Environmental Robustness in Populations Adapted to Constant Environments and Heterogeneous Environments

by

Ao Wang

A thesis submitted in conformity with the requirements for the degree of Master of Science

Graduate Department of Ecology and Evolutionary Biology
University of Toronto

© Copyright by Ao Wang 2017
Environmental Robustness in Populations Adapted to Constant Environments and Heterogeneous Environments

Ao Wang

Master of Science
Graduate Department of Ecology and Evolutionary Biology
University of Toronto
2017

Abstract

Environmental robustness is the reduction of phenotypic variation under environmental perturbation and is essential for populations to persist in stressful conditions, but few empirical studies have examined the conditions under which robustness evolve. Here we consider two hypotheses, robustness evolves via environmental heterogeneity or as a by-product of adaption to specific constant environments. We assessed environmental robustness in juvenile viability of 20 *D. melanogaster* populations evolving under constant or heterogeneous environments. Contrary to prediction, in the original selective environments, populations from heterogeneous regimes were not the most robust ones. In the novel environments, populations from the constant salt regime were more robust than populations from the constant cadmium regime. Populations from heterogeneous regimes had intermediate level of robustness. Our results suggest that environmental robustness is an indirect product of adaptation to specific environments and environmental heterogeneity plays a relative less important role in inducing robustness than the details of adaptation.
Acknowledgments

This thesis would not have been possible without the contribution from many people. First, I would like to thank my supervisor Aneil Agrawal for giving me this opportunity to start my journey in the field of science. No matter what direction this path takes, the beginning will always remain memorable. Aneil provided me with necessary guidance throughout my MSc program. Not only did he help greatly with my academic writing and presentation, which I do struggle a lot with, he also demonstrated to me that science is a rigorous way of examining the phenomena of nature and should be carried out in the most careful and responsible manner. Next, I want to thank my supervisor committee members Asher Cutter and Nicole Mideo, who have provided me with very useful comments and feedbacks along the course of this project. Especially I am very grateful for the discussion with Nicole about my potential choices after completing this program.

Through stimulating discussions with current and past members of the Agrawal lab, Amardeep Singh, Yuheng Huang, Eddie Ho and Li Yun, I have improved my conceptual grasp on this project and become more efficient at presenting my ideas. They have also given me many practical advices on various graduate-life related issues. I am eternally grateful for the labour contribution from the lab manager and volunteers, Malak Bayoumi, Diane Andreas, Amy Yeung and many others. Without their help, I would certainly not be able to obtain these data within the timeline of my program.

During my first year of graduate school I had the opportunity to meet many intelligent and kind people from EEB who I share the same passion with. My appreciation goes to those who have helped improving my grad-experience in many ways: Rebecca
Schalkowski, Megan Greischar, David Punzalan, Stephanie Penk, Rebecca Batstone, Abby Daigle, Vanessa Luzuriaga and Tessa Brinklow.

Last but not least, I am extremely grateful for the moral support from my family, friends from undergrad, and my cat Nugget. They are always there for me during time of stress, and it is their presence that I can’t do without.

This research was supported by Natural Sciences and Engineering Research Council (NSERC) Discovery Grants to Aneil F. Agrawal and by the Department of Ecology and Evolution at the University of Toronto.
Table of Contents

Abstract .................................................................................................................................ii

Acknowledgments ................................................................................................................ i

Table of Contents .................................................................................................................. v

List of Tables ........................................................................................................................ vi

List of Figures ......................................................................................................................... vii

Introduction ............................................................................................................................ 1

Methods .................................................................................................................................. 8

Selective History of Experimental Populations ................................................................. 9

Test Environments ................................................................................................................ 9

Viability Assay ....................................................................................................................... 9

Data Analysis .......................................................................................................................... 10

Results ................................................................................................................................... 14

Viability in Original Environments ................................................................................... 14

Viability in Novel Environments ......................................................................................... 14

Discussion ............................................................................................................................. 16

Original Environments ........................................................................................................ 16

Novel Environments ............................................................................................................. 17

Tables and Figures ............................................................................................................... 24

References ............................................................................................................................. 34
List of Tables

Table 1 ........................................................................................................24
Table 2 ........................................................................................................25
Table 3 ........................................................................................................26
List of Figures

Figure 1a.........................................................................................................................27
Figure 1b..........................................................................................................................28
Figure 1c..........................................................................................................................29
Figure 2...............................................................................................................................30
Figure 3...............................................................................................................................31
Figure 4...............................................................................................................................32
Figure 5...............................................................................................................................33
Introduction

Genetic variation is universally known as the fuel of adaptive evolution and the rate of change in the trait value is proportional to the additive genetic variance of that trait (Price 1972). Which factors shape the pattern of genetic variation among and within populations is one of the central questions in evolutionary biology (Leffler et al. 2012). In nature, environments change over space and time and an individual’s phenotype and fitness is linked to the environment it experiences. Therefore, environmental heterogeneity is one of the most important processes maintaining genetic variation in fitness (Dempster 1955; Hedrick 1986). This can occur if alternative alleles at ecologically relevant sites are selected in different environments. Consequently, standing genetic variation maintained by environmental heterogeneity could potentially assist adaptation to novel environments if variation maintained by heterogeneity is relevant to the novel selection pressure. One study has found that *D. melanogaster* populations maintained in heterogeneous environments have larger adaptive response to novel selection in female fecundity than populations maintained in homogeneous environments (Huang et al. 2016). Although this suggests that environmental heterogeneity may facilitate adaptation in the long-term, little is known about its short-term impact on population’s ability to persist in stressful environments. This is important because adaptation can be a relatively slow process and individuals may need to cope with abrupt changes in either internal or external conditions within their life time.

Phenotypic plasticity and environmental robustness are two important mechanisms that help organisms respond to immediate environmental perturbations. Phenotypic plasticity describes a genotype’s ability to produce different phenotypes due to environmental variation,
while robustness is the reduction of phenotypic variation under perturbation (the latter was initially coined ‘canalization’ by Waddington in 1940s and I will use both terms interchangeably). Whether a trait evolves plasticity or robustness depends on the optimal phenotypic values across environments. If the optimal values are the same or similar in different environments, then robustness is favoured; otherwise plasticity is favoured. Although these two phenomena may seem opposite to each other, robustness at one trait level often results in plasticity at another level and vice-versa. They both fundamentally describe the same process: the dependence of the phenotype on the environment (De Visser et al. 2003; Flatt 2005). For example, the maintenance of constant body temperatures in endotherms is likely coupled with plasticity in basal metabolic rate and dilation of blood vessels.

In the case of stressful conditions, canalization for the most basic physiological and developmental processes to properly function is necessary for populations to survive in the new environment (Waddington 1953; Scharloo 1991; Ghalambor et al. 2007). Therefore, environmental buffering of homeostasis and development is an essential first step for populations to cope with external stress, followed by a second stage of adaptive evolution responding to the selection pressure. Besides environmental robustness that buffers against non-heritable perturbations there is also genetic robustness that buffers against heritable perturbations such as mutation. However, natural mutation rate is very low for most species except for bacteria and viruses and most mutations have no detectable effect on protein structure or phenotype (Guo 2004), whereas environmental perturbations typically have greater impacts on fitness than genetic perturbation. Therefore, there is likely a stronger selection for environmental robustness than genetic robustness. This is supported by the
observation that environmental variance is almost always greater than genetic variance of a trait, thus producing the benefit of canalization (Meiklejohn and Hartl 2002; Hallsson and Bjorklund 2012a). Importantly, robustness can occur at different levels within biological organization, ranging from gene expression and protein production at the micro level to trait development and even individual fitness at the macro level. Consequently, different buffering mechanisms are adopted at different levels such as gene redundancy or behavioral change (De Visser et al. 2003).

The concept of “canalization” was first established by Waddington and Schmalhausen in the 1940s to describe the constancy of wild types. However, both theoretical and empirical research on this topic declined following their initial work due to the lack of a clear definition and thorough understandings of the molecular mechanisms underlying canalization. Recently, interest in canalization has surged among evolutionary biologists, spurred by the developments in molecular genetics and theoretical biology (Flatt 2005). However, the mechanisms and evolutionary causes and consequences of canalization remain elusive (Siegal and Bergman 2002). One major question centers on the evolutionary forces and conditions that favor the evolution of canalization. Most classic studies including those by Waddington himself, suggest that a trait will evolve robustness under long-term stabilizing selection. Their argument is that buffering of developmental systems evolves as a result of stabilizing selection on phenotype towards an intermediate optimum and any mechanism that reduces deviation from that optimum will be favored by stabilizing selection (Waddington 1942, 1957; Rendel 1967; Layzer 1980; A. Wagner 1996; G.P. Wagner et al. 1997; Gibson and Wagner 2000). Others show that certain types of fluctuating selection (e.g., environmental heterogeneity) favor the evolution of robustness that reduce phenotypic
expression of maladapted genotypes in mismatched environment (Kawecki 2000).

Alternatively, canalization could exist in complex genetic networks without being subjected to selection as an intrinsic property of such networks (Nijhout 2002; Siegal and Bergman 2002). Although most of those efforts focus on the evolution of genetic robustness, the same challenge and arguments are also relevant to environmental robustness as buffering against genetic perturbation may be a by-product of buffering against environmental perturbation (G.P. Wagner et al. 1997; Gibson and Wagner 2000; Meiklejohn and Hartl 2002). Evidence comes from the role of heat shock protein Hsp90 in buffering both mutational perturbation and environmental perturbation, and the demonstration that selection for environmental insensitivity of RNA secondary structure also confers mutational robustness (Rutherford and Lindquist 1998; Ancel and Fontana 2000).

To help address the general question of how different types of environments shape the evolution of environmental robustness, we assess the degree of robustness among D. melanogaster populations evolving in constant environments (and likely experiencing stabilizing selection) and populations evolving in heterogeneous environments (and likely experiencing fluctuating selection) for the last ~190 generations. The experimental fly populations studied here come from four selection regimes: a cadmium-enriched environment; a salt-enriched environment; a temporally heterogeneous environment with respect to cadmium and salt and a spatially heterogeneous environment with respect to cadmium and salt. Our focus is on robustness in a major fitness component, juvenile viability. Because environmental robustness is defined as the reduction in phenotypic variation under nongenetic perturbation (Flatt 2005), so the variance of a trait among environments could be used as an estimate for the robustness of that trait. We assay juvenile viability in a set of
novel environmental conditions distinct from the environments they have evolved in and assess environmental robustness for each population. Because the novel environments vary with respect to a range of abiotic conditions that are known to be stressful for D. melanogaster, populations that express less variance in mean viability among environments indicate a higher level of environmental robustness.

We consider two hypotheses on the evolution of robustness. First, robustness will be greater in populations that have evolved with environmental heterogeneity. This is because under heterogeneous regimes, populations experiencing fluctuating selection may be more likely to develop canalizing mechanisms (i.e., regulatory/canalizing genes and gene products) to cope with multiple stressors. While populations evolving in the constant environment will be closer to the optimum value in that environment than populations from other selection regimes, such populations may do very poorly in another environment. Populations evolved with environmental heterogeneity may be reasonably robust across environments. This robust phenotype may not be perfectly matched to every constant environment but overall has higher average fitness as well as lower variance in fitness and this robustness might carry into novel environmental conditions. Although theory has shown both temporal and spatial heterogeneity could play an important role in the evolution of canalization (Proulx and Phillips 2005), which form of heterogeneity is more likely to select for environmental robustness is unclear.

The second hypothesis is somewhat less intuitive and has two parts: (i) robustness will be greater in populations under one constant selection regime than under a different constant regime; and (ii) populations evolving in heterogeneous environments will have intermediate levels of robustness. The first part of the hypothesis arises simply because
adaptation to one constant environment involves mechanisms that, by chance, confer robustness across a broader set of environments than adaptation to some other constant environment. Thus, robustness is a by-product of adaptation, not the target of direct selection.

For example, if adaptation to Environment A is multifunctional while adaptation to Environment B is very specialized, populations adapted to Environment A will likely be robust with higher average fitness across a range of environments whereas populations adapted to Environment B will be sensitive to external perturbation. Populations facing both Environment A and B (i.e., heterogeneous environments) will acquire both adaptations but neither to the full extent (assuming they are costly) and thus will have an intermediate level of robustness.

Hypothesis 1 is most likely to apply to the original selective condition (cadmium and salt). This is because populations under the heterogeneous regimes are specifically selected to cope with the adverse effects on fitness caused by both salt and cadmium, thus likely to have high level of robustness across both environments compared to populations from either constant environment. Therefore, we have a strong expectation for hypothesis 1 to hold for the original environments but any one of the hypotheses may be true in the novel environments that have never been experienced by our experimental fly populations.

Whether fitness, as a trait, should evolve robustness is a complex question without clear answers (De Visser et al. 2003). Nonetheless we chose to focus on a major fitness component in our study because there is likely selection for buffering of the most important physiological functions and developmental pathways that determine the survival of an individual (Ghalambor et al. 2007). If such buffering does exist, it could be demonstrated by the reduced sensitivity of fitness under a range of environmental stressors. Thus, by
measuring robustness of fitness, we can indirectly assess the extent of robustness for those fundamental traits. However, a high level of robustness of fitness does not necessarily mean high average fitness across environments under the definition that robustness is inversely proportional to variance. For example, high robustness could be achieved by either displaying consistently low fitness in most environments or consistently high fitness in most environments. Therefore, we also assess the correlation between robustness of viability and average viability among environments. If canalization is a beneficial phenomenon, then highly canalized populations should also have higher mean fitness across environments than less canalized populations (G.P. Wagner, et al. 1997).
Methods

Selective History of Experimental Populations

All *Drosophila melanogaster* populations used in this study were derived from the grand ancestor population, which was originally collected from the Similkameen Valley, British Columbia in 2005. This population was maintained in a benign lab condition (cornmeal-yeast media; 25°C; 12L:12D photoperiod; 50% RH) at a population size of 2000-4000 flies.

In July 2007, 12 replicate populations each of size ~1000 were drawn from the grand ancestor population and maintained in a cadmium-enriched environment based on the cornmeal-yeast media. In the following April, those 12 populations were pooled together to establish the ancestral cadmium population (approximate 2000 flies). In August 2008, a subset of flies from the grand ancestor population was used to establish the ancestral salt population (approximate 2000 flies) by maintaining those flies in a salt-enriched environment based on the cornmeal-yeast media. The concentration of salt and cadmium in the two environments were gradually increased to 75 μg/ml and 29 mg/ml respectively before the establishment of the experimental evolution lines which are described below.

In October 2009, 448 virgin males and 448 virgin females were collected from both the ancestral salt population and the ancestral cadmium population and were crossed to flies from the other population via mass mating. Offspring from the next generation were randomly assigned to one of the four selective regimes: a constant salt environment (hereafter *Salt*), a constant cadmium environment (*Cad*), a temporally heterogeneous regime...
(Temp) that alternate every generation between salt and cadmium environment and a spatially heterogeneous regime (Spatial) that split the population between salt and cadmium environment within each generation. Each selective regime had five replicate populations that were maintained in 14 vials with 16 males and 16 females in each vial. At generation 158 the maintenance scheme was switched from vials to bottles where each population was held in 2 bottles (~250 flies per bottle). Those populations were evolving independently in their respective selection regime for ~190 generations before the start of this study which commenced in January 2017, and were referred as the experimental populations. Prior to the actual experiment the population size was doubled for each replicate population. All populations were maintained on a two-week non-overlapping generation cycle.

**Test Environments**

The two original selection environments (cadmium and salt) and 16 novel environments were chosen as test environments in which robustness in juvenile viability of the experimental populations were measured. The concentration of cadmium and salt were 76 \( \mu \text{g/ml} \) and 27 mg/ml respectively, which is slightly different than concentration at the beginning of the evolution experiment. Two of the 16 novel environments were benign conditions for the flies whereas the other 14 were stressful conditions that occurred outside the range of ideal conditions for rearing *D. melanogaster*. The novel environments varied with respect to a range of abiotic factors including food quality, temperature, water availability and chemical substance of the medium (further details on the composition of the novel environments can be found in Table 1.)

**Viability Assay**
Beginning in January 2017 till July 2017, larva-to-adult viabilities were measured in test environments conducted in separate blocks, meaning all experimental populations were tested at the same time for a given environment but different environments were tested at different times. To reduce the potential impact of maternal effect on the results, all experimental populations were reared on a common benign food media (cornmeal-yeast) for one generation before being tested in each environment. Offspring from each population were pooled into a large cage to mate and lay eggs (one cage per population with ~2000 flies). To standardize the developmental stage of egg and larvae, three large grape juice lay-plates with live yeast paste added on top were placed in each cage for ~90 mins to allow females to oviposit before being discarded. Immediately afterwards 10 small grape fruit lay-plates with live yeast paste were placed in each cage for ~8 hours. Plates with eggs were then incubated for ~16 hours at 25 °C before larvae collection. For every population in each test environment, 25 replicate vials were collected each with 30 first instar larvae (some populations had less than 25 vials). Larvae were collected from the lay-plates using a small metal rod with a sharp tip. The number of surviving adult flies in each vial were counted on the day when it was estimated that more than half of flies had emerged from the pupae; vials were then counted again a few days afterwards (i.e., vials were typically counted on days 11 and day 14). All the vials for a given environmental assay were counted on the same days. In addition to the experimental populations, a single population of the grand ancestor stock was also tested in the 6 novel environments in a similar fashion.

In addition to larval viability, egg-to-adult viability was tested for each population in the original selective environments. The assay protocol was identical to the one used in the larval viability assay except 50 eggs were transferred into each vial and each population had
~15 replicate vials.

**Data Analysis**

Larva-to-adult viability was calculated as the number of surviving adults in each vial summed across both scoring days divided by 30. Results from the original selection environments (cadmium and salt) and the 16 novel environments were analyzed separately and all analysis was performed in R 3.2.2. (R Core Team 2015).

Results from the novel environments were first analyzed using a linear mixed effect model to assess the treatment effect on viability using the lmer function in the lme4 package. Viability was treated as the dependent variable whereas selection history as a fixed effect, population as a random effect nested under selection history and environment as another random effect. The full model was \( \text{Viability} \sim \text{Selection\_History} + (1|\text{Environment}) + (1|\text{Selection\_History} : \text{Population}) \). Post-hoc comparison of mean viability was carried out using least-squared means to measure differences between selection regimes. Furthermore, the among-environment variance of mean viabilities was calculated for each population and was used as the response variable in a one-way ANOVA to examine heterogeneity in the among-environment variance between selection regimes. The among-environment variance of each population was regressed against the overall mean viability of that population (the average of mean viabilities in all novel environments) to detect any relationship between those two parameters. The model is \( \text{Among\_environment\_Variance} \sim \text{Mean\_Viability} + \text{Selection\_History} \). Because the data are based on means of proportions (which are greater than 50%), a slight negative relationship between mean and variance is expected due to sampling. To test if this sampling effect plays a dominate role in producing observed results,
we simulated the mean viabilities of each population in each novel environment by assuming all the means are equal to the observed overall mean viability of that population. This was done in R with `rbinom` function for 5000 times. For example, for population A with an overall mean $X_A$, the command `mean(rbinom(25, 30, X_A/30))` would simulate the mean viability in one environment for population A. We then calculated and regressed the simulated among-environment variances against the overall means in the same way as we did with the real data. If the sampling effect does not play a dominant role in deciding the relationship between mean and variance, then we expect to see the observed regression slope being steeper than simulated slopes.

Theory has predicted that hidden genetic variation will be released in stressful conditions resulting increased phenotypic variation (Hoffmann and Parsons 1991; Scharloo 1991; Hoffmann and Merila 1999). This might be caused by the breakdown of canalizing mechanisms buffering mutational effects. To assess the effect of environmental stress on unmasking genetic variation among divergent populations, the among-population variance of mean viabilities within selection regime was regressed against regime-level mean viability in the novel environments. An environment with low average viability is interpreted as being more stressful than environments with high average viability. The model is $\text{Among-population Variance} \sim \text{Mean_Viability} + \text{Selection_History}$; here $\text{Mean_Viability}$ is the average viability across populations within a regime for a single environment. Similarly, because the data are based on means of proportions (which are greater than 50%), a slight negative relationship between mean and variance is expected due to sampling. To test if this sampling effect plays a dominate role in producing observed results, we simulated the mean viabilities of each population in each environment by assuming all the means are equal to the
observed selection regime mean in that environment. For example, if regime A had an overall mean $X_A$ in a given environment, the command `mean(rbinom(25, 30, X_A/30))` would simulate the mean viability of one replicate population under regime A in that environment. 

We then calculated and regressed the simulated among-population variances against the regimes means in the same way as we did with the real data. If the sampling effect does not play a dominant role in deciding the relationship between mean and variance, then we expect to see the observed regression slope being steeper than simulated slopes.

Results from the cadmium and salt assay environments were analyzed separately using two linear mixed effect models with viability as the dependent variable, selection regimes and population as explanatory factors. A one-way ANOVA was used to test the difference of among-environment variance of mean viability between selection regimes.

Because there is only a single Grand Ancestor population its mean viability (where measured) and among-environment variance were used as a qualitative comparison with the experimental populations but not in the formal statistical analysis.
Results

Viability in Original Environments

Larva-to-adult viabilities in both salt and cadmium assay environments were significantly affected by the different selection regimes ($F_{3,16} = 11.841, p = 2.457 \times 10^{-4}$, $F_{3,16} = 16.306, p = 4.016 \times 10^{-5}$). In the cadmium assay environment, *Salt* had the lowest mean viability whereas in the salt assay environment, *Cad* had the lowest mean viability (Figure 1a). One-way ANOVA testing for differences in among-environment variance showed a significant effect of selection regime ($F_{3,16} = 15.993, p = 4.504 \times 10^{-5}$). *Salt* had the lowest variance and *Cad* had the highest variance. The egg-to-adult viability assay yielded qualitatively similar results (Figure 1b). Results from an earlier study obtained from the same populations were different than ours (Figure 1c).

Viability in Novel Environments

Viability in the 16 novel environments is summarized in figure 2. Using a linear mixed effect model, we found the mean viability differs significantly among selection regimes, ($F_{3,16} = 23.77, p = 3.936 \times 10^{-6}$). Post-hoc comparison of least-square means showed populations from the *Cad* regime had significantly lower mean viability than the other three regimes (Table 2). The mean viabilities were not statistically different among *Salt*, *Spatial*, and *Temp* regimes though the *Salt* regime had the highest point estimate of mean viability (Table 3). To quantify environment robustness of viability for each selection regime, we calculated the among-environment variance of viability for each population as an estimate of robustness. Based on a one-way ANOVA, there is significant heterogeneity among regimes in among-environment variance ($F_{3,16} = 3.5777, p = 0.038$). Populations from *Salt*, *Spatial*,
and *Temp* regimes had smaller among-environment variance than *Cad* regime (Figure 3, Table 3).

To detect if there was any relationship between robustness and average fitness, the among-environment variance of each population was regressed against the population mean viability and there was a strong correlation between those two factors (Figure 4), where high robustness (i.e., low variance) was associated with high average fitness (*t* = -3.479, *p* = 0.00336, *R*^2^ = 0.6688). Furthermore, the observed slope was steeper than the slopes of all 5000 simulations designed to account for sampling effects (see Methods for details).

In addition, we tested the relationship between environmental stress and among-population variation within a given regime and there was a negative relationship between mean viability of the environments and the among-population variance (Figure 5, *t* = -3.385, *p* = 0.00127, *R*^2^ = 0.2192). Results from the simulated data showed that the observed slope was steeper than expected by sampling; none of the 5000 simulated regression slopes were steeper than the observed slope.

The grand ancestor population tends to have low viability in stressful conditions. It had the lowest mean viability and the highest among-environment variance compared to populations from other regimes in the 6 assays we tested.
Discussion

*Original Environments*

We expected that populations under constant environments should have the highest viability in their home environment, whereas populations from a different constant environment should have the lowest viability and populations from heterogeneous environments should have intermediate level of viability. In terms of among-environment variance, *Spatial* and *Temp* populations should be more robust than *Cad* and *Salt* populations. This prediction was confirmed by Huang (2014) using the same populations with a slightly different fitness assay at an earlier time point (Fig. 1c). However, two new assays conducted in the *original* selective environments were inconsistent with our expectation and the earlier results (Fig 1a, b). In all three assays, populations showed evidence of local adaptation in the sense that *Cad* had significantly higher viability than *Salt* in the cadmium assay and vice versa in the salt assay. However, in the two current assays (Fig 1a, b), absolute viability (rather than viability relative to the best regime) is reasonably high for the *Salt* regime in both salt and cadmium. Consequently, *Salt* has the smallest change in viability across the two diets (i.e., lowest among-environment variance) of all four regimes. The observed change in viability between Huang (2014) and this study can be attributed mostly to the increased survival in all populations in the cadmium assay and reduced survival of *Salt* in the salt assay. Because data from Huang (2014) were obtained at generation ~50 and our results were collected at generation ~190, there might have been considerable amount of adaptive change that led to increased survival in the cadmium assay. There were also differences in how the assays were conducted which may be important for the different patterns. For example, the
concentration of salt used in the assay from Huang (2014) was 22 mg/ml which is lower than what we used.

**Novel Environments**

Results from the novel environments were consistent with hypothesis 2. *Salt* populations had the highest mean viability and lowest variance whereas *Cad* populations had the lowest mean viability and highest variance. *Spatial* and *Temp* populations were intermediate with respect to both of mean viability and among-environment variance compared to populations from constant regimes. Notably, average variance from *Temp* was as low as average variance from *Salt*. This indicates that environmental robustness in populations from the four selection regimes follow this order: \( \text{Salt} = \text{Temp} > \text{Spatial} > \text{Cad} \) (Table 3). If adaptation to salt can be used to buffer against other environmental stressors whereas adaptation to cadmium has no benefit to other types of perturbation, then it follows quite straightforwardly that populations evolved in constant salt regime will enjoy higher average fitness as well as robustness of fitness than populations evolved in constant cadmium regime across a range of external conditions. Populations exposed to both salt and cadmium during adaptation will have intermediate values of fitness and robustness. Indeed, empirical studies have found salt tolerance is correlated with drought and cold tolerance in both plants (Hare and Cress 1997; Munns 2002) and insects (Shimada and Riihimaa 1990; Fields et al. 1998; Misener *et al.* 2001; Yerushalmi 2016). This is largely due to the presence of osmotic stress in all three types of perturbation, thus selecting similar developmental and metabolic responses in exposed individuals such as the production of proline. On the other hand, adaptation to cadmium in fruit flies involves duplication of metallothionein gene which is metal-specific and less likely to contribute to buffering of other environmental stressors.
(Posthuma and Straalen 1993). Furthermore, we found that robustness was positively correlated with average viability across environments, implying robustness could be advantageous in variable stressful conditions. The grand ancestral population, which was maintained in a benign environment, had the lowest average viability and highest among-environment variance in 6 novel environments tested. This further supports the idea that evolving with environmental stress could select for increased environmental robustness as an adaptive strategy.

Theory has predicted that environmental stress could lead to genetic decanalization and subsequent release of cryptic genetic variation that is not expressed in normal conditions (Hoffmann and Parsons 1991; Scharloo 1991; Hoffmann and Merila 1999). One well known example is the chaperon protein Hsp90. When its function is inhibited by either genetic or environmental factors, previously hidden genetic variation is phenotypically expressed (Rutherford and Lindquist 1998). Most studies on cryptic genetic variation have focused on among-individual variation within a population. Based on the assumption that stress could also exaggerate the differences among populations, we took another perspective by examining the among-population variance of replicate populations with similar evolutionary histories. Because those replicate populations were derived from a common ancestral population, they started with similar amount of genetic variation. As they evolve independently, the divergence among them is largely due to genetic drift and the accumulation of cryptic mutations. We found that environmental stress was positively correlated with among-population variation (Fig. 5), confirming the expectation that stress could induce the release of hidden genetic variation upon genetic decanalization.
One of the major aims of evolutionary biology is to understand how the diversity of life is produced and maintained by natural selection. Traditionally, patterns of selection and mutation have been the main focuses in the attempt to explain this diversity (Pelabon et al. 2010). After the establishment of a conceptual distinction between variation and variability, where the former describes the level of differentiation already present and the latter refers to a trait’s capacity to vary (Wagner and Altenberg 1996), more recent works have redirected some attention to whether variability itself can respond to selection and evolve (Kawecki 2000; Pelabon et al. 2010; Hallsson and Bjorklund 2012a; Le Rouzic et al. 2013). It is widely accepted that variability is determined by both sources of variation (mutation and environmental effects) and regulatory processes such as canalization, phenotypic plasticity and developmental stability (Debat and David 2001). Furthermore, these regulatory mechanisms are shown to be under genetic control and exhibit considerable amount of genetic variation (Flatt 2005), providing opportunity for selection to shape the pattern of variability under different scenarios.

Although much effort has been made towards understanding the genetic architecture and evolutionary causes and implications of these processes (Gibson and Wagner 2000; Siegal and Bergman 2002), results from studies investigating the effect of different types of selection on canalization are equivocal. Before discussing the implications of these findings, I want to clarify two major differences between these studies and ours. First, the majority of the past studies have focused on the effect of stabilizing and/or fluctuating selection on robustness of a specific trait (physiological/morphological). Here I looked at how the sensitivity of a fitness component, juvenile viability, is affected in populations evolving with
and without environmental heterogeneity. Although some underlying traits are probably under stabilizing selection in the constant regimes and under fluctuating selection in the heterogeneous regimes, fitness itself is under directional selection and high viability is likely favoured in all situations. Thus, there is probably selection for adaptive robustness of high fitness (De Visser et al. 2003). The second difference is that most authors quantified environmental robustness as among-individual variance within a given environment whereas we assessed among-environment variance of viability. This is due to differences in the definition of the concept, while some refer to environmental robustness as buffering against all sources of external perturbations, others define environmental robustness as buffering against external micro-environmental variation. However, the difference between macro- and micro-environments is an arbitrary choice on a continuous scale of environmental effects and there is no clear distinction faced by natural populations (Sztepanacz et al. 2017; Debat and David 2001). Therefore, environmental canalization can refer to buffering against both macro- and micro-environmental variation (Flatt 2005). With these distinctions in mind, now we review both theoretical and empirical findings on how the evolution of robustness is affected by different selection regimes.

Most authors modeled the evolution of canalization under stabilizing selection as suggested by Waddington (Rendel 1967; Layzer 1980; A. Wagner 1996; G.P. Wagner et al. 1997; Gibson and Wagner 2000; Zhang 2005). While environmental canalization generally increases with increasing strength of stabilizing selection (G.P. Wagner et al. 1997; Meiklejohn and Hartl 2002), genetic canalization may be weakened if the strength of selection is too strong that genetic variation is purged below an optimal level for the evolution of canalization (G.P. Wagner et al. 1997). Fluctuating selection on the other hand,
is considered by most to favor de
canalization and plasticity (Flatt 2005; Zhang and Hill 2005; Lande 2009), but Kawecki (2000) found that canalizing alleles were favored under fluctuations with periodicities below 10 generations. Le Rouzic et al. (2013) also showed that canalization can occur under both stabilizing and fluctuating selections but the latter was less efficient at producing canalized equilibria.

Compared to the large body of theoretical literature on this topic, there are few empirical studies testing how different types of selection influence the evolution of canalization, especially environmental canalization. One study examined the effect of stabilizing, fluctuating and disruptive selection on the among-individual variation of a wing shape character in *Drosophila melanogaster*. After subjecting two outbred lab populations to the three types of selection for 20 generations, they found that phenotypic variation strongly increased under disruptive selection but decreased under fluctuating and stabilizing selection relative to the controls, and fluctuating asymmetry (within-individual variation) increased in all selection treatments (Pelabon et al. 2010). This implies that only disruptive selection favors de
canalization and the underlying genetic mechanisms for canalization and development stability are distinct. Another study exposed seed beetles *Callosobruchus maculatus* to continuous linear and/or fluctuating increase in temperature for 18 generations, and tested how these different selection regime affect among-individual variation within a given environment. They found both genetic variation and environmental variation decreased in some life-history traits under fluctuating selection (Hallsson and Bjorklund 2012a). This result hinted at the possibility of a common buffering mechanism against both genetic and environmental perturbation as suggested by theory (Meiklejohn and Hartl 2002). Overall, both studies showed support for increased genetic and environmental canalization under
fluctuating selection. Although these results provided insight into the evolution of canalization, they are not conclusive tests for the systematic difference between stabilizing and fluctuating selections as 20 generations is a relatively short time scale for selection to operate (Le Rouzic et al. 2013). Both studies examined canalization as among-individual variation within a macro-environment. However, the effect of different selection on among-environment variance of the mean is unclear.

In our study, the experimental populations have been evolving in different selection regimes for ~190 generations and significant genetic differentiation has accumulated among populations under different regimes (Huang 2014). This provided a more powerful comparison between selection in constant environments and selection in heterogeneous environments and their effect on the evolution of canalization. More importantly, past selection experiments investigating the role of environmental heterogeneity only focused fitness in the original selective environments. Our study looked at the effect of environmental heterogeneity on robustness of fitness both in the original environments and under novel conditions. Our results suggest that environmental robustness is an indirect product of adaptation to specific environments and environmental heterogeneity plays a relatively less important role in inducing robustness than the details of adaptation.

This study also has its limitations. Because only robustness of one important fitness trait (juvenile viability) was investigated, we do not have knowledge of the actual set of underlying traits affected by environmental perturbations. Thus, a better understanding of how different types of selection affect robustness can be obtained by measuring traits that are targets of selection at the physiological/morphological level; and/or by investigating other fitness components such as fecundity so that we can infer whether the pattern of robustness
of different fitness related traits are concordant or if there exist trade-offs in buffering against perturbations. Furthermore, using outbred populations increases the realism of this study but reduces our ability to detect the relative contribution of different factors that affect among-environment variance of viability. For example, the change in mean viability could be affected by both environmental and genetic canalization. In some stressful environments tested here, hidden genetic variation might be released in one selection regime more easily than another, thus affecting larval viability. We do not know how much of the change in viability is due to environmental canalization or the effect of cryptic genetic variation being exposed. Genetic canalization is also a very interesting topic because of its debated role in the evolution of evolvability (De Visser et al. 2003; Flatt 2005). On one hand, it reduces the phenotypic effect of mutation, thus constrains the efficacy of selection; on the other hand, genetic canalization is shown to increase the amount of cryptic genetic variation that could be used for future adaptation. Further research could investigate empirically if environmental canalization and genetic canalization are correlated responses by measuring the degree of genetic buffering in populations used here.
Tables and Figures

| A | A standard sugar-yeast-agar recipe. 10 mL food per vial housed at 25 C. |
| B | A high-nutrition medium based on cornmeal, sugar and yeast |
| C | A reduced volume of cornmeal-based food, only 2 mL/vial |
| D | Cornmeal medium housed at 17 C |
| E | Cornmeal medium housed at 31 C |
| F | 25% less water in cornmeal medium |
| G | 25% more water in cornmeal medium |
| H | Acetic acid (50 ml/L) added to cornmeal medium |
| I | Caffeine (0.75 g/L) added to cornmeal medium |
| J | Ethanol (100 ml 95% ethanol per 1 L of food) added to cornmeal medium |
| K | Urea (4g/L) added to cornmeal medium |
| L | 50% less water in cornmeal medium |
| M | 50% more water in cornmeal medium |
| N | A poor-nutrition medium made with cornstarch |
| O | A cold shock at 3 degrees for 4 hours is applied to second instar larvae |
| P | A heat shock at 39 degrees for 4 hours is applied to second instar larvae |

**Table 1.** Details of the 16 novel environmental assays.
<table>
<thead>
<tr>
<th>Contrast</th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t ratio</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cad-Salt</td>
<td>-0.119</td>
<td>0.01491771</td>
<td>15.98</td>
<td>-7.995</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Cad-Spatial</td>
<td>-0.082</td>
<td>0.01491798</td>
<td>15.98</td>
<td>-5.499</td>
<td>0.0003</td>
</tr>
<tr>
<td>Cad-Temp</td>
<td>-0.093</td>
<td>0.01491785</td>
<td>15.98</td>
<td>-6.214</td>
<td>0.0001</td>
</tr>
<tr>
<td>Salt-Spatial</td>
<td>0.037</td>
<td>0.01491245</td>
<td>15.95</td>
<td>2.497</td>
<td>0.0987</td>
</tr>
<tr>
<td>Salt-Temp</td>
<td>0.027</td>
<td>0.01491231</td>
<td>15.95</td>
<td>1.782</td>
<td>0.3172</td>
</tr>
<tr>
<td>Spatial-Temp</td>
<td>-0.011</td>
<td>0.01491258</td>
<td>15.95</td>
<td>-0.715</td>
<td>0.8898</td>
</tr>
</tbody>
</table>

Table 2. Post-hoc comparison of least-square means among selection regimes. The estimate is the difference in mean viability between two regimes.
<table>
<thead>
<tr>
<th>Selection History</th>
<th>Mean Viability ± SE</th>
<th>Average Among-environment Variance ± SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cad</td>
<td>0.736 ± 0.015</td>
<td>0.0105 ± 0.001</td>
</tr>
<tr>
<td>Salt</td>
<td>0.857 ± 0.004</td>
<td>0.0038 ± 0.0011</td>
</tr>
<tr>
<td>Spatial</td>
<td>0.82 ± 0.013</td>
<td>0.0063 ± 0.0029</td>
</tr>
<tr>
<td>Temp</td>
<td>0.831 ± 0.007</td>
<td>0.0038 ± 0.0006</td>
</tr>
</tbody>
</table>

**Table 3.** Mean larval viability in the *novel* environments for each selection regime and the average among-environment variance of mean calculated from replicate populations within selection regime.
Figure 1a. Larval-to-adult viability of each selection regime in the *original* selective environments. Error bars are SE among replicate population means. The ancestral population was also included for reference.
Figure 1b. Egg-to-adult viability of each selection regime in the original selective environments. Error bars are SE among replicate population means.
Figure 1c. Data obtained from Huang (2014). Larva-to-adult viability of each selection regime in the original selective environments. Error bars are SE among replicate population means.
Figure 2. Mean larval viability of each selection regime in the novel environments. Each panel is a distinct environment. Error bars are SE among replicate population means. For some of the environments, larval viability of the ancestral population (black) was also measured.
**Figure 3.** Boxplots of mean viabilities of each population across the 16 novel environments.

Populations from the same selection regime share identical color. Populations with bigger boxes have higher among-environment variance of mean viability.
Figure 4. Overall mean viability averaged from 16 novel environments regressed against among-environment variance of mean viabilities for populations (stars with red regression line) and selection regimes (solid circles with black regression line). Each star represents a single population plotted with respect to its average viability among the 16 novel environments and its among-environment variance in viability. Solid points are calculated combining all the data among populations for within each regime.
**Figure 5.** Mean viability at the selection regime level regressed against variance of means among replicate populations. Here each point represents the data associated with a single selection regime in one *novel* environment; for each point, the x-axis value is the average viability calculated among populations within a regime measured in a single environment and the y-axis value is the variance among populations (within a regime) in viability for that environment.
Reference


De Visser, JA; Herisson, J; Wagner, GP; Ancel Meyers, L; Bagheri-Chaichian, H; Blanchard, JL; Chao, L; Cheverud, JM; et al. (2003) "Perspective: Evolution and detection of genetic robustness". Evolution; international journal of organic evolution. 57 (9): 1959–72.


