Title: How phenotypic convergence arises in experimental evolution

Running Title: Experimental Evolution of Convergence

Authors: Pedro Simões[1#*] & Inês Fragata[1#*], Josiane Santos[1], Marta A. Santos[1], Mauro Santos[3], Michael R. Rose[2] & Margarida Matos[1]

Affiliations: [1] cE3c – Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal, Tel: +351 21 75 000 00**;** [2] Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA.; [3] Departament de Genètica i de Microbiologia, Grup de Genòmica, Bioinformàtica i Biologia Evolutiva (GGBE), Universitat Autonòma de Barcelona, Spain.

[*]Corresponding authors: PS: pmsimoes@fc.ul.pt, IF: irfragata@gmail.com
# both authors contributed equally to this work.

Author Contributions: P.S. and I.F. participated in the conception and design of the study, collected and analysed the data and drafted the manuscript. J.S. participated in the conception and design of the study, data collection and revised critically the manuscript. M.A.S. participated in data collection and revised critically the manuscript. M.S. and M.R.R. participated in the conception and design of the study, as well as substantively revising the manuscript. M.M. conceived and designed the study, coordinated the study and helped draft the manuscript. All authors gave final approval for publication.

Acknowledgments: The authors thank Laurence D. Mueller for the first version of the script that was later adapted for the analyses of the present study. This study was partially financed by "Fundação para a Ciência e a Tecnologia" (FCT) projects POCTI/BSE/33673/2000, POCI-PPCDT/BIA-BDE/55853/2004 and PTDC/BIA-BDE/65733/2006 and by the cE3c FCT Unit UID/BIA/00329/2013. PS and IF have postdoctoral FCT grants (SFRH/BPD/86186/2012 and within the project JPIAMR/0001/2016 respectively). MS is funded by grants CGL2017-89160P from Ministerio de Economía y Competitividad (Spain), and 2017SGR 1379 from Generalitat de Catalunya.

Data Accessibility Statement: Raw data will be available at the Dryad Digital Repository upon acceptance.

Keywords: Convergent Evolution; Experimental Evolution; *Drosophila subobscura*; Life-History traits; Long-term Evolution; Adaptation.

Abstract

Evolutionary convergence is a core issue in the study of adaptive evolution, as well as a highly debated topic at present. Few studies have analyzed this issue using a "real-time" or evolutionary trajectory approach. Do populations initially differentiated converge to a similar adaptive state when experiencing a common novel environment? *Drosophila subobscura* populations founded from different locations and years showed initial differences and variation in evolutionary rates in several traits during short-term (~20 generations) laboratory adaptation. Here we extend that analysis to 40 more generations to analyze (1) how much differences in evolutionary dynamics between populations change between shorter and longer time spans, and (2) whether evolutionary convergence occurs after sixty generations of evolution in a common environment. We found substantial variation in longer-term evolutionary trajectories and differences between short and longer-term evolutionary dynamics. Though we observed pervasive patterns of convergence between populations towards the values of long-established populations, populations still remain differentiated for several traits at the final generations analyzed. This pattern might involve transient divergence, as we report in some cases, indicating that more generations should lead to final convergence. These findings highlight the importance of longer-term studies for understanding convergent evolution.

Introduction

Understanding how populations adapt to environmental challenges is becoming increasingly important in both evolutionary biology and conservation (Franks and Hoffmann 2012; Botero et al. 2015). However, we are still unsure how predictable adaptation to novel environments is (Wiser et al. 2013; Lachapelle et al. 2015; Lenski et al. 2015; Orgogozo 2015; Lässig et al. 2017). Unpredictability in evolution can be caused by different genetic backgrounds due to prior evolutionary history (see Barton and Keightley 2002; Barrett and Schluter 2008; Hansen 2013), and stochastic events such as founder events, genetic drift, bottlenecks, etc. (see Lenormand et al. 2009). Furthermore, interactions between selection and genetic drift may also increase variation in evolutionary responses (e.g. Cohan 1984; Cohan and Hoffmann 1986; Santos et al. 2012).

An important question when different populations adapt to new environmental challenges is whether they will diverge or converge through time. Convergent evolution is expected to arise through the action of natural selection, erasing differences between populations (Endler 1986; Losos 2011; Stern 2013). Alternatively, differentiated populations could conceivably evolve increased differentiation when placed under similar selective regimes (Wright 1931; Cohan 1984; Whitlock et al. 1995). Discovering the constraints that produce either evolutionary convergence or evolutionary divergence is fundamental to ultimately understanding the foundations of adaptive evolution.

Experimental evolution is a powerful tool with which to address this problem, especially by studying the real-time evolutionary trajectories of different populations subjected to the same selective challenge. Several studies have observed convergent evolutionary responses in a new common environment (e.g. Travisano et al. 1995;

Teotónio and Rose 2000; 2002; Joshi et al. 2003; Simões et al. 2007, 2008; Teotónio et al. 2009; Santos et al. 2012; Fragata et al. 2014; Burke et al. 2016). Nevertheless, divergent evolutionary responses have also been observed (e.g. Cohan 1984; Cohan and Hoffmann 1986; Melnyk and Kassen 2011). Furthermore, several studies support the notion that the impact of evolutionary contingencies varies (history and chance) between traits closely or loosely related to fitness (Travisano et al. 1995; Teotónio et al. 2002; Joshi et al. 2003; Simões et al. 2008, 2017). It is thus clear from experimental evidence that evolutionary contingencies have a role in shaping evolutionary responses.

An important question, seldom addressed in the literature (but see Burke et al. 2016), is the effect of initial differentiation between populations on their long-term evolution. In particular, it is expected that different initial genetic backgrounds will have a higher impact during short-term evolution in a constant environment (Joshi et al. 2003; Fragata et al. 2014; Burke et al. 2016). On the other hand, at longer evolutionary scales, the cumulative effects of genetic drift and other stochastic events acting on the evolving populations will likely have a higher impact on the evolutionary trajectories observed (e.g. see Brito et al. 2005; Lenormand et al. 2009). Furthermore, different levels of standing genetic variation and/or epistatic interactions can have an important impact on long-term evolution (Barrett and Schluter 2008; Goodnight 2015; Paixão and Barton 2016; see empirical examples in Barton and Keightley 2002; Hansen 2013; Wiser et al. 2013; Good and Desai 2015). This might produce differences between populations, even in populations subject to similar selective pressures, possibly through different timings in the deceleration of the evolutionary response over time, for example (Teotónio and Rose 2000; Gilligan and Frankham 2003; Simões et al. 2007; Khan et al. 2011; Schoustra et al. 2012).

Long-term evolutionary dynamics have been mostly studied in microbial experimental evolution systems rather than in sexual organisms, due to the shorter generation time of the former. In the *E. coli* long-term evolution experiment performed in Lenski's lab, recent evidence indicates a deceleration of the evolutionary rate over 50000 generations (Wiser et al. 2013; Lenski et al. 2015). Furthermore, and perhaps surprisingly, heterogeneity in evolutionary trajectories is still present after so many generations, in part due to differences in mutation rates (Lenski et al. 2015). Several studies with sexual organisms, though involving fewer generations, have also observed the slowing down of evolutionary responses to newly imposed selection regimes (e.g. Gilligan and Frankham 2003; Rose et al. 2004; Simões et al. 2007, see below). The expectation of a deceleration of laboratory evolutionary trajectories in sexual organisms is sometimes justified in terms of temporal exhaustion of additive genetic variance, although genomic scans in experimentally evolved *Drosophila* populations have found only limited evidence of fixed alleles following selection (Burke et al. 2010; Burke and Long 2012; Orozco-Terwengel et al. 2012; Long et al. 2015; Phillips et al. 2016; Seabra et al. 2018). In a previous study by our team, we found evidence for a deceleration in the evolutionary trajectory of fecundity in populations of *Drosophila subobscura* evolving for more than 80 generations in the lab environment (Simões et al. 2007). Teotónio and Rose (2000) also found this pattern of response in several *D. melanogaster* lines undergoing reverse selection to their ancestral environment. Gilligan and Frankham (2003) also reported a slowing down of the rate of adaptation to captivity after 87 generations in the lab by comparing *Drosophila* populations in different stages of adaptation. However, this pattern is not universal, as other experimental studies have not found such deceleration of the evolutionary response, even after a higher number of generations. One example of this is the work of Chippindale et al. (1997), who imposed

selection for accelerated development time in *D. melanogaster*. Nevertheless, studies of long-term experimental evolution in sexual species are scarce, and have not specifically addressed the variation in evolutionary dynamics that might occur during evolution in short-term evolution, relative to longer evolutionary time periods.

We have previously shown variation in the evolutionary response of several populations of *Drosophila subobscura* during the first 20 generations of evolution in a new environment, the lab (Simões et al. 2008). These populations were founded from different nearby locations over several years. We observed higher variation in the evolutionary response for female starvation resistance, a trait likely more loosely related to fitness in our experimental setting. By contrast, patterns for fecundity traits, which are expected to be closer to fitness, were more repeatable. Importantly, the different starvation resistance patterns led in fact to convergence between populations. In this study we extend the earlier analysis to cover around forty additional generations. We address the following questions: (1) How much do evolutionary rates vary between short and longer-term evolution? (2) Do differences in evolutionary dynamics between populations change in the transition from earlier to later generations? (3) Is convergence observed after sixty generations of evolution in a common environment?

We expect that, during short-term evolution, variation in the initial genetic backgrounds will lead to disparate rates of adaptation to the new environment. Over the longer term, as the evolutionary response decelerates, differences between populations of contrasting initial genetic composition are likely to be reduced relative to those observed during short-term evolution, particularly if populations are evolving towards the same phenotypic optimum.

Materials and Methods

Founding and Maintenance of the Laboratory populations

Five sets of wild-caught samples of *Drosophila subobscura* were analyzed in this study. These populations were founded in 1998 (NW populations; see Matos et al. 2002), 2001 (AR and TW populations; see (Simões et al. 2007), and 2005 (FWA and NARA; see (Simões et al. 2008). NW, TW and FWA populations were collected from a pinewood in Adraga (Portugal), whereas AR and NARA populations were collected from a pinewood in Arrábida (also from Portugal, some 50 Km from Adraga, on the other margin of the Tagus river; see Simões et al. 2007, 2008). All populations were three-fold replicated two generations after founding (e.g., FWA_{1-3} designating the three populations of FWA). A set of long-established laboratory populations (called "NB", founded in 1990 from Adraga) was used as a control for all the experimental populations. NB populations were at their 90th, 136th and 181st laboratory generations at the time of foundation of the 1998, 2001 and 2005 collections, respectively.

All populations were maintained under the same laboratory environment with discrete generations of 28 days, reproduction close to peak fecundity, controlled temperature of 18ºC, with a 12-h L: 12-h D photoperiod. Flies were kept in vials, with controlled densities for both adult (around 50 individuals per vial) and larval stages (around 80 per vial). At each generation, emergences from the several vials of each replicate population were randomized using CO_2 anesthesia. Adult population sizes ranged between 600 and 1200 individuals. To study the evolutionary trajectories during laboratory adaptation, all experimental populations and the controls were periodically assayed for several phenotypic traits (see below).

Phenotypic Assays and Generations analyzed

For the phenotypic assays, mated pairs of flies were transferred daily to fresh medium and the number of eggs laid per female was counted during the first 12 days since emergence. After the fecundity assay, each pair of flies was transferred to a vial containing plain agar medium to measure starvation resistance (with deaths checked every 6 h). Five characters were analyzed: age of first reproduction (number of days between emergence and the day of first egg laying), early fecundity (total number of eggs laid during the first week), peak fecundity (total number of eggs laid between days 8 and 12), and female and male starvation resistance. Sample sizes ranged between 14 and 24 pairs per replicate population and assay. All assays involved synchronous analyses with NB populations.

Periodical phenotypic assays were performed starting at generation 3 or 4 up to generation 58-60. All generations assayed for the several populations are presented in Table S1. We analyze here both short-term - ~20 generations - and a longer-term period – between ~20 and ~60 generations, here designated "long-term" - of laboratory evolution of these populations. We also analyzed the entire evolutionary trajectory, spanning the complete data set. The short-term data was studied in Simões et al. (2008) for a larger number of populations, the five sets of populations referred to above and an extra set of populations in each of the 2005 locations (details in Simões et al. 2008). Moreover, for NW there were five replicate populations with data on short term, but here we only analyze three replicate populations, for both short and long-term, as only these have data for more advanced generations. Finally, we expand our analyses to include male starvation resistance data, which was not analyzed in Simões et al. (2008).

In order to calculate the initial or final state for each replicate population, we calculated the mean value of the 2 (or 3) first (or last) generations by choosing 15

random individual data points (with replacement) of each generation involved. The initial generations used where the following: 4,6 and 7 for AR and TW; 4 and 8 for NW; 3,6 and 10 for NARA and FWA. The final generations analyzed were: 48, 55 and 60 for AR and TW; 52, 53 and 58 for NW; 49 and 58 for NARA and FWA.

Statistical Methods

To estimate the evolutionary trajectories for each population, in each assayed generation, we used the differences between individual data and the mean of the same-numbered NB replicate population (assayed synchronously with experimental populations; e.g. AR_1-average NB_1, see (Simões et al. 2008). This was done to remove the effect of possible evolutionary changes not related to laboratory adaptation, and to minimize the effects of environmental heterogeneity between non-synchronous assays (see also Matos et al. 2002; Simões et al. 2007, 2008).

Linear and linear-log models were tested in both periods and in the whole evolutionary trajectory of the populations (around 60 generations) and we chose the model with the highest R^2. The linear model was chosen for each period and the linear-log for the whole trajectory, as it generally presented a better fit than the linear model (see Table S2).

Bootstrap Techniques

Variation in the slope of evolutionary response between sets of populations and periods was studied using bootstrap techniques as in Simões et al. (2008). Briefly, for each replicate population we estimated the intercept ($\hat{\beta}_0$), evolutionary slope ($\hat{\beta}_1$) and the residuals of each point (ε) using a simple linear regression. In each iteration of the bootstrap, a new vector of phenotypic data was created by resampling the residuals,

with replacement (ε^*) and employing the following formula to calculate a new phenotypic value for each data point used:

1) $$y^* = \hat{\beta}_0 + \hat{\beta}_1 x \text{ Generation} + \varepsilon^*$$

After this, a new slope (β_1^*) and intercept (β_0^*) were estimated through a linear regression. For the linear-log model the same analysis was applied using the natural logarithm of the generation. A total of 10000 slopes were generated for each replicate population. All analyses testing differences between slopes were done using these values.

To compare two sets of populations from the same location in different years, we calculated the mean of each set involved in the comparison (by randomly sampling one slope from each replicate population) and the difference between them (e.g. comparison Arrábida 2001 vs. Arrábida 2005: $((AR_1\beta_1^* + AR_2\beta_1^* + AR_3\beta_1^*)/3) - ((NARA_1\beta_1^* + NARA_2\beta_1^* + NARA_3\beta_1^*)/3)$. This process was repeated 10000 times. Statistical significance was assessed by estimating the fraction of these 10000 differences that were greater than zero. Two times this fraction or 1 minus two times this fraction (whichever is less) corresponds to the *P*-value. To compare differences between the two locations we used all 2001 and 2005 replicate populations from each location. NW data was not included as there were no corresponding populations founded in Arrábida in 1998. We calculated the location means using data of six replicate populations (e.g. $FWA_{1\text{-}3}$ and $TW_{1\text{-}3}$ for the Adraga slopes), again using random samples of slopes from each replicate population, as above. Differences between the short and long-term evolutionary response for each set of populations were also assessed (e.g. comparison of $TW_{1\text{-}3}$ short-term slopes vs $TW_{1\text{-}3}$ long-term slopes). We further analyzed whether differences between periods varied between populations founded from distinct years or locations. These comparisons followed the same

rationale as above (e.g. comparison short vs long-term for Arrábida 2001 vs Arrábida 2005: $((AR_1\beta_1{*}_{\mathbf{S}} + AR_2\beta_1{*}_{\mathbf{S}} + AR_3\beta_1{*}_{\mathbf{S}})/3 - (NARA_1\beta_1{*}_{\mathbf{S}} + NARA_2\beta_1{*}_{\mathbf{S}} + NARA_2\beta_1{*}_{\mathbf{S}})/3) - ((AR_1\beta_1{*}_{\mathbf{L}} + AR_2\beta_1{*}_{\mathbf{L}} + AR_3\beta_1{*}_{\mathbf{L}})/3 - (NARA_1\beta_1{*}_{\mathbf{L}} + NARA_2\beta_1{*}_{\mathbf{L}} + NARA_2\beta_1{*}_{\mathbf{L}})/3)$. This analysis was performed with 10000 random samples and tested as described above.

To test whether populations differed in the initial or final performance, 10000 comparisons between years and locations were assessed using the same rationale as above.

All analyses were performed using R version 3.3.1 (R Core Team 2018), package reshape2 (Wickham 2007) and visualization was done using ggplot2 package (Wickham 2009).

Results

Initial differences between populations

Populations were clearly differentiated from the control populations in the initial performance of fecundity traits, though less so for starvation resistance (Fig 1 and 2). NW populations performed significantly better than the other Sintra populations, both in age of first reproduction and early fecundity, whereas they performed worse for male starvation resistance (see Table S5 and Figs 1 and 2). Most populations from different years showed significant differences in the initial performance for peak fecundity and female starvation resistance. On the other hand, no significant differences were found between locations for any trait (see Table S5).

Short-term Evolutionary Dynamics

In general, fecundity-related traits, particularly early fecundity, show a clear evolutionary increase in performance during short-term evolution across populations, with a tendency to converge to control values, although at different rates (see Fig 1 and below; see also Fig S1, for data on the mean and variation of slopes of replicate populations). In contrast, patterns for starvation resistance are less consistent. In fact, male starvation resistance does not show a noticeable evolutionary response, although there is a suggestion of increased starvation across generations for all sets of populations except AR (see Fig 2 and Fig S1). The evolutionary response of female starvation resistance varies greatly among sets of populations with patterns of stasis (AR and TW), decreased (FWA and NARA) and increased performance (for NW). In spite these differences, the patterns are again of convergence to control values (see Fig 1, S1 and below).

When comparing the evolutionary response among populations, we observe significant differences of slopes between years (see Table 1). This variation is particularly evident for female starvation resistance, in agreement with our previous analysis (Simões et al. 2008). On the other hand, for male starvation resistance no significant variation in the evolutionary response was found. Significant differences between locations were only observed for peak fecundity (see Table 1).

Interestingly, of the nine comparisons showing significant differences between populations in short-term dynamics across all assayed traits (Table 1), six of these showed also significant variation in the initial performance (cf. Table 1 and Table S5). This concordance corresponded to a reduction of differences between populations through time for age of first reproduction and female starvation resistance. In contrast, for early fecundity, a higher initial performance of NW relative to TW or FWA was followed by faster improvement through time increasing the initial differences, leading at least to transient divergence between populations (Table 1 and Table S5, Fig 1 and 2).

Long-term Evolutionary Dynamics

In each set of populations there was a clear variation of evolutionary rates (slopes) between the short-term and the long-term period for age of first reproduction, early fecundity and female starvation resistance (Table S3). This corresponded to a general slowing down of the evolutionary response as populations tended to converge to the control values (see Figs 1, 2 and S1; Fig S2 shows the same indications in the evolutionary trajectories using all generations).

Differences in evolutionary dynamics between sets of populations were more evident in the long-term than in the short-term evolutionary response for several traits - age of first reproduction, early fecundity and male starvation resistance (see Table 1; Figs 1, 2

and S1). For early fecundity, besides a significant effect of location with higher evolutionary rate in Sintra populations, almost all comparisons showed significant effects of year, partly due to a more accentuated slowing down of the response in the 2005 populations. For male starvation resistance, significant differences in evolutionary response were obtained when comparing TW (stasis) with the other two Sintra sets of populations (both showing slight decreases over time). On the other hand, differences between populations in the evolutionary response of female starvation resistance decreased in this period with only two significant effects in five comparisons (in contrast with 4 in the short-term response) – see Table 1. These significant effects involved comparisons with NW, which showed a clear drop in performance during this later period (see Fig 2).

When comparing the variation in evolutionary rates of the different sets of populations between the two periods (short vs. long-term evolution), early fecundity and female starvation resistance showed the higher differences between populations, due to the above mentioned differential slowing down of response for early fecundity during long-term evolution and to the reported high variation in evolutionary rates seen in the short term evolution of female starvation resistance (Figs 1 and 2, Table 2). Importantly, for early fecundity populations with higher short-term evolutionary rates (NW and the two 2005 populations) were also those with a stronger slowing down in the long-term period (Fig 1), which is expected under convergent evolution (see below).

Final differences between populations

In more advanced generations, there was a loss of the initial differences between populations for several comparisons, as expected if full convergence occurs (see Table S5). This was observed between NW and TW for all fecundity traits, between NW and

FWA for age of first reproduction, and between the 2001 and 2005 populations for female starvation resistance (Table S5 and Figs. 1 and 2). Nevertheless, significant differences in final performance were also found for several comparisons (see Table S5). In some cases, differentiation was also present at the start. Three temporal patterns were observed taking into account initial, intermediate and final values (see Table S6): 1- continuous reduction of differences (NW versus TW for male starvation resistance); 2- increased differences through time (TW versus FWA for peak fecundity); 3- differentiation at the initial and final generations but with intermediate loss of differentiation (NW versus FWA for early fecundity and female starvation resistance). Finally, in other comparisons there was a significant differentiation between populations at the later stage of adaptation, not present at the start (Table S5). In this case two temporal patterns were observed (see Fig 1 and Table 1, S5 and S6): 1- higher differences at intermediate generations than at the end of the study due to a differential slowing down of the evolutionary rate (between the two sets of Arrábida populations for age of first reproduction and early fecundity); 2- higher absolute differences at the end than at intermediate or initial generations (Arrabida versus Sintra populations for early fecundity and NW versus FWA for peak fecundity).

Overall Evolutionary Dynamics

Evolutionary trajectories across the entire time span confirm a general deceleration of the evolutionary response through time, as populations evolved towards the control values (see Fig S2). This led to a general better fit of the overall evolutionary trajectory to a linear-log model relative to a linear one, particularly for fecundity related data (see Table S2). Differences between sets of populations in the overall evolutionary response corroborate the patterns described above for the two periods (see Table 1 and S4).

Discussion

Evolutionary convergence is a core expectation for adaptive evolution in a similar environment (Losos 2011; Stern 2013). With a smooth fitness landscape, that lacks multiple peaks, populations will tend to evolve to the same outcome (Wright 1931). Under such conditions, the outcome of evolution will be highly predictable, an issue of much debate at present (e.g. de Visser and Krug 2014; Orgogozo 2015). Experimental evolution is a great tool for testing whether adaptive evolution involves smooth or rugged landscapes, as it allows us to study the fate of populations initially differentiated when subject to similar selective pressures, especially whether they evolve towards similar or different fitness values (Matos et al. 2015; Orgogozo 2015; Fragata et al. 2018). Here we add to the previous Simões et al. (2008) study the analysis of c. 40 more generations of laboratory adaptation, in order to determine whether: 1) longer term evolution leads to similar outcomes as short term evolution; 2) populations will ultimately tend to converge or show more complex evolutionary patterns.

In this study we found a general pattern of convergent evolution, with clear changes in the evolutionary rates between the short (~20 generations) and longer-term (~60 generations) periods. We observed a slowing down of the evolutionary response through time for several traits as populations approached evolutionary equilibrium of long-established populations. Empirical evidence for deceleration of evolutionary rate has been observed in other experimental studies using both asexual (Wiser et al. 2013; Lenski et al. 2015) and sexual organisms (Gilligan and Frankham 2003; Simões et al. 2007).

In this study we also observed that the differences between short and longer-term dynamics were trait and population specific. Whereas differences in the early fecundity

response between sets of populations increased from short to long-term evolution, the inverse pattern was observed for female starvation resistance. The source of differences between populations also varied between traits. In the case of early fecundity, trajectory variation was due to a continuous increase in performance of the 2001 populations, even during long-term evolution, contrasting with the 1998 and 2005 populations, where the quicker short-term evolution was followed by an accentuated slowing down of the evolutionary response after generation 20. These differences are consistent with convergent evolution, as faster evolution in an earlier period is followed by a plateauing, while slower evolution corresponds to a steadier evolutionary rate throughout generations. Such contrasting evolutionary dynamics led to an interesting pattern: an intermediate phase of transient divergence was followed in the long-term by a partial convergence among evolving populations. In contrast, for female starvation resistance there were striking differences in the evolutionary trajectories during short-term evolution, with increase, decrease, or stasis contingent on the degree of initial differentiation from controls (see also Simões et al. 2008). For this trait, convergence was fast between all populations. These patterns were followed in general by a reduction of differences between evolutionary trajectories over the longer time period analyzed. The exception was the NW populations, which presented an initial positive trend, unique across populations (see also Matos et al. 2004), followed by a negative long-term trend. Nevertheless, despite the different underlying evolutionary dynamics, both early fecundity and female starvation resistance show a general pattern that suggests convergence in longer-term periods.

It is an inherent expectation of convergent evolution that there will be a negative association between initial state and subsequent evolutionary rates of populations adapting to a new environment (Simões et al. 2007). This expectation was confirmed

for *D. subobscura* populations with clear initial historical differentiation, founded from contrasting latitudes of the European cline (Fragata et al. 2014). In that study fast convergence was observed after only 14 generations in a common environment. In our study, evidence of such an association was only found for age of first reproduction and female starvation resistance for the short-term dynamics. Even so, for female starvation resistance the overall trend was not of convergence in the case of NW populations (see above). The lack of universal association in our study might be due to the smaller degree of initial differentiation of these populations, with greater sampling effects (Santos et al. 2012).

If full convergence occurs, an obvious corollary is that populations will not be differentiated as an outcome of evolution in a common environment. This expectation was not entirely met in our study, as several populations still remained differentiated for different traits even after sixty generations of evolution. In this context, several patterns emerged when comparing dynamics between different populations: (1) a phase of transient divergence, followed by partial convergence (for age of first reproduction and early fecundity); (2) continuous reduction of differences indicating partial convergence (for male starvation resistance); (3) continuous divergence between populations (for early and peak fecundity) or (4) transient convergence followed by later divergence (for early fecundity and female starvation resistance). Teotónio and his collaborators (Teotónio and Rose 2000; Teotónio et al. 2002) performed a reverse evolution study during 50 generations involving several genetically differentiated *Drosophila melanogaster* populations. They found that populations converged to ancestral values, but this trend was not general as it varied with the previous history and the trait studied. They concluded that populations converged to similar fitness values to a larger extent than other characters did. In contrast, in our study we did not see any clear relation

between the extent of convergence and how the traits analyzed were presumed to determine fitness. In fact, several populations remained differentiated for early fecundity, a trait that is under strong selection in our environment with clear and consistent improvement across many independent studies (Matos et al. 2002, 2004, Simões et al. 2007, 2008; Fragata et al. 2014). Given our interpretation of transient divergence and partial convergence in some of these populations, it is possible that they will ultimately converge over more generations of evolution.

We observed a high variation between short and longer-term dynamics in all our populations, which questions the ability to predict long-term from short-term evolution. This contrasts with the study of Burke et al. (2016), which suggests that short-term evolution is predictive of longer evolutionary time periods. In that study recently selected *D. melanogaster* populations converged to the trait values of other independently derived populations evolving in a similar selection regime for a longer time scale, regardless of the history of the populations studied. However, the shorter-term evolutionary responses of that study were sometimes more than 100 generations in duration. In general, the fact that our study showed such differentiated outcomes and complex evolutionary patterns highlights the importance of characterizing extended periods of experimental evolution and the possible pitfalls of predicting evolution from short-term adaptive patterns.

Conclusions

We here show that even after 60 generations of evolution in a common environment, *Drosophila subobscura* populations remain differentiated for several traits. Noticeably,

this was observed even for life-history traits that are clearly under selection in our regime. In this context, we found evidence for transient divergence, as a result of heterogeneity in evolutionary rates through time, occurring under a general scenario of convergence. Ultimately, we conclude that extrapolating from short-term evolutionary patterns to longer evolutionary periods might be risky, particularly if one is interested in predicting the outcomes of evolution.

Literature Cited

Barrett, R. D. H., and D. Schluter. 2008. Adaptation from standing genetic variation. Trends Ecol. Evol. 23:38–44.

Barton, N. H., and P. D. Keightley. 2002. Understanding quantitative genetic variation. Nat. Rev. Genet. 3:11–21.

Botero, C. A., F. J. Weissing, J. Wright, and D. R. Rubenstein. 2015. Evolutionary tipping points in the capacity to adapt to environmental change. Proc. Natl. Acad. Sci. USA 112:184–189.

Brito, R. A., L. S. Pletscher, and J. M. Cheverudi. 2005. the Evolution of Genetic Architecture. I. Diversification of Genetic Backgrounds By Genetic Drift. Evolution 59:2333–2342.

Burke, M., J. Dunham, P. Shahrestani, K. Thornton, M. Rose, and A. Long. 2010. Genome-wide analysis of a long-term evolution experiment with Drosophila. Nature 467:587–590.

Burke, M. K., T. T. Barter, L. G. Cabral, J. N. Kezos, M. A. Phillips, G. A. Rutledge, K. H. Phung, R. H. Chen, H. D. Nguyen, L. D. Mueller, and M. R. Rose. 2016. Rapid

divergence and convergence of life-history in experimentally evolved *Drosophila melanogaster*. Evolution 70:2085–2098.

Burke, M. K., and A. D. Long. 2012. What paths do advantageous alleles take during short-term evolutionary change? Mol. Ecol. 21:4913–4916.

Chippindale, A. K., J. A. Alipaz, H.-W. Chen, and M. R. Rose. 1997. Experimental Evolution of Accelerated Development in Drosophila. 1. Developmental Speed and Larval Survival. Evolution 51:1536–1551.

Cohan, F. M. 1984. Can Uniform Selection Retard Random Genetic Divergence Between Isolated Conspecific Populations? Evolution 38:495–504.

Cohan, F. M., and A. A. Hoffmann. 1986. Genetic divergence under uniform selection. II. Different responses to selection for knockdown resistance to ethanol among Drosophila melanogaster populations and their replicate lines. Genetics 114:145–164.

de Visser, J. A. G. M., and J. Krug. 2014. Empirical fitness landscapes and the predictability of evolution. Nat. Rev. Genet. 15:480–490.

Endler, J. A. 1986. Natural Selection in the Wild. Princeton University Press.

Fox, C. W., J. D. Wagner, S. Cline, F. A. Thomas, and F. J. Messina. 2011. Rapid Evolution of Lifespan in a Novel Environment: Sex-Specific Responses and Underlying Genetic Architecture. Evol. Biol. 38:182–196.

Fragata, I., P. Simões, M. Lopes-Cunha, M. Lima, B. Kellen, M. Bárbaro, J. Santos, M. R. Rose, M. Santos, and M. Matos. 2014. Laboratory selection quickly erases historical differentiation. PLoS One 9:e96227.

Fragata, I., P. Simões, M. Matos, E. Szathmáry, and M. Santos. 2018. Playing evolution

in the laboratory: From the first major evolutionary transition to global warming. Epl 122.

Franks, S. J., and A. A. Hoffmann. 2012. Genetics of Climate Change Adaptation. Annu. Rev. Genet 46:185–208.

Garland, T., and M. R. Rose (eds). 2009. Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments. University of California Press, Los Angeles, USA.

Gilligan, D. M., and R. Frankham. 2003. Dynamics of genetic adaption to captivity. Conserv. Genet. 4:189–197.

Good, B. H., and M. M. Desai. 2015. The impact of macroscopic epistasis on long-term evolutionary dynamics. Genetics 199:177–190.

Goodnight, C. J. 2015. Long-term selection experiments: Epistasis and the response to selection. Pp. 1–18 *in* J. H. Moore and S. M. Williams, eds. Epistasis: Methods and protocols. Springer Science+Business Media New York.

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

Joshi, A., R. B. Castillo, and L. D. Mueller. 2003. The contribution of ancestry, chance, and past and ongoing selection to adaptive evolution. J. Genet. 82:147–162.

Kawecki, T. J., and F. Mery. 2006. Genetically idiosyncratic responses of *Drosophila melanogaster* populations to selection for improved learning ability. J. Evol. Biol. 19:1265–1274.

Khan, A. I., D. M. Dinh, D. Schneider, R. E. Lenski, and T. F. Cooper. 2011. Negative

epistasis between beneficial mutations in an evolving bacterial population. Science 332:1193–1196.

Lachapelle, J., J. Reid, and N. Colegrave. 2015. Repeatability of adaptation in experimental populations of different sizes. Proc. R. Soc. B 282:20143033.

Lässig, M., V. Mustonen, and A. M. Walczak. 2017. Predicting evolution. Nat. Ecol. Evol. 1:0077.

Lenormand, T., D. Roze, and F. Rousset. 2009. Stochasticity in evolution. Trends Ecol. Evol. 24:157–165.

Lenski, R. E., M. J. Wiser, N. Ribeck, Z. D. Blount, R. Maddamsetti, A. R. Burmeister, E. J. Baird, and J. Bundy. 2015. Sustained fitness gains and variability in fitness trajectories in the long-term evolution experiment with *Escherichia coli*. Proc. R. Soc. B 282:20152292.

Long, A., G. Liti, A. Luptak, and O. Tenaillon. 2015. Elucidating the molecular architecture of adaptation via evolve and resequence experiments. Nat. Rev. Genet. 16:567–582.

Losos, J. B. 2011. Convergence, adaptation, and constraint. Evolution 65:1827–1840.

Matos, M., T. Avelar, and M. R. Rose. 2002. Variation in the rate of convergent evolution: adaptation to a laboratory environment in *Drosophila subobscura*. J. Evol. Biol. 15:673–682.

Matos, M., P. Simões, A. Duarte, C. Rego, T. Avelar, and M. R. Rose. 2004. Convergence to a novel environment: comparative method versus experimental evolution. Evolution 58:1503–1510.

Matos, M., P. Simões, M. A. Santos, S. G. Seabra, G. S. Faria, and F. Vala. 2015. History, chance and selection during phenotypic and genomic experimental evolution: replaying the tape of life at different levels. Front. Genet. 6:71.

Melnyk, A. H., and R. Kassen. 2011. Adaptive landscapes in evolving populations of *pseudomonas fluorescens*. Evolution 65:3048–3059.

Orgogozo, V. 2015. Replaying the tape of life in the twenty-first century. Interface Focus 5:20150057.

Orozco-Terwengel, P., M. Kapun, V. Nolte, R. Kofler, T. Flatt, and C. Schlötterer. 2012. Adaptation of Drosophila to a novel laboratory environment reveals temporally heterogeneous trajectories of selected alleles. Mol. Ecol. 21:4931–4941.

Paixão, T., and N. H. Barton. 2016. The effect of gene interactions on the long-term response to selection. Proc. Natl. Acad. Sci. 113:4422–4427.

Phillips, M. A., A. D. Long, Z. S. Greenspan, L. F. Greer, M. K. Burke, B. Villeponteau, K. C. Matsagas, C. L. Rizza, L. D. Mueller, and M. R. Rose. 2016. Genome-wide analysis of long-term evolutionary domestication in *Drosophila melanogaster*. Sci. Rep. 6:39281.

R Core Team. 2018. R: A Language and Environment for Statistical Computing.

Rose, M. R., H. B. Passananti, and M. Matos. 2004. Methuselah Flies: A Case Study in the Evolution of Aging. World Scientific.

Santos, J., M. Pascual, P. Simões, I. Fragata, M. Lima, B. Kellen, M. Santos, A. Marques, M. R. Rose, and M. Matos. 2012. From nature to the laboratory: the impact of founder effects on adaptation. J. Evol. Biol. 25:2607–2622.

Schoustra, S. E., D. Punzalan, R. Dali, H. D. Rundle, and R. Kassen. 2012. Multivariate Phenotypic Divergence Due to the Fixation of Beneficial Mutations in Experimentally Evolved Lineages of a Filamentous Fungus. PLoS One 7(11): e50305.

Seabra, S. G., I. Fragata, M. A. Antunes, G. S. Faria, M. A. Santos, V. C. Sousa, P. Simões, and M. Matos. 2018. Different genomic changes underlie adaptive evolution in populations of contrasting history. Mol. Biol. Evol. 35:549–563.

Simões, P., I. Fragata, S. Seabra, G. S. Faria, M. Santos, M. R. Rose, M. Santos, and M. Matos. 2017. Predictable phenotypic, but not karyotypic, evolution of historically differentiated populations. Sci. Rep. 7:913.

Simões, P., M. R. Rose, A. Duarte, R. Gonçalves, and M. Matos. 2007. Evolutionary domestication in *Drosophila subobscura*. J. Evol. Biol. 20:758–766.

Simões, P., J. Santos, I. Fragata, L. D. Mueller, M. R. Rose, and M. Matos. 2008. How repeatable is adaptive evolution? The role of geographical origin and founder effects in laboratory adaptation. Evolution 62:1817–1829.

Stern, D. L. 2013. The genetic causes of convergent evolution. Nat. Rev. Genet. 14:751–764.

Teotónio, H., I. M. Chelo, M. Bradić, M. R. Rose, and A. D. Long. 2009. Experimental evolution reveals natural selection on standing genetic variation. Nat. Genet. 41:251–257.

Teotónio, H., M. Matos, and M. Rose. 2002. Reverse evolution of fitness in *Drosophila melanogaster*. J. Evol. Biol. 15:608–617.

Teotónio, H., and M. R. Rose. 2000. Variation in the reversibility of evolution. Nature

408:463–466.

Travisano, M., J. A. Mongold, A. F. Bennett, and R. E. Lenski. 1995. Experimental Tests of the Roles of Adaptation, Chance, and History in Evolution. Science 267:87–90.

Whitlock, M. C., P. C. Phillips, F. B. Moore, and S. J. Tonsor. 1995. Multiple Fitness Peaks and Epistasis. Annu. Rev. Ecol. Syst. 26:601–629.

Wickham, H. 2009. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.

Wickham, H. 2007. Reshaping Data with the reshape Package. J. Stat. Softw. 21:1–20.

Wiser, M. J., N. Ribeck, R. E. Lenski, J. S. Littell, C. J. Muller, K. A. Dunne, A. V Vecchia, L. R. Leung, Y. Qian, Y. Zhang, J. L. Betancourt, and G. J. Mccabe. 2013. Long-Term Dynamics of Adaptation in Asexual Populations. Science 342:1364–1367.

Wright, S. 1931. Evolution in Mendelian Populations. Genetics 16:97–159.

Figure Legends

Figure 1 - Short and long-term evolutionary trajectories for fecundity related traits for the 5 sets of populations studied. A1R - Age of first reproduction (number of days), F17 - Early fecundity (number of eggs), F812 - Peak fecundity (number of eggs). Points represent mean values for each replicate at each generation. Dashed lines indicate short term period and full line indicates long-term period. Shaded area represent 95% confidence intervals estimated from the regression, using mean replicate values.

Figure 2 - Short and longer term evolutionary trajectories for female and male starvation resistance for the 5 sets of populations studied. RM - Male starvation resistance (in hours), RF - Female starvation resistance (in hours). Points represent mean value for each replicate at each generation. Dashed lines indicate short term period and full line indicates long-term period. Shaded area represent 95% confidence intervals estimated from the regression, using mean replicate values.

Figures

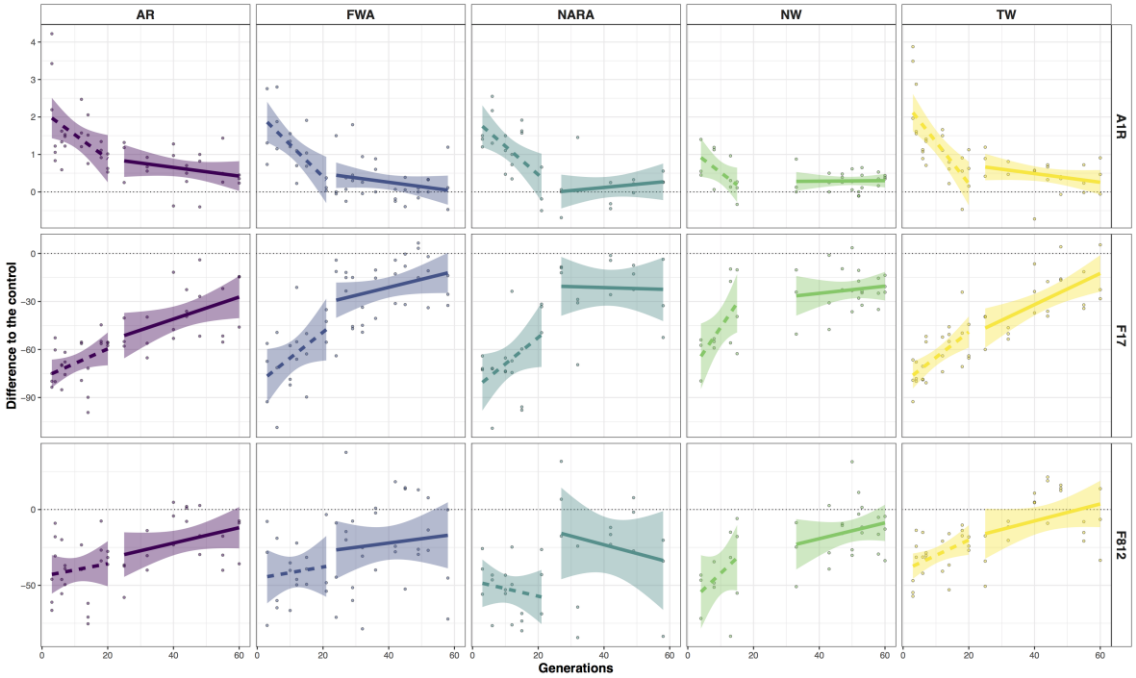

Figure 1

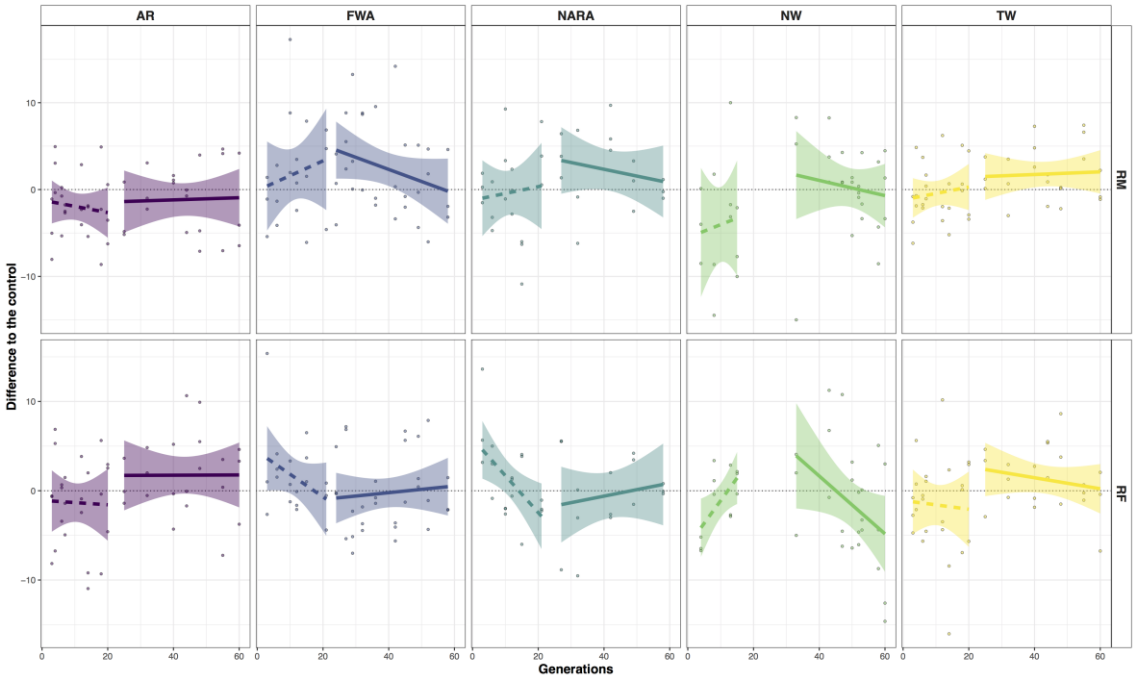
AR
FWA
NARA
NW
TW
RM
RF
Difference to the control
Generations
10
0
−10
0
20
40
60

Figure 2

Table 1 - Comparison of evolutionary rates between different years or locations for short or longer periods.

	Comparison	A1R	F17	F812	RF	RM
Short	Arr_0105	0.7534	0.1462	0.05	**0.0004**	0.3682
	Sin_9801	**0.0408**	**0.0066**	0.185	**0.0014**	0.6638
	Sin_9805	0.3634	**0.0292**	**0.0392**	**0**	0.9202
	Sin_0105	0.1888	0.6282	0.189	**0.0294**	0.4228
	Arr_Sin	0.063	0.2862	**0.0496**	0.4724	0.1286
Long	Arr_0105	**0.0084**	**0.002**	**0.0004**	0.2442	0.191
	Sin_9801	0.069	**0.0014**	0.1594	**4.00E-04**	**0.0478**
	Sin_9805	0.1188	0.0938	0.7108	**2.00E-04**	0.9632
	Sin_0105	0.894	**0.0396**	0.2322	0.0862	**0.0106**
	Arr_Sin	**0.0382**	**0.0036**	**0.0114**	0.3736	0.6238

Note: P-values were obtained by residual bootstraping of 10000 samples and estimated the fraction of these samples that were greater than 0 (see Material and Methods for more details). Significant results are indicated in bold

Table 2 - Comparison of short and long term evolutionary rates between years and locations.

Comparison	A1R	F17	F812	RF	RM
Arr_0105	0.2376	**0.0042**	0.7132	**0**	0.1444
Sin_9801	0.1402	**0**	0.069	**0**	0.2288
Sin_9805	0.6716	**0.006**	**0.048**	**0**	0.9166
Sin_0105	0.238	0.1714	0.5928	**0.0078**	0.0702
Arr_Sin	0.2628	0.6254	0.7414	0.2986	0.1144

Note: P-values were obtained by residual bootstraping of 10000 samples and estimated the fraction of these samples that were greater than 0 (see Material and Methods for more details). Significant results are indicated in bold