1. Isaac Salazar-Ciudad

How predictable is the genotype-phenotype map: combining developmental biology and quantitative genetics

Institute of Biotechnology, 3 years (09/2018-08/2021)

2. Rationale in brief

One vision of the future is an era of 'personal genomics', wherein '-omic' data will predict the phenotype of an individual (e.g. disease susceptibility, morphological anomalies, etc…). This requires a mechanistic description of how genetic variation leads to specific phenotypic variation (and why to that phenotypic variation and not to other). This requires, in other words, to understand the relationship between genetic and phenotypic variation, also called the genotype-phenotype map or GP map (Lewontin, 1974; Alberch, 1991). How to develop predictive models of such relationship is an open question.

One interesting kind of phenotypes are multivariate quantitative phenotypes. These are phenotypes that are only adequately described by multiple traits (i.e. multivariate) that take quantitative continuous or nearly-continuous values. An example is morphology. The morphology of an organ, for example the wing of a fly, can not be described by a one trait, for example its length, nor even by two, its length and width. A quantitative description of wing morphology, necessarily requires several traits (i.e. the length of each of its veins and its perimeter, Figure 1A). There are several theoretical approaches to the GP map of multivariate quantitative phenotypes. These can be aligned along an axis going from the very statistical approach of quantitative genetics to the mechanistic approach of mathematical models of developmental biology and evolution.

The first one, quantitative genetics, is perceived to lay, together with population genetics, at the heart of classic evolutionary theory and animal and plant breeding (Falconer & Mackay, 1996). One of the aims of quantitative genetics is to disentangle which proportion of the phenotypic variance in a trait is due to genetic variance, which to additive genetic variance and which to the environment and other sources. Additive genetic variance is the variance due to alleles that have a fixed effect on the value of a phenotypic trait. This is variance due to alleles whose effect does not depend on the other alleles in the same locus or, more in general, to alleles in other loci. There are several ways to estimate such variance but most relay in comparing the phenotypic trait values between individuals of known genetic relatedness (e.g regression between parent and offspring phenotypic trait values). From the additive variance of a trait it is possible to estimate how a trait average will change in a population through the breeder's equation:

$$R = h^2 S$$

Where h2, the trait's narrow sense heritability, is the ratio between additive genetic variance and total phenotypic variance, R is the response to selection (the difference between the trait average in one generation and the next) and S, the selection differential, is a measure of how much does fitness increase with increases in a trait value (Falconer & Mackay 1996). Fitness is a measure of the genetic contribution of an individual to the next generation. With truncation selection, where a fixed proportion of the population is chosen to reproduce and found the next generation, S is equal to the difference in mean trait values between the selected individuals and the entire population.

A multivariate extension of breeder's equation has been developed (Lande, 1979; Lande & Arnold, 1983). This extension, Lande's equation, takes into account that in populations, traits may co-vary and that, then, selection in one trait may result in indirect selection to other traits:

$$\Delta \check{z} = G P^{-1} s$$

Where G is the matrix of additive genetic covariances between traits, P is the matrix of phenotypic covariances between traits, $\Delta \check{z}$ is the response to selection (the change in each trait average) and s is,

again, how fitness changes with changes in individual's trait values. There is a very large number of studies on the G matrix (e.g: estimating such matrix (Roff, 2007) or how much G changes over generations (Arnold *et al.*, 2008; Roff *et al.*, 2012; Björklund *et al.*, 2013; Björklund & Gustafsson, 2015; Chebib and Guillaume, 2017)). There are not many studies where such matrix is used to predict phenotypic change, or dedicated to measuring the accuracy of these predictions. Lande's equation, however, makes a number of assumptions. Most importantly, it assumes a normal distribution of both parental genotypes and offspring phenotypes, implicitly assuming a large number of fixed effect loci and a GP map that can be adequately described linearly, at least locally (Pigliucci & Schlichting, 1997; Pigliucci, 2006, 2007). These assumptions are not always in agreement with what is known about the often very non-linear gene product interactions in the embryonic development that produces morphology and its variation (Pigliucci & Schlichting, 1997; Pigliucci 2006, 2007). This may be specially relevant in the context of complex multivariate phenotypes (Salazar-Ciudad and Jernvall, 2005).

The quantitative genetics approach is statistical. The association between genetic and phenotypic variation in the GP map can be accurately described but the mechanisms by which such phenotypic variation is produced from interactions between gene products (for example in development) are not explicitly considered. Then, it is difficult to understand, from this approach alone, why the GP map is the way it is or how it may evolve (Piglucci 2007; Salazar-Ciudad & Jernvall, 2005). There are some models in evolutionary genetics (Rice, 2002; Johnson and Barton 2005; Hansen *et al.*, 2006) that do not consider, as such, the mechanisms by which the phenotype is constructed (e.g. development) but consider that the phenotypic effect of an allele may depend on that of the alleles in other loci (this is epistasis) and that an allele may have an effect in several traits (pleiotropy). These models have been applied to generic questions on the importance of pleiotropy in evolution but tend to either be too abstract to be applied to specific phenotypes or make assumptions that are not very compatible with how genes interact in development (Pigliucci, 2007).

The second approach to the GP map of multivariate morphology is evolutionary developmental biology (or evo-devo). This approach is devoted to understand how the way the morphology is built in development affects the way it can vary and then morphological evolution. There is a rich theoretical literature in the field, most notably since the 80s (Oster & Alberch, 1981; Maynard Smith et al., 1985; Newman & Comper, 1990; Raff, 1996; Arthur, 2004; Forgacs and Newman, 2005; Salazar-Ciudad 2006). While classical population and quantitative genetics are based, at least in its origins, on mendelian genetics, this approach is based on developmental genetics and biology. The point here is to understand how gene products interact in such a way that the morphology of organisms is produced. This involves also considering cells, their cell-cell signalling and their mechanical interactions in the construction of the phenotype. Since any change in morphology is first a change in the developmental processes that produce it, understanding those processes should inform about which morphological variation can arise and for which genetic variation. Traditionally, developmental biology has not been interested in population-level variation but the advent of evo-devo from the 80s has changed that (Alberch, 1982). Development, however, is a quite complex process. Because of that, most model systems are not understood to the point where predictions on possible phenotypic variation in the short-term and evolution would be possible. Even when detailed information about a developmental process is available, it is not obvious how to integrate such knowledge to predict specific morphological variation. Mathematical models are helpful in this sense. There are currently several emerging theories about how to mathematically represent gene and cell interactions during development to understand, or even predict, how embryonic morphology changes over development (Graner & Glazier, 1992; Honda et al., 2004; Newman, 2005, Marin-Riera et al., 2016). These are also being integrated with ongoing theories in evo-devo (Forgacs and Newman, 2005; Salazar-Ciuadd 2006; Jaeger *et al.*, 2012).

One of the few examples where multivariate morphological variation in a natural population can be partially predicted from a mathematical model of development is my own work on teeth (Salazar-Ciudad and Jernvall, Nature, 2010). This is a mechanistic model that mathematically implements how genes interact, how cell mechanically interact and cell adhesion and proliferation.

The model includes an hypothesis, based on previous experimental work over the years, about the basic gene regulatory network in tooth development and about how it regulates cell signaling, cell adhesion and proliferation. The model is mechanistic in the sense that from this hypothesis and some very simple initial conditions (i.e. a flat epithelium representing the initiation of tooth development) the model reproduces, without further intervention, how the morphology and patterns of gene expression in 3D arise during tooth development until an adult morphology is reached. This is the distribution of each cell in 3D space so as to conform a specific tooth morphology. The model has a number of developmental parameters, such as the strength of genetic interactions and diffusion rates of signalling molecules. We found that most of the subtle morphological tooth variation (that is 3D and multivariate) in a natural population was reproducible from changes in just one of these developmental parameters. We have also integrated this model with a model of natural selection in populations to address questions in the evolution of morphology (Salazar-Ciudad and Marín-Riera, Nature, 2013).

Mathematical models of development promise a more mechanistic understanding of the GP map. These models, however, are only possible after years of experimental developmental biology in specific organs (i.e. the wing, teeth, etc…). In contrast, quantitative genetics require only measuring the genetic relatedness between individuals in a population and their phenotypic variation (although this should be done in large samples). Mathematical models of development, on other hand, are free from the general assumptions of multinormality, linearity and fixed allelic effects but make also their own specific assumptions. The questions would then be which approach is better and in which situations, whether they can inform one each other.

To study and quantify the limitations of quantitative genetics, we have been testing the accuracy of Lande's equation in predicting phenotypic change in virtual populations. We have developed simulations combining development and evolution. In this algorithm we have a population consisting of a fixed number of diploid individuals. Their genotypes are modelled as a number of loci -the genes- that additively determine the values of each developmental parameter. These developmental parameters are given to the developmental model and the latter gives the resulting morphology, from which trait values are extracted (as in Salazar-Ciudad and Marín-Riera, Nature, 2013) parents are later selected based on their morphology, and mated to create the next generation. (just like in many artificial selection experiments Weber, 1990, 1992; Bolstad et al, 2015). Then we apply Lande's equation at each step and compare its predictions with the real change in traits means.

We found that Lande's equations will often incur in rather large errors in their predictions of trait change, to the extent that a trait will decrease when Lande's predicts it should increase or vice versa. We found these errors to arise from the linear mutinormal assumptions of Lande's equations not matching the complex non-linear GP map arising from tooth development. We were able to identify which aspects of developmental dynamics lead to morphological variation departing the most from Lande's predictions. These errors, however, do not always occur. In many cases Lande's equations work just fine. Most commonly, we find both regimes in a single evolutionary simulation, that is, predictions are accurate for a number of generations, until the local characteristics of the GP map complex become enough for problems to arise. When this happens, Lande's predictions start to fail, either because the expected phenotypic variation is not produced, it is not produced in the expected direction or this direction changes relatively fast between generations. Ultimately, what we found is that the GP map of the tooth model is different for different areas of the parameter space. If the optimal morphology is very different from the average morphology in the population, then the population has to move long distances along the parameter space and, then, the chances to encounter zones with complex GP maps are high. This suggests that given enough time, Lande's equations would start to fail. For some morphologies this will happen all the time.

Being able to identify when, how much, and why the predictions of quantitative genetics may fail is important for a better understanding the GP map. A better understanding of development, together with its congruences and discrepancies with quantitative genetics, is also pivotal in this aim. Our results are, however, limited by the fact that the tooth model, although mechanistic and

based in the state-of-the-art tooth developmental biology, is still a theoretical model. The ultimate test is to compare the tooth model and Lande's equations in respect to real artificial selection experiments. For this study, mammalian teeth are not the optimal system since since mammals have quite long generation times and artificial selection experiments with significant amount of generations become infeasible.

3. Scientific objectives and expected impact

3 A. Scientific objectives

This project aims to do exactly that. We aim to get a better, more predictive, understanding of the GP map by comparing quantitative genetics and developmental biology theoretical models in their capacity to predict how a phenotype changes in actual artificial selection experiments.

The phenotype we chose to study is the wing morphology of the model species *Drosophila melanogaster*. This is because *D. melanogaster* has very short generation times (it is a very popular animal species for artificial selection experiments), it is easy to rear, its genomics and populational genomics are relatively well understood, and its development is well understood as well. The wings of the fly are an ideal phenotype to study because it is easy to automatically measure many morphological traits in them (see Figure 1A, Houle *et al.*, 2003) and, most importantly, because its developmental biology is understood to the extent that a mathematical model of its morphogenesis exists. This model was constructed by my own group (Ray *et al.*, 2015) and is similar to my previous tooth model in the sense that it reproduces an organ morphology – i.e.a spatial distribution of cells that form the shape of a wing in 2D;- from very simple initial conditions: the early pupal wing or the early wing imaginal disc depending on the version of the model being used. The model only specifies these initial conditions and how genes and cells interact in a specific hypothesis network based on what is currently known about wing development. As a result of these initial conditions and the hypothesis the wing morphogenesis arises from the model dynamics (and without further intervention from the researcher). In that sense this model is mechanistic and represents the predictions on wing morphology that can be made from our understanding of its development. The parameters of such model quantitatively specify how strongly specific genes positively or negatively regulate each other expression, the diffusivity of the extracellular molecules, the affinity for their receptors and how strongly specific genes positively or negatively regulate cell growth, cell division, cell adhesion and cell polarity. Changes in these parameters lead to changes in the wing morphology being produced by the model, i.e. to phenotypic trait variation.

Artificial selection experiments (Weber, 1990; Weber, 1992; Bolstad *et al.*, 2015) consist in rearing a population of organisms and, in each generation, selecting some of them as the founders of the next generation. By selecting these founders based on some specific phenotypic traits, the population-mean values of such traits may change over generations. Artificial selection experiments study which phenotypic variation is possible in a system over generations from standing genetic variation. Since one does not know before-hand the exact phenotypes that will arise in such experiments, they can be used to test the predictive capacity of existing models without being biased, or over-fitted, to variation already known to exist. This is why we chose this approach for comparatively studying the predictive capacity of quantitative genetics and developmental biology. The wing morphogenesis model could be tested for its capacity to reproduce the phenotypes of lab mutants or mutation-accumulation lines instead. We are also planning to examine this as a different project that is not specially suitable to compare quantitative genetics and mathematical models of development.

In this study we will 1) perform artificial selection experiments on populations of Drosophila melanogaster, selecting for several traits on the wing morphology; 2) test and compare the predictive power of quantitative genetics and developmental models. In each generation we will estimate the G matrix and use it to estimate the values of these traits in the next generation using Lande's quantitative genetics equations. We will then measure the deviation between these estimations and the observed trait changes. We will do a similar thing for the

wing model. In each generation we will search for the parameters of the model that produce the wings of each selected male and female. We will then recombine the parameters of each pair and run each recombination in the wing model to get each pair's offspring. For each generation we will compare the results of the model with the wings observed in the artificial selection experiments.

Questions and hypotheses we aim to address from that are:

Q1.1 How well does Lande's equation predict the phenotypic change between generations?
Q1.2 When does it fail?
Q1.3 Can we attribute these failures to some specific aspects of development?
Q1.4 Can the wing model help to improve the prediction in these cases? More specifically, can the wing model predict when there would not be response to natural selection and when Lande's equations will fail?

Q2.1 How well does the wing mathematical model predict the phenotypic change between generations?
Q2.2 When does it fail?
Q2.3 Can we learn from those failures? What does need to be changed in the wing model need to improve its predictions?
Q2.4 Can the G matrix be used to improve the wing model predictions in these cases?

Our hypothesis is that Lande's equations will, in general, work rather well. This hypothesis is based in the generally perceived success of quantitative genetics in animal and plant breeding. Most of this research, specially in artificial selection, is focused in selecting variants that increase any of a set of traits (e.g. fat content in milk, milk production, etc…). In here, in contrast, we will be selecting for individuals to have a specific combination of trait values, this is a specific morphology. This is selecting for the vector of trait values to be as close as possible, in euclidean distance, to an arbitrary optimal trait vector phenotype and not just for the sum of trait values, as in the selection indexes used in quantitative genetics, to be as large as possible. In the wing, the experiments closest to ours consider only 2 traits (Weber, 1990). Typically, artificial selection experiments show an initial linear response that gradually de-accelerates and eventually stops, although usually experiments are stopped when the rate of change per generation is very small (Alberch, 1982). This de-acceleration is often interpreted as an exhaustion of the standing genetic variation for the traits studied due to the selection.

An additional reason to expect Lande's to work well are estimations of the G matrix for 20 morphological traits in the fly wing from one of the collaborators in this project (Mezey and Houle, 2005). This study shows that the number of significant independent directions of change in the wing are equal to the number of traits measured. This is interpreted as wing morphology being able to change in any direction, at least for these 20 traits, from the average trait values of the studied population. In other words, wing morphology should be able to change into any morphology we select for as long as this morphology is not very different from the average morphology in the population. This interpretation, however, presupposes that if a population shows variation in a given morphological direction (this is in a given eigenvector, a specific weighted combination of traits) then variation should be possible in this specific direction in the long-term. Breeder's and Lande's equation do a similar thing, they linearly interpolate that evolution in a given morphological direction is possible based on the fact that there is variation in this direction in the current population. This may be quite reasonable hypothesis when not much is known about how phenotypic variation is produced (e.g. development), and for a small number of generations. This assumption of linearity, however, is in plain contrast with what we know about development. The complex and non-linear nature of development may make that after some generations, limits to variation in a direction arise (and, thus, de-accelerations on the response to selection in the

directions estimated from a G matrix estimated in the past). These limitations may preclude linear interpolations from past variation based on Lande's from being predictive (since they predict variation in a phenotypic direction in which there is not, after some threshold, no possible variation). In other words, Lande's prediction can over- or understimate the change in traits' means through the generations because developmental can allow for faster transitions than expected, as well as impose limits to change (developmental constraints). These characteristics of developmental cannot be summarized in the G matrix.

It is generally accepted that the G matrix of a system may change over time (Arnold *et al.*, 2008; Roff *et al.*, 2012; Björklund *et al.*, 2013; Björklund & Gustafsson, 2015; Chebib and Guillaume, 2017), although there is an on-going controversy about how often and how fast. Development is one of the processes that can help resolve how often this happens, in which way G changes and how the limitations or de-accelerations on the responses to natural selection may arise.

Our hypothesis is also, thus, that at some generations of selection for a specific morphology, Lande's equations will start to fail significantly. This is what happens in our tooth theoretical artificial experiments and that is what we expect from developmental systems having non-lineal dynamics (and we expect most developmental systems to be like that (Alberch, 1982; Salazar-Ciudad 2006, Salazar-Ciudad and Jernvall, 2010)). In fact, the existence and necessity of these non-linearities in development and their effect in evolution is one of the foundational topics of developmental evolutionary biology (Alberch, 1982; Salazar-Ciudad 2006). Understanding why this should be the case is not specially easy so I will do it based on an simple example from wing development:

One important process in late wing morphogenesis is hinge contraction (Ray et al., 2015, Etournay *et al.*, 2015). In hinge contraction the most proximal part of the wing contracts. The most distal wing margin is attached to the cuticle and does not move. As a results of this contraction and margin attachment, all regions of the wing between this margin and the hinge get displaced in the proximal and medial directions (this is towards a central line going from the distal to the proximal part of the wing, see Figure 2d). The closer a wing region is to the hinge the stronger and more medial is this displacement. As a result the wing transforms its shape from that of an rounded square to that of a club (see Figure 2d). From that knowledge one can expect that the more the hinge contracts, the narrower the region of attachment of the wing with the rest of the body. This suggests that adult wings with narrow hinges should also have regions between the hinge and the distal margin that are, compared to wings with wider hinges, displaced medially and proximally. This suggests also that the contrary should not be possible or occur only rarely, wings with wider hinges should not have regions between the hinge and the distal margin that are, compared to wings with wider hinges, displaced medially and proximally. Thus, variation that increases the distance between wing landmarks 11 and 12 in figure 1 while displacing landmarks 1 and 4 in the medial and proximal direction should not occur.

From this perspective hinge contraction may seem to preclude variation in one specific direction but this would be slightly misleading. From a theoretical developmental perspective one has to propose a mechanism explaining why some variation exists, not why it does not exist. Phenotypes and their variation exist because genes and cells interact during development not in spite of that. This is without development there would be no phenotypic variation. For example, it is because of hinge contraction that a specific morphology, club shaped wings, is possible and it is because of this contraction that variation in the hinge width (and associated changes) is possible. Without hinge contraction such variation would not exist.

This is different in quantitative genetics. In quantitative genetics, what happens between the genotype and the phenotype (e.g. development) is treated as a black box and then, in principle, all aspects of the phenotype are equally likely to exhibit variation without having to propose a mechanisms of why is that so. The developmental perspective is based on specific mechanistic hypotheses and as such it may be wrong or incomplete: processes other than hinge contraction, that are currently unknown to us, may exist that produce variation in the directions we claim that are not possible. Lande's equation does not make any specific hypotheses at those levels but makes some

general assumptions of linearity, multinormality and fixed allelic effects. These assumptions do not preclude to estimate G and apply Lande's equations but could greatly decrease its accuracy. Lets explain how can this happen based on the hinge contraction example.

The hinge contraction example suggest that variation should be possible in one direction in trait space but not in others. This should be captured from the estimation of the G matrix (there would be simply some eigenvalue that would not be significantly different from zero) and, thus, be predictable from Lande's equations. Hinge contraction is a major event that has effects over the whole wing morphology and, thus, variation in these not-allowed directions should be limited. By limited we mean that the population can change in this direction for a short while until it is not any more possible: a limit is reached and imposed by hinge contraction. The existence directions of change that are possible for a range of values and then stop after some threshold value can lead to large errors in Lande's estimations. This is because what Lande's equations do in practice is to regress each combination of traits between parents and offspring (or other more sophisticated linear interpolations between individuals of known genetic relatedness) and linearly interpolate that there is going to be a response in the direction of a trait combination if their regression is significantly different from zero (and if negative regressions between other traits do not preclude that). These interpolations would pass over the limits imposed by hinge contraction and predict a continuation of a response to selection in a direction in which this is not going to happen because of the limit. This will induce a large systematic error that will remain even if G is estimated in each generation (since both the limit and the significative regression will remain). We have found in the tooth model other more complex but similar situations where Lande's equations fail because of their linear approach. From this perspective, the quantitative genetics result (Mezey and Houle, 2005) that that there is variation in all possible directions in the wing trait space, would only hold as long as the population does not start to move along a specific direction in the trait space. If the population moves it would likely encounter the above limits.

Expected research results and their anticipated scientific impact, potential for scientific breakthroughs and for the renewal of science and research:

Our expected results are the confirmation of the above hypotheses. If our hypothesis turn out to be wrong, this would also be a very interesting significant result. It would tell us, that as far as we can see, the approaches of quantitative genetics work just fine for relatively complex morphologies with rather complex and mechanistically understood GP map. Another expected result is that the wing model can predict when Lande's equations will fail. They will predict, for example, that there is no actual variation in a morphological direction after some threshold trait value, even if there is variation in this direction in the population before this threshold value. Since we understand, in the wing model, why variation happens, we will be able to pinpoint which aspects of the GP map are lead to errors in Lande's estimations and which developmental dynamics produce those.

Developmental evolutionary biology and quantitative genetics represent the two main approaches to understand morphological variation, evolution and, more in general, the GP map. Being able to precisely quantify where one works and one fails, and vice versa, would be major breakthrough in our understanding the GP map. This is not only to delimit where one approach is best and where it is not, but also, and most importantly, studying why is this the case, from our understanding of development, should give us significant deep insights about how is the GP map and why. For a proper comparison of both approaches we needed to apply them both to the same model system but the insights acquired about the GP map should be of general validity for many other systems and, in fact, acquiring those insights in any other system would be much more difficult given the special suitability of the fly wing for these approaches.

Exploring whether the wing model can reproduce the morphological variation arising in the artificial selection experiments would also be very informative about wing development. If the model succeeds, it means that the model summarizes well wing development. If it does not, it means that the current understanding of wing development that the model encapsulates is not

enough. This is also interesting and informative since we will need to explore what does need to be changed in the model to produce the observed variation and, in the process, we will learn a lot about wing development. Notice that the variation in the artificial experiments is a special one, is not as extreme as the one in lab mutants but not as subtle and smooth as the one in natural populations, that specially for fly wings is specially subtle. Thus, we will get a better wing model and a better understanding of wing development. All these insights will be discussed and contrasted with our experimental collaborators.

3 B. Effects and impact beyond academia

The 20th century can be seen as the century of genetics. We have learned that phenotypes and most of its variation have, ultimately, a specific genetic basis. We know what is at the bottom, the genome, and what is at the top, the phenome, but we do not understand well enough the processes in between to predict and explain which changes at the gene level lead to which specific changes at the phenotypic level (and why to those changes and not to others). Improving our understanding of this GP map is an important step that would make the molecular biology and -omics data that has been accumulated over the years more easy to interpret to approach applied problems such as complex diseases, animal and plant breeding and genetic manipulation, organoid and regenerative medicine. It is my perception, and that of many others, that the early 21th century biology would largely be about understanding this GP map. This involves an integration between genomics, molecular biology, cell biology, biophysics, evolutionary biology and developmental biology. Theoretical and computational biology will play a prominent role in this integration.

This project is directly addressing how is the GP map and why, from a mechanistic developmental perspective. The project also aims to test the predictive capacity of the two main modelling approaches to the GP map. Identifying in which situations, and how much, these two approaches fail and why, should be of general applied interest. This is because one of these approaches, quantitative genetics, is massively used, both in purely scientific questions and in applied topics such as animal and plant breeding, GP maps in cancer research and the genetic and biological bases of complex diseases. The other approach is likely to become more and more popular over time as developmental systems become better studied and this leads to large amounts of complex data that can benefit from mechanistic mathematical models (again with future implications in breeding, cancer and complex diseases). This project could, thus, easily become a reference for the limits, and compatibilities and incompatibilities between these approaches.

3 C. Publication plan

Identifying when quantitative genetics fail and why based on a specific understanding of development is likely to be of general interest and, thus, I expect these results of our project to be publishable in an article in a general science high impact journal (Nature, Science, PNAS, etc…) as I have done in the past. Another article would be the insights learned on development based on comparing the models with the data arising from the artificial selection experiments. This may be publishable also in a general science high impact journal or in a more specialized high impact journal in developmental biology and evolution (Developmental Cell, Development, Evolution and Development, etc…). The general outcomes of the research should be of interest to anybody interested in the phenotype, how it is produced and how is the GP map. That is developmental and evolutionary biologists, quantitative and populational geneticists, animal and plant breeders, researchers in complex diseases and systems biologists. All articles will be published in green open access and we will be ready to any interviews or media requests we may get from them, as we have done in the past. The results of this project will also be discussed in scientific meetings by me and members of my team. These are the ICQG2020 in quantitative genetics, the eseb2019 in evolutionary biology and the euro-evo-devo2020 in evo-devo.

4. Research methods and material, support from research environment

Artificial selection experiments: The initial population will be made by intercrossing 30 inbred lines obtained from the Drosophila Genetic Reference Panel (Mackay et al., 2012). This population will be allowed to mate freely for one generation before starting the experiment. Then, we will have two artificial selection experiments with two replicate populations each. Typically for these experiments we can run 25 generations per year. In each generation and population, 100 virgin females and 100 males will be chosen haphazardly and imaged. From those we will select 25 flies of each sex to produce the next generation. Each pair and offspring will be reared in an different vial. This will allow us to know which are the parent of each of the 200 imaged flies in the next generation (these 200 flies are taken at random between all the vials of a population in a given generation). This information will allow us to estimate the G matrix in each generation through parent offspring regression for each measured trait. The S vector, also required for Lande's equation (see Eq 2), can be analytically derived since it is us who chose the selection criterion based on an optimal morphology. The P matrix is simply the total phenotypic covariance and can be directly estimated from the data. By using this estimation of G, P and S we will predict how the average phenotypes should be in the next generation (note that although the actual population will be much larger, all genetic parameters are estimated on these 200 individuals). These estimations will be contrasted with the trait measurements made in the 200 individuals haphazardly chosen for imaging in the next generation. Selection will be based on the 20 landmarks will be measured in each individual left wing by using a simple experimental set up that allows to image wings automatically, the WingMachine and accompanying software (Houle *et al.*, 2003). Setting up the wing machine is relatively easy an cheap (https://www.youtube.com/watch?v=Nq-OIGKzdLk). Essentially a picture will be taken of each wing and the landmark coordinates (the actual morphological traits), will be extracted automatically. This takes around 1-3 minutes per fly. This set up was designed by the main external collaborator in this project (David Houle). The 25 flies of each sex select in each generation will be the ones with the smallest Euclidian distance to an optimal combination of trait values (see later). As explained above, previous wing studies in quantitative genetics (Mezey and Houle, 2005) suggest that there should be a response to selection in any direction while studies in development suggest this should not be the case or that, at least, Lande's predictions will fail. To maximize the chances of this occurring we will chose an optimum that could be expected to difficult to arise from what we know about wing development (we will start with the ones described in the objectives).

The wing model: The wing model is already published (Ray *et al.*, 2015). It has been used to predict how the phenotypes of some lab mutants develop into having a specific aberrant phenotype. Mathematically the wing model is a vertex model (Farhadifar et al., 2007) in which we have added cell-cell signalling, signal diffusion, regulated cell proliferation, regulated cell polarization and specific wing initial and boundary conditions (see Figure 2). The model predicts adult wing morphology from these initial conditions and from a specific hypothesis of how a specific gene network regulates gene expression, cell-cell signalling, cell proliferation, cell adhesion and cell polarization (Ray *et al.*, 2015). The initial conditions of the published model are those of the early pupal wing. We want to improve the model's descriptive capacity by making it start from the early imaginal disc. Fortunately we have a yet unpublished, but under preparation, model of the development of the imaginal disc (see Figure 2a,b and f for results) and another one, submitted, for the disc unfolding that occurs between the disc and the pupal wing (Fristrom & Fristrom, 1993). All these models put together simulate the whole wing morphogenesis and its variation.

The model will be used for the following: 1) In a low stringency test, we would explore if by freely tuning the model parameters we can reproduce the morphological variation observed over generations. This is the average trait values for each generation, and ideally also, the actual trait values for each of the 200 flies imaged per generation. 2) In a more stringent test we will, in each generation, automatically search for parameter values in the model that will produce the wings of each male and female selected. We will then recombine the parameters of such males and females to produce the parameters of each offspring individual. This is, each offspring parameter would be either that of the father, that of the mother or the average of the two. We will run all these possible

combinations in the model and compare the produced wing morphologies with the trait averages and variation of the wings in each generation. An exact match is unlikely to be found even with a perfect model. This is because each developmental parameter may, in fact, depend on several genes and those may segregate in complex ways in the population, while we are essentially assuming a single locus per parameter with either total dominance or co-dominance. If the model is good enough, however, the wings in the actual offspring should be a subset of the wing morphologies produced in the model from a given pair.

Critical points for success, alternative implementation strategies: Each experiment line will be stopped when we encounter significant errors in the estimations provided by Lande's equations for at least 10 generations in a row, because that is what we want to find, or when there is no response to selection for 10 generations in a row, also interesting in itself. Both results would be considered a success. After any of these events we will re-inititate that line with a new artificial selection experiment with a different selection criterion. If this does not occur, thing that we do not expect, this is still and interesting result, as explained in the objectives. It is also possible that we fail to reproduce the wing morphologies with the developmental model. The model, however, should not be seen as a result but as a process. The model is built from what is currently understood about development of the wing. If the model does not work, we will explore which aspects of it need to be changed for it to accurately reproduce the observed morphologies, while keeping it compatible with experimental evidence coming from developmental biology. The artificial selection experiments will be pivotal to clarify what morphological variation escapes the model, if any, and will allow to improve the wing model and our understanding of wing development and its GP map.

Tangible support from research environments: The Salazar-Ciduad group belongs to the Center of Excellence in Experimental and Developmental Biology of the Developmental Biology program at the Institute of Biotechnology, University of Helsinki. The Institute provides lab and office spaces. The basic equipment has been set up for *Drosophila* genetics (six benches with stereomicroscopes and CO_2 gas, 25^{o}C and 18^{o}C fly incubators). Light Microscopy Unit at the Institute of Biotechnology provides four confocal microscopies (Leica SP5 and Zeiss LSM700). The *Drosophila* stocks in this proposal are available from a stock center (Bloomington *Drosophila* _Stock Center, http://flystocks.bio.indiana.edu/; or obtained from the other labs. As a researcher at the BI I also have access to the CSC computing center and sufficient computing power from computers in my office's group (around 100 cpus in total).

5. Ethical issues

This project involves animals, an insect, but no actual direct experimental manipulation of their genotype or phenotype. So there are no ethical issues.

6. Implementation: schedule, budget, distribution of work

The artificial selection experiments will be started as soon as possible, once, the Wingmachine set up is established. This would allow to reach up tp 70 generations for 4 lines (although the actual number of experiments may vary depending on what is explained in section 4). Estimating the response to selection based on the estimated G matrix would be done in each generation while the experiments are underway since, as explained above, we will stop each experimental line when Lande's equations fail for a number of generations. The analysis with the wing model can be done at any moment during the three years since all morphological data will be stored in our computer cluster and the continuation of the experiments does not depend on it. Computational analysis will be carried out mostly during the last year of the project, after most of the data has been collected.

Most of the costs of the project are on personnel because most of the required equipment, computers and basic Drosophila equipment, are already provided by the Biotechnology Institute. The actual artificial selection experiments are technically not very demanding and somehow too tedious, on its experimental side, for a PhD student. These consist mostly in imaging wings and

putting selected flies in new growth media. The plan is that the bulk of these experiments will be implemented by two undergraduate students under my supervision and the assistance of a graduate student (we only ask two years of salary for him because he has already started with me a theoretical topic closely related to this project; the merely theoretical study based on the tooth model). He, Lisandro Milocco (a biologist and mathematician), will be the person working with the quantitative genetics estimations of the project. The actual simulations with the wing model will be done by the postdoc hired for a year, towards the end of the project when a significant proportion of the data will be available (note, that as explained, we need to be doing the quantitative genetics in each generation but the modelling is not needed to decide on when to stop or not ongoing experiments). Other costs include 2000 euros per year for fly's food (we are dealing with a relatively small number of flies; 3000-6000 flies at any given time) and small expenses to set the WingMachine (1000 euros), travel expenses to visit the collaborator David Houle and participate in three congresses 4000 euros per year) and 10000 euros for open access publication. This project includes the salary for the PI for the last 4,5 months simply because the PI salary, from the present forecast, would not necessarily be covered by other sources. The PI task will be analysing the results of the project (compare the two approaches undertaken), interpret the results and write the articles arising from the project (and supervise the whole group). Before these 4,5 months the salary will be covered from the Center of Excellence in Experimental and Computational Developmental Biology (2018-2019) and funding associated with it arising from Helsinki University for the whole of 2020 and the first 3,5 months of 2021.

7. Research team and collaborative partners

The main collaborators in this project are David Houle, from Florida State University, USA and Osamu Shimmi, from the same Center of Excellence in Experimental and Computational Developmental Biology in which I am. David Houle is renowned quantitative and evolutionary geneticist with which I have already collaborated on the fly wing (Ray et al., 2015; Matamoro et al., 2015). Houle has a long experience in artificial selection in the wing and he developed the WingMachine and a big part of the accompanying software to automate the measurement of wing morphological traits. He will be sharing with his knowledge and opinions in technical (fly artificial experiments) and theoretical (e.g. G matrix estimation) aspects of the project and he will likely be a co-author in the resulting articles (not the last or first author).
Osammu Shimmi is a fly wing developmental biology. He has been working on individual vein morphogenesis in the wing. He will assist us in technical and theoretical aspects of the project and he may be a co-author in some of the resulting articles (not the last or first author).

8. Research careers and researcher training

Being awarded this grant will allow me to enlarge my research from evolution and development approaches based on the mathematical modelling of development to quantitative genetics also. This will provide me and my cv a broader and more diverse bases to tackle the problem of phenotypic evolution and the GP map. This project includes the supervision of two master students, one PhD student and a postdoc. These will all supervised by me and will lead to the completion of two master degrees and one PhD thesis. The most experimental part of the project is simple and involves dealing with basic animal manipulation and automatic imaging. The students will also learn the theoretical underpinnings of quantitative genetics, developmental biology and its modelling and their comparisons in the study of the GP map. Recruitment of the personnel in the project will be done according to the Helsinki guidelines to promote equality

9. Mobility plan for the funding period

This project includes three major visits to or from our collaborator David Houle in Florida State University, USA. The first one will be a visit by me and the persons involved in the experiments to David Houle to learn an optimal way to use of the WingMachine and the accompanying software of

wing scanning and morphometrics. After the first year or after the first example of failure of Lande's equations in predicting the response to selection in the artificial selection experiments, there would be a visit to David Houle or from him to Finland (depending on the respective schedules), to discuss the ongoing results. This visit will be repeated towards the end of the project to coordinate and write the papers (in which I will be the senior author). Each visit will last between three weeks and a month.

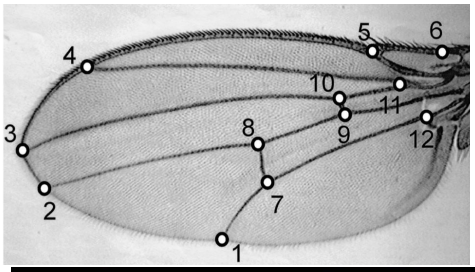

Figure 1: Landmarks that will be measured in this project. Each landmark X and Y position are a phenotypic trait.

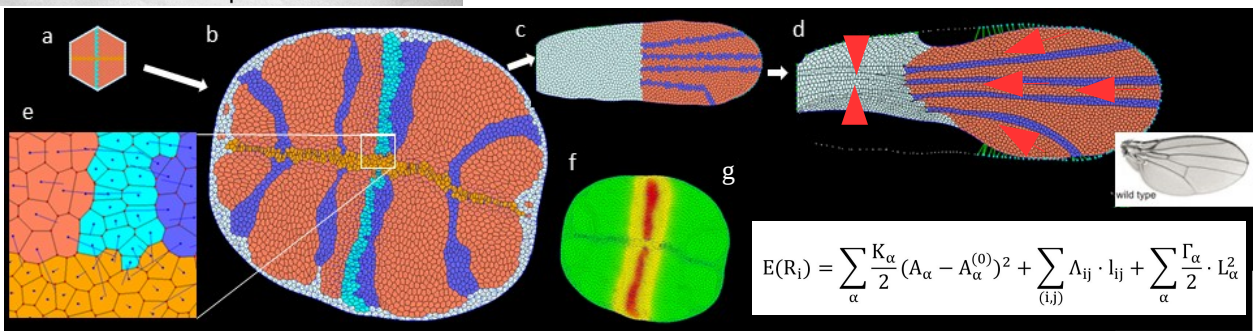

Figure 2: Schema of the wing development mathematical model: (a,b,e,f) Imaginal disc development. **(a)** Initial conditions of the model, a simple flat epithelium representing the first instar larval wing imaginal disc. The light line in the middle is the anterio-posterior boundary where cells express the extracellular signal dpp. The orange line are the cells in the dorso-ventral boundary, they express the extracellular signal wg. This latter signal is expressed also in the margin of the disc, cells in white. Cells get polarized in the direction of the concentration gradients of both signals, growth in their direction of polarization and divide orthogonally to it. **(b)** Late imaginal disc as predicted from the mathematical model of development, dark blue cells are pro-vein cells, cells that will give rise to the veins of the wing. **(c)** Initial condition for the wing model proper, the initial condition represents the shape of the pupal wing just after the unfolding of the wing imaginal disc. We are currently finnishing a model to go from **b**, the predictions of the disc part of the model, to **c** (so that the spacing between veins in **c** would be the result of the spacing between veins in **b** and the perimeter and anterio-posterior asymmetry of the wing margin in **c** will be the result of the length and anterio-posterior margin in **b** (light blue line of cells in the middle of the disc blade). **(d)** Wild-type wing obtained so far. Red arrows indicate the lines of stretch resulting from the hinge contraction. The hinge is shown in white, veins in blue **(e)** More detailed depiction of the structure of the model, each polygon is a cell, a line from each cell center indicates the direction of cell polarization. Calculations are made only at the points of intersection between three cells (the vertices). These points move in 2D as determined by the equation in **g. (f)** The same than in b but where each cell is painted with a colour that is proportional to concentration of the signal dpp it receives. This is pretty much the dpp concentration gradient. **(g)** Basic equations of the vertex model. Nodes are moved at random but movements that decrease the energy $E(R_i)$ of a node are more likely to occur. This energy depends on how much the cells to which a vertex belongs are at their ideal size (first term in the equation), on on how much contact they have with other cells (second and third terms). In addition, cells growth in the direction of the signal gradients arising from signal diffusion in the space of the model. After a threshold size cells divide. The surface tension, or adhesion, between cells depends on the kind of cell (vein, margin, hinge, etc...) and hinge cells contract by decreasing their ideal size.

10. Bibliography (2 pages)

-Alberch P. 1982. Developmental constrains in evolutionary processes. In: Bonner JT, editor. Evolution and development. Dahlem Konferenzen, Heidelberg: Springer. P 313–32.

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

-Arnold SJ, Bürger R, Hohenlohe PA, Ajie BC, Jones AG.Understanding the evolution and stability of the G-matrix. Evolution. 2008 Oct;62(10):2451-61.

-Arthur W. 2004. The effect of development on the direction of evolution: toward a twenty-first century consensus. Evol Dev 6:282–288.

-Bolstad et al., 2015. Complex constraints on allometry revealed by artificial selection on the wing of Drosophila melanogaster. PNAS. 11, 13284–13289.

-Björklund M, Gustafsson L. The stability of the G-matrix: The role of spatial heterogeneity. Evolution. 2015 Jul;69(7):1953-8

-Björklund M, Husby A, Gustafsson L. Rapid and unpredictable changes of the G-matrix in a natural bird population over 25 years. J Evol Biol. 2013 Jan;26(1):1-13.

-Chebib J, Guillaume F. What affects the predictability of evolutionary constraints using a G-matrix? The relative effects of modular pleiotropy and mutational correlation. Evolution. 2017. 29.

-Etournay R, Popović M, Merkel M, Nandi A, Blasse C, Aigouy B, Brandl H, Myers G, Salbreux G, Jülicher F, Eaton S. Interplay of cell dynamics and epithelial tension during morphogenesis of the Drosophila pupal wing. Elife. 2015 Jun 23;4:e07090.

-Falconer, D. S., and T. F. C. Mackay. 1996. Quantitative genetics. Longman Group, Essex, U.K.

-Graner, F. and Glazier, J. 1992. Simulation of biological cell sorting using a two-dimensional extended potts model. Physical Review Letters, 69, 2013–2016.

-Farhadifar, R., Röper, J.C., Aigouy, B., Eaton, S., and Jülicher, F. 2007. The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing. Curr. Biol. 17, 2095–2104.

-Forgacs, G. and Newman, S. 2005. Biological physics of the developing embryo. Cambridge University Press.

-Fristrom D, Fristrom J. 1993. The metamorphic development of the adult epidermis. In: Bate M, Martinez-Arias AM, editors. Cold Spring Harbour, NY: Cold Spring Harbour Press. P 843–897.

-Hansen TF, Alvarez-Castro JM, Carter AJ, Hermisson J, Wagner GP. Evolution of genetic architecture under directional selection. Evolution. 2006 Aug;60(8):1523-36.

-Honda, H., Tanemura, M., and Nagai, T. 2004. A three-dimensional vertex dynamics cell model of space-filling polyhedra. Journal of Theoretical Biology, 226, 439–453.

-Houle D, Mezey J, Galpern P, Carter A 2003. Automated measurement of Drosophila wings. BMC Evol Biol 3(1):25.

-Johnson T, Barton N. Theoretical models of selection and mutation on quantitative traits. Philos Trans R Soc Lond B Biol Sci. 2005 Jul 29;360(1459):1411-25

-Mackay T. F. C., Richards S., Stone E. A., Barbadilla A., Ayroles J. F., et al., 2012 The Drosophila melanogaster Genetic Reference Panel. Nature 482: 173–178.

-Marin-Riera, M., Brun-Usan, M, Zimm, Välikangas, T. and Salazar-Ciudad, I. 2016. Computational modeling of development by epithelia, mesenchyme and their interactions: a unified model. Bioinformatics, 1-7, 212-217

Matamoro, A, Salazar-Ciudad, I, Houle, D. Making quantitative morphological variation from basic developmental processes: where are we? The case of the Drosophila wing. Developmental Dynamics, 244; 9:1058-1073

-Maynard Smith J, Burian R, Kauffman SA, Alberch P, Campbell J, Goodwin BC, Lande R, Raup D, Wolpert L. 1985. Developmental constrains and evolution. Quart Rev Biol 60:265–287.

-Mezey, J. G., and D. Houle. 2005. The dimensionality of genetic variation for wing shape in Drosophila melanogaster. Evolution 59: 1027–1038.

-Newman, T. 2005. Modelling multi-cellular systems using sub-cellulars elements. Mathematical Biosciences and Engineering, 2, 611–622.

-Newman SA, Comper WD. 1990. 'Generic' physical mechanisms of morphogenesis and pattern formation. Development 110:1–18.
-Lande, R. 1979. Quantitative genetic analysis of multivariate evolution applied to brain:body allometry. *Evolution* **33,** 402-416.
-Lande, R. & S. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* **37,** 1210-1226.
-Lewontin, Richard C. 1974. The genetic basis of evolutionary change ([4th printing.] ed.). New York: Columbia University Press.
-Oster GF, Alberch P. 1981. Evolution and bifurcation of developmental programs. Evolution 36:444–59.
-Pigliucci, M. 2006. Genetic variance-covariance matrices: A critique of the evolutionary quantitative genetics research program. Biol. & Philos. 21: 1-23
-Pigliucci M. Finding the way in phenotypic space: the origin and maintenance of constraints on organismal form. Ann Bot. 2007 Sep;100(3):433-8.
-Raff RA. 1996. The shape of life: genes, development, and the evolution of animal form. Chicago: University of Chicago Press.
-Ray, R.P., Matamoro-Vidal, A., Ribero, P.S., Tapon, N., Houle, D., Salazar-Ciudad*, I & Thompson*, B.J.. 2015. Genetic control of global tissue mechanics determines appendage shape in Drosophila. Developmental Cell. 34; 3:310-322. *Co- senior authors and corresponding authors
-Rice SH. 2002. A general population genetic theory for the evolution of developmental interactions. Proc Natl Acad Sci U S A. 2002 Nov 26;99(24):15518-23.
-Roff DA. A centennial celebration for quantitative genetics. Evolution. 2007 May;61(5):1017-32.
-Roff DA, Prokkola JM, Krams I, Rantala MJ. There is more than one way to skin a G matrix. J Evol Biol. 2012 Jun;25(6):1113-26.
-Salazar-Ciudad I. 2006a. Developmental constraints vs. variational properties: how pattern formation can help to understand evolution and development. J Exp Zool B Mol Dev Evol 306:107–25.
-Salazar-Ciudad I., 2010 Morphological evolution and embryonic developmental diversity in metazoa. Development 137: 531–9.
-Salazar-Ciudad I.* and Jernvall J. Graduality and innovation in the evolution of complex phenotypes: insights from development. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution. vol 304B No 6: 619-631
-Weber, K.E. 1990. Selection on wing allometry in Drosophila melanogaster. Genetics 126:975–989.
-Weber, K.E. 1992. How small are the smallest selectable domains of form? Genetics 130:345–353.