Evolvability and Robustness
A Paradox in Evolutionary Theory

Christine Syrowatka

University of Oslo
EvoGene, CEES, CEDE

Thesis submitted for the degree of Philosophiæ Doctor

September 11, 2020

Remember to look up at the stars,

and not down at your feet.

Try to make sense of what you see,

and wonder about what makes the universe exist.

Be curious.

And however difficult life may seem,

there is always something you can do,

and succeed at.

It matters that you don't just give up.

Stephen Hawking

CONTENTS

i

ACKNOWLEDGMENTS

First of all, I would like to thank my supervisor Thomas Hansen, for always having an open door and sharing his knowledge with me. Your unwavering enthusiasm kept me constantly engaged with my research. Thank you for your endless patience during the many inspiring scientific discussions we had. I would not have managed without the great guidance of my co-supervisors. I would like to thank Arnaud Le Rouzic for all his wonderful ideas, support in coding and running my scripts, and the many times where he translated Thomas to me. Thank you for all your encouragement and motivation. I am grateful to Lee Hsiang Liow for her guidance and support. Thank you for all your advice and for always being there for a chat when needed. Last but not least, Kamran Shalchian-Tabrizi thank you for all the stimulating discussions, your inspiration, and your support. Thank you all for paving the way for me to become an independent researcher.

I would also like to thank all my amazing colleagues and friends for being who they are and for all the inspiring discussions, coffee breaks, cigarette breaks, lunches, and especially wine and cheeses. Delphine, I think I would not have survived the time of my Ph.D. without all the fun, support, and wine we shared. Franzi, thanks for being an amazing office-mate over the years. You shared all my laughs, frustration and cheered me up when needed. Kristof, over all the years and the endless scientific discussions, frustrations, laughs, and silliness you still have not given up on me. Julia, thanks for all the fun we had, even if we sometimes only had three carrots and toilet water. Jostein, thanks for all the stimulating discussions about science and life, and making cigarette breaks less boring. Many thanks to Masahito, Micha, Shrinidi, Jon, Øistein, Kjetil, Trond, Leana, and to all my friends from the Austria Stammtisch in Oslo, especially Dirk, Martina, Maria, and Alex.

I am grateful to the Department of Theoretical Biology at the University of Vienna to always have a place to work for me when I was there. Thanks to Philipp, Gerd, Brian, Mihaela, and all the others for the amazing scientific discussions we had. Heidi and Elisabeth for all the support during the years. Hans Nemeschkal

SUMMARY

Investigating the relationship between genotype and phenotype is a topic of major interest in evolutionary biology. The mapping from genotype to phenotype shape the space that is explored through evolutionary processes when an organism evolves. Theoretical rules that specify the mapping are commonly known as the genotype-phenotype map.

This thesis introduces and expands on two different mathematical models of the genotype-phenotype map to study the paradoxical relationship between evolvability and robustness. Evolvability is the ability to respond to selection and produce heritable and selectable phenotypic variation. On the other hand, robustness is the ability to persist against perturbations. Based on that definition, a system cannot be evolvable and robust at the same time. However, evolvability and robustness are both important for the evolution of complex traits, which generates a paradox in evolutionary theory. How can organisms be evolvable, while maintaining robustness against perturbations?

Paper I resolves the paradox between evolvability and robustness using a network approach. Here, I use a Boolean genotype-phenotype map that defines genotypes and phenotypes as binary states comparable to an on-off switch. I demonstrate that the relationship between evolvability and robustness is dependent on the topology and structure of the map. I found a large amount of genotype-phenotype maps that gave a positive correlation between evolvability and robustness, but I also found maps with negative correlations. Negative correlations were more common in maps with higher degree of pleiotropy. In Papers II and III, I use a reaction-diffusion model as a mapping function for the genotype-phenotype map. The reaction-diffusion model is motivated by the development of butterfly eyespots in *Bicyclus anynana*. In Paper II, I illustrate that interactions between parameters in the model lead to sudden jumps in evolution, epistasis, and the emergence of novel structures. In Paper III, I demonstrate that evolvability and robustness are indeed negatively correlated on that particular genotype-phenotype map of the used reaction-diffusion model.

Finally in Paper IV, I introduce a statistical tool to analyze two-dimensional embryonic images to study interindividual differences in shape and tissue densities between different developmental stages of rainbow trout. This work resulted in a new method to quantify developmental variation as color and shape variations in images. The method allows to study the genotype-phenotype map experimentally by combining geometric morphometrics and image analysis. Such an approach allows us to quantify morphological change during development, study developmental variation, and the evolution of novelties.

LIST OF PAPERS

Paper I.   **Mayer C** and Hansen TF (2017). "Evolvability and Robustness: A paradox restored" In: *Journal of Theoretical Biology.* 430, pp. 78-85.

Paper II.   **Syrowatka C**, Le Rouzic A and Hansen TF (in prep). "Genotype-phenotype maps and developmental dynamics: insights from a simple reaction-diffusion model".

Paper III.   **Syrowatka C**, Le Rouzic A and Hansen TF (in prep). "Evolvability and robustness of a simple reaction-diffusion model".

Paper IV.   **Mayer C**, Metscher B, Mueller GB and Mitteroecker P (2014). "Studying developmental variation with geometric morphometric image analysis (GMIA)" In: *PLoS One.* 9(12).

Additional publications not directly linked to thesis topic:

- **Mayer C**, Windhager S, Schaefer K and Mitteroecker P (2017). "BMI and WHR are Reflected in Female Facial Shape and Texture: A Geometric Morphometric Image Analysis". In: *PLoS One.* 12(1).

- Tsuboi M, Kopperud BT, **Syrowatka C**, Grabowski M, Voje KL, Pélabon C and Hansen TF (2020). "Measuring Complex Morphological Traits with 3D Photogrammetry: A Case Study with Deer Antlers". In: *Evolutionary Biology.* 47, pp 175-186.

1 INTRODUCTION

The relationship between genotype and phenotype is a major topic of interest in evolutionary biology (Alberch, 1991; Gerhart and Kirschner, 1997; Maynard Smith, 1970; Oster and Alberch, 1982; Raff, 1996; A. Wagner, 2005; G. P. Wagner and Altenberg, 1996). The Modern Synthesis (Dobzhansky, 1937; Huxley, 1932; Mayr, 1982) proposed that genes determine the phenotype as a 'genetic blueprint' metaphor (Pigliucci, 2010), but the processes that connect them are neglected and treated as a 'black-box' (Hendrikse et al., 2007). However, it is necessary to understand how genotypes translate into phenotypes to broaden our understanding of evolution.

1.1 What is the Genotype-Phenotype Map?

The relationship between genotypes – an organism's genetic information – and phenotypes – an organism's 'observable' structure – can be defined by a set of theoretical rules that describe the process of the mapping between genotype space and phenotype space. This set of rules is commonly named genotype-phenotype map (Alberch, 1991; Hendrikse et al., 2007; Lewontin, 1970; Maynard Smith, 1970; Pigliucci, 2010). The genotype-phenotype map is a metaphor to describe the process that spans the space that is explored through evolution when an organism adapts to its environment (Wright, 1932, 1967). Genotypic information is passed onto the next generation as heritable information in accordance with Mendelian laws. The selection itself does not act on the level of the genotype, instead, it acts on the phenotype – the product that is defined by the genotype of an organism. The process in-between these two levels of the genotype-phenotype map is commonly argued to be development (fig. 1). However, development was long treated as a 'black-box' and not seen as an important factor to understand evolutionary processes (Arthur, 1997, 2000; Hall, 1999; Hall, 2003; Pigliucci, 2010). Genes do not directly generate phenotypes. Instead, they regulate gene expression relevant for a phenotypic structure to emerge, change enzyme activity

and structural proteins, etc. Hence, the mapping itself is a combination of genetic and developmental interactions (Alberch, 1991) defined by a genetic architecture that in turn determines the structure of the genotype-phenotype map.
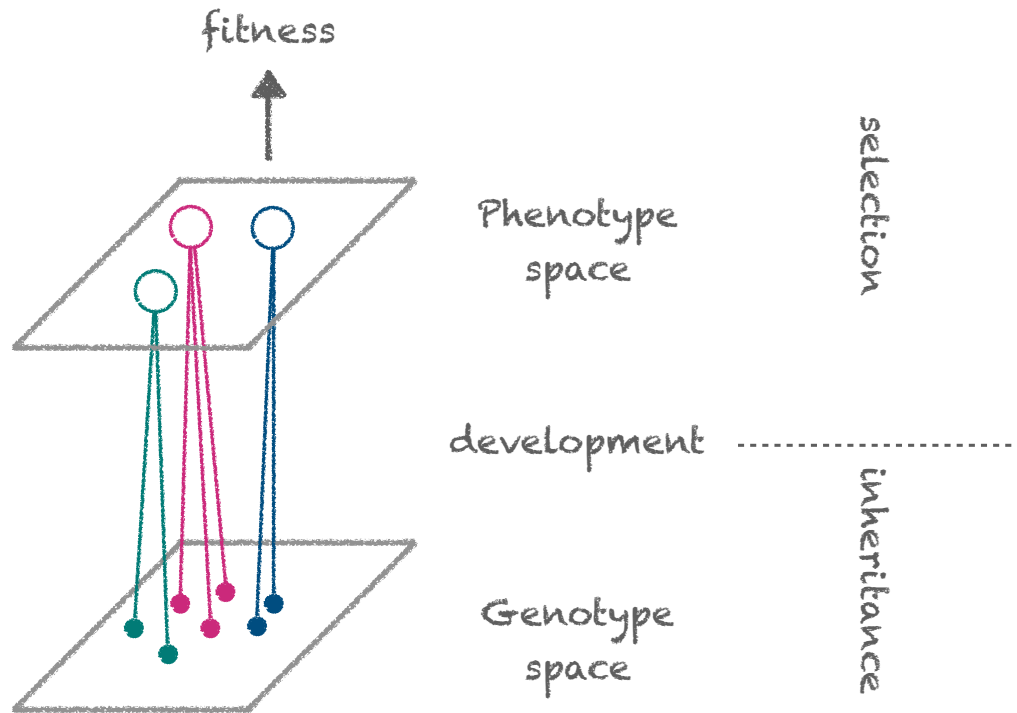
Figure 1: Illustration of the genotype-phenotype map. The filled circles represent different genotypes, the open circles represent different phenotypes and the lines in between denote which genotype realizes which phenotype. The colors correspond to a set of genotypes that realize one and the same phenotype (same color). The genotype-phenotype map is a set of functions that explains the relationship between genotype and phenotype space. Often this function is interpreted as a developmental process. The heritable information is processed through the genotype, but selection acts on the phenotype that determines the fitness of an organism.

Evolutionary biologists tend to assume that development does not play a major role (Laland et al., 2014; Mayr, 1961; Pigliucci and Müller, 2010; Wallace, 1986) for evolution and can be easily neglected. They assume that the genotype-phenotype map can best be approximated through a linear relationship (DiFrisco

8

and Jaeger, 2019; Fisher, 1930; Orr, 2000). Hence, phenotypic variation can be expressed as genotypic variation and evolutionary dynamics can be described by the latter one. On the contrary, evolutionary developmental biologists argue that the genotype-phenotype map is highly complex and non-linear (Alberch, 1991; DiFrisco and Jaeger, 2019; Gjuvsland, Vik, et al., 2013; Minelli and Fusco, 2012; Newman and Müller, 2000; Rice, 2002, 2004; Riedl, 1978; G. P. Wagner and Zhang, 2011). Here, the genotype-phenotype map is structured through developmental dynamics based on genetic and environmental interactions.
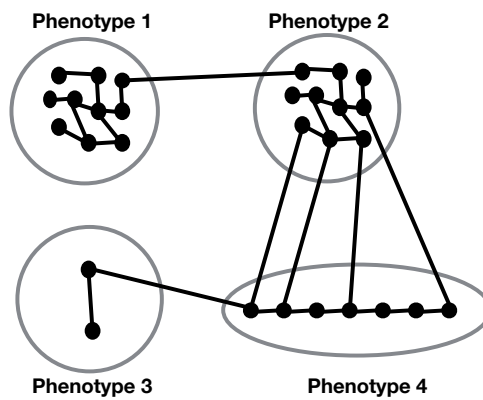
Figure 2: Schematic illustration of a genotype-phenotype map based on discrete genotypes and phenotypes. The black dots represent genotypes that are mutationally connected (black lines). A set of associated genotypes realise the same phenotype. Each of the grey open circles represent a phenotype.

Today the concept of the genotype-phenotype map is widely used in evolutionary biology to understand evolutionary processes and their underlying mechanisms and dynamics. There are different models used to understand the structure of the genotype-phenotype map in different systems, e.g. in RNA-secondary structure, transcription-factor binding sites, and computational models of evolution (Altenberg, 1994; Cotterell and Sharpe, 2010; Fontana, 2002; Greenbury et al., 2010; Payne, Moore, et al., 2014; Salazar-Ciudad and Marín-Riera, 2013; Schuster et al., 1994; A. Wagner, 2008). Many of these models have in common that the genotype-phenotype map consists of a network-like structure and assume that discrete genotypes map to discrete phenotypes. The idea behind that goes back to Maynard Smith's idea of protein spaces (Maynard Smith, 1970). He ar-

gued that proteins build a network that can be walked through by single mutational steps to evolve from one protein into another one. Each of the proteins occupies a connected area of the genotype space that makes it easier for an evolutionary process, driven by random mutations, to find them. The genotypes are mutationally connected with the underlying assumption that a certain set of genotypes realize the same protein or phenotype. Hence, that many genotypes map to a few phenotypes (fig. 2). The set of genotypes that realize one particular phenotype is called a neutral network (Fontana, 2002; Fontana et al., 1993; Gavrilets, 2004; Grüner et al., 1996; S. A. Kauffman, 1993; Schuster et al., 1994) or genotype network (Payne and A. Wagner, 2014; A. Wagner, 2011). Neutral, because mutations inside a certain neutral network do not change the phenotype.
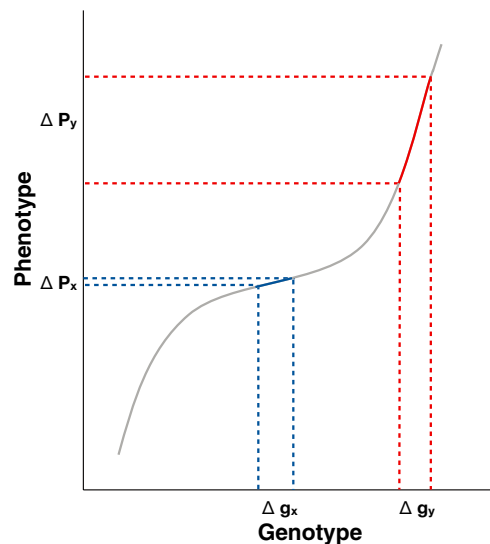
Figure 3: The function between the genotype and the phenotype defines the slope of the genotype-phenotype map. The larger the genetic effect on the phenotype, the steeper the slope. The absolute difference of the genotypes between the blue and the red areas is the same, but the effect on the phenotype is different. The change in genotype g_y has a larger effect on the phenotype than the change in genotype g_x, resulting in a steeper slope. (modified from Hansen, 2006).

The view of the genotype-phenotype map as a network that maps discrete characters to each other is somewhat simplistic. Many morphological quantitative

traits are continuous, like measurements of e.g. height, area, weight, and angles. Genotype-phenotype maps based on discrete characters cannot capture the properties that arise through the developmental processes of continuous traits. If we are interested in observing and quantifying traits that are on a continuous scale, we need to make different assumptions about the underlying genetic architecture to get different insights into properties that influence evolvability, robustness, and epistasis. Continuous morphological traits can be used to study the effect of a genetic change on a phenotype. The effect size of the genetic change defines the steepness of the slope of the map (fig. 3). Robusntess of a phenotype is determined through the slope that differs throughout the map. The slope is flatter in areas with more robust and canalized phenotypes, because the effect of a genetic change is smaller (Hansen, 2006).

The structure of the genotype-phenotype map is defined by its genetic architecture. However, we also need to take into account the impact of developmental dynamics on the relationship between genetic and phenotypic variation. Comparing genotype-phenotype maps based on phenotypes on different measurement scales and increasingly realistic assumptions about selection and adaptation will help us to understand how organisms are evolving and adapting to a constantly changing environment.

1.2 Waddington's Epigenetic Landscape: a Metaphor for using the Genotype-Phenotype Map

The illustration of the epigenetic landscape (Waddington, 1959; Waddington, 1957) can be a piece of the puzzle to understand the processes that connect the genotype and the phenotype and their underlying genetic architecture. The metaphor is useful to provide insights into biological processes that shape the surface of the genotype-phenotype map and to make inferences about properties like evolvability and robustness (Alberch, 1991; Jaeger and Monk, 2014).

The epigenetic landscape was Waddington's way to represent his concept of
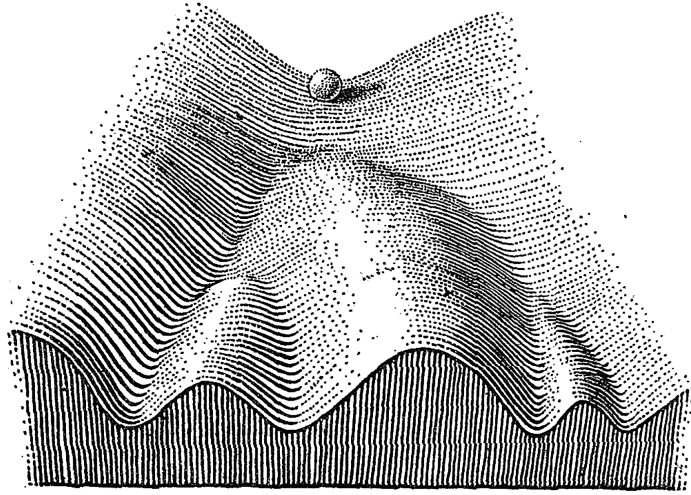
Figure 4: C. H. Waddingtons illustration of the epigenetic landscape. The cell fate is represented by a ball that rolls down a landscape of ridges and valleys. Depending on the way the ball goes the cell ends up in different states. (originally in Waddington, 1957)

embryonic development (fig. 4). The idea behind is the illustration of the way a cell takes to differentiate. The landscape is marked by ridges and valleys that represent the possible paths a cell can take. The cell itself is illustrated as a ball that is rolling down the landscape. The valleys are the developmental pathways or 'chreods', like Waddington called them, that lead the ball (or cell) to its final state of tissue. The surface of the landscape is shaped by the interaction of genes and their regulation that structure the valleys and ridges. The process of regulation through genes represents the canalization of the developmental process. Some mutations or environmental perturbations would be strong enough to push the ball over a ridge into another valley (Gilbert, 2000).

Waddington suggested that canalization evolved by natural selection and would be enhanced through evolution towards an optimum. Populations under stabilizing selection around an optimum tend to end up in an area of the genotype-phenotype map that is most robust against perturbations. There still would be plenty of un-

derlying genetic variation, but stabilizing selection would reduce the effect of the genetic variation onto the phenotype and stabilize the phenotypic outcome by reducing developmental noise (Ancel and Fontana, 2000; Gavrilets and Hastings, 1994; Rice, 1998; Siegal and Bergman, 2002; A. Wagner, 1996; G. P. Wagner, Booth, et al., 1997).

We need to expand on the currently used models (Altenberg, 1994; Cotterell and Sharpe, 2010; Fontana, 2002; Greenbury et al., 2010; Omholt, 2013; Payne, Moore, et al., 2014; Salazar-Ciudad and Marín-Riera, 2013; Schuster et al., 1994; A. Wagner, 2008) by defining the genetic architecture as developmental pathways. Then we will be able to explore genotype-phenotype maps based on developmental models and assess their impact on evolutionary processes.

1.3 Epistasis, Pleiotropy and the Landscape of the Genotype-Phenotype Map

Most quantitative morphological traits are affected by many genes (Barton and Turelli, 2004; Boyle et al., 2017; Hermisson et al., 2003; Phillips, 1998, 2008). Their combined genetic effect determines the topology of the genotype-phenotype map and with that influences evolvability and robustness measures (Carter et al., 2005).

Bateson et al. (1909) were the first to introduce the term epistasis. They used the term epistasis to explain that the effect of an allele was covered by the effect of another allele. The principle was later picked up again by Ronald A. Fisher (1918), who first used the term 'epistacy'. He argued that contributions of allelic effects on a phenotype are additive and that epistasis is the deviation from that additive effect.

Theoretical models of the genotype-phenotype map demonstrate that epistasis is a crucial factor in e.g. the evolution of sex and recombination (Barton, 1995; Kimura, 1956; Maynard Smith, 1978), in Muller's ratchet (Lynch et al., 1995), in models of speciation (Gavrilets, 2004; Templeton, 1979; Templeton, 1980) and

the evolution of canalization and robustness (Hansen, Álvarez-Castro, et al., 2006; Hermisson et al., 2003; Le Rouzic et al., 2013; A. Wagner, 2005). However, quantitative geneticists fail to show empirical evidence of epistasis (Barton and Turelli, 2004; Cheverud and Routman, 1996; Goodnight, 1988; Hansen, 2013).

The debate about the evolutionary importance of epistasis originated in the discussions between Ronald A. Fisher and Sewall Wright (Fisher, 1918; Wright, 1931). Fisher argued that the effect of epistasis is irrelevant to evolution. In large populations, the epistatic effect would average out over the population and can be treated as statistical noise. On the contrary, epistasis is known to play a major role in shaping "Wrightian" adaptive landscapes by introducing multiple fitness peaks that complicate evolutionary dynamics (Cheverud and Routman, 1995; Hansen, 2013; S. Kauffman and Levin, 1987; Wright, 1932; Wright, 1980). The multiple peaks are generated through intermediate effects between interacting genes that have larger fitness than their independent effects. In a multipeak landscape a population tends to evolve to the closest fitness peak that would necessarily not be the one with the highest fitness. To solve the problem of a population crossing a fitness valley to a higher peak, Wright introduced the shifting-balance theory that consists of a combination of genetic drift, natural selection, and migration (Wright, 1932).

Quantitative geneticists base their assumptions on the infinitesimal model and argue that epistasis has only weak effects on natural and artificial selection of quantitative traits. The model assumes that most quantitative traits are built upon an infinitely large number of loci where each locus has an infinitely small effect on the phenotype. Compared to genetic drift selection on each of the loci is weak enough to not affect response to selection, which is also true for interaction effects (Barton, 2017; Barton and Turelli, 2004; Cheverud and Routman, 1995). The argument is based on Ronald A. Fisher's (1918) statistical model of epistasis. Such models are based on statistical properties as population averages and regression coefficients. Epistasis is approximated as interaction terms in regressions on allelic effects in a population and therefore, epistasis is a population property (Cockerham, 1954; Kempthorne, 1954).

14

Cheverud and Routman (1995) introduced a physiological model of epistasis that measured the effect of a genetic change regardless of the population they are measured in. Hansen and G.P. Wagner (2001) extended the physiological model and called it the functional model. Here, a genetic effect is dependent on its genetic background and relative to a reference genotype. To assess the evolutionary importance of epistasis they introduced a model of a multilinear genotype-phenotype map (multilinear model).

The multilinear model describes the phenotype z as a multilinear function of genotypic reference effects $g = \{^1y, ^2y, ..^ny\}$, where the reference ^iy is defined as a phenotypic effect of a genetic change in locus i. For a one trait one locus case this gives

$$z = z_r + {}^1y, \tag{1}$$

where z is the resulting phenotype, z_r the reference phenotype and ^1y the genetic effect of a change in locus 1. In the case of one trait two loci, we need to add an interaction term – the epistasis coefficient ε –

$$z = z_r + {}^1y + {}^2y + {}^{12}\varepsilon\, {}^1y\, {}^2y \tag{2}$$

or as general formula for n loci

$$z = z_r + \sum_i {}^iy + \sum_i \sum_{j>i} {}^{ij}\varepsilon\, {}^iy\, {}^jy. \tag{3}$$

The epistasis coefficient is used to quantify the strength and directionality of epistasis between the corresponding loci. It is defined in inverse phenotypic units

15

and difficult to interpret and compare between different traits. Therefore, Hansen and Wagner (2001) suggest to use the unitless epistatic factor that describes how much a genetic change in locus 1 is affected by a genetic change in locus 2 and vice versa. The epistatic factor is formulated as

$$f_1 = 1 + {}^{12}\varepsilon^2 y \qquad (4)$$

$$f_2 = 1 + {}^{12}\varepsilon^1 y. \qquad (5)$$

In case $f = 1$ there is no epistasis, $f < 1$ negative epistasis, $f > 1$ positive epistasis and $f < 0$ sign epistasis. Positive epistasis increases the genetic effect compared to the additive effect in the presence of another genetic change and negative epistasis decreases the genetic effect. Sign epistasis changes the sign of the effect, e.g. a deleterious mutation can increase the beneficial effect of another mutation.

However, we need to consider a third model of epistasis - called mechanistic or mechanical epistasis (DiFrisco and Jaeger, 2019; Hansen, 2008; G. Wagner et al., 2000). Here the genotype-phenotype map is defined through models of developmental or physiological systems, gene-regulatory networks or metabolic control systems (Bagheri-Chaichian et al., 2003; Gjuvsland, Hayes, et al., 2007; Jernvall, 2000; Salazar-Ciudad and Jernvall, 2004; A. Wagner, 1996).

The development of quantitative traits requires complex underlying genetic architectures consisting of interacting genes that build complex regulatory networks (Mackay, 2014). Thus, complex quantitative traits can only develop as a result of many genes interacting with each other and not based on a single gene. The complex underlying genetic architecture will generate large amounts of epistasis. We need to understand the patterns of epistasis that are generated through developmental systems of complex quantitative traits to enhance our knowledge on epistasis affecting evolutionary dynamics.

The landscape of the genotype-phenotype map is not only shaped through

epistatic interactions of genes, but also by pleiotropy and the fact that one gene can affect many traits. Ronald A. Fisher (1930) explained the effect of pleiotropy on evolvability in his geometric model with a metaphor. He interpreted mutations as the tuning of a microscope, whereas large changes have only a small probability to improve the image. In a phenotype with many traits mutations with smaller effect sizes have a higher probability to bring the organism towards his fitness optimum. Fisher concluded that a phenotype that consists of more traits needs smaller mutations to improve.

1.4   The Paradox: Evolvability and Robustness

Evolvability is the ability to respond to selection and adapt to changing environments through mutation. Robustness, on the other hand, is the ability to persist against perturbations as mutations. Evolvability and robustness are both thought to be properties that are essential for evolving complex traits (Conrad, 1990; Gerhart and Kirschner, 1997; S. A. Kauffman, 1993; A. Wagner, 2005), but they are contradictory (Lesne, 2008; Visser et al., 2003). Neutral mutations that occur in highly robust phenotypes can be hidden from selection and induce the accumulation of hidden genetic variation that is not accessible for selection. If the selection pressure on the phenotype changes – e.g. a changing environment – hidden genetic variation can suddenly be available for selection. This sudden increase in variation can facilitate evolvability and is the reason for the argument that organisms need to be highly evolable and robust at the same time to evolve complex phenotypes (Conrad, 1990; Gerhart and Kirschner, 1997; S. A. Kauffman, 1993; Masel and Trotter, 2010; A. Wagner, 2005). However, this is a paradox in itself, because based on the definition of evolvability and robustness a phenotype cannot be robust and evolable at the same time.

Andreas Wagner (2008) resolved the paradox of the relationship between evolvability and robustness by arguing that the paradox originated by thinking of evolvability and robustness as properties of the genotype. Here, evolvability and robustness indeed conflict with each other. By his definition a mutation of the genotype

17

either changes the phenotype or not. He argued that the paradox can be resolved by seeing evolvability and robustness as a property of the phenotype and not as a property of the genotype. Andreas Wagner (2008) argued that most phenotypes would have large and highly connected underlying neutral networks. Such phenotypes would consist of many different genotypes so that a mutation in a particular genotype would often result in the same phenotype. On the other hand, such phenotypes are connected with many different neighboring neutral networks – other phenotypes – and mutations in many cases would indeed change the phenotype. Hence, such phenotypes would be highly evolvable, and at the same time highly robust (fig. 5, upper left corner). He could confirm his hypothesis using simulation studies of RNA-secondary structure, where he defined the secondary structure as phenotype and its nucleotide sequence as genotype. Nucleotide sequences that realize the same secondary structure define a neutral network.

This argument is based on the structure and topology of the genotype-phenotype map and is only true for either large neutral networks that are highly connected and have many neighbors or small neutral networks that are poorly connected (fig. 5). Here, the relationship between evolvability and robustness would be positive. However, theoretically one could think about structures of the neutral network that lead to negative relationships between evolvability and robustness. The network could either be poorly connected and have many neighbors (fig. 5 lower-left corner) or highly connected but not have many neighbors (fig. 5 upper right corner). In other words, the structure of the genotype-phenotype map influences the relationship between evolvability and robustness. Thus, the relationship between evolvability and robustness of phenotypes is not necessarily positive and instead defined by the structure and topology of the underlying genotype-phenotype map.

The above assumptions and consequences only hold for network-like structured genotype-phenotype maps that are based on discrete characters. They do not describe the complexity that exists in most quantitative morphological traits of continuous character. Small changes in a genotype of a continuous character can have a huge effect on the phenotype itself and its fitness. These models also neglect the impact of developmental dynamics on the genotype-phenotype map

and its relationship between evolvability and robustness.

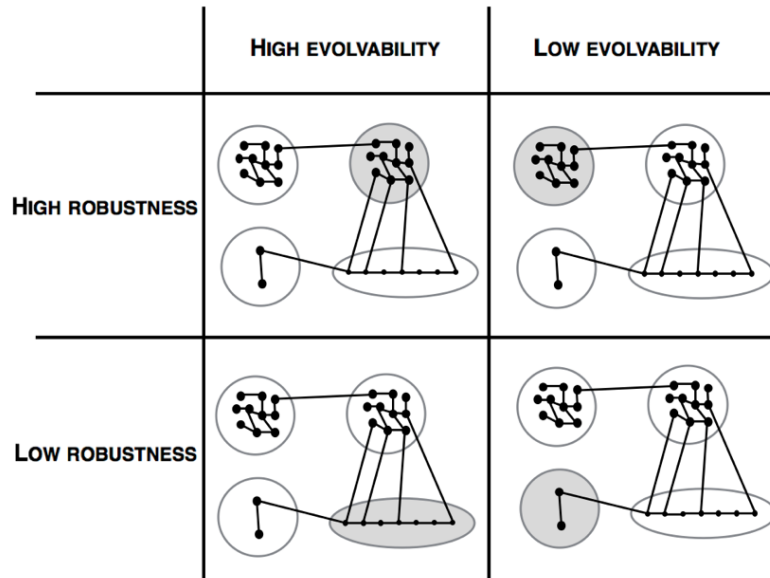Figure 5: Evolvability and robustness of a phenotype (circles) depend on the size of its underlying neutral network (black dots are genotypes) and the number of its neighboring phenotypes. In each quadrant evolvability and robustness are assigned to the phenotype colored in grey, e.g. for the upper left quadrant the phenotype marked as a filled grey circle is highly evolvable and highly robust.

2 THE CONTRIBUTION OF THIS THESIS

Evolvability and robustness are argued to be important to evolve complex organisms (Conrad, 1990; Gerhart and Kirschner, 1997; S. A. Kauffman, 1993; A. Wagner, 2005) and are influenced by the topology and structure of the genotype-phenotype map. Our knowledge about epistasis and pleiotropy is essential to gain insight into the mechanisms that structure the genotype-phenotype map. To make general implications about the epistatic and pleiotropic structure, we need to study different representations of the genotype-phenotype map (Hansen, 2013).

Several studies of the relationship between evolvability and robustness analyze the structure and topology of genotype-phenotype maps (Altenberg, 1994; Fontana, 2002; S. Kauffman and Levin, 1987; Payne and A. Wagner, 2014; P. F. Stadler and B. M. R. Stadler, 2006; A. Wagner, 2008). Many of them are based on discrete characters in a neutral space and use examples of special cases to make general implications. Nevertheless, studies of genotype-phenotype maps based on developmental models remain underrepresented.

The focus of my thesis is to assess the properties of the genotype-phenotype map that influence the relationship between evolvability and robustness. Compared to the studies mentioned above I will use a general abstract representation of the map, a map that is based on a developmental model, and morphological data to solve the following questions:

1. *Does robustness facilitate evolvability?* **Paper I**
   I will introduce a general abstract model of the genotype-phenotype map to challenge Andreas Wagner's (2008) results on evolvability and robustness.

2. *Is the relationship between evolvability and robustness the same for a genotype-phenotype map that is based on a developmental process?* **Paper II & III**
   I will present a framework to study a realistic genotype-phenotype map based on a developmental model to investigate the structure and topology of the map and examine if robustness does facilitate evolvability.

3. *How does genetic variation translates into phenotypic variation?* **Paper IV**
   I will introduce a method to quantify the developmental variation of experimentally gained morphological data to study a real genotype-phenotype map.

To approach the above questions I expand the genotype-phenotype map literature by using two different mathematical models to study the structure of the genotype-phenotype map that influences properties like evolvability, robustness, epistasis and pleiotropy and how they are connected. The first model is based on a Boolean genotype-phenotype map that describes a static network. The second model is based on a simple reaction-diffusion model of a dynamic genotype-phenotype map that mimics development. These two models allow us to explore the topology of the landscape in two different types of genotype-phenotype maps that are on different measurement scales. The Boolean genotype-phenotype map defines genotypes and phenotypes on a discrete scale as binary states, whereas the map built upon the reaction-diffusion model defines genotypes and phenotypes on a continuous scale as quantitative morphological traits. Last, I introduce a method called Geometric Morphometric Image Analysis (GMIA) to quantify the developmental variation of morphological data to study how genetic variation maps onto phenotypic variation.

3 THEORY

The key to understand the topology and structure of the genotype-phenotype map is to understand the genetic architecture that defines the mapping of the genotype space to the phenotype space. The existing challenge consists of developing a framework to study the different properties of the genotype-phenotype map in different contexts. It is not possible to design one single framework or model to study all possible genotype-phenotype maps. We need to distinguish between network-like maps (Aldana et al., 2007; Draghi et al., 2010; Payne and A. Wagner, 2014; A. Wagner, 2008) and dynamic maps that are defined through biological processes (Jaeger, Irons, et al., 2012; Milocco and Salazar-Ciudad, 2020; Salazar-Ciudad and Marín-Riera, 2013; Verd et al., 2014). However, both of these types of genotype-phenotype maps have in common that the basic approach to define their underlying genetic architecture consists of the definition of two parameter spaces – genotype space and phenotype space – and a function that connects both of these spaces. Therefore, the genotype-phenotype map can be defined as a mathematical function linking the phenotype and genotype (fig. 6).

3.1 Boolean Genotype-Phenotype Maps

The Boolean genotype-phenotype map is characterized through discrete genotypes and phenotypes that consist of vectors of Boolean variables – coded as 0 and 1. Such coding can be interpreted as true and false or on and off, e.g. of gene expression or a trait that is either present or absent. The mapping function itself is based on a combination of Boolean logic operators like AND, OR or NOT (fig. 7). Boolean algebra is a convenient way to model abstract forms of genotype-phenotype maps to make general conclusions about their properties (Aldana et al., 2007; Ebner et al., 2001; Espinosa-Soto et al., 2004; Frank, 1999; S. A. Kauffman, 1969, 1993).

Boolean genotype-phenotype maps can be used to explore the properties of a
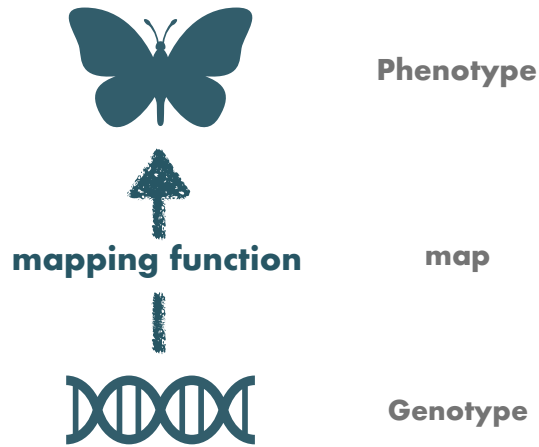
Figure 6: The genotype-phenotype map is the connection between genotype and phenotype space through a mapping function.

network-like structured map with discrete characters. In **Paper I** I used Boolean maps to explore the relationship between evolvability and robustness in maps with different degrees of pleiotropic structure. This setup allows for an exhaustive analysis of genotype-phenotype maps with different topologies.

The Boolean genotype-phenotype map is also a special case of the multilinear epistatic model of Hansen and G.P. Wagner (2001), where the logic operator can be expressed as a multilinear form in Boolean variables. The multilinear form of the three Boolean operators AND, OR and NOT are defined as follows:

$$AND \quad z = {}^1y{}^2y{}^3y...{}^ny \tag{6}$$

$$OR \quad z = {}^1y + {}^2y + {}^3y - {}^1y{}^2y - {}^1y{}^3y - {}^2y{}^3y + {}^1y{}^2y{}^3y \tag{7}$$

$$NOT \quad z = 1 - y \tag{8}$$

Here, z is the phenotype and y is the genotype locus. The *AND* operation results in in a phenotype z that is 1, if all the genotype loci y are true or 1. By using the *OR* operator at least one of the genotype loci y need to be 1 to result in a phenotype z that is 1. The *NOT* operator changes the phenotype z from either 1 to 0 or from 0 to 1. The advantage of expressing the Boolean genotype-phenotype map in the multilinear framework is the possibility to explore properties like epistasis.

0 1 1   **Phenotype**

Boolean logic operators   **map**

0 1 1 0 0 1   **Genotype**
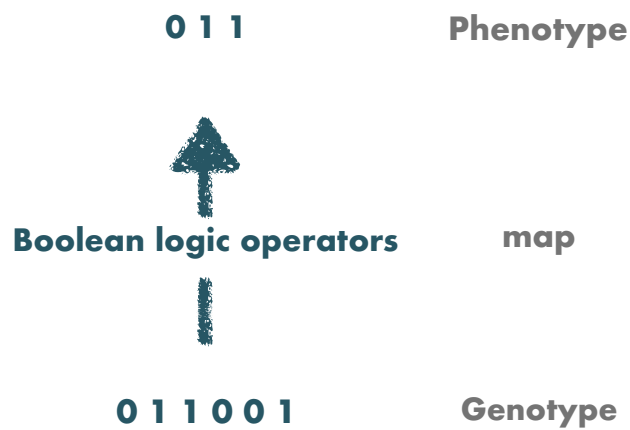
Figure 7: The Boolean genotype-phenotype map is defined through a genotype and a phenotype that consist of Boolean variables. The mapping function are Boolean logic operators – like AND, OR and NOT.

3.2 Reaction-Diffusion Models as Genotype-Phenotype Map

Alan Turing (1952) proposed a theory of morphogenesis to study how natural patterns are generated out of a homogenous state. Today the theory of Turing patterns is widely used to understand embryonic development (Jaeger, Surkova,

et al., 2004; Kondo and Miura, 2010; Kondo and Asai, 1995; Lange et al., 2018; Murray, 2002; Raspopovic et al., 2014; Salazar-Ciudad and Jernvall, 2004). Turing's idea is based on reaction-diffusion models, where two morphogenes diffuse over an area and react with each other. Here, one reagent acts as an activator and the other one as an inhibitor. Both initially have a homogenous concentration that over time changes to a periodic pattern like stripes, spots, and spirals (Meinhardt, 1983; Murray, 2002).
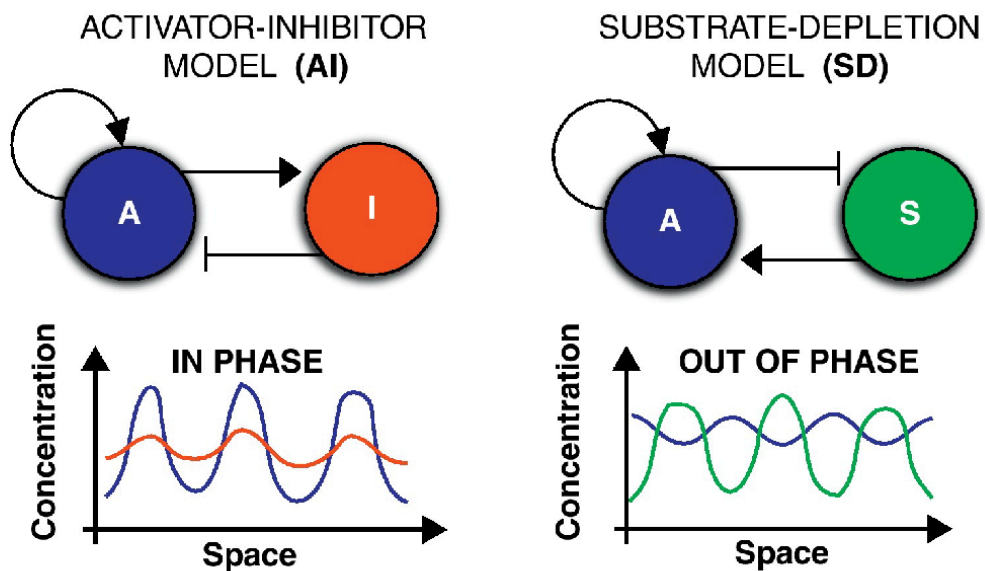
Figure 8: Two mechanisms that generate Turing patterns. The activator-inhibitor model (A = activator, I = inhibitor) on the left and the substrate-depletion model (A = activator, S = substrate) on the right (originally in Marcon & Sharpe, 2012).

Marcon and Sharpe (2012) distinguish two different types of interaction networks that can generate such periodic patterns (fig. 8) – the activator-inhibitor model and the substrate-depletion model. The first is characterized through an activator that activates itself and its inhibitor to form a periodic pattern where both morphogens are in phase. In contrast, the second has an activator that consumes

the substrate to activate itself and forms a periodic pattern where both morphogenes are out of phase with each other. Both types of models can be described using a system of nonlinear partial differential equations. For two morphogenes in two-dimensions the general form is:

$$\frac{\partial A}{\partial t} = f(A, I) + D_A \left(\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} \right) \tag{9}$$

$$\frac{\partial I}{\partial t} = g(A, I) + D_I \left(\frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \right) \tag{10}$$

In eq. 9 and eq. 10, A and I are the two morphogenes. The first part of the equations $\frac{\partial A}{\partial t}$ and $\frac{\partial I}{\partial t}$ describe the change of the concentration of the morphogenes over time. The second part $f(A, I)$ and $g(A, I)$ are functions of the morphogenes that define the reaction between both of them. The last part $D_A(\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2})$ and $D_I(\frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2})$ describes the diffusion of the morphogenes over an area, where D_A and D_I are the diffusion constants that give the speed of the diffusion for A and I, respectively. The width and length of the area are defined by x and y.

In **Paper II** and **Paper III** I use a modified version of a substrate-depletion model (Connahs et al., 2019) that is motivated by the development of the center of an eyespot in the butterfly *Bicyclus anynana* (fig. 9). The model is defined as the following:

$$\frac{\partial U}{\partial t} = KU^2V - k_U U + D_U \left(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \right) + \sigma \frac{\partial \varepsilon}{\partial t} \tag{11}$$

$$\frac{\partial V}{\partial t} = \alpha - KU^2V - k_V V + D_V \left(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} \right) + \sigma \frac{\partial \varepsilon}{\partial t} \tag{12}$$

In the above equation, U is the activator, and V is the substrate that gets consumed by U. In eq. 11 and eq. 12 the terms KU^2V and $\alpha - KU^2V$ are the reaction terms and $D_U(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2})$ and $D_V(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2})$ are the diffusion terms.
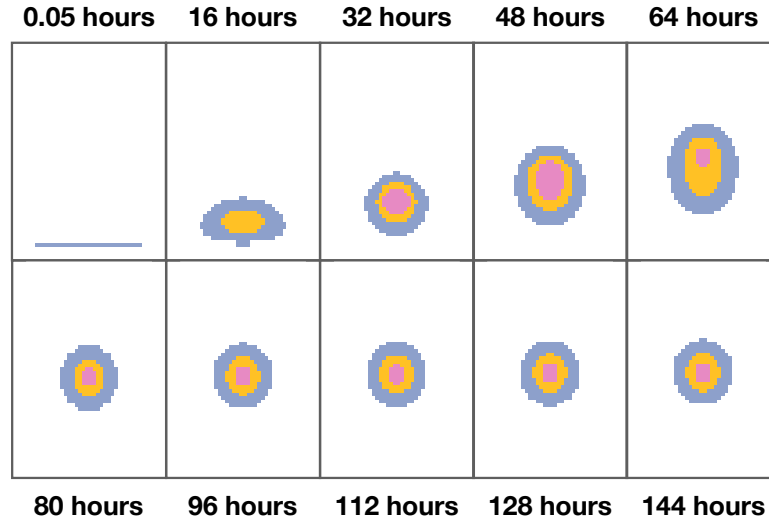
Figure 9: Example development of the butterfly eyespot generated through the reaction-diffusion model.

The degradation rates of the morphogenes are denoted by the expressions $k_U U$ and $k_V V$. The last term $\sigma \frac{\partial \varepsilon}{\partial t}$ describes developmental noise as a stochastic change per time step, where $\partial \varepsilon$ is an independent, normally distributed random variable with mean zero and variance ∂t. The σ denotes the standard deviation of the stochastic changes in the model.

Reaction-diffusion models can be used to model actual developmental processes to map the genotype onto the phenotype and hence they are dynamic (fig. 10). The genotype is defined through the model parameters – in our case α, K, k_U, k_V, D_U, and D_V. The mathematical description of the model maps the parameters to the phenotype through the modeled process over time. The resulting phenotype is a quantitative morphological trait: the relative size of the concentric rings that define the butterfly eyespot in our case.

Phenotype

$$\frac{\partial U}{\partial t} = \quad KU^2V - k_U U \ + D_U \Big(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \Big)$$
$$\frac{\partial V}{\partial t} = \alpha - KU^2V - k_V V \ + D_V \Big(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} \Big)$$

map

K α **k$_U$ k$_V$ D$_U$ D$_V$**

Genotype

Figure 10: The genotype-phenotype map that is generated by the reaction-diffusion model is defined through a genotype that consists of the parameters of the model and a phenotype that is an actual quantitative morphological trait – here a butterfly eyespot in *Bicyclus anynana*.

4 THE PAPERS

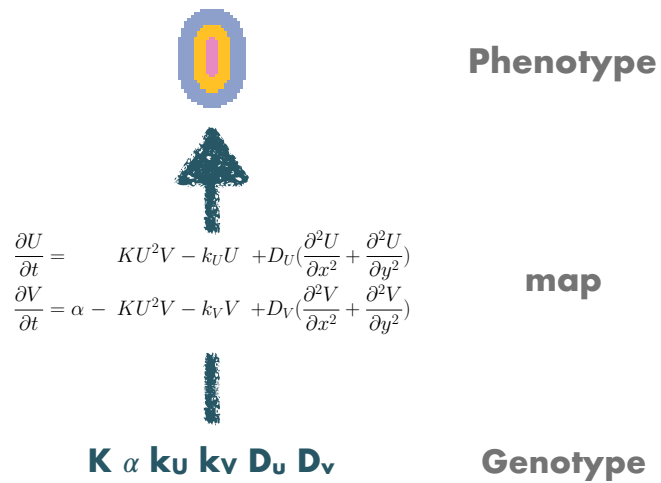## 4.1 Paper I: Evolvability and robustness: A paradox restored

Evolvability and robustness are crucial for the origin and maintenance of complex organisms, but may not be simultaneously achievable as robust traits are also hard to change. Andreas Wagner has proposed a solution to this paradox by arguing that the many-to-few aspect of genotype-phenotype maps creates neutral networks of genotypes coding for the same phenotype. Phenotypes with large networks are genetically robust, but they may also have more neighboring phenotypes and thus higher evolvability. In this paper, we explore the generality of this idea by sampling large numbers of random genotype-phenotype maps for Boolean genotypes and phenotypes. We show that there is indeed a preponderance of positive correlations between the evolvability and robustness of phenotypes within a genotype-phenotype map, but also that there are negative correlations between average evolvability and robustness across maps. We interpret this as predicting a positive correlation across the phenotypic states of a character, but a negative correlation across characters. We also argue that evolvability and robustness tend to be negatively correlated when phenotypes are measured on ordinal or higher scale types. We conclude that Wagner's conjecture of a positive relation between robustness and evolvability is based on strict and somewhat unrealistic biological assumptions.

4.2 Paper II: Genotype-phenotype maps and developmental dynamics: Insights from a simple reaction-diffusion model

The variational properties of development determine the highly complex and non-linear structure of the genotype-phenotype map. However, models in quantitative genetics do not account for the influence of developmental dynamics on the structure of the genotype-phenotype map. Statistical representations of the

genotype-phenotype map cannot capture properties like sudden jumps in evolution, generation of epistasis, and the emergence of novelties. Reaction-diffusion models are a powerful tool to study developmental processes and their variation. Here, we are using a reaction-diffusion model as a mechanical representation of the genotype-phenotype map that describes the development of eyespots on the wings in *Bicyclus anynana*. We can demonstrate that even if the genotype space is continuous, the phenotype space shows discontinuities, especially towards the boundaries. We conclude that a mechanistic representation of the genotype-phenotype map can be used to find sudden jumps in evolution, epistatic patterns and the emergence of novelties.

4.3   Paper III: Evolvability and robustness of a simple reaction-diffusion model

Evolvability is the ability to respond to selection through heritable and selectable phenotypic variation, while robustness is the ability of a phenotype to persist against perturbations. Hence, evolvability and robustness may conflict. They are, however, both important for the evolution of complex phenotypes. Development plays an important role in determining how genetic variation translates into phenotypic variation and thus affecting the relationship between evolvability and robustness. We use a simple reaction-diffusion process to model the development of eyespots in the butterfly species *Bicyclus anynana* to explore the relationship between evolvability and robustness on a biological realistic genotype-phenotype map. We investigate the quantitative morphological change of eyespots under selection and show that evolvability and robustness are negatively correlated.

4.4 Paper IV: Studying developmental variation with geometric morphometric image analysis (GMIA)

The ways in which embryo development can vary across individuals of a population determine how genetic variation translates into adult phenotypic variation. The study of developmental variation has been hampered by the lack of quantitative methods for the joint analysis of embryo shape and the spatial distribution of cellular activity within the developing embryo geometry. By drawing from the strength of geometric morphometrics and pixel/voxel-based image analysis, we present a new approach for the biometric analysis of two-dimensional and three-dimensional embryonic images. Well-differentiated structures are described in terms of their shape, whereas structures with diffuse boundaries, such as emerging cell condensations or molecular gradients, are described as spatial patterns of intensities. We applied this approach to microscopic images of the tail fins of larval and juvenile rainbow trout. Inter-individual variation of shape and cell density was found highly spatially structured across the tail fin and temporally dynamic throughout the investigated period.

5 DISCUSSION

The paradox of the relationship between evolvability and robustness is still a puzzling question in evolutionary biology. Establishing an understanding of the nature of the paradox is one of the key solutions to gain insight into the evolution of complex organisms. Is it necessary to be highly evolvable and highly robust to evolve complex organisms? How can it be that an organism is highly evolvable and therefore, easy to change by a mutation and at the same time highly robust, and not easy to change by a mutation?

5.1 Network-like Genotype-Phenotype Maps

Andreas Wagner (2008) showed in his paper 'Robustness and evolvability: a paradox resolved' that robustness indeed enhances evolvability when thinking about them as properties of the phenotype. However, he could not find the same results on the level of the genotype, where evolvability and robustness contradict each other.

A negative relationship between the so-called genotype evolvability and genotype robustness lies already in Andreas Wagner's (2008) definition of them. Robust mutations are not changing the phenotype and the rest of possible mutations change the phenotype. This will always give a negative correlation because the more robust mutations are possible, the less evolvable mutations are possible and vice versa. This argument is not as clear for phenotype evolvability and phenotype robustness. I made the theoretical argument that the relationship between evolvability and robustness of phenotypes depends on the topology of the underlying genotype-phenotype network. The positive correlation between them only holds for phenotypes that consist of large and highly connected neutral networks as found in RNA-secondary structure or TF-binding sites (Aldana et al., 2007; Draghi et al., 2010; Payne, Moore, et al., 2014; A. Wagner, 2008). Mutations on such large neutral networks would often stay on the network. Yet, they would also
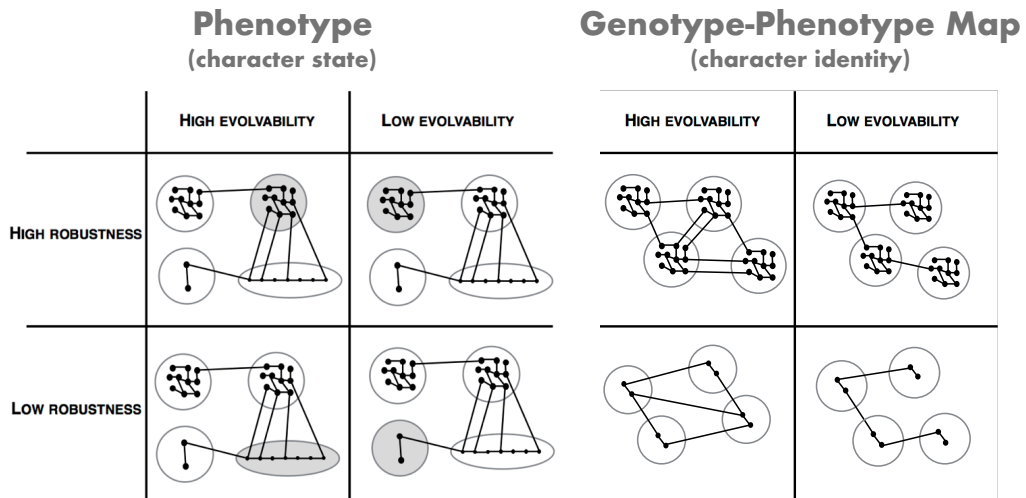
Figure 11: Evolvability and robustness can be determined by the topology of either the phenotype (left) or the map (right). The size of the neutral network (black dots) and its neighboring phenotypes determine evolvability and robustness of the phenotype, whereas evolvability and robustness of the map are expressed as their average overall phenotypes.

often leave the network and hence, change the phenotype, because large networks tend to have more neighbors (fig. 11 – left panel, upper left corner). Thus, the relationship between evolvability and robustness is defined through the topology of the genotype-phenotype map. Theoretically, it is possible to construct large neutral networks with only a few phenotypic neighbors (fig. 11 – left panel, upper right corner) or small neutral networks with many neighbors (fig. 11 – left panel, lower left corner).

Based on the above theoretical considerations I challenge Andreas Wagner's (2008) general results on a positive correlation between evolvability and robustness. In **Paper I** I establish a mathematical model of a Boolean genotype-phenotype map and show that Andreas Wagner's (2008) result is an artifact of the structure of the genotype-phenotype map found in systems of RNA-secondary structures. Interestingly, I indeed found a large amount of genotype-phenotype maps that gave

a positive correlation between phenotype evolvability and phenotype robustness. I also found maps with negative correlations (up to 10%) between phenotype evolvability and robustness. Negative correlations were more common in genotype-phenotype maps with complex topologies and larger amounts of pleiotropy.

It is argued (Payne, Moore, et al., 2014; Schuster et al., 1994; A. Wagner, 2012) that genotype-phenotype maps often consist of one large and highly connected neutral network that has many neighbors. These neighbors are mostly smaller neutral networks that themselves have fewer neighbors. By considering the average phenotype evolvability and phenotype robustness over a genotype-phenotype map, we avoid the overrepresentation of one large neutral network. The average phenotype evolvability and phenotype robustness is a property of the genotype-phenotype map itself and can be used to directly compare different structures and topologies. Here, I found consistently negative relationships across the used averages (fig. 11, right panel). Therefore, evolvability and robustness indeed tend to be positively correlated across phenotypic states but are negatively correlated across characters. This argument holds for network-like genotype-phenotype maps based on discrete phenotypes.

5.2 Genotype-Phenotype Maps of Development

How genetic variation maps onto phenotypic variation is crucial for the understanding of the structure of the genotype-phenotype map. Development is the mechanism that defines how genotypes and phenotypes are connected (Alberch, 1991; Pigliucci and Müller, 2010). Phenotypes are products of their development, and evolutionary change happens through modifying developmental mechanisms (Amundson, 2005; Brigandt, 2015). Despite, the obvious importance of development for evolutionary theory, the Modern Synthesis assumes that 'black-boxing' of developmental processes does not cause any explanatory problems for evolutionary processes (Hendrikse et al., 2007; Mayr, 1961; Pigliucci, 2010; Wallace, 1986). This argument is based on the assumption that the genotype-phenotype map is linear and therefore, the effect of development on its structure can be ne-

glected (DiFrisco and Jaeger, 2019). However, we need to understand how developmental processes shape the genotype-phenotype map to gain insight into their impact on evolutionary change.

In **Paper II** and **Paper III** I introduce and expand a model of the genotype-phenotype map that is based on a developmental process – the development of eyespots on the wing of *Bicyclus anynana*. The developmental process is implemented as reaction-diffusion processes. I used this model to explore and investigate the structure of the genotype-phenotype map. Followed, by asking how evolvability and robustness can be measured and how they are related to each other. The differences of that model to the Boolean genotype-phenotype map is the use of a biological realistic genotype-phenotype map of continuous traits.

In **Paper II** I investigate the impact developmental interactions may have on the topology of the genotype-phenotype map and thereupon, on evolutionary dynamics. Polly (2008) argued that three consequences would arise through neglecting developmental interactions. First, that even if the genotype is quasi-continuous the phenotypic change can be discontinuous and lead to a rugged genotype-phenotype map and sudden jumps in evolution. Second, developmental interactions lead to epistasis. Third, developmental interactions produce novelties – which can be as well defined as the loss or gain of structures (Müller and G. P. Wagner, 1991). I indeed found evidence for all of them, but mostly at the boundary of the phenotype space. The first consequence was not as clear and mostly only found in the rings the eyespot consisted of. I defined the rings as relative to the total size that caused constraints between them. In case if one ring got larger, one of the other rings had to get smaller. This causes a slightly rugged pattern of the genotype-phenotype map. I could find epistatic patterns as proposed in the second consequence as deviations from additivity. This implies that continuous genotypic changes do not map additively onto the phenotype. I could as well demonstrate the emergence of novel patterns as the gain and loss of rings.

The results of **Paper II** indicate that the studied genotype-phenotype map is complex and non-linear. However, the genotype-phenotype map appears most non-linear at the boundaries of the investigated phenotype space. These non-

linearites complicate the prediction of evolution (Milocco and Salazar-Ciudad, 2020).

In **Paper III** I used the same genotype-phenotype map to explore the relationship between evolvability and robustness. First, I investigated the three rings of the eyespot and its total size under directional selection. Second, I examined the eyespot as a whole structure under stabilizing selection. Locally evolvability is defined as mutational evolvability and should shed light on the influence of the local structure of the genotype-phenotype map around an eyespot and short-term evolution. Globally evolvability is defined as realized evolvability and should give insight into the connection between robustness and long-term evolution.

As hypothesized I found negative correlations between robustness and evolvability throughout all analyses. The result for mutational evolvability and robustness is determined through its definition. Robustness is the average change caused by a mutation, except for the mutation that caused the largest fitness advantage. The change that is caused by the mutation that gives the largest fitness advantage is mutational evolvability. This definition implies a negative correlation between evolvability and robustness, arguably similar to genotype evolvability and genotype robustness defined by A. Wagner (2008). Additionally, robustness and realized evolvability are negatively correlated. This finding confirms the results of **Paper I**, where I could show that robustness hinders the evolvability of sequential adaptations.

The main implications for evolvability and robustness are holding for the particular model I used. There are limitations to the used simulations. It could be argued that I used only one particular phenotype and ignored the surrounding phenotypes, which makes it difficult to compare it with the original results of A. Wagner (2008). Of course, there is potential for future exploration and to extend the analysis to the whole morphospace. Capturing the gain and loss of traits still causes difficulties using modern morphometrics, because morphospaces are constructed by variables that appear in all phenotypes. This does not capture the emergence or loss of traits (Polly, 2008).

5.3 Quantifying Embryological Development

Modern developmental biology has led to a large body of knowledge about the qualitative effects of developmental genes and their regulatory function in the formation of tissue and organs. Much less is known about the quantitative properties of development. Quantitative developmental studies are crucial to a detailed understanding of mechanisms that lead to a normal versus perturbed development. The variational properties of development determine how genetic and environmental variation translates into phenotypic variation in postnatal and adult individuals (Hallgrímsson and Lieberman, 2008; Mitteroecker, 2009; Müller and Newman, 2003; Pigliucci and Müller, 2010; Salazar-Ciudad, 2006; G. P. Wagner and Altenberg, 1996; West-Eberhard, 2003). In turn, the population pool of phenotypic variation is the substrate for natural selection, and hence, a major determinant of organismal evolution (Müller and Newman, 2003). The lack of quantitative studies of developmental variation has impeded the long-expected connection of developmental biology with the formal core of evolutionary theory.

Population models in evolutionary theory, genetics, and epidemiology are based on quantitative representations of phenotypic and genetic variation. Yet variation in embryological development and growth – the processes translating genetic variation into phenotypic variation – is still poorly understood. It has been argued on theoretical grounds that the properties of development assume to play a major role in shaping phenotypic variation within and across populations (Hallgrímsson and Lieberman, 2008; Mitteroecker, 2009; Müller, 2007).

In **Paper IV** I introduce a method – *Geometric Morphometric Image Analysis (GMIA)* – to measure the geometry of a developing embryo together with spatial patterns of tissue formation. The method consists of two steps. For structures with sharp boundaries and well-differentiated tissue, morphological variation is quantified as variation in shape, position, size, and orientation of the structure. Structures with diffuse boundaries are described as spatial patterns of intensities or directions within the organism. This approach results in two sets of data, representing well-differentiated anatomical structures and imaged tissue properties.
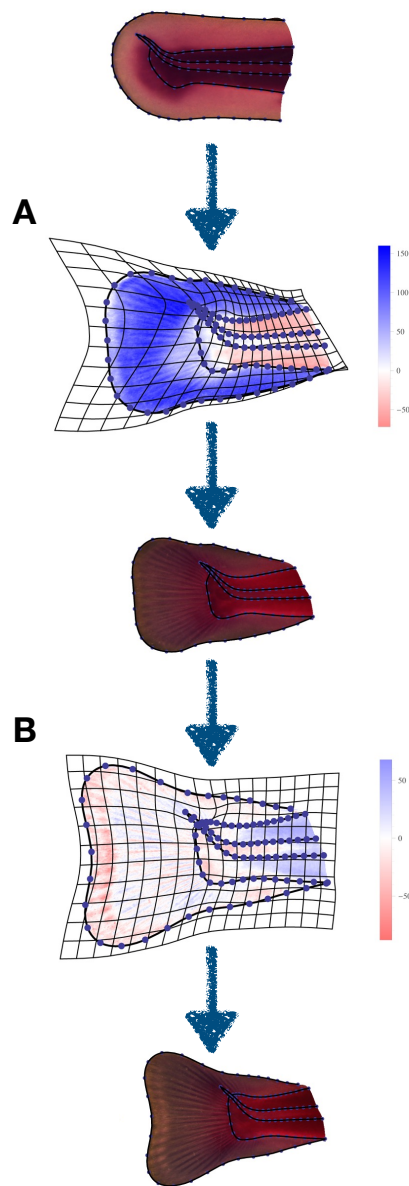
Figure 12: Visualization of shape change and change of tissue density (A) between 21 days post fertilization and 40 days post fertilization and (B) between 40 days post fertilization and 56 days post fertilization. The shape of the fin changes from a round to a more triangular shape in both steps. The change in tissue density shows an increase (blue) in the fin between 21 days post fertilization and 40 days post fertilization, a decrease (red) of tissue density in musculature and notochord, and an increase of tissue density in the fin rays between 40 days post fertilization and 56 days post fertilization.

Statistics can be calculated separately for shape and texture, or both together.

I applied the method onto two-dimensional microscopic images of a sample of rainbow trout (*Oncorhynchus mykiss*) tail fins and could illustrate average changes in shape and cell density over a developmental period of 35 days. By combining modern morphometrics and image analysis, I quantified variation in the shape of developmentally homologous structures (e.g., fin outline, notochord) and used this information to register microscopic images. The RGB values (the values of red, green, and blue color channels of a pixel) of the registered images were analyzed to observe spatial changes in tissue densities and the emergence of novel structures, such as the fin rays. Both shape and tissue densities can be averaged and analyzed by multivariate statistical methods. The visualization of statistical results (fig. 12) can include shape differences, tissue densities or both at the same time.

The method allows us to quantify how genetic variation translates into phenotypic variation. It can be applied to study the properties of the genotype-phenotype map that are caused by developmental mechanisms. However, there are limitations to the method as the manual setting of landmarks (shape variables) to define the boundaries of the studied structure. Hence, it is not possible yet to use the method in large-scale evolutionary simulations.

5.4   Concluding Remarks

"Nothing in biology makes sense except in the light of evolution."

– Theodosius Dobzhansky (1973)

Dobzhansky (1973) argued that the observed diversity on earth only makes sense if life on earth has a shared history. To understand life's history and how diversity is generated it is crucial to expand our knowledge on the mechanisms behind evolutionary processes. The debate around how it is possible to evolve complex organisms is a major driving force for research in evolutionary biology. On

the ground of that debate stands the question about the relationship between genotypes and phenotypes. The Modern Synthesis proposes that genes determine phenotypes, but neglects the importance of developmental processes. Nevertheless, it has been shown (Hallgrímsson and Lieberman, 2008; Mitteroecker, 2009; Müller and Newman, 2003; Pigliucci and Müller, 2010; Salazar-Ciudad, 2006; G. P. Wagner and Altenberg, 1996; West-Eberhard, 2003) that development plays a major role in shaping the genotype-phenotype map. Large amounts of genetic variation are arguably concentrated by a limited number of developmental pathways into a lower-dimensional space of phenotypic variation (Hallgrímsson and Lieberman, 2008; Martinez-Abadias et al., 2012). C.H. Waddington (1942) showed that developmental canalization or robustness is an important factor when it comes to buffering genetic variation and the accumulation of cryptic variation. Hence, the actual underlying developmental mechanisms are still poorly understood.

The genotype-phenotype map is the metaphorical construct that refers to the underlying processes that explain the translation of genetic into phenotypic variation. However, we still lack an understanding of the processes that shape and structure the genotype-phenotype map. This thesis aims to shed light on properties that shape the genotype-phenotype map like epistasis and pleiotropy and with that influence the relationship between evolvability and robustness. But, it also emphasizes the need to understand the impact of developmental mechanisms onto the genotype-phenotype map. We need to gain knowledge about the mechanistic properties that build the genotype-phenotype map (G. Wagner et al., 2000).

The arguments made here are progressing from being based on a simple static network-like genotype-phenotype map (**Paper I**) to the introduction of a framework to build a biological realistic genotype-phenotype map based on an actual model of development (**Paper II, Paper III**). I show that the relationship between evolvability and robustness is dependent on the structure of the genotype-phenotype map. The structure of the genotype-phenotype map is not general and seems to vary over different systems dependent on its epistatic and pleiotropic structure.

A full understanding of the influence of the structure of the genotype-phenotype

43

map on evolutionary dynamics can only be achieved considering all different approaches that are proposed so far. There is a need to combine statistical, functional and mechanical representations of the genotype-phenotype map in a common framework.

Finally, I completed my argumentation by introducing a powerful biometric tool for studying the variation of a developing organism (**Paper IV**). *Geometric Morphometric Image Analysis* is a statistical method to study developmental variation that is central to gain insight into the mechanisms that structure the genotype-phenotype map.

The contribution of this thesis provides a piece of the puzzle in connecting evolutionary and developmental biology. Combining evolutionary and developmental models can promote the integration of developmental biology and Evolutionary Developmental Biology (EvoDevo) together with population and quantitative genetics. It may thus provide an additional contribution to an extended evolutionary synthesis (Pigliucci and Müller, 2010).

6 FUTURE PERSPECTIVE

All of the work done here expands our knowledge about the structure of the genotype-phenotype map. The structure influences the ability to evolve novel phenotypes and adapt to a changing environment. The next step to achieve full integration of population genetics with EvoDevo is to understand how evolution changes the distribution of phenotypes in a population. We need to understand how heritable variation is passed onto the next generation (Barton and Turelli, 1989) – a process that is influenced by the underlying genetic architecture and therefore, the structure of the genotype-phenotype map.
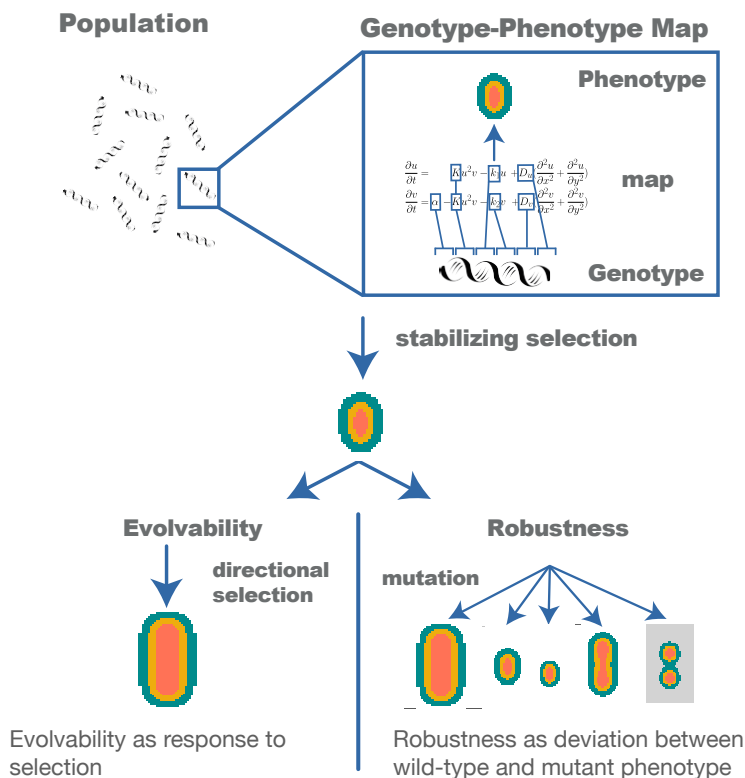

Figure 13: Illustration of the individual-based simulations to calculate evolvability and robustness in a population under selection.

Therefore, we need to expand the model presented in **Paper II** and **Paper III**. The reaction-diffusion model can be used in simulations of populations. The parameters again are used as genotype and the resulting pattern as phenotype. By improving the introduced *GMIA* method to work landmark-free it is possible to use it to analyze the simulated patterns in a population. The advantage would be the potential to quantify novelties and emerging structures.

The relationship between evolvability and robustness is analysed using populations that are in mutation-selection balance at first. Evolvability is defined as response to directional selection and measured as the average change of the trait – e.g. size – per generation in the population. Robustness is calculated as mutational effect and measured as average change between a mutant population and the wild-type population (fig. 13).

Here, we would explore the properties of a statistical representation of the genotype-phenotype map. Such an approach would allow us to study the connection between developmental variation, and genetic and phenotypic variation. In the case of adding developmental noise and select for a reduction of it, it would be possible to explore the relationship between developmental robustness (canalization) and evolvability. This approach would broaden our insight into the mechanisms that shape the connection between evolvability and robustness.

REFERENCES

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

Aldana, M., E. Balleza, S. A. Kauffman, and O. Resendiz (2007). Robustness and evolvability in genetic regulatory networks. *Journal of Theoretical Biology* 245 (3):433–48.

Altenberg, L. (1994). The Evolution of Evolvability in Genetic Programming. *Advances in Genetic Programming*. Edited by K. E. Kinnear. MIT Press, p. 47–74.

Amundson, R. (2005). *The changing role of the embryo in evolutionary thought: roots of evo-devo*. Cambridge: Cambridge University Press.

Ancel, L. W. and W. Fontana (2000). Plasticity, evolvability, and modularity in RNA. *Journal of Experimental Zoology* 288 (3):242–283.

Arthur, W. (1997). *The origin of animal body plans: a study in evolutionary developmental biology.* Cambridge: Cambridge University Press.

Arthur, W. (2000). The concept of developmental reprogramming and the quest for an inclusive theory of evolutionary mechanisms. *Evolution & Devlopment* 2:49–57.

Bagheri-Chaichian, H., J. Hermisson, J. R. Vaisnys, and G. P. Wagner (2003). Effects of epistasis on phenotypic robustness in metabolic pathways. *Mathematical Biosciences* 184 (1):27–51.

Barton, N. H. (1995). A general model for the evolution of recombination. *Genetical Research* 65 (2):123–144.

Barton, N. H. (2017). How does epistasis influence the response to selection? *Heredity* 118 (1):96–109.

Barton, N. H. and M. Turelli (2004). Effects of genetic drift on variance components under a general model of epistasis. *Evolution* 58 (10):2111–2132.

Barton, N. H. and M. Turelli (1989). Evolutionary quantitative genetics: How little do we know? *Annual Review of Genetics* 23 (1):337–370.

Bateson, W. (1909). *Mendel's principles of heredity*. Cambridge: Cambridge University Press.

Boyle, E. A., Y. I. Li, and J. K. Pritchard (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* 169 (7):1177–1186.

Brigandt, I. (2015). Evolutionary Developmental Biology and the Limits of Philosophical Accounts of Mechanistic Explanation. *Explanation in Biology*. Edited by P.-A. Braillard and C. Malaterre. Dordrecht: Springer, p. 135–173.

Carter, A. J., J. Hermisson, and T. F. Hansen (2005). The role of epistatic gene interactions in the response to selection and the evolution of evolvability. eng. *Theoretical Population Biology* 68 (3):179–196.

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

Cheverud, J. M. and E. J. Routman (1996). Epistasis as a source of increased additive genetic variance at population bottlenecks. *Evolution* 50 (3):1042–1051.

Cockerham, C. C. (1954). An Extension of the Concept of Partitioning Hereditary Variance for Analysis of Covariances among Relatives When Epistasis Is Present. *Genetics* 39 (6):859–82.

Connahs, H., S. Tlili, J. van Creij, T. Y. J. Loo, T. D. Banerjee, T. E. Saunders, and A. Monteiro (2019). Activation of butterfly eyespots by Distal-less is consistent with a reaction-diffusion process. *Development* 146:dev169367.

Conrad, M. (1990). The geometry of evolution. *Bio Systems* 24 (1):61–81.

Cotterell, J. and J. Sharpe (2010). An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients. *Molecular Systems Biology* 6 (1):2115–2124.

DiFrisco, J. and J. Jaeger (2019). Beyond networks: mechanism and process in evo-devo. *Biology and Philosophy* 34 (6):1–24.

Dobzhansky, T. (1937). *Genetics and the origin of species.* New York: Columbia University Press.

Dobzhansky, T. (1973). Nothing in Biology Makes Sense Except in the Light of Evolution. *American Biology Teacher* 35 (3):125–129.

Draghi, J. A., T. L. Parsons, G. P. Wagner, and J. B. Plotkin (2010). Mutational robustness can facilitate adaptation. *Nature* 463 (7279):353–5.

Ebner, M., M. Shackleton, and R. Shipman (2001). How neutral networks influence evolvability. *Complexity* 7 (2):19–33.

Espinosa-Soto, C., P. Padilla-Longoria, and E. R. Alvarez-Buylla (2004). A gene regulatory network model for cell-fate determination during Arabidopsis thalianal flower development that is robust and recovers experimental gene expression profiles. *Plant Cell* 16 (11):2923–2939.

Fisher, R. A. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh* 52 (2):399–433.

Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford, UK: Clarendon.

Fontana, W. (2002). Modelling 'evo-devo' with RNA. *Bioessays* 24:1164–1177.

Fontana, W., D. A. M. Konings, P. F. Stadler, and P. Schuster (1993). Statistics of RNA secondary structures. *Biopolymers* 33 (9):1389–1404.

Frank, S. A. (1999). Population and quantitative genetics of regulatory networks. *Journal of Theoretical Biology* 197 (3):281–294.

Gavrilets, S. (2004). *Fitness Landscapes and the Origin of Species*. Princeton (New Jersey): Princton University Press.

Gavrilets, S. and A. Hastings (1994). A quantitative-genetic model for selection on developmental noise. *Evolution* 48 (5):1478–1486.

Gerhart, J. and M. Kirschner (1997). *Cells, Embryos and Evolution*. Malden: Blackwell Science Ltd.

Gilbert, S. F. (2000). Diachronic Biology Meets Evo-Devo: C. H. Waddington's Approach to Evolutionary Developmental Biology. *American Zoologist* 40 (5):729–737.

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

Gjuvsland, A. B., J. O. Vik, D. A. Beard, P. J. Hunter, and S. W. Omholt (2013). Bridging the genotype-phenotype gap: What does it take? *Journal of Physiology* 591 (8):2055–2066.

Goodnight, C. J. (1988). Epistasis and the effect of founder events on the additive genetic variance. *Evolution* 42 (3):441–454.

Greenbury, S. F., I. G. Johnston, M. A. Smith, J. P. K. Doye, and A. A. Louis (2010). The effect of scale-free topology on the robustness and evolvability of genetic regulatory networks. *Journal of Theoretical Biology* 267 (1):48–61.

Grüner, W., R. Giegerich, D. Strothmann, C. Reidys, J. Weber, I. L. Hofacker, P. F. Stadler, and P. Schuster (1996). Analysis of RNA sequence structure maps by exhaustive enumeration I. Neutral networks. *Monatshefte für Chemie/Chemical Monthly* 127 (4):355–374.

Hall, B. K. (1999). *Evolutionary Developmental Biology*. Dordrecht: Kluwer Academic Publishers.

Hall, B. K. (2003). Unlocking the Black Box between Genotype and Phenotype: Cell Condensations as Morphogenetic (modular) Units. *Biology & Philosophy* 18 (2):219–247.

Hallgrímsson, B. and D. E. Lieberman (2008). Mouse models and the evolutionary developmental biology of the skull. *Integrative and Comparative Biology* 48 (3):373–384.

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. *Evolutionary Biology* 35 (5):182–185.

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

Hansen, T. F., J. M. Álvarez-Castro, A. J. R. Carter, J. Hermisson, and G. P. Wagner (2006). Evolution of genetic architecture under directional selection. eng. *Evolution* 60 (8):1523–1536.

Hansen, T. F. and G. P. Wagner (2001). Modeling Genetic Architecture: A Multilinear Theory of Gene Interaction. en. *Theoretical Population Biology* 59 (1):61–86.

Hendrikse, J. L., T. E. Parsons, and B. Hallgrímsson (2007). Evolvability as the proper focus of evolutionary developmental biology. *Evolution & Development* 9 (4):393–401.

Hermisson, J., T. F. Hansen, and G. P. Wagner (2003). Epistasis in polygenic traits and the evolution of genetic architecture under stabilizing selection. *The American naturalist* 161 (5):708–34.

Huxley, J. S. (1932). *Problems of Relative Growth*. London: Methuen and Co.

Jaeger, J., D. Irons, and N. Monk (2012). The Inheritance of Process: A Dynamical Systems Approach. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 318 (8):591–612.

Jaeger, J. and N. Monk (2014). Bioattractors: dynamical systems theory and the evolution of regulatory processes. *Journal of Physiology* 592:2267–2281.

Jaeger, J., S. Surkova, et al. (2004). Dynamic control of positional information in the early Drosophila embryo. TL - 430. *Nature* 430:368–371.

Jernvall, J. (2000). Linking development with generation of novelty in mammalian teeth. *Proceedings of the National Academy of Sciences of the United States of America* 97 (6):2641–5.

Kauffman, S. A. (1969). Homeostasis and Differentiation in Random Genetic Control Networks. *Nature* 224 (5215):177–178.

Kauffman, S. A. (1993). *The Origins of Order: Self Organization and Selection in Evolution*. Oxford University Press.

Kauffman, S. and S. Levin (1987). Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology* 128 (1):11–45.

Kempthorne, O. (1954). The correlation between relatives in a random mating population. *Proceedings of the Royal Society of London. Series B - Biological Sciences* 143 (910):103–113.

Kimura, M. (1956). A Model of a Genetic System which Leads to Closer Linkage by Natural Selection. *Evolution* 10 (3):278.

Kondo, S. and T. Miura (2010). Reaction-Diffusion Model as a Framework for Understanding Biological Pattern Formation. *Science* 329 (5999):1616–1620.

Kondo, S. and R. Asai (1995). A reaction-diffusion wave on the skin of the marine angelfish Pomacanthus. *Nature* 376 (6543):765–768.

Laland, K. et al. (2014). Does evolutionary theory need a rethink?

Lange, A., H. L. Nemeschkal, and G. B. Müller (2018). A threshold model for polydactyly. *Progress in Biophysics and Molecular Biology* 137:1–11.

Le Rouzic, A., J. M. Álvarez-Castro, and T. F. Hansen (2013). The Evolution of Canalization and Evolvability in Stable and Fluctuating Environments. *Evolutionary Biology* 40 (3):317–340.

Lesne, A. (2008). Robustness: Confronting lessons from physics and biology. *Biological Reviews* 83 (4):509–532.

Lewontin, R. C. (1970). The units of selection. *Annual Review of Ecology and Systematics* 1:1–18.

Lynch, M., J. Conery, and R. Burger (1995). Mutation accumulation and the extinction of small populations. *American Naturalist* 146 (4):489–518.

Mackay, T. F. (2014). Epistasis and quantitative traits: Using model organisms to study gene-gene interactions. *Nature Reviews Genetics* 15 (1):22–33.

Marcon, L. and J. Sharpe (2012). Turing patterns in development: what about the horse part? *Current Opinion in Genetics & Development* 22 (6):578–584.

Martinez-Abadias, N., P. Mitteroecker, T. E. Parsons, M. Esparza, T. Sjovold, C. Rolian, J. T. Richtsmeier, and B. Hallgrímsson (2012). The Developmental Basis of Quantitative Craniofacial Variation in Humans and Mice. *Evolutionary Biology* 39 (4):554–567.

Masel, J. and M. V. Trotter (2010). Robustness and evolvability. *Trends in Genetics* 26 (9):406–14.

Maynard Smith, J. (1970). Natural Selection and the Concept of a Protein Space. *Nature* 225 (5232):563–564.

Maynard Smith, J. (1978). *The Evolution of sex*. Cambridge, U.K.: Cambridge University Press.

Mayr, E. (1961). Cause and effect in biology. *Science* (134):1501–1506.

Mayr, E. (1982). *The Growth of Biological Thought*. Cambridge, MA.: Harvard University Press.

Meinhardt, H. (1983). Cell determination boundaries as organizing regions for secondary embryonic fields. *Developmental Biology* 96:375–385.

Milocco, L. and I. Salazar-Ciudad (2020). Is evolution predictable? Quantitative genetics under complex genotype-phenotype maps. *Evolution*:evo.13907.

Minelli, A. and G. Fusco (2012). On the Evolutionary Developmental Biology of Speciation. *Evolutionary Biology* 39 (2):242–254.

Mitteroecker, P. (2009). The Developmental Basis of Variational Modularity: Insights from Quantitative Genetics, Morphometrics, and Developmental Biology. *Evolutionary Biology* 36:377–385.

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

Müller, G. B. and S. A. Newman (2003). *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology*. Edited by G. B. Müller and S. A. Newman. Cambridge: MIT Press.

Müller, G. B. and G. P. Wagner (1991). Novelty in evolution: Restructuring the concept. *Annual Review of Ecology and Systematics* 22 (1):229–256.

Murray, M. E. (2002). *Mathematical Biology : I . An Introduction , Third Edition*. New York: Springer-Verlag.

Newman, S. A. and G. B. Müller (2000). Epigenetic mechanisms of character origination. *The Journal of experimental zoology* 288 (4):304–17.

Omholt, S. W. (2013). From sequence to consequence and back. *Progress in Biophysics and Molecular Biology* 111 (2-3):75–82.

Orr, H. A. (2000). Adaptation and the cost of complexity. *Evolution* 54 (1):13–20.

Oster, G. F. and P. Alberch (1982). Evolution and bifurcation of developmental programs. *Evolution* 36 (3):444–459.

Payne, J. L., J. H. Moore, and A. Wagner (2014). Robustness, evolvability, and the logic of genetic regulation. *Artificial life* 20 (1):111–26.

Payne, J. L. and A. Wagner (2014). The robustness and evolvability of transcription factor binding sites. *Science* 343 (6173):875–7.

Phillips, P. C. (1998). The Language of Gene Interaction. *Genetics* 149 (3):1167–71.

Phillips, P. C. (2008). Epistasis - the essential role of gene interactions in the structure and evolution of genetic systems. *Nature Reviews Genetics* 9 (11):855–867.

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.

Pigliucci, M. and G. B. Müller (2010). *Evolution - The Extended Synthesis*. Cambridge: MIT Press.

Polly, P. D. (2008). Developmental Dynamics and G-Matrices: Can Morphometric Spaces be Used to Model Phenotypic Evolution? *Evolutionary Biology* 35 (2):83–96.

Raff, R. (1996). *The Shape of Life: Genes, Development, and the Evolution of Animal Form.* Chicago: Univ Chicago Press.

Raspopovic, J., L. Marcon, L. Russo, and J. Sharpe (2014). Modeling digits. Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients. *Science* 345 (6196):566–70.

Rice, S. H. (1998). The evolution of canalization and the breaking of von Baer's Laws: Modeling the evoltution of development with epistasis. *Evolution* 52 (3):647–656.

Rice, S. H. (2002). A general population genetic theory for the evolution of developmental interactions. eng. *Proceedings of the National Academy of Sciences of the United States of America* 99 (24):15518–15523.

Rice, S. H. (2004). Developmental Associations Between Traits: Covariance and Beyond. *Genetics* 166:513–526.

Riedl, R. J. (1978). *Order in Living Organisms*. New York: John Wiley and Sons.

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

Salazar-Ciudad, I. and J. Jernvall (2004). How different types of pattern formation mechanisms affect the evolution of form and development. *Evolution and Development* 6 (1):6–16.

Salazar-Ciudad, I. and M. Marín-Riera (2013). Adaptive dynamics under development-based genotype-phenotype maps. *Nature* 497 (7449):361–4.

Schuster, P., W. Fontana, P. F. Stadler, and I. L. Hofacker (1994). From sequences to shapes and back: a case study in RNA secondary structures. *Proceedings of the Royal Society B: Biological Sciences* 255 (1344):279–84.

Siegal, M. L. and A. Bergman (2002). Waddington's canalization revisited: developmental stability and evolution. *Proceedings of the National Academy of Sciences of the United States of America* 99 (16):10528–10532.

Stadler, P. F. and B. M. R. Stadler (2006). Genotype-Phenotype Maps. *Biological Theory* 1 (3):268–279.

Templeton, A. R. (1979). The unit of selection in Drosophila mercatorum. II. Genetic revolution and the origin of coadapted genomes in parthenogenetic strains. *Genetics* 92 (4):1265–1282.

Templeton, A. R. (1980). The theory of speciation via the founder principle. *Genetics* 94 (4):1011–1038.

Turing, A. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 237 (641):37–72.

Verd, B., A. Crombach, and J. Jaeger (2014). Classification of transient behaviours in a time-dependent toggle switch model. *BMC Systems Biology* 8 (1):43.

Visser, J. A. G. M. de et al. (2003). Perspective:Evolution and Detection of Genetic Robustness. *Evolution* 57 (9):1959.

Waddington, C. H. (1942). The canalization of development and the inheritance of acquired characters. *Nature* 150:563–565.

Waddington, C. H. (1959). Canalization of development and genetic assimilation of acquired characters. *Nature* 183:1654–1655.

Waddington, C. H. (1957). *The strategy of the genes: a discussion of some aspects of theoretical biology.* London: Allen & Unwin.

Wagner, A. (1996). Does Evolutionary Plasticity Evolve? *Evolution* 50 (3):1008–1023.

Wagner, A. (2005). *Robustness and evolvability in living systems.* Princeton (New Jersey): Princeton University Press.

Wagner, A. (2008). Robustness and evolvability: A paradox resolved. *Proceedings of the Royal Society B: Biological Sciences* 275 (1630):91–100.

Wagner, A. (2011). *The Origins of Evolutionary Innovations: A Theory of Transformative Change in Living Systems.* New York: Oxford University Press.

Wagner, A. (2012). The role of robustness in phenotypic adaptation and innovation. *Proceedings of the Royal Society B: Biological Sciences* 279 (1732):1249–1258.

Wagner, G., C. Chiu, and M. D. Laubichler (2000). Developmental evolution as a mechanistic science: the inference from developmental mechanisms to evolutionary processes. *American Zoologist* 40 (5):819–831.

Wagner, G. P. and L. Altenberg (1996). Complex adaptations and the evolution of evolvability. *Evolution* 50 (3):967–976.

Wagner, G. P., G. Booth, H. Bagheri-Chaichian, and H. Homayoun-Chaichia (1997). A Population Genetic Theory of Canalization. *Evolution* 51 (2):329–347.

Wagner, G. P. and J. Zhang (2011). The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nature Reviews Genetics* 12:204–213.

Wallace, B. (1986). Can embryologists contribute to an understanding of evolutionary mechanisms? *Integrating scientific disciplines*. Edited by W. Bechtel. Dordrecht: Martinus Nijhoff, p. 149–163.

West-Eberhard, M. J. (2003). *Developmental Plasticity and Evolution*. New York: Oxford University Press.

Wright, S. (1931). Evolution in Mendelian Populations. *Genetics* 16 (2):97–159.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the sixth international congress of genetics* 1:356–366.

Wright, S. (1967). "Surfaces" of selective value. *Proceedings of the National Academy of Sciences of the United States of America* 58 (1):165.

Wright, S. (1980). Genic and Organismic Selection. *Evolution* 34 (5):825.

PAPER I

I

Evolvability and robustness: A paradox restored

Christine Mayer*, Thomas F. Hansen

Department of Biosciences, CEES, EvoGene & CEDE, University of Oslo, PB 1066, Blindern, 0316 Oslo, Norway

ABSTRACT

Evolvability and robustness are crucial for the origin and maintenance of complex organisms, but may not be simultaneously achievable as robust traits are also hard to change. Andreas Wagner has proposed a solution to this paradox by arguing that the many-to-few aspect of genotype–phenotype maps creates neutral networks of genotypes coding for the same phenotype. Phenotypes with large networks are genetically robust, but they may also have more neighboring phenotypes and thus higher evolvability. In this paper, we explore the generality of this idea by sampling large numbers of random genotype–phenotype maps for Boolean genotypes and phenotypes. We show that there is indeed a preponderance of positive correlations between the evolvability and robustness of phenotypes within a genotype–phenotype map, but also that there are negative correlations between average evolvability and robustness across maps. We interpret this as predicting a positive correlation across the phenotypic states of a character, but a negative correlation across characters. We also argue that evolvability and robustness tend to be negatively correlated when phenotypes are measured on ordinal or higher scale types. We conclude that Wagner's conjecture of a positive relation between robustness and evolvability is based on strict and somewhat unrealistic biological assumptions.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Evolvability and robustness are both thought essential for the origin and maintenance of complex, well-adapted organisms (Conrad, 1990; Kauffman, 1993; Gerhart and Kirschner, 1997; Wagner, 2005), but these two properties conflict as evolvability depends on the ability to generate new potentially adaptive phenotypes through mutation while robustness depends on the ability to maintain the same phenotype in the face of mutation. The easier it is to change the phenotype through genetic change, the more evolvable and less robust is the genotype. This creates a paradox for the evolution of complex phenotypes.

Andreas Wagner (2008) proposed an ingenious solution to this paradox. He argued that the paradox originated in thinking about evolvability and robustness as properties of the genotype. In this case, the two properties indeed conflict. His resolution of the paradox was to consider evolvability and robustness not as properties of the genotype, but as properties of the phenotype. A phenotype may be realized by a number of different genotypes, which then forms a neutral network (Fontana et al., 1993; Kauffman, 1993; Schuster et al., 1994; Gruener et al., 1996; Fontana, 2002), also called a genotype network by Wagner (2011) and Payne and Wagner (2014). By defining phenotype evolvability as the number of different phenotypes that can be generated through mutations of any member of the corresponding neutral network, and phenotype robustness as the average probability that the phenotype will not change through mutation of a member of the corresponding neutral network, Wagner argued that phenotypes with large, connected neutral networks could be both robust and evolvable, because they would have many neighbors at the same time as most mutations would leave them unchanged.

Wagner (2008) confirmed the general possibility of a positive correlation between phenotype evolvability and robustness in simulation studies in which the phenotypes were secondary structures of RNA and the associated neutral networks were sets of nucleotide sequences that generated the same secondary structure. Wagner (2005, 2012) and Payne and Wagner (2014) discuss other examples in which robustness facilitates evolvability including protein folding and transcription-factor binding, while Ebner et al., (2002), Aldana et al., (2007), Lesne (2008) and Draghi et al., (2010) have made similar arguments.

Here we explore the generality of these findings and discuss properties of the genotype–phenotype map (GP map) that generate either a positive or a negative relation between evolvability and robustness. We are using a general abstract representation of genotype–phenotype maps in which the genotype and the phenotype are represented by Boolean variables connected through Boolean operators. We demonstrate that both positive and negative relations between evolvability and robustness are possible de-

---

* Corresponding author.
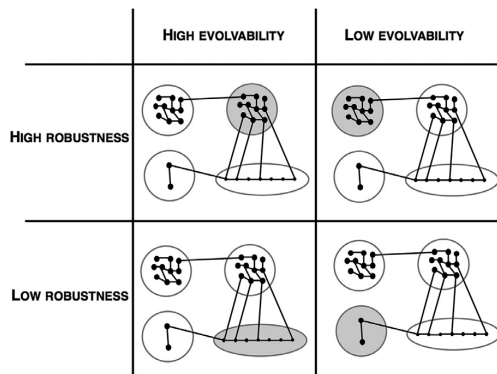 *E-mail address:* christine.mayer@ibv.uio.no (C. Mayer).

Fig. 1. How evolvability and robustness may vary across phenotypes within a GP map. The black dots represent genotypes with lines representing possible mutations, and the circles represent phenotypes. In each quadrant evolvability and robustness are assigned to the focal phenotype colored in grey. Evolvability is the number of phenotypic neighbors, and robustness is the fraction of mutations that do not leave the phenotype. Note how all combinations of phenotype evolvability and robustness are possible.
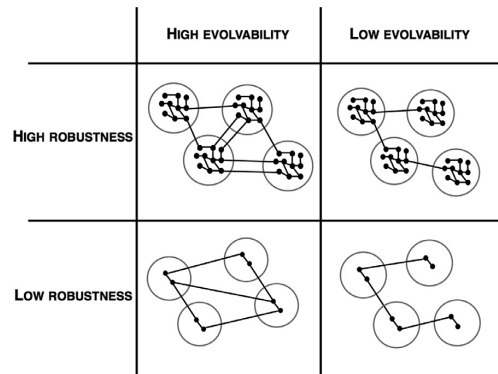

Fig. 2. How robustness and evolvability may vary across GP maps. Each quadrant shows a genotype–phenotype map with a different combination of robustness and evolvability averaged over the four phenotypes each map generates. Notation as in Fig. 1.

pending on the logical structure of the map. Initially we follow the definitions of Wagner (2008) closely, but later we relax some of the inherent assumptions and debate whether they are biologically reasonable.

2. Theory and methods

2.1. General considerations

Wagner (2008) proposed the following definitions (rendered in our terminology):

- Genotype robustness is the number (or fraction) of mutations of a genotype that do not lead to an alternative phenotype.
- Genotype evolvability is the number of alternative phenotypes that can be generated by a single mutation of the genotype.
- Phenotype robustness is the average of genotype robustness over all genotypes within a neutral network.
- Phenotype evolvability is the number of alternative phenotypes that can be generated from the phenotype through mutation.

While genotype evolvability and robustness are negatively related almost by definition, phenotype evolvability and robustness need not be. Positive relationships can arise, as explained by Wagner (2008), if evolvable networks with many neighbors tend to be internally highly connected as illustrated in the upper-left quadrant of Fig. 1, and if networks with few neighbors also tend to have few internal connections as in the lower-right quadrant of Fig. 1. However, as illustrated by the two other quadrants in Fig. 1, inverse relationships can also arise because internally connected networks do not necessarily have many neighbors and vice versa. Some genotype–phenotype maps may also tend to generate more evolvable or robust phenotypes than others. As illustrated in Fig. 2, any combination of evolvability and robustness is possible across maps. Phenotype evolvability and phenotype robustness are thus not automatically positively related, and the relationship needs to be established on a case-by-case basis. To explore the relative occurrence of the different relationships, we will do an exhaustive investigation of all possible genotype–phenotype maps up to a certain level of complexity.

2.2. The Boolean genotype–phenotype map

In our model both genotypes and phenotypes are vectors of Boolean variables coded as 0 and 1, and the genotype–phenotype map is defined as a combination of Boolean logic operators. Boolean variables and functions have often been used as a convenient way of modeling genotype–phenotype relationships in the abstract, or to approximate threshold responses (e.g. Kauffman, 1969, 1993; Frank, 1999; Gavrilets, 1999; Thieffry and Romero, 1999; Ebner et al., 2002; Albert and Othmer, 2003; Espinosa-Soto et al., 2004; Quayle and Bullock, 2006; Aldana et al., 2007; Fierst and Phillips, 2015). It is also worth noting that any Boolean function or logic operator can be represented as a multilinear form in Boolean variables and are thus special cases of the multilinear epistatic model of Hansen and Wagner (2001). Here we will generate large numbers of random Boolean maps. Pleiotropy is modeled by allowing elements in the genotype vector to affect more than one element in the phenotype vector, and mutation is modeled by changing single elements of the genotype vector.

The evolvability and robustness of phenotypes generated by a Boolean genotype–phenotype map can be calculated from the following formula:

$$\mathbf{B}_{zz} = \mathbf{B}_{zg}\mathbf{B}_{gg}\mathbf{B}_{gz}, \tag{1}$$

where $\mathbf{B}_{gg}$ is a genotypic adjacency matrix where the ijth element is 1 if the ith and jth genotypes are connected by a single mutation, $\mathbf{B}_{gz}$ is a matrix where the ijth element is 1 if the genotype i is associated with the phenotype j, and 0 if not, and the matrix $\mathbf{B}_{zg} = \mathbf{B}_{gz}{}^T$ is thus a description of the genotype–phenotype map. The resulting symmetric matrix $\mathbf{B}_{zz}$ describes the connections of phenotypes by mutation in their corresponding genotypes. Its diagonal elements give the number of ways the corresponding phenotype can mutate into itself, and its off-diagonal elements, ij, give the number of ways phenotype i can mutate into phenotype j. The robustness of a phenotype is thus given by the corresponding diagonal element of $\mathbf{B}_{zz}$, or alternatively as a fraction by dividing this with the sum of the corresponding row (or column). The evolvability of a phenotype in Wagner's sense is the number of non-zero off-diagonal elements in the corresponding row (or column).

If we arrange the Boolean genotypes according to the value of their corresponding binary number, then the matrix $\mathbf{B}_{gg}$ has a characteristic pattern, as illustrated here for a genotype of length three:

$$\mathbf{B}_{gg} = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \quad (2)$$

The corresponding list of genotypes is {000, 001, 010, 011, 100, 101, 110, 111}, so that the entry of "1" in position 13 shows that the genotype 010 can mutate into 000, while the entry "0" in position 23 shows that the genotype 010 cannot mutate into 001. In general, this matrix can be constructed by an iterative procedure. If $\mathbf{B}_{gg}$ is the matrix of dimension n (i.e. the genotype is based on an n-digit Boolean number) then $\mathbf{B}_{gg}$ for dimension $n+1$ is given by

$$\begin{bmatrix} \mathbf{B}_{gg} & \mathbf{I} \\ \mathbf{I} & \mathbf{B}_{gg} \end{bmatrix}, \quad (3)$$

where $\mathbf{I}$ is an $2^n \times 2^n$ identity matrix. The main diagonal is always zero as a genotype cannot mutate to itself. Note that this operation is similar to constructing an n-dimensional hypercube by joining all vertices of two $n-1$-dimensional hypercubes.

The matrix $\mathbf{B}_{zg}$ for mapping a three-digit genotype to a single Boolean phenotype ordered as {0,1} could be:

$$\mathbf{B}_{zg} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

Hence, the genotypes 000, 001, 011, 101 and 110 code for phenotype 0, and genotypes 010, 100 and 111 code for phenotype 1. There are no constraints on this matrix except that each column must have one and only one non-zero entry, so that each genotype corresponds to one and only one phenotype.

Calculating $\mathbf{B}_{zz}$ using Eqs. (1), (2) and (4) yields:

$$\mathbf{B}_{ZZ} = \begin{bmatrix} 6 & 9 \\ 9 & 0 \end{bmatrix}, \quad (5)$$

which means that there are 6 mutations that change phenotype 0 into itself, 9 mutations that change phenotype 0 into phenotype 1 and vice versa, and 0 mutations that change phenotype 1 into itself. In this case both phenotypes have evolvability 1, since they can each change into one and only one other phenotype. Meaningful variation in evolvability across phenotypes requires maps that generate more than two phenotypes.

2.3. Generating random maps

The phenotypes in our model consist of three Boolean traits (digits), each calculated from one or more loci in the genotype. Thus, we have a total of eight different phenotypes. The genotypes consist of five or six loci. We chose the sizes of the phenotype and genotype to be small enough to create complete sets of genotype–phenotype maps, but also large enough to generate a range of different map topologies.

We defined six different topologies with varying degrees of pleiotropy to test the influence of complexity on the relationship between phenotype evolvability and robustness. Within each topology, we generated all possible maps in the sense of generating all possible combinations of all possible truth tables for each of the three components of the phenotype. Some of these maps are pseudoreplicates in the sense that they are equivalents with just the order of the loci swapped.

Topology 1 – no pleiotropy: Here each of the three components of the phenotype is determined independently by two loci each,

and no locus has an effect on more than one phenotype component. It turned out that all phenotypes generated within such a topology have exactly the same numbers of neighbors, and thus the same evolvability. Different genotype–phenotype maps may show different levels of evolvability, however.

Topology 2 – low pleiotropy: Here we use the same setup with two loci affecting each phenotypic component, but additionally we allow one locus to affect two traits so that one phenotype component is affected by three loci and the other two components by two loci.

Topology 3 – medium pleiotropy: Same as topology two except that the one pleiotropic locus now affects all three phenotype components. Here, two phenotype components are affected by three loci and one phenotype component is affected by two loci.

Topology 4 – high pleiotropy: Same as topology three except that one additional genotype locus can affect two phenotype components. All three phenotype components are affected by three loci.

Topology 5 – modular: This topology has two modules with three genotype loci jointly affecting two phenotype components and the three other genotype loci jointly affecting the third phenotype component.

Topology 6 – integrated modules: Same as topology 5, but the modules get connected through one genotype locus, which thus affects all phenotype components. In addition, one genotype locus is dropped so that each "module" is determined by two loci plus the shared one. There are thus only five genotype loci in this topology.

We also sampled ten million $\mathbf{B}_{zg}$ matrices out of all possible combinations for a genotype of length six and a phenotype of length three. Due to computational constraints, it was not possible to calculate the whole set of all $\mathbf{B}_{zg}$ matrices.

3. Results

3.1. Phenotype evolvability and robustness within genotype–phenotype maps

We computed evolvability and robustness across the individual phenotypes generated by each specific genotype–phenotype map (i. e. a specific $\mathbf{B}_{zg}$ matrix). Many of these maps, and particularly those from the simpler topologies, do not generate any variation in evolvability among their phenotypes. Among the "nondegenerate" cases with variation in evolvability, the results in Fig. 3 show that the majority generate a positive relationship between evolvability and robustness across phenotypes. Cases of negative correlations are always a minority, but become more common in the more complex topologies.

3.2. Character evolvability and robustness across genotype–phenotype maps

The evolvability and robustness of a particular genotype–phenotype map can be measured as the average of the evolvability and robustness of the phenotypes it generates. We suggest interpreting this as the evolvability and robustness of a character (or character identity sensu GP Wagner, 2014) as opposed to the evolvability and robustness of individual character states (i. e. phenotypes). As illustrated in Fig. 4, evolvability and robustness tend to be negatively correlated across genotype–phenotype maps. This holds true both within the specific topologies and across the whole sample. Note, however, that some of this is driven by genotype–phenotype maps that generate only one or two phenotypes. These necessarily have high robustness and low evolvability. Indeed, it makes sense that genotype–phenotype maps that can generate many phenotypes, or characters with many possible states, are considered more evolvable. Our results show that such characters also tend to be less robust.
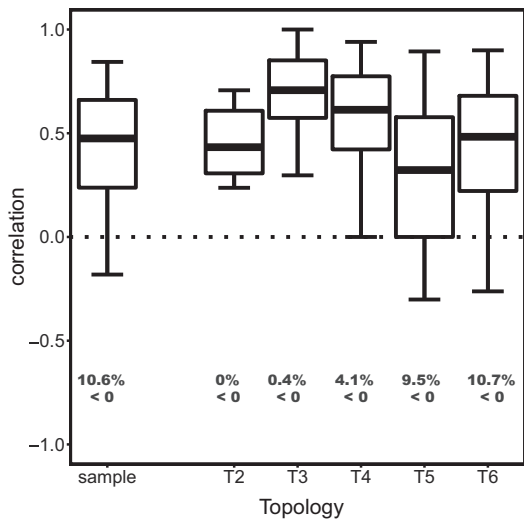
63

Fig. 3. Boxplot of correlation coefficients between phenotype evolvability and robustness across phenotypes within a GP map. The boundaries of the boxes in the plot represent the first and the third quartile. The line in the middle of the box is the median. The whiskers are the 5th and 95th percentiles. "Sample" refers to a random sample of 10 mill maps from the complete set, "T2" to topology 2, "T3" to topology 3, "T4" to topology 4, "T5" to topology and "T6" to topology 6 as defined in the main text. The numbers below the boxes are the percentage of negatively correlated maps found in the corresponding topology. While the majority of maps yield a positive correlation, the more complex maps have a larger probability of generating a negative correlation.

3.3. Robustness and evolvability of sequential adaptations

An inherent assumption in Wagner's (2008) definitions is that evolvability depends on finding a single random phenotype through a single random mutation. This stands in sharp contrast to the traditional neo-Darwinian view of adaptations being built sequentially through many contingent steps. It seems likely that the positive relations between evolvability and robustness found for RNA secondary structures and transcription-factor binding depend crucially on this assumption. In this section, we introduce a model to explore the effects of phenotype robustness on evolvabil-
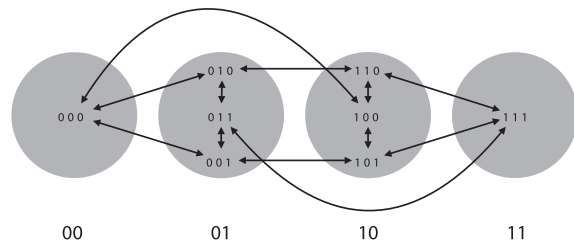

Fig. 5. Transversing genotypes on ordinal scale. The grey circles represent 4 different phenotypes (denoted by the bold binary numbers below) of increasing fitness from left to right. The small binary numbers are the genotypes and the black arrows represent mutational connections between genotypes.

ity when the latter is operationalized as the time to transverse a sequence of phenotypic steps.

Let an adaptation require transversing a sequence of phenotypes ordered by their Boolean numbers as illustrated in Fig. 5. This introduces an ordinal scale for phenotypes. Specifying the relative fitness of the phenotypes will lead to a stronger interval scale type where the selection coefficient for a transition between two Boolean phenotypes equals the sum of the selection coefficients for each step between them.

Now consider the rate of transition between two genotypes in this system. We assume that mutation is sufficiently rare for the population to be fixed at one genotype before the next mutation appears. If the mutation rate between the genotypes is u, then the transition rate for the population will also be u if the genotypes code for the same phenotype (Kimura, 1983). If the genotypes code for different phenotypes, and the change from one phenotype to the next is associated with a relative fitness advantage of s, then the rate of transition will be $b \approx 4sNu$, where N is effective population size, and we have assumed an approximate fixation probability of $2s$ (e. g. Haldane, 1927; Bürger and Ewens, 1995).

If the pairwise transition rates between all genotypes are specified in a matrix **A**, then the time development of the vector of probabilities, **x**, of being in each state can be described with a system of ordinary differential equations: $d\mathbf{x} = \mathbf{A}\mathbf{x}dt$. From this the distribution of times to pass through the system can be derived with standard methods as detailed in the Appendix. For a given genotype–phenotype map on Boolean variables as specified by a
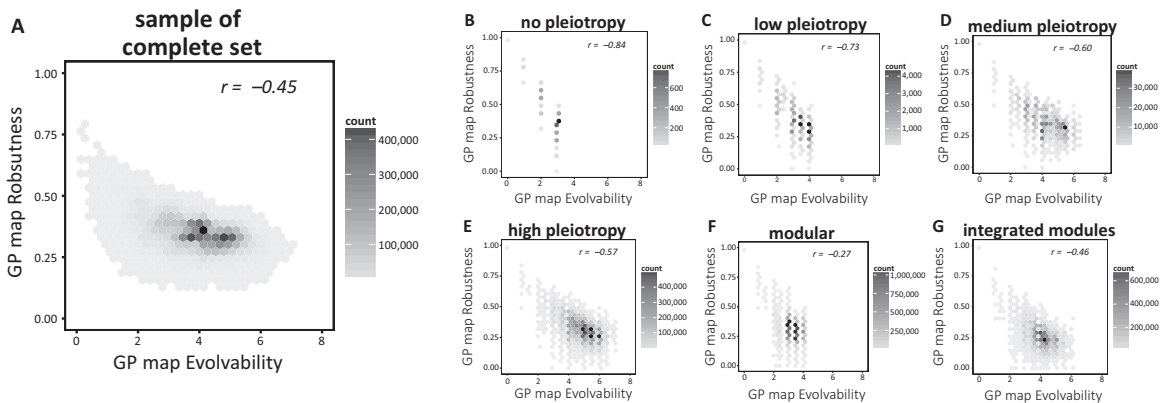

Fig. 4. Correlation between evolvability and robustness across GP maps. (A) Sample of complete set. (B) Topology 1: no pleiotropy. (C) Topology 2: low pleiotropy. (D) Topology 3: medium pleiotropy. (E) Topology 4: high pleiotropy. (F) Topology 5: modular. (G) Topology 6: integrated modules. All topologies show consistent negative correlations between GP map evolvability and robustness.

$\mathbf{B}_{zg}$ matrix, the transition matrix can be constructed as

$$\mathbf{A} = \mathbf{B}_{gz}\mathbf{T}\mathbf{B}_{zg} \odot \mathbf{B}_{gg} - (\mathbf{B}_{gg}\mathbf{B}_{gz}\mathbf{T}\mathbf{B}_{zg}) \odot \mathbf{I}, \tag{6}$$

where $\odot$ signify Hadamard product, $\mathbf{I}$ is an identity matrix, and the matrix $\mathbf{T}$ is a transition matrix for the phenotypes. The ij element of $\mathbf{T}$ is the rate of transition from a specific genotype in phenotype j to a specific genotype in phenotype i if there were a mutation that made this transition possible. For example, for the system illustrated in Fig. 5 of four phenotypes coded by two-digit Boolean numbers ordered as {00, 01, 10, 11}, and with constant mutation rates and selection coefficients, the T-matrix is

$$\mathbf{T} = \begin{bmatrix} u & 0 & 0 & 0 \\ b & u & 0 & 0 \\ 2b & b & u & 0 \\ 3b & 2b & b & u \end{bmatrix}, \tag{7}$$

where the u's on the main diagonal give transition rates between connected genotypes within a phenotype, and the multiples of b below the diagonal specify rates between connected genotypes belonging to different phenotypes as functions of their fitness differences. For example, the $3b$ entry in row 4 and column 1 gives the rate of transition from a genotype in phenotype 00 to a mutationally-connected genotype in 11 (regardless of whether such a mutation exists or not). This will have a selection coefficient that is three times the selection coefficient for a single step in the sequence, and hence a transition rate of $3b$. This ensures that the fitness gain of moving from one phenotype to another is the same regardless of the number of steps that were made to achieve it. The zeros above the diagonal mean that it is impossible to move backwards in the system of phenotypes. The off-diagonal elements of the matrix $\mathbf{B}_{gz}\mathbf{T}\mathbf{B}_{zg}$ give the transition rates between genotypes if all transitions were mutationally possible, and the Hadamard multiplication with $\mathbf{B}_{gg}$ removes all transitions that are not possible. The second term in the equation constructs the diagonal so as to make $\mathbf{A}$ a proper probability transition matrix.

For the genotype–phenotype map illustrated in Fig. 5,

$$\mathbf{B}_{zg} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{8}$$

and using this in Eq. (6) yields:

$$\mathbf{A} = \begin{bmatrix} -4b & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ b & -u-b & 0 & u & 0 & 0 & 0 & 0 \\ b & 0 & -u-b & u & 0 & 0 & 0 & 0 \\ 0 & u & u & -2u-2b & 0 & 0 & 0 & 0 \\ 2b & 0 & 0 & 0 & -2u & u & u & 0 \\ 0 & b & 0 & 0 & u & -u-b & 0 & 0 \\ 0 & 0 & b & 0 & u & 0 & -u-b & 0 \\ 0 & 0 & 0 & 2b & 0 & b & b & 0 \end{bmatrix}, \tag{9}$$

Using the method in the Appendix, we compute the mean passage time from genotype state 000 to state 111 to be

$$\bar{t} = \frac{b^2 + 11bu + 14u^2}{4bu(2u+b)} \approx \frac{1}{4u}\left(1 + \frac{11}{4sN}\right), \tag{10}$$

where the approximation is due to ignoring terms in u^2. This shows that the passage time is inversely proportional to mutation rates and approaches a lower limit of $1/4u$ when selection is strong ($sN \gg 1$).

To explore the relations between robustness and evolvability it is useful to consider more complex maps. We will use a sequence of eight phenotypes coded by three-digit Boolean numbers and different genetic architectures. The T-matrix for this system, given in
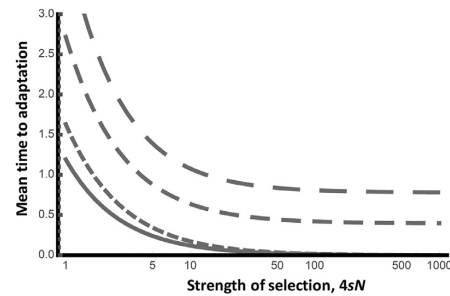

Fig. 6. Mean times to adaptation as a function of the strength of selection for four degrees of robustness. Each graph shows the mean passage time through a sequence of 8 phenotypes in units per mutation rate, u. Hence, a value of 1.0 means that the passage time is the inverse of the mutation rate (e.g. a million generations if $u = 10^{-6}$/generation). The four functions show genetic architectures based on three-, four-, five- and six-digit Boolean genotypes with subsequently coarser dashing for higher dimensions. Note how the more robust higher-dimensional architectures have longer passage times and are thus less evolvable.
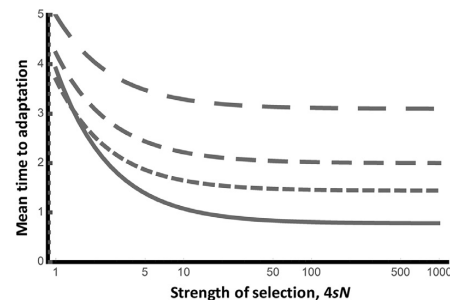

Fig. 7. Mean times to adaptation as a function of the strength of selection with different numbers of selectively distinct phenotypes. The whole line shows the case where all 8 phenotypes are selectively distinct as in Fig. 6. The dashed lines show cases where the six intermediate phenotypes are grouped into, from bottom up, three, two or one selectively-distinct phenotypes. In all cases, there are 64 six-digit Boolean genotypes. Note how the cases with fewer and thus more robust phenotypes have longer passage times and are thus less evolvable. Time in units of mutation rate as in Fig. 6.

the supplement, has the same structure as above in keeping with the idea that the fitness gain for transversing the entire system should be the same regardless of the number of steps it takes. The genotype–phenotype maps will always have the first and the last phenotypes (000 and 111) coded by a single genotype (all zeros and all ones, respectively), and we will then vary the genetic architectures of the six intervening phenotypes by altering the $\mathbf{B}_{zg}$ matrices as detailed in the supplement.

In Fig. 6 we show that the mean passage time is increasing with the number of genotypes that code for each phenotype and hence with robustness. This is implemented by increasing the dimensionality of the genotypes from a three-digit to a six-digit Boolean number. The $\mathbf{B}_{zg}$ matrices are built by assigning as similar as possible number of genotypes to each phenotype with genotypes grouped according to the similarity of their Boolean number. In Fig. 7 we illustrate the negative relation between evolvability and robustness in another way by collapsing phenotypes while keeping the number of genotypes constant. We do this starting from a map from 64 genotypes to 8 phenotypes, and then combining the six intervening phenotypes into sets of three, two or one phenotype(s). This shows that the mean passage time is increasing with the increased size and hence intraconnectedness and robustness of the phenotypes despite the fact that there are fewer phenotypic steps and stronger selection for each step. This can also

65

be illustrated with the simpler map in Fig. 5 by collapsing the two intermediate phenotypes (01 and 10) into one, and letting the selection coefficient for the step from this to the phenotype 11 be 2*s*. Then the mean passage time is

$$\bar{t} = \frac{8u + 3b}{6bu} \approx \frac{1}{2u}\left(1 + \frac{2}{3sN}\right), \tag{11}$$

which is longer than the time computed above for the less robust system provided $sN > 1/12$; i.e. for all cases with a meaningful strength of selection. Generally, we may say that robustness hampers the evolvability of sequential adaptations because there are more genotypes for which no advantageous mutation exists, and when the population hits these it has to wait for slow neutral evolution to bring it to a genotype that allows advantageous mutations.

4. Discussion

Evolvability and robustness are crucial to the evolution of complex phenotypes but they have been thought to conflict with each other (reviewed in de Visser et al., 2003; Lesne, 2008). Wagner (2005, 2008), however, argued that evolvability and robustness understood as properties of phenotypes may be positively related if the genotype–phenotype map is many-to-few, so that each phenotype can be realized by a network of mutationally-connected genotypes (see also Ebner et al., 2002). This conjecture has been supported by models of RNA secondary structure, transcription-factor binding and gene regulation (Aldana et al., 2007; Wagner, 2008, 2012; Draghi et al., 2010; Payne and Wagner, 2014). Robustness has also been argued to facilitate evolvability through capacitance mechanisms allowing the accumulation of hidden genetic variation that can be released to fuel adaptive evolution (Rutherford, 2000; Hermisson and Wagner, 2004).

While we do not deny that large and connected neutral networks can be simultaneously robust and evolvable in the sense of Wagner (2008), we have shown that this is not a general result. As illustrated in Fig. 1 the correlation between phenotype evolvability and phenotype robustness depends on the specific architecture of the genotype–phenotype map. In our systematic investigation of all possible genotype–phenotype maps up to a certain level of complexity we found that maps with a positive relation between phenotype evolvability and robustness are indeed more common, but we also found many maps with negative correlations. In a more dramatic contradiction of Wagner's conjecture we found that average evolvability and robustness were consistently negatively related across genotype–phenotype maps leading to the prediction that evolvability and robustness may be consistently negatively related across characters even if they tend to be positively related across phenotypic states of the same character.

Schuster et al. (1994), Wagner (2012), and Payne and Wagner (2014) argue in their studies that one larger phenotype with high phenotype robustness and high phenotype evolvability tend to be surrounded by many smaller phenotypes that are less robust and evolvable. Looking at evolvability and robustness on the level of the map itself averages out the massive influence of the one large phenotype that drives the negative correlation between evolvability and robustness across phenotypes within a map.

A crucial inherent assumption of Wagner's argument and all the model systems used to support it is that the evolvability of a character state can be understood as the total number of other character states or phenotypes accessible from this character state. The evolvability of a RNA secondary structure is measured as the number of other secondary structures that can be reached by a single mutation from the focal structure. Although this can be rephrased as a probability of being able to find a specific structure, we think it represents a restricted and unusual view of evolution. In ef-

fect, it assumes that evolution happens on a nominal scale with no ordering of phenotypes. More commonly, complex adaptations are thought to involve gradual improvement through specific sequences of innovation or the near-continuous improvement of a quantitative trait. For evolution on ordinal or higher scale types it seems more likely with a tradeoff between evolvability and robustness, as we have illustrated with a simple model of adaptation through a sequence of phenotypes.

Another inherent assumption of Wagner's argument is that the mapping from genotype to phenotype is many-to-few. This is indeed a common assumption in many evolutionary arguments depending on systems drift (e.g. Schuster et al., 1994; Gavrilets and Gravner, 1997; True and Haag, 2001; Gavrilets, 2004; Wagner, 2005; Haag, 2007; Fierst and Hansen, 2010). But, as Hansen (2006a) argued in his review of Wagner (2005), this is not self-evident, because phenotypes can also be complex and many dimensional (see also Houle, 2010). Many of the models concocted by "geneticists" achieve their many-to-few mappings through extremely simplified assumptions as to what constitutes a phenotype. Even the relatively realistic models of secondary RNA structure provide a seriously depauperate view of the potential phenotypic complexity of RNA molecules. Although many nucleotide or even amino-acid substitutions may have none or insignificant phenotypic effects, we may ask whether such redundancy really matters for either evolvability or robustness on the phenotype level. This would require some form of epistatic interaction where neutral changes alter the probability of the changes that matter. If non-neutral changes are not affected by the state of the neutral component of the phenotype, then the latter simply does not matter for either evolvability or robustness at the phenotype level.

All of these arguments leave out the effects of within-population variation. In both Wagner's and our models, it is assumed either that population variation is insignificant or that it is uniformly distributed among the genotypes in a neutral network. In reality the population distribution would be determined by population-genetical processes such as mutation-selection balance that should generate a non-uniform distribution clustering in the more strongly connected parts of a network. Van Nimwegen et al. (1999) found that the mutation load of a neutral network decreases with its average connectedness. The population robustness of the phenotype can be equated with the inverse of the mutation load, and is hence predicted to increase with connectedness. This assumes that all other phenotypes have the same low fitness however, and it remains to be studied how this will pan out in complex genotype–phenotype maps generating several connected phenotypes with varying fitness. Evolvability is also crucially dependent on population variation. While the genetic variation stored in a neutral network may be "hidden" on the phenotypic level, it can facilitate access to other phenotypes through systems drift, and may also be revealed by decanalization due to environmental or genetic changes if the neutrality is caused by canalization in the first place (Hermisson and Wagner 2004).

Modularity and integration have long been thought crucial determinants of evolvability in complex characters (e.g. Wagner and Altenberg 1996; Hansen 2006b; Pavličev and Cheverud 2015). In our models, integration and complexity increase with increasing pleiotropy between the components of the phenotype. The more pleiotropic topologies indeed seem to present more variation in evolvability and robustness, as well as in the relation between them. Hansen (2003) and Pavličev and Hansen (2011) have previously found that an intermediate level of pleiotropy and a high degree of variation in types of pleiotropy may facilitate average evolvability over different directions in morphospace even if it may not optimize evolvability along particular directions of change. Robustness has not been investigated in this perspective, and future

66

studies may benefit from considering robustness as a trait that may vary across different morphological directions.

In conclusion, we have partially restored the "paradox" that Wagner (2008) resolved. This does not detract from the importance of the conceptual innovations of Wagner (2005, 2008), or questions the specific relationships identified by him and others, it simply shows that biology is complex and case dependent. While a positive relation between robustness and evolvability across phenotypic states may indeed be common in systems conforming to Wagner's assumptions, we suspect that trade offs between phenotype robustness and evolvability may be more common across characters as opposed to character states, or when adaptation requires transversing phenotypes along ordinal or continuous scale types.

Acknowledgments

Appendix. The mean time to adaptation

Given a matrix of transition probabilities, the time evolution of the probabilities of being in each state can be found with standard techniques from stochastic process theory (e.g. Karlin and Taylor, 1981). To find the distribution of the time to pass through the system we use the approach from Appendix B of Hansen et al. (2000). Briefly, amend the system with an additional absorbing state and add a transition at a high rate k to this from the final state of our system so that the final transition does not add significant time. Let t be the time to reach the additional absorbing state. If $f(t)$ is the probability density of t, then we can note that the amount of probability mass that has arrived at the terminal state is

$$y(t) = \int_0^t f(\tau)d\tau, \tag{A.1}$$

which means that $f(t) = dy/dt$ and thus that $f(t) = kx_n(t)$, where x_n is the probability mass of the final genotype in our system (e.g. of 111 for the system in Fig. 5). Hence, if we solve the system of differential equations for the x-vector, we can use the solution for $x_n(t)$ to find $f(t)$. Hence, we need to solve

$$\frac{d}{dt}\mathbf{x} = \mathbf{A}\mathbf{x}, \tag{A.2}$$

where $\mathbf{A}$ is amended with adding $-k$ to its nnth element. The solution to this system is $x(t) = e^{\mathbf{A}t}x(0)$. The vector Laplace transform of this is $\mathbf{L}(z) = -(\mathbf{A} - z\mathbf{I})^{-1}\mathbf{x}(0)$, and the moment generating function for t is hence the last element of the vector

$$\mathbf{M}(z) = k\mathbf{L}(-z) = -k(\mathbf{A} + z\mathbf{I})^{-1}\mathbf{x}(0), \tag{A.3}$$

where $\mathbf{x}(0)$ is the vector $\{1, 0, .., 0\}$, which puts all probability mass on the first element (e.g. 000 in Fig. 5). The mean of t can then be found by taking the derivative of $\mathbf{M}(z)$ in $z = 0$, or also calculated directly as

$$\bar{t} = \int_0^\infty \tau f(\tau)d\tau = \int_0^\infty \tau kx_n(\tau)d\tau$$

$$= k\int_0^\infty \tau\left[e^{\mathbf{A}\tau}\mathbf{x}(0)\right]_n d\tau = k\left[\mathbf{A}^{-2}\mathbf{x}(0)\right]_n. \tag{A.4}$$

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.jtbi.2017.07.004.

References

Albert, R., Othmer, H.G., 2003. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in Drosophila melanogaster. J. Theor. Biol. 223 (1), 1–18.

Aldana, M., Balleza, E., Kauffman, S., Resendiz, O., 2007. Robustness and evolvability in genetic regulatory networks. J. Theor. Biol. 245 (3), 433–448.

Bürger, R., Ewens, W., 1995. Fixation probabilities of additive alleles in diploid populations. J. Math. Biol. 33 (5), 557–575.

Conrad, M., 1990. The geometry of evolution. BioSystems 24 (1), 61–81.

de Visser, J.A.G.M., Hermisson, J., Wagner, G.P., Ancel Meyers, L., Bagher-i-Chaichian, H., Blanchard, J.L., Chao, L., Cheverud, J.M., Elena, S.F., Fontana, W., Gibson, G., Hansen, T.F., Krakauer, D., Lewontin, R.C., Ofria, C., Rice, S.H., von Dassow, G., Wagner, A., Whitlock, M.C., 2003. Evolution and detection of genetic robustness. Evolution 57 (5), 1959–1972.

Draghi, J.A., Parsons, T.L., Wagner, G.P., Plotkin, J.B., 2010. Mutational robustness can facilitate adaptation. Nature 463 (7279), 353–355.

Ebner, M., Shackleton, M., Shipman, R., 2002. How neutral networks influence evolvability. Complexity 7 (2), 19–33.

Espinosa-Soto, C., Padilla-Longoria, P., Alvarez-Buylla, E.R., 2004. A gene regulatory network model for cell-fate determination during Arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. Plant Cell 16 (11), 2923–2939.

Fierst, J.L., Hansen, T.F., 2010. Genetic Architecture and postgotic reproductive isolation: evolution of Bateson–Dobzhansky–Muller incompatibilities in a polygenic model. Evolution 64 (3), 675–693.

Fierst, J.L., Phillips, P.C., 2015. Modeling the evolution of complex genetic systems: The gene network family tree. J. Exp. Zool. Part B: Mol. Dev. Evol. 324 (1), 1–12.

Fontana, W., 2002. Modelling 'Evo-devo' with RNA. Bioessays 24 (12), 1164–1177.

Fontana, W., Konings, D.A., Stadler, P.F., Schuster, P., 1993. Statistics of RNA secondary structures. Biopolymers 33 (9), 1389–1404.

Frank, S.A., 1999. Population and quantitative genetics of regulatory networks. J. Theor. Biol. 197 (3), 281–294.

Gavrilets, S., 1999. A dynamical theory of speciation on holey adaptive landscapes. Am. Nat. 154 (1), 1–22.

Gavrilets, S., 2004. Fitness landscapes and the origin of species. Monographs in Population Biology. Princton University Press.

Gavrilets, S., Gravner, J., 1997. Percolation on the fitness hypercube and the evolution of reproductive isolation. J. Theor. Biol. 184 (1), 51–64.

Gerhart, J., Kirschner, M., 1997. Cells, Embryos and Evolution. Blackwell Science.

Gruener, W., Giegerich, R., Strothmann, D., Reidys, C., Weber, J., Hofacker, I.L., Stadler, P.F., Schuster, P., 1996. Analysis of RNA sequence structure maps by exhaustive enumeration I. Neutral Netw. Monatsh. Für Chem. Chem. Monthly 127 (4), 355–374.

Haag, E.S., 2007. Compensatory vs pseudocompensatory evolution in molecular and developmental interactions. Genetica 129 (1), 45–55.

Haldane, J.B.S., 1927. A mathematical theory of natural and artificial selection. V. selection and mutation. Math. Proc. Cambridge Philos. Soc. 23 (7), 838–844.

Hansen, T.F., 2003. Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evolvability. BioSystems 69 (2–3), 83–94.

Hansen, T.F., 2006a. The origins of robustness. Evolution 60 (2), 418–420.

Hansen, T.F., 2006b. The evolution of genetic architecture. Annu. Rev. Ecol. Evol. Syst. 37 (1), 123–157.

Hansen, T.F., Carter, A.J.R., Chiu, C.H., 2000. Gene conversion may aid adaptive peak shifts. J. Theor. Biol. 207 (4), 495–511.

Hansen, T.F., Wagner, G.P., 2001. Modeling genetic architecture: a multilinear theory of gene interaction. Theor. Popul. Biol. 59 (1), 61–86.

Hermisson, J., Wagner, G.P., 2004. The population genetic theory of hidden variation and genetic robustness. Genetics 168 (4), 2271–2284.

Houle, D., 2010. Numbering the hairs on our heads: the shared challenge and promise of phenomics. Proc. Natl. Acad. Sci. 107 (suppl 1), 1793–1799.

Karlin, S., Taylor H., E., 1981. A Second Course in Stochastic Processes. Elsevier.

Kauffman, S.A., 1969. Homeostasis and differentiation in random genetic control networks. Nature 224 (5215), 177–178.

Kauffman, S.A., 1993. The Origins of Order: Self Organization and Selection in Evolution. Oxford University Press.

Kimura, M., 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press.

Lesne, A., 2008. Robustness: confronting lessons from physics and biology. Biol. Rev. 83 (4), 509–532.

Pavličev, M., Cheverud, J.M., 2015. Constraints evolve: context dependency of gene effects allows evolution of pleiotropy. Annu. Rev. Ecol. Evol. Syst. 46 (1), 413–434.

Pavličev, M., Hansen, T.F., 2011. Genotype-phenotype maps maximizing evolvability: modularity revisited. Evol. Biol. 38 (4), 371–389.

Payne, J.L., Wagner, A., 2014. The robustness and evolvability of transcription factor binding sites. Science 343 (6173), 875–877.

Quayle, A.P., Bullock, S., 2006. Modelling the evolution of genetic regulatory networks. J. Theor. Biol. 238 (4), 737–753.

Rutherford, S.L., 2000. From genotype to phenotype: buffering mechanisms and the storage of genetic information. BioEssays 22 (12), 1095–1105.

Schuster, P., Fontana, W., Stadler, P.F., Hofacker, I.L., 1994. From sequences to shapes and back: a case study in RNA secondary structures. Proc. R. Soc. B 255 (1344), 279–284.

Thieffry, D., Romero, D., 1999. The modularity of biological regulatory networks. BioSystems 50 (1), 49–59.

True, J.R., Haag, E.S., 2001. Developmental system drift and flexibility in evolutionary trajectories. Evol. Dev. 3 (2), 109–119.

van Nimwegen, E., Crutchfield, J.P., Huynen, M., 1999. Neutral evolution of mutational robustness. Proc. Natl. Acad. Sci. 96 (17), 9716–9720.

Wagner, A., 2005. Robustness and Evolvability in Living Systems. Princeton University Press.

Wagner, A., 2008. Robustness and evolvability: a paradox resolved. Proc. R. Soc. B 275 (1630), 91–100.

Wagner, A., 2011. The Origins of Evolutionary Innovations: A Theory of Transformative Change in Living Systems. Oxford University Press.

Wagner, A., 2012. The role of robustness in phenotypic adaptation and innovation. Proc. R. Soc. B 279 (1732), 1249–1258.

Wagner, G.P., 2014. Homology, Genes, and Evolutionary Innovation. Princeton University Press.

Wagner, G.P., Altenberg, L., 1996. Complex adaptations and the evolution of evolvability. Evolution 50 (3), 967–976.

Supplement

To compute the passage times in Figure 6 we used the following matrices:

$$\mathbf{T} = \begin{bmatrix} u & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ b & u & 0 & 0 & 0 & 0 & 0 & 0 \\ 2b & b & u & 0 & 0 & 0 & 0 & 0 \\ 3b & 2b & b & u & 0 & 0 & 0 & 0 \\ 4b & 3b & 2b & b & u & 0 & 0 & 0 \\ 5b & 4b & 3b & 2b & b & u & 0 & 0 \\ 6b & 5b & 4b & 3b & 2b & b & u & 0 \\ 7b & 6b & 5b & 4b & 3b & 2b & b & u \end{bmatrix}, \tag{S.1}$$

and different $\mathbf{B}_{zg}$ matrices for the different dimensions of the genotype. For the 3-digit genotype we used a 8 x 8 identity matrix. For the 4-digit genotype we used

$$\mathbf{B}_{zg} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{S.2}$$

For the 5-digit genotype we used

$$\mathbf{B}_{zg} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{S.3}$$

and for the 6-digit genotype we used a 8 x 64 matrix that is too big to illustrate, but with similar structure to earlier cases except that it has eleven ones in rows 2 and 3, and ten ones in rows 4 to 7.

To make Figure 7 we used the same $\mathbf{B}_{zg}$ matrices for the 6-digit genotype, but modified

the T-matrix to collapse the phenotypes. To make three intermediate phenotypes we

used

$$
\mathbf{T} = \begin{bmatrix}
u & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
b & u & u & 0 & 0 & 0 & 0 & 0 \\
b & u & u & 0 & 0 & 0 & 0 & 0 \\
3b & 2b & 2b & u & u & 0 & 0 & 0 \\
3b & 2b & 2b & u & u & u & 0 & 0 \\
5b & 4b & 4b & 2b & 2b & u & u & 0 \\
5b & 4b & 4b & 2b & 2b & u & u & 0 \\
7b & 6b & 6b & 4b & 4b & 2b & 2b & u
\end{bmatrix},
$$
(S.4)

to implement the two intermediate phenotypes, we used

$$
\mathbf{T} = \begin{bmatrix}
u & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
b & u & u & u & 0 & 0 & 0 & 0 \\
b & u & u & u & 0 & 0 & 0 & 0 \\
b & u & u & u & u & 0 & 0 & 0 \\
4b & 3b & 3b & 3b & u & u & u & 0 \\
4b & 3b & 3b & 3b & u & u & u & 0 \\
4b & 3b & 3b & 3b & u & u & u & 0 \\
7b & 6b & 6b & 4b & 4b & 2b & 2b & u
\end{bmatrix},
$$
(S.5)

and for a single intermediate phenotype we used

$$
\mathbf{T} = \begin{bmatrix}
u & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
b & u & u & u & u & u & u & 0 \\
b & u & u & u & u & u & u & 0 \\
b & u & u & u & u & u & u & 0 \\
b & u & u & u & u & u & u & 0 \\
b & u & u & u & u & u & u & 0 \\
b & u & u & u & u & u & u & 0 \\
7b & 6b & 6b & 6b & 6b & 6b & 6b & u
\end{bmatrix}.
$$
(S.6)

2

70

PAPER IV

IV

Studying Developmental Variation with Geometric Morphometric Image Analysis (GMIA)

Christine Mayer, Brian D. Metscher, Gerd B. Müller, Philipp Mitteroecker*

Department of Theoretical Biology, Faculty of Life Sciences, University of Vienna, Althanstraße 14, A-1090, Vienna, Austria

*philipp.mitteroecker@univie.ac.at

Abstract

The ways in which embryo development can vary across individuals of a population determine how genetic variation translates into adult phenotypic variation. The study of developmental variation has been hampered by the lack of quantitative methods for the joint analysis of embryo shape and the spatial distribution of cellular activity within the developing embryo geometry. By drawing from the strength of geometric morphometrics and pixel/voxel-based image analysis, we present a new approach for the biometric analysis of two-dimensional and three-dimensional embryonic images. Well-differentiated structures are described in terms of their shape, whereas structures with diffuse boundaries, such as emerging cell condensations or molecular gradients, are described as spatial patterns of intensities. We applied this approach to microscopic images of the tail fins of larval and juvenile rainbow trout. Inter-individual variation of shape and cell density was found highly spatially structured across the tail fin and temporally dynamic throughout the investigated period.

Introduction

Despite the rapidly growing knowledge of the mechanisms underlying embryological development, little is known about how development varies across the individuals of a population. The variational properties of development determine how genetic and environmental variation translate into phenotypic variation in postnatal and adult individuals [1–8]. In turn, the population pool of phenotypic variation is the substrate for natural selection and, hence, a major determinant of organismal evolution [8, 9]. The lack of quantitative studies of

developmental variation has impeded the long-expected connection of developmental biology with the formal core of evolutionary theory. In addition, modeling developmental variation is key for understanding the multifactorial etiology of many diseases. Genetic and environmental factors that alter the pattern of developmental variation may increase the probability of individuals to pass a threshold towards pathological development [10,11].

The study of developmental variation has been hampered by the difficulties of measuring the geometry of developing embryos jointly with the spatial patterns of tissue formation and cellular activity. Yet an integrated understanding of organ formation and evolutionary change requires the coordinated study of gene expression, cellular activity, and organismal geometry [1,7,12]. In this paper, we present a novel approach that integrates geometric morphometrics and pixel- or voxel-based image analysis into a combined biometric method, allowing for the joint analysis of embryological shape and spatial patterns of tissue properties. For a demonstration, we apply this approach to a set of two-dimensional microscopic images of the tail fins of rainbow trout, but the approach can equally be applied to other imaging and staining methods as well as to three-dimensional images obtained from embryonic specimens [13,14].

Geometric morphometrics is the state-of-the-art method for biological shape analysis [15–19]. It is based on the representation of homologous point locations, curves, and surfaces by landmarks and semilandmarks (two- or three-dimensional measurement points). Semilandmarks are points on curves or surfaces for which the exact position along the curve or surface cannot be determined using anatomical criteria. They are estimated in the course of the analysis, establishing geometric homology within the sample [20,21]. The careful – usually manual – setting of the landmarks and semilandmarks, based on criteria of anatomical homology, leads to biologically interpretable estimates of means and variances and allows for an effective visualization of such statistical results as actual shapes or shape deformations [15]. However, this limits the application of geometric morphometrics to structures that are present and clearly visible in all individuals of the studied sample. With this method it is not possible to investigate the emergence or loss of structures, which is characteristic for embryological development. Nor does standard geometric morphometrics permit the assessment of structures with unclear boundaries, such as cell condensations or molecular gradients.

Statistical image analysis based on the gray values or color values of image elements (two-dimensional "pixels" or three-dimensional "voxels") is frequently used in medical imaging [22–24]. The variety of image analysis methods differ, among other aspects, in the way images are registered in order to yield correspondence across the compared pixels or voxels. Usually, the registration is an automatic or non-label based (without manual specification of points or curves), non-affine (non-linear) transformation that minimizes some measure of overall dissimilarity across the images [23,25]. Shape differences between individuals are often considered as nuisance rather than signal and hence are not explicitly estimated. These kinds of approaches have proven powerful for object

168

classification and computer vision, but the imperfect registration of the boundaries of homologous anatomical structures in different individuals may lead to sample averages and variance patterns that are not biologically meaningful [26]. For example, an average of well-delineated structures tends to have fuzzy boundaries, so that this average may no longer represent an actual anatomical structure. The variance of gray values or color values typically is concentrated at the misaligned edges of structures.

In the new method we term *Geometric Morphometric Image Analysis* (GMIA), we take advantage of the strengths of both approaches. It consists of two steps that represent two complementary ways in which developmental differences typically are described in biology. For structures with sharp boundaries, such as organs, bones, and other well-differentiated tissues, morphological variation is described in terms of variation in the shape and in the relative position, size, and orientation of these structures. Structures with diffuse boundaries, such as emerging cell condensations or molecular gradients, instead are described as spatial patterns of intensities or directions (scalar or vector fields) within the organism or within selected organs.

GMIA thus starts with a careful, manual or semi-automatic representation of homologous, well-defined anatomical point locations, curves, and surfaces by the assignment of a dense set of landmarks and semilandmarks. The positions of the semilandmarks are estimated by the sliding landmark algorithm, which minimizes the "bending energy" of the thin-plate spline interpolation, a measure of local form difference, between the specimens and their sample average [20, 21, 27]. The image registration consists of two steps. First, the landmark configurations are superimposed by Generalized Procrustes Analysis (a least squares-based rigid registration plus scaling; [28]), which standardizes for variation in overall position, scale, and orientation. The coordinates of the registered landmarks and semilandmarks, the so-called Procrustes shape coordinates, parameterize the shape of the digitized structures. Second, the actual images are all registered to the sample average shape of the landmark configurations by another use of the thin-plate spline interpolation [27]. In these registered images, the anatomical structures (point locations, curves, and surfaces represented by landmarks and semilandmarks) thus have exactly the same shape in all specimens, and the spaces in between the landmarks are interpolated as "smoothly" as possible (minimizing the integral of the squared second derivatives; see Methods section for details). In the vicinity of the aligned anatomical structures, the registered pixels are likely to represent homologous tissue locations within the sample. Depending on the actual imaging and staining methods, the gray values or RGB values (the values of red, green, and blue color channels) of these pixels represent the tissue properties of the imaged structures.

This approach yields two complementary sets of data: (1) the Procrustes shape variables, describing variation of well-differentiated anatomical structures, and (2) the texture of the registered images (i.e., the pixel or voxel values), representing variation in the spatial distribution of imaged tissue properties. Statistics and

169

resulting visualizations can be computed separately for shape and texture, and also jointly for both.

This two-step approach resembles the separate parameterization of shape and "shape-free" texture in active appearance models and related techniques, which have found wide application in face recognition and some areas of medical imaging [29, 30]. But while active appearance models are aimed at identifying or classifying objects from images that display variation in viewpoint, lighting, and other conditions, geometric morphometric image analysis is a biometric method for the joint analysis of shape and tissue properties in a biological context.

To illustrate this approach, we use a sample of 20 larval and juvenile specimens of rainbow trout (*Oncorhynchus mykiss*), seven of them fixed at 21 days post fertilization (dpf), eight at 40 dpf, and five at 56 dpf. All specimens were stained in the same way with Mayer's hematoxylin and were stored in 75% glycerol. Microscopic images, prepared under identical conditions, were taken of the tail fins or their corresponding precursors (see Methods section for details). Hematoxylin stains the cell nuclei, hence the color intensity in the microscopic images is taken to correspond with local cell density. We recorded the two-dimensional coordinates of four anatomical landmarks and 95 semilandmarks on each specimen in order to quantify the shape of the developing fin fold, the notochord, and the musculature (Fig. 1).

Materials and Methods

Staining and imaging

Our sample consists of 20 rainbow trout specimens (*Oncorhynchus mykiss*) from the zoological collection of the University of Vienna. The fish were collected from one hatchery in the year 2000. Specimens were euthanized by overdose of MS222, fixed in buffered formalin, and stored in 70% ethanol. Seven of them were fixed at 21 days post fertilization (dpf), eight at 40 dpf, and five at 56 dpf. In 2013, the specimens were transferred into distilled water before they were stained with Mayer's hematoxylin for 10 minutes. After keeping them in tap water for 15 minutes, they were transferred to a Scott Solution (tap water, sodium bicarbonate, magnesium sulfate) for 10 minutes, and then again rinsed in tap water for 10 minutes. The stained specimens were stored in 75% glycerol. The tail fins of all specimens were photographed with a Leica MZ16F stereomicroscope. All images were taken under identical light conditions and standardized specimen orientation with a 5× magnification (see S1, S2, S3 Figures). All use of trout specimens are in compliance with EU guidelines for the treatment of vertebrate animals in laboratory research (DIRECTIVE 2010/63/EU on the protection of animals used for scientific purposes). No specific ethical approval was necessary for the use of these collection specimens.
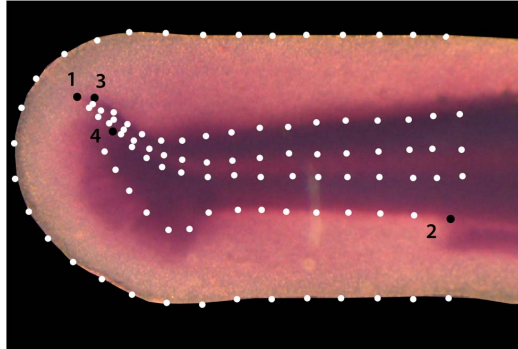
170

Fig. 1. Landmark configuration. Tail fin of a 21 dpf *O. mykiss* specimen with four anatomical landmarks (black points) and 95 semilandmarks (white points). The anatomical landmarks represent the tip of the notochord (1), the posterior end of the anal fin (2), and the two points where the dorsal and ventral outlines of the musculature meet the outline of the notochord (3, 4). The semilandmarks represent the outlines of the notochord, the musculature, and the fin fold. These landmarks were digitized on microscopic images of 20 *O. mykiss* specimens, seven of them fixed at 21 dpf, eight at 40 dpf, and five at 56 dpf.

doi:10.1371/journal.pone.0115076.g001

Landmarks

We recorded 4 anatomical landmarks and 95 semilandmarks on each of the 20 images (Fig. 1) using the software TPSdig (James Rohlf). The anatomical landmarks represent the tip of the notochord, the posterior end of the anal fin, and the two points where the dorsal and ventral outlines of the musculature meet the outline of the notochord. The semilandmarks represent the outlines of the notochord, the musculature, and the fin fold.

The positions of the semilandmarks along their corresponding curves were computed by the sliding landmark algorithm [15, 20, 21, 31], which iteratively "slides" the semilandmarks along tangents to the curve in order to minimize the bending energy of the thin plate spline (TPS) function between each individual and the sample average configuration. The tangent to the curve at each landmark was estimated by a vector connecting the two neighboring landmarks. The TPS function between two sets of landmarks is a non-affine interpolation that maps the landmarks exactly and the in-between space as "smoothly" as possible by minimizing the integral (in all two or three dimensions) of the squared second derivatives, a quantity referred to as bending energy [15, 27].

In most applications of the sliding landmark algorithm, the curves start and end with anatomical landmarks that constrain the sliding of the semilandmarks. As the curves in our application were all open, we computed the average landmark configuration only once and iteratively slid the semilandmarks against this average. This guaranteed convergence to a non-degenerate mean shape.

171

Shape analysis

The 20 configurations of landmarks and semilandmarks were superimposed by Generalized Procrustes Analysis [18, 28]. Thereby the configurations are translated to a common origin (the centroid – the coordinate-wise average of each configuration – is sent to the origin of the coordinate system), scaled to unit centroid size (square root of summed squared distances between every landmark and their centroid), and iteratively rotated to minimize the sum of squared distances among the homologous landmarks. The resulting Procrustes shape coordinates only describe the shape of each landmark configuration because variation in position, scale, and orientation was removed by the Procrustes registration. The Euclidean distance (square root of summed squared differences) between two sets of shape coordinates approximates the Procrustes distance, a measure of shape difference between the corresponding landmark configurations [16, 18].

Group mean shapes are estimated by averages of the shape coordinates. A low-dimensional ordination of shape space was computed by a between-group principal component analysis [32], which is an orthogonal projection of the individual shapes on the principal components of the group means. The between-group PCs thus constitute an orthonormal basis that optimally (in a least-squares sense) represents the Procrustes distances between the group mean. This technique typically leads to a better group separation than ordinary PCA and does not have the problems associated with discriminant analysis [32]. Shape differences between group means are visualized by thin-plate spline deformation grids [15]. These deformation grids are computed by applying the TPS interpolation between the two mean shapes to the vertices of a regular grid that is superimposed onto the template configuration [15, 18, 27].

Texture analysis

All images were registered to the sample mean shape based on the measured landmarks and semilandmarks by using the thin-plate spline interpolation [27]. In the registered images, the structures digitized by landmarks all have the same shape within the sample, and the space in between the landmarks is interpolated by minimizing the bending energy of the TPS function. The pixels of the registered images hence are assumed to represent corresponding anatomical structures. In order to avoid empty image elements (in areas of expansion) or overlapping pixels (in areas of compression), the images were subject to backward warping (also referred to as unwarping) instead of forward warping. In the typology of image registration techniques devised by Maintz & Viergever [23] this is an intrinsic (landmark-based), non-affine (or curved), local, semi-automatic, monomodal, intersubject registration with directly computed transformation parameters.

Mean cell density was computed as the average of the RGB values of the registered images, computed separately for every color channel of every pixel. For calculating group mean differences and variances of image texture, we

172

transformed the RGB values of each pixel into a scalar value by using the average of the RGB values of each pixel. This average corresponds to the brightness of the pixel and was interpreted as cell density in our application. Group mean differences and variances of image texture (cell density) were visualized by color maps. As for shape, the pattern of individual and group mean differences in texture were ordinated by between-group principal component analysis. Principal components of images are also referred to as eigenimages in the image analysis literature [24, 33] and are least-squares ordinations of the Euclidean distances between the image textures.

Joint analysis of shape and texture

Group mean shape and mean texture were jointly visualized by unwarping the average texture to the corresponding average shape. Group mean differences for both shape and texture were displayed by superimposing a TPS deformation grid and a color map of texture differences.

A joint ordination analysis of shape and texture is not possible via PCA because the scaling of the pixel values relative to the shape coordinates is ambiguous. While both shape and texture can be separately equipped with a Euclidean metric (but see also [34]), the arbitrary scaling between shape and texture prevents the use of a common Euclidean metric. To approach a common ordination of shape and texture, we instead used a two-block partial least squares analysis (PLS [35, 36]). PLS yields two linear combinations (each with squared coefficients summing up to 1), one for shape and one for texture, that have the maximum possible covariance. When scaling the two linear combinations via major axis regression so that they also correspond in magnitude, the combined scaled coefficient vectors can be interpreted as a common factor of joint variation (for details see [37]). After projecting both shape and texture in the subspaces perpendicular to these vectors, a second pair of vectors can be computed in the same way, and similarly for further dimensions. We used the first two dimensions of such a scaled PLS analysis to represent the joint variation of shape and texture.

All analyses and visualizations of shape and texture were computed in Mathematica 9.0 (Wolfram Research Inc., Champaign, IL, USA)

Results

After estimating the semilandmark positions and superimposing the configurations by Generalized Procrustes Analysis, the resulting shape coordinates were averaged for each of the three age groups in order to estimate the *age-specific mean shapes*. Fig. 2a illustrates how the average size of the fin fold and the musculature increased relative to the notochord, and how the fin changed from a rounded to a more triangular shape.

After registering all images to the same shape, texture can likewise be averaged for the three age groups in order to estimate age-specific image texture, which, in
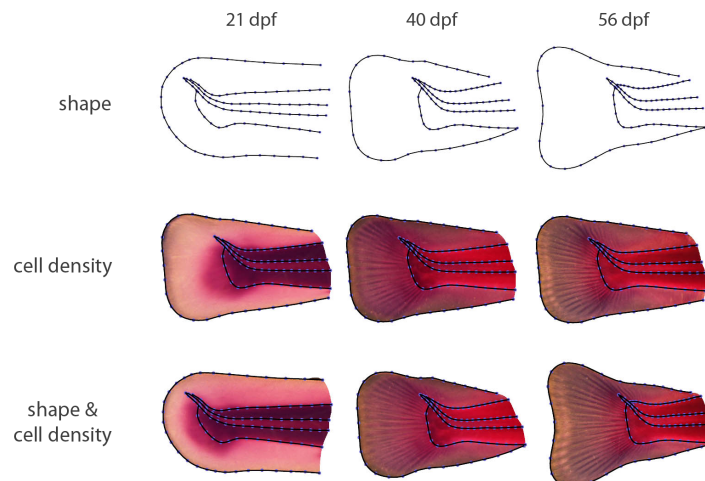
173

Fig. 2. Averages of fin shape, cell density, and both together. Average fin shape (first row), average cell density (second row), and average fin shape together with average cell density (third row) for each of the three age groups (the three columns). The structures measured by (semi)landmarks – the outlines of the fin fold, the musculature, and the notochord – are perfectly registered, but note also how precisely the fin rays, which emerge at 40 dpf, are registered by the TPS interpolation, even though they are not measured by landmarks.

our case, represents the *spatial pattern of average cell density.* While no fin rays were visible at 21 dpf, they were partly formed at 40 dpf, and were extended towards anterior at 56 dpf (Fig. 2b). The structures measured by (semi)landmarks – the outlines of the fin fold, the musculature, and the notochord – necessarily are perfectly registered. Note, however, how well the fin rays are registered by the TPS interpolation, even though they were not measured by landmarks (compare these averages with the individual images in S1, S2, S3 Figures).

In Fig. 2c, the average image textures are warped to the corresponding average shape, allowing for the joint visualization of age-specific average shape and average cell density.

The differences between the age groups, i.e., the *developmental transformations*, can explicitly be visualized by superimposing a thin-plate spline deformation grid, depicting developmental shape change, and a color map that represents increase or decrease of cell density. For this purpose, the full color information for each pixel was reduced to brightness (average of the RGB values). The higher the cell density (number of projected stained cell nuclei), the darker the pixel will appear and the lower the brightness will be. Fig. 3a shows the transformation from a rounded to a more triangular fin shape between 21 dpf and 40 dpf, which was associated with an average increase of cell density in the fin fold and a decrease of cell density in the musculature and the notochord. Between 40 dpf and 56 dpf, average cell density decreased in the musculature, the notochord, and also in between the fin rays, while density increased inside the fin rays (Fig. 3b).
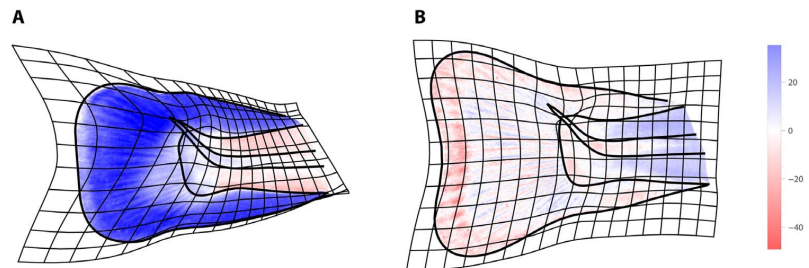
174

Fig. 3. Visualization of average shape change together with average change of cell density. (A) Between 21 dpf and 40 dpf and (B) between 40 dpf and 56 dpf. The deformation grid shows how the shape of the fin changed from a rounded to a more triangular shape during both periods. Changes of cell density are represented by the color map. From 21 to 40 dpf, cell density increased (blue) in the fin, whereas it decreased (red) in the musculature and the notochord. From 40 to 56 dpf, cell density increased in the fin rays and decreased in between.

The spatial pattern of *individual variation of cell density* is also visualized via a color map (Fig. 4). Clearly, variation in cell density was not uniform across the tail fin and differed between the age groups. This reflects different ongoing developmental processes with varying degrees of canalization. At 21 dpf, the variance in cell density was mainly located in the fin fold, whereas cell density in the notochord and in the musculature was relatively similar in all individuals (Fig. 4a). At 40 dpf, variation had increased in the musculature and, after the fin rays had formed, cell density was more variable in between fin rays than within the fin rays (Fig. 4b). At 56 dpf, variance of cell density in the musculature had reduced again, whereas variation in between the fin rays had increased. Note that due to the precise registration, variance is located within the assessed structures, not at their boundaries.

The multivariate pattern of individual differences can be assessed by an ordination analysis that yields a low-dimensional diagram in which the distances between the individuals approximate a certain measure of multivariate (dis)similarity. The shape metric typically used in geometric morphometrics is Procrustes distance (approximated by the Euclidean distance between the sets of shape coordinates [16]). As a metric for image texture, we likewise used the Euclidean distance between the sets of RGB values [33], even though other metrics, e.g. based on mutual information, are equally possible.

We used a between-group principal component analysis [32] to ordinate the multivariate shape differences among the specimens (Fig. 5a). In the scatter plot of the first two between-group principal components (PCs), each symbol represents one individual, and the distance between two symbols approximates the magnitude of overall shape difference between the respective individuals. The first two PCs represent the shape distances among the three group means exactly and account for 91% of total variation between the 20 individual shapes. The shape features corresponding to the two axes are visualized by reconstructed shapes along the corresponding axis locations. The scatter plot shows that –
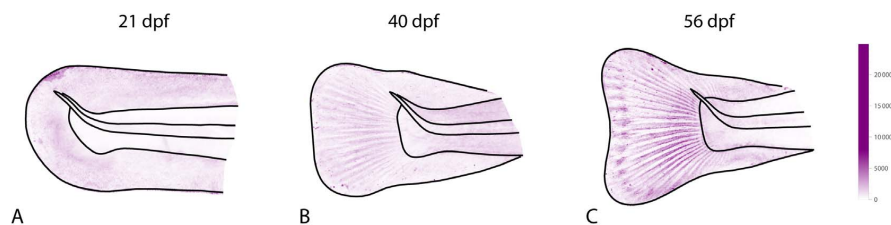
175

21 dpf   40 dpf   56 dpf

Fig. 4. Individual variation in cell density. The spatial pattern of variation in cell density is shown by a color map for each age group. (A) 21 dpf, (B) 40 dpf, (C) 56 dpf. Variation in cell density was not uniform across the tail fin and differed between the age groups. At 21 dpf, the variance was concentrated in the fin fold, whereas cell density was very similar across all individuals in the notochord and in the musculature. At 40 dpf and 56 dpf, variance between the fin rays was higher than variance of cell density in the fin rays.

doi:10.1371/journal.pone.0115076.g004

despite considerable individual variation – the three age groups differed in average fin shape and were even separated without individual overlap. The groups differ along PC 1, mainly representing the relative size increase of the fin fold, associated with a shape change from a rounded to a triangular shape. More of these changes had occurred between 21 dpf and 40 dpf as compared to the second period. The second principal component, representing the shape of the fin fold independent of its size, appeared considerably more variable at 21 dpf as compared with the two later stages.

Fig. 5b shows the first two between-group principal components of the registered images, which account for 58% of individual variation in image texture (cell density) and for all the variation between the three group means. In the scatter plot, the symbols represent individuals and the distances between them approximate the magnitude of overall difference in texture. As for shape, the age groups were clearly separated with regard to the spatial pattern of cell density. Within the three age groups, the individuals varied mostly along PC1, which resembles the average differences between 21 dpf and 40 dpf: a shift of cell density from the musculature and the notochord towards the fin fold, along with the formation of fin rays.

Fig. 5c provides a joint ordination of both shape and image texture, based on a scaled partial least squares analysis (see Methods section). The visualizations along the corresponding axes hence comprise differences in fin shape as well as in cell density.

Quantitative developmental studies, such as our example, typically are exploratory; statistical tests of the usual null-hypotheses do not bear much biological relevance here. However, as already obvious from the ordination analyses, permutation tests indicated significant group mean differences in shape as well as in image texture between the age groups ($P<0.0013$ for each of the pairwise comparisons, using the Euclidean distance between group means as test statistic).
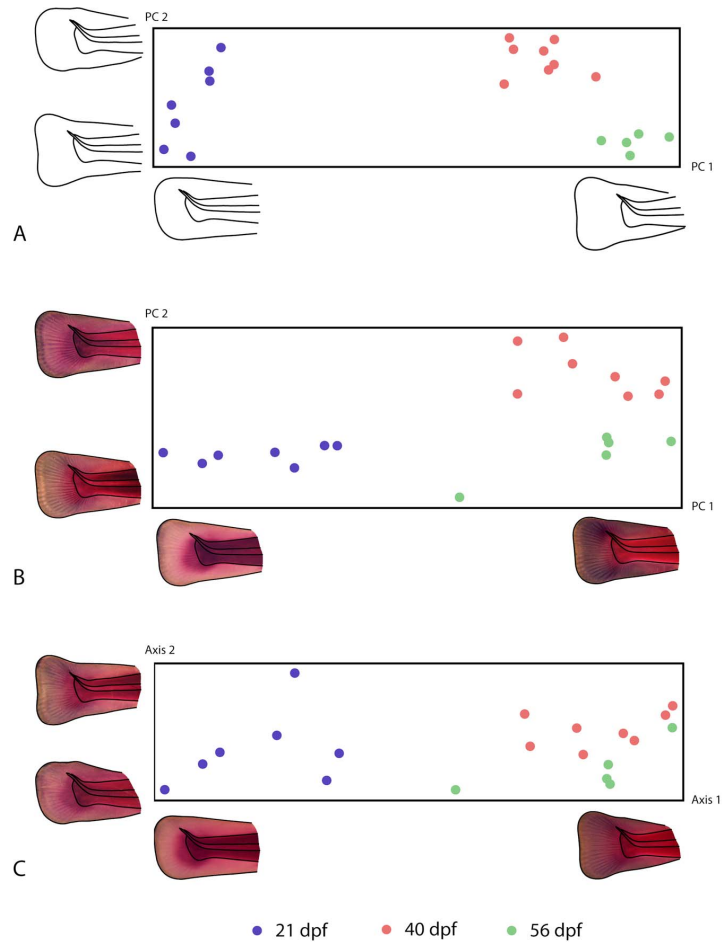
Fig. 5. Principal component analyses. Principal component analyses of (A) fin shape and (B) cell density, as well as (C) a joint ordination of both fin shape and cell density based on a scaled partial least squares analysis [37]. Each symbol in the scatter plots corresponds to one individual, and the distance between individuals approximates the overall amount of shape difference or difference in cell density. The axes of these plots correspond to patterns of shape change, to patterns of change in cell density, and to a combination of both, respectively. These patterns are visualized by reconstructed shapes and cell density patterns along the axes that correspond to the limits of variation occurring in the sample (approximately 3 standard deviations from the mean).

Discussion

Population models in evolutionary theory, genetics, and epidemiology are based on quantitative representations of phenotypic and genetic variation. Yet variation in embryological development and growth – the processes translating genetic variation into phenotypic variation – is still poorly understood. In fact, it has

177

rarely been addressed empirically. Nevertheless, it has been argued on theoretical grounds that the properties of development assume a major role in shaping phenotypic variation within and across populations [1–8]. For example, the vast amount of genetic variation is assumed to be "funneled" and structured by the limited number of possible developmental pathways into a much lower-dimensional pattern of phenotypic variation [2, 38]. Strong developmental constraints on the expression of genetic variation may even cause phenotypic stasis [39]. Already in the mid-20th century, C. H. Waddington [40] emphasized the importance of developmental canalization for buffering genetic and environmental variation and the accumulation of cryptic genetic variation, but the actual molecular and developmental mechanisms underlying canalization are still not well understood. The pivotal role of the pattern of developmental variation in humans has recently been adopted by the "Developmental Origin of Health and Disease" paradigm [10, 11].

The study of developmental variation has been hampered by difficulties with the quantification of embryological traits. Our morphometric approach combines the strengths of geometric morphometrics and pixel- or voxel-based image analysis. Well-defined tissue structures are described by the shape of their boundaries, whereas diffuse spatial patterns, such as cell condensations or molecular gradients, are described as scalar fields that are extracted from the texture of the registered images. We demonstrated how to separately analyze embryological shape and image texture (cell density), and we also outlined a strategy for their joint analysis. The anatomical structures are perfectly registered by our method and, hence, group averages of anatomical structures have well-defined boundaries. Variance in image texture is concentrated within the structures, not at their misaligned boundaries [26]. This separate parameterization of shape and image texture resembles the approach in active appearance models and related techniques [29, 30], but the biometric strategy originates from geometric morphometrics [18, 41] and statistical parametric mapping in voxel-based image analysis [22, 24].

In our application of GMIA we analyzed two-dimensional microscopic images. The method can equally be used with other imaging methods, including 3D imaging such as confocal microscopy and micro-CT [13, 14, 42]. Most importantly, it can also be applied to staining methods more specific than the hematoxylin used in our study. Other developmental techniques, such as in situ hybridization or antibody staining, can also be combined with GMIA and will permit the statistical analysis of spatial patterns of gene expression within a population of developing embryos.

Several components of the GMIA approach, most notably the registration of the images in between the landmarks, were based on the TPS interpolation. This algorithm has proved powerful in multiple morphometric contexts, and it worked excellently for registering the fin rays in our application. But the algorithm originated in material physics [27] and may not be optimal in all biological contexts. Likewise, the way in which we transformed RGB values and differences

178

between RGB values into scalar values worked well for our images, but it may need adjustment for other staining and imaging techniques.

By applying GMIA to a sample of rainbow trout tail fins, we were able to demonstrate how average fin shape and average cell density changed within a period of 35 days. These changes include the emergence of the fin rays as novel tissue structures between 21 and 40 dpf, a process associated with a general shift of cell density from the musculature to the fin fold (Figs. 2 and 3). Inter-individual variation of cell density was highly spatially structured and temporally dynamic throughout the investigated period (Fig. 4; compare also [43]). For example, the average increase of cell density within the fin fold between 21 and 40 dpf was accompanied by a reduction (canalization) of variation in cell density in the fin fold and with a temporary increase of variation in the musculature. Between 40 and 56 dpf, variation in the musculature decreased again. The ordination analyses (Fig. 5) demonstrate that the individuals clearly separated between the age groups, even though this separation was more pronounced for shape than for cell density. Within the limits of the small sample, these analyses further indicate that variation along PC2 of shape (i.e., rounded versus triangular shape of the fin fold, independent of its relative size) canalized within the observed age range, whereas no such general trend was apparent for cell density.

In order to verify these results, we produced histological sections along the frontal plane for one fish of each age class. Between 21 and 40 dpf, the tail fin became thicker and the density of mesenchymal cells increased. Also, the number of mucous glands increased. The myotomes decreased in width relative to the notochord. This corresponds well to our results of increasing (projected) cell density in the fin fold and decreasing density in the musculature. Between 40 and 56 dpf, the relative thickness of the myotomes increased again, associated with an increase of visible muscle fibers by approximately 50%. The fin rays and the in-between connective tissue were well differentiated at 40 and 56 dpf. Again, this is represented by our results (Fig. 3).

Highly variable embryonic structures, resulting from variation in the onset, tempo, or mode of developmental processes, are particularly responsive to environmental or genetic disturbances and, hence, are promising candidates for experimental studies. These loci of developmental variation constitute the pattern of phenotypic variation that is subject to natural selection during development and adulthood. But because tissues differ in their degrees of canalization, variation generated at early developmental stages is not equally maintained during later stages and may induce different evolutionary dynamics (see also [44, 45]).

GMIA presents itself as a powerful biometric tool for studying variation in organismal shape in concert with variation in the spatial patterns of various tissue properties. This approach may foster research in an emerging field of biomedical science, the study of developmental variation, which, at the same time, is central to any formal connection between evolutionary and developmental biology. More generally, the inherent strategy underlying GMIA can be used to study the statistical properties of various scalar fields, vector fields, or even tensor fields embedded within in the geometries of different organisms. This includes color

179

patterns, such as in butterfly wings or in human facial skin, or even biomechanical properties, such as the distribution of stress and strain in samples of adult or subadult individuals (e.g., [46, 47]). This capacity of GMIA – the combined registration of variation in embryonic shape, gene expression, and physical properties of cell and tissue masses – can foster the integration of developmental biology and EvoDevo with the population genetic account of evolutionary theory. It may thus provide an important element of an expanded formal framework, or an Extended Synthesis [8] of evolutionary theory.

Supporting Information

S1 Figure. Images of the 21 dpf specimens.
doi:10.1371/journal.pone.0115076.s001 (TIF)

S2 Figure. Images of the 40 dpf specimens.
doi:10.1371/journal.pone.0115076.s002 (TIF)

S3 Figure. Images of the 56 dpf specimens.
doi:10.1371/journal.pone.0115076.s003 (TIF)

Acknowledgments

Author Contributions

Conceived and designed the experiments: PM. Performed the experiments: CM. Analyzed the data: CM PM. Contributed reagents/materials/analysis tools: BDM GBM. Wrote the paper: CM BDM GBM PM.

References

1. **Müller GB** (2007) Evo-devo: extending the evolutionary synthesis. Nat Rev Genet 8: 943–949. doi:nrg2219 [pii] 10.1038/nrg2219

2. **Hallgrimsson B, Lieberman DE** (2008) Mouse models and the evolutionary developmental biology of the skull. Integr Comp Biol 48: 373–384.

3. **Mitteroecker P** (2009) The Developmental Basis of Variational Modularity: Insights from Quantitative Genetics, Morphometrics, and Developmental Biology. Evol Biol 36: 377–385.

4. **Wagner GP, Altenberg L** (1996) Complex adaptations and the evolution of evolvability. Evolution (N Y) 50: 967–976.

5. **West-Eberhard MJ** (2003) Developmental Plasticity and Evolution. New York: Oxford University Press.

6. **Salazar-Ciudad I** (2006) Developmental constraints vs. variational properties: How pattern formation can help to understand evolution and development. J Exp Zool Part B-Molecular Dev Evol 306B: 107–125. doi:Doi 10.1002/Jez.B.21078

7. **Müller GB, Newman SA** (2003) Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology. Vienna Ser Theor Biol. Cambridge: MIT Press.

180

8. **Pigliucci M, Müller GB** (2010) Evolution – The Extended Synthesis. Cambridge: MIT Press.

9. **Roff DA** (1997) Evolutionary quantitative genetics. New York: Chapman & Hall.

10. **Gluckman PD, Hanson MA, Low FM** (2011) The role of developmental plasticity and epigenetics in human health. Birth Defects Res C Embryo Today 93: 12–18.

11. **Low FM, Gluckman PD, Hanson MA** (2012) Developmental Plasticity, Epigenetics and Human Health. Evol Biol 39: 650–665. doi:10.1007/s11692-011-9157-0

12. **Hallgrimsson B, Hall BK** (2011) Epigenetics: Linking Genotype and Phenotype in Development and Evolution. University of California Press.

13. **Louie AY, Huber MM, Ahrens ET, Rothbacher U, Moats R, et al.** (2000) In vivo visualization of gene expression using magnetic resonance imaging. Nat Biotechnol 18: 321–325. doi:10.1038/73780

14. **Metscher BD, Müller GB** (2011) MicroCT for Molecular Imaging: Quantitative Visualization of Complete Three-Dimensional Distributions of Gene Products in Embryonic Limbs. Dev Dyn 240: 2301–2308.

15. **Bookstein FL** (1991) Morphometric tools for landmark data: geometry and biology. Cambridge (UK); New York: Cambridge University Press.

16. **Dryden IL, Mardia KV** (1998) Statistical Shape Analysis. New York: John Wiley and Sons.

17. **Slice D** (2007) Geometric Morphometrics. Annu Rev Anthr 36: 261–281.

18. **Mitteroecker P, Gunz P** (2009) Advances in geometric morphometrics. Evol Biol 36: 235–247.

19. **Klingenberg CP** (2010) Evolution and development of shape: integrating quantitative approaches. Nat Rev Genet 11: 623–635.

20. **Bookstein FL** (1997) Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Med Image Anal 1: 225–43.

21. **Gunz P, Mitteroecker P** (2013) Semilandmarks: A Method for Quantifying Curves and Surfaces. Hystrix 24: 103–109.

22. **Ashburner J, Friston KJ** (2000) Voxel-Based Morphometry—The Methods. Neuroimage 11: 805–821.

23. **Maintz JBA, Viergever MA** (1998) A survey of medical image registration. Med Image Anal 2: 1–36. doi:10.1016/S1361-8415(01)80026-8

24. **Penny WD, Friston KJ, Ashburner JT, Kiebel SJ, Nichols TE** (2007) Statistical Parametric Mapping: The Analysis of Functional Brain Images. London: Academic Press.

25. **Zitová B, Flusser J** (2003) Image registration methods: a survey. Image Vis Comput 21: 977–1000. doi:10.1016/S0262-8856(03)00137-9

26. **Bookstein FL** (2001) "Voxel-Based Morphometry" Should Not Be Used with Imperfectly Registered Images. Neuroimage 14: 1454–1462.

27. **Bookstein FL** (1989) Principal warps: Thin plate splines and the decomposition of deformations. IEEE Trans Pattern Anal Mach Intell 11: 567–585.

28. **Rohlf FJ, Slice DE** (1990) Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst Zool 39: 40–59.

29. **Lanitis A, Taylor C, Cootes T** (1995) Automatic face identification system using flexible appearance models. Image Vis Comput 13: 393–401.

30. **Cootes T, Edwards G, Taylor C** (2001) Active appearance models. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(6):681–685.

31. **Gunz P, Mitteroecker P, Bookstein FL** (2005) Semilandmarks in three dimensions. In: Slice DE, editor. Modern Morphometrics in Physical Anthropology. New York: Kluwer Press. pp. 73–98.

32. **Mitteroecker P, Bookstein FL** (2011) Linear discrimination, ordination, and the visualization of selection gradients in modern morphometrics. Evol Biol 38: 100–114.

33. **Zhang D, Chen S, Liu J** (2005) Representing Image Matrices: Eigenimages Versus Eigenvectors. Lect Notes Comput Sci 3497.

34. **Huttegger S, Mitteroecker P** (2011) Invariance and Meaningfulness in Phenotype Spaces. Evol Biol 38: 335–352.

181

35. **Bookstein FL** (1994) Partial Least Squares: A Dose-response Model for Measurement in the Behavioral and Brain Sciences. Psycoloquy 5.

36. **Rohlf FJ, Corti M** (2000) The use of two-block partial least-squares to study covariation in shape. Syst Biol 49: 740–753.

37. **Mitteroecker P, Bookstein FL** (2007) The Conceptual and Statistical Relationship between Modularity and Morphological Integration. Syst Biol 56: 818–836.

38. **Martinez-Abadias N, Mitteroecker P, Parsons TE, Esparza M, Sjovold T, et al.** (2012) The Developmental Basis of Quantitative Craniofacial Variation in Humans and Mice. Evol Biol 39: 554–567.

39. **Maynard Smith J, Burian R, Kauffman S, Alberch P, Campbell J, et al.** (1985) Developmental Constraints and Evolution: A Perspective from the Mountain Lake Conference on Development and Evolution. Q Rev Biol 60: 265–287.

40. **Waddington CH** (1942) The canalization of development and the inheritance of acquired characters. Nature 150: 563.

41. **Bookstein FL** (1991) Morphometric tools for landmark data: geometry and biology. New York: Cambridge University Press.

42. **Sensen CW, Hallgrimsson B** (2009) Advanced Imaging in Biology and Medicine: Technology, Software Environments, Applications. Berlin: Springer.

43. **Mitteroecker P, Bookstein F** (2009) The ontogenetic trajectory of the phenotypic covariance matrix, with examples from craniofacial shape in rats and humans. Evolution (N Y) 63: 727–737. doi:EVO587 [pii] 10.1111/j.1558-5646.2008.

44. **Mitteroecker P, Gunz P, Neubauer S, Muller G** (2012) How to Explore Morphological Integration in Human Evolution and Development? Evol Biol 39: 536–553.

45. **Bookstein FL, Mitteroecker P** (2014) Comparing Covariance Matrices by Relative Eigenanalysis, with Applications to Organismal Biology. Evol Biol.

46. **Newman SA, Müller GB** (2000) Epigenetic mechanisms of character origination. J Exp Zool 288: 304–317. doi:10.1002/1097-010X(20001215)288:4<304::AID-JEZ3>3.0.CO;2-G

47. **O'Higgins P, Fitton LC, Phillips R, Shi JF, Liu J, et al.** (2012) Virtual Functional Morphology: Novel Approaches to the Study of Craniofacial Form and Function. Evol Biol 39: 521–535. doi:Doi 10.1007/S11692-012-9173-8

182