How phenotypic convergence arises in experimental evolution

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.

Data Accessibility Statement: Raw data is available at the Dryad Digital Repository https://doi.org/10.5061/dryad.k69gf52.

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

Accepted for publication in the journal Evolution. The final authenticated version is available online at: https://doi.org/10.1111/evo.13806

Citation: Simões, P. & Fragata, I., Santos, J., Santos, M. A., Santos, M., Rose, M. R. & Matos, M. 2019. How phenotypic convergence arises in experimental evolution. Evolution 73-9: 1839–1849 https://doi.org/10.1111/evo.13806.

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 that are 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 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 towards the character 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 (Botero et al. 2015; Franks and Hoffmann 2012). However, we are still unsure how predictable adaptation to novel environments is (Lachapelle et al. 2015; Lässig et al. 2017; Lenski et al. 2015; Orgogozo 2015; Wiser et al. 2013). 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; Rebolleda-Gómez and Travisano 2019). 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 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

88 (Teotónio and Rose 2000; Gilligan and Frankham 2003; Simões et al. 2007; Khan et al.
89 2011; Schoustra et al. 2012).

90 Long-term evolutionary dynamics have been mostly studied in microbial
91 experimental evolution systems rather than in sexual organisms, due to the shorter
92 generation time of the former. In the *E. coli* long-term evolution experiment performed
93 in Lenski's lab, recent evidence indicates a deceleration of the evolutionary rate over
94 50000 generations (Wiser et al. 2013; Lenski et al. 2015). Furthermore, and perhaps
95 surprisingly, heterogeneity in evolutionary trajectories is still present after so many
96 generations, in part due to differences in mutation rates (Lenski et al. 2015). Several
97 studies with sexual organisms, though involving fewer generations, have also observed
98 the slowing down of evolutionary responses to newly imposed selection regimes (e.g.
99 Gilligan and Frankham 2003; Rose et al. 2004; Simões et al. 2007, see below). The
100 expectation of a deceleration of laboratory evolutionary trajectories in sexual organisms
101 is sometimes justified in terms of temporal exhaustion of additive genetic variance,
102 although genomic scans in experimentally evolved *Drosophila* populations have found
103 only limited evidence of fixed alleles following selection (Burke et al. 2010; Burke and
104 Long 2012; Orozco-Terwengel et al. 2012; Long et al. 2015; Phillips et al. 2016; Seabra
105 et al. 2018). In a previous study by our team, we found evidence for a deceleration in
106 the evolutionary trajectory of fecundity in populations of *Drosophila subobscura*
107 evolving for more than 80 generations in the lab environment (Simões et al. 2007).
108 Teotónio and Rose (2000) also found this pattern of response in several *D.*
109 *melanogaster* lines undergoing reverse selection in their ancestral environment. Gilligan
110 and Frankham (2003) also reported a slowing down of the rate of adaptation to captivity
111 after 87 generations in the lab by comparing *Drosophila* populations in different stages
112 of adaptation. However, this pattern is not universal, as other experimental studies have

113 not found such deceleration of the evolutionary response, even after a higher number of
114 generations. One example of this is the work of Chippindale et al. (1997), who imposed
115 selection for accelerated development time in *D. melanogaster*. Nevertheless, studies of
116 long-term experimental evolution in sexual species are scarce and have not specifically
117 addressed the variation in evolutionary dynamics that might occur during evolution over
118 the short term, relative to longer evolutionary time periods.

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

132

133 We expect that, during short-term evolution, variation in the initial genetic
134 backgrounds will lead to disparate rates of adaptation to the new environment. Over the
135 longer term, as the evolutionary response decelerates, differences between populations
136 of contrasting initial genetic composition are likely to be reduced relative to those
137 observed during short-term evolution, particularly if populations are evolving towards
138 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 near Sintra (Portugal), whereas AR and NARA populations were collected from a pinewood in Arrábida (also from Portugal, some 50 Km from Sintra, 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., $FWA_{1\text{-}3}$ designating the three populations of FWA). A set of long-established laboratory populations (called "NB", founded in 1990 from Sintra) was used as a control for all the experimental populations. NB populations were at their 90^{th}, 136^{th} and 181^{st} 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. Census population sizes ranged between 600 and 1200 adults. To study the evolutionary trajectories during laboratory adaptation, all experimental populations and the controls were periodically assayed for several phenotypic traits (see below).

167

168

169 *Phenotypic Assays and Generations analyzed*

170 For the phenotypic assays, mated pairs of flies were transferred daily to fresh
171 medium and the number of eggs laid per female was counted during the first 12 days
172 since emergence. After the fecundity assay, each pair of flies was transferred to a vial
173 containing plain agar medium to measure starvation resistance (with deaths checked
174 every 6 h). Five characters were analyzed: age of first reproduction (number of days
175 between emergence and the day of first egg laying), early fecundity (total number of
176 eggs laid during the first week), peak fecundity (total number of eggs laid between days
177 8 and 12), and female and male starvation resistance. Sample sizes ranged between 14
178 and 24 pairs per replicate population and assay. All assays involved synchronous
179 analyses with NB populations.

180 Periodical phenotypic assays were performed starting at generation 3 or 4 up to
181 generation 58-60. All generations assayed for the several populations are presented in
182 Table S1. We analyze here both short-term - ~20 generations - and a longer-term period
183 - between ~20 and ~60 generations, here designated "long-term" - of laboratory
184 evolution of these populations. We also analyzed the entire evolutionary trajectory,
185 spanning the complete data set. The short-term data was studied in Simões et al. (2008)
186 for a larger number of populations, the five sets of populations referred to above and an
187 extra set of populations in each of the 2005 locations (details in Simões et al. 2008).
188 Moreover, for NW there were five replicate populations with data on short term, but
189 here we only analyze three replicate populations, for both short and long-term, as only
190 these have data for more advanced generations. Finally, we expand our analyses to
191 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 were 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. AR1-average NB1), see (Simões et al. 2008). This was done to remove the effect of possible temporal changes not related to laboratory adaptation such as trends due to environmental variation or to inadvertent evolutionary changes not intended in the study (e.g. due to slight changes of conditions in lab). This procedure also minimizes the effects of environmental heterogeneity between non-synchronous assays (see also Matos et al. 2002; Simões et al. 2007, 2008). Temporal performance of the control populations was generally quite stable across traits, allowing us to rule out undesirable sources of variation such as those due to further laboratory adaptation or inbreeding (see Fig S1).

Linear and linear-log models were tested for both periods separately and over the whole evolutionary trajectory of the populations (around 60 generations). Models were chosen according to their fit to the data based on R^2 values (see Table S2). For the separate analyses of short-term and long-term periods, we chose the linear over the linear-log model as a compromise across populations and periods, since the same model

had to be applied to allow for direct comparisons between periods (e.g. for the tests in Table 1 and 2). For the analysis of overall trajectories, the linear-log model was chosen, 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 Sintra 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.

When testing for differences between populations statistical significance is presented both with and without False Discovery Rate (FDR) correction for five tests (theorem 1.3 Benjamini and Yekutieli 2001). Marginally significant results after FDR correction will also be considered when the general reading justifies, i.e. if there are consistent patterns across populations. This is a compromise, as being too conservative also has drawbacks, given that the focus of this study is not on single tests but rather to analyze patterns across comparisons.

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

The experimental 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 S3 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 S3).

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 S2, 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 S2). 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 of these differences, the patterns are again of convergence to control values (see Fig 2, S2 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 eight comparisons showing significant (or at least marginally significant, after FDR correction) differences between populations in short-term dynamics across all assayed traits (Table 1), six of these showed also significant variation in initial performance (cf. Table 1 and Table S3). 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 S3, 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 S4). This corresponded to a general slowing down of the evolutionary response as populations tended to converge to the

316 control values (see Figs 1, 2 and S2; Fig S3 shows the same pattern in the evolutionary
317 trajectories using all generations).

318 Differences in evolutionary dynamics between sets of populations were more evident
319 in the long-term than in the short-term evolutionary response for several traits,
320 particularly for early fecundity (see Table 1; Figs 1, 2 and S2). For this trait, a
321 significant effect of location was due to a higher evolutionary rate in Sintra populations.
322 Also, several comparisons showed significant effects of year, due in part to a lower
323 slowing down of the response of the 2001 populations. On the other hand, differences
324 between populations in the evolutionary response of female starvation resistance
325 decreased in this period with only two significant effects in five comparisons– see Table
326 1. These significant effects involved comparisons with NW, which showed a clear drop
327 in performance during this later period (see Fig 2).

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

337

338 *Final differences between populations*

339 In more advanced generations, there was a loss of the initial differences between
340 populations for several comparisons, as expected if full convergence occurs (see Table

341 S3). This was observed between NW and TW for all fecundity traits, between NW and
342 FWA for age of first reproduction, and between the 2001 and 2005 populations for
343 female starvation resistance (Table S3 and Figs. 1 and 2). Nevertheless, significant
344 differences in final performance were also found for several comparisons (see Table
345 S3). In some cases, differentiation was also present at the start. Three temporal patterns
346 were observed taking into account initial, intermediate, and final values (see Table S5):
347 1- continuous reduction of differences (NW versus TW for male starvation resistance);
348 2- increased differences through time (TW versus FWA for peak fecundity); 3-
349 differentiation at the initial and final generations but with intermediate loss of
350 differentiation (NW versus FWA for early fecundity and female starvation resistance).
351 Finally, in other comparisons there was a significant (or at least marginally significant
352 after FDR correction) differentiation between populations at the later stage of
353 adaptation, not present at the start (Table S3). In this case two temporal patterns were
354 observed (see Fig 1 and Table 1, S3 and S5): 1- higher differences at the end than at
355 intermediate or initial generations (Arrábida versus Sintra populations for early
356 fecundity and NW versus FWA for peak fecundity); 2 - - higher differences at
357 intermediate generations than at the end of the study due to a differential slowing down
358 of the evolutionary rate (between the two sets of Arrábida populations for age of first
359 reproduction and early fecundity; in both cases differences are marginally significant
360 after FDR correction).
361
362 *Overall Evolutionary Dynamics*
363 Evolutionary trajectories across the entire time span confirm a general deceleration
364 of the evolutionary response through time, as populations evolved towards the control
365 values (see Fig S3). This led to a generally better fit of the overall evolutionary

366 trajectory to a linear-log model relative to a linear one, particularly for fecundity-related
367 data (see Table S2). Differences between sets of populations in the overall evolutionary
368 response were due to variable changes between short and long periods, leading to
369 pervasive contrasts, particularly for early fecundity (see Table 1 and S6).
370

371 ## Discussion

372 Evolutionary convergence is a core expectation for adaptive evolution in a
373 similar environment (Losos 2011; Stern 2013). With a smooth fitness landscape, that
374 lacks multiple peaks, populations will tend to evolve to the same outcome (Wright
375 1931). In such cases, the outcome of evolution will be predictable. The predictability of
376 evolution is an issue of much interest at present (e.g. de Visser and Krug 2014;
377 Orgogozo 2015). Experimental evolution is a great tool for testing whether adaptive
378 evolution involves smooth or rugged landscapes, as it allows us to study the fate of
379 populations initially differentiated when subject to similar selective pressures,
380 especially whether they evolve towards similar or different fitness values (Fragata et al.
381 2018; Matos et al. 2015; Orgogozo 2015; Rebolleda-Gómez and Travisano 2019). Here
382 we add to the previous Simões et al. (2008) study the analysis of c. 40 more generations
383 of laboratory adaptation, in order to determine whether: 1) longer-term evolution leads
384 to similar outcomes as short-term evolution; 2) populations will ultimately tend to
385 converge or show more complex evolutionary patterns.

386 In this study we found a general pattern of convergent evolution, with clear changes
387 in the evolutionary rates between the short-term (~20 generations) and longer-term (~60
388 generations) periods. We observed a slowing down of the evolutionary response
389 through time for several traits as populations approached the evolutionary equilibria of
390 long-established populations. Empirical evidence for deceleration of evolutionary rate

391 has been observed in other experimental studies using both asexual (Wiser et al. 2013;
392 Lenski et al. 2015) and sexual organisms (Gilligan and Frankham 2003; Simões et al.
393 2007).

394 We also observed that the differences between short-term and longer-term dynamics
395 were trait and population specific. Whereas differences in the early-fecundity response
396 between sets of populations increased from short- to long-term evolution, the inverse
397 pattern was observed for female starvation resistance. The source of differences
398 between populations also varied between traits. In the case of early fecundity, trajectory
399 variation was due to a continuous increase in performance of the 2001 populations, even
400 during long-term evolution, contrasting with the 1998 and 2005 populations, where
401 quicker short-term evolution was followed by a slowing of the evolutionary response
402 after generation 20. These differences are consistent with convergent evolution, as faster
403 evolution in an earlier period is followed by a plateauing, while slower evolution
404 corresponds to a steadier evolutionary rate throughout generations. Such contrasting
405 evolutionary dynamics led to an interesting pattern: an intermediate phase of transient
406 divergence was followed in the long-term by a partial convergence among evolving
407 populations. In contrast, for female starvation resistance there were striking differences
408 in the evolutionary trajectories during short-term evolution, with increase, decrease, or
409 stasis contingent on the degree of initial differentiation from controls (see also Simões
410 et al. 2008). For this trait, convergence was fast between all populations. These patterns
411 were followed in general by a reduction of differences between evolutionary trajectories
412 over the longer time period analyzed. The exception was the NW populations, which
413 presented an initial positive trend, unique across populations (see also Matos et al.
414 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 relative lack of such overall and rapid convergence 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 remained differentiated for some traits after sixty generations of evolution. In this context, several patterns emerged when comparing dynamics between different populations: (1) continuous reduction of differences indicating partial convergence (for male starvation resistance); (2) continuous divergence between populations (for early and peak fecundity); (3) transient divergence followed by partial convergence (for age of first reproduction and early fecundity) or (4) transient convergence followed by later divergence (for early fecundity and female starvation resistance). Teotónio and his collaborators (Teotónio et al. 2002; Teotónio and Rose 2000) performed a reverse evolution study during 50 generations

440 involving many genetically differentiated *Drosophila melanogaster* populations. They
441 found that populations converged to ancestral values, but this trend was not general as it
442 varied with the previous history and the trait studied. They concluded that populations
443 converged to similar fitness values to a larger extent than other characters did. In
444 contrast, in our study we did not see any clear relation between the extent of
445 convergence and how the traits analyzed were presumed to determine fitness. In fact,
446 several populations remained differentiated for early fecundity, a trait that is under
447 strong selection in our environment with clear and consistent improvement across many
448 independent studies (Fragata et al. 2014; Matos et al. 2002; Matos et al. 2004; Simões et
449 al. 2007; Simões et al. 2008). Given our interpretation of transient divergence and
450 partial convergence in some of these populations, it is possible that the time span of the
451 study was not sufficient to allow for full convergence in some cases, convergence that
452 might ultimately occur over more generations of evolution.
453 We observed considerable differences between short-term and longer-term dynamics
454 in all our populations, which raises questions about predicting long-term evolution from
455 short-term evolution. This contrasts with the study of Burke et al. (2016), which
456 suggests that short-term evolution is predictive of longer evolutionary time periods. In
457 that study recently selected *D. melanogaster* populations converged to the trait values of
458 other independently derived populations evolving in a similar selection regime for a
459 longer time scale, regardless of the evolutionary history of the populations studied.
460 However, different time scales were involved, as the shorter-term evolutionary
461 responses of that study were sometimes more than 100 generations in duration, with
462 long-term evolution approaching 1,000 generations. In general, the fact that our study
463 showed such differentiated outcomes and complex evolutionary patterns highlights the

464 importance of characterizing extended periods of experimental evolution and the
465 possible pitfalls of predicting evolution from short-term adaptive patterns.
466
467 **Conclusions**
468 We here showed that after 60 generations of evolution in a common environment,
469 *Drosophila subobscura* populations remain differentiated for several traits. Noticeably,
470 this was observed even for life-history traits that are clearly under selection in our lab.
471 In this context, we found evidence for transient divergence, as a result of heterogeneity
472 in evolutionary rates through time, occurring under a general scenario of convergence.
473 Ultimately, we conclude that extrapolating from short-term evolutionary patterns to
474 longer evolutionary periods might be risky, particularly if one is interested in predicting
475 the outcomes of evolution.
476
477 ## Author Contributions
478 PS and IF participated in the conception and design of the study, collected and analyzed
479 the data, and drafted the manuscript. JS participated in the conception and design of the
480 study, data collection, and revised critically the manuscript. MAS participated in data
481 collection and revised critically the manuscript. MS and MRR participated in the
482 conception and design of the study, as well as substantively revising the manuscript.
483 MM conceived and designed the study, coordinated the study, and helped drafting the
484 manuscript. All authors gave final approval for publication. PS and IF contributed
485 equally to this work.
486
487 ## Acknowledgments
488 The authors thank Laurence D. Mueller for the first version of the script that was later
489 adapted for the analyses of the present study. Two anonymous reviewers provided

helpful comments on an earlier draft. 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.

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.

Benjamini, Y., and D. Yekutieli. 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Stat. 29:1165–1188.

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 (N. Y). 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.

513 Nature 467:587–590.

514 Burke, M. K., T. T. Barter, L. G. Cabral, J. N. Kezos, M. A. Phillips, G. A. Rutledge, K.
515 H. Phung, R. H. Chen, H. D. Nguyen, L. D. Mueller, and M. R. Rose. 2016. Rapid
516 divergence and convergence of life-history in experimentally evolved Drosophila
517 melanogaster. Evolution (N. Y). 70:2085–2098.

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

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

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

525 Cohan, F. M., and A. A. Hoffmann. 1986. Genetic divergence under uniform selection.
526 II. Different responses to selection for knockdown resistance to ethanol among
527 Drosophila melanogaster populations and their replicate lines. Genetics 114:145–
528 164.

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

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

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

535 Fragata, I., P. Simões, M. Matos, E. Szathmáry, and M. Santos. 2018. Playing evolution
536 in the laboratory: From the first major evolutionary transition to global warming.
537 Epl 122.

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

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

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

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

547 Hansen, T. F. 2013. Why epistasis is important for selection and adaptation. Evolution
548 (N. Y). 67:3501–3511.

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

551 Khan, A. I., D. M. Dinh, D. Schneider, R. E. Lenski, and T. F. Cooper. 2011. Negative
552 epistasis between beneficial mutations in an evolving bacterial population. Science
553 (80-.). 332:1193–1196.

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

556 Lässig, M., V. Mustonen, and A. M. Walczak. 2017. Predicting evolution. Nat. Ecol.

557 Evol. 1:1–9. Macmillan Publishers Limited.

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

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

564 Long, A., G. Liti, A. Luptak, and O. Tenaillon. 2015. Elucidating the molecular
565 architecture of adaptation via evolve and resequence experiments. Nat. Rev. Genet.
566 16:567–582.

567 Losos, J. B. 2011. Convergence, adaptation, and constraint. Evolution (N. Y). 65:1827–
568 1840.

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

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

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

578 Melnyk, A. H., and R. Kassen. 2011. Adaptive landscapes in evolving populations of

579 pseudomonas fluorescens. Evolution (N. Y). 65:3048–3059.

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

582 Orozco-Terwengel, P., M. Kapun, V. Nolte, R. Kofler, T. Flatt, and C. Schlãtterer.
583 2012. Adaptation of Drosophila to a novel laboratory environment reveals
584 temporally heterogeneous trajectories of selected alleles. Mol. Ecol. 21:4931–
585 4941.

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

588 Phillips, M. A., A. D. Long, Z. S. Greenspan, L. F. Greer, M. K. Burke, B.
589 Villeponteau, K. C. Matsagas, C. L. Rizza, L. D. Mueller, and M. R. Rose. 2016.
590 Genome-wide analysis of long-term evolutionary domestication in Drosophila
591 melanogaster. Sci. Rep. 6:1–12. Nature Publishing Group.

592 Rebolleda-Gómez, M., and M. Travisano. 2019. Adaptation, chance, and history in
593 experimental evolution reversals to unicellularity. Evolution (N. Y). 73:73–83.

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

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

599 Schoustra, S. E., D. Punzalan, R. Dali, H. D. Rundle, and R. Kassen. 2012. Multivariate
600 Phenotypic Divergence Due to the Fixation of Beneficial Mutations in

601 Experimentally Evolved Lineages of a Filamentous Fungus. PLoS One 7:1–7.

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

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

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

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

613 Stern, D. L. 2013. The genetic causes of convergent evolution. Nat. Rev. Genet.
614 14:751–64. Nature Publishing Group.

615 Teotónio, H., I. M. Chelo, M. Bradić, M. R. Rose, and A. D. Long. 2009. Experimental
616 evolution reveals natural selection on standing genetic variation. Nat. Genet.
617 41:251–257.

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

620 Teotónio, H., and M. R. Rose. 2000. Variation in the reversibility of evolution. Nature
621 408:463–466.

622 Travisano, M., J. A. Mongold, A. F. Bennett, and R. E. Lenski. 1995. Experimental

623 Tests of the Roles of Adaptation, Chance, and History in Evolution. Science (80-.
624). 267:87–90.

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

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

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

630 Wiser, M. J., N. Ribeck, R. E. Lenski, J. S. Littell, C. J. Muller, K. A. Dunne, A. V
631 Vecchia, L. R. Leung, Y. Qian, Y. Zhang, J. L. Betancourt, and G. J. Mccabe.
632 2013. Long-Term Dynamics of Adaptation in Asexual Populations. Science (80-.).
633 342:1364–1367.

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

635 2018. R Core Team. R: A Language and Environment for Statistical Computing.

636

Tables

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

	Comparison	Age First Reprod	Early Fecundity	Peak Fecundity	Fem Starv Resist	Male Starv Resist
Short	Arrábida 2001 vs 2005	0.7534	0.1462	0.05	**0.0004** ***	0.3682
	Sintra 1998 vs 2001	**0.0408** m.s.	**0.0066** *	0.185	**0.0014** **	0.6638
	Sintra 1998 vs 2005	0.3634	**0.0292** m.s.	**0.0392** m.s.	**0** ***	0.9202
	Sintra 2001 vs 2005	0.1888	0.6282	0.189	**0.0294** m.s.	0.4228
	Arrábida vs Sintra	0.063	0.2862	**0.0496** n.s	0.4724	0.1286
Long	Arrábida 2001 vs 2005	**0.0084** **	**0.002** **	**0.0004** ***	0.2442	0.191
	Sintra 1998 vs 2001	0.069	**0.0014** **	0.1594	**0.0004** ***	**0.0478** n.s.
	Sintra 1998 vs 2005	0.1188	0.0938	0.7108	**0.0002** ***	0.9632
	Sintra 2001 vs 2005	0.894	**0.0396** m.s.	0.2322	0.0862	**0.0106** *
	Arrábida vs Sintra	**0.0382** m.s.	**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). When $p<0.05$ (indicated in bold) significance levels after FDR correction are also presented (in superscript): *** $p<0.00044$ ($\alpha=0.001$); ** $0.00044<p<0.0044$ ($\alpha=0.01$); * $0.0044<p<0.022$ ($\alpha=0.05$); m.s. $0.022<p<0.044$ ($\alpha=0.1$); n.s. $p>0.044$ ($\alpha=0.1$)

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

Comparison	Age First Reprod	Early Fecundity	Peak Fecundity	Fem Starv Resist	Male Starv Resist
Arrábida 2001 vs 2005	0.2376	**0.0042** **	0.7132	**0** ***	0.1444
Sintra 1998 vs 2001	0.1402	**0** ***	0.069	**0** ***	0.2288
Sintra 1998 vs 2005	0.6716	**0.006** *	**0.048** n.s.	**0** ***	0.9166
Sintra 2001 vs 2005	0.238	0.1714	0.5928	**0.0078** *	0.0702
Arrábida vs Sintra	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. When $p<0.05$ (indicated in bold) significance levels after FDR correction are also presented (in superscript): *** $p<0.00044$ ($\alpha=0.001$); ** $0.00044<p<0.0044$ ($\alpha=0.01$); * $0.0044<p<0.022$ ($\alpha=0.05$); m.s. $0.022<p<0.044$ ($\alpha=0.1$); n.s. $p>0.044$ ($\alpha=0.1$)

Figures

Figure 1 - Short and long-term evolutionary trajectories for fecundity related traits for the 5 sets of populations studied. Age of first reproduction (number of days), Early fecundity (number of eggs), Peak fecundity (number of eggs) are represented. 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 represents 95% confidence intervals estimated from the regression, using mean replicate population values.

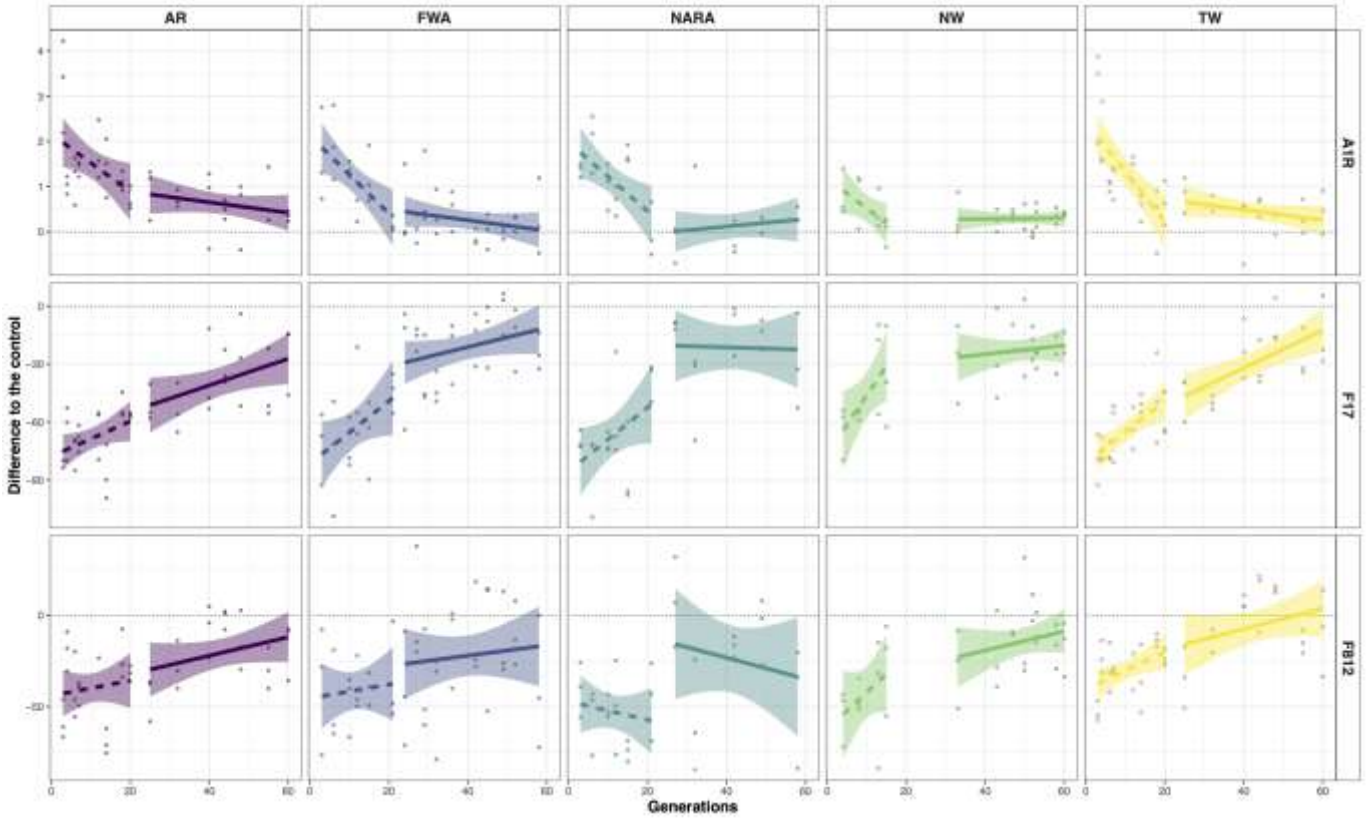

Figure 2 - Short and long-term evolutionary trajectories for female and male starvation resistance for the 5 sets of populations studied. Male starvation resistance (in hours) and Female starvation resistance (in hours) are represented. Points represent mean value for each replicate at each generation.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