studies will be indispensable (Clutton-Brock and Sheldon 2010). More specifically, we need field data on individual phenotypes and performance, on ecological drivers affecting these individuals, and on the relatedness between individuals. Such data sets are already available, for various populations of birds, e.g. great tits (Lack 1954; Mulder et al. 2016) and Darwin's finches (Grant and Grant 2002), and mammals, e.g. Soay sheep (Ozgul et al. 2009), yellow-bellied marmots (Ozgul et al. 2010) and chimpanzees (Lawick-Goodall van 1968), and have provided important insights in ecological and evolutionary questions (Clutton-Brock and Sheldon 2010). Future efforts should set up more such long-term studies, also focusing on other taxonomic (invertebrate) groups. This will further improve our understanding in how genetic variation, phenotypic plasticity and intra-genotypic variability shape the ability of natural populations to persist in the face of ever-changing biotic and abiotic conditions.

Bibliography

Acar, M., Mettetal, J. T., and Oudenaarden, A. van (2008). Stochastic switching as a survival strategy in fluctuating environments. *Nature Genetics* 40(4), 471–475.

Adler, P. B., Ellner, S. P., and Levine, J. M. (2010). Coexistence of perennial plants: an embarrassment of niches. *Ecology Letters* 13(8), 1019–1029.

Agrawal, A. A., Underwood, N., and Stinchcombe, J. R. (2004). Intraspecific variation in the strength of density dependence in aphid populations. *Ecological Entomology* 29(5), 521–526.

Ansel, J., Bottin, H., Rodriguez-Beltran, C., Damon, C., Nagarajan, M., Fehrmann, S., François, J., and Yvert, G. (2008). Cell-to-cell stochastic variation in gene expression is a complex genetic trait. *PLoS Genetics* 4(4), e1000049.

Arnold, S. J. (1992). Constraints on phenotypic evolution. *The American Naturalist* 140, S85–S107.

Arnold, S. J., Pfrender, M. E., and G., J. A. (2001). The Apaptive Landscape as a conceptual bridge between mico- and macroevolution. *Genetica* 112(1), 9–32.

Ashander, J., Chevin, L. M., and Baskett, M. L. (2016). Predicting evolutionary rescue via evolving plasticity in stochastic environments. *Proceedings of the Royal Society B: Biological Sciences* 283(1839), 20161690.

Atkinson, D. (1994). Temperature and organism size—a biological law for ectotherms? *Advances in Ecological Research* 25, 1–58.

Atkinson, D. (1995). Effects of temperature on the size of aquatic ectotherms: exceptions to the general rule. *Journal of Thermal Biology* 20(1), 61–74.

Auclerc, A., Libourel, P. A., Salmon, S., Bels, V., and Ponge, J. F. (2010). Assessment of movement patterns in *Folsomia candida* (Hexapoda: Collembola) in the presence of food. *Soil Biology and Biochemistry* 42(4), 657–659.

Auld, J. R., Agrawal, A. A., and Relyea, R. A. (2009). Re-evaluating the costs and limits of adaptive phenotypic plasticity. *Proceedings of the Royal Society of London B: Biological Sciences* 277(1681), 503–511.

Ayroles, J. F., Buchanan, S. M., O'Leary, C., Skutt-Kakaria, K., Grenier, J. K., Clark, A. G., Hartl, D. L., and Bivort, B. L. de (2015). Behavioral idiosyncrasy reveals genetic control of phenotypic variability. *Proceedings of the National Academy of Sciences* 112(21), 6706–6711.

Azur, M. J., Stuart, E. A., Frangakis, C., and Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *International Journal of Methods in Psychiatric Research* 20(1), 40–9.

Bánszegi, O., Kosztolányi, A., Bakonyi, G., Szabo, B., and Dombos, M. (2014). New method for automatic body length measurement of the collembolan, *Folsomia candida* Willem 1902 (Insecta: Collembola). *PLoS ONE* 9(6), e98230.

Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O'Shea, E., Pilpel, Y., and Barkai, N. (2006). Noise in protein expression scales with natural protein abundance. *Nature Genetics* 38(6), 636–643.

Barfield, M., Holt, R. D., and Gomulkiewicz, R. (2011). Evolution in stage-structured populations. *The American Naturalist* 177(4), 397–409.

Barnich, O. and Van Droogenbroeck, M. (2011). ViBe: A Universal Background Subtraction Algorithm for Video Sequences. *IEEE Transactions on Image Processing* 20(6), 1709–1724.

Bartoń, K. (2016). MuMIn: Multi-Model Inference. R package version 1.15.6. https://CRAN.R-project.org/package=MuMIn, https://cran.r–project.org/package=MuMIn.

Bassar, R. D., Lopéz-Sepulcre, A., Walsh, M. R., Turcotte, M. M., Torres-Mejia, M., Reznick, D. N., and Bassar, R. (2010a). Bridging the gap between ecology and evolution: integrating density regulation and life-history evolution. *Annals of the New York Academy of Sciences* 1206(1), 17–34.

Bassar, R. D., Marshall, M. C., López-Sepulcre, A., Zandonà, E., Auer, S. K., Travis, J., Pringle, C. M., Flecker, A. S., Thomas, S. A., Fraser, D. F., and Reznick, D. N. (2010b). Local adaptation in Trinidadian guppies alters ecosystem processes. *Proceedings of the National Academy of Sciences* 107(8), 3616–3621.

Bassar, R. D., Heatherly, T., Marshall, M. C., Thomas, S. A., Flecker, A. S., and Reznick, D. N. (2015). Population size-structure-dependent fitness and ecosystem consequences in Trinidadian guppies. *Journal of Animal Ecology* 84(4), 955–968.

Becks, L., Ellner, S. P., Jones, L. E., and Hairston, N. G. (2012). The functional genomics of an eco-evolutionary feedback loop: linking gene expression, trait evolution, and community dynamics. *Ecology Letters* 15(5), 492–501.

Benthem, K. J. van, Bruijning, M., Bonnet, T., Jongejans, E., Postma, E., and Ozgul, A. (2017). Disentangling evolutionary, plastic and demographic processes underlying

trait dynamics: A review of four frameworks. *Methods in Ecology and Evolution* 8(1), 75–85.

Bishop, A. L., Rab, F. A., Sumner, E. R., and Avery, S. V. (2007). Phenotypic heterogeneity can enhance rare-cell survival in 'stress-sensitive' yeast populations. *Molecular Microbiology* 63(2), 507–520.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. New York: Springer Science & Business Media.

Blake, W. J., Balázsi, G., Kohanski, M. A., and Isaacs, F. J. (2006). Phenotypic consequences of promoter-mediated transcriptional noise. *Molecular Cell* 24(6), 853–865.

Bolker, B. M., Gardner, B., Maunder, M., Berg, C. W., Brooks, M., Comita, L., Crone, E., Cubaynes, S., Davies, T., Valpine, P. de, Ford, J., Gimenez, O., Kéry, M., Kim, E. J., Lennert-Cody, C., Magnusson, A., Martell, S., Nash, J., Nielsen, A., Regetz, J., Skaug, H., and Zipkin, E. (2013). Strategies for fitting nonlinear ecological models in R, AD Model Builder, and BUGS. *Methods in Ecology and Evolution* 4(6), 501–512.

Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., Rudolf, V. H. W., Schreiber, S. J., Urban, M. C., and Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution* 26(4), 183–92.

Botero, C., Weissing, F., and Wright, J. (2015). Evolutionary tipping points in the capacity to adapt to environmental change. *Proceedings of the National Academy of Sciences* 112(1), 184–189.

Bourgeois, Y., Roulin, A. C., Müller, K., and Ebert, D. (2017). Parasitism drives host genome evolution: Insights from the *Pasteuria ramosa - Daphnia magna* system. *Evolution* 71(4), 1106–1113.

Bouwmans, T. (2014). Traditional and recent approaches in background modeling for foreground detection: An overview. *Computer Science Review* 11, 31–66.

Bradshaw, A. D. (1965). Evolutionary significance of phenotypic plasticity in plants. *Advances in genetics* 13, 115–155.

Brambilla, D. J. (1983). Microsporidiosis in a *Daphnia pulex* population. *Hydrobiologia* 99(3), 175–188.

Breton, L. M. and Addicott, J. F. (1992). Density-dependent mutualism in an aphid-ant interaction. *Ecology* 73(6), 2175–2180.

Brooks, M. E., Mugabo, M., Rodgers, G. M., Benton, T. G., and Ozgul, A. (2016). How well can body size represent effects of the environment on demographic rates? Disentangling correlated explanatory variables. *Journal of Animal Ecology* 85(2), 318–328.

Bruijning, M., Visser, M. D., Muller-Landau, H. C., Wright, S. J., Comita, L. S., Hubbell, S. P., Kroon, H. de, and Jongejans, E. (2017a). Surviving in a cosexual world: A cost-benefit analysis of dioecy in tropical trees. *The American Naturalist* 189(3), 297–314.

Bruijning, M., Visser, M. D., Hallmann, C. A., and Jongejans, E. (2017b). *trackdem*: Particle Tracking and Demography, https://cran.r–project.org/web/packages/trackdem/i.

Bruijning, M., ten Berge, A. C. M., and Jongejans, E. (2018a). Data for: Population-level responses to temperature, density and clonal differences in *Daphnia magna* as revealed by Integral Projection Modeling. *DANS EASY Archive (DOI: https://doi.org/10.17026/dans-2b8-gx7j)*.

Bruijning, M., ten Berge, A. C. M., and Jongejans, E. (2018b). Population-level responses to temperature, density and clonal differences in *Daphnia magna* as revealed by Integral Projection Modeling. *Functional Ecology* 32(10), 2407–2422.

Bruijning, M., Visser, M. D., Hallmann, C. A., and Jongejans, E. (2018c). *trackdem*: Automated particle tracking to obtain population counts and size distributions from videos in R. *Methods in Ecology and Evolution* 9(4), 965–973.

Bull, J. J. (1987). Evolution of phenotypic variance. *Evolution* 41(2), 303–315.

Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*. New York: Springer.

Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research* 33(2), 261–304.

Buuren, S. van and Groothuis-Oudshoorn, K. (2011). *mice*: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* 45(3), 1–67.

Cameron, T. C., O'Sullivan, D., Reynolds, A., Piertney, S. B., and Benton, T. G. (2013). Eco-evolutionary dynamics in response to selection on life-history. *Ecology Letters* 16(6), 754–763.

Canale, C. I., Ozgul, A., Allainé, D., and Cohas, A. (2016). Differential plasticity of size and mass to environmental change in a hibernating mammal. *Global Change Biology* 22(10), 3286–3303.

Carius, H. J., Little, T. J., and Ebert, D. (2001). Genetic variation in a host-parasite association: Potential for coevolution and frequency-dependent selection. *Evolution* 55(6), 1136–1145.

Carvalho, G. R. (1987). The clonal ecology of *Daphnia magna* (Crustacea: Cladocera): II. Thermal differentiation among seasonal clones. *Journal of Animal Ecology* 56(2), 469–478.

Caswell, H. (1978). A general formula for the sensitivity of population growth rate to changes in life history parameters. *Theoretical Population Biology* 14(2), 215–230.

Caswell, H. (2001). *Matrix population models: Construction, Analysis and Interpretation*. Sunderland, Massachusetts: Sinauer Associates.

Caswell, H. (1989). Analysis of life table response experiments I. Decomposition of effects on population growth rate. *Ecological Modelling* 46(3-4), 221–237.

Chalancon, G., Ravarani, C. N., Balaji, S., Martinez-Arias, A., Aravind, L., Jothi, R., and Babu, M. M. (2012). Interplay between gene expression noise and regulatory network architecture. *Trends in Genetics* 28(5), 221–232.

Chang, F., Chen, C. J., and Lu, C. J. (2004). A linear-time component-labeling algorithm using contour tracing technique. *Computer Vision and Image Understanding* 93(2), 206–220.

Chang, W., Cheng, J., Allaire, J. J., Xie, Y., and McPherson, J. (2017). shiny: Web Application Framework for R, https://cran.r–project.org/package=shiny.

Charmentier, A., Garant, D., and Kruuk, L. E. B. (2014). *Quantitative Genetics in the Wild*. Oxford: Oxford University Press.

Charpentier, A., Anand, M., and Bauch, C. T. (2012). Variable offspring size as an adaptation to environmental heterogeneity in a clonal plant species: integrating experimental and modelling approaches. *Journal of Ecology* 100(1), 184–195.

Chenouard, N., Smal, I., Chaumont, F. de, Maska, M., Sbalzarini, I. F., Gong, Y., Cardinale, J., Carthel, C., Coraluppi, S., Winter, M., Cohen, A. R., Godinez, W. J., Rohr, K., Kalaidzidis, Y., Liang, L., Duncan, J., Shen, H., Xu, Y., Magnuson, K. E. G., Jaldén, J., Blau, H. M., Paul-Gilloteaux, P., Kervrann, C., Waharte, F., Tivinez, J.-Y., Shorte, S. L., Willemse, J., Celler, K., Wezel, G. P. van, Dan, H.-W., Tsai, Y.-S., Solorzano, C. O. de, Olivo-Marin, J.-C., and Meijering, E. (2014). Objective comparison of particle tracking methods. *Nature Methods* 11(3), 281–289.

Chevin, L. M., Collins, S., and Lefèvre, F. (2013). Phenotypic plasticity and evolutionary demographic responses to climate change: taking theory out to the field. *Functional Ecology* 27(4), 967–979.

Chevin, L. M. (2015). Evolution of adult size depends on genetic variance in growth trajectories: a comment on analyses of evolutionary dynamics using integral projection models. *Methods in Ecology and Evolution* 6(9), 981–986.

Chiavelli, D. A., Mills, E. L., and Threlkeld, S. T. (1993). Host preference, seasonality, and community interactions of zooplankton epibionts. *Limnology and Oceanography* 38(3), 574–583.

Childs, D. Z., Coulson, T. N., Pemberton, J. M., Clutton-Brock, T. H., and Rees, M. (2011). Predicting trait values and measuring selection in complex life histories: reproductive allocation decisions in Soay sheep. *Ecology Letters* 14(10), 985–992.

Childs, D. Z., Metcalf, C. J. E., and Rees, M. (2010). Evolutionary bet-hedging in the real world: empirical evidence and challenges revealed by plants. *Proceedings of the Royal Society B: Biological Sciences* 277(1697), 3055–3064.

Childs, D. Z., Sheldon, B. C., and Rees, M. (2016). The evolution of labile traits in sex- and age-structured populations. *Journal of Animal Ecology* 85(2), 329–342.

Cleasby, I. R., Nakagawa, S., and Schielzeth, H. (2015). Quantifying the predictability of behaviour: statistical approaches for the study of between-individual variation in the within-individual variance. *Methods in Ecology and Evolution* 6(1), 27–37.

Clutton-Brock, T. and Sheldon, B. C. (2010). Individuals and populations: the role of long-term, individual-based studies of animals in ecology and evolutionary biology. *Trends in Ecology & Evolution* 25(10), 562–573.

Colbourne, J. K., Pfrender, M. E., and Gilbert, D. (2011). The ecoresponsive genome of *Daphnia pulex*. *Science* 331(6017), 555–561.

Collins, F. S. and Tabak, L. A. (2014). NIH plans to enhance reproducibility. *Nature* 505(7485), 612.

Coulson, T. and Tuljapurkar, S. (2008). The dynamics of a quantitative trait in an age-structured population living in a variable environment. *The American Naturalist* 172(5), 599–612.

Coulson, T., Tuljapurkar, S., and Childs, D. Z. (2010). Using evolutionary demography to link life history theory, quantitative genetics and population ecology. *Journal of Animal Ecology* 79(6), 1226–1240.

Coulson, T., MacNulty, D. R., Stahler, D. R., Wayne, R. K., and Smith, D. W. (2011). Modeling effects of environmental change on wolf population dynamics, trait evolution, and life history. *Science* 334(6060), 1275–1278.

Coulson, T., Plard, F., Schindler, S., and Ozgul, A. (2015). Quantitative genetics meets integral projection models: unification of widely used methods from ecology and evolution. *arXiv preprint arXiv: http://arxiv.org/abs/1509.01351*.

Coulson, T., Kendall, B. E., Barthold, J., Plard, F., Schindler, S., Ozgul, A., and Gaillard, J.-M. (2017). Modeling adaptive and nonadaptive responses of populations to environmental change. *The American Naturalist* 190(3), 313–336.

Cropper Jr, W. P., Holm, J. A., and Miller, C. J. (2012). An inverse analysis of a matrix population model using a genetic algorithm. *Ecological Informatics* 7(1), 41–45.

Crowley, P. H., Ehlman, S. M., Korn, E., and Sih, A. (2016). Dealing with stochastic environmental variation in space and time: bet hedging by generalist, specialist, and diversified strategies. *Theoretical Ecology* 9(2), 149–161.

Czypionka, T., Reeves, G., Vanhamel, M., and De Meester, L. (2016). Assessing hatching rates and the timing of hatching from plankton resting stages-an accurate and cost effective high throughput approach. *Limnology and Oceanography: Methods* 14(11), 718–724.

Darwin, C. (1859). *On the Origin of Species*. London: Murray.

David, O., Garnier, A., Larédo, C., and Lécomte, J. (2010). Estimation of plant demographic parameters from stage-structured censuses. *Biometrics* 66(3), 875–882.

De Barba, M., Miquel, C., Lobréaux, S., Quenette, P. Y., Swenson, J. E., and Taberlet, P. (2017). High-throughput microsatellite genotyping in ecology: improved accuracy, efficiency, standardization and success with low-quantity and degraded DNA. *Molecular Ecology Resources* 17(3), 492–507.

De Meester, L. (1993). Genotype, fish-mediated chemical, and phototactic behavior in *Daphnia magna*. *Ecology* 74(5), 1467–1474.

De Meester, L., Weider, L. J., and Tollrian, R. (1995). Alternative antipredator defenses and genetic-polymorphism in a pelagic predator-prey system. *Nature* 378(6556), 483–485.

De Meester, L., Van Doorslaer, W., Geerts, A., Orsini, L., and Stoks, R. (2011). Thermal genetic adaptation in the water flea *Daphnia* and its impact: an evolving metacommunity approach. *Integrative and Comparative Biology* 51(5), 703–718.

Decaestecker, E., Lefever, C., De Meester, L., and Ebert, D. (2004). Haunted by the past: Evidence for dormant stage banks of microparasites and epibionts of *Daphnia*. *Limnology and Oceanography* 49(4part2), 1355–1364.

Decaestecker, E., Gaba, S., Raeymaekers, J. A. M., Stoks, R., Van Kerckhoven, L., Ebert, D., and De Meester, L. (2007). Host-parasite 'Red Queen' dynamics archived in pond sediment. *Nature* 450(7171), 870–873.

Dekel, E. and Alon, U. (2005). Optimality and evolutionary tuning of the expression level of a protein. *Nature* 436(7050), 588–592.

Dell, A. I., Bender, J., Branson, K., Couzin, I. D., Polavieja, G. G. de, Noldus, L. P. J. J., Pérez-Escudero, A., Perona, P., Straw, A. D., Wikelski, M., and Brose, U. (2014). Automated image-based tracking and its application in ecology. *Trends in Ecology & Evolution* 29(7), 417–428.

DeWitt, T. J., Sih, A., and Wilson, D. S. (1998). Cost and limits of phenotypic plasticity. *Trends in Ecology & Evolution* 13(97), 77–81.

Diaz, S. A. and Viney, M. (2014). Genotypic-specific variance in *Caenorhabditis elegans* lifetime fecundity. *Ecology and Evolution* 4(11), 2058–2069.

Doak, D., Kareiva, P., and Klepetka, B. (1994). Modeling population viability for the desert tortoise in the western Mojave Desert. *Ecological Applications* 4(3), 446–460.

Donelson, J. M., Salinas, S., Munday, P. L., and Shama, L. N. S. (2018). Transgenerational plasticity and climate change experiments: Where do we go from here? *Global Change Biology* 24(1), 13–34.

Duchet, C., Coutellec, M.-A., Franquet, E., Lagneau, C., and Lagadic, L. (2010). Population-level effects of spinosad and *Bacillus thuringiensis israelensis* in *Daphnia pulex* and *Daphnia magna*: comparison of laboratory and field microcosm exposure conditions. *Ecotoxicology* 19(7), 1224–1237.

Dudycha, J. L. and Tessier, A. J. (1999). Natural genetic variation of life span, reproduction, and juvenile growth in *Daphnia*. *Evolution* 53(6), 1744–1756.

Easterling, M. R., Ellner, S. P., and Dixon, P. M. (2000). Size-specific sensitivity: applying a new structured population model. *Ecology* 81, 694–708.

Ebert, D. (1994). Virulence and local adaptation of a horizonally transmitted parasite. *Science* 265(5175), 1084–1086.

Ebert, D. (1995). The ecological interactions between a microsporidian parasite and its host *Daphnia magna*. *Journal of Animal Ecology* 64(3), 361.

Ebert, D. (2005). *Ecology, epidemiology, and evolution of parasitism in Daphnia*. Bethesda (MD): National Library of Medicine. National Center for Biotechnology Information.

Eddelbuettel, D. and Sanderson, C. (2014). RcppArmadillo: Accelerating R with high-performance C++ linear algebra. *Computational Statistics and Data Analysis* 71, 1054–1063.

Einum, S. and Fleming, I. A. (2000). Highly fecund mothers sacrifice offspring survival to maximize fitness. *Nature* 405(6786), 565–567.

Ellers, J., Rog, S., Braam, C., and Berg, M. P. (2011). Genotypic richness and phenotypic dissimilarity enhance population performance. *Ecology* 92(928), 1605–1615.

Ellner, S. P. and Rees, M. (2006). Integral projection models for species with complex demography. *The American Naturalist* 167(3), 410–428.

Ellner, S. P., Rees, M., Ellner, S. P., and Rees, M. (2007). Stochastic stable population growth in integral projection models: theory and application. *Journal of Mathematical Biology* 54(2), 227–256.

Ellner, S. P., Geber, M. A., and Hairston, N. G. (2011). Does rapid evolution matter? Measuring the rate of contemporary evolution and its impacts on ecological dynamics. *Ecology Letters* 14(6), 603–614.

Ellner, S. P. (2013). Rapid evolution: from genes to communities, and back again? *Functional Ecology* 27(5), 1087–1099.

Ellner, S. P., Childs, D. Z., and Rees, M. (2016). *Data-driven Modelling of Structured Populations. A Practical Guide to the Integral Projection Model*. Switzerland: Springer International Publishing.

Engen, S., Kvalnes, T., and Sæther, B.-E. (2014). Estimating phenotypic selection in age-structured populations by removing transient fluctuations. *Evolution* 68(9), 2509–2523.

Errington, T. M., Iorns, E., Gunn, W., Tan, F. E., Lomax, J., and Nosek, B. A. (2014). An open investigation of the reproducibility of cancer biology research. *eLife* 3, e04333.

Evans, M. E. K., Ferrière, R., Kane, M. J., and Venable, D. L. (2007). Bet hedging via seed banking in desert evening primroses (*Oenothera*, Onagraceae): demographic evidence from natural populations. *The American Naturalist* 169(2), 184–194.

Ezard, T. H. G., Côté, S. D., and Pelletier, F. (2009). Eco-evolutionary dynamics: disentangling phenotypic, environmental and population fluctuations. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364(1523), 1491–1498.

Færøvig, P. J., Andersen, T., and Hessen, D. O. (2002). Image analysis of *Daphnia* populations: non-destructive determination of demography and biomass in cultures. *Freshwater Biology* 47(10), 1956–1962.

Falconer, D. S. (1960). *Introduction to Quantitative Genetics*. Edinburgh and London: Oliver and Boyd.

Falconer, D. S. (1965). "Maternal effects and selection response". *Genetics today, Proceedings of the XI International Congress on Genetics, Vol. 3*. Oxford: Pergamon, 763–774.

Ficke, A. D., Myrick, C. A., and Hansen, L. J. (2007). Potential impacts of global climate change on freshwater fisheries. *Reviews in Fish Biology and Fisheries* 17(4), 581–613.

Fielding, A. (1999). *Machine Learning Methods for Ecological Applications*. Norwell, Massachusetts: Kluwer Academic Publishers.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Oxford University Press.

Fitzsimmons, J. M. and Innes, D. J. (2006). Inter-genotype variation in reproductive response to crowding among *Daphnia pulex*. *Hydrobiologia* 568(1), 187–205.

Formoso-Rafferty, N., Cervantes, I., Ibáñez-Escriche, N., and Gutiérrez, J. P. (2016). Genetic control of the environmental variance for birth weight in seven generations of a divergent selection experiment in mice. *Journal of Animal Breeding and Genetics* 133(3), 227–237.

Fowler, C. W. (1981). Density dependence as related to life history strategy. *Ecology* 62(3), 602–610.

Fox, G. A. and Kendall, B. E. (2002). Demographic stochasiticy and the variance reduction effect. *Ecology* 83(7), 1928–1934.

Frank, P., Boll, C., and Kelly, R. (1957). Vital statistics of laboratory cultures of *Daphnia pulex* DeGeer as related to density. *Physiological Zoology* 30(4), 287–305.

Frank, S. A. (2012). Natural selection. IV. The Price equation. *Journal of Evolutionary Biology* 25(6), 1002–1019.

Fraser, H. B., Hirsh, A. E., Giaever, G., Kumm, J., and Eisen, M. B. (2004). Noise minimization in eukaryotic gene expression. *PLoS Biology* 2(6), e137.

Fritsch, S. and Guenther, F. (2016). neuralnet: Training of Neural Networks, https://cran.r–project.org/package=neuralnet.

Furness, A. I., Lee, K., and Reznick, D. N. (2015). Adaptation in a variable environment: Phenotypic plasticity and bet-hedging during egg diapause and hatching in an annual killifish. *Evolution* 69(6), 1461–1475.

Fussmann, G. F., Loreau, M., and Abrams, P. A. (2007). Eco-evolutionary dynamics of communities and ecosystems. *Functional Ecology* 21(3), 465–477.

Gabsi, F., Glazier, D. S., Hammers-Wirtz, M., Ratte, H. T., and Preuss, T. G. (2014). How do interactive maternal traits and environmental factors determine offspring size in *Daphnia magna*? *International Journal of Limnology* 50(1), 9–18.

García-Roger, E. M., Serra, M., and Carmona, M. J. (2014). Bet-hedging in diapausing egg hatching of temporary rotifer populations - A review of models and new insights. *International Review of Hydrobiology* 99(1-2), 96–106.

Garreau, H., Bolet, G., Larzul, C., Robert-Granié, C., Saleil, G., SanCristobal, M., and Bodin, L. (2008). Results of four generations of a canalising selection for rabbit birth weight. *Livestock Science* 119(1-3), 55–62.

Geerts, A. N., Vanoverbeke, J., Vanschoenwinkel, B., Van Doorslaer, W., Feuchtmayr, H., Atkinson, D., Moss, B., Davidson, T. A., Sayer, C. D., and De Meester, L. (2015). Rapid evolution of thermal tolerance in the water flea *Daphnia*. *Nature Climate Change* 5(7), 665–668.

Gerum, R. C., Richter, S., Fabry, B., and Zitterbart, D. P. (2017). *ClickPoints*: an expandable toolbox for scientific image annotation and analysis. *Methods in Ecology and Evolution* 8(6), 750–756.

Ghosh, S., Gelfand, A., and Clark, J. (2012). Inference for size demography from point pattern data using integral projection models. *Journal of Agricultural, Biological, and Environmental Statistics* 17(4), 641–677.

Gibson, G. and Dworkin, I. (2004). Uncovering cryptic genetic variation. *Nature Reviews Genetics* 5(9), 681–690.

Giebelhausen, B. and Lampert, W. (2001). Temperature reaction norms of *Daphnia magna*: the effect of food concentration. *Freshwater Biology* 46(3), 281–289.

Gienapp, P., Teplitsky, C., Alho, J. S., Mills, J. A., and Merilä, J. (2008). Climate change and evolution: disentangling environmental and genetic responses. *Molecular Ecology* 17(1), 167–178.

Gliwicz, Z. M. (1990). Food thresholds and body size in cladocerans. *Nature* 343(6259), 638.

González, E. J., Martorell, C., and Bolker, B. M. (2016). Inverse estimation of integral projection model parameters using time series of population level data. *Methods in Ecology and Evolution* 7(2), 147–156.

Goser, B. and Ratte, H. T. (1994). Experimental evidence of negative interference in *Daphnia magna*. *Oecologia* 98(3-4), 354–361.

Gotthard, K., Nylin, S., and Nylin, S. (1995). Adaptive plasticity and plasticity as an adaptation: a selective review of plasticity in animal morphology and life history. *Oikos* 74(1), 3–17.

Govaert, L., Pantel, J. H., and De Meester, L. (2016). Eco-evolutionary partitioning metrics: assessing the importance of ecological and evolutionary contributions to population and community change. *Ecology Letters* 19(8), 839–853.

Graham, J. K., Smith, M. L., and Simons, A. M. (2014). Experimental evolution of bet hedging under manipulated environmental uncertainty in *Neurospora crassa*. *Proceedings of the Royal Society B-Biological Sciences* 281(1787), 20140706.

Grant, P. R. and Grant, B. R. (1995). Predicting microevolutionary responses to directional selection on heritable variation. *Evolution* 49(2), 241.

Grant, P. R. and Grant, B. R. (2002). Unpredictable evolution in a 30-year study of Darwin's finches. *Science* 296(5568), 707–11.

Grantham, M. E., Antonio, C. J., O'Neil, B. R., Zhan, Y. X., and Brisson, J. A. (2016). A case for a joint strategy of diversified bet hedging and plasticity in the pea aphid wing polyphenism. *Biology Letters* 12(10), 20160654.

Green, R. M., Fish, J. L., Young, N. M., Smith, F. J., Roberts, B., Dolan, K., Choi, I., Leach, C. L., Gordon, P., Cheverud, J. M., Roseman, C. C., Williams, T. J., Marcucio, R. S., and Hallgrímsson, B. (2017). Developmental nonlinearity drives phenotypic robustness. *Nature Communications* 8(1), 1970.

Gremer, J. R. and Venable, D. L. (2014). Bet hedging in desert winter annual plants: optimal germination strategies in a variable environment. *Ecology Letters* 17(3), 380–387.

Gremer, J. R., Kimball, S., and Venable, D. L. (2016). Within-and among-year germination in Sonoran Desert winter annuals: bet hedging and predictive germination in a variable environment. *Ecology Letters* 19(10), 1209–1218.

Gross, K., Craig, B. A., and Hutchison, W. D. (2002). Bayesian estimation of a demographic matrix model from stage-frequency data. *Ecology* 83(12), 3285–3298.

Grueber, C. E., Nakagawa, S., Laws, R. J., and Jamieson, I. G. (2011). Multimodel inference in ecology and evolution: challenges and solutions. *Journal of Evolutionary Biology* 24(4), 699–711.

Guisande, C. (1993). Reproductive strategy as population density varies in *Daphnia magna* (Cladocera). *Freshwater Biology* 29(3), 463–467.

Gust, K. A., Kennedy, A. J., Melby, N. L., Wilbanks, M. S., Laird, J., Meeks, B., Muller, E. B., Nisbet, R. M., and Perkins, E. J. (2016). *Daphnia magna*'s sense of competition: intra-specific interactions (ISI) alter life history strategies and increase metals toxicity. *Ecotoxicology* 25(6), 1126–1135.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R Package. *Journal of Statistical Software* 33(2), 1–22.

Hadfield, J. D., Wilson, A. J., Garant, D., Sheldon, B. C., and Kruuk, L. E. B. (2010). The misuse of BLUP in ecology and evolution. *The American Naturalist* 175(1), 116–125.

Hadfield, J. D., Wilson, A. J., and Kruuk, L. E. B. (2011). Cryptic evolution: does environmental deterioration have a genetic basis? *Genetics* 187(4), 1099–1113.

Hairston, N. G. and Munns, W. R. (1984). The timing of copepod diapause as an evolutionarily stable strategy. *The American Naturalist* 123(6), 733–751.

Hairston, N. G., Lampert, W., Cáceres, C. E., Holtmeier, C. L., Weider, L. J., Gaedke, U., Fisher, J. M., Fox, J. A., and Post, D. M. (1999). Lake ecosystems: rapid evolution revealed by dormant eggs. *Nature* 401(6752), 446–446.

Hairston, N. G., Ellner, S. P., Geber, M. A., Yoshida, T., and Fox, J. A. (2005). Rapid evolution and the convergence of ecological and evolutionary time. *Ecology Letters* 8(10), 1114–1127.

Hamilton, W. (1966). The moulding of senescence by natural selection. *Journal of Theoretical Biology* 12(1), 12–45.

Harris, K. D. M., Bartlett, N. J., and Lloyd, V. K. (2012). *Daphnia* as an emerging epigenetic model organism. *Genetics Research International* 2012, 1–8.

Hart, S. P., Schreiber, S. J., and Levine, J. M. (2016). How variation between individuals affects species coexistence. *Ecology Letters* 19(8), 825–838.

Harvell, C., Mitchell, C. E., Ward, J. R., Altizer, S., Dobson, A. P., Ostfeld, R. S., and Samuel, M. D. (2002). Climate warming and disease risks for terrestrial and marine biota. *Science* 296(5576), 2158–2162.

Henderson, C. (1950). Estimation of genetic parameters. *Annals of Mathematical Statistics* 21(2), 309–310.

Henderson, C. (1976). A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32(1), 69–83.

Henning-Lucass, N., Cordellier, M., Streit, B., and Schwenk, K. (2016). Phenotypic plasticity in life-history traits of *Daphnia galeata* in response to temperature–a comparison across clonal lineages separated in time. *Ecology and Evolution* 6(4), 881–891.

Heywood, J. S. (2005). An exact form of the breeder's equation for the evolution of a quantitative trait under natural selection. *Evolution* 59(11), 2287–2298.

Hickling, R., Roy, D. B., Hill, J. K., Fox, R., and Thomas, C. D. (2006). The distributions of a wide range of taxonomic groups are expanding polewards. *Global Change Biology* 12(3), 450–455.

Hill, W. G. (2007). The variance of the variance - genetic analysis of environmental variability. *Journal of Animal Breeding and Genetics* 124(2), 49–49.

Hill, W. G. and Mulder, H. A. (2010). Genetic analysis of environmental variation. *Genetic Research* 92(5-6), 381–395.

Hoefnagel, K. N., Vries, E. H. J. L. de, Jongejans, E., and Verberk, W. C. E. P. (2018). The temperature-size rule in *Daphnia magna* across different genetic lines and ontogenetic stages: Multiple patterns and mechanisms. *Ecology and Evolution* 8(8), 3828–3841.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). Bayesian Model Averaging: A Tutorial. *Statistical Science* 14(4), 382–401.

Hoffmann, A. A. and Sgrò, C. M. (2011). Climate change and evolutionary adaptation. *Nature* 470(7335), 479–485.

Holdich, D. M. and Reeve, I. D. (1991). Distribution of freshwater crayfish in the British Isles, with particular reference to crayfish plague, alien introductions and water quality. *Aquatic Conservation: Marine and Freshwater Ecosystems* 1(2), 139–158.

Hooper, H. L., Connon, R., Callaghan, A., Maund, S. J., Liess, M., Duquesne, S., Hutchinson, T. H., Moggs, J., and Sibly, R. M. (2006). The use of image analysis to estimate population growth rate in *Daphnia magna*. *Journal of Applied Ecology* 43(4), 828–834.

Horton, N. J. and Kleinman, K. P. (2007). Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. *The American Statistician* 61(1), 79–90.

Huey, R. and Berrigan, D. (2001). Temperature, demography, and ectotherm fitness. *The American Naturalist* 158(2), 204–210.

Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Underwood, N., and Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology Letters* 11(6), 609–623.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine* 2(8), e124.

IPCC (2014). *IPCC climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland: IPCC.

Janeiro, M. J., Coltman, D. W., Festa-Bianchet, M., Pelletier, F., and Morrissey, M. B. (2017). Towards robust evolutionary inference with integral projection models. *Journal of Evolutionary Biology* 30(2), 270–288.

Jansen, M., Coors, A., Stoks, R., and De Meester, L. (2011). Evolutionary ecotoxicology of pesticide resistance: a case study in *Daphnia*. *Ecotoxicology* 20(3), 543–551.

Jaqaman, K., Loerke, D., Mettlen, M., Kuwata, H., Grinstein, S., Schmid, S. L., and Danuser, G. (2008). Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods* 5(8), 695–702.

Jensen, J. L. W. V. (1906). Sur les fonctions convexes et les inegalites entre les valeurs moyennes. *Acta Mathematica* 30(1), 175–193.

Jeyasingh, P. D., Weider, L. J., and Sterner, R. W. (2009). Genetically-based trade-offs in response to stoichiometric food quality influence competition in a keystone aquatic herbivore. *Ecology Letters* 12(11), 1229–1237.

Jimenez-Gomez, J. M., Corwin, J. A., Joseph, B., Maloof, J. N., and Kliebenstein, D. J. (2011). Genomic analysis of QTLs and genes altering natural variation in stochastic noise. *PLoS Genetics* 7(9), e1002295.

Jiménez-Melero, R., Ramírez, J. M., and Guerrero, F. (2013). Seasonal variation in the population growth rate of a dominant zooplankter: what determines its population dynamics? *Freshwater Biology* 58(6), 1221–1233.

Jongejans, E., Shea, K., Skarpaas, O., Kelly, D., and Ellner, S. P. (2011). Importance of individual and environmental variation for invasive species spread: a spatial integral projection model. *Ecology* 92(1), 86–97.

Jongejans, E., Silverman, E. J., Skarpaas, O., and Shea, K. (2015). Post-dispersal seed removal of *Carduus nutans* and *C. acanthoides* by insects and small mammals. *Ecological Research* 30(1), 173–180.

Kain, J. S., Zhang, S., Akhund-Zade, J., and Samuel, A. D. T. (2015). Variability in thermal and phototactic preferences in *Drosophila* may reflect an adaptive bet-hedging strategy. *Evolution* 69(12), 3171–3185.

King, O. D. and Masel, J. (2007). The evolution of bet-hedging adaptations to rare scenarios. *Theoretical Population Biology* 72(4), 560–575.

Kingsolver, J. G., Hoekstra, H. E., Hoekstra, J. M., Berrigan, D., Vignieri, S. N., Hill, C. E., Hoang, A., Gibert, P., and Beerli, P. (2001). The strength of phenotypic selection in natural populations. *The American Naturalist* 157(3), 245–61.

Kingsolver, J. G. and Huey, R. B. (2008). Size, temperature, and fitness: three rules. *Evolutionary Ecology Research* 10(2), 251–268.

Kleiven, O. T., Larsson, P., and Hobæk, A. (1992). Sexual reproduction in *Daphnia* magna requires three stimuli. *Oikos* 65(2), 197–206.

Klinkhamer, P. G. L., Meelis, E., Jong, T. J. de, and Weiner, J. (1992). On the analysis of size-dependent reproductive output in plants. *Functional Ecology* 6(3), 308–316.

Koons, D. N., Pavard, S., Baudisch, A., and Metcalf, C. J. E. (2009). Is life-history buffering or lability adaptive in stochastic environments? *Oikos* 118(7), 972–980.

Krebs, M., Routtu, J., and Ebert, D. (2017). QTL mapping of a natural genetic polymorphism for long-term parasite persistence in *Daphnia* populations. *Parasitology* 144(13), 1686–1694.

Kroon, H. de, Plaisier, A., Groenendael, J. van, and Caswell, H. (1986). Elasticity: the relative contribution of demographic parameters to population growth rate. *Ecology* 67(5), 1427–1431.

Kroon, H. de, Groenendael, J. van, and Ehrlén, J. (2000). Elasticities: a review of methods and model limitations. *Ecology* 81(3), 607–618.

Kruuk, L. E. B. (2004). Estimating genetic parameters in natural populations using the 'animal model'. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 359(1446), 873–890.

Kruuk, L. E. B., Slate, J., Pemberton, J. M., Brotherstone, S., and Clutton-Brock, T. H. (2002). Antler size in red deer: hertability and selection but no evolution. *Evolution* 56(8), 1683–1695.

Kuno, G. (1973). Biological notes of *Amoebidium parasiticum* found in Puerto Rico. *Journal of Invertebrate Pathology* 21(1), 1–8.

Kuss, P., Rees, M., AEgisdóttir, H. H., Ellner, S. P., and Stöcklin, J. (2008). Evolutionary demography of long-lived monocarpic perennials: a time-lagged integral projection model. *Journal of Ecology* 96(4), 821–832.

Labaude, S., Rigaud, T., and Cézilly, F. (2017). Additive effects of temperature and infection with an acanthocephalan parasite on the shredding activity of *Gammarus*

fossarum (Crustacea: Amphipoda): the importance of aggregative behavior. *Global Change Biology* 23(4), 1415–1424.

Lack, D. (1954). *The Natural Regulation of Animal Numbers*. Oxford: The Clarendon Press.

Lampert, W. (1978). A field study on the dependence of the fecundity of *Daphnia spec.* on food concentration. *Oecologia* 36(3), 363–369.

Lande, R. (1976). Natural selection and random genetic drift in phenotype evolution. *Evolution* 30(2), 314–334.

Lande, R. (1979). Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* 33(1Part2), 402–416.

Lande, R. and Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution* 37(6), 1210.

Lavergne, S., Mouquet, N., Thuiller, W., and Ronce, O. (2010). Biodiversity and climate change: integrating evolutionary and ecological responses of species and communities. *Annual Review of Ecology, Evolution, and Systematics* 41, 321–350.

Lawick-Goodall van, J. (1968). The behaviour of free-living chimpanzees in the Gombe Stream Reserve. *Animal Behaviour Monographs* 1, 161–IN12.

Lehner, B. (2008). Selection to minimise noise in living systems and its implications for the evolution of gene expression. *Molecular Systems Biology* 4(1), 170.

Lion, S. (2018). Theoretical approaches in evolutionary ecology: environmental feedback as a unifying perspective. *The American Naturalist* 191(1), 21–44.

Lukacs, P. M., Burnham, K. P., Anderson, D. R., Lukacs, P. M., Burnham, K. P., and Anderson, D. R. (2010). Model selection bias and Freedman's paradox. *Annals of the Institute of Statistical Mathematics* 62(1), 117–125.

Lürling, M. and Tolman, Y. (2010). Effects of lanthanum and lanthanum-modified clay on growth, survival and reproduction of *Daphnia magna*. *Water Research* 44(1), 309–319.

Lynch, M. (1977). Fitness and optimal body size in zooplankton population. *Ecology* 58(4), 763–774.

Lynch, M. and Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland, Massachusetts: Sinauer Associates.

Lynch, M. and Walsh, B. (2014). *Evolution and Selection of Quantitative Traits*. Sunderland, Massachusetts: Sinauer Associates.

Ma, Z., Yu, L., and Chan, A. B. (2015). Small instance detection by integer programming on object density maps. *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 3689–3697.

MacArthur, J. W. and Baillie, W. H. T. (1929). Metabolic activity and duration of life. I. Influence of temperature on longevity in *Daphnia magna*. *Journal of Experimental Zoology* 53(2), 221–242.

Mackay, T. F. and Lyman, R. F. (2005). *Drosophila* bristles and the nature of quantitative genetic variation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 360(1459), 1513–1527.

Madigan, D. and Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. *Journal of the American Statistical Association* 89(428), 1535–1546.

Mallard, F., Bourlot, V. L., and Tully, T. (2013). An automated image analysis system to measure and count organisms in laboratory microcosms. *PLoS One* 8(5), e64387.

Manly, J. and Seyb, A. (1989). A comparison of three maximum likelihood models for stage-frequency data. *Researches on Population Ecology* 31(2), 367–380.

Martin, J. G. A., Pirotta, E., Petelle, M. B., and Blumstein, D. T. (2017). Genetic basis of between-individual and within-individual variance of docility. *Journal of Evolutionary Biology* 30(4), 796–805.

Martínez-García, R. and Tarnita, C. E. (2017). Seasonality can induce coexistence of multiple bet-hedging strategies in *Dictyostelium discoideum* via storage effect. *Journal of Theoretical Biology* 426, 1339–1351.

Matthews, B., Aebischer, T., Sullam, K. E., Lundsgaard-Hansen, B., and Seehausen, O. (2016). Experimental evidence of an eco-evolutionary feedback during adaptive divergence. *Current Biology* 26(4), 483–489.

McLean, N., Lawson, C. R., Leech, D. I., and Pol, M. (2016). Predicting when climate-driven phenotypic change affects population dynamics. *Ecology Letters* 19(6), 595–608.

Menu, F., Roebuck, J.-P., and Viala, M. (2000). Bet-Hedging diapause strategies in stochastic environments. *The American Naturalist* 155(6), 724–734.

Merilä, J. and Sheldon, B. C. (1999). Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. *Heredity* 83(2), 103.

Merilä, J. and Hendry, A. P. (2014). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications* 7(1), 1–14.

Merow, C., Dahlgren, J. P., Metcalf, C. J. E., Childs, D. Z., Evans, M. E., Jongejans, E., Record, S., Rees, M., Salguero-Gómez, R., and McMahon, S. M. (2014). Advancing population ecology with integral projection models: a practical guide. *Methods in Ecology and Evolution* 5(2), 99–110.

Metcalf, C. J. E., Rose, K. E., and Rees, M. (2003). Evolutionary demography of monocarpic perennials. *Trends in Ecology & Evolution* 18(9), 471–480.

Metcalf, C. J. E. and Pavard, S. (2007). Why evolutionary biologists should be demographers. *Trends in Ecology & Evolution* 22(4), 205–212.

Metcalf, C. J. E., Rose, K. E., Childs, D. Z., Sheppard, A. W., Grubb, P. J., and Rees, M. (2008). Evolution of flowering decisions in a stochastic, density-dependent environment. *Proceedings of the National Academy of Sciences* 105(30), 10466–10470.

Metcalf, C. J. E., Burghardt, L. T., and Koons, D. N. (2015). Avoiding the crowds: the evolution of plastic responses to seasonal cues in a density-dependent world. *Journal of Ecology* 103(4), 819–828.

Metcalf, C. J. E., Graham, A. L., Martinez-Bakker, M., and Childs, D. Z. (2016). Opportunities and challenges of Integral Projection Models for modelling host-parasite dynamics. *Journal of Animal Ecology* 85(2), 343–355.

Metzger, B. P. H., Yuan, D. C., Gruber, J. D., Duveau, F., and Wittkopp, P. J. (2015). Selection on noise constrains variation in a eukaryotic promoter. *Nature* 521(7552), 344–347.

Miller, T. E., Williams, J. L., Jongejans, E., Brys, R., and Jacquemyn, H. (2012). Evolutionary demography of iteroparous plants: Incorporating non-lethal costs of reproduction into integral projection models. *Proceedings of the Royal Society B: Biological Sciences* 279(1739), 2831–2840.

Miner, B. E., De Meester, L., Pfrender, M. E., Lampert, W., and Hairston, N. G. (2012). Linking genes to communities and ecosystems: *Daphnia* as an ecogenomic model. *Proceedings of the Royal Society B: Biological Sciences* 279(1735), 1873–1882.

Mitchell, S. E. and Lampert, W. (2000). Temperature adaptation in a geographically widespread zooplankter, *Daphnia magna*. *Journal of Evolutionary Biology* 13(3), 371–382.

Morgante, F., Sørensen, P., Sorensen, D. A., Maltecca, C., and Mackay, T. F. C. (2015). Genetic architecture of micro-environmental plasticity in *Drosophila melanogaster*. *Scientific Reports* 5(1), 9785.

Morley, N. J. and Lewis, J. W. (2014). Temperature stress and parasitism of endothermic hosts under climate change. *Trends in Parasitology* 30(5), 221–227.

Morris, T. P., White, I. R., and Royston, P. (2014). Tuning multiple imputation by predictive mean matching and local residual draws. *BMC Medical Research Methodology* 14(1), 75.

Morrissey, M. B., Kruuk, L. E. B., and Wilson, A. J. (2010). The danger of applying the breeder's equation in observational studies of natural populations. *Journal of Evolutionary Biology* 23(11), 2277–88.

Morrissey, M. B., Walling, C. A., Wilson, A. J., Pemberton, J. M., Clutton-Brock, T. H., and Kruuk, L. E. B. (2012a). Genetic analysis of life-history constraint and evolution in a wild ungulate population. *The American Naturalist* 179(4), E97–E114.

Morrissey, M. B., Parker, D. J., Korsten, P., Pemberton, J. M., Kruuk, L. E. B., and Wilson, A. J. (2012b). The prediction of adaptive evolution: empirical application of the secondary theorem of selection and comparison to the breeder's equation. *Evolution* 66(8), 2399–2410.

Mouritsen, K. N., Tompkins, D. M., and Poulin, R. (2005). Climate warming may cause a parasite-induced collapse in coastal amphipod populations. *Oecologia* 146(3), 476–483.

Mulder, H. A., Gienapp, P., and Visser, M. E. (2016). Genetic variation in variability: Phenotypic variability of fledging weight and its evolution in a songbird population. *Evolution* 70(9), 2004–2016.

NEN6503 (1980). Methodebeschrijvingen voor de beoordeling van verontreinidgde waterbodems volgens de TRIADE benadering.

Newman, J. R., Ghaemmaghami, S., Ihmels, J., Breslow, D. K., Noble, M., DeRisi, J. L., and Weissman, J. S. (2006). Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* 441(7095), 840.

Noldus, L. P. J. J., Spink, A. J., and Tegelenbosch, R. A. J. (2002). Computerised video tracking, movement analysis and behaviour recognition in insects. *Computers and Electronics in Agriculture* 35(2), 201–227.

Olofsson, H., Ripa, J., and Jonzén, N. (2009). Bet-hedging as an evolutionary game: the trade-off between egg size and number. *Proceedings of the Royal Society of London B: Biological Sciences* 276(1669), 2963–2969.

Oostra, V., Saastamoinen, M., Zwaan, B. J., and Wheat, C. W. (2018). Strong phenotypic plasticity limits potential for evolutionary responses to climate change. *Nature Communications* 9(1), 1005.

Ordas, B., Malvar, R. A., and Hill, W. G. (2008). Genetic variation and quantitative trait loci associated with developmental stability and the environmental correlation between traits in maize. *Genetics Research* 90(5), 385–395.

Ozgul, A., Tuljapurkar, S., Benton, T. G., Pemberton, J. M., Clutton-Brock, T. H., and Coulson, T. (2009). The dynamics of phenotypic change and the shrinking sheep of St. Kilda. *Science* 325(5939), 464–467.

Ozgul, A., Childs, D. Z., Oli, M. K., Armitage, K. B., Blumstein, D. T., Olson, L. E., Tuljapurkar, S., and Coulson, T. (2010). Coupled dynamics of body mass and population growth in response to environmental change. *Nature* 466(7305), 482–485.

Ozgul, A., Coulson, T., Reynolds, A., Cameron, T. C., and Benton, T. G. (2012). Population responses to perturbations: the importance of trait-based analysis illustrated through a microcosm experiment. *The American Naturalist* 179(5), 582–594.

Pantel, J. H., Duvivier, C., and De Meester, L. (2015). Rapid local adaptation mediates zooplankton community assembly in experimental mesocosms. *Ecology Letters* 18(10), 992–1000.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics* 37, 637–669.

Pauwels, K., Stoks, R., Decaestecker, E., and De Meester, L. (2007). Evolution of heat shock protein expression in a natural population of *Daphnia magna*. *The American Naturalist* 170(5), 800–805.

Pellegrini, S., Ess, A., Schindler, K., and Van Gool, L. (2009). You'll never walk alone: Modeling social behavior for multi-target tracking. *2009 IEEE 12th International Conference on Computer Vision*, 261–268.

Pelletier, F., Clutton-Brock, T., Pemberton, J., Tuljapurkar, S., and Coulson, T. (2007). The evolutionary demography of ecological change: linking trait variation and population growth. *Science* 315(5818), 1571–1574.

Pelletier, F., Garant, D., and Hendry, A. P. (2009). Eco-evolutionary dynamics. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364(1523), 1483–1489.

Pelletier, F., Moyes, K., Clutton-Brock, T. H., and Coulson, T. (2012). Decomposing variation in population growth into contributions from environment and phenotypes in an age-structured population. *Proceedings of the Royal Society of London B: Biological Sciences* 279(1727), 394–401.

Pennekamp, F. and Schtickzelle, N. (2013). Implementing image analysis in laboratory-based experimental systems for ecology and evolution: a hands-on guide. *Methods in Ecology and Evolution* 4(5), 483–492.

Pennekamp, F., Schtickzelle, N., and Owen, P. L. (2015). BEMOVI, software for extracting behavior and morphology from videos, illustrated with analyses of microbes. *Ecology and Evolution* 5(13), 2584–2595.

Pérez-Escudero, A., Vicente-Page, J., Hinz, R. C., Arganda, S., and De Polavieja, G. G. (2014). idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nature Methods* 11(7), 743–748.

Perry, G. M. L., Nehrke, K. W., Bushinsky, D. A., Reid, R., Lewandowski, K. L., Hueber, P., and Scheinman, S. J. (2012). Sex modifies genetic effects on residual variance in urinary Calcium excretion in rat (*Rattus norvegicus*). *Genetics* 191(3), 1003–1013.

Petrovskii, S. and Petrovskaya, N. (2012). Computational ecology as an emerging science. *Interface Focus* 2(2), 241–254.

Philippi, T. and Seger, J. (1989). Hedging one's evolutionary bets, revisited. *Trends in Ecology & Evolution* 4(2), 41–44.

Philippi, T. (1993). Bet-hedging germination of destert annuals: Beyond the first year. *The American Naturalist* 142(3), 474–487.

Piersma, T. and Drent, J. (2003). Phenotypic flexibility and the evolution of organismal design. *Trends in Ecology & Evolution* 18(5), 228–233.

Pietrzak, B. (2011). Interclonal differences in age-specific performance in *Daphnia magna*. *Journal of Limnology* 70(2), 345–352.

Pimentel, D. (1961). Animal population regulation by the genetic feed-back mechanism. *The American Naturalist* 95(881), 65–79.

Plummer, M. (2016). rjags: Bayesian Graphical Models using MCMC, https://cran.r–project.org/package=rjags.

Ponti, B., Piscia, R., Bettinetti, R., and Manca, M. (2010). Long-term adaptation of *Daphnia* to toxic environment in Lake Orta: the effects of short-term exposure to copper and acidification. *Journal of Limnology* 69(2), 217–224.

Post, D. M. and Palkovacs, E. P. (2009). Eco-evolutionary feedbacks in community and ecosystem ecology: interactions between the ecological theatre and the evolutionary play. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364(1523), 1629–1640.

Postma, E. (2006). Implications of the difference between true and predicted breeding values for the study of natural selection and micro-evolution. *Journal of Evolutionary Biology* 19(2), 309–320.

Postma, E. (2014). "Four decades of estimating heritabilities in wild vertebrate populations: improved methods, more data better estimates?" *Quantitative Genetics in the Wild*. Oxford: Oxford University Press.

Prenter, J., MacNeil, C., Dick, J. T. A., and Dunn, A. M. (2004). Roles of parasites in animal invasions. *Trends in Ecology & Evolution* 19(7), 385–390.

Price, G. R. (1970). Selection and covariance. *Nature* 227, 520–521.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R
    Foundation for Statistical Computing. Vienna, Austria.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R
    Foundation for Statistical Computing. Vienna, Austria.

Råberg, L., Sim, D., Read, A. F., and Raberg, L. (2007). Disentangling genetic variation
    for resistance and tolerance to infectious diseases in animals. *Science* 318(5851),
    812–814.

Raftery, A. E., Madigan, D., and Hoeting, J. A. (1997). Bayesian model averaging for
    linear regression models. *Journal of the American Statistical Association* 92(437),
    179–191.

Raj, A., Rifkin, S. A., Andersen, E., and Oudenaarden, A. V. (2010). Variability in gene
    expression underlies partial penetrance in *Caenorhabditis elegans*. *Nature* 12(7283),
    913–918.

Rajon, E., Desouhant, E., Chevalier, M., Débias, F., and Menu, F. (2014). The evolution
    of bet hedging in response to local ecological conditions. *The American Naturalist*
    184(1), E1–E15.

Raser, J. M. and O'Shea, E. K. (2005). Noise in gene expression: Origins, consequences,
    and control. *Science* 309(5743), 2010–2013.

Rausher, M. D. (1992). The measurement of selection on quantitative traits: biases due to
    environmental covariances between traits and fitness. *Evolution* 46(3), 616–626.

Réale, D., McAdam, A. G., Boutin, S., and Berteaux, D. (2003). Genetic and plastic
    responses of a northern mammal to climate change. *Proceedings of the Royal Society
    of London B: Biological Sciences* 270(1515), 591–596.

Rebke, M., Coulson, T., Becker, P. H., and Vaupel, J. W. (2010). Reproductive improve-
    ment and senescence in a long-lived bird. *Proceedings of the National Academy of
    Sciences* 107(17), 7841–7846.

Rebke, M. (2012). From the Price equation to a decomposition of population change.
    Journal of Ornithology 152(2), 555–559.

Reed, T. E., Robin, S. W., Schindler, D. E., Hard, J. J., and Kinnison, M. T. (2010).
    Phenotypic plasticity and population viability: The importance of environmental
    predictability. *Proceedings of the Royal Society B: Biological Sciences* 277(1699),
    3391–3400.

Rees, M., Childs, D. Z., Rose, K. E., and Grubb, P. J. (2004). Evolution of size-dependent
    flowering in a variable environment: Partitioning the effects of fluctuating selection.
    Proceedings of the Royal Society B: Biological Sciences 271(1538), 471–475.

Rees, M., Childs, D. Z., Metcalf, C. J. E., Rose, K. E., Sheppard, A. W., and Grubb,
    P. J. (2006). Seed dormancy and delayed flowering in monocarpic plants: selective
    interactions in a stochastic environment. *The American Naturalist* 168(2), E53–E71.

Rees, M. and Ellner, S. P. (2009). Integral projection models for populations in temporally
    varying environments. *Ecological Monographs* 79(4), 575–594.

Rees, M., Metcalf, C. J. E., and Childs, D. Z. (2010). Bet-hedging as an evolutionary
    game : the trade-off between egg size and number. *Proceedings of the Royal Society
   of London B: Biological Sciences* 277(1685), 1149–1151.

Rees, M. and Ellner, S. P. (2016). Evolving integral projection models: Evolutionary demography meets eco-evolutionary dynamics. *Methods in Ecology and Evolution* 7(2), 157–170.

Reznick, D., Nunney, L., and Tessier, A. (2000). Big houses, big cars, superfleas and the costs of reproduction. *Trends in Ecology & Evolution* 15(10), 421–425.

Reznick, D. N. and Ghalambor, C. K. (2001). The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* 112, 183–198.

Reznick, D. (2013). A critical look at reciprocity in ecology and evolution: Introduction to the symposium. *The American Naturalist* 181(S1), S1–S8.

Ripa, J., Olofsson, H., and Jonzén, N. (2010). What is bet-hedging, really? *Proceedings of the Royal Society B: Biological Sciences* 277(1685), 1153–1154.

Robertson, A. (1966). A mathematical model of the culling process in dairy cattle. *Animal Science* 8(1), 95–108.

Roff, D. A. (2007). A centennial celebration for quantitative genetics. *Evolution* 61(5), 1017–1032.

Ronget, V., Garratt, M., Lemaître, J.-F., and Gaillard, J.-M. (2017). The 'Evo-Demo' implications of condition-dependent mortality. *Trends in Ecology & Evolution* 32(12), 909–921.

Rönnegård, L. and Valdar, W. (2011). Detecting major genetic loci controlling phenotypic variability in experimental crosses. *Genetics* 188(2), 435–447.

Rönnegård, L. and Valdar, W. (2012). Recent developments in statistical methods for detecting genetic loci affecting phenotypic variability. *BMC Genetics* 13(1), 63.

Ros, M., Sorensen, D., Waagepetersen, R., Dupont-Nivet, M., SanCristobal, M., Bonnet, J. C., and Mallard, J. (2004). Evidence for genetic control of adult weight plasticity in the snail *Helix aspersa*. *Genetics* 168(4), 2089–2097.

Rossiter, M. C. (1991). Maternal effects generate variation in life history: Consequences of egg weight plasticity in the gypsy moth. *Functional Ecology* 5(3), 386.

Sae-Lim, P., Kause, A., Janhunen, M., Vehviläinen, H., Koskinen, H., Gjerde, B., Lillehammer, M., and Mulder, H. A. (2015). Genetic (co)variance of rainbow trout (*Oncorhynchus mykiss*) body weight and its uniformity across production environments. *Genetics Selection Evolution* 47(1), 46.

Saether, B. E. and Bakke, O. (2000). Avian life history variation and contribution of demographic traits to the population growth rate. *Ecology* 81(3), 642–653.

Sæther, B. E. and Engen, S. (2015). The concept of fitness in fluctuating environments. *Trends in Ecology & Evolution* 30(5), 273–281.

Salguero-Gómez, R. and Plotkin, J. B. (2010). Matrix dimensions bias demographic inferences: implications for comparative plant demography. *The American Naturalist* 176(6), 710–722.

Samson, D. A. and Werk, K. S. (1986). Size-dependent effects in the analysis of reproductive effort in plants. *The American Naturalist* 127(5), 667–680.

Savin, T. and Doyle, P. S. (2005). Static and dynamic errors in particle tracking microrheology. *Biophysical Journal* 88(1), 623–638.

Sbalzarini, I. F. and Koumoutsakos, P. (2005). Feature point tracking and trajectory analysis for video imaging in cell biology. *Journal of Structural Biology* 151(2), 182–195.

Schaub, M. and Abadi, F. (2011). Integrated population models: a novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology* 152(1), 227–237.

Scheiner, S. M. (1993). Genetics and evolution of phenotypic plasticity. *Annual Review of Ecology and Systematics* 24(1), 35–68.

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9(7), 671.

Schoener, T. W. (2011). The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. *Science* 331(6016), 426–429.

Shen, X., Pettersson, M., Rönnegård, L., and Carlborg, Ö. (2012). Inheritance beyond plain heritability: Variance-controlling genes in *Arabidopsis thaliana*. *PLoS Genetics* 8(8), e1002839.

Shreif, Z. and Periwal, V. (2014). A network characteristic that correlates environmental and genetic robustness. *PLoS Computational Biology* 10(2), e1003474.

Simons, A. M. (2011). Modes of response to environmental change and the elusive empirical evidence for bet hedging. *Proceedings of the Royal Society of Londen B: Biological Sciences* 278(1712), 1601–1609.

Smallegange, I. M. and Coulson, T. (2013). Towards a general, population-level understanding of eco-evolutionary change. *Trends in Ecology & Evolution* 28(3), 143–148.

Smith, C. C. and Fretwell, S. D. (1974). The optimal balance between size and number of offspring. *The American Naturalist* 108(962), 499–506.

Snell, T. W. and King, C. E. (1977). Lifespand and fecundity in rofiters: the cost of reproduction. *Evolution* 31(4), 882–890.

Soleymani, A., Pennekamp, F., Petchey, O. L., and Weibel, R. (2015). Developing and integrating advanced movement features improves automated classification of ciliate species. *PLoS ONE* 10(12), e145345.

Sommer, S., Piscia, R., Manca, M. M., Fontaneto, D., and Ozgul, A. (2016). Demographic cost and mechanisms of adaptation to environmental stress in resurrected *Daphnia*. *Journal of Limnology* 75(s2), 30–35.

Stearns, S. C. (1989). Trade-offs in life-history evolution. *Functional Ecology* 3(3), 259–268.

Stinchcombe, J. R., Simonsen, A. K., and Blows, M. W. (2014). Estimating uncertainty in multivariate responses to selection. *Evolution* 68(4), 1188–1196.

Stirnadel, H. A. and Ebert, D. (1997). Prevalence, host specificity and impact on host fecundity of microparasites and epibionts in three sympatric *Daphnia* species. *Journal of Animal Ecology* 66(2), 212–222.

Stoks, R., Govaert, L., Pauwels, K., Jansen, B., and De Meester, L. (2016). Resurrecting complexity: the interplay of plasticity and rapid evolution in the multiple trait response to strong changes in predation pressure in the water flea *Daphnia*. *Ecology Letters* 19(2), 180–190.

Stollewerk, A. (2010). The water flea *Daphnia* - a 'new' model system for ecology and evolution? *Journal of Biology* 9(2), 21.

Strauss, S. Y. (2014). Ecological and evolutionary responses in complex communities: implications for invasions and eco-evolutionary feedbacks. *Oikos* 123(3), 257–266.

Szeliski, R. (2011). *Computer Vision: Algorithms and Applications*. London: Springer Science & Business Media.

Tarazona, E., García-Roger, E. M., and Carmona, M. J. (2017). Experimental evolution of bet hedging in rotifer diapause traits as a response to environmental unpredictability. *Oikos* 126(8), 1162–1172.

Thomas, C. D. and Lennon, J. J. (1999). Birds extend their ranges northwards. *Nature* 399(6733), 213.

Thompson, J. N. (1998). Rapid evolution as an ecological process. *Trends in Ecology & Evolution* 13(8), 329–332.

Threlkeld, S. T. and Willey, R. L. (1993). Colonization, interaction, and organization of cladoceran epibiont communities. *Limnology and Oceanography* 38(3), 584–591.

Tilman, D., Wedin, D., and Knops, J. (1996). Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* 379(6567), 718–720.

Traill, L. W., Schindler, S., and Coulson, T. (2014). Demography, not inheritance, drives phenotypic change in hunted bighorn sheep. *Proceedings of the National Academy of Sciences* 111(36), 13223–13228.

Turcotte, M. M., Reznick, D. N., and Hare, J. D. (2011a). Experimental assessment of the impact of rapid evolution on population dynamics. *Evolutionary Ecology Research* 13(2), 113–131.

Turcotte, M. M., Reznick, D. N., and Hare, J. D. (2011b). The impact of rapid evolution on population dynamics in the wild: experimental test of eco-evolutionary dynamics. *Ecology Letters* 14(11), 1084–1092.

Turcotte, M. M., Reznick, D. N., and Daniel Hare, J. (2013). Experimental test of an eco-evolutionary dynamic feedback loop between evolution and population density in the green peach aphid. *The American Naturalist* 181(S1), S46–57.

Tylianakis, J. M., Didham, R. K., Bascompte, J., and Wardle, D. A. (2008). Global change and species interactions in terrestrial ecosystems. *Ecology Letters* 11(12), 1351–1363.

Urbanek, S. (2013). png: Read and write PNG images, https://CRAN.R–project.org/package=png.

Van Doorslaer, W., Stoks, R., Duvivier, C., Bednarshka, A., and De Meester, L. (2009a). Population dynamics determine genetic adaptation to temperature in *Daphnia*. *Evolution* 63(7), 1867–1878.

Van Doorslaer, W., Vanoverbeke, J., Duvivier, C., Rousseaux, S., Jansen, M., Jansen, B., Feuchtmayr, H., Atkinson, D., Moss, B., Stoks, R., and De Meester, L. (2009b). Local adaptation to higher temperatures reduces immigration success of genotypes from a warmer region in the water flea *Daphnia*. *Global Change Biology* 15(12), 3046–3055.

Van Doorslaer, W., Stoks, R., Swillen, I., Feuchtmayr, H., Atkinson, D., Moss, B., and De Meester, L. (2010). Experimental thermal microevolution in community-embedded *Daphnia* populations. *Climate Research* 43(1-2), 81–89.

VanDerWal, J., Falconi, L., Januchowski, S., Shoo, L., and Storlie, C. (2014). SDMTools: Species Distribution Modelling Tools: Tools for processing data associated with species distribution modelling exercises, https://CRAN.R–project.org/package=SDMTools.

Vezzani, R. and Cucchiara, R. (2010). Video surveillance online repository (ViSOR): an integrated framework. *Multimed Tools and Applications* 50(2), 359–380.

Villellas, J., Doak, D. F., García, M. B., and Morris, W. F. (2015). Demographic compensation among populations: what is it, how does it arise and what are its implications? *Ecology Letters* 18(11), 1139–1152.

Vindenes, Y., Engen, S., and Sæther, B.-E. (2011). Integral projection models for finite populations in a stochastic environment. *Ecology* 92(5), 1146–1156.

Vindenes, Y. and Langangen, Ø. (2015). Individual heterogeneity in life histories and eco-evolutionary dynamics. *Ecology Letters* 18(5), 417–432.

Viney, M. and Reece, S. E. (2013). Adaptive noise. *Proceedings of the Royal Society B: Biological Sciences* 280(1767), 20131104.

Visser, M. E. and Both, C. (2005). Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings of the Royal Society of London B: Biological Sciences* 272(1581), 2561–2569.

Visser, M. D., Jongejans, E., Breugel, M. van, Zuidema, P. A., Chen, Y.-Y., Kassim, A. R., and Kroon, H. de (2011). Strict mast fruiting for a tropical dipterocarp tree: a demographic cost-benefit analysis of delayed reproduction and seed predation. *Journal of Ecology* 99(4), 1033–1044.

Visser, M. D., McMahon, S. M., Merow, C., Dixon, P. M., Record, S., and Jongejans, E. (2015). Speeding up ecological and evolutionary computations in R; essentials of high performance computing for biologists. *PLoS Computational Biology* 11(3), e1004140.

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J.-M., Hoegh-Guldberg, O., and Bairlein, F. (2002). Ecological responses to recent climate change. *Nature* 416(6879), 389–395.

White, I. R. and Carlin, J. B. (2010). Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. *Statistics in Medicine* 29(28), 2920–2931.

White, J. W., Nickols, K. J., Malone, D., Carr, M. H., Starr, R. M., Cordoleani, F., Baskett, M. L., Hastings, A., and Botsford, L. W. (2016). Fitting state-space integral projection models to size-structured time series data to estimate unknown parameters. *Ecological Applications* 26(8), 2675–2692.

Wielgus, J., Gonzalez-Suárez, M., Aurioles-Gamboa, D., and Gerber, L. R. (2008). A noninvasive demographic assessment of the sea lions base on stage-specific abundance. *Ecological Applications* 18(5), 1287–1296.

Wilber, M. Q., Langwig, K. E., Kilpatrick, A. M., McCallum, H. I., and Briggs, C. J. (2016). Integral Projection Models for host-parasite systems with an application to amphibian chytrid fungus. *Methods in Ecology and Evolution* 7(10), 1182–1194.

Williams, J. L., Miller, T. E. X., and Ellner, S. P. (2012). Avoiding unintentional eviction from integral projection models. *Ecology* 93(9), 2008–2014.

Wilson, A. J., Réale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., Kruuk, L. E. B., and Nussey, D. H. (2009). An ecologist's guide to the animal model. *Journal of Animal Ecology* 79(1), 13–26.

Wood, S. N. (1994). Obtaining birth and mortality patterns from structured population trajectories. *Ecological Monographs* 64(1), 23–44.

Yampolsky, L. Y., Schaer, T. M. M., and Ebert, D. (2013). Adaptive phenotypic plasticity and local adaptation for temperature tolerance in freshwater zooplankton. *Proceedings of the Royal Society of London B: Biological Sciences* 281(1776), 20132744.

Yang, Y. (2007). Prediction/Estimation with simple linear models: is it really that simple? *Econometric Theory* 23(1), 1–36.

Yoshida, T., Jones, L. E., Ellner, S. P., Fussmann, G. F., and Hairston Jr, N. G. (2003). Rapid evolution drives ecological dynamics in a predator–prey system. *Nature* 424(6946), 303–306.

Yuan, Y. C. (2010). Multiple imputation for missing data: Concepts and new development (Version 9.0). *SAS Institute Inc, Rockville, MD* 49, 1–11.

Zhang, Z., Qian, W., and Zhang, J. (2009). Positive selection for elevated gene expression noise in yeast. *Molecular Systems Biology* 5(1), 299.
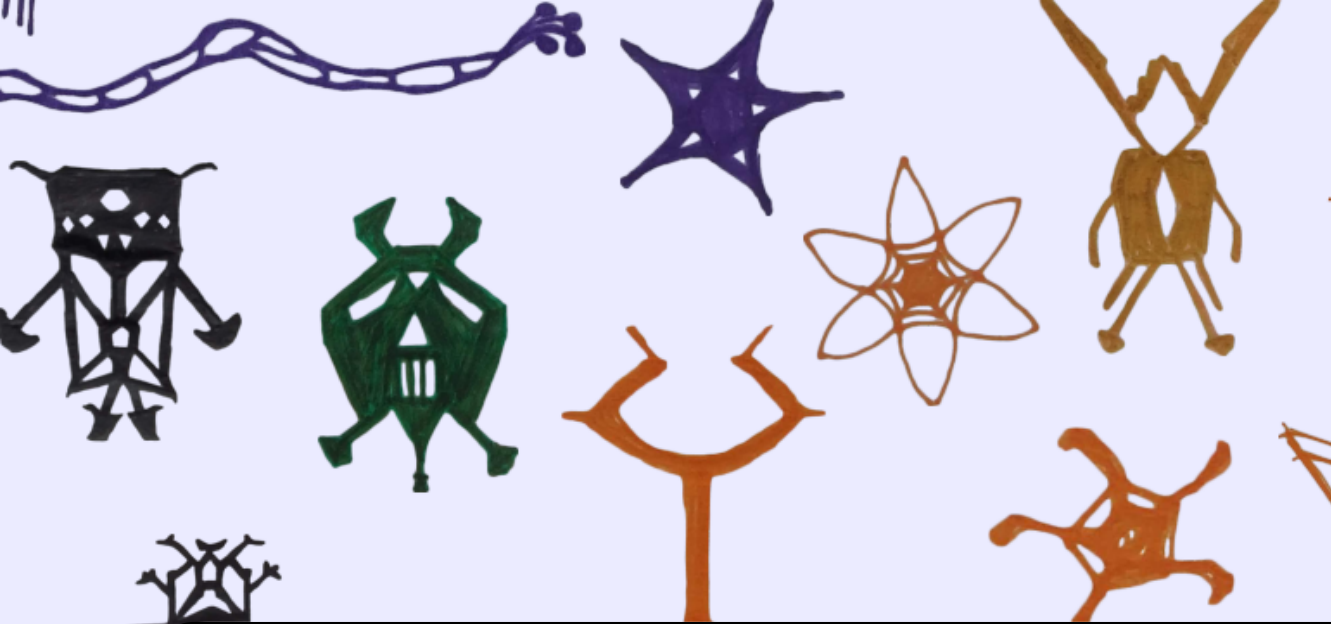
Appendices

A   Supplementary Information for Chapter 2

Supplementary Information can be found at https://tinyurl.com/bruijning2019-appA.

B   Supplementary Information for Chapter 3

Supplementary Information can be found at https://tinyurl.com/bruijning2019-appB.

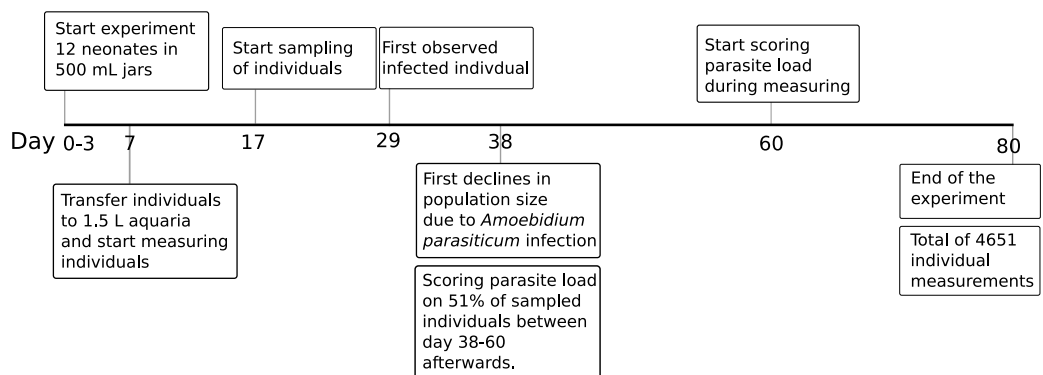## C   Supplementary Information for Chapter 4

C.1   Timeline

Figure C1: Timeline of the experiment.

C.2 Genotype frequencies

For each of the 54 mixed populations (including only between-population comparisons, so exluding competing genotypes from the same population), we identified the sampled individuals using genetic markers. Results are presented in the figures below. To look at temporal changes in genotype frequencies, we fitted a logistic model per population. The intercept was fixed at 0.5, as this was the known proportion at day 0. Per population, three different models were tested: 1) a model in which frequencies remained constant at 0.5, 2) a model with an additional effect of day (allowing for directional changes in genotype frequencies), and 3) a model including an additive effect of day and one of day squared (allowing for changes in the direction of effects, if the sign of both coefficients is different). We chose the model with the lowest AIC and used it to predict genotype frequencies through time. Note that in Fig. C2 we show temporal changes including *only* the observed genotypes. For the vital rate analysis, we imputed missing values (see Appendix C.3). In order to calculate the experienced competition C, the same logistic regressions were used, while also including the imputed values.

C.3 Impute missing observations on parasite load and genotype
Imputation procedure

To handle missing data in one of more covariates, we imputed values by Multiple Imputation by Chained Equations, using R-package MICE (Buuren and Groothuis-Oudshoorn 2011). Imputation is generally recommended when data are missing at random (MAR). MAR implies that missing data depends on the observed data, but, conditional on the observed data, the probability of missing data does not depend on unobserved data (Horton and Kleinman 2007; Morris et al. 2014). This in contrast to data which is missing completely at random (MCAR), in which the probability of missing data depends neither on the observed nor unobserved data (Horton and Kleinman 2007; Morris et al. 2014).

Fifteen percent of the observations on parasite load were missing, and seventeen percent of the observations on genotype. For parasite load, missing data were mainly from days 30-60. Although we were able to score infection on sampled individuals after the experiment, we could obviously not do so for individuals that were not sampled (which were dead individuals, or whenever populations were too small to sample from). The same is true for genotype identities; these could not be determined whenever individuals were not sampled (including all individuals from the first 16 days, as sampling started on day 17). We thus expect data to be missing at random. In this case, performing complete-case-analysis (i.e. removing all rows with one or more missing values, which would be 30% in our case) can lead to a loss of statistical power and biased estimates (White and Carlin 2010).

In short, the imputation procedure was as follows. Two imputation models were used, fitting parasite load and clone as a function of aquarium, day, squared day, temperature, and body size. Using MICE, missing values were subsequently imputed, based on predictive mean matching (Morris et al. 2014), which takes the observed value that is closest to the prediction (to ensure that imputed values are within the range of observations). This way, we created 10 imputed data sets. Note that imputation takes into account uncertainty in the expected imputed values: whenever there is a high uncertainty, this will be reflected across the 10 imputed datasets and in subsequent estimated coefficients and variance components. See e.g. Azur et al. (2011) for more information on MICE. For each vital rate, the selected model (based on AIC) was fitted on each of the imputed datasets. Subsequently, coefficients and variances were averaged, the latter taking into account both the between-imputation variance, and the variance associated with each parameter (Yuan 2010). We compared models fitted on the imputed dataset with models fitted on the original dataset in which we deleted all rows with missing values (Fig. C3).

Figure C2: Identified genotypes through time. Each graph shows a different population. Green line shows predictions of the best model based on a comparison of three different models: 1) no temporal trend, 2) an effect of day, and 3) a squared effect of day. The intercept was fixed at 0.5. Different colours denote different populations (darkblue: N1, lightblue: N2, red: B1, orange: B2.)

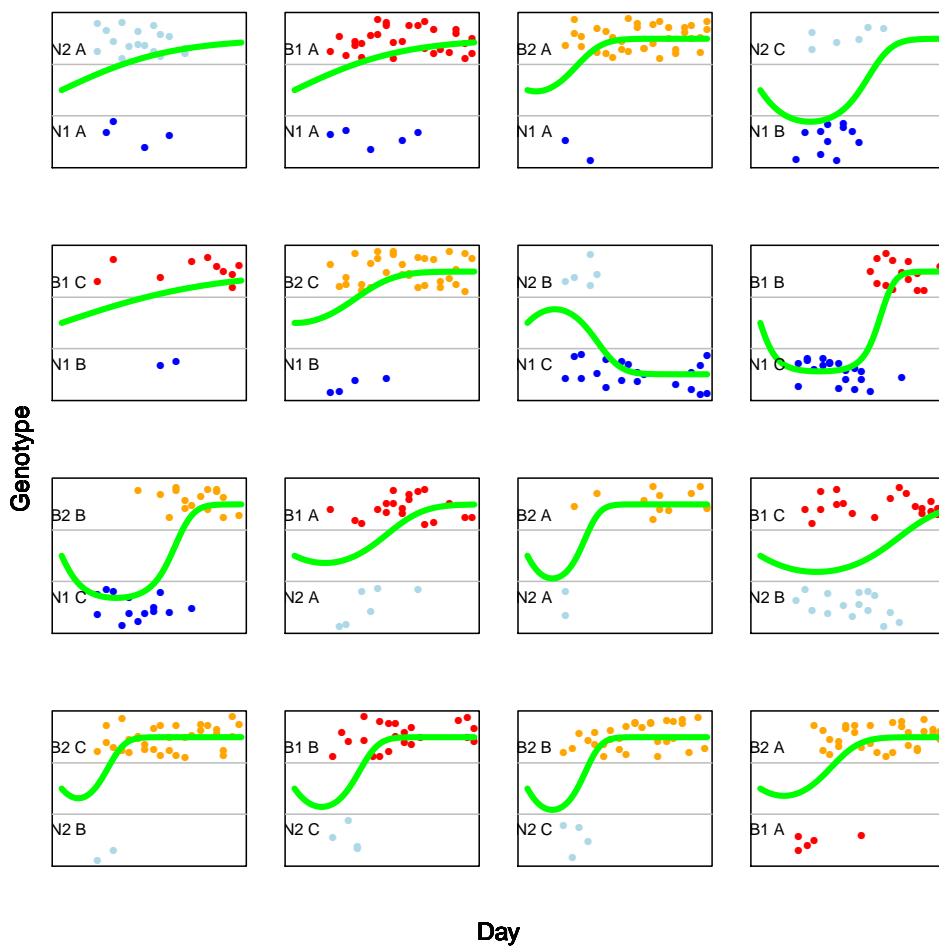**Figure C2:** *(Continued)* Identified genotypes through time. Each graph shows a different population. Green line shows predictions of the best model based on a comparison of three different models: 1) no temporal trend, 2) an effect of day, and 3) a squared effect of day. The intercept was fixed at 0.5. Different colours denote different populations (darkblue: N1, lightblue: N2, red: B1, orange: B2.)

Figure C2: *(Continued)* Identified genotypes through time. Each graph shows a different population. Green line shows predictions of the best model based on a comparison of three different models: 1) no temporal trend, 2) an effect of day, and 3) a squared effect of day. The intercept was fixed at 0.5. Different colours denote different populations (darkblue: N1, lightblue: N2, red: B1, orange: B2.)
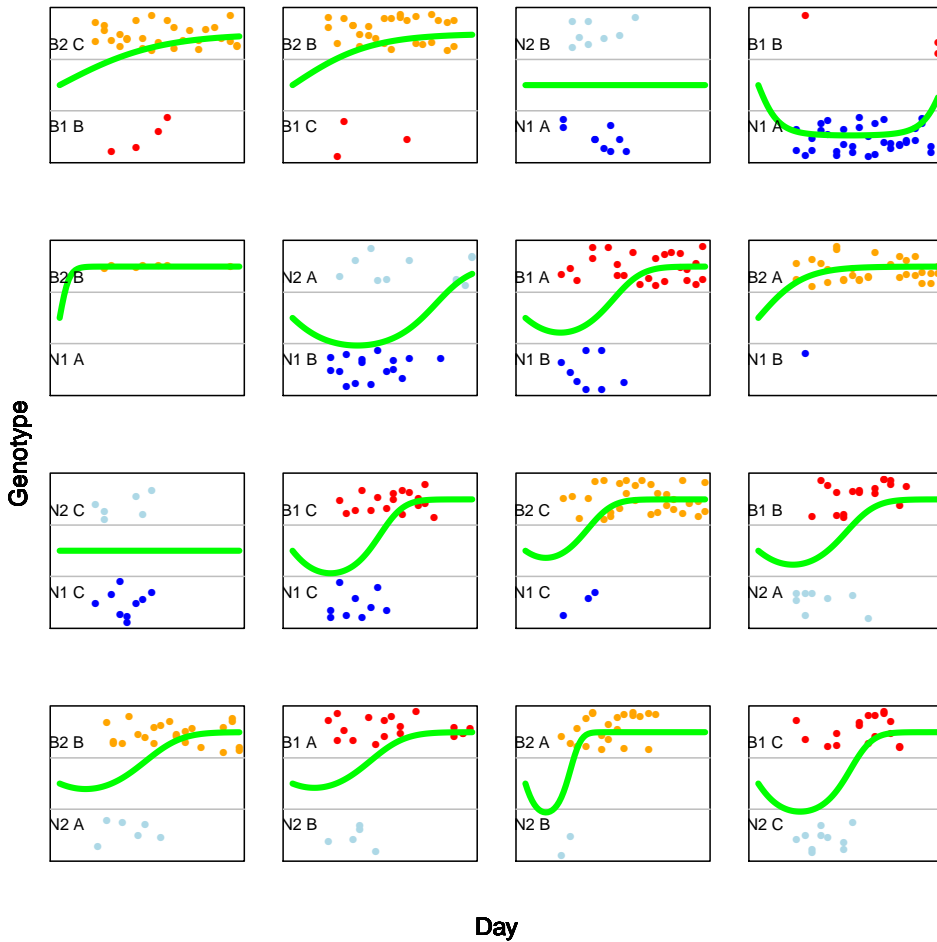
Figure C2: *(Continued)* Identified genotypes through time. Each graph shows a different population. Green line shows predictions of the best model based on a comparison of three different models: 1) no temporal trend, 2) an effect of day, and 3) a squared effect of day. The intercept was fixed at 0.5. Different colours denote different populations (darkblue: N1, lightblue: N2, red: B1, orange: B2.)
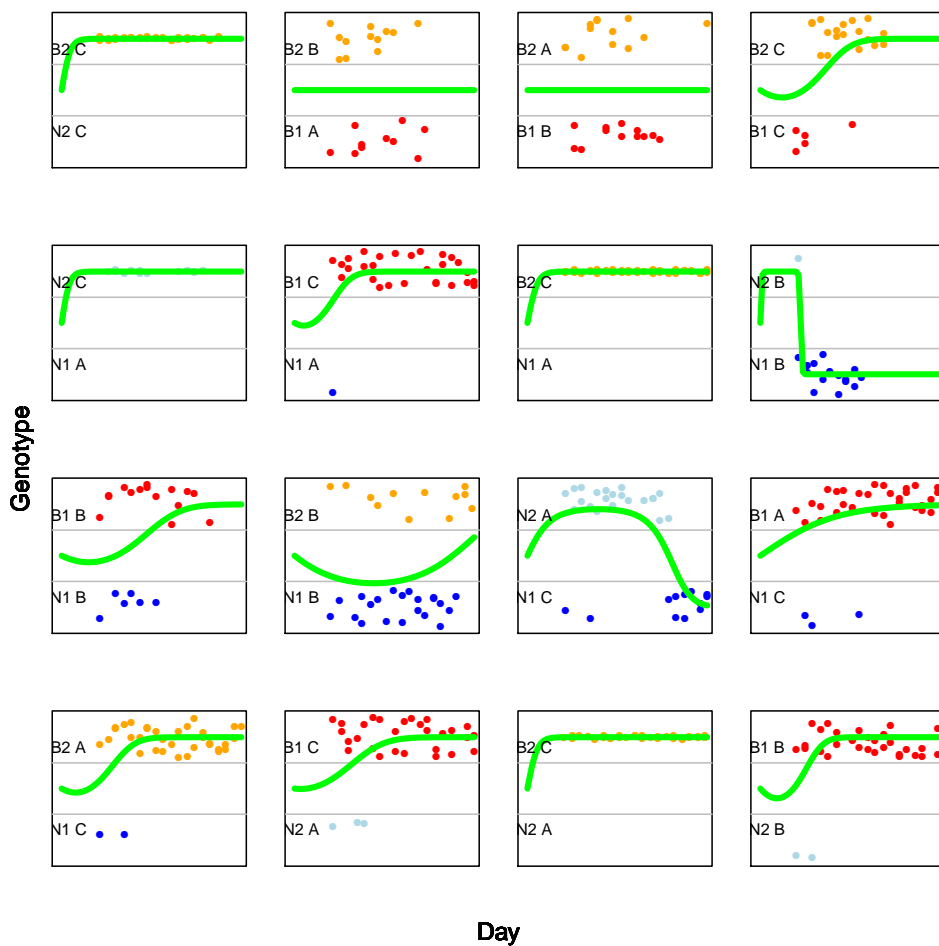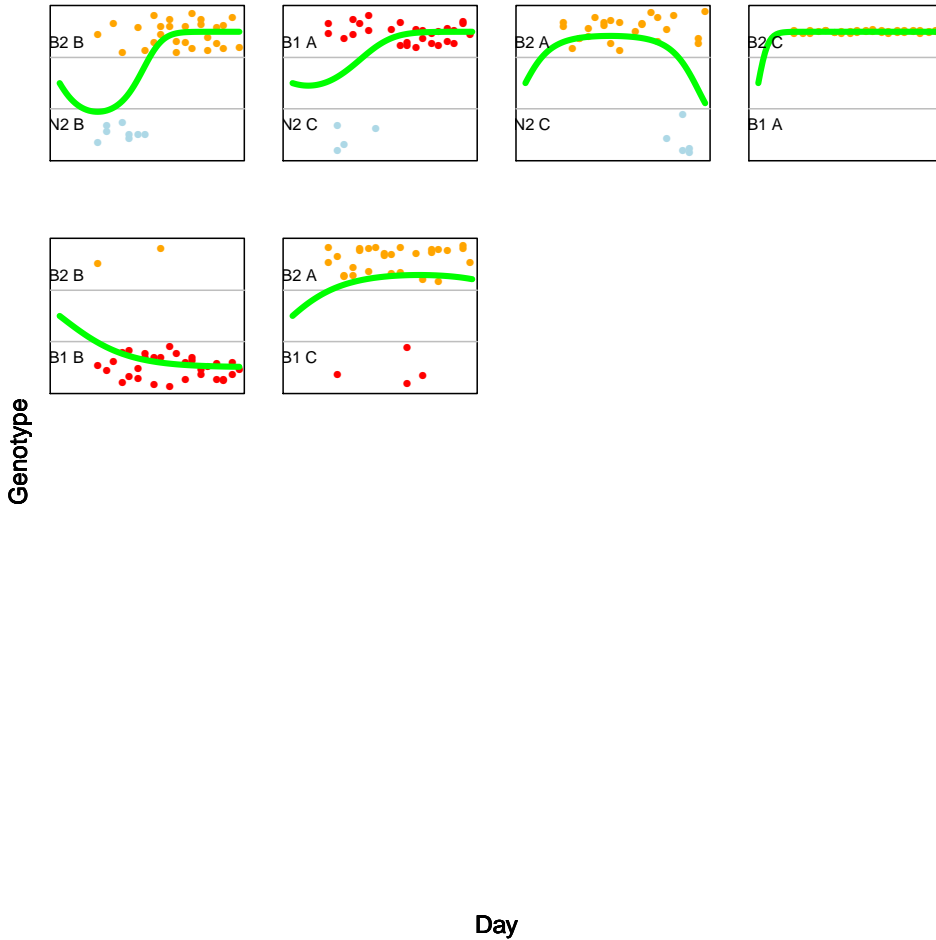
Vital rates models using complete-case analysis
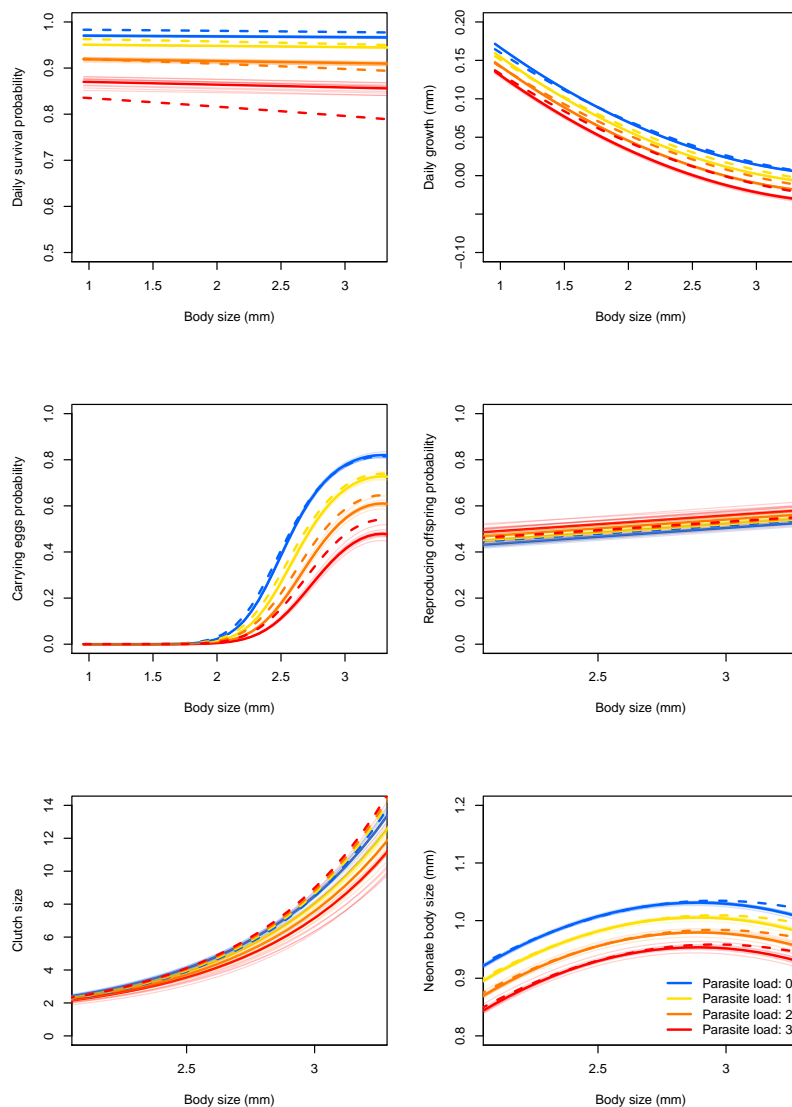
Figure C3: Comparison of vital rates predictions as a function of body size for imputed datasets, and the data in which we deleted all rows that contained missing values in one or more covariates (resulting in a removal of 30% of the data). Different colours indicate different parasite load classes. Thick solid lines show predictions according to the weighed model across imputed datasets, which was used for all subsequent analysis. Thick dotted lines show predictions based on complete-case analysis. Thin lines show predictions for each of the imputed dataset, to indicate the stability.

C.4 Model selection for vital rates

As explained in the manuscript, we tested different models and continued with the model with the lowest AIC. In the tables below, we show ΔAIC values of the ten best models, for each vital rate. As the best model was chosen based on the performance on each of the ten imputed datasets, we additionally provide ΔAIC values per imputed dataset.

Model	ΔAIC	Imp 1	Imp 2	Imp 3	Imp 4	Imp 5	Imp 6	Imp 7	Imp 8	Imp 9	Imp 10
$z+C+T+P+L+L*P+L*T+L*C+P*T$	0	0	0	0	0	0.232	0	0.071	0	0	0
$z+C+T+P+L+z^2+L*P+L*T+L*C+P*T$	0.768	0.831	0.823	0.934	0.666	1.469	0.303	0.944	0.942	0.729	0.342
$z+C+T+P+L+L*P+L*T+L*z+L*C+P*T$	1.301	0.364	0.09	0.459	2.515	3.91	0.195	0	4.086	1.646	0.046
$z+C+T+P+L+L*P+L*T+L*C$	1.47	2.241	2.054	1.788	0.665	2.395	2.386	1.257	0.138	0.266	1.812
$z+C+T+P+L+z^2+L*P+L*T+L*z+L*C+P*T$	2.208	1.134	1.146	1.658	3.278	5.368	0.729	0.749	5.232	2.497	0.593
$z+C+T+P+L+z^2+L*P+L*T+L*C$	2.386	3.221	3.036	2.897	1.46	3.762	2.863	2.272	1.191	1.091	2.369
$z+T+P+L+L*P+L*T+L*z+L*C$	2.781	2.739	2.058	2.466	3.179	6.028	2.584	1.333	4.273	1.871	1.581
$z+C+T+P+L+L*P+L*T+P*T$	3.14	6.425	0.991	3.282	3.088	0	3.669	1.257	0.429	2.684	9.876
$z+C+T+P+L+z^2+L*P+L*T+P*T$	3.265	7.016	1.938	3.854	3.32	1.472	3.294	1.009	0.542	3.536	6.968
$z+C+T+P+L+z^2+L*P+L*T+L*z+L*C$	3.823	3.664	3.25	3.815	4.063	7.592	3.281	2.234	5.516	2.805	2.313

Table C1: Model selection for models describing survival probability

Model	ΔAIC	Imp 1	Imp 2	Imp 3	Imp 4	Imp 5	Imp 6	Imp 7	Imp 8	Imp 9	Imp 10
$z+C+T+N+P+L+z^2+L*P+L*T+L*n$	0	0.318	0.594	0	0.72	2.879	0	0.168	0	0	0
$z+C+T+N+P+L+z^2+L*T+L*n$	1.205	0	3.49	2.797	0	4.71	1.791	0	2.111	1.806	0.029
$z+C+T+N+P+L+z^2+L*P+L*T+L*n+n+L*C$	1.533	3.447	0	1.854	3.915	0	2.689	0.788	0.534	4.227	2.553
$z+C+T+N+P+L+z^2+L*P+L*T+L*n+P*T$	1.903	2.296	2.581	1.889	2.718	4.87	1.795	1.779	1.916	1.951	1.92
$z+C+T+N+P+L+z^2+L*P+L*T+L*n+L*C$	2.282	2.903	2.126	3.881	2.896	0.953	4.066	0.352	2.056	5.887	2.383
$z+C+T+N+P+L+z^2+L*P+L*T+L*n+P*T$	3.12	1.977	5.49	4.746	1.999	6.677	3.593	1.589	4.042	3.795	1.97
$z+C+T+N+P+L+z^2+L*P+L*T+L*n+n+L*z$	3.401	2.292	1.999	3.901	5.455	6.202	0.989	4.213	3.995	4.585	5.062
$z+C+T+N+P+L+z^2+L*P+L*T+L*n+n+L*C+P*T$	3.447	5.438	1.98	3.744	5.913	1.987	4.517	2.404	2.482	6.182	4.499
$z+C+T+N+P+L+z^2+L*T+L*n+n+L*C+P*T$	4.21	4.894	4.121	5.821	4.895	2.898	5.92	1.98	4.021	7.879	4.354
$z+C+T+N+P+L+z^2+L*T+L*n+L*z$	4.438	1.788	4.538	6.579	4.597	7.696	2.635	3.947	5.927	6.449	4.907

Table C2: Model selection for models describing somatic growth

Table C3: Model selection for models describing carrying eggs probability

Model	ΔAIC	Imp 1	Imp 2	Imp 3	Imp 4	Imp 5	Imp 6	Imp 7	Imp 8	Imp 9	Imp 10
$z + T + N + P + L + z^2 + L*T + L*T + L*z + P*T$	**0**	0	0	0	0	0	0	4.823	0	0	1.546
$z + C + T + N + P + L + z^2 + L*T + L*T + L*z + P*T$	**1.237**	0.488	0.632	1.644	1.668	1.977	0.916	5.465	1.589	1.088	3.277
$z + T + N + P + L + z^2 + L*P + L*T + L*z + n + L*z + P*T$	**1.891**	3.339	0.565	1.175	4.685	4.076	2.138	2.954	3.452	2.893	0
$z + C + T + N + P + L + z^2 + L*P + L*T + L*z + n + L*z + C + P*T$	**2.227**	2.196	3.152	2.263	3.101	3.637	0.011	2.573	5.836	2.393	3.472
$z + C + T + N + P + L + z^2 + L*T + L*n + L*z + P*T$	**3.237**	3.957	1.372	2.921	6.401	6.07	3.171	3.833	5.095	4.093	1.826
$z + T + N + P + L + z^2 + L*T + L*n + L*z$	**3.594**	4.282	3.671	2.031	0.779	4.163	4.892	7.835	5.277	2.79	6.591
$z + C + T + N + P + L + z^2 + L*P + L*T + L*z + L*n + L*z + L*C + P*T$	**4.058**	5.628	3.918	3.392	7.558	7.513	2.204	0	9.223	5.08	2.433
$z + C + T + N + P + L + z^2 + L*T + L*n + L*z$	**4.821**	4.746	4.269	3.684	2.458	6.15	5.718	8.509	6.831	3.906	8.303
$z + C + T + N + P + L + z^2 + L*T + L*n + L*z + L*C$	**5.668**	6.256	6.637	4.225	3.928	7.653	4.71	5.43	10.987	5.005	8.214
$z + T + N + P + L + z^2 + L*T + L*n + P*T$	**6.128**	3.007	5.868	4.587	9.303	7.552	4.657	11.7	8.539	6.041	6.398

Table C4: Model selection for models describing reproduction probability

Model	ΔAIC	Imp 1	Imp 2	Imp 3	Imp 4	Imp 5	Imp 6	Imp 7	Imp 8	Imp 9	Imp 10
$z + T + N + P + L + L*P + L*T$	**0**	0.782	1.244	0.315	1.669	0.498	1.385	0.714	0	1.433	1.76
$z + T + N + P + L + L*P + L*T + P*T$	**0.743**	0.163	1.372	0.729	2.264	1.318	2.599	0	1.846	0.908	2.224
$z + T + N + P + L + L*P + L*T + L*n$	**0.988**	0.899	1.998	1.969	0	0	1.134	1.926	3.202	1.346	1.213
$z + C + T + N + P + L + L*P + L*T$	**1.498**	2.071	1.658	1.923	3.174	2.275	1.558	1.876	1.9	1.349	1.59
$z + T + N + P + L + z^2 + L*P + L*T$	**1.587**	2.245	2.482	0.393	3.397	2.263	3.895	2.148	1.713	2.249	3.93
$z + T + N + P + L + z^2 + L*P + L*T + L*n + P*T$	**1.59**	0	2.815	1.395	0.245	0.715	2.829	1.004	4.966	2.686	3.3
$z + T + N + P + L + z^2 + L*P + L*T + P*T$	**2.206**	1.416	3.244	1.719	3.876	2.992	2.575	1.203	3.526	2.805	3.008
$z + C + T + N + P + L + L*P + L*T + P*T$	**2.299**	1.577	2.965	2.268	3.833	3.141	4.062	1.316	3.753	2.182	3.79
$z + T + N + P + L + z^2 + L*P + L*T + L*n$	**2.525**	2.307	3.372	2.452	1.719	1.735	3.328	3.328	4.871	2.481	3.625
$z + C + T + N + P + L + L*P + L*T + L*n$	**2.616**	2.355			1.728	1.889	3.723	3.382	5.128		

Table C5: Model selection for models describing clutch size

Model	ΔAIC	Imp 1	Imp 2	Imp 3	Imp 4	Imp 5	Imp 6	Imp 7	Imp 8	Imp 9	Imp 10
$z + C + T + N + P + L + L*P + L*n + L*z + L*C$	**0**	0.632	2.398	2.107	0	11.727	0.27	1.459	0.371	0.409	1.16
$z + C + T + N + P + L + L*P + L*n + L*z + L*C + P*T$	**0.308**	0	2.917	1.785	0.66	10.246	2.225	1.579	2.145	0	2.127
$z + C + T + N + P + L + L*P + L*n + L*z + L*C$	**1.263**	2.554	4.272	3.756	1.626	12.388	1.856	3.128	0	0.995	1.9
$z + C + T + N + P + L + z^2 + L*P + L*n + L*z + L*C + P*T$	**1.687**	1.986	4.678	3.59	2.419	11.306	5.285	2.755	1.893	1.624	3.051
$z + C + T + N + P + L + L*P + L*T + L*z + L*C$	**2.162**	1.702	5.464	0.5	4.449	6.395	4.992	4.319	3.599	3.861	6.278
$z + C + T + N + P + L + z^2 + L*P + L*T + L*z + L*C + P*T$	**2.243**	0.876	5.717	2.117	4.855	3.911	7.252	4.37	5.259	4.06	7.103
$z + C + T + N + P + L + z^2 + L*P + L*T + L*n + L*z + L*C$	**3.371**	3.58	7.307	1.787	6.08	6.703	6.868	6.012	4.868	4.91	7.006
$z + C + T + N + P + L + z^2 + L*P + L*n + L*z + L*C + P*T$	**3.585**	2.844	7.423	17.129	6.632	4.711	18.662	25.127	11.857	5.338	8.028
$z + C + T + N + P + L + z^2 + L*P + L*n + L*z$	**9.994**	17.868	0	16.723	8.601	6.676	18.594	26.574	13.652	12.682	0
$z + C + T + N + P + L + z^2 + L*P + L*n + L*z + P*T$	**10.322**	17.304	0.446		9.134	5.296				13.222	0.937

Model	ΔAIC	Imp 1	Imp 2	Imp 3	Imp 4	Imp 5	Imp 6	Imp 7	Imp 8	Imp 9	Imp 10
$z + C + T + N + P + L + z^2 + L*T + L*C + P*T$	**0**	0	0	2.706	0	0	1.145	1.364	0	1.893	0
$z + C + T + N + P + L + z^2 + L*T + L*C$	**0.404**	0.148	0.397	2.484	0.965	1.146	0.874	2.158	0.142	2.404	0.428
$z + T + N + P + L + z^2 + L*T + P*T$	**0.912**	1.92	4.169	0.646	1.254	6.272	0.208	0	0.935	0	0.829
$z + C + T + N + P + L + z^2 + L*T + P*T$	**1.19**	2.924	2.982	0.01	1.366	6.965	0	0.288	2.457	0.816	1.2
$z + T + N + P + L + z^2 + L*T$	**1.591**	2.144	4.964	0.64	2.694	7.825	0.308	0.896	1.303	0.649	1.594
$z + C + T + N + P + L + z^2 + L*T$	**1.86**	3.136	3.783	0	2.754	8.482	0.08	1.135	2.835	1.487	2.014
$z + C + T + N + P + L + z^2 + L*P + L*T + L*C + P*T$	**3.542**	3.277	3.455	6.97	3.474	2.132	4.971	3.965	3.831	6.451	4
$z + C + T + N + P + L + z^2 + L*P + L*T + L*C$	**3.792**	3.307	3.996	6.944	4.419	2.566	4.443	4.227	4.031	6.711	4.387
$z + C + T + N + P + L + z^2 + L*T + L*n + L*C + P*T$	**3.865**	2.955	4.563	6.552	4.974	2.164	6.447	5.418	3.217	5.789	3.675
$z + C + T + N + P + L + z^2 + L*C + P*T$	**3.975**	5.003	4.325	8.105	4.636	5.739	3.883	6.204	1.409	5.515	2.039

Table C6: Model selection for models describing neonate body size
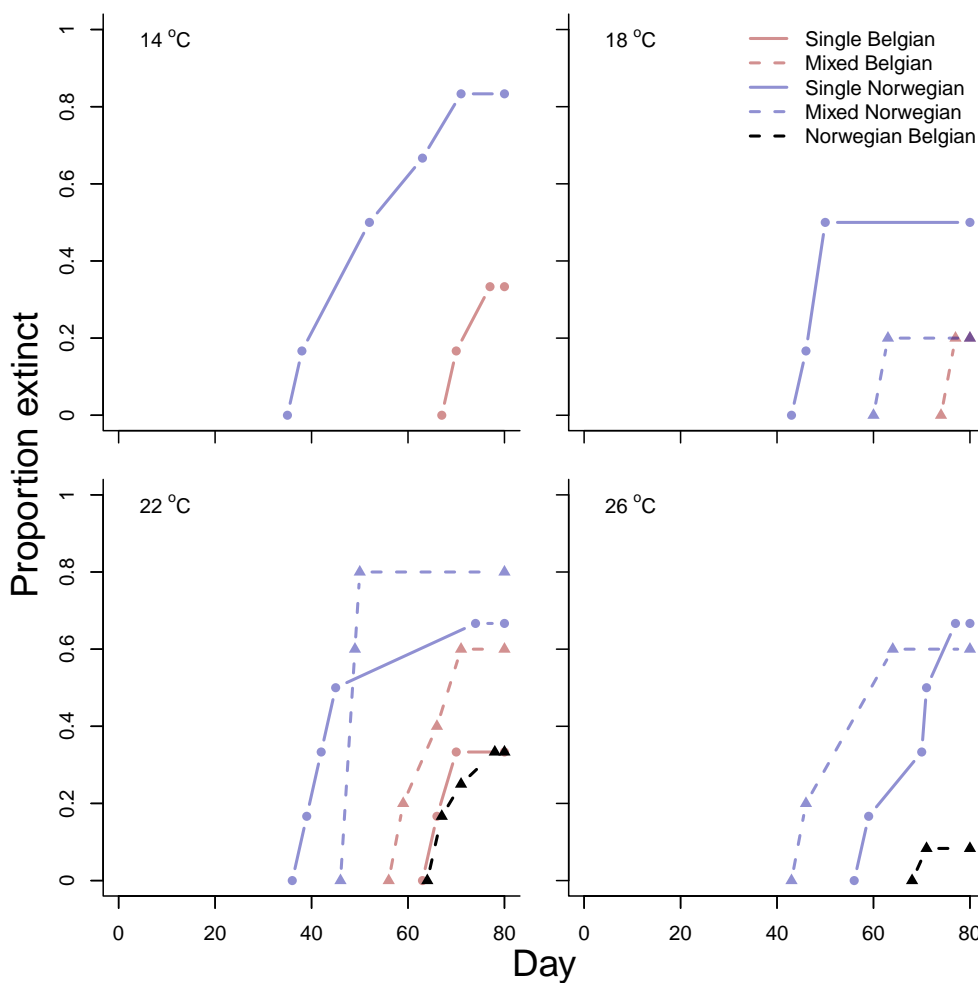
C.5 Extinctions

Figure C4: Proportion of populations that is extinct over time. Each graph shows a different temperature. Different lines show different clonal treatments; blue: Norwegian clones, red: Belgian clones, black: Combinations of Norwegian and Belgian clones. Solid lines show single clone populations, dotted lines show mixed populations. Note that at 14 °C there were no mixed populations. At 18, 22 and 26 °C, missing lines indicate that none of the populations from this specific treatment went extinct.

C.6 Parasite dynamics

The fitted matrix **A** was used to project parasite dynamics for each of the four experimental temperatures and for each location, keeping total population sizes equal. At day 0, we simulated the start of the infection, starting with 3.4% of the population being in parasite class 1 (which is the estimated proportion according to the fitted model, see the manuscript for more details). After 60 days, we compared the observed parasite load structure with the predictions (Fig. C5).
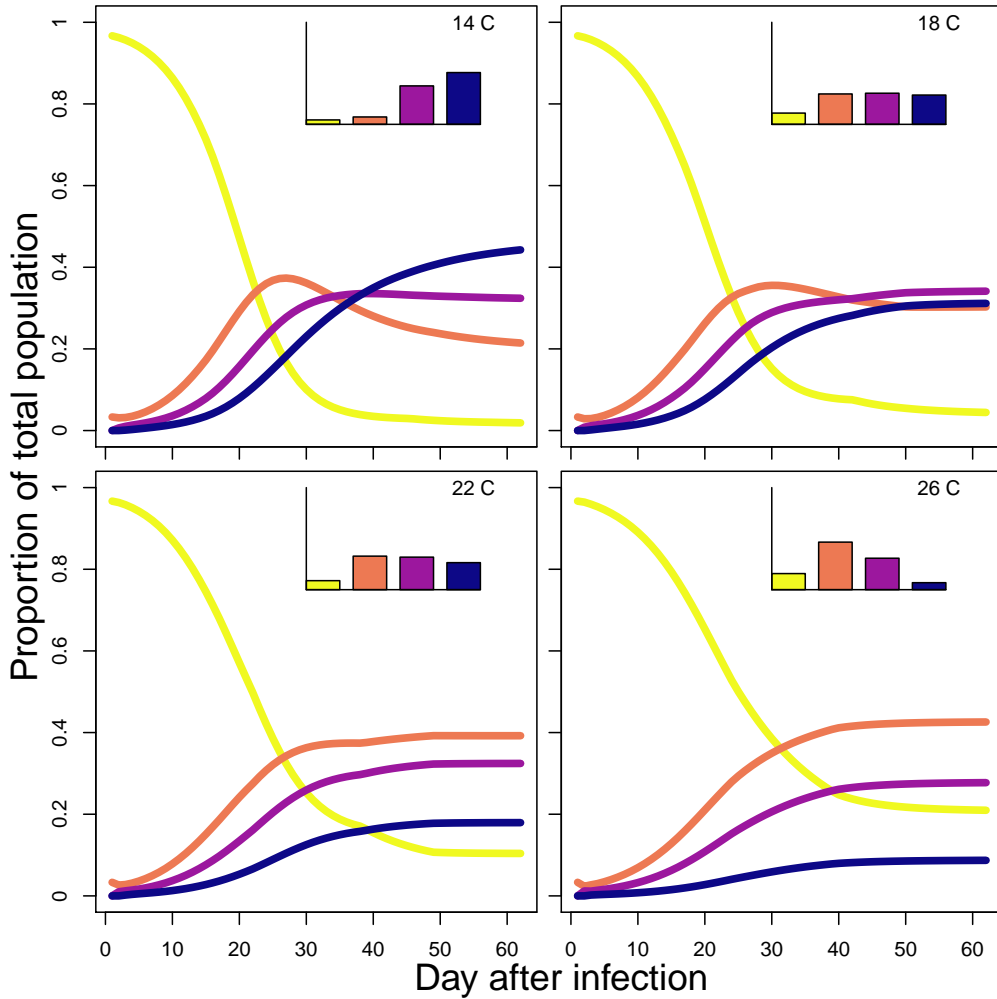
Figure C5: Simulated changes in the proportion of individuals in each parasite class for different temperatures, after the start of infection at day 0. Proportions are averaged over the four locations. Different colours indicate different parasite load classes (yellow: 0, red: 1, purple: 2, blue: 3). Bar plots show the observed frequencies, averaging over the last 15 days of the experiment.

C.7 Temperature effects on host-parasite dynamics

We quantified the contribution of multiple processes to the different host-parasite dynamics at 14 °C compared to 26 °C. In Figure 4.5 of the manuscript, results were shown for location B2. Below we show the results for the other locations, which were qualitatively very similar.
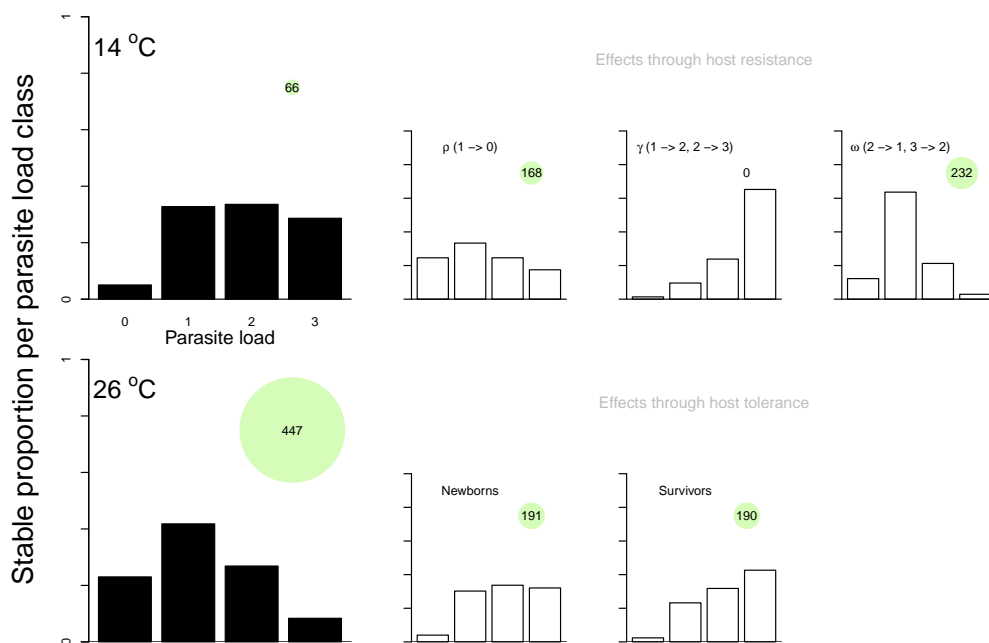

Figure C6: Host-parasite dynamics at 14 °C and 26 °C, based on a host-parasite integral projection model. Barplots on the left show the projected proportion of individuals in each parasite class after 200 time steps. Green circles show stabilized densities. Starting with the host-parasite dynamics at 14 °C, white barplots show proportions and densities when one function is replaced by the relevant function at 26 °C. Upper row shows results when changing functions associated with the parasite dynamics: recovery probability, probability of moving to the next parasite class, and probability of moving to the previous parasite class (see Eq. 4.5). Bottom row shows the interactive effects of temperature and infection effects on the dynamics of newborns (reproduction and neonate body size) and survivors (survival and growth). Results shown for location N1.
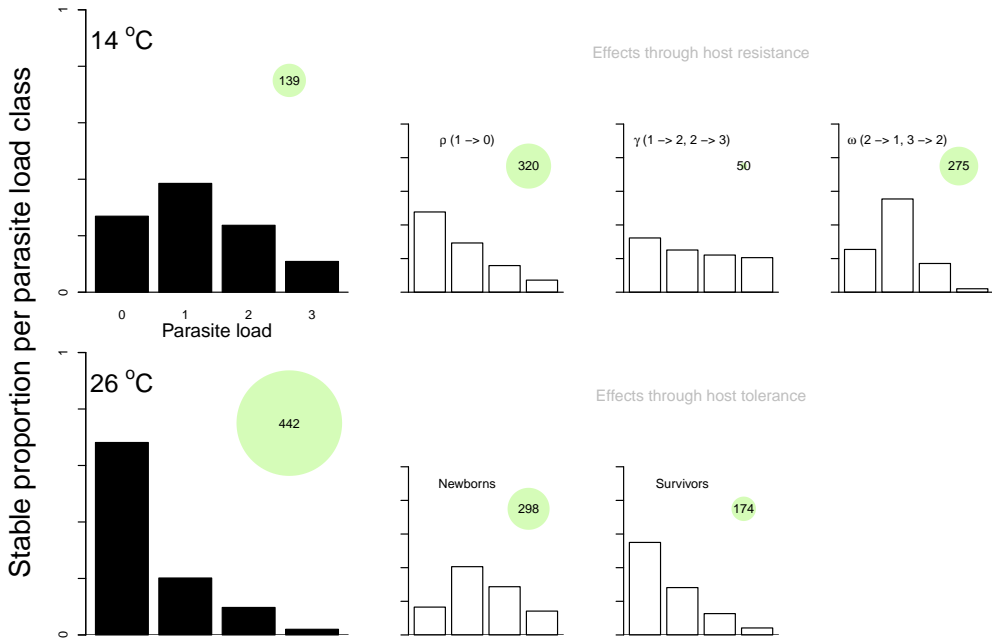
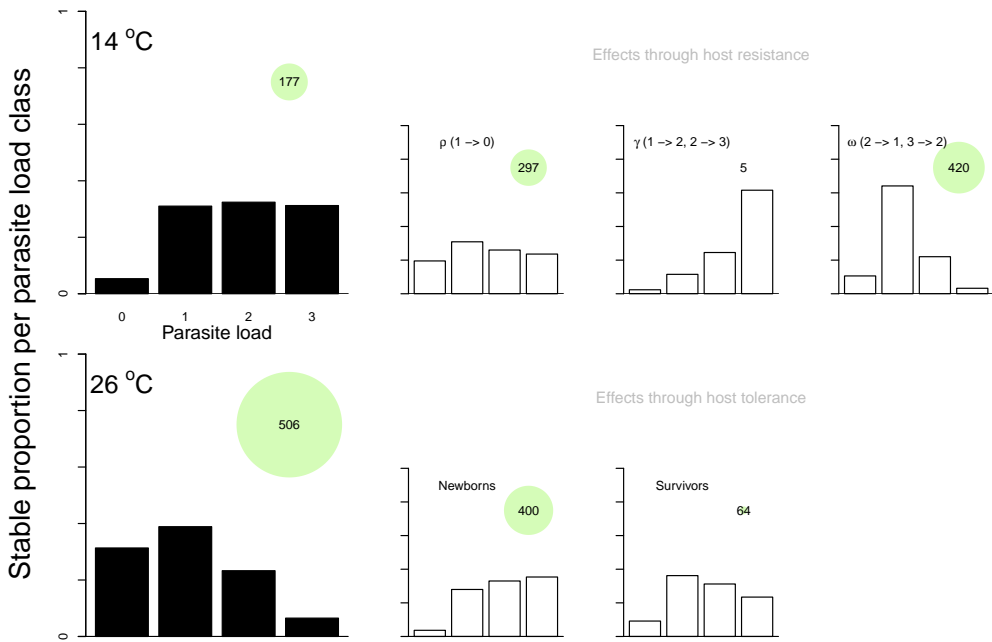**Figure C6:** *(Continued)* Results shown for location N2. See previous page for more information.

Figure C6: *(Continued)* Results shown for location B1. See previous page for more information.

C.8 Invasion population growth rates

In Figure 4.2 of the manuscript, we projected equilibrium densities for combinations of temperature and parasite load. Below, we show results when considering invasion growth rate instead of equilibruim density, which result in the same patterns. The invasion growth rate is the growth rate when density is set at 1 individual.
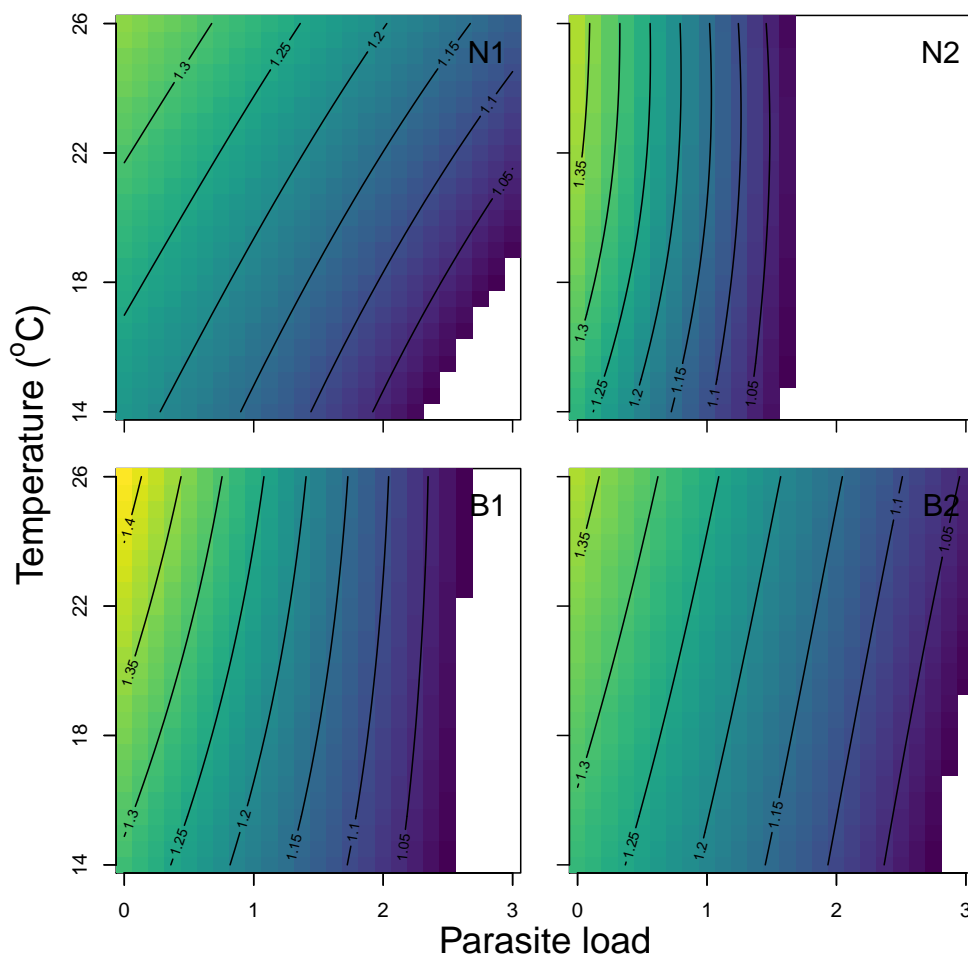

Figure C7: Projected invasion growth rates for different combinations of temperature and parasite load, based on an integration of all vital rates using Integral Projection Models. Different panels show results for different locations of origin. Values smaller than 1 are not shown.

D   Supplementary Information for Chapter 5

D.1   Plant sizes

Plant size was included as a covariate in the vital rate models (either directly, or in combination with population size to obtain densities). We used the number of leaves as a proxy for plant size. To be able to implement daily plant sizes, we fitted smooth functions, as data was collected on a three or four days basis. Generalized additive models were fitted, for each plant (containing a population aphids) separately. In Fig. D1 the daily model predictions are shown. Predictions below 1 were set to 1, in order to include a minimum plant size of 1 leaf.
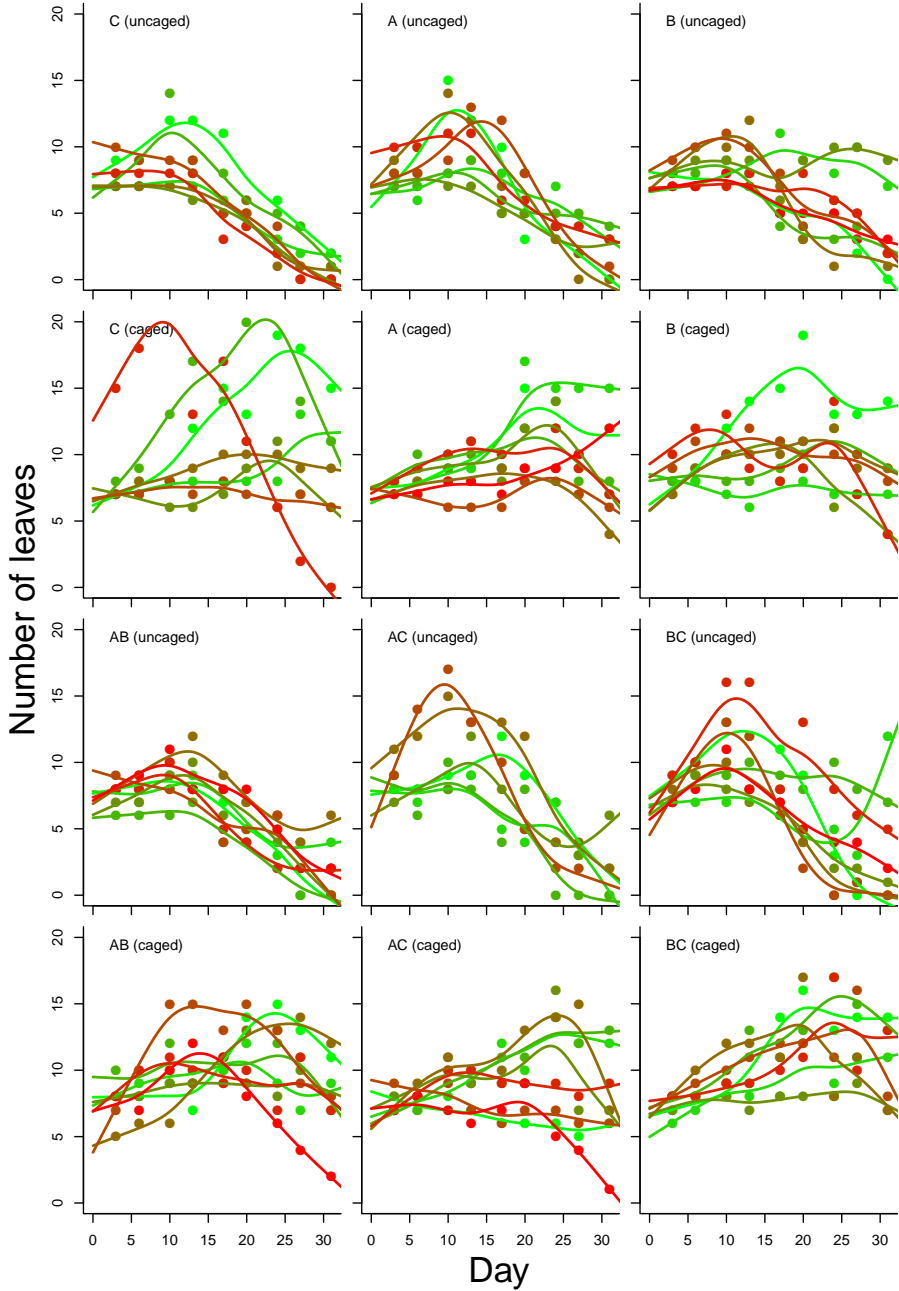
Figure D1: Changes in the number of leaves over time for each plant. Each graph shows replicates for a different aphid treatment. Different colours indicate different plants (replicates). Dots show the data and lines show model predictions based on generalized additive models, shown per experimental treatment.
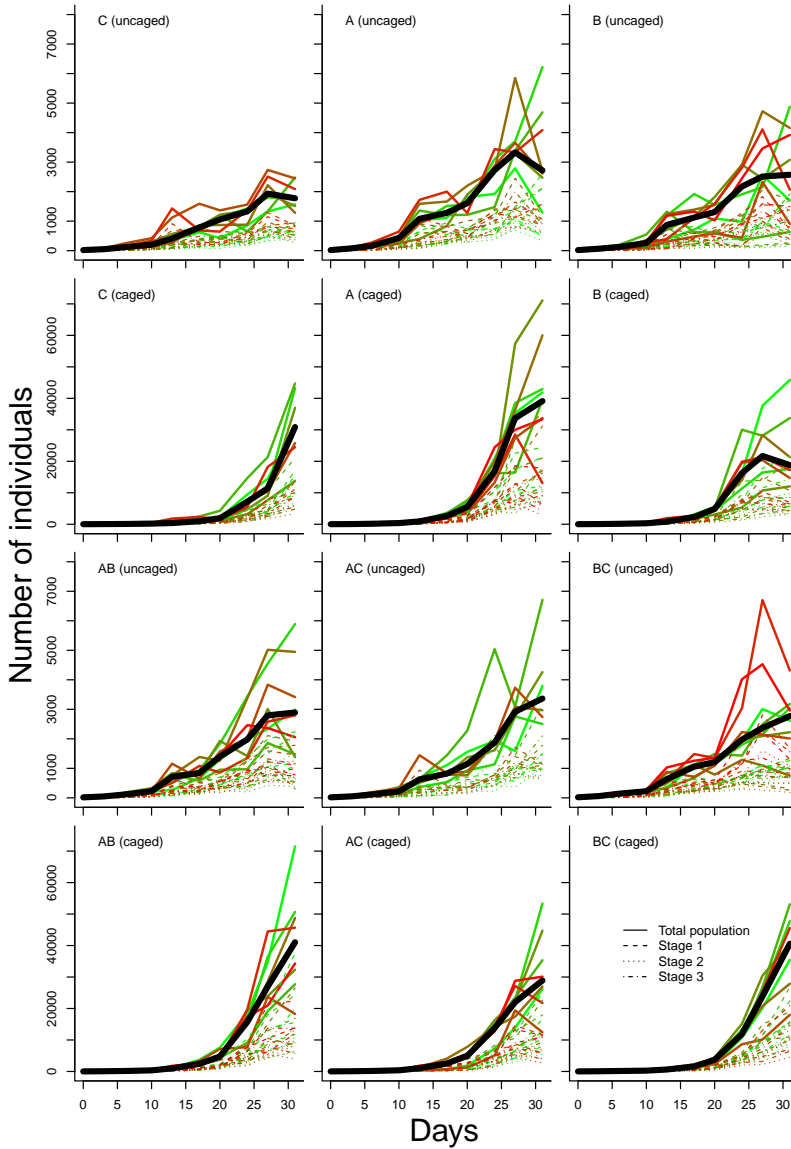
D.2 Population sizes

Figure D2: Changes in the individuals over time for each population, shown per experimental treatment. Solid lines show the total population sizes, different dotted lines show the numbers per developmental stage. Thick black line shows the median total population size per treatment. Different colours indicate different replicates, and correspond to the colours shown in Fig. D1. Note the different y-axis for the caged (row 2,4) and uncaged (row 1,3) treatment.

D.3 Fitting the seven basic models

In order to evaluate which covariate resulted in the best model fit, we started by testing seven different 'basic' models. These models were fitted for each of the six aphid treatments ('A', 'B', 'C', 'AB', 'AC', 'BC') separately to assess the performance of different modelling structures for each aphid treatment independently. These models were: a null model (including only a survival, growth and reproduction intercept), and six models including intercepts, together with additive effects of developmental stage enabling different 1) survival and 2) growth for different stages x, 3) caging (C), 4) population size (N), 5) plant size (P) and 6) population size divided by plant size to obtain densities (number of individuals leaf^{-1}; D). We refer to these models as Models 0-6 (see Table 5.1 of the manuscript for an overview). Caging was included as a factor with two levels (caged and uncaged). For the model including a caging effect, parameters for the uncaged treatments were estimated including data from day 14 and onwards only, since cages were removed at day 13. Observations until day 13 were used to estimate parameters for the caged treatments. Population size was included as a continuous variable, and we transformed values by Z-score normalization (using the mean and standard deviation calculated across the full dataset). We used the number of leaves as a proxy for plant size, included as a continuous covariate after Z-score normalization. Finally, densities were obtained by dividing population size by plant size, to obtain number of individuals per leaf, and values were normalized.

To assess the predictive ability of a fitted model in order to compare different models, we used K-fold cross-validation. Each basic model was fitted eight times, each time removing one replicate per caging treatment for validation. This implied that two populations were removed per fit (one from both caging treatments), as parameters were fitted per aphid treatment and each aphid treatment was exposed to both caging treatments. Model predictions were calculated for the validation data, and the cross-validation score was calculated as the mean squared error between the predictions and observations. The best model was considered the model resulting in the lowest overall mean error (i.e. the highest out-of-sample predictive ability), across aphid treatments.

As can be seen in Table 5.1, the model including density resulted in the lowest mean error overall. We then constructed a full model including density, fitted across clonal treatments on the complete data set. As explained in the manuscript, we additionally included an effect of caging. For the seven basic models, the likelihood was based on the goodness-of-fit between estimated and observed population-level data, and not on the goodness-of-fit between the predicted and observed individual-level life table data. This was done to keep the models as flexible as possible. For this full model, we estimated parameter values that optimized the likelihoods for both the population-level (i.e. the counts per stage) and the individual-level life table data (see the manuscript for details on the likelihoods).
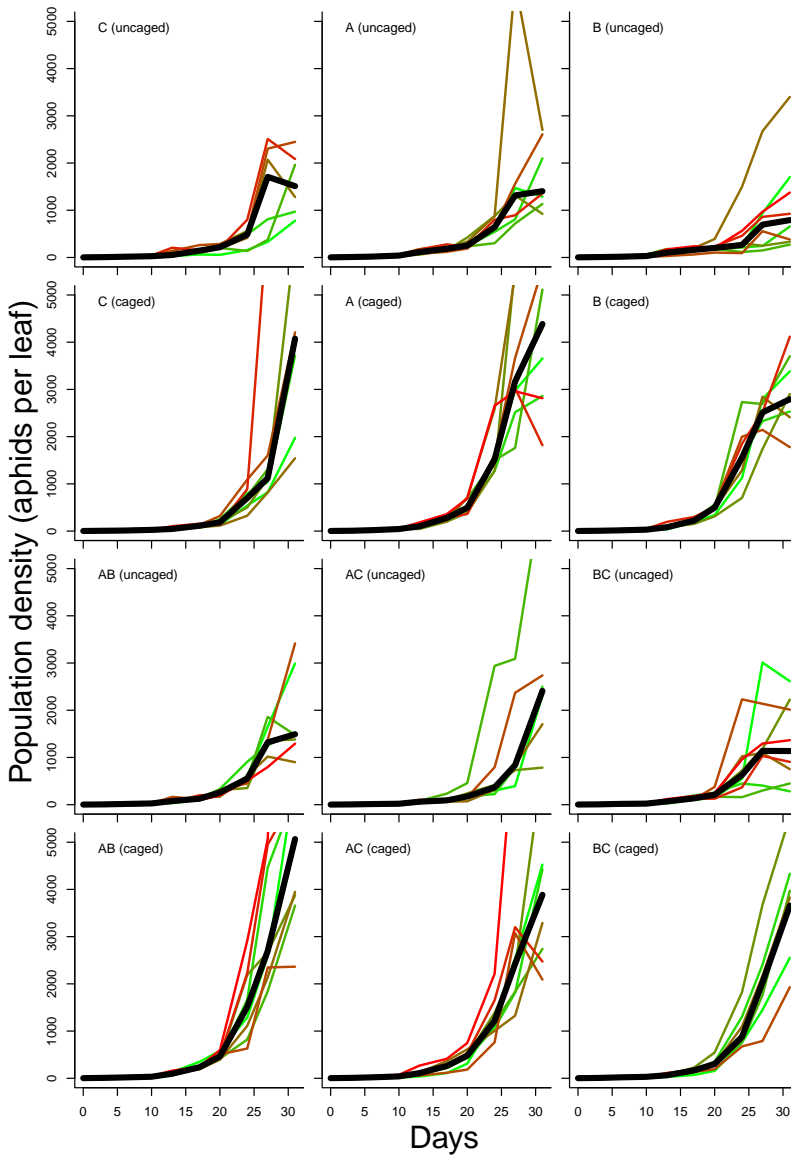
D.4 Population densities

Figure D3: Changes in the densities (number of individuals per leaf) over time for each population, shown per aphid treatment. These results are obtained by dividing observed population numbers (Fig. D2) by the number of leaves (Fig. D1). Lines show the total population densities. Thick black line shows the median density per treatment. Different colours indicate different replicates, and correspond to the colours shown in Fig. D1.

D.5 Implementation in JAGS

All models were fitted in a Bayesian framework, implemented in JAGS software using the R-package *rjags* (Plummer 2016). Three chains were run in parallel, and we checked convergence by

Gelman and Rubin's convergence diagnostic (using 1.05 as a threshold for each parameter). We used a burn-in period of at least 50,000 (which was extended if convergence was not yet achieved), and we took 50,000 samples from the posterior distributions after convergence.

D.6 Clonal frequencies

We quantified which processes were most important in determining transient daily population growth rates of the evolving populations (results presented in Fig. 5.5 of the manuscript). To assess the importance of changes in clonal frequencies (evolution), we used the observed clonal frequencies. On day 13, 20 and 31, between 16 and 32 aphids from each plant were genotyped at three microsatellite loci (for more details, see Turcotte et al. 2011b).

We performed a linear regression between frequency and day, for each aphid treatment separately (results below). Subsequently, the predicted frequencies at each measuring day, were used to calculate weighted vital rates of the evolving populations (see the manuscript for more details).
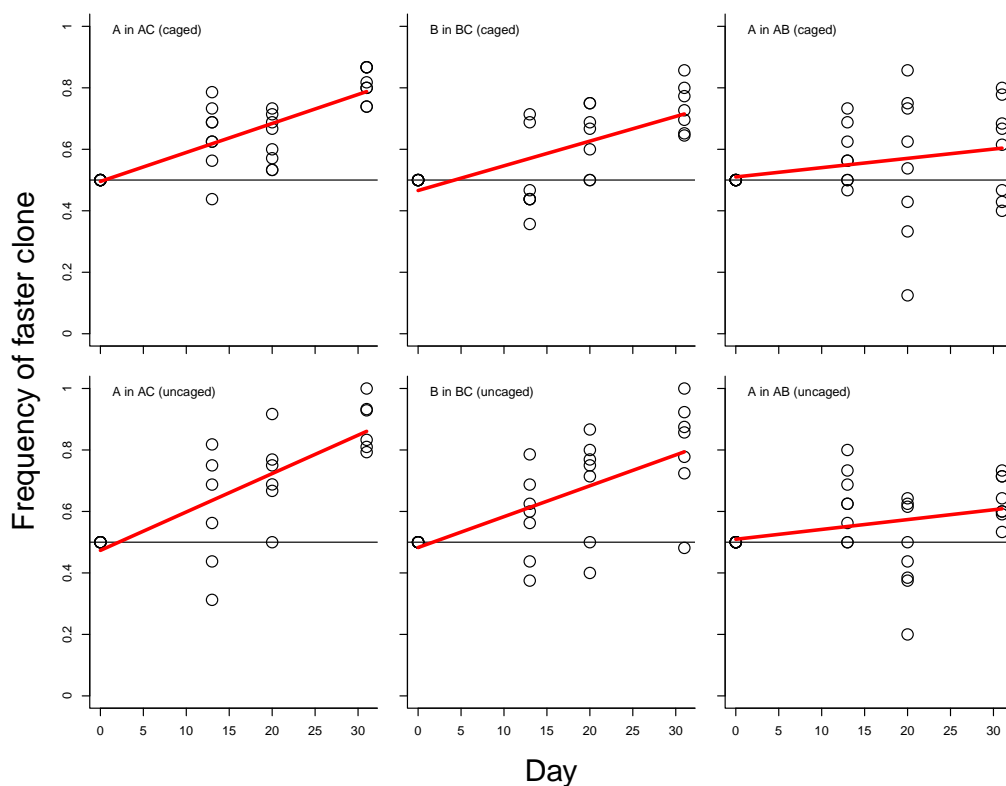

Figure D4: Frequency of the fastest clone in each evolution treatment. Dots represent the observed frequencies for different replicates, and red lines indicate the fit based on a linear regression.

D.7 Simulated data

To ensure that the defined models were identifiable and evaluate whether the 'correct' model could successfully be selected, we tested our inverse modelling approach with simulated data (for which the true relationships are known). We simulated data according to three different basic modelling structures. These were the null model (model 0), the model including stage-specific coefficients for survival (model 1), and the population size-dependent model (model 4 or model 6, which is equal to model 4 if plant size remains constant). Random survival, growth and reproduction parameters were drawn, as well as stage effects and density coefficients when relevant; the latter constrained to be below zero (in order to only simulated negative density dependence). Starting with 20 individuals in stage 2, we simulated (density-dependent) dynamics for 31 days. This was repeated three times to create three replicates. Per replicate, normally distributed noise was added to each of the parameters (standard deviation 0.05), leading to some variation between replicates. We recorded population size and stage composition on the same days as in the original experiment (i.e. every three or four days). We then continued with the simulations whenever population size of all replicates at day 31 was between 634 and 71512 individuals, which was the range of observed population sizes at day 31, to include only those parameter combinations that were realistic in the context of this experiment.

Using these data, models 0, 1 and 4 were all fitted in the same way as the experimental data, as explained in Section D.3. Based on K-fold cross-validation, we calculated the predictive ability of each model, and compared estimated parameters with the real, simulated, values. This was repeated ten times for each modelling structure, each time with a different parameter set. In total, this resulted in 30 simulated datasets, and three fitted models per dataset. To keep this computationally feasible, we created just one validation dataset (i.e. one out of the three replicates) and assessed predictive ability for only this replicate.

The true model resulted in the highest cross validation score in 10%, 60% and 100% of the simulations, when simulating according to model 0, 1 and 4, respectively (see Table 5.2 of the manuscript). For models 0 (including only intercepts) and 1 (stage-specific survival rates), differences in mean error between the correct model and the best-performing model were small (less than 5) suggesting practically equal model performance (Table 5.2). When simulating density-dependent dynamics, fitting models ignoring density (0 and 1) resulted in a large mean error and the correct model always resulted in a higher predictive ability (bottom row in Table 5.2). Estimated parameters (of the correct model) were close to the simulated parameters for the entire range of parameters and unbiased (Fig. 5.1). This is true for each of the three modelling structures (indicated by different symbols in Fig. 5.1). Credible intervals were generally small, and largest for the stage-specific survival estimates (red and blue dots in Fig. 5.1a).
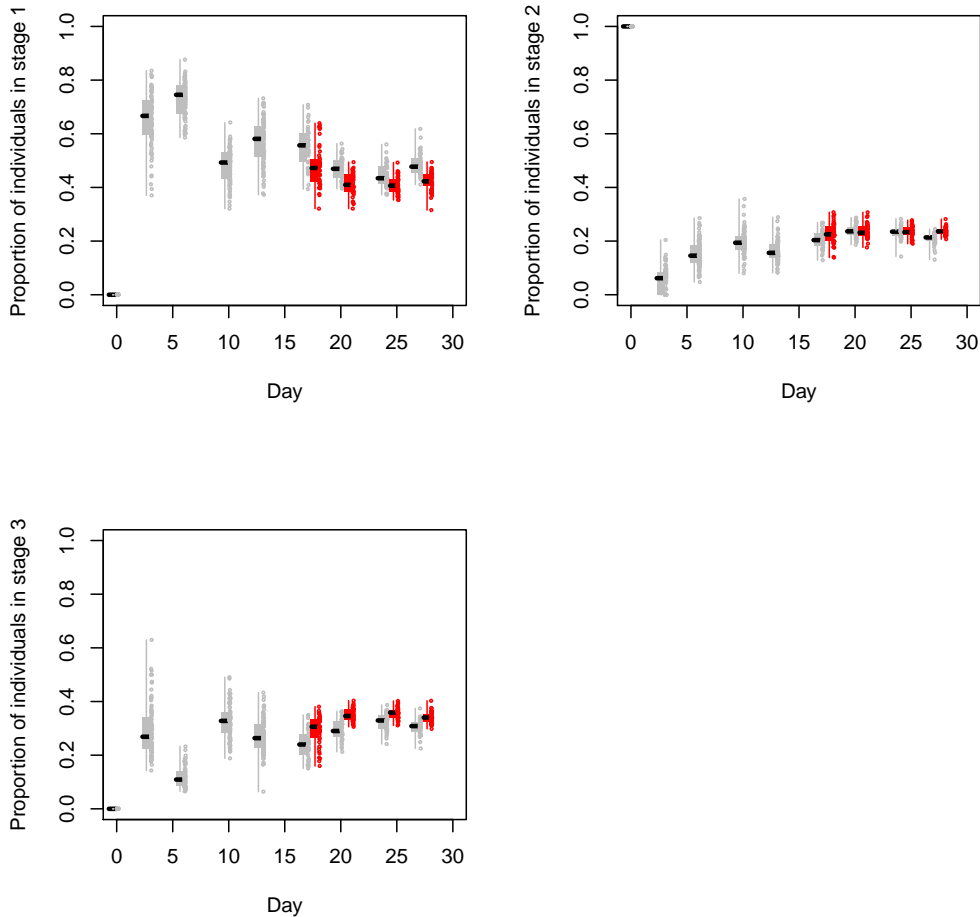
D.8 Population structure through time

Figure D5: Boxplots showing the proportion of individuals in each stage in the population counts through time. In grey, caged populations are shown and in red, uncaged populations are shown. Dots show individual data points, across all treatments.

D.9 LTRE for uncaged treatments

The analyses presented in Fig. 5.3 of the manuscript were performed for the caged treatments. Below we show the results when done for the uncaged treatments.
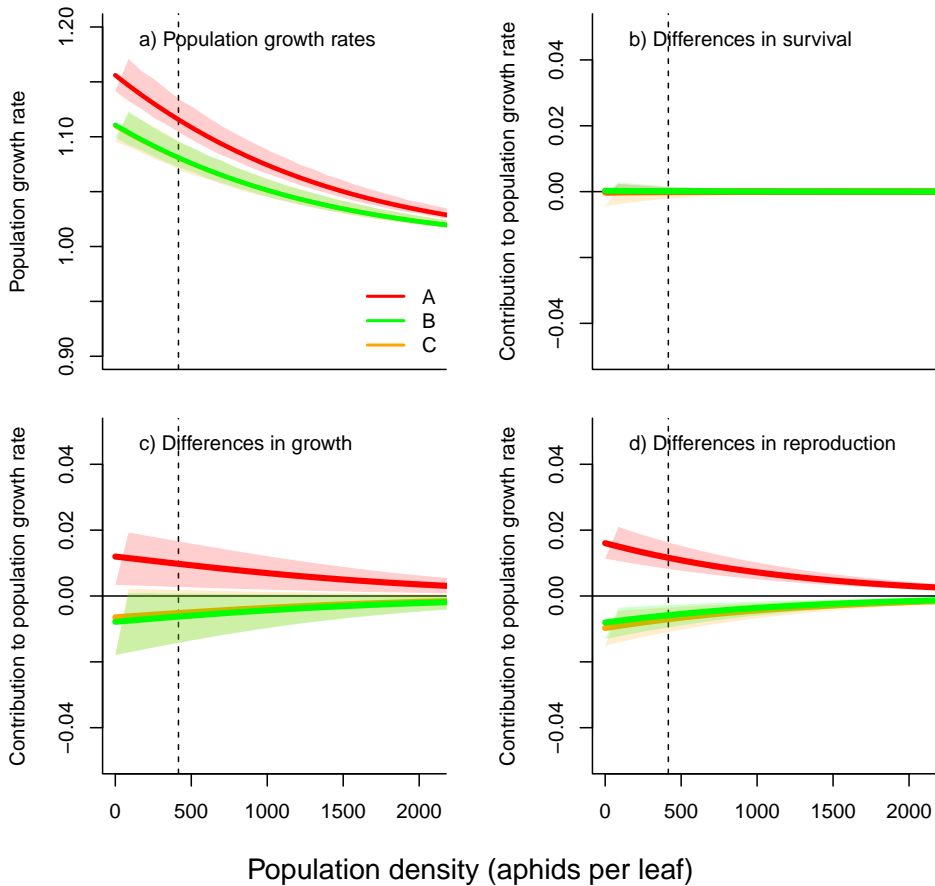

Figure D6: a) Projected population growth rate as a function of density for the three pure clones, under uncaged conditions. Population growth rate of clone C was essentially the same as population growth rate of clone B, and is therefore not visible. b-d) Life Table Response Experiment comparing different clones, as a function of density (individuals per leaf). Clone A-C were compared to the average matrix across the three clones. Densities range between 0 and the 95% quantile of observed densities, under uncaged conditions. Different colours indicate different clones and vertical lines indicate the average density under uncaged conditions. Shaded polygons show 95% confidence intervals in the predictions, obtained by simulating 1000 transition matrices by drawing coefficients from the posterior distributions of the clonal effects.

E Supplementary Information for Chapter 6

E.1 Integral Projection Modelling framework

To obtain long-term fitness measures for a life history shaped by a continuous underlying trait (x), we used Integral Projection Models (IPM; Box 6.2). In an IPM, the dynamics of a population are tracked in discrete time, the demography being driven by a continuous underlying trait (Easterling et al. 2000; Ellner and Rees 2006). An array of fitness measures can then be extracted, including in a stochastically varying environment (Ellner et al. 2007).

Three demographic functions define the IPM that we constructed: reproduction $R(x,k)$, or number of offspring produced, as a function of trait x and environment k; inheritance $D(x',m,s)$, which defines the probability density function of offspring having phenotype x', given genotype m and s; and, finally, survival S, which is set as a constant across size and phenotypes. Given that in our framework, x does not change during an individual's life, we do not require a function defining transitions resulting from growth; the diagonal of the discretized survival matrix was therefore set at the survival probabilities, with off-diagonal at 0. The three functions were combined accordingly:

$$n(t+1,x',k,m,s) = \int [S(x,k) + R(x,k) \cdot D(x',m,s)] n(t,x,k,m,s) dx \tag{E1}$$

See Table E1 for an overview of parameterization of these functions for the four scenarios. In scenarios A-C (Fig. 6.2a-c in the manuscript), offspring i is born with a phenotype x'_i randomly drawn from a normal distribution with mean m_i and variance s_i (Table E1). Both m_i and s_i are genetically determined values and are perfectly inherited (we assume asexual reproduction for simplicity and a heritability of 1). Larger s values result in a wider offspring phenotype distribution (reflecting intra-genotypic variability). In scenario D (Fig. 6.2d in the manuscript), we define two possible phenotypes. Offspring i obtains phenotype x'_1 with probability s and x'_2 with probability $1-s$. Here, if s is either 0 or 1, only one of the two traits is expressed. A value in between indicates intra-genotypic variability, and results in the expression of a combination of both phenotypes.

For scenarios A-C, we constructed IPMs for a total of 400 combinations; 20 values ranging between -2 and 2 describing mean genotype value m, and 20 values ranging between 0.01 and 4 describing variance s. For scenario D, where s defines the ratio of both traits being expressed (see above), we varied s between 0 and 1, for a total of 20 values. Each constructed IPM, as defined by equation E1, was discretized into a 200$\times$200 matrix (i.e. 200 classes of trait x). To avoid eviction (Williams et al. 2012) (in particular for higher s values), we let x values range between -8 and 8.

E.2 Long-term fitness landscapes

To explore the evolutionary dynamics associated with particular combinations of genotype for mean and variance, i.e. for a given genotype [$m*,s*$] we initiated a population vector consisting of the offspring size distribution $D(x',m*,s*)$ (see Table E1). We let the environment k vary over time (see below), and constructed environment-specific IPMs. These constructed IPMs were used to project population dynamics (i.e. project population size at $t+1$, $N(t+1)$, given $N(t)$) for 2000 time steps, which led to robust results (further increasing the number of time steps did not affect the outcome). Long-term intrinsic growth rate $\tilde{r}$ was calculated as:

$$\tilde{r} = \frac{\sum_{t=1}^{T} ln[(N(t+1)/N(t))]}{T} \tag{E2}$$

This enabled us to quantify how different combinations of m and s result in different long-term population growth rates, for a focal reproduction kernel $R(x,k)$ (scenarios A-D, depicted in Fig.

Scenario	Reproduction	Survival	Inheritance
A) Linear relation between trait x and reproduction.	$R(x,k) = 2 + 0.1xk$	$S(x,k) = 0.5$	$D(x',m,s) = N(x', \mu = m, \sigma^2 = s)$
B) Exponential relation between trait and reproduction.	$R(x,k) = \exp(0.3xk)$	$S(x,k) = 0.5$	$D(x',m,s) = N(x', \mu = m, \sigma^2 = s)$
C) Highest reproduction for intermediate x values.	$5 \cdot \exp(-((x-k)^2/0.5))$	$S(x,k) = 0.5$	$D(x',m,s) = N(x', \mu = m, \sigma^2 = s)$
D) A threshold relation between trait and reproduction.	If $x < 0$ and $k < 0$: $R(x,k) = 3$ If $x \geq 0$ and $k \geq 0$: $R(x,k) = 3$ else: $R(x,k) = 0.3$	$S(x,k) = 0.5$	$D(x',s) = Bernoulli(x', p = s)$

Table E1: Parameterization of the three demographic functions that together form the IPM (Eq. E1), for the four different scenarios.

6.2a-d in the manuscript) and environmental condition (depicted in Fig. 6.2e-f in the manuscript). We evaluated two environmental conditions: 1) a stable environment set at 0.5, and 2) a fluctuating environment randomly drawn from normal distribution with a mean of 0.5 and a variance of 1 (and no temporal correlation).

Because of the non-linear relations between long-term fitness, and m and s, we used thin plate splines to fit $\tilde{r}$ as a function of m and s for scenarios A-C. We confirmed that these captured the observed dynamics (Fig. E1). We used the fitted adaptive landscape to find the m and s values resulting in the highest long-term fitness ($\hat{m}$ and $\hat{s}$). For scenario D, we used smooth functions based on generalized additive models to fit $\tilde{r}$ as a function of s, finding the s value optimizing $\tilde{r}$.

E.3 Interactive effects of environmental variance and life span

Core analyses were performed with environmental variance set at 1 and survival set at 0.5. This survival probability sets the lower boundary for the yearly intrinsic growth rate: since we are keeping survival constant, yearly population growth rate will never drop below 0.5, even when reproduction is zero due to a mismatch between the phenotype and the environment. As long-term fitness is a multiplicative process, it is proportionally more affected by years with low growth rates. To expand our investigation beyond our core analyses, we varied survival from 0.0001 to 0.99, while holding reproduction either constant, or changing it accordingly while keeping average population growth rates similar. Additionally, we varied the environmental variance from 0.01 to 2 (reflecting an almost constant environment to a strongly fluctuating environment). This was done for both scenario C (optimum relation) and scenario D (discrete relation), as in these scenarios we found stabilizing selection on the variance (see Fig. 6.3 in the manuscript).

Given a survival probability and an environmental variance, we calculated the long-term fitness for different values of s. For scenario C, we hereby set the mean trait value m at its optimal value (0.5), to only assess the effect of variance, and we let s range between 0.01 and 4. For scenario D, we varied s between 0 and 1. To find the values for s optimizing long-term fitness, we used models proposed by Bull (1987) which are essentially equal to the reproduction kernel of our constructed IPMs. When we add (the constant) survival to Bull's models, we end up with the full IPM (as was confirmed; compare dots to lines in Fig. 6.4 of the manuscript). Optimal variance $\hat{s}$ decreases with an increasing survival probability, with the effects being stronger if reproduction is changed accordingly (blue versus red lines in Fig. 6.4), both for scenario C and D. Furthermore, increasing the environmental variance leads to an increase the optimal variance.
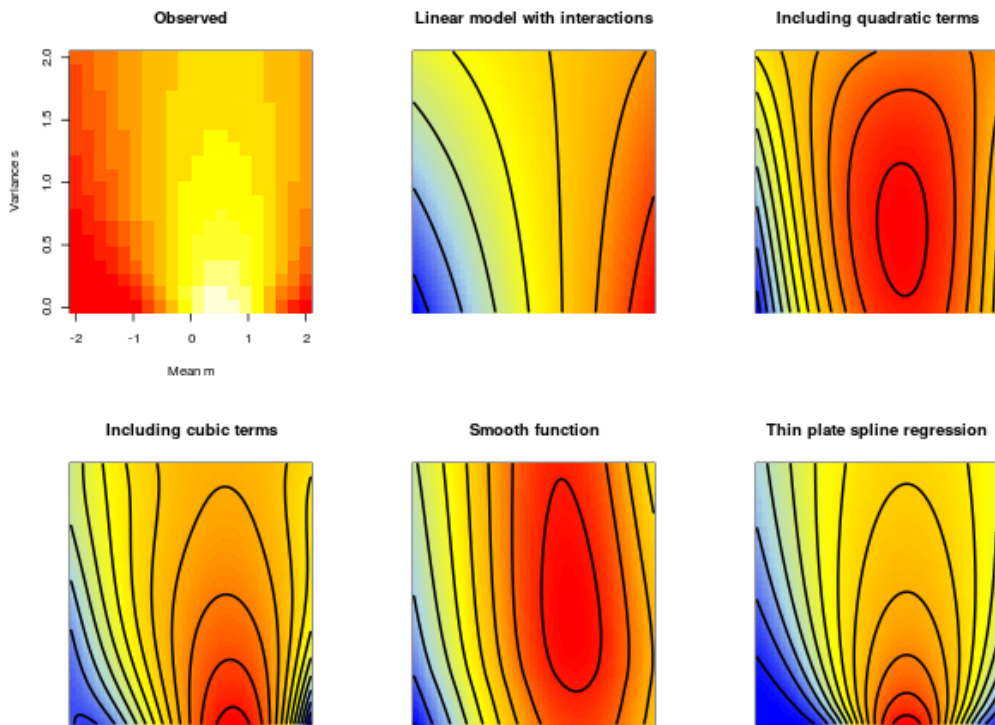
Figure E1: Overview of different approaches to fit the adaptive landscape. Graph in the upper left corner depicts an observed landscape for a total of 400 combinations of *m* and *s*. Other graphs show predictions of different regression methods, fitting long-term fitness as a function of *m* and s. We continued with the thin plate spline regression as this captured the observed patterns and allowed for most flexibility.

F Supplementary Information for Chapter 7

Supplementary Information can be found at https://tinyurl.com/bruijning2019-appF.

Summary

The world around us is constantly changing. All organisms, whether animal, plant or microbe must in one way or another deal with change. If not, they risk going extinct. How successful an organism is, or its "fitness", depends on its phenotype. A phenotype includes all characteristics of an organism such as body size, colour, behaviour and metabolism. These traits often vary in a population, even when it concerns individuals of the same species. In other words, there is often *phenotypic variation* within a population. This dissertation is about this phenotypic variation. Which mechanisms create phenotypic variation? How much variation is there within and between natural populations? How does a phenotype influence an individual's fitness, and does this change with environmental factors such as temperature? To find answers to the above questions, the dissertation presents a combination of theoretical and empirical studies, using water fleas and aphids as model species. As fitness consists of multiple components, for example clutch size, reproduction frequency and lifespan, this dissertation places an emphasis on the *integration* of such single fitness components, in order to arrive at a meaningful measure of fitness.

   Chapter 1 introduces three ways that all can result in phenotypic variation: phenotypic plasticity, genetic variation and intra-genotypic variability. *Phenotypic plasticity* is the ability of organisms to change, to a certain extent, their phenotype in response to the environment. For example, different bird species lay eggs earlier in the season, due to increasing temperatures caused by climate change. In addition, the phenotype of an individual is partly determined by its genotype (DNA), and hence *genetic variation* provides phenotypic variation. When some individuals are more successful than others because their phenotype is better suited to the environment, due to their genotype, an evolutionary response may occur. An evolutionary response is a change in the genetic composition of a population over time. Finally, the genotype of an individual cannot only influence the *average* phenotype, but also the amount of *variation* in a phenotype. For example, in artificial selection experiments it is often seen that some lineages show more variation in certain traits (such as birth weight) than other lineages. This *intra-genotypic variability* leads to phenotypic variation,

even between genetically identical individuals, experiencing the same environment.

Phenotypes in a population can change over time through both phenotypic plasticity and evolution. Understanding and predicting these phenotype dynamics is crucial for predicting the fate of natural populations that, for example, are facing climate change. **Chapter 2** compares four established frameworks to quantify the importance of evolution and plasticity. Simulated datasets are used that reflect different scenarios, such as ones with strong and weak selection pressure. It becomes clear that the frameworks differ in what they exactly quantify and in the definitions of different terms. For example, two frameworks are blind to the genetic component of a phenotype, and can therefore not distinguish between selection on heritable and non-heritable traits. The chapter emphasizes that researchers must be aware of these differences between the frameworks, and that the most suitable framework depends on the research question.

Chapters 3 and **4** present two empirical studies on water fleas (*Daphnia magna*). Water fleas are small crustaceans, often used in toxicological, ecological, genetic and evolutionary studies. Because they alternate sexual reproduction with asexual (clonal) reproduction, and this can be experimentally manipulated, the same genotypes can be used for different treatments in a controlled manner. **Chapter 3** investigates how five genotypes differ in individual growth, reproduction and survival (fitness components), how this results in differences in fitness, and how this depends on the temperature. For this purpose, four genotypes are used that are collected from a pond in Hilversum, and a fifth genotype that has been living under lab conditions for years. In a laboratory experiment, each of these genotypes is exposed to eight temperatures, varying between 10 and 26 °C, and the dynamics of these 40 populations are followed for 80 days. The experimental setup allows individuals to be followed while they are part of the population. This is important, for example, in order to estimate effects of population density on individual parameters. The collected data on individual fitness components are then integrated with a so-called *Integral Projection Model* (IPM), to obtain a measure of fitness. The five genotypes differ in fitness and in their thermal optimum. Moreover, fitness decreases with increasing population density in all genotypes, for example caused by competition for food. Interestingly, genotypes optimize fitness through different routes: one genotype mainly benefits from higher survival, while another benefits from higher reproduction. This study shows that natural water flea populations can have a considerable degree of phenotypic variation as a result of genotypic variation, and that genotypes react differently to temperature. This underlines the importance of genetic variation for natural populations to deal with climate change.

Chapter 4 does not look at variation *within* a natural population, but rather at variation *between* natural populations of water fleas. A total of twelve water flea genotypes are used that originate from four different locations, two of which are in Norway and two in Belgium. In this experiment genotypes are again exposed to different (four in this case) temperatures. In addition, pairwise combinations of genotypes are made, so that genotypes compete with each other for space and food. This reflects an expected scenario in which northern populations will face migrating southern populations, due to climate change. There is a similar setup as for **Chapter 3**, in which demographic data are collected from individuals from the populations. Halfway through the experiment, all populations became infected with the ectoparasite *Amoebidium parasiticum*, which causes a strong decline in numbers of water fleas. Interestingly, it is mainly the Norwegian genotypes that suffer from the parasite. Moreover, parasite effects show a strong temperature-dependence, with water fleas being more resistant at higher temperatures. This study shows that genotypes originating from different locations can vary greatly in demographic rates, fitness and parasite sensitivity. Future climate change will not only result in temperature changes, but will also change the dynamics of existing parasites and give new parasites a chance. This study suggests that this can have major consequences for local populations, especially if migrants carry along

parasites to which they are better adapted.

An evolutionary response can affect the number of individuals in the population. This change in population size can in turn influence the evolutionary response, when different genotypes react differently to density (as was shown in **Chapter 3**). This can result in an eco-evolutionary feedback loop, in which evolution and ecology (where population size is an ecological factor) influence each other. Indications for such a feedback loop were found in populations of aphids (*Myzus persicae*), and these results were published in 2011. In a field experiment, three genotypes were followed separately and in each pairwise combination, for four weeks. Differences in population growth among treatments suggested the presence of an eco-evolutionary feedback loop. **Chapter 5** presents a study in which these data are reanalyzed to investigate through which demographic rates this feedback loop occurs. Because no individuals have been explicitly followed in this study, these demographic data are not immediately available. Through *inverse* modeling, using data on changes in numbers and population structure over time, attempts are made to infer these rates. The inverse estimation of demographic rates is a complex problem, because there is often no unique solution. The reliability and robustness of the method used in this chapter is therefore checked using simulations. Modelling results suggest that the dynamics of all populations are largely influenced by plant size, as a measure for available resources. Moreover, the three genotypes differ in estimates of growth, reproduction and survival, and an integration of these rates shows that the relative fitness of the genotypes changes with density. In addition, evolving populations are able to realize higher population growth rates compared to non-evolving populations, caused by increased individual growth and reproduction. This study confirms the presence of an eco-evolutionary feedback loop in this system. Moreover, it underscores that mechanistic insight into eco-evolutionary feedback loops requires knowledge of the underlying demographic parameters, and how these vary among genotypes.

Chapter 6 explores in depth a third mechanism for phenotypic variation: how intra-genotypic variability creates phenotypic variation. The chapter starts by providing an overview of phenotypic traits and species in which intra-genotypic variability has been observed. Increasing numbers of empirical studies document that the degree of phenotypic variation is genetically determined, and identify locations on the DNA that regulate this variation. However, the conditions under which intra-genotypic variability is beneficial are imprecisely defined. Using demographic models, this chapter explores the theoretical fitness consequences of intra-genotypic variability. Results show that these fitness consequences depend on the relationship between the phenotypic trait and fitness (for example the relation between plant size and number of seeds produced). When this relationship is linear, the degree of phenotypic variation does not matter, and such a strategy can therefore not evolve through selection. When the relationship is convex, however, increased variation is beneficial. When there is an optimum phenotype, intra-genotypic variability is advantageous if this optimum also changes unpredictably over time (due to a changing environment). In the latter case, this is called *bet hedging*. Bet hedging is a risk-minimizing strategy, and follows the idea of 'do not put all your eggs in one basket'. Bet hedging has for example been observed in seed germination of plants. Optimum germination time depends on the year and is difficult to predict (e.g. the last day of frost varies yearly). In such a case, it may be beneficial to have seeds germinate across a longer period of time, instead of all seeds at the same time. Although in this way the number of new seedlings is never maximal, it ensures that there are always some seeds that germinate at the right time.

In **Chapters 3** and **4**, it was necessary to follow many populations of water fleas, which consisted of counting up to hundreds of individuals per population. Manually counting such numbers is unfeasible in practice, and possible only with automation. **Chapter 7** tackles this problem by introducing the openly available R-package *trackdem*. This package uses short videos

on populations of for instance water fleas as input. Code was developed to identify moving particles and reconstructed each particle's trajectory. With machine learning, the relevant particles (e.g. water fleas) can be distinguished from noise (e.g. air bubbles, debris). Moreover, the methodology is tested on simulated data and on various other systems such as insects, fish and moving cars. The methodology results in highly accurate estimates of population sizes and population structure in a very short time which makes the sampling of large numbers of populations possible.

Finally, **Chapter 8** presents a synthesis of the earlier chapters. Central to the chapter is a conceptual framework that shows the relationship between phenotypes, fitness components, population density and the environment. How natural populations respond to environmental change, depends on the amount of phenotypic variation within the population, and on how this phenotypic variation translates into variation in fitness. This chapter outlines possibilities to use population models to gain more insight into the mechanisms and circumstances that create this phenotypic variation. To conclude, as shown in this dissertation, genetic variation, phenotypic plasticity and intra-genotypic variability can all provide phenotypic variation. Predicting eco-evolutionary responses to future (climate) change requires disentangling these processes, for example separating phenotypes into a genetic and an environmental component. Moreover, population density is an important ecological factor to which genotypes can react differently. These density effects must therefore be explicitly considered, in addition to the effects of other environmental variables. Finally, this dissertation emphasizes that it is essential to estimate fitness based on integration over multiple fitness components, as each of these components can vary among genotypes. Altogether, taking into account the above points improves our understanding of how natural populations persist in the face of ever-changing conditions.

Nederlandse samenvatting

De wereld om ons heen verandert voortdurend. Alle organismen, of het nou gaat om dieren, planten of microben, moeten op de een of andere manier met deze veranderende omgeving om kunnen gaan. Zo niet, dan sterft de populatie uit. Hoe succesvol een organisme is (de "fitness") hangt af van al zijn/haar kenmerken tezamen: het fenotype. Een fenotype omvat kenmerken als lichaamsgrootte, kleur, gedrag en stofwisseling. Dit soort kenmerken varieert dikwijls in een populatie, ook wanneer het individuen van dezelfde soort betreft. Anders gezegd, er is vaak *fenotypische variatie* binnen een populatie. Dit proefschrift gaat over deze fenotypische variatie. Via welke mechanismen ontstaat fenotypische variatie? Hoeveel variatie is er binnen en tussen natuurlijke populaties? Hoe beïnvloedt het fenotype de fitness van een individu, en hoe verandert dit met omgevingsfactoren zoals temperatuur? Om antwoorden op bovenstaande vragen te vinden, presenteert dit proefschrift een combinatie van theoretische en empirische studies, gebruikmakend van watervlooien en bladluizen als modelsoorten. Aangezien fitness uit meerdere componenten bestaat, denk bijvoorbeeld aan het aantal jongen per worp, de frequentie waarop wordt gereproduceerd en de levensduur, ligt de nadruk in alle hoofdstukken op een *integratie* van dergelijke losse fitnesscomponenten, om zo tot een zinnige maat voor fitness te komen.

Hoofdstuk 1 introduceert drie manieren die allen kunnen zorgen voor fenotypische variatie en die in dit proefschrift worden onderzocht: fenotypische plasticiteit, genetische variatie en intra-genotypische variabiliteit. *Fenotypische plasticiteit* is het vermogen van organismen om, tot op zekere hoogte, hun fenotype te veranderen in reactie op de omgeving. Verschillende vogelsoorten leggen bijvoorbeeld steeds eerder in het seizoen eieren, als reactie op stijgende temperaturen door klimaatverandering. Daarnaast wordt het fenotype van een individu deels bepaald door zijn/haar genotype (DNA) en kan *genetische variatie* dus zorgen voor fenotypische variatie. Wanneer sommige individuen succesvoller zijn dan anderen (doordat ze een geschikter fenotype hebben voor de omgeving waarin ze zitten) als gevolg van hun genotype, kan er een evolutionaire respons optreden. Een evolutionaire respons is niets anders dan een verandering

door de tijd, in de genetische samenstelling van een populatie. Ten slotte, het genotype van een individu kan niet alleen het *gemiddelde* fenotype beïnvloeden, maar ook de hoeveelheid *variatie* in het fenotype. Zo wordt in artificiële selectie-experimenten vaak gezien dat sommige lijnen meer variatie in bepaalde eigenschappen (zoals geboortegewicht) tonen, dan andere lijnen. Deze *intra-genotypische variabiliteit* leidt tot fenotypische variatie, zelfs tussen individuen die genetisch identiek zijn en dezelfde omgeving ervaren.

Fenotypes in een populatie kunnen door de tijd heen veranderen door zowel fenotypische plasticiteit als door evolutie. Het begrijpen en kunnen voorspellen van deze fenotype-dynamiek is cruciaal om de toekomst van natuurlijke populaties, die bijvoorbeeld te maken krijgen met klimaatverandering, te kunnen voorspellen. **Hoofdstuk 2** vergelijkt vier gevestigde methoden om het belang van evolutie en plasticiteit te kwantificeren. Om deze methoden te kunnen vergelijken wordt gebruik gemaakt van gesimuleerde datasets die verschillende scenario's weerspiegelen, zoals een scenario met sterke en minder sterke selectiedruk. Er wordt duidelijk dat de methoden verschillen in wat ze precies kwantificeren en in de definities van verschillende termen. Zo zijn twee methoden blind voor de genetische component van een fenotype, met als gevolg dat ze geen onderscheid maken tussen selectie op erfelijke en op niet-erfelijke eigenschappen. Het hoofdstuk benadrukt dat onderzoekers zich bewust moeten zijn van deze verschillen tussen de methoden en dat de meest geschikte methode afhangt van de onderzoeksvraag.

Hoofdstuk 3 en **4** behandelen twee empirische studies over watervlooien (*Daphnia magna*). Watervlooien zijn kleine kreeftachtigen, die vaak worden gebruikt in zowel toxicologische, ecologische, genetische als evolutionaire studies. Doordat ze seksuele reproductie afwisselen met aseksuele (klonale) reproductie, en dit experimenteel gemanipuleerd kan worden, kunnen dezelfde genotypen op een gecontroleerde manier gebruikt worden voor verschillende behandelingen. **Hoofdstuk 3** onderzoekt hoe vijf genotypen verschillen in individuele groei, reproductie en overleving (fitnesscomponenten), hoe dit resulteert in verschillen in fitness en hoe dit afhangt van de temperatuur. Hiertoe worden vier genotypen gebruikt die afkomstig zijn uit een sloot in Hilversum en een vijfde genotype dat al jarenlang onder laboratoriumomstandigheden leeft. In een laboratoriumexperiment wordt elk van deze genotypen blootgesteld aan acht temperaturen, variërend tussen de 10 en 26 °C, en wordt de dynamiek van deze 40 populaties gedurende 80 dagen gevolgd. De experimentele opstelling maakt het mogelijk dat individuen, terwijl ze deel uitmaken van de populatie, worden gevolgd. Dit is bijvoorbeeld belangrijk om effecten van populatiedichtheid op individuele parameters te kunnen schatten. De verzamelde gegevens over losse fitnesscomponenten worden vervolgens geïntegreerd met een zogeheten *Integral Projection Model* (IPM), om een schatting van fitness te verkrijgen. De vijf genotypen verschillen in fitness en hebben verschillende temperatuuroptima. Bovendien neemt in alle genotypen de fitness af met toenemende populatiedichtheid, bijvoorbeeld veroorzaakt door competitie voor voedsel. Interessant is dat de genotypen fitness optimaliseren via verschillende routes: zo profiteert één genotype vooral van hogere overleving, terwijl een ander profiteert van hogere reproductie. Deze studie laat zien dat natuurlijke watervlo-populaties een behoorlijke mate van fenotypische variatie kunnen hebben als gevolg van genotypische variatie en dat genotypen verschillend reageren op temperatuur. Dit onderstreept het belang van genetische variatie voor natuurlijke populaties, om met klimaatverandering om te kunnen gaan.

Hoofdstuk 4 kijkt vervolgens niet naar variatie *binnen* een populatie, maar juist naar variatie *tussen* natuurlijke populaties watervlooien. Er wordt een totaal van twaalf watervlo-genotypen onderzocht, afkomstig van vier verschillende locaties, waarvan twee in Noorwegen en twee in België. Bij dit experiment worden opnieuw de genotypen aan verschillende (vier in dit geval) temperaturen blootgesteld. Bovendien worden hier paarsgewijze combinaties van genotypen gecreëerd, zodat ze met elkaar concurreren voor ruimte en voedsel. Dit weerspiegelt een verwacht

scenario waarin noordelijke populaties te maken krijgen met migrerende zuidelijke populaties, als gevolg van klimaatverandering. Er is een soortgelijke opzet als in **Hoofdstuk 3**, waarin demografische gegevens aan individuen binnen de populaties worden verzameld. Halverwege het experiment raakten alle populaties geïnfecteerd met de ectoparasiet *Amoebidium parasiticum*, wat zorgt voor een sterke achteruitgang in aantallen watervlooien. Interessant genoeg zijn het vooral de Noorse genotypen die lijden onder de parasiet. Bovendien blijken de gevolgen van de parasiet sterk af te hangen van temperatuur, waarbij watervlooien bij hogere temperaturen meer resistent zijn. Deze studie laat zien dat er tussen genotypen afkomstig van verschillende locaties, grote variatie kan zijn in demografie, fitness en gevoeligheid voor parasieten. Toekomstige veranderingen in het klimaat zullen niet alleen veranderingen in temperatuur teweegbrengen, maar zullen ook de dynamiek van aanwezige parasieten veranderen en nieuwe parasieten een kans geven. Deze studie suggereert dat dit grote gevolgen kan hebben voor lokale populaties, zeker als migranten parasieten meenemen waartegen zij beter bestand zijn.

Een evolutionaire respons kan een effect hebben op het aantal individuen in de populatie. De populatiegrootte kan vervolgens de evolutionaire respons beïnvloeden, wanneer verschillende genotypen anders reageren op populatiegrootte, zoals aangetoond in **Hoofdstuk 3**. Hierdoor kan er een eco-evolutionaire terugkoppeling ontstaan, waar evolutie en ecologie (waarbij populatiegrootte een ecologische factor is) elkaar beïnvloeden. Aanwijzingen voor een dergelijke terugkoppeling werden gevonden in populaties bladluizen (*Myzus persicae*) en deze resultaten zijn in 2011 gepubliceerd. In een veldexperiment werden drie genotypen afzonderlijk en in elke paarsgewijze combinatie, gedurende vier weken gevolgd. Verschillen in populatiegroei tussen deze behandelingen suggereerden een terugkoppeling tussen de evolutionaire respons en de populatiegrootte. In **Hoofdstuk 5** worden de gegevens opnieuw geanalyseerd om te onderzoeken via welke demografische parameters deze terugkoppeling optreedt. Omdat er in deze studie geen individuen expliciet zijn gevolgd, zijn deze demografische gegevens niet direct beschikbaar. Via *invers* modelleren, gebruikmakend van gegevens over de veranderingen in aantallen en populatiestructuur door de tijd heen, wordt geprobeerd deze parameters af te leiden. Het invers schatten van demografische parameters is een ingewikkeld probleem, omdat er vaak geen unieke oplossing is. De betrouwbaarheid en robuustheid van de in dit hoofdstuk gebruikte methode wordt gecontroleerd aan de hand van simulaties. Resultaten suggereren dat de dynamiek van alle populaties voor een groot deel wordt beïnvloed door plantgrootte, als maat voor aanwezig voedsel. Tevens verschillen de drie genotypen in schattingen van groei, reproductie en overleving. Een integratie over deze parameters laat zien dat de relatieve fitness van de genotypen verandert met dichtheid. Daarbij suggereren de resultaten dat de evoluerende populaties in staat zijn sneller te groeien, veroorzaakt door verhoogde individuele groei en reproductie. Deze studie bevestigt de aanwezigheid van een eco-evolutionaire terugkoppeling in deze populaties. Bovendien onderstreept de studie dat het kennen van de onderliggende demografische parameters en de variatie hierin tussen genotypen, meer inzicht kan bieden in eco-evolutionaire terugkoppelingen.

Hoofdstuk 6 verkent vervolgens een derde mechanisme dat kan zorgen voor fenotypische variatie: intra-genotypische variabiliteit. Het hoofdstuk begint met een overzicht van fenotypische eigenschappen en van soorten waarin intra-genotypische variabiliteit is geobserveerd. Zo zijn er meer en meer empirische voorbeelden bekend waarin de mate van fenotypische variatie inderdaad genetisch bepaald is en worden er plekken op het DNA gevonden die deze variatie reguleren. Onder welke omstandigheden intra-genotypische variabiliteit voordelig is, is echter nog grotendeels onontdekt terrein. Dit hoofdstuk onderzoekt de theoretische fitnessgevolgen van intra-genotypische variabiliteit, door gebruik te maken van een theoretisch model. Resultaten tonen dat de fitnessgevolgen van intra-genotypische variabiliteit afhangen van de relatie tussen de fenotypische eigenschap en een of meer fitnesscomponenten (bijvoorbeeld de relatie tussen

plantgrootte en aantal geproduceerde zaden). Wanneer deze relatie lineair is, doet de mate van fenotypische variatie er niet toe, en kan een dergelijke strategie dus niet door selectie evolueren. Wanneer de relatie convex is, is toegenomen variatie daarentegen wel voordelig. Wanneer er een optimum fenotype bestaat, is intra-genotypische variabiliteit voordelig als dit optimum tevens ook onvoorspelbaar fluctueert door de tijd (door een veranderende omgeving). In dit laatste geval, is er sprake van *bet hedging*. Bet hedging is een risico-minimaliserende strategie, en volgt het idee 'niet op één paard wedden'. Bet hedging is onder andere in zaadontkieming van planten geobserveerd. Afhankelijk van het jaar, kan het optimale ontkiem-moment variëren (bijvoorbeeld afhankelijk van wanneer de laatste vorst is). Dit optimale moment is echter moeilijk te voorspellen. In een dergelijk geval kan het gunstig zijn om zaden altijd verspreid over een langere periode te laten ontkiemen, in plaats van alle zaden op één moment. Op deze manier is de opbrengst (het aantal zaailingen) dan wel nooit maximaal, er zijn in ieder geval altijd wat zaden die ontkiemen op het juiste moment.

Voor **Hoofdstuk 3** en **4** was het nodig om vele populaties watervlooien door de tijd heen te volgen. Het handmatig tellen van aantallen (tot in de honderden per populatie) was praktisch onhaalbaar. Om dit te automatiseren wordt in **Hoofdstuk 7** een R-package *trackdem* gepresenteerd dat openlijk beschikbaar is voor andere gebruikers. Dit package gebruikt korte video's als input. Via een aantal functies worden bewegende deeltjes in de video herkend, om vervolgens het afgelegde traject van ieder deeltje te reconstrueren. Met machinaal leren kunnen de relevante deeltjes (bijvoorbeeld watervlooien) van ruis (bijvoorbeeld luchtbubbels, vuil) worden onderscheiden. Ook wordt de methode getest op gesimuleerde data en op diverse andere systemen zoals insecten, vissen en rijdende auto's. Deze methodologie resulteert in zeer accurate schattingen van populatiegroottes en populatiestructuur, en maakt het mogelijk om in zeer korte tijd grote aantallen populaties te tellen en op te meten.

Hoofdstuk 8 presenteert tenslotte een synthese van de eerdere hoofdstukken. Centraal in het hoofdstuk staat een conceptueel kader dat de samenhang tussen fenotypes, fitnesscomponenten, populatiedichtheid en de omgeving weergeeft. Hoe natuurlijke populaties reageren op een verandering in de omgeving, hangt af van de hoeveelheid fenotypische variatie binnen de populatie, en van hoe deze variatie resulteert in variatie in fitness. Het hoofdstuk schetst een aantal mogelijkheden voor hoe populatiemodellen gebruikt kunnen worden om meer inzicht te krijgen in de mechanismen en omstandigheden die zorgen voor deze fenotypische variatie. Concluderend, dit proefschrift laat zien dat zowel genetische variatie, fenotypische plasticiteit als intra-genotypische variabiliteit kunnen zorgen voor fenotypische variatie. Voor het voorspellen van eco-evolutionaire responsen, moet het belang van elk van deze processen worden ontrafeld, bijvoorbeeld door het scheiden van fenotypes in een genetische en een omgevingscomponent. Bovendien is populatiedichtheid een belangrijke ecologische factor, waarop genotypen verschillend kunnen reageren. Deze dichtheidseffecten moeten daarom expliciet worden meegenomen, naast de effecten van andere omgevingsvariabelen. Tot slot is het belangrijk om fitness te schatten door verschillende componenten te integreren, aangezien elk van deze componenten kan variëren tussen genotypen. Al met al, rekening houden met bovenstaande punten zorgt voor een beter begrip van de manier waarop natuurlijke populaties omgaan met veranderingen in de omgeving.

Dankwoord

Ik heb een ontzettend fijne, leuke en leerzame tijd gehad op de ecologie-afdeling in Nijmegen, waar ik sinds november 2010 rondloop. Eerst als student, later als promovenda. Het is dan ook met gemengde gevoelens dat ik nu vertrek. Enerzijds is het goed om ergens anders heen te gaan en kijk ik uit naar wat er komen gaat. Anderzijds vind ik het ook verdrietig om Nijmegen te verlaten, en ik ga alles en iedereen erg missen. Het is hier waar ik heb ontdekt hoe leuk en spannend ik het doen van onderzoek vind. Ik ben iedereen die hier gedurende de afgelopen jaren aan heeft bijgedragen heel erg dankbaar.

Henk, ik dank je voor de mogelijkheid die je me hebt gegeven om binnen de afdeling Dierecologie mijn promotie-onderzoek uit te voeren. Bedankt voor het vertrouwen om mijn eigen richting aan mijn onderzoek te geven, en ook voor de ruimte om deel van mijn onderzoek bij andere labs uit te voeren. Ik dank je ook voor je hulp de laatste maanden om alles vlot af te kunnen ronden. Eelke, ik heb veel van je geleerd na zes jaar intensieve samenwerking. Over analyseren van data, statistiek, populatiemodellen, wetenschappelijk schrijven, alsook over het hele submit- en review-proces. Ik waardeer je eerlijkheid als onderzoeker, je oprechte interesse in mijn werk, en je plezier in het doen van onderzoek. Het is fijn dat ik het gevoel had dat de deur altijd open stond, en dat ik alles kon vragen. Ik dank ook jou dat je me de kans hebt gegeven samenwerkingen met andere labs op te zetten, en me verder te ontwikkelen in Leuven en Princeton.

Graag wil ik ook Philippine en Joop bedanken, die tijdens mijn eerste masterstage mijn interesse in ecologisch onderzoek hebben aangewakkerd. Hans, ik heb veel van je geleerd tijdens mijn tweede masterstage. Hartelijk dank ook voor je interesse in mijn promotietraject. Beginnen als stagiaire bij de ecologie-afdeling voelde als een warm bad, en dat komt voor een groot deel door

de medestudenten van destijds. Annieke, je bent een van de meest loyale mensen die ik ken, en je hebt mijn promotietraject van erg dichtbij gevolgd (en ik dat van jou). Ooit begonnen we als 'echte meisjes in de jungle' in Panama, waar we lief en (ook zeker) leed deelden. Het is bijzonder dat we nu ongeveer tegelijk weggaan uit Nijmegen, en ik dank je voor je vriendschap en betrokkenheid. Onno en Isabella, het is fijn ook met jullie het hele promotietraject te hebben gedeeld, en ook jullie bedankt voor alle interesse en betrokkenheid. Michiel, Ruud, Lisanne, Hilke, Laura en Clara, ik heb hele warme herinneringen aan onze tijd op de afdeling als student (jong en optimistisch als we waren), en het is bijzonder om nog steeds contact te hebben!

Ik heb er erg van genoten om met verschillende groepen samen te werken, en ik ben dankbaar voor alles wat ik hiervan heb geleerd. Luc, ik dank je voor de mogelijkheid het grote *Daphnia*-experiment in Leuven uit te voeren. Ik vind de dynamiek in je groep erg inspirerend, en ik dank je voor de kans om hier een paar maanden deel van uit te maken. Jessica and Julien, thank you for giving me the opportunity to come back to Princeton, as a postdoc. I'm very much looking forward to continue the collaboration! Koen, ik waardeer je erg als collega en als mens. Ik hoop dat we in de toekomst nog eens aan een gezamenlijk project zullen werken. Arpat, Erik, Timothée, Martin, Erlend, Joost and Lynn, thank you all for the insightful and inspiring collaborations.

Iets waar ik veel plezier in heb gehad, energie van kreeg en zelf ook veel van heb geleerd, is het begeleiden van studenten. Ik wil graag mijn elftal aan studenten hiervoor bedanken. Marit en Robin, dankjulliewel voor het allereerste uitzoek-werk met de *Daphnia* (hoe zien ze er eigenlijk uit? Wat eten ze?). Selwyn en Sven, ik dank jullie voor het zetten van de eerste stappen met de automatische beeldanalyse, wat later tot het R-package *trackdem* heeft geleid. Lisenka, dankje voor het verzamelen van de eerste demografische data aan de *Daphnia* genotypen. Anne, het was leuk om samen met jou het eerste watervlo-experiment uit te voeren, en ik ben je dankbaar voor je enthousiasme, toewijding en gezelligheid. Maarten and Tanya, thank you for the hard work and your independence when I wasn't in Nijmegen to help. Héléne, we hadden geluk met jou als student om te helpen met het experiment in Leuven. Ik wens je alle succes nu met je eigen promotie-onderzoek. Finally, Marta and Pauline, I'm very grateful to both of you for all the hard work in analyzing 2000 genetic samples.

Daarnaast zijn er in Nijmegen een heleboel mensen in meer of mindere mate betrokken geweest bij mijn onderzoek. Wilco, bedankt voor je belangstelling in mijn onderzoek. Marij, bedankt voor je hulp bij de eerste *Daphnia* die in Nijmegen arriveerden. Tom, bedankt voor je goede adviezen en hulp rondom de labopstelling. Niels, ik ben je erg dankbaar voor je geduldige hulp bij de genetische analyses. Peter, José (A, B en J), Daisy, Wim en Jan, ook jullie bedankt voor hulp omtrent praktische zaken.

Mede-promovendi, dankjulliewel voor het delen van alle leuke, bijzondere, frustrerende en grappige momenten. Nils, rug-aan-rug hebben we een groot deel van ons onderzoek volbracht, dankjewel voor je interesse en het meedenken. Valérie, ik vond het fijn de laatste maanden samen met jou op een kantoor te zitten en deze laatste fase te delen. Natan, het was leuk jou als dierecologie-maatje te hebben. Caspar, dankjewel voor al je inhoudelijke ideeën en suggesties. Laura, Janneke, Marloes en Maartje, toen ik als student begon, liepen jullie net rond als promovendi. Het was leuk via jullie al mee te krijgen wat promoveren allemaal behelst en het hele proces alvast te zien. Ik dank

jullie voor de interesse en wijze raad. Nick, Dina, Yingying, Qian, Ralf, Ralph, Marlous, Marloes, Stefan, Eva, Renske, Tamara, Ernandes, Sarah Faye, Jacqueline, Tom, en alle anderen, ook jullie bedankt voor het luisterend oor, de gezelligheid en de borrels! De goede sfeer in Vleugel 1.1 wordt gecreëerd door iedereen die daar werkt. Jan Willem, Roy, Germa, Hannie, Peter, Eric, Annelies, Titi, Leon, Annemiek, Heidi, Jeroen, Christian, Sarian, en iedereen die niet in dit lijstje staat, allen bedankt voor jullie bijdrage daaraan!

Een groot deel van mijn leven speelt zich nog af naast de universiteit, en dat is maar goed ook. Lieve Anne, Dorien, Femke, Imke, Jolien, Leonie, Lieke, Martine en Michelle, ik ben dankbaar voor onze hechte vriendschap die er is sinds het begin van onze studententijd. Ik denk oprecht dat mijn proefschrift er zonder jullie niet zou zijn. Het was heerlijk om wekelijks het wel en wee van de watervlooien met jullie te kunnen bespreken, de boel te relativeren en mijn verhaal kwijt te kunnen. We hebben al zoveel met elkaar meegemaakt en ik koester ontelbaar mooie, onvergetelijke herinneringen aan woensdagavonden, feestjes, weekendjes, 21-diners en skivakanties. Anne, dankjewel voor je relativerend vermogen en je nuchtere kijk op de dingen. Dorien, het is altijd fijn om je te zien en ik dank je voor je interesse. Femke, ovum-maatje, tennis-maatje, en trouwe vriendin vanaf het begin. Dankjewel! Imke, een bijzonder plekje in mijn hart hebben onze reizen naar Thailand en Australië. Ik ben dankbaar voor onze vriendschap. Jolien, dankjewel voor je eerlijkheid, je scherpe blik en je trouwe vriendschap. Leonie, dankjewel voor de gezelligheid, interesse en betrokkenheid wanneer we elkaar zien. Lieke, je bent altijd voor alles in, altijd vrolijk, en niks is te veel. Bedankt voor je belangstelling in mijn onderzoek. Martine, er is niemand die zo heeft meegeleefd gedurende mijn hele promotie-traject, als jij. Duizendmaal dank hiervoor! Michelle, ik weet dat ik altijd bij je terecht kan, en ik dank je voor je vriendschap. Martine en Femke, bedankt dat jullie mijn paranimf willen zijn! Ook al gaat iedereen nu meer haar eigen weg (en zijn we al een boel prachtige jaarclub-baby's rijker, met nog meer op komst, zelfs in duplo), jullie zijn heel belangrijk voor me. Ik hoop dat er nog veel nieuwe herinneringen zullen volgen.

Mike, ik ben blij met onze vriendschap, die er is vanaf de eerste dag dat ik in Nijmegen woon. Ik bewonder je energie en je enthousiasme. Dankjewel voor je trouwe bezoekjes in Rheden, Montpellier, de Dordogne en Panama. Deze zomer New York, samen met Kristy? Kim, we kennen elkaar meer dan ons halve leven, en als ik aan mijn middelbare schooltijd denk, denk ik onmiddellijk aan jou. Ik ben blij dat we nog steeds contact hebben!

Mijn familie heeft een belangrijke rol gespeeld gedurende het hele promotietraject. Helena, dankjewel voor je betrokkenheid, je zorgzaamheid en je support. Janneke en Hamed, ik wens jullie alle goeds en hoop dat we elkaar snel weer zien. Sieger en Ducky, bedankt voor jullie belangstelling en voor jullie hartelijkheid. Catrien, dankjewel dat je er bent, voor Jeroen, maar ook voor ons. Opa, we delen het enthousiasme voor de biologie, en ik luister altijd met veel plezier en interesse naar de verhalen over je belevenissen als docent. Lijs, het is fijn om zo'n bijzondere oma te hebben. Miriam, ik heb bewondering voor wat je voor elkaar krijgt en hoe je in het leven staat. Ik ben je ook super dankbaar voor de mooie tekeningen die je hebt gemaakt voor dit proefschrift. Als laatste, Carla en Jeroen, ik ben jullie ontzettend dankbaar voor alles. Het is door hoe jullie me hebben grootgebracht, met vertrouwen, geborgenheid en ruimdenkendheid, dat ik dit proefschrift heb kunnen (en überhaupt willen) schrijven.

Ten slotte, lieve Marco. Ik heb ongelooflijk veel aan je gehad, zowel persoonlijk als professioneel. Zonder jou was ik als onderzoeker niet geweest waar ik nu ben. Dankjewel voor je steun en voor je liefde, maar ook dat je me continu uitdaagt het beste uit mezelf te halen. Dankjewel voor je vertrouwen in mijn kunnen, als mij dat soms even ontbreekt. Twee jaar met een oceaan tussen ons in was lang, maar ik ben blij met de manier waarop we het hebben doorstaan. Ik ben trots op waar we nu zijn. En ik verheug me op onze toekomst, samen.

Curriculum Vitae

Marjolein Bruijning was born in Leiden on the 17th of May 1988. She grew up in Voorburg, together with her parents and sister, and finished high school in 2006. In 2007, she started studying biology at Radboud University Nijmegen, where her internships in particular contributed to her enthusiasm in doing research. As an intern she studied the effects of shade stress on phenotypic and epigenetic patterns in *Scabiosa columbaria*, under supervision of Philippine Vergeer. In a second internship, Marjolein performed field work on Barro Colorado Island (Panama), under supervision of Marco Visser. She collected field data on the gender identity of dioecious trees, and combined this with data on the performance of seeds, seedlings and trees to investigate how dioecious species compensate for the presence of males. After obtaining her masters in 2012, Marjolein was given the opportunity to work as a junior researcher for one year on the dynamics of geese populations. She started her PhD-project at the Animal Ecology and Physiology department at Radboud University in 2013, under supervision of Eelke Jongejans. Here she applied theoretical and statistical models to empirical data, quantifying how both plasticity and genetics shape phenotypes and eventually fitness of, for instance, water fleas. Marjolein is fascinated by evolution, and has a deep interest in using models and quantitative methods to understand how evolution shapes life around us, across species and systems. After graduating, Marjolein will start a post-doctoral position at Princeton University (USA), in Jessica Metcalf's lab. Here, she hopes to take steps in combining the fields of demography and genetics.

Publications

- **Bruijning M.**, ten Berge A.C.M. & Jongejans E. (2018). Population-level responses to temperature, density and clonal differences in *Daphnia magna* as revealed by integral projection modelling. *Functional Ecology*, 32.10: 2407-2422.
- **Bruijning M.**, Visser M.D., Hallmann C.A. & Jongejans, E. (2018). *trackdem*: Automated particle tracking to obtain population counts and size distributions from videos in R. *Methods in Ecology and Evolution*, 9.4: 965-973.
- **Bruijning M.**, Visser M.D., Hallmann C.A. & Jongejans, E. (2018). *trackdem*: Particle Tracking and Demography. R-package version 0.1-0.4.2, https://CRAN.R-project.org/package=trackdem.
- **Bruijning M.**, Visser M.D., Muller-Landau H., Wright S.J., Comita L., S.P. Hubbell, de Kroon H. & Jongejans E. (2017). Surviving in a cosexual world: a cost-benefit analysis of dioecy in tropical trees. *The American Naturalist*, 189.3: 297-314.
- Van Benthem K.*, **Bruijning M.***, Bonnet T.*, Jongejans E., Postma E. & Ozgul A. (2017). Disentangling evolutionary plastic and demographic processes underlying trait dynamics: A review of four frameworks. *Methods in Ecology and Evolution*, 8.1: 75-85.
- Visser M.D., **Bruijning M.**, Wright S.J., Muller-Landau H., Jongejans E., Comita L. & de Kroon H. (2016). Functional traits as predictors of vital rates across the life-cycle of tropical trees. *Functional Ecology*, 30.2:168–180.
- * Contributed equally as first authors.

International presentations

- **Invited seminar speaker at the Eawag** (Swiss Federal Institute of Aquatic Science and Technology), Dübendorf (Switzerland), 2018. The importance of plastic and genetic factors for *Daphnia* populations as revealed by Integral Projection Modelling.
- **Seminar speaker at the KU Leuven** (Belgium), Seminar Series in Ecology, Evolution & Biodiversity, 2017. How can fitness be optimized?
- Oral presentation at the **European Society for Evolutionary Biology** (ESEB), Groningen (the Netherlands), 2017. The importance of plastic and genetic factors for *Daphnia* populations as revealed by Integral Projection Modelling.
- Invited speaker at the **Ecological Society of America Annual Meeting** (ESA), Baltimore (USA), 2015. A cost benefit analysis of dioecy in tropical trees.
- Lightning talk and poster presentation at the **Evolutionary Demography Society Annual Meeting**, Lunteren (the Netherlands), 2015. Using automated particle tracking to obtain demographic information from video sequences in R.
- Poster presentation at the **Evolutionary Demography Society Annual Meeting**, Odense (Denmark), 2013. Quantifying the importance of plasticity and evolution in aquatic ectotherms.
- Oral presentation at the **Netherlands Annual Ecology Meeting**, Lunteren (the Netherlands), 2013. Living in a cosexual world: dioecy as a reproductive strategy.
- Oral presentation at the **Netherlands Annual Ecology Meeting**, Lunteren (the Netherlands), 2012. The effect of shade stress on the epigenetic and phenotypic patterns in inbred and outbred lines of the plant species *Scabiosa columbaria*.

Grants and Awards

- **Christine Mohrmann Grant**, 2016. Annual prize for promising female PhD candidates, awarded by the Executive Board of Radboud University, used to visit Princeton University between January and May 2017.
- **Van Nieuwenhoven Prize**, 2014 Annual prize for the best master student in biology and medical biology at the Radboud University Nijmegen.
- **Unilever Research Prize**, in biology, 2012. Annual prize for outstanding master thesis research at Dutch universities.
- **Smithsonian Tropical Research Institute**, 2012. Short-term fellowship awarded for the study: Surviving in a cosexual world: dioecy as a reproductive strategy.