Doctoral Programme in Integrative Life Science
Faculty of Biological and Environmental Sciences
University of Helsinki
Finland

Quantitative genetics in nonlinear genotype-phenotype maps

Lisandro Milocco

Doctoral dissertation

To be presented for public examination with the permission of the Faculty of Biological and Environmental Sciences of the University of Helsinki, in room 2402, Biocenter 3, on the 30th of September, 2022 at 13 o'clock.

Helsinki 2022

Supervisor
Isaac Salazar-Ciudad, Autonomous University of Barcelona, Spain

Thesis advisory committee
Juha Merilä, University of Helsinki, Finland
Heikki Helanterä, University of Oulu, Finland

Pre-examiners
Jon Brommer, University of Turku, Finland
P. David Polly, Indiana University, United States of America

Opponent
Kjetil Lysne Voje, University of Oslo, Norway

Custos
Craig Primmer, University of Helsinki, Finland

The Faculty of Biological and Environmental Sciences uses the Ouriginal system (plagiarism recognition) to examine all doctoral dissertations.

ISSN 2342-3161 (print)
ISSN 2342-317X (online)
ISBN 978-951-51-8487-0 (paperback)
ISBN 978-951-51-8488-7 (PDF)
Helsinki 2022
Unigrafia

Abstract

Understanding the genotype-phenotype map (GPM), or how specific genetic variation relates to specific phenotypic variation, is a fundamental objective of biology. Particularly for the study of evolution, the GPM is important since it bridges genes, the heritable units, with the traits that interact with the environment and determine fitness. The applied field of quantitative genetics approximates the GPM using linear statistical models. This approximation is the basis of many important applications, including methods to predict the response to selection in a population. The other major field concerned with the GPM is evolutionary developmental biology or evo-devo, which studies the gene and cell interactions by which the phenotype is built through the process of development. A main result from this field is that the GPM is notoriously complex and nonlinear, with genetic effects being highly dependent on genetic background, biophysical factors and the environment. This can be in conflict with the linear approximation of quantitative genetics. Because evo-devo and quantitative genetics have developed independently, there is a gap in the understanding of how these conceptualizations of the GPM fit with each other, what are their limitations, and if there is a potential to improve our ability to make prediction about evolutionary systems by combining insights from both approaches. This thesis deals with these issues.

The thesis can be divided in two parts. The first part consists of theoretical work studying in detail how quantitative genetics predictions and models behave under complex and nonlinear GPMs that arise from development. For this, I performed large-scale evolutionary simulations using a realistic representation of the GPM given by a computational model of tooth development. Using this set of simulations, I first studied how well the linear approximation of quantitative genetics is able to predict the response to selection. I found that predictions using the linear approximation can be biased when the GPM is nonlinear. In other words, there can be a significant part of the response to selection that is missed by the linear approximation. Using the same set of simulations, I studied how the linear

approximation evolves in a complex GPM. I found that the evolutionary dynamics of the linear approximation are highly dependent on the GPM, and differ substantially from what is expected for a linear GPM. These dynamics are not purely stochastic, but rather deterministic ways in which the linear approximation changes as a reflection of the curvature of the GPM. In the second part of my thesis, I use the insight obtained in the first part to develop an applied method that uses techniques of quantitative genetics, combined with insight from evo-devo, to provide better predictions of evolutionary response to selection. The method uses a Kalman filter to combine information from selection acting on each generation, with information from the evolutionary time-series. I test the method with the simulated dataset using the model of tooth development, and also with an artificial selection experiment on the wing of the fruit fly, where 16 000 flies were measured. The new method is able to improve predictions and is a promising path to combine knowledge from both fields studying the GPM. In this way, the work in this thesis shows that rather than being a nuisance, the nonlinear nature of the GPM contains information that can improve our understanding of evolutionary processes.

Acknowledgements

Completing a PhD is not an easy task, and I would not have been able to do it without the help of many people. I can not put into words my gratitude to everyone who has been a part of my life in the last years and that has, directly or indirectly, contributed to the completion of this thesis.

The work presented here was carried out at the Institute of Biotechnology, University of Helsinki, HiLife with the financial support from the Integrative Life Science (ILS) doctoral programme, the Academy of Finland, the Finnish Cultural Foundation and the Center of Excellence in Experimental and Computational Developmental Biology. I thank my supervisor Isaac Salazar-Ciudad for sharing his passion for development and evolution and for allowing me to explore scientific questions with freedom. I want to express my deep gratitude to Jukka Jernvall not only for being a scientific inspiration, but also for being a pivotal institutional support during the years working on my thesis. Thank you Jukka for always finding the time to help me and for always making me feel at home with your lab. I would also like to thank my thesis committee members Juha Merilä and Heikki Helanterä for their feedback and support in every step of this journey. I thank the pre-examiners P. David Polly and Jon Brommer for their insightful reviews of the thesis. I also thank Kjetil L. Voje for agreeing to be my Opponent.

I am thankful to Niko Björkell, Shuja Wahid, Ardit Fejzullahi, and Sara Goncalves for their help carrying out the fruit fly experiments, and for making the measuring of 45 000 wings as fun as possible. I am also thankful to Osamu Shimmi and his lab, Heini Lassila and the Hi-Fly core facility for invaluable technical support to carry out the experiments. I further thank the CSC for providing the computational power required for the simulations included in this thesis. I thank Asko Mäki-Tanila for his predisposition and advice on my quantitative-genetic questions.

I would like to thank the members of the Salazar-Ciudad group, all very smart people with whom it has been a pleasure to discuss ideas with. Thank you Pascal Hagolani for becoming such a close friend, an excellent person to

discuss science and otherwise, Jhon Alves, for all the times we shared at the beginning of my PhD, and Renske Vroomans, for your positive outlook and always-insightful comments. Thank you Miguel, Miquel, Irepan, Roland, Carlos, Hugo and Aleksa. I thank the members of the Jernvall group. Thank you Mona Christensen for your openness and for your readiness to answer my questions about tooth development at any time of the day, Vilma Väänänen, for our amazing swimming-and-therapy sessions at Töölö urheilutalo, Outi Hallikas, for being an all-around great person, Susanna Sova and Aida Kaffash Hoshiar, for the fun conversations and for always helping me out in times of distress. Thank you Yoland, Rishi, Ona, Mia, Otto, Ana-Marija, Yue.

I am grateful to other members of the research community at the University of Helsinki who have shared this experience with me. Thank you Jacqueline Moustakas-Verho, for being a great friend, for your career advice, and for pushing for diversity inside and outside the University. Thank you Isabel Mogollón for sharing both fun and stressful times during these years. Thank you Anamaria Balic for your assertiveness, career advice and help at various points in this journey. Thank you Alison, Ewelina, Otto M. and Marja Mikkola.

I have been lucky enough to meet amazing people outside of the University. These are the people that have suffered my neurosis the most, who have helped me survive the Finnish winters and have magical times. I thank my dear friends from Oulu, particularly Johannes and Polina, for making me feel so welcome already on my third day after arriving in Finland. I thank the Mechelininkatu family, particularly Matti, Aapo and Markus for the fun times living together all those years. I thank my closest friends, Boodi, David, Juls, Kari, Meli, Noe, Tati, Vili and Vilkku, for their unconditional support, for hanging with me in the most stressful times, and for all the buffoonery. Thank you Ville for your loving support in the last stretch of the PhD, which was maybe the hardest, and for bringing a refreshing new perspective to my life.

Finally, my biggest thanks go to my parents Ruben and Maria, and my sister Florencia. Thank you for your unconditional love and support in every way imaginable, I could not have completed this thesis without you. Thank you dad for sharing your excitement for science, and for always being up to discuss dynamical systems with me.

Helsinki, September 2022
Lisandro Milocco

List of publications

This thesis includes the following original publications which are referred to in the text by their Roman numerals.

I **Milocco, L.***, & Salazar-Ciudad, I.* (2020). Is evolution predictable? Quantitative genetics under complex genotype-phenotype maps. Evolution, 74(2), 230-244.

II **Milocco, L.***, & Salazar-Ciudad, I.* (2022). The evolution of the G-matrix under nonlinear genotype-phenotype maps. The American Naturalist, 199(3), 420-435.

III **Milocco, L.***, Salazar-Ciudad, I.* (2022). A method to predict the response to directional selection using a Kalman filter. Proceedings of the National Academy of Sciences, 119(28), e2117916119.

* Corresponding author

Author's contributions:

I LM and ISC were associated in the conceptualization, funding acquisition and writing (review & editing). LM wrote the original draft, developed the software, performed the data analysis and visualization.

II LM and ISC were associated in the conceptualization, funding acquisition and writing (review & editing). LM wrote the original draft, performed the data analysis and visualization.

III LM conceptualized and developed the method, with input from ISC. LM carried out the experiments and wrote the original draft. LM and ISC designed the artificial selection experiments and participated in the editing of the manuscript.

Contents

Chapter 1

Introduction

Darwin's most revolutionary contribution was, arguably, to propose a theory of evolution that is variational (Levins and Lewontin 1985). Indeed, previous theories of evolution were transformational, as they proposed that changes in populations occurred as the result of each individual in the population experiencing change in the same direction. In Lamarck's theory of evolution, for example, the length of the giraffe's neck is the result of each individual giraffe attempting to stretch their neck to reach the top of the trees. In contrast, in Darwin's variational theory of evolution, each individual of the population differs from each other in some properties, and the system evolves by changing the distribution of these different types, with new types possibly arising. For giraffes, that would mean that there is individuals with necks of different lengths in a population, and that the distribution of neck length changes in time towards longer necks. Thus, evolution is the conversion of variation among individuals in a population into variation between groups in time and space. Without variation, then, there is no biological evolution.

Despite the central role that phenotypic variation plays in biological evolution, the question of *how* phenotypic variation arises in a population has not been the focus of dominant evolutionary theory for the last century. Indeed, during most of the 20th century, the dominant framework to study evolution was the Modern Synthesis (Huxley 1942, Mayr 1963, Beatty 1986). This paradigm resulted from the integration of Darwinian natural selection, Mendelian inheritance and population-level mathematical models. An underlying assumption of the Modern Synthesis is that the process that generates phenotypes, known generally as development, can be treated as an unknown black box (Pigliucci 2010, Hall 2012), which ultimately maps random genetic changes due to mutations and recombination into small, gradual changes in phenotype. In this way, understanding devel-

opment is inconsequential to understanding evolution, as it merely serves the role of a "change of scale" from genetic to phenotypic. This conceptualization allows to describe evolution entirely in terms of changes at the level of genes. These ideas are formalized in the statistical branch of the Modern Synthesis, known as quantitative genetics, and in the closely related field of population genetics.

Quantitative genetics describes phenotypic variation in a population in terms of linear statistical models. The field was largely developed in the context of animal and plant breeding sciences to solve applied problems. The field provides methods to describe variation in a population, to measure selection acting on a population, and even to predict how the mean of a set of traits in a population will change in time as a result of selection.

During the second half of the 20th century, a growing body of work known as evolutionary-developmental biology or evo-devo started to reveal details of the process of development and its relationship with evolutionary dynamics. A main result from the field of evo-devo is that, because phenotypes are constructed through the process of development, it is development that determines what phenotypic variation arises (Alberch 1991). Development, then, plays a critical role in determining evolutionary dynamics, as evolution can only occur in the directions were phenotypic variation is possible. Development is a complex dynamical processes with many parts interacting with each other at different levels of biological organization, including genes, signalling molecules, cells, tissues and environment. This results in the relationship between genotypes and phenotypes to be complex and show features such as nonlinearity, which can be missed by linear statistical models.

An apparent conflict arises from the ways quantitative genetics and evo-devo approach the study of the evolution of phenotypes. The first is based on the assumptions that development can be treated a black box and that the relationship between genotypes and phenotypes can be approximated with linear, statistical models for practical applications. On the other hand, evo-devo suggests that development is more than a simple transformation of small genetic changes to small phenotypic changes, resulting in a complex and nonlinear relationship between genotypes and phenotypes with certain features that cannot be captured by linear models.

This thesis deal with the conflict mentioned above. In the following sections I will briefly describe the relevant conceptual underpinnings of quantitative genetics and evo-devo, which were somewhat caricatured in this brief introduction. Then I will discuss the important differences between the two fields, specifically in terms of how they describe the relationship between

genetic and phenotypic variation. Finally, I give the specific objectives of this thesis.

1.1 Quantitative genetics

Quantitative genetics is a statistical branch of genetics that studies the evolution and inheritance of characters that show continuous variation, such as height. The approach is based on the idea that continuous characters are affected by many genes, as well as nongenetic environmental factors. Much of the development in the field has been done in the context of animal and plant breeding, and pushed by the need to solve practical problems like how to design breeding schemes to increase traits of economic interest.

The backbone of quantitative genetics is the concept of the additive effect of an allele. For an allele B_1, and assuming random mating, this effect can be intuitively understood as the average deviation of an individual having allele B_1 with respect to the population mean. The focus on additive effects is typically justified by the fact that, in sexually-reproducing organisms, offspring inherit only one allele at each locus from each parent, and not complete genotypes. Then, the additive genetic effect of an allele attempts to give a quantitative measure of the expected phenotypic effect associated to that allele, which allows to make predictions about the next generation.

Additive effects can be calculated using linear regression, the methodological powerhouse of quantitative genetics. For a diallelic locus, the additive genetic effects of the two alleles B_1 and B_2 in a population can be estimated with the following linear regression (Lynch and Walsh 1998 Ch. 4):

$$\mathcal{G}_{jk} = \mu_{\mathcal{G}} + \alpha_1 N_{1jk} + \alpha_2 N_{2jk} + \delta_{jk}, \tag{1.1}$$

where $\mathcal{G}_{jk}$ is the *genotypic value* of genotype B_jB_k, defined as the expected phenotype for genotype B_jB_k; $\mu_{\mathcal{G}}$ is the mean genotypic value in the population; N_{1jk} and N_{2jk} are the *gene content*, defined as the number of copies of alleles B_1 and B_2 respectively (i.e. N_{1jk} is 0, 1 or 2, and N_{2jk} is 2, 1, 0 for genotypes B_2B_2, B_1B_2/B_2B_1, B_1B_1, respectively); δ_{jk} are the residuals of the regression; and, finally, the slopes α_1 and α_2 are the additive genetic effects of alleles B_1 and B_2 respectively.

There are two important features of equation (1.1) which are common to all statistical models in quantitative genetics. First, all terms on both sides are expressed in units of phenotype. That is, even though the name *genotypic value* suggests a measure of genetic variation, it is in units of phenotype. A second important feature of equation (1.1) is that it is defined

in terms of deviation from the mean of the population, highlighting the fact that it is a local description, dependent on genotypic and environmental frequencies of a given population at a given time. This means that the additive values α_1 and α_2 depend on population frequencies.

The residuals δ_{jk} of the regression (1.1) are called the dominance effects, as they describe the deviation of genotypic value $\mathcal{G}_{jk}$ from the one expected by total additivity. Indeed, if the residual δ_{jk} is zero then $\mathcal{G}_{jk}$ is simply the sum of additive effects and the population mean. Note that this definition of dominance as deviations from additivity is statistical and depends on allelic frequencies. This statistical definition of dominance is not equivalent to the genetic effect known as dominance, where the phenotypic value of the heterozygote is not midway between the phenotypic values of the two homozygotes, a definition that is frequency-independent. Some authors have proposed to define this latter type of dominance as "physiological" (Cheverud and Routman 1995) or "functional" (Hansen 2006, 2013), to differentiate it from statistical dominance.

It is important to note that there is no one-to-one relationship between functional and statistical dominance, meaning that the presence of functional dominance does not translate exclusively as statistical dominance. Indeed, functional dominance affects not only the δ_{jk} terms of equation (1.1) and the variance that they explain, but it also affects the additive genetic effects α_1 and α_2 and the variance that they explain. This occurs because functional dominance affects the phenotypes associated with the different allelic combinations, and will therefore affect the slope of the regression. An immediate consequence of this is that one cannot asses the causal importance of functional dominance by measuring the size of statistical dominance (Cheverud and Routman 1995, Huang and Mackay 2016), since statistical dominance can be low but functional dominance can still have a strong effect on the additive values.

The genotypic value $\mathcal{G}_{jk}$ of each locus given by equation (1.1) can be added together across loci to obtain the total genotypic value of an individual, defined as the expected phenotype for a given genotype. If there are interactions among different loci, additional interaction terms can be included in the linear regression. These terms are known as statistical epistasis and they capture the deviation from additivity for pairs of alleles at different loci, much like statistical dominance captures the deviation from sum of alleles in one loci.

Statistical epistasis, just like its dominance counterpart, is a statistical abstraction representing deviation from additivity that is compatible with the linear models of quantitative genetics. It is different to the functional

definition of epistasis, where the phenotypic effect of an allelic substitution depends on alleles in other loci. As explained before for dominance, the size of statistical epistasis does not reflect the causal importance of functional epistasis in the phenotype, and functional epistatic interactions influence additive, dominance and epistatic statistical components (Hansen 2013, Huang and Mackay 2014).

The linear models in quantitative genetics treat dominance and epistasis as deviations from additivity. Indeed, these effects are collectively known as *non*additive genetic effects. This parametrization minimizes the variance in phenotypes explained by genetic interactions (Lynch and Walsh 1998, Hansen 2006, Nelson et al. 2013, Huang and Mackay 2016). This occurs because the parametrization is centered around additive effects, and linear regression by least squares then maximizes the amount of variance explained by additive effects (Lynch and Walsh 1998). Then, only marginal variance is left to be explained by nonadditive genetic effects. Huang and Mackay 2016 show that alternative parametrizations centered around dominance (or epistasis) result in most variation being statistically explained by dominance (or epistatic) statistical effects. This ultimately occurs because, as explained above for dominance, there is no one-to-one correspondence between the functional genetic effects of additivity, dominance and epistasis, and their statistical counterparts. This is the case because genetic effects are not independent of each other, so a linear model cannot disentangle them.

In applied quantitative genetics, and as mentioned above, the focus is on the additive effects since these values can be used for applied purposed such as predicting the response to selection as will be explained below. In the model given in equation (1.1) we have the additive genetic effects α_1 and α_2 for two alleles in a given loci. The sum of the additive genetic effects of all alleles at all loci for a given individuals is known as the breeding value of the individual. The variance of the breeding values in a population is known as additive genetic variance, V_A.

V_A is a central concept in quantitative genetics. It represents the part of the total phenotypic variance (V_P) that can be used for prediction of the phenotypes in the next generations, since it is associated with the additive genetic effects. An important feature of V_A is that it can be estimated from the resemblance between relatives. The idea is that different relatives share different number of genes through inheritance, and this will result in different degrees of phenotypic similarity (Lynch and Walsh 1998). For example, assuming (as it is typically done) that epistasis terms are negligible, V_A can be estimated as twice the covariance of parent-offspring trait value.

The ratio V_A/V_P is known as (narrow-sense) heritability and it plays an

central role in the breeder's equation (Lush 1937, Hill 2014), which allows to predict the change in the mean of a trait in a population ($\bar{z}$) as

$$\Delta\bar{z} = \frac{V_A}{V_P}s = h^2 s, \tag{1.2}$$

where $\Delta\bar{z}$ is the change in the mean of the trait from one generation to the next, h^2 is the heritability and s is the selection differential acting on that trait, measured as the covariance between the trait and relative fitness (Lynch and Walsh 2018, Falconer and Mackay 1996).

The breeder's equation is a pillar of quantitative-genetic theory, and its influence in animal and plant breeding cannot be understated (Hill 2010, Hill and Bunger 2004, Hill and Kirckpatrick 2010).

The derivation of equation (1.2) follows from Price's theorem under the assumption that there is no change in trait mean without selection and that the joint distribution of parent and offspring is bivariate normal (Rice 2004, Walsh and Lynch 2018). Indeed, the assumption of multivariate normality is central to much of quantitative genetics (Ch. 2 Lynch and Walsh 1998, Rice 2004). This is the case because this distribution has several qualities that make it convenient for mathematical analysis. For example, assuming joint bivariate normality between two variables allows to characterize the relationship between them with a linear regression plus noise. Moreover, the distribution is symmetrical and entirely defined by its mean and variance.

Evolutionary biologists recognized the potential of the breeder's equation to predict the response to selection in natural settings. Natural selection, however, acts on several traits, and evolutionary biologist are in general interested in studying all of them. To this end, Russell Lande (Lande 1979, Arnold and Lande 1983) developed the multivariate version of the breeder's equation under the assumption of joint multivariate normal distribution for all traits between parents and offspring. The multivariate breeder's equation is given by,

$$\Delta\bar{\mathbf{z}} = GP^{-1}\mathbf{s}, \tag{1.3}$$

where $\Delta\bar{\mathbf{z}}$ is the vector of changes in the mean and $\mathbf{s}$ is the multivariate selection differential. G and P are the additive genetic and phenotypic variance-covariance matrices, respectively. These are square and symmetric matrices, with diagonal elements equal to the additive genetic (or phenotypic) variances of the traits, and offdiagonal elements equal to the additive genetic (or phenotypic) covariances between pairs of traits.

The analogy between equations (1.3) and (1.2) is clear, with G and P being the multivariate extensions of V_A and V_P. The G-matrix, just like its univariate counterpart V_A, can be estimated using phenotypic data of

known relatives. Statistical requirements are, however, much larger, since several elements have to be estimated (for n traits, G has $n(n+1)/2$ distinct elements to be estimated). This has stimulated the development and application of several statistical techniques to estimate G. The most important of these is the "animal model", a special type of linear mixed models that allows to incorporate information from complex pedigrees (Lynch and Walsh 1998, Kruuk 2004).

There is surprisingly not many studies that directly address the question of how well the multivariate breeder's equation (1.3) predicts the response to selection. Some studies have, however, reported inconsistencies between the predictions and the observed change in trait mean (e.g. Sheridan and Barker 1974, Campo and Raya 1986, Roff 2007, Pujol et al 2018, Shaw 2019, Pélabon et al. 2021). Ultimately, errors in the prediction of the breeder's equation must arise due to violations in the assumptions underlying the equation. These include the parent-offspring regression being nonlinear (Rice 2012, Walsh and Lynch 2018), but also other sources of error that are more technical rather than conceptual, such as using wrong estimates of G or $\mathbf{s}$, or selecting a set of traits that do not capture all the selection acting on the population (Walsh and Lynch 2018).

The multivariate breeder's equation is commonly found in the literature in the form

$$\Delta\bar{\mathbf{z}} = G\boldsymbol{\beta}, \tag{1.4}$$

where $\boldsymbol{\beta} = P^{-1}\boldsymbol{s}$ is known as the selection gradient. $\boldsymbol{\beta}_l$ represents the change in relative fitness given a one unit change in trait l while holding all other traits constant (Hill and Kirckpatrick 2010, Lande and Arnold 1983, Walsh and Lynch 2018).

Equation (1.4) predicts that the change in the mean of a set of traits will be the compromise between what selection favors, given by $\boldsymbol{\beta}$, and what direction has more additive genetic variation, given by G (Maynard-Smith et al. 1985, Schluter 1996). In this way, the G-matrix takes a central role in the study of evolution in quantitative genetics, as it summarizes what directions can more readily respond to selection.

The study of G has become a subfield of its own within evolutionary quantitative genetics (Steppan et al 2002, Macguigan 2005) and the concept of G has shaped the way much research in evolutionary biology is carried out. A large amount of work has been devoted to comparing G matrices among populations (e.g. Cano et al. 2004, Arnold et al. 2008, Doroszuk et al. 2008, Eroukhmanoff and Svensson 2011, Wood and Brodie 2015, Assis et al. 2016, Delahaie et al. 2017, Walter et al. 2018, Hangartner et al. 2020, Chakrabarty and Schielzeth 2020), study how G possibly constrains

evolution (e.g. Cheverud 1988, Arnold 1992, Schluter 1996, Hansen and Houle 2008, Kirkpatrick 2009, Eroukhmanoff 2009, Walsh and Blows 2009), retrospectively reconstruct selection (e.g. Merilä et al. 1994), or predict responses to selection (e.g. Lande and Arnold 1983, Campo and Raya 1986, Roff 2007). This large body of empirical work studying and comparing G-matrices is a mixture of studies in wild and laboratory populations, the latter using artificial selection where the experimenter imposes the selective pressure.

It is important to highlight that the G-matrix, just like V_A, is a statistical abstraction that depends both on the genetic basis and development underlying the traits, and on the distribution of genotypes and environments in a given population. Therefore G has some features that should be stressed.

The first important feature of G is that it cannot be used to infer the underlying genetic architecture of a set of traits (i.e. the patterns of pleiotropy, dominance, and epistasis behind the set of traits). That is, there is no simple one-to-one relationship between underlying genetic architecture and G. This is tightly related to the discussion above about functional and statistical genetic effects, and exemplified by Houle 1991 who uses a trade off model where a limited amount of a resource is allocated into two life-history traits. In the model, two traits are determined by two processes: acquisition, which determines the total amount of limiting resource that an individual acquires, and allocation, which determines how the resource is allocated between the two traits. He assumes a simplified architecture where each loci affects either allocation or acquisition. One would expect that the presence of allocation would manifest as a negative genetic covariance between the two traits in question (i.e. the offdiagonal element of G). However, Houle 1991 finds that covariance can be positive depending on the relative number of loci that participate in acquisition and allocation. In this way, the sign of the covariance between traits cannot be used to infer the existence of a trade-off between the traits, given in this case by the allocation process. Other studies have further shown that multiple underlying genetic mechanisms can produce the same G-matrix (Gromko 1995, Pigliucci 2006, Chebib and Guillaume 2016).

Another fundamental feature of G is that it is local. That is, any estimate of G is specific to the particular population and a given generation. However, G is typically used to make inferences for longer time frames of several generations. The validity of this is tightly linked to the question of how fast G changes. There is no theoretical expectation of how G should change, except under strong simplifying assumptions such as the infinitesi-

mal model, which assumes infinitely many loci that contribute additively to the trait (Arnold et al. 2008, Walsh and Lynch 2018 Ch. 16). In this way, the question of how G changes with generations and environments has been treated empirically, with studies showing both that G can change rapidly (Cano et al. 2004, Doroszuk et al. 2008, Eroukhmanoff and Svensson 2011, Wood and Brodie 2015, Walter et al. 2018, Chakrabarty and Schielzeth 2020) or slowly (Delahaie et al. 2017, Hangartner et al. 2020, and others reviewed in Arnold et al. 2008). The only emergent conclusion from this body of work is that the change in G depends on the traits under study. In any case, it is clear that the locality of G introduces serious limitations in using G to explain evolution in more than one generation, and specially for macroevolutionary trends (Steppan 2002). Ultimately, because estimating G is costly and time-consuming, understanding how fast it can change is important since change reduces the predictive value of G (Eroukhmanoff 2009).

1.2 Evo-devo

Evolutionary-developmental biology, or evo–devo, is a field of research concerned with the two-way interaction between phenotypic change that happens between generations during evolution, and individual development that happens in each generation (Oster and Alberch 1982, Alberch 1982, 1991, Maynard Smith et. al 1985, Raff 1996, Müller 2007). In the most general sense, development can be understood as all the changes that occur in a given organism from its beginning as an egg to adulthood (or death, Oyama 2000, Gilbert and Barresi 2016). Development is then the process by which the phenotype of each individual is generated, and therefore the process that creates phenotypic variation in each generation. The concept of development can be applied to behavioral traits (Lickliter 2007, Hall 2013), but in this thesis and much of the evo-devo literature, I focus on morphological traits.

During development, genes produce gene products that can regulate the expression of other genes, or regulate cellular behaviours like cell division or migration, and cell mechanical properties, like elastic and shear modulus (Garcia and Garcia 2018). Together with generic biophysical properties such as surface tension (Newman and Muller 2000) and affected by random variations, these signals, behaviors and biomechanical properties determine how the developing tissue interacts with its local environment, leading to changes in its geometry. This changes in the spatial distribution of cells will themselves affect gene expression and the interaction between

cells by, for example, changing the relative distances between cells (and therefore limiting or stimulating cell-cell signaling) or by changing the mechanical state of cells. In this way, development is a complex dynamical process that occurs by a series of temporal and spatial interactions and feedbacks between different parts at different levels of biological organization (Alberch 1982, Newman and Comper 1990, Newman and Muller 2000, Salazar-Ciudad 2006).

It is useful to conceptualize development as composed of several developmental mechanisms. These mechanisms can be defined as the combination of a genetic network with a given topology and the set of signals, cell behaviors and mechanical properties that are regulated by the network to produce a given morphological transformation (Salazar-Ciudad 2006). In this way, genes and gene networks are an important part of developmental mechanisms, but do not act in isolation to produce a phenotype. Genes participate in generating phenotypes only by interacting with other genes and cell behaviors in a complex developmental mechanism. Therefore, genetic substitutions can only result in phenotypic changes by modifying the dynamical process of development. In this way, from the perspective of evo-devo, the effect of a gene cannot be understood outside of a specific developmental mechanism. Then, there is no inherit phenotypic effect than can be assigned to a genetic change.

Because development is the process that generates morphologies, any change in morphology must result from a change in development. Indeed, genetic variation is produced by mutations, recombination and other chromosomal rearrangements. However, these can only result in phenotypic variation through development.

Phenotypic variation produced by a given developmental mechanism can be described in a useful way by the concept of variational properties (Salazar-Ciudad 2006). These properties can be defined as the set of morphologies that can be generated by the developmental mechanism through small changes in the intensities of the interactions among genes, cell behaviors and mechanical properties that compose the developmental mechanism. In this way, the variational properties represent the set of possible morphologies that can be produced by the developmental mechanism and that can interact with, for example, natural selection during adaptive evolution.

The concept of variational properties describes phenotypic variability (i.e. the ability to produce variation) as opposed to statistical phenotypic variance which is a way to describe realized variation. Variational properties are operationally more complex to calculate than measuring variance, but they incorporate all the information on the possible variation generated by

development.

The concept of variational properties has been most successfully applied when using computational models of development. Indeed, if enough about the development of a given system is known, a computational model can be created to represent the mechanistic hypotheses in mathematical terms. These models require initial conditions and the network of interactions between genes, cellular behaviors and mechanics that make the developmental mechanism. By modifying the parameters of that model, the variational properties of the developmental mechanism can be simulated. A paradigmatic example that is particularly important for this thesis is the model of tooth development that has been used to study natural variation in seals (Salazar-Ciudad and Jernvall 2010). The computational model summarizes tooth development by incorporation a spatial context for cells, their mechanics, and a gene network that regulates the cell behaviours of proliferation and differentiation. The dynamics of the model are determined by a set of parameters. Variation in those parameters is able to reproduce the morphological variation found in a natural population, and allows to infer what changes at the level of parameters can explain morphological changes seen in the population.

Variational properties can help understand how development introduces directionality to the evolutionary process. Indeed, evolution can only proceed in the directions where development produces phenotypic variation, as natural selection can only act on existing phenotypic variation (Alberch 1982, Mayr 1982). In this way, it is both development and natural selection that determine the direction of morphological evolution.

Historically, evo-devo has focused on studying patterns of conservation and change over relatively large evolutionary timescales, such as between species or larger taxonomic groups (Nunes et al 2013). The paradigmatic example of this is the discovery that the developmental "toolkit" of genes is mostly conserved across distantly related taxa and that phenotypic change across such broad scales is often accompanied by spatial or temporal changes in the expression of these conserved genes (Davidson 2001, Wilkins 2002, Carroll et al. 2013). This body of work shows how rewiring an already-existing developmental mechanism can result in large phenotypic changes and allow for evolutionary divergence. On this line, there is also a large body of work describing the genetic and developmental bases of the differences in morphology between several species (e.g. Stern 1998, Loehlin and Werren 2012, Mallarino et. al 2012, Arif et. al 2013).

Most work in evo-devo has focused on macroevolutionary trends because questions in evo-devo require detailed understanding of development, which

in turn can only be achieved by the accumulation of data from experiments in developmental biology. These experiments are largely carried out in a few species known as "model species" which include mice and fruit fly. The reason for this is that these animals are compatible with laboratory conditions, and much technological development has occurred to ease experimentation with these animals. In this way, the development of only a few organisms is understood to enough degree of detail and therefore comparisons are made between these evolutionarily distant organisms.

Evo-devo at the population level is in its infancy (Nunes et al 2013). A main reason for this is that developmental mechanisms are for the most part not understood to the level required to make predictions of how subtle changes in development give rise to subtle, quantitative variation observed in populations (Nunes et al., 2013, Parsons & Albertson, 2013). In this way, there are no quantitative measures of how the theory described in this section interacts with population-level phenomena such as natural selection.

1.3 The genotype-phenotype map (GPM) and the conflict

A key conceptual difference between the Modern Synthesis, embodied in quantitative genetics, and evo-devo, is how they treat the relationship between genotypes and phenotypes. In this section, I will briefly summarize the ways the two fields conceptualize this relationship and show how these conceptualizations can be conflictive with each other. This will lead to the research questions of this thesis.

From the perspective of evo-devo, genes are part of developmental mechanisms that jointly determine phenotypes, together with biomechanics, noise, interactions with the environment, cellular behaviors and signalling. In this way, development is a complex system that is composed of many parts interacting with each other. Alberch 1991 uses the metaphor of the genotype-phenotype map (GPM) to summarize the complex process of development. In his conceptualization, Alberch proposes the following equation

$$\frac{d\boldsymbol{\Phi}}{dt} = \boldsymbol{f}(\boldsymbol{\Phi}, \boldsymbol{\theta}_i), \tag{1.5}$$

where $\boldsymbol{\Phi}$ is the phenotype, $d\boldsymbol{\Phi}/dt$ is the change in time of the phenotype during development, $\boldsymbol{\theta}_i$ is a set of developmental parameters capturing the dynamics of the developmental process and $\boldsymbol{f}$ is a developmental function. Examples of developmental parameters are the diffusion rates of signaling molecules and the strength of regulatory interactions between genes.

Equation (1.5) says that the change in the phenotype at a given time is a function of the phenotype at that time, and the developmental parameters. This conceptualization allows to build a mapping from initial condition for the phenotype and a set of developmental parameters, to a final phenotype. Because developmental parameters are, at least in part, determined by the genes, this results is a genotype-phenotype map: a function that assigns a phenotype to each genotype.

Because development is a complex dynamical process that can, at least in principle, be expressed with differential equations of the type given in expression (1.5), the resulting GPM possesses distinct features (Alberch 1991). One key property of the GPMs associated with development is nonlinearity, meaning that the same input (i.e. change in genotype or developmental parameter) results in different outputs (i.e. changes in phenotype) depending on the state or context. In practical terms, this means that a small perturbation may cause a large effect, a proportional effect, or even no effect at all. In linear systems, the effect is *always* directly proportional to cause.

In quantitative genetics, the complex mapping of genotypes to phenotypes is approximated with a statistical construct that maps the mean of the genotypes to mean of the phenotypes (Feldman and Lewontin 1975, see section *1.1 Quantitative genetics*). This local linear description is given in terms of additive genetic effects, as shown in equation (1.1), and summarized by the concepts of additive genetic variance and the G-matrix. This is the best local linear description possible, in the sense of least-squares, and it has sometimes been compared to first-order approximation of a Taylor series around the mean (Lewontin 1974, Hansen 2013). Assuming that the GPM is smooth, the approximation will be arbitrarily good, in terms of residuals, for *some* neighborhood around the mean. How small that neighborhood is will depend on how good we want the approximation to be (it is tautological on the mean) and on the GPM.

Importantly, what I here call *linearity of the GPM* is analogous to the *additivity of genetic effects* explained in the section *1.1 Quantitative genetics*. That is, assuming that genes act additively and assuming a linear GPM are equivalent. The concept of additivity is far more common in the quantitative genetics literature, but in this section I use the concept of linearity as it is framed in terms of the GPM.

An important feature of the linear approximation of quantitative genetics is that it does not require knowledge of the underlying development that generates phenotypes. Indeed, even if the GPM is not linear, one can estimate the best linear approximation around the mean. However, as mentioned above, this description is purely local and, strictly speaking,

nothing can be said about the suitability of the description outside the specific conditions in which it was calculated. Moreover, as explained before, it is important to remember that the statistical description does not reveal anything about the underlying biology that generates a phenotype, as discussed in section *1.1 Quantitative genetics*. The statistical description is, thus, a black box that admittedly does not give causal explanation for phenotypes, but that can be useful for certain applications, such as predicting the response to selection in the short term.

The linear approximation of the GPM given in quantitative genetics can be more than a local description only if we assume that the *real* GPM is linear. In this case, the statistical abstraction fully describes the map, and the local description is a global description. It is not surprising, then, that much (but certainly not all) of quantitative genetics assumes implicitly or explicitly that the *real* GPM is linear (e.g. Hill 2010). The assumption of linearity is sometimes justified by the thought that in the absence of any evidence, the linear function is the simplest hypothesis, so assuming a linear map follows by Occam's razor (Lewontin 1974). However, from what we know from molecular and developmental biology, a linear map is an unlikely option and no known developmental process results in a linear GPM. The assumption that the *real* GPM is linear is then wrong, and in direct conflict with what we know from evo-devo.

Without assuming that the *real* GPM is linear, and taking the statistical description as a local approximation, some researches propose that the framework is useful for some time, as long as one remains around the population mean. Note that here I refer to the framework as "useful" for a given objective, such as predicting the response to selection. This is indeed the metric used in quantitative genetics, a historically applied research field, where many assumptions are justified because "they work" for a given purpose (e.g. Hill 2014). This transforms a qualitative question into a quantitative one: how close to the mean should we be for the local description based on additive effects to be useful? This question is tightly linked to the question of how does the linear description changes with generations. Indeed, if the local description changes slowly, then for all practical reasons we can rely on it for several generations without fear of severe mistakes. These two question are central to my thesis, and are addressed explicitly in Publications I and II. To be clear, the questions are:

Q1. How good is the local statistical description when the genotype-phenotype map is nonlinear?

Q2. How does the local statistical description change in time when the genotype-phenotype map is nonlinear?

Despite the importance of these question, there is a scarcity of work in evolutionary biology directly addressing them. I will briefly go through the most relevant works below.

Question 1 can be addressed by comparing the predictions of the breeder's equation with the observed change in the mean of the traits under study. Indeed this is a practical measure of usefulness of the framework: if predictions fail, this indicates inadequacy of the framework. To associate prediction error with the nonlinearity of the GPM, however, all other sources of error must be controlled for. These sources include using wrong estimates of G, P or $\boldsymbol{s}$, random noise, or having an incomplete picture of the traits under selection. As explained in section *1.1 Quantitative genetics*, the emerging picture from empirical work is that for multivariate traits, responses inconsistent with the breeder's equation are common. However, the sources of these errors are rarely identified, and the possibility that errors may be associated with the nonlinearity of the GPM is rarely explored.

From the theoretical side, Sean Rice's framework to study evolution (Rice 2002, 2004) shows that the breeder's equation is not sufficient to predict the response to selection for certain GPMs. Rice's mathematical framework is based on a complete description of the GPM using a Taylor series, as an infinite sum of terms that are expressed in terms of the map's partial derivatives with respect to underlying developmental parameters, evaluated at a single point. Further, he describes the (arbitrary) distribution of underlying parameters in terms of its statistical moments. He shows that if the GPM is nonlinear or if the underlying parameters do not have a multivariate normal distribution, the change in the mean of a set of traits is not accurately represented by the breeder's equation. Carter et al. 2005 arrive to a similar conclusion using the multilinear model (Hansen and Wagner 2001), which is less general than Rice's framework but allows for certain types of nonlinear GPMs while remaining mathematically tractable. These works show that deviations from the breeder's equation can happen due to the nonlinear nature of the GPM, but do not directly address the questions of how large and common we should expect these deviations to be, particularly for the type of GPMs associated with the development of complex phenotypes described earlier in this section.

Question 2 can be addressed directly by studying the evolution of the G-matrix, which summarizes the statistical local description of quantitative genetics. From the empirical side, as explained in section *1.1 Quantitative genetics*, there is a large body of work comparing G-matrices, but there is no clear picture of how we should expect G to evolve. Moreover, this body of empirical work does not focus on the GPM, and empirical comparisons

are affected by multiple confounding effects which are rarely disentangled. From the theoretical side, virtually all work studying the evolution of G has been done under strong assumptions, including linearity of the GPM (Turelli 1985, Slatkin and Frank 1990, Reeve 2000, Jones et al. 2003, 2012). Under this assumption, the general consensus is that the shape of G is relatively stable and ultimately determined by selection, with G fluctuating randomly in each generation around this expected shape.

The situation is, then, that Questions 1 and 2 remain open and this gap is the starting point for this thesis. Publications I and II answer Questions 1 and 2 respectively. For this, the model of tooth development briefly introduced in section *1.2 Evo-devo* was used to build large set of evolutionary simulations. The model provides a realistic representation of the developmental process involved in tooth development, and therefore a realistic GPM. Simulations allow to obtain enough data to build the best linear models at each time point, which is the best local description from quantitative genetics possible without statistical limitations. The results from the first two publications provided insight to build a novel method that improves the predictions of the breeder's equation when the GPM is nonlinear, but also when other assumptions of the breeder's equation are violated, a common occurrence as explained in section *1.1 Quantitative genetics*. The novel prediction method combines predictions of the breeder's equation with past records of the mean. The method is developed and applied to data of a 20-generation artificial selection experiment in the wing of the fruit fly in Publication III.

1.4 Aims

The general aim of the thesis is to advance in the understanding of how the genotype-phenotype map (GPM) affects evolutionary dynamics. The specific aims of the thesis are:

1. Study how well the breeder's equation predicts the response to selection for a set of traits associated to a complex GPM, using tooth development as a case study.

2. Study how the G-matrix and other relevant statistics of quantitative genetics evolve for a set of traits associated to a complex GPM, using tooth development as a case study.

3. Develop a method that improves the predictions of the breeder's equation using insight from the results of Aims 1 and 2. Test the method

in simulated data using a complex GPM, and on experimental data that includes other possible sources of prediction errors.

Chapter 2

Methods

In this chapter I will briefly explain the main methods used in the thesis. Details are given in the *Methods* sections of the original publications included in this thesis.

2.1 The tooth development model

The complex genotype-phenotype map used for the evolutionary simulations in this thesis is generated using a model of tooth development (Salazar-Ciudad and Jernvall, 2010), schematically shown in Figure 2.1A. The model is a computational representation of the development of teeth that includes a spatial representation of cells, their mechanics, and a gene network that regulates cell behaviors such as proliferation and differentiation.

The tooth model allows to simulate the dynamics of tooth development. The simulation starts with a small, flat epithelium over a block of mesenchyme, closely resembling the earliest stages in tooth development (Figure 2.1A, Jernvall and Thesleff 2000). Cells produce signaling molecules that diffuse in the extracellular space. There are three of these signaling molecules. The first is an *activator*, which promotes its own synthesis and secretion on epithelial cells. It also induces the differentiation of epithelial cells into enamel knots, which are important signaling centers in tooth development (Jernvall and Thesleff 2000). These enamel knots do not divide, and produce the other two types of signaling molecules implemented in the model: an *inhibitor* that inhibits the secretion of the *activator* in the cells that receive it, preventing the formation of another enamel knot close to an already-existing one, and a *secondary signal* that induces cell differentiation, and in this way affects cell proliferation. The model also includes growth biases to represent the fact that the tooth grows at different rates

at the anterior and posterior borders, and generic forces of epithelial tissues like cell adhesion. As a result of model dynamics the epithelium starts growing downwards and soon the first enamel knot appears. As the tooth epithelium keeps growing, new knots appear and the growing epithelium between them becomes the valleys between the cusps of the forming tooth. Thus, signaling and induction are taking place on the tissue, at the same time as the tissue itself changes its shape, as shown in Figure 2.1B.

The model defines a number of parameters that quantify certain aspects of the cellular and signalling interactions, such as the growth rate of cells, mechanical properties of cells, diffusion rates of signals and strength of regulatory interactions between genes. There is a total of 21 of these developmental parameters, and different values of these parameters will change the dynamics of development during the simulation and thus alter the final phenotype generated (i.e. the final distribution of cells in three-dimensional space, Figure 2.1D). In this way, the values of the developmental parameters, together with an initial condition are enough to run the model and produce a final phenotype.

2.2 Evolutionary simulations

The tooth development model was embedded in a population model to simulate evolution on a complex GPM. Data from these evolutionary simulations where used in all the publications of the thesis. Details of the evolution model are given in the *Methods* section of Publication I. The algorithm of the evolutionary simulations is schematically shown in Figure 2.1C. In brief, the evolution model considers four steps per generation:

1. The mapping between genotypes and developmental parameters. Each individual has a diploid genotype with a fixed number of loci. Each allele in a loci has a specific quantitative value. These genetic values are added together to give the value of one of the 21 developmental parameters required to define the dynamics of tooth development (see section *The tooth development model*).

2. The mapping between developmental parameters and phenotypes. This is the tooth development model, which takes the value of the 21 developmental parameters of each individual as input and provides the 3D morphology of a tooth as an output. From the tooth morphology, five traits are measured. These are the coordinates of three landmarks located on the three tallest cusps of the tooth (see Figure

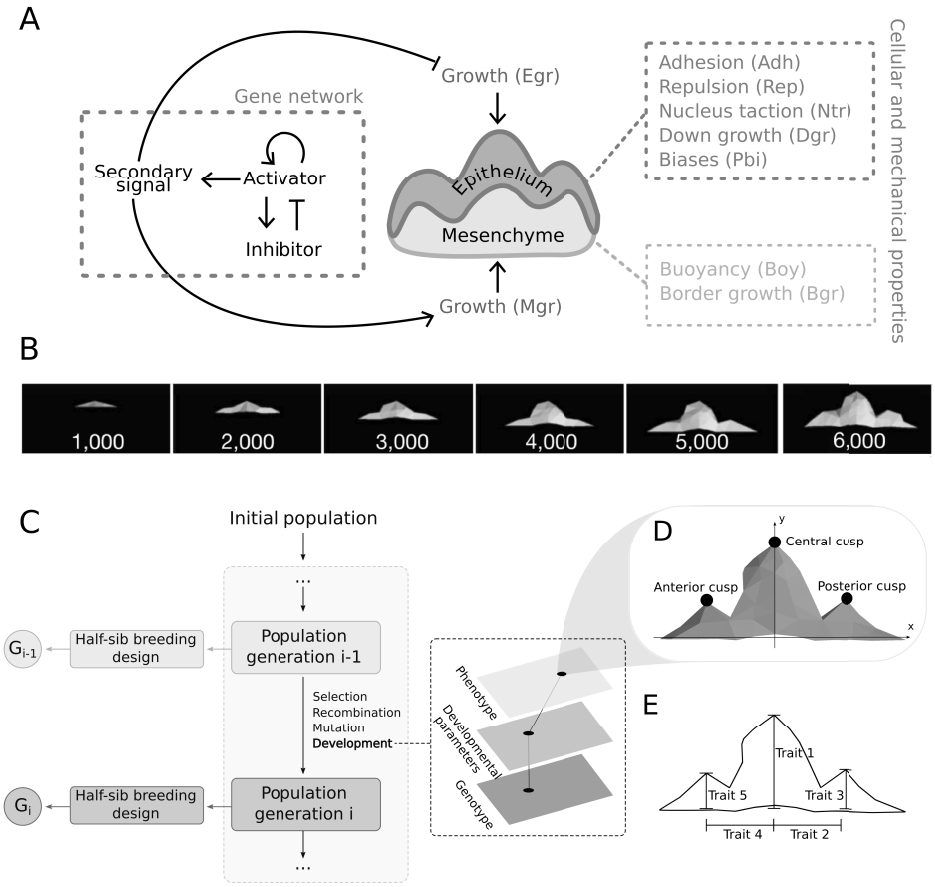

Figure 2.1: **A** shows a scheme of the tooth development model. The model explicitly simulates a sheet of epithelial cells, together with the genetic network and the set of signals, cell behaviors and mechanical properties that are regulated by the network during tooth development. **B** shows the development of a an example tooth, where the eruption of the cusps can be seen during 6000 iterations. **C** shows the algorithm of the evolutionary simulations. Each individual in the population is modeled explicitly and has a set of genetic values, which additively determine the values of the developmental parameters. These parameters are mapped to a phenotype using the tooth model. Selection, recombination and mutation are applied in each generation. The *G*-matrix is calculated in every generation. **D** shows an example of a phenotype produced by the developmental model, and the location of the three landmarks on the three tallest cusps of the tooth. **E** shows the five measured traits.

2.1D-E). Note that each individual has a set of genetic values, a set of developmental parameters, and a set of trait values.

3. Selection of parents for the next generation. Selection is implemented by keeping the 50% males and females with trait values closest to a predefined optimum. The optimum is defined at the beginning of each simulation, and each simulation has a different optimum. To generate the optima, each of the five morphological traits was chosen to increase or decrease in respect to the initial reference morphology. This leads to a total of $2^5 = 32$ trait combinations as optima.

4. Reproduction, with recombination and mutation on the genotypes. Once the parents are selected, they produce one gamete each by randomly selecting one of the two alleles for each loci with equal probability. The gametes of the parents then fuse to form the diploid genome of the offspring, with a probability of alleles being mutated by adding a normally-distributed random amount to it. Each parental couple generates two male and two female offspring for the next generation. This keeps the population size constant and results in all selected parents having exactly the same fitness.

By iterating steps 1 to 4 in each generation, we simulated how the genotypes and phenotypes of the population change over generations.

2.3 Artificial selection with the fruit fly

Artificial selection experiments on the wing of *Drosophila melanogaster* were carried out and used to test the method developed in Publication III. To build the starting population, 250 female flies were captured in Groningen, The Netherlands during the Summer of 2017 by the Billeter's lab. From each isofemale line, 25 virgin females and males were collected and merged to make a single large, outbred population that was maintained in laboratory conditions. From this large population, 400 females and 400 males were collected as virgins and randomly assigned to one of the four experimental lines (i.e. 100 males and 100 females in each line). One line was used as a control without selection, while the other three were subject to selection.

In each generation of the experiment, 100 male and 100 female virgins were collected. The automatic system known as the *wing machine* (Houle et al. 2003) was used to image the left wing of each anesthetized fly and locate the position of five landmarks, as shown in Figure 2.2. In the control line, 50 females and 50 males were randomly chosen to be parents of the next

generation. In the three lines with selection, the 50 females and 50 males with wings closest to the optimum morphology were selected as parents. The optimum morphology (see Figure 2.2D) was defined at the start of the experiment and was the same for the three lines.

Selected parents were paired randomly. Two males and two females were randomly collected from the offspring of each couple, resulting in the 100 males and 100 females of the next generation. The process of image processing and selection was repeated in each generation for a total of 20 generations. Siblings were not allowed to mate to reduce inbreeding. If some of the formed couples failed to produce offspring for the next generation, we measured more offspring from other couples to complete the 200 individuals per generation. We also formed three extra couples in each generation, to provide extra individuals in case some of the original 50 couples failed to produce offspring. Throughout the experiment the flies developed at 25°C.

As mentioned above, the phenotypic data in each generation are the x- and y-coordinates of five landmarks on the wing, resulting in a total of ten traits. However, because the data was aligned using Procrustes least squares superimposition (Houle et al. 2003), four degrees of freedom were lost (i.e. one to estimate wing size and three to standardize the orientation of wing shapes). In this way, six independent traits remain in the data.

2.4 Quantitative genetics

Estimation of various statistics of quantitative genetics is central to the thesis. Of most importance is the estimation of the additive genetic variances and covariances that make the G-matrix. As explained in the *Introduction*, additive genetic variances and covariances can be estimated from correlations between relatives. Classical methods include parent-offspring regression and analyses of variance (ANOVA) applied to data of full-sib and half-sib families (Lynch and Walsh 1998). However, the use of what is known as "animal model" has become the standard for the estimation of additive genetic variances and covariances (Kruuk 2004, Kruuk et al. 2008, Postma and Charmantier 2007), in large part because it can incorporate information from any type of relatives through complex and possibly incomplete pedigrees (Kruuk 2004). The animal model also has other desirable properties, such as providing unbiased estimates of variance components by any effects of finite population size, assortive mating, selection or inbreeding that may occur during generations included in the pedigree. This is therefore the method used in this thesis to estimate the G-matrix.

The animal model is a special type of linear mixed model, a linear re-

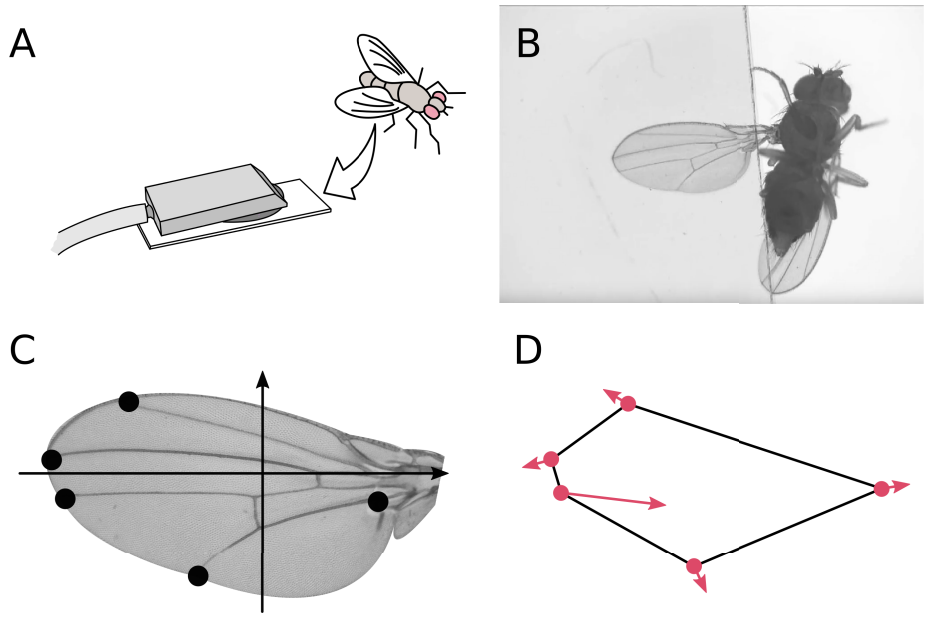

Figure 2.2: **A** shows a scheme of the *wing machine* device (Houle et al. 2003) used to measure the wings of anaesthetised fruit flies. The *wing machine* is connected to a pump that creates vacuum, sucking the wing of the fly into the space between two microscope slides and allowing for a picture of the extended wing to be taken using a microscope, as shown in **B**. **C** shows the five landmarks used in the wings. **D** shows the direction of selection (i.e. the optimum shape) for the five landmarks in the artificial selection experiments.

gression where the explanatory terms are a mixture of "fixed" and "random" effects. An effect is modeled as fixed if the different levels can be described as constants that change the mean of the distribution of phenotypes. A classic example of fixed effect is sex. An effect is modelled as random if the levels found in the data are a sample from a larger population, for which the analysis provides an estimate of the variance associated to the effect rather than a parameter for each factor level. Random effects therefore influence the variance of the trait (Pinheiro and Bates 2000). The distinct feature of the animal model is that it fits the breeding values of the individuals as random effects. This allows to obtain an estimate of the variance in breeding values which is defined as the additive genetic variance V_A in the univariate case, and the G-matrix in the multivariate case (see section *1.1 Quantitative genetics*). Note that a pedigree is essential for this, as it gives the expected covariation of breeding values between pairs of individuals.

A general animal model has the form

$$\boldsymbol{z} = X\boldsymbol{\beta} + Z\boldsymbol{u} + \boldsymbol{e}, \tag{2.1}$$

where $\boldsymbol{z}$ is the vector of phenotypes for all individuals, $\boldsymbol{\beta}$ is the vector of fixed effects, X is the design matrix relating the fixed effects to each individual, $\boldsymbol{u}$ is the vector of random effects, Z is the design matrix for random effects and $\boldsymbol{e}$ is a vector of residuals. As explained above, $\boldsymbol{u}$ must include at least the breeding values, but possibly also other random effects such as maternal effects. The simplest animal model possible then has X as a vector of 1's, $\boldsymbol{\beta} = \mu$ the population mean, $Z = I$ the identity matrix and $\boldsymbol{u}$ the vector of breeding values. This is the animal model used to analyze the data from the simulations using the tooth model. For the estimations of the artificial selection experiments in the fruit fly, additional fixed effects were included to account for sexual dimorphism and the fact that multiple people participated in measuring the wings.

With a model of the type given by equation (2.1), the record of phenotypes and a pedigree, additive genetic variance can be estimated using a maximum likelihood algorithm (Lynch and Walsh 1998). The algorithm works by finding the parameters that maximize the likelihood of observing the data. More specifically, a restricted maximum likelihood (REML) algorithm was used, which maximizes only the portion of the likelihood that does not depend on the fixed effects (Meyer 1989). The development of the equations of the REML algorithm can be found in Lynch and Walsh 1998 (Ch. 28). First, an expression of the log-likelihood of the parameters given the data is obtained by assuming that the random effects and the residuals have a multivariate normal distribution. Then, the partial derivatives of

this expression with respect to the parameters are obtained. The equations for the REML estimators are then obtained by setting these derivatives equal to zero. These equations have to be solved numerically through a recursive algorithm, since they are coupled and nonlinear. In this thesis, this was done using the software WOMBAT (Meyer 2007, Houle and Meyer 2015).

2.5 The Kalman filter

The method developed in Publication III to predict the response to selection uses the Kalman filter, which is a hallmark of control theory (Kalman 1960, Åström and Wittenmark 1997). The Kalman filter is a general algorithm to estimate the value of a set of n state variables of interest, represented at time i by the state vector $\boldsymbol{x}_i \in \mathbb{R}^n$. For this, we need a *transition model* that describes how we expect the states to change over time,

$$\boldsymbol{x}_{i+1} = A_i \boldsymbol{x}_i + \boldsymbol{w}_i, \tag{2.2}$$

where A_i is a squared matrix describing how we expect $\boldsymbol{x}_i$ to change and $\boldsymbol{w}_i$ is process noise vector. We also need an *observation model*, with relates the state vector to a vector of measured variables, $\boldsymbol{y}_i$,

$$\boldsymbol{y}_i = H_i \boldsymbol{x}_i + \boldsymbol{v}_i, \tag{2.3}$$

where H_k describes the relationship between the measurements and the states, and $\boldsymbol{v}_i$ is the measurement noise vector.

The Kalman filter works in two steps, prediction and correction. In the prediction step the states are projected forward in time from $i-1$ to i, using the transition equation (2.2). This gives the *a priori* state estimate at time i. It is called *a priori* because it has not been combined with the measurements at time, $\boldsymbol{y}_i$. This combination of information occurs during the correction step. The important thing here is that we have two independent sources of information about the underlying states at time i. On the one hand, we have an *a priori* estimate. On the other hand, we have the measurements of the system that we obtain at generation i, and that are related to the states through the observation model given by equation (2.3). What the Kalman filter does is a linear combination of these two sources of information to obtain an *a posteriori* estimates of the states at time i, symbolized $\hat{\boldsymbol{x}}_i$. The linear combination is determined by a weight matrix called the gain, symbolized K_i. This matrix is calculated as the trade off between the confidence we have on the *a priori* estimate of the states and the

confidence we have on our measurements at time i. If the measurements are to be trusted, then the gain will give more weight to the measurements. If the *a priori* state estimates are to be trusted, then the gain will assign more weight to it. The "trust" is quantified by the associated error covariance matrices of the *a priori* estimate and the measurements. More specifically, K_i is found by minimizing the error covariance matrix $E(\boldsymbol{e}_i \boldsymbol{e}_i^T)$, where $E()$ is the expected value, ^{T} is the transpose operator and $\boldsymbol{e}_i = \boldsymbol{x}_i - \hat{\boldsymbol{x}}_i$ is the *a posteriori* estimation error (Åström and Wittenmark 1997).

Assuming that the equations (2.2) and (2.3) are the true representation of the dynamical system, and that measurement noise is independent and gaussian stochastic process, the Kalman filter provides the optimal solution for estimation of the state vectors, in the sense that it is unbiased and of minimum variance (Åström and Wittenmark 1997). In Publication III of this thesis, we introduce a novel method to predict the response to selection that combines the Kalman filter with the breeder's equation.

Chapter 3

Main results and discussion

3.1 Local GPMs

A key feature of the evolutionary simulations in this thesis is that the genotype-phenotype map (GPM) emerges from explicitly simulating the development of each individual's phenotype in the population, using the tooth model. The tooth model is composed of coupled differential equations that define the change in time of position, gene expression, cellular behaviors and mechanical properties for each cell during development, until the final phenotype is obtained (see equation (1.5) and *Methods*). The equations use a set of developmental parameters, such as diffusion rates, that are additively determined by genes. These parameters specify the dynamics of development and ultimately, the phenotype.

As shown in Figure 3.1, evolution occurs in three spaces: phenotype space, developmental-parameter space and genotype space. The genotype of an individual is a point in genotype space. This point is mapped additively to a point in developmental-parameter space (see *Methods*). Finally, this point in developmental-parameter space is mapped to a point in phenotype space through the tooth development model. As selection acting on phenotypes pushes the population to move in the phenotypic space in the direction determined by the optimum, it also results in the distribution of genotypes and developmental parameters to change.

The dynamics of the developmental process are different for different combinations of developmental parameters (see equation (1.5)). In this way, as the population evolves, the relationship between genotypes and phenotypes changes because the dynamics of development themselves change. Thus, the population experiences different local GPMs with different characteristics during its evolutionary trajectory. In this way, the local GPM

for a population at a given time is given by the distribution of genotypes in the population at that time, and by the developmental dynamics that map it to a distribution of phenotypes.

A population experiencing different local GPMs in its evolutionary trajectory is given in Figure 3.1 for different time points 1, 2, 3 and 4. At each time point, the local characteristics of the GPM are different. This is reflected in changes in the slope of the best local linear description used to make predictions with quantitative genetics, which is represented in Figure 3.1 as the plane that is tangent to the surface of the GPM at the population mean in each time point (see also Rice 2008).

The work in this thesis shows that experiencing different local GPMs has important consequences in evolutionary dynamics. Most importantly, as explained in the sections below, it results in the breeder's prediction sometimes failing to predict the response to selection and in different ways in which the G-matrix evolves.

3.2 The local GPM and the response to selection

The main finding of Publication I is that single-generation predictions using the breeder's equation can be biased when the local GPM is nonlinear, and that this happens often for GPMs based on development. The word *bias* is used here to emphasize that the prediction errors are systematic, not purely stochastic.

We found that the prediction bias when using the breeder's equation was larger when the local GPM was more nonlinear (Figure 3.2A). This means that the breeder's equation correctly predicted change when the map was locally linear (e.g. 1 and 3 in Figure 3.1), but could be biased when the map was locally nonlinear (e.g. points 2 and 4 in Figure 3.1). Indeed, the local linear description of quantitative genetics is arbitrarily good for a neighborhood of the mean, but how small that neighbourhood is will depend on the nonlinearty of the GPM itself. If the GPM is very nonlinear (e.g. point 4 of Figure 3.1), then this neighborhood may be very small and nonlinearities will affect the dynamics of a population distributed around that mean. Note that the breeder's prediction is given as a change in the mean of the population from one generation to the next. The process of taking the population mean implies integrating the trait values over their distribution, which is effectively equivalent to smoothing the GPM. Even with this smoothing effect, the predictions can be biased.

The largest nonlinearities in the GPM, which were associated with the largest prediction biases, occurred where small genetic changes led to large

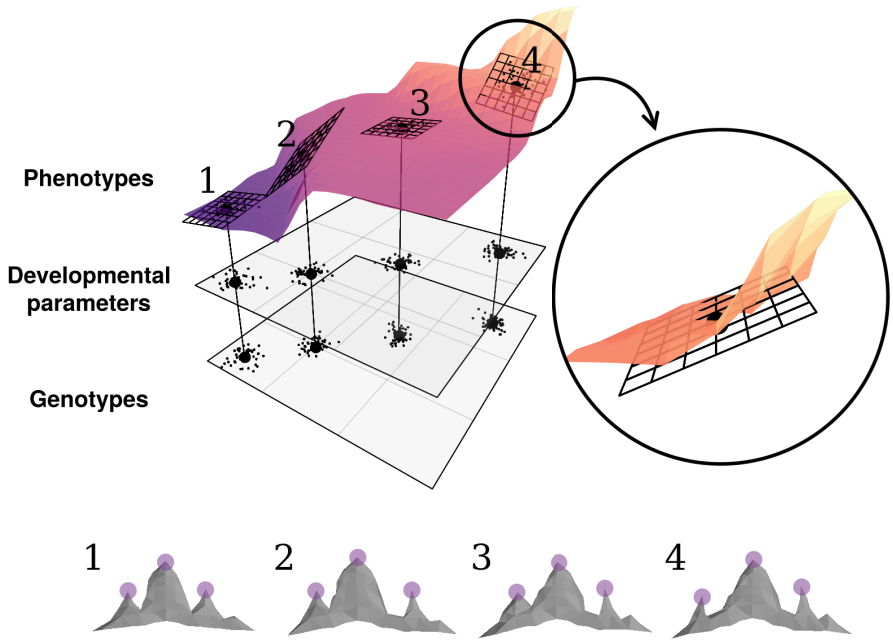

Figure 3.1: The local characteristics of the genotype-phenotype map (GPM) change as the population evolves. The figure shows genotypic space, developmental-parameter space and phenotypic space. The population at four timepoints is shown as a cloud of points in each of these spaces, with each point representing an individual. The linear approximations to the GPM that is used in G-matrix models is represented as a tangent plane plotted in black. The plane fits the local GPM well at some points (1 and 3) and not well at others (2 and 4, shown with a different angle in the inset). The latter are associated with prediction biases when using the breeder's equation. Note that the orientation of the plane changes dramatically depending on the where the population is located. This is reflected in the ways in which the G-matrix evolves. The tooth morphology closest to the population mean at each time point is shown below.

phenotypic changes (e.g. time point 4 of Figure 3.1). This can be seen as discontinuities in the GPM, where the linear approximation is particularly bad. This was most frequently the case for the traits related to the distance between cusps (see traits 2 and 4 in Figure 2.1), where a small change in the developmental parameters can lead to a relatively large change in the positioning of the cusps as shown with the morphologies in Figure 3.1.

An extreme case of large nonlinear behavior was the loss of one of the lateral cusps during evolution, which can be classified as a novelty (Wagner and Lynch 2010). In these situations, the predictions using the breeder's equation was heavily biased. Even though it can be argued that the study of novelties is not within the scope of quantitative-genetic methods (Polly 2008, Hansen 2008), it is worth noting that the loss of cusp was not a rare event in our simulations. Indeed, it occurred at some point of the 30 generations in 10 out of the 32 evolutionary simulations, each with a different optimum. This was particularly the case for simulations in which the lateral cusps were selected to become smaller and more distant to the central cusp. In these cases, small changes in the diffusion of the molecules as a response to selection for smaller and distant cusps led to some individuals not forming lateral cusps, as not enough of the *activator* molecule could accumulate for the formation of an enamel knot, which later in development becomes a cusp (see *Methods*). In this way, a small quantitative change in the diffusion of a molecule lead to a qualitative change in the final morphology. For a complex GPM like the one we use, then, the difference between quantitative and qualitative becomes blurred, raising the question of how separable these types of changes really are.

Previous theoretical work (Rice 2002, 2004, Carter et al. 2005, Heywood 2005) shows that the response to selection may not be correctly predicted by the breeder's equation when the GPM shows certain types of nonlinearities. However, this body of work does not address the question of how common and large we should expect these prediction errors to be for realistic GPMs, so it is not clear if they are of empirical importance. Indeed, if nonlinear behavior of the type that leads to prediction errors is uncommon, then one can effectively rely on the linear model for most practical applications (Hansen 2013). We show in our simulations that for a realistic GPM based on development, these nonlinearities are common and that the resulting prediction errors can be large. Indeed, as shown in Figure 3.2B, we found that for traits related to the distance between cusps (i.e. traits 2 and 4), the median of the prediction bias for single generation prediction was 13% of the observed change, with an interquartile range of 7-32 % of relative bias (see trait 4 in Figure 3.2B). Note that this is for single-generation predictions,

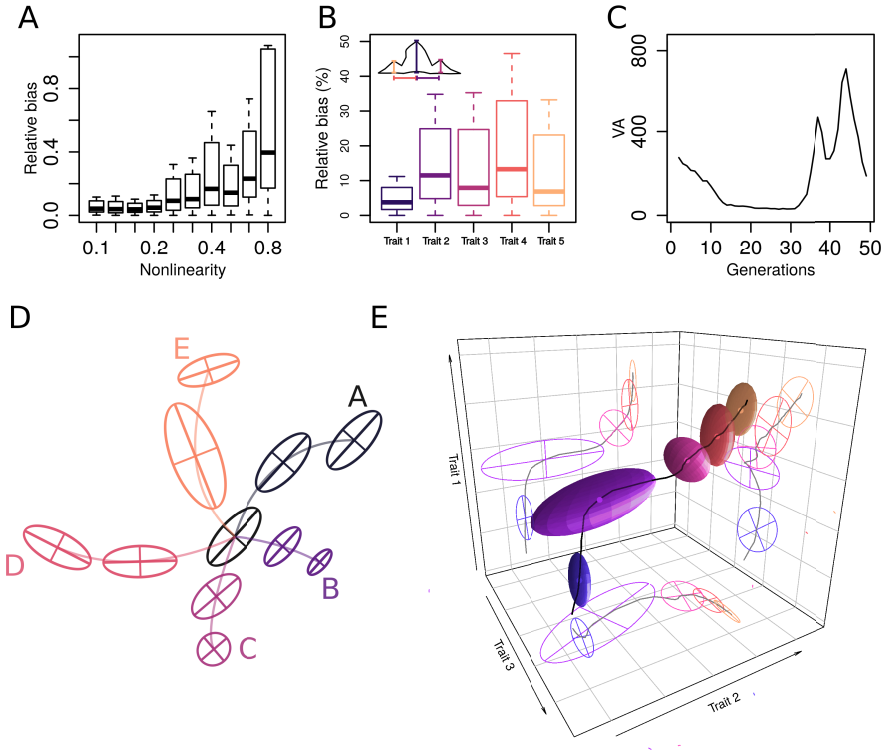

Figure 3.2: **A** shows that the relative prediction bias (i.e. difference between observed and predicted change, divided by observed change) is larger when the local GPM is more nonlinear for the simulations using the tooth model. **B** shows the distribution of prediction bias when predicting the response to selection for each trait in the tooth morphology. **C** shows the amount of additive genetic variance in the direction of selection for example simulation 10. **D** shows shows a schematic representation of the types of changes that can be observed in G in our simulations. The G-matrix is plotted here as ellipses with axes in the direction of the first two eigenvectors, and the length of each axis proportional to the corresponding eigenvalue. **E** shows the evolution of the G-matrix in example simulation 6 at five time points. Details of A-B and C-D-E are given in Publication I and II, respectively.

with the best G-matrix possible in each generation, and excluding the cases where the cusp is lost, which would seriously inflate the median prediction error.

Prediction errors when using the multivariate breeder's equation have been reported empirically (e.g. Sheridan and Barker 1974, Campo and Raya 1986, Roff 2007, Pujol et al 2018, Shaw 2019, Pélabon et al. 2021) and seem to be quite common. However, multiple sources of these errors exist, including wrong estimates of G and omitting traits in the analysis (Shaw 2019, Pujol et al 2018). These sources of error are rarely disentangled, and studies within the framework of quantitative genetics rarely consider the nonlinearities of the GPM as a possible source of error for short term prediction. This is partly rooted in the fact that the genetic interactions that are part of development have been largely regarded as having a negligible contribution to the response to selection, and therefore rarely considered to explain the deviations from predictions (Hill et al. 2008, Crow 2008 and section *The nonlinearity of the GPM is not noise* below). By realistically simulating these interactions, the results from our simulations show that the nonlinearity of the GPM should be more often considered as a possible source of errors when predicting the response to selection using the framework of quantitative genetics.

3.3 The local GPM and the evolution of the G-matrix

The main result from Publication II is that, as the population experiences different local GPMs in our simulations, this leads to different patterns of variation and covariation in the morphologies, and thus to different G-matrices. In this way, if the GPM changes rapidly in the evolutionary trajectory of the population, so will the G-matrix. In turn, if the local GPM stays roughly the same in the trajectory, G will remain constant.

We performed an extensive study of how G evolves in the simulations, using multiple metrics of change. We classified the change in G in five categories, based on how the eigenvectors and eigenvalues of G change. For this, we measured the change in time of the size of G (i.e. the sum of the eigenvalues), the direction of the main eigenvector of G, and the eccentricity of G (i.e. how much of the total additive genetic variance is distributed along the main eigenvector). The five categories of change for G we found are:

A. Constant G: the eccentricity and size of G remain constant over generations.

B. Proportional G: the size of G changes while keeping the same eccentricity. This means G either shrinks of expands.

C. Disproportional G: the eccentricity of G changes. The size may or may not vary.

D. Rotation of G: the direction of the main eigenvector of G changes.

E. Sudden changes in G: relatively rapid changes in eccentricity, size or orientation of G can occur in nonlinear regions of the GPM.

These changes are shown schematically in Figure 3.2D. In a single simulation, the G-matrix can experience several different kinds of changes, as shown for an example in Figure 3.2E. While category A was the most frequent in the simulations, all other categories were found in significant amounts accounting for at least 42% of the changes (details in Publication II). Note that these changes are deterministic, in the sense that they are not stochastic fluctuations but rather structured ways in which G evolves as a reflection of how the local GPM changes.

The multiple ways in which G can change that we find in our simulations have been reported empirically in various systems (e.g. category A in Hangartner et al. 2020, category B in Blows and Higgie 2003, category C in Doroszuk et al. 2008, category D in Walter et al. 2018, category E in Björklund et al. 2013) but have not generally been linked to the GPM. Further, most of these types of changes in G have never been reported in theoretical work. This is because the large majority of previous theoretical work has assumed that the GPM is linear (Turelli 1985, Slatkin and Frank 1990, Reeve 2000, Jones et al. 2003, 2012, 2014, Arnold et al. 2008). The general consensus coming from this body of work using linear GPMs is that the expected shape of G is ultimately determined by selection, with G fluctuating randomly in each generation around this expected shape. These stochastic fluctuations around the expected G can be larger or smaller depending on certain population parameters such as population size and the strength of selection. Much of the previous body work is focused on stabilizing selection, but similar dynamics are reported with a moving optimum (i.e. when there is directional selection, Jones et al. 2012). In this later case with a moving optimum, the authors additionally found that immediately after a shift in the optimum, the strength of directional selection spikes and G has a subtle increase in size and eccentricity, as the population responds to this directional selection. However, the size and eccentricity of G decreases again once the population reaches the optimum. The authors find that this pattern is repeated in the same way each time the optimum moves.

There are several differences between how G evolves in our study and in previous theoretical studies. In our study the changes in G are not stochastic fluctuations nor small, transient and repeatable changes following a shift in the optimum. On the contrary, the types of changes in G that we observe in our simulations have a large deterministic component as observed in changes of the eccentricity, size and orientation as the population evolves. That is, they are not random fluctuations that can be controlled by modifying population parameters, but actual changes in variation and covariation that arise from the fact that the population experiences different local GPMs as it evolves. These changes in G cannot be observed if the GPM is assumed linear.

The high complexity of our GPM also allows to see changes in G that can be much larger than the changes reported under the assumption of a linear GPM. Indeed, we find that different measurements of G such as its size can change in more than 100% in only a few generations (see Figure 3.2E for an example). With the inclusion of genes of large (additive) effects, Agrawal et al. 2001 also report fast and large changes in G that are not stochastic in nature. However, a key difference is that the changes in Agrawal et al. 2001 are reversible and transient, as they depend on the large-effect gene to be segregating. In our case, the changes observed in the G-matrix are not necessarily reversible, as they are determined by the local GPM the population is experiencing.

An aspect of G that has received special attention is its projection on the direction of selection, which has been called evolvability when properly standardized (Hansen and Houle 2008). Some authors suggest that this quantity should increase in evolution as a direct result of selection (Jones 2012, 2014, Pavlicev 2011). This would occur by selection increasing the frequency of the alleles associated with favorable correlations among traits, and decreasing the frequency of those associated with unfavorable correlations. As a result, G would reorient to increase its projection in the direction of selection (Jones et al. 2014, Pavlicev 2011). This implicitly requires a GPM that can change in any direction. On the contrary, we find that V_A in the direction of selection can increase, decrease or remain constant depending on the trajectory of the population across the GPM in our simulations, as shown for an example in Figure 3.2C. Indeed, the amount of V_A in a given direction will be given by how the genetic variation maps to phenotypic variation. If the population enters a region of the GPM that produces much adaptive phenotypic variation given small genetic changes, as in generation 40 in the example given in Figure 3.2C, this will result in an increase in the amount of V_A in the direction of selection. On the contrary,

if the population enters a region where genetic variation is mapped to little adaptive phenotypic variation, then V_A will decrease. These results further extend the insights from simpler models that include nonlinear GPMs and find that V_A in the single-trait case has different dynamics than those expected by linear GPMs (Carter et al. 2005, Hansen 2006, see also Rice 2004).

3.4 The nonlinearity of the GPM is not noise

When included, the interactions between genes enter applied quantitative-genetic models as statistical deviations from additivity, akin to noise (Fisher 1958, Hill et al. 2008, Crow 2008, 2010, Hansen 2013, Nelson et al. 2013, and section *1.1 Quantitative genetics*). Modeling genetic interactions as noise is compatible with the linear regressions that form the basis of quantitative genetics, but effectively erases the structure of the GPM. This occurs because noise cannot represent specific patterns of interactions that result in curvature of the GPM (Rice 2002, 2004, Hansen 2006). Modelling interactions as noise thus obscures the influence of the curvature of the GPM in evolutionary dynamics.

Previous work shows that if the GPM includes certain patterns of interactions, this has systematic effects on evolutionary dynamics that are not captured when modelling interactions as noise (Rice 2004, Carter et al. 2005, Hansen 2006). For example, as explained in the *Introduction*, Rice (2002, 2004) uses his framework to show that for certain functional relationships between underlying developmental parameters and traits, the breeder's equation is not enough to predict the response to selection. Moreover, Carter et al. 2005 use the multilinear model (Hansen and Wagner 2001) and find that despite deviations from the breeder's equations being negligible for single-generation predictions, the system quickly deviates from the linear prediction through rapid changes in the amount of additive genetic variance in the presence of signed epistasis (i.e. when epistatic interactions have an average tendency to reinforce or diminish each other's effects). The work in this thesis further proves these claims, and importantly shows that GPMs with such characteristics are to be expected when using a realistic model of development.

In our simulations, the deterministic and causal role of the GPM in evolutionary dynamics was evidenced by the finding of systematic bias using the breeder's equation for single-generation predictions, as well as in the different deterministic ways in which we found the G-matrix to evolve. Both effects on evolutionary dynamics are a consequence of the structure of the

GPM.

The discussion above highlights the fact that the way in which we model the GPM will determine the evolutionary roles we can attribute to the GPM. In this sense, the work in this thesis is valuable because instead of proposing an arbitrary GPM, we model the development of each individual using a realistic model of the development of a complex organ, and from this a realistic GPM arises. Previous work (Omholt et al. 2000, Gjuvsland et al. 2007, 2011, 2013) uses a similar approach to represent the GPM using gene networks. This body of work is of relevance because it connects properties of genetic networks to summary statistics of quantitative genetics, and highlights the importance of describing the GPM as arising from a dynamical system.

Dynamical systems present a set of generic properties such as context dependency and nonlinearity. These generic properties result in the interesting evolutionary dynamics reported in this thesis, so we should expect these dynamics to be common for other GPMs associated with dynamical systems, including the development of other organs (e.g. Newman and Müller, Urdy 2012), RNA folding (e.g. Schuster et al. 1994, Aguirre et al. 2011) and gene networks (e.g. Wagner 1996, Cotterel and Sharpe 2010). In this way, even though the results presented in this thesis are specific to the GPM generated by tooth development, the results should be regarded as indicative of the general properties of GPMs that arise from complex dynamical systems. Importantly, this realistic representation of the GPM uncovers important implications of the GPM on evolutionary dynamics that remain hidden with simplified models.

3.5 Using the local GPM to improve predictions

As explained above, when the GPM is nonlinear, the breeder's equation can be biased for single-generation predictions and the G-matrix can change relatively fast. Both of these features negatively affect our ability to predict evolution using the breeder's equation. Indeed, because G is so cumbersome to estimate, in practice one uses a single estimate of G to predict change in several generations under the assumption that it remains constant (Walsh and Lynch 2018). If the G-matrix does change in that period, this will negatively affect the predictive capacity of the breeder's equation. Further problems arise when applying the breeder's equation in real populations since, as mentioned in the *Introduction*, any violation in the assumptions underlying the breeder's equation can result in prediction errors. Common violations include the fact that ultimately we only have estimates of G,

which can be wrong, particularly when G is estimated with data from a different population (Pigliucci 2006) or when not accounting for relevant effects (e.g. maternal effects, Pujol et al 2018, Walsh and Lynch 2018). Another common violation occurs when the traits chosen for study do not account for all selection (Pujol et al. 2018, Shaw 2019).

An important feature of the error arising from these violations in the assumptions of the breeder's equation is that it is systematic. For example, if a trait under selection is missing from the analysis, the prediction using the breeder's equation can be biased because there is an indirect effect of selection that is systematically omitted in the prediction (Merilä et al. 2001). Moreover, if the G-matrix has changed rapidly, for example because the local GPM has changed, predictions will also be biased because the G used for predictions has changed in a systematic way, as explained in the previous section.

The presence of a systematic bias in the predictions means that the error is not purely stochastic, but somewhat structured. In other words, the error at a given generation i is informative of the error at generation $i + 1$. This indicates that there is potential to improve predictions by incorporating this bias, if one could hold the information of past generations as a "memory".

The insight above motivated the development of a novel method to predict the response to selection in Publication III. The method estimates the change in the mean of the traits in each generation as the the breeder's equation plus a bias term. Change is then estimated at generation i by combining information from the previous recorded change in the mean at $i-1$ and the breeder's equation, as schematically shown in Figure 3.3A. The combination of information is done using a Kalman filter, with parameters fitted in each generation using a machine-learning algorithm on the record of past changes. Importantly, the novel method will correct any type of prediction bias in the breeder's equation.

The novel prediction method was applied to the simulations using the tooth development model, and to a 20-generation experiment using the wing of *Drosophila melanogaster* where a total of 16000 flies, 200 per generation, where measured in three selection lines and a control (see *Methods*). The experiments with the fruit fly include the full complexity of predicting the response to selection. We show that in both sets of experiments, the new method outperforms the breeder's equation in its ability to predict the response to selection when some of the assumptions of the breeder's equation do not hold.

Figure 3.3B compares the prediction of the breeder's equation and the new method for the change in the mean of one of the traits in one of the

selection lines of the artificial selection experiments with the fly (see Publication III for details). The figure shows that the new method provides better prediction than the breeder's equation (i.e. closer to the observed change).

Figure 3.3C gives a summary of the total prediction error for the breeder's equation and the new method, for the three selection lines of experimental evolution on the wing, and using G-matrices estimated at the start of the experiment with different degrees of precision. Precision was given by the number of generations of the control line used to make estimates of the G-matrix, which we call the pedigree depth. Naturally, a more precise estimate of G is more expensive to obtain as it requires more generations of data. The figure shows that the new method gives better predictions regardless of how accurate the G-matrix is. For example, for a pedigree depth of 2, the new method reduces the prediction error of the breeder's equation by 27%, 60% and 52% for replicate lines 1, 2 and 3 respectively. Further, the figure shows that the new method using an inaccurate G-matrix yields better predictions than the breeder's equation using a very accurate, and expensive to estimate, G-matrix. Moreover, the method is able to provide good prediction even when G is not estimated at all, which corresponds to pedigree depth 1 in the x-axis of Figure 3.3G.

The method is mostly developed using tools of quantitative genetics, including the breeder's equation. However, it's development is rooted in insight from evo-devo. Therefore, the method can be considered to combine knowledge from both fields. The method also combines different types of information: past and present. Past information is incorporated by the recursive nature of the model, as it forecasts the variables of interest using past estimates of the variables (see Figure 3.3A). This is combined dynamically in each generation with present information given by selection acting in each generation. The combination of approaches and types of information distinguishes the method from other valuable recent efforts to use recursive models to make predictions of future evolution (Le Rouzic et al 2011, Nosil et al. 2018, Rescan et al. 2020, Nosil et al. 2020, Rescan et al. 2021). Indeed, several authors advocate the combination of multiple sources of information and approaches to improve predictions (Nosil et al 2020). This reflects the current needs in much of the biological sciences, where the amount of data has surpassed the ability of existing methods to integrate all the information. The method in Publication III is a contribution in this direction.

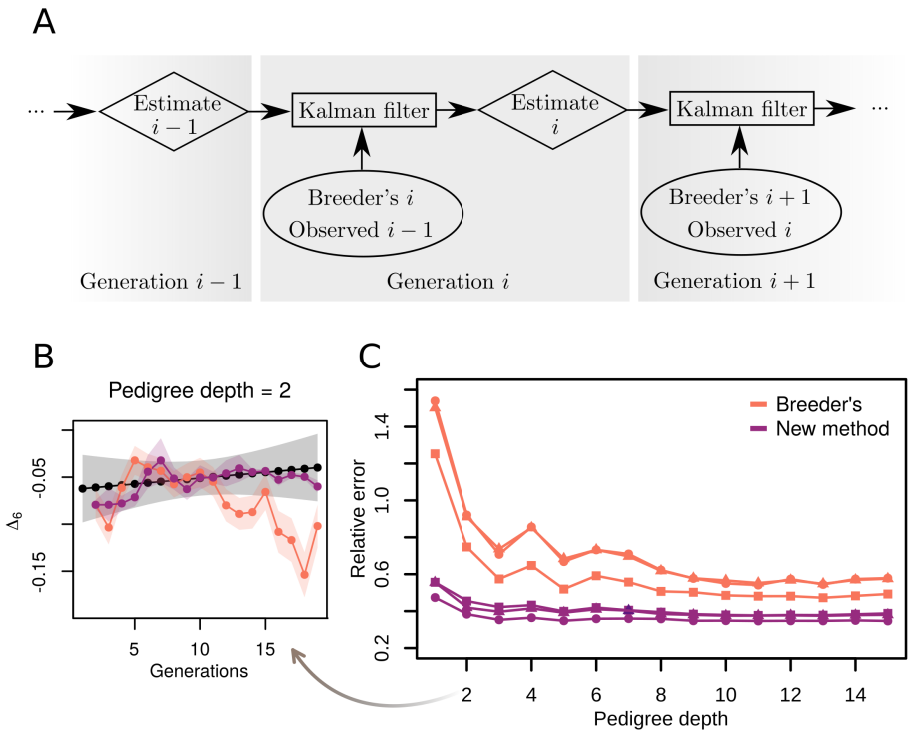

Figure 3.3: **A** shows a simplified version of the algorithm used in the new method to predict the response to selection developed in Publication III. In each generation, the prediction of the breeder's equation is combined with the last observed change and used to provide the best possible estimate of the change in the mean of the trait for the next generation. **B** shows the change in the mean of one of the traits (symbolized Δ_6) in line 1 of the artificial selection experiments using the wing of the fruit fly. The observed change is shown in black, and the predictions using the breeder's equation and the new method are shown in orange and purple respectively. **C** shows the total error of the prediction using the breeder's equation (orange) and the new method (purple) for the 3 selection lines of the experiments with the fruit fly (squares correspond to line 1, circles to line 2 and triangles to line 3). The new method outperforms the breeder's equation for all pedigree depths.

Chapter 4

Conclusions

The fact that the mapping between genotypes and phenotypes, or GPM, is complex and nonlinear is widely accepted among biologists. How this nonlinearity affects evolutionary dynamics, however, remains a central open question in evolutionary biology. This is particularly the case when studying microevolution at the population level. In quantitative genetics, the GPM is approximated with a linear statistical abstraction which is assumed to capture all the information about the GPM that is relevant for evolutionary dynamics, at least in the short term. We find that this is not necessarily the case when then GPM is nonlinear, by answering Questions 1 and 2 posed in the *Introduction*:

Q1. How good is the local statistical description when the genotype-phenotype map is nonlinear?

Q2. How does the local statistical description change in time when the genotype-phenotype map is nonlinear?

First, we find that the prediction of the change in the mean of the traits in response to selection using the linear abstraction can be biased when the GPM is nonlinear. In other words, there can be a significant part of the response to selection that is missed by the linear approximation when the GPM is very nonlinear. Second, we find that the dynamics of how the local linear description changes in time are also highly dependent on the GPM, and differ substantially from what is expected for a linear GPM. These dynamics, which we classify and study in detail, are not purely stochastic, but rather deterministic ways in which the linear approximation changes as a reflection of changes in the local GPM. All in all, these results highlight how the GPM determines evolutionary dynamics, even for short-term evolution at the level of population. Importantly, this insight about the GPM inspired the development of a novel prediction method for the response to selection that provides improved predictions.

The work in this thesis shows that rather than being a nuisance, the structure of the GPM contains information that has the potential to improve our understanding of evolutionary processes. This insight can be exploited for applied purposes such as improving our ability to predict evolution, but potentially also for other applications that require understanding how genetic information relates to phenotypes, such as personalized medicine and risk assessment. Understanding the GPM should then be a central objective of 21st century biology.

References

Agrawal, A. F., Brodie, E. D., & Rieseberg, L. H. (2001). Possible consequences of genes of major effect: Transient changes in the G-matrix. Genetica, 112–113, 33–43.

Aguirre, J., Buldú, J. M., Stich, M., & Manrubia, S. C. (2011). Topological structure of the space of phenotypes: the case of RNA neutral networks. PloS one, 6(10), e26324.

Alberch, P. (1991). From genes to phenotype: dynamical systems and evolvability. Genetica, 84(1), 5-11.

Arif, S., Hilbrant, M., Hopfen, C., Almudi, I., Nunes, M. D. S., Posnien, N., Kuncheria, L., Tanaka, K., Mitteroecker, P., Schlötterer, C., & McGregor, A. P. (2013). Genetic and developmental analysis of differences in eye and face morphology between Drosophila simulans and Drosophila mauritiana. Evolution & Development, 15(4), 257–267.

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

Arnold, S. J., Bürger, R., Hohenlohe, P. A., Ajie, B. C., & Jones, A. G. (2008). Understanding the evolution and stability of the G-matrix. Evolution, 62(10), 2451-2461.

Assis, A. P. A., Patton, J. L., Hubbe, A., & Marroig, G. (2016). Directional selection effects on patterns of phenotypic (co)variation in wild populations. Proceedings of the Royal Society B: Biological Sciences, 283(1843).

Åström, K., & Wittenmark, B. (1997). Computer-controlled systems: theory and design. 3rd Edition. Englewood Cliffs, NK: Prentice-Hall.

Beatty, J. (1986). The Synthesis and the Synthetic Theory. In Integrating Scientific Disciplines, ed. W. Bechtel, pp. 125–136. The Hague, Netherlands: Nijhoff.

Björklund, M., Husby, A., & Gustafsson, L. (2013). Rapid and unpredictable changes of the G-matrix in a natural bird population over 25 years. Journal of Evolutionary Biology, 26(1), 1–13.

Blows, M. W., & Higgie, M. (2003). Genetic constraints on the evolution of mate recognition under natural selection. American Naturalist, 161(2), 240–253.

Campo, J. L., & Raya, L. G. (1986). Realized genetic parameters from an antagonistic selection index in Tribolium castaneum. Canadian journal of genetics and cytology, 28(3), 358-364.

Cano, J. M., Laurila, A., Palo, J., & Merilä, J. (2004). Population differentiation in g matrix structure due to natural selection in Rana temporaria. Evolution, 58(9), 2013–2020.

Careau, V., Wolak, M. E., Carter, P. A., & Garland, T. (2015). Evolution of the additive genetic variance–covariance matrix under continuous directional selection on a complex behavioural phenotype. Proceedings of the Royal Society B: Biological Sciences, 282(1819).

Carroll, S. B., Grenier, J. K., & Weatherbee, S. D. (2013). From DNA to diversity: molecular genetics and the evolution of animal design. John Wiley & Sons.

Carter, A. J. R., Hermisson, J., & Hansen, T. F. (2005). The role of epistatic gene interactions in the response to selection and the evolution of evolvability. Theoretical Population Biology, 68, 179–196.

Chakrabarty, A., & Schielzeth, H. (2020). Comparative analysis of the multivariate genetic architecture of morphological traits in three species of Gomphocerine grasshoppers. Heredity, 124(2), 367–382.

Chebib, J., & Guillaume, F. (2017). What affects the predictability of evolutionary constraints using a G-matrix? The relative effects of modular pleiotropy and mutational correlation. Evolution, 71(10), 2298-2312.

Cheverud, J. M. (1988). A comparison of genetic and phenotypic correlations. Evolution, 42(5), 958–968.

Cheverud, J. M., & Routman, E. J. (1995). Epistasis and its contribution to genetic variance components. Genetics, 139(3), 1455-1461.

Cotterell, J., & Sharpe, J. (2010). An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients. Molecular systems biology, 6(1), 425.

Crow, J. F. (2008). Maintaining evolvability. J. Genet. 87: 349-353.

Crow, J. F. (2010). On epistasis: why it is unimportant in polygenic directional selection. Phil. Trans. R. Soc. B 365: 1241-1244.

Davidson, E. H. (2001). Genomic regulatory systems: in development and evolution. Elsevier.

Delahaie, B., Charmantier, A., Chantepie, S., Garant, D., Porlier, M., & Teplitsky, C. (2017). Conserved G-matrices of morphological and life-history traits among continental and island blue tit populations. Heredity, 119, 76–87.

Doroszuk, A., Wojewodzic, M. W., Gort, G., & Kammenga, J. E. (2008). Rapid divergence of genetic variance-covariance matrix within a natural population. American Naturalist, 171(3), 291–304.

Eroukhmanoff, F., & Svensson, E. I. (2011). Evolution and stability of the G-matrix during the colonization of a novel environment. Journal of Evolutionary Biology, 24(6), 1363–1373.

Eroukhmanoff, F.. (2009). Just how much is the G-matrix actually constraining adaptation? Evolutionary Biology, 36(3), 323–326.

Falconer, D. S. and Mackay, T. F. C. (1996). Introduction to Quantitative Genetics. Pearson, Fourth Edition.

Feldman, M. W., & Lewontin, R. C. (1975). The heritability hang-up. Science, 190(4220), 1163-1168.

Fisher, R. A., (1958). The Genetical Theory of Natural Selection. Dover, New York.

Garcia, P. D., & Garcia, R. (2018). Determination of the Elastic Moduli of a Single Cell Cultured on a Rigid Support by Force Microscopy. Biophysical Journal, 114(12), 2923–2932.

Gilbert S. F. and Barresi M. J. F. (2016). Developmental biology. Sunderland, MA: Sinauer Associates.

Gimelfarb, A., & Willis, J. H. (1994). Linearity Versus nonlinearity of offspring-parent regression: an experimental study of Drosophila melanogaster. Genetics, 138(2), 343-352.

Gjuvsland, A. B., Hayes, B. J., Omholt, S. W., & Carlborg, O. (2007). Statistical epistasis is a generic feature of gene regulatory networks. Genetics, 175(1), 411-420.

Gjuvsland, A. B., Vik, J. O., Woolliams, J. A., & Omholt, S. W. (2011). Order-preserving principles underlying genotype-phenotype maps ensure high additive proportions of genetic variance. Journal of Evolutionary Biology, 24(10), 2269–2279.

Gjuvsland, A. B., Wang, Y., Plahte, E., & Omholt, S. W. (2013). Monotonicity is a key feature of genotype-phenotype maps. Frontiers in genetics, 4, 216.

Gromko, M. H. (1995). Unpredictability of correlated response to selection: pleiotropy and sampling interact. Evolution, 49(4), 685-693.

Hall, B. K. (2012). Evolutionary developmental biology (Evo-Devo): Past, present, and future. Evolution: Education and outreach, 5(2), 184-193.

Hall, B. K. (2013). Homology, homoplasy, novelty, and behavior. Developmental Psychobiology, 55(1), 4–12.

Hangartner, S., Lasne, C., Sgrò, C. M., Connallon, T., & Monro, K. (2020). Genetic covariances promote climatic adaptation in Australian Drosophila. Evolution, 74(2), 326–337.

Hansen, T. F., & Houle, D. (2008). Measuring and comparing evolvability and constraint in multivariate characters. Journal of Evolutionary Biology, 21(5), 1201–1219.

Hansen, T. F., & Wagner, G. P. (2001). Modeling genetic architecture: A multilinear theory of gene interaction. Theoretical population biology, 59(1), 61-86.

Hansen, T. F. (2006). The Evolution of Genetic Architecture. Annual Review of Ecology, Evolution, and Systematics, 37(1), 123–157.

Hansen, T. F. (2008). Macroevolutionary quantitative genetics? A comment on Polly (2008). Evolutionary Biology, 35(3), 182-185.

Hansen, T. F. (2013). Why epistasis is important for selection and adaptation. Evolution, 67(12), 3501-3511.

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.

Hill, W. G. (2014). Applications of population genetics to animal breeding, from Wright, Fisher and Lush to genomic prediction. Genetics, 196(1), 1-16.

Hill, W. G. (2010). Understanding and using quantitative genetic variation. Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1537), 73-85.

Hill, W. G., & Bunger, L. (2004). Inferences on the genetics of quantitative traits from long-term selection in laboratory and domestic animals. Plant breeding reviews, 24(2), 169-210.

Hill, W. G., & Kirkpatrick, M. (2010). What Animal Breeding Has Taught Us about Evolution. Annu. Rev. Ecol. Evol. Syst, 41, 1–19.

Hill, W. G., Goddard, M. E., & Visscher, P. M. (2008). Data and theory point to mainly additive genetic variance for complex traits. PLoS genetics, 4(2), e1000008.

Houle, D. (1991). Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. Evolution, 45(3), 630-648.

Houle, D., & Meyer, K. (2015). Estimating sampling error of evolutionary statistics based on genetic covariance matrices using maximum likelihood. Journal of Evolutionary Biology, 28(8), 1542-1549.

Houle, D., Mezey, J., Galpern, P., & Carter, A. (2003). Automated measurement of Drosophila wings. BMC Evolutionary Biology, 3.

Huxley, J. (1942). Evolution. The Modern Synthesis. George Alien & Unwin Ltd.

Huang, W., & Mackay, T. F. (2016). The genetic architecture of quantitative traits cannot be inferred from variance component analysis. PLoS genetics, 12(11), e1006421.

Jernvall, J., & Thesleff, I. (2000). Reiterative signaling and patterning during mammalian tooth morphogenesis. Mechanisms of Development, 92(1), 19–29.

Jones, A. G., Bürger, R., & Arnold, S. J. (2014). Epistasis and natural selection shape the mutational architecture of complex traits. Nature communications, 5(1), 1-10.

Jones, A. G., Bürger, R., Arnold, S. J., Hohenlohe, P. A., & Uyeda, J. C. (2012). The effects of stochastic and episodic movement of the optimum on the evolution of the G -matrix and the response of the trait mean to selection. Journal of Evolutionary Biology, 25(11), 2210–2231.

Jones, Adam G., Arnold, S. J., & Bürger, R. (2003). Stability of the g-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. Evolution, 57(8), 1747–1760.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Journal of Fluids Engineering, Transactions of the ASME, 82(1), 35–45.

Kirkpatrick, M. (2009). Patterns of quantitative genetic variation in multiple dimensions. Genetica, 136(2), 271–284.

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

Kruuk, L. E., Slate, J., & Wilson, A. J. (2008). New answers for old questions: the evolutionary quantitative genetics of wild animal populations. Annual Review of Ecology, Evolution, and Systematics, 39, 525-548.

Lande, R. (1979). Quantitative Genetic Analysis of Multivariate Evolution, Applied to Brain: Body Size Allometry. Evolution, 33(1), 402.

Lande, R., & Arnold, S. J. (1983). The measurement of selection on correlated characters. Evolution, 1210-1226.

Lickliter, R. (2007). The dynamics of development and evolution: Insights from behavioral embryology. Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology, 49(8), 749-757.

Le Rouzic, A., Houle, D., & Hansen, T. F. (2011). A modelling framework for the analysis of artificial-selection time series. Genetics research, 93(2), 155-173.

Levins, R., & Lewontin, R. (1985). The dialectical biologist. Harvard University Press.

Lewontin, R. C. (1974). The analysis of variance and the analysis of causes. American journal of human genetics, 26(3), 400.

Loehlin, D. W., & Werren, J. H. (2012). Evolution of shape by multiple regulatory changes to a growth gene. Science, 335(6071), 943-947.

Lush, J. L. (1937). Animal breeding plans. In Animal breeding plans. (Issue Edn 2). Iowa State College Press.

Lynch, M., & Walsh, B. (1998). Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA.

Mallarino, R., Campàs, O., Fritz, J. A., Burns, K. J., Weeks, O. G., Brenner, M. P., & Abzhanov, A. (2012). Closely related bird species demonstrate flexibility between beak morphology and underlying developmental programs. Proceedings of the National Academy of Sciences, 109(40), 16222-16227.

Maynard Smith, J., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B., ... & Wolpert, L. (1985). Developmental constraints and evolution: a perspective from the Mountain Lake conference on development and evolution. The Quarterly Review of Biology, 60(3), 265-287.

Mayr, E. (1963). Animal species and evolution. Harvard University Press, Cambridge, Mass.

Mayr, E. (1982). The growth of biological thought: Diversity, evolution, and inheritance. Harvard University Press.

McGuigan, K., & Blows, M. W. (2007). The phenotypic and genetic covariance structure of drosphilid wings. Evolution, 61(4), 902–911.

Merilä, J., Björklund, M., & Gustafsson, L. (1994). Evolution of morphological differences with moderate genetic correlations among traits as exemplified by two flycatcher species (Ficedula; Muscicapidae). Biological Journal of the Linnean Society, 52(1), 19–30.

Merilä, J., Sheldon, B. C., & Kruuk, L. E. B. (2001). Explaining stasis: microevolutionary studies in natural populations. Genetica, 112, 199–222.

Meyer, K. (1989). Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm. Genetics selection evolution, 21(3), 317-340.

Meyer, K. (2007). WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). Journal of Zhejiang University. Science. B, 8(11), 815.

Milocco, L., & Salazar-Ciudad, I. (2020). Is evolution predictable? Quantitative genetics under complex genotype-phenotype maps. Evolution, 74(2), 230-244.

Milocco, L., & Salazar-Ciudad, I. (2022). Evolution of the G Matrix under Nonlinear Genotype-Phenotype Maps. The American Naturalist, 199(3), 000-000.

Milocco, L., & Salazar-Ciudad, I. (2021). Learning from your mistakes: a novel method to predict the response to directional selection. EcoEvoRxiv.

Morrissey, M. B., Kruuk, L. E. B., & Wilson, A. J. (2010). The danger of applying the breeder's equation in observational studies of natural populations. Journal of Evolutionary Biology, 23(11), 2277–2288.

Müller, G. B. (2007). Evo-devo: Extending the evolutionary synthesis. Nature Reviews Genetics, 8(12), 943-949.

Nelson, R. M., Pettersson, M. E., & Carlborg, Ö. (2013). A century after Fisher: time for a new paradigm in quantitative genetics. Trends in Genetics, 29(12), 669–676.

Newman, S. A., & Comper, W. D. (1990). 'Generic'physical mechanisms of morphogenesis and pattern formation. Development, 110(1), 1-18.

Newman, S. A., & Müller, G. B. (2000). Epigenetic mechanisms of character origination. Journal of Experimental Zoology, 288(4), 304–317.

Newman, S. A., & Müller, G. B. (2005). Origination and innovation in the vertebrate limb skeleton: an epigenetic perspective. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 304B(6), 593–609.

Nosil, P., Villoutreix, R., Carvalho, C. F. de, Farkas, T. E., Soria-Carrasco, V., Feder, J. L., Crespi, B. J., & Gompert, Z. (2018). Natural selection and the predictability of evolution in Timema stick insects. Science, 359(6377), 765–770.

Nosil, P., Villoutreix, R., de Carvalho, C. F., Feder, J. L., Parchman, T. L., & Gompert, Z. (2020). Ecology shapes epistasis in a genotype–phenotype–fitness map for stick insect colour. Nature Ecology & Evolution, 4(12), 1673-1684.

Nunes, M. D. S., Arif, S., Schlötterer, C., & McGregor, A. P. (2013). A Perspective on Micro-Evo-Devo: Progress and Potential. Genetics, 195(3), 625.

Omholt, S. W., Plahte, E., Øyehaug, L., & Xiang, K. (2000). Gene regulatory networks generating the phenomena of additivity, dominance and epistasis. Genetics, 155(2), 969-980.

Oster, G., & Alberch, P. (1982). Evolution and Bifurcation of Developmental Programs. Evolution, 36(3), 444–459.

Oyama, S. (2000). Evolution's eye. Duke University Press.

Parsons, K. J., & Albertson, R. C. (2013). Unifying and generalizing the two strands of evo-devo. Trends in ecology & evolution, 28(10), 584-591.

Pavlicev, M., Cheverud, J. M., & Wagner, G. P. (2011). Evolution of adaptive phenotypic variation patterns by direct selection for evolvability. Proceedings of the Royal Society B: Biological Sciences, 278(1713), 1903–1912.

Pélabon, C., Albertsen, E., Rouzic, A. Le, Firmat, C., Bolstad, G. H., Armbruster, W. S., & Hansen, T. F. (2021). Quantitative assessment of observed versus predicted responses to selection. Evolution, 75(9), 2217–2236.

Penna, A., Melo, D., Bernardi, S., Oyarzabal, M. I., & Marroig, G. (2017). The evolution of phenotypic integration: How directional selection reshapes covariation in mice. Evolution 71(10), 2370-2380.

Phelan, J. P., Archer, M. A., Beckman, K. A., Chippindale, A. K., Nusbaum, T. J., & Rose, M. R. (2003). Breakdown in correlations during laboratory evolution. I. Comparative analyses of drosophila populations. Evolution, 57(3), 527–535.

Pigliucci, M. (2006). Genetic Variance–covariance Matrices: A Critique of the Evolutionary Quantitative Genetics Research Program. Biology and Philosophy 2006 21:1, 21(1), 1–23.

Pigliucci, M. (2010). Genotype–phenotype mapping and the end of the 'genes as blueprint'metaphor. Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1540), 557-566.

Pinheiro, J. C., & Bates, D. M. (2000). Linear mixed-effects models: basic concepts and examples. Mixed-effects models in S and S-Plus, 3-56.

Polly, P. D. (2008). Developmental dynamics and G-matrices: Can morphometric spaces be used to model phenotypic evolution?. Evolutionary biology, 35(2), 83-96.

Postma, E., & Charmantier, A. (2007). What 'animal models' can and cannot tell ornithologists about the genetics of wild populations. Journal of Ornithology, 148(2), 633-642.

Pujol, B., Blanchet, S., Charmantier, A., Danchin, E., Facon, B., Marrot, P., Roux, F., Scotti, I., Teplitsky, C., Thomson, C. E., & Winney, I. (2018). The Missing Response to Selection in the Wild. Trends in Ecology & Evolution, 33(5), 337–346.

Raff, R. A. (1996). The shape of life: genes, development, and the evolution of animal form. University of Chicago Press.

Reeve, J. P. (2000). Predicting long-term response to selection. Genetics Research, 75(1), 83–94.

Rescan, M., Grulois, D., Aboud, E. O., de Villemereuil, P., & Chevin, L. M. (2021). Predicting population genetic change in an autocorrelated random environment: Insights from a large automated experiment. PLoS genetics, 17(6), e1009611.

Rescan, M., Grulois, D., Ortega-Aboud, E., & Chevin, L. M. (2020). Phenotypic memory drives population growth and extinction risk in a noisy environment. Nature ecology & evolution, 4(2), 193-201.

Rice, S. H. (2004). Evolutionary Theory: Mathematical and Conceptual Foundations. Sinauer Associates.

Rice, S. H. (2008). The G-Matrix as one piece of the phenotypic evolution puzzle. Evolutionary Biology, 35(2), 106–107.

Rice, S. H. (2012). The place of development in mathematical evolutionary theory. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 318(6), 480-488.

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

Rice, S. H. (2002). A general population genetic theory for the evolution of developmental interactions. Proceedings of the National Academy of Sciences, 99(24), 15518-15523.

Salazar-Ciudad, I., & Jernvall, J. (2010). A computational model of teeth and the developmental origins of morphological variation. Nature 2010 464:7288, 464(7288), 583–586.

Salazar-Ciudad, I. (2006). Developmental constraints vs. variational properties: how pattern formation can help to understand evolution and development. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 306B(2), 107–125.

Schluter, D. (1996). Adaptive radiation along genetic lines of least resistance. Evolution, 50(5), 1766-1774.

Schuster, P., Fontana, W., Stadler, P. F., & Hofacker, I. L. (1994). From sequences to shapes and back: a case study in RNA secondary structures. Proceedings of the Royal Society of London. Series B: Biological Sciences, 255(1344), 279-284.

Shaw, F. H., Shaw, R. G., Wilkinson, G. S., & Turelli, M. (1995). Changes in genetic variances and covariances: G whiz! Evolution, 49(6), 1260–1267.

Shaw, R. G. (2019). From the past to the future: Considering the value and limits of evolutionary prediction. American Naturalist, 193(1), 1–10.

Sheridan, A. K., & Barker, J. S. F. (1974). Two-trait selection and the genetic correlation II. Changes in the genetic correlation during two-trait selection. Australian journal of biological sciences, 27(1), 89-102.

Slatkin, M., & Frank, S. A. (1990). The Quantitative Genetic Consequences of Pleiotropy Under Stabilizing and Directional Selection. Genetics, 125(1), 207-213.

Steppan, S. J., Phillips, P. C., & Houle, D. (2002). Comparative quantitative genetics: Evolution of the G matrix. Trends in Ecology and Evolution, 17(7), 320-327.

Stern, D. L. (1998). A role of Ultrabithorax in morphological differences between Drosophila species. Nature, 396(6710), 463-466.

Turelli, M. (1985). Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. Genetics, 111(1), 165-195.

Urdy S. (2012). On the evolution of morphogenetic models: mechano-chemical interactions and an integrated view of cell differentiation, growth, pattern formation and morphogenesis. Biol Rev Camb Philos Soc. 2012 Nov;87(4):786-803.

Wagner, G. P., & Lynch, V. J. (2010). Evolutionary novelties. Current Biology, 20(2).

Wagner, A. (1996). Does evolutionary plasticity evolve?. Evolution, 50(3), 1008-1023.

Walsh, B., & Blows, M. W. (2009). Abundant genetic variation+ strong selection= multivariate genetic constraints: a geometric view of adaptation. Annual Review of Ecology, Evolution, and Systematics, 40, 41-59.

Walsh, B., & Lynch, M. (2018). Evolution and selection of quantitative traits. Oxford University Press.

Walter, G. M., Aguirre, J. D., Blows, M. W., & Ortiz-Barrientos, D. (2018). Evolution of genetic variance during adaptive radiation. The American Naturalist, 191(4), E108-E128.

Wilkins, A. S. (2002). The evolution of developmental pathways. Sinauer Associates.

Wood, C. W., & Brodie III, E. D. (2015). Environmental effects on the structure of the G-matrix. Evolution, 69(11), 2927-2940.

Zhang, X. S., & Hill, W. G. (2005). Genetic variability under mutation selection balance. Trends in Ecology & Evolution, 20(9), 468-470.