Divergent evolution during an experimental adaptive radiation

R. Craig MacLean\textsuperscript{1*} and Graham Bell\textsuperscript{1,2}

\textsuperscript{1}Department of Biology, McGill University, 1205 Drive Penfield, Montreal, Quebec H3A 1B1, Canada
\textsuperscript{2}Redpath Museum, McGill University, 859 Sherbrooke Street West, Montreal, Quebec H3A 2K6, Canada

How repeatable a process is evolution? Comparative studies of multicellular eukaryotes and experimental studies with unicellular prokaryotes document the repeated evolution of adaptive phenotypes during similar adaptive radiations, suggesting that the outcome of adaptive radiation is broadly reproducible. The goal of this study was to test this hypothesis by using phenotypic traits to infer the genetic basis of adaptation to simple carbon-limited environments in an extensive adaptive radiation. We used a clone of the bacterium \textit{Pseudomonas fluorescens} to find two sets of experimental lines. The first set of lines was allowed to adapt to one of 23 novel environments for 1100 generations while the second set of lines was allowed to accumulate mutations by drift for 2000 generations. All lines were then assayed in the 95 environments provided by Biolog microplates to determine the phenotypic consequences of selection and drift. Replicate selection lines propagated in a common environment evolved similar adaptive components of their phenotype but showed extensive variation in non-adaptive phenotypic traits. This variation in non-adaptive phenotypic traits primarily resulted from the ascendance of different beneficial mutations in different lines. We argue that these results reconcile experimental and comparative approaches to studying adaptation by demonstrating that the convergent phenotypic evolution that occurs during adaptive radiation may be associated with radically different sets of beneficial mutations.

\textbf{Keywords:} \textit{Pseudomonas}; Biolog; selection experiment; adaptive radiation; contingency; correlated response

1. INTRODUCTION

If chance and contingency are the dominant features of adaptive evolution, repeated instances of adaptation to a common environment should never lead to the same evolutionary outcome (Gould 1989). Recent comparative studies of adaptive radiation over a wide range of temporal and phylogenetic scales have inferred that the same niche-specialist phenotypes, or ‘ecomorphs’, have evolved in independent lineages inhabiting similar environments in a wide range of taxa (Losos et al. 1998; Rüber et al. 1999; Bossuyt & Milinkovich 2000; Madsen et al. 2001). The repeated evolution of niche specialists is also a common feature of experimentally induced adaptive radiations in laboratory microcosms (Rainey & Travisano 1998; Treves et al. 1998; Travisano & Rainey 2000). Collectively, these studies suggest that strong directional selection is the dominant feature of adaptive radiation and that chance and contingency may play only a minor role in determining the eventual phenotypic outcome of adaptive radiation. If we accept that the phenotypic outcome of adaptive radiation is broadly reproducible, the repeatability of adaptive radiation at the genotypic level must be determined by the complexity of the relationship between genotypes and phenotypes. If a limited number of genotypes can produce a niche-specialist phenotype, then adaptive radiation should be reproducible at the genotypic level; if many genotypes can produce the same phenotype, then chance and contingency may play a dominant role in determining the genotypic diversity that evolves during adaptive radiation.

Although the genetic basis of diversification during adaptive radiations is generally not known, there is some evidence that many different genotypes can produce even the ‘simple’ specialist phenotypes that evolve in laboratory microcosms over a period of days (Spiers et al. 2002). This suggests that the exact path by which specialization evolves may be complicated and that each instance of adaptive radiation may produce a genetically unique assemblage of specialists. This assertion is supported by quantitative studies of artificial selection that have inferred that closely related populations often respond to common selective pressures through different genetic changes (Cohan 1984; Cohan & Hoffman 1989; Gromko et al. 1991).

The goal of the experiments presented in this paper was to determine the reproducibility of adaptive radiation at the genotypic level in a large-scale experimentally induced adaptive radiation. A clonal isolate of \textit{Pseudomonas fluorescens} was used to found 92 selection lines, which were propagated in one of 23 novel environments for \textit{ca}. 1100 generations. Our selection environments were provided by Biolog plates, commercially available 96 well microtitre plates in which 95 wells contain unique growth-limiting carbon sources. In this experiment, selection lines became adapted to their respective environments while the ancestral clone remained frozen in a non-evolving state (MacLean & Bell 2002). In a second experiment, the same clonal isolate was used to found 20 lines, which were propagated in the absence of effective selection for \textit{ca}. 2000 generations. We then assayed the evolved lines, the unselected lines and the ancestral clone in the 95 environments provided by the Biolog microplates.

*Author for correspondence (rmacle2@po-box.mcgill.ca).
In clonal populations of bacteria, selection acts on spontaneously arising mutations. Adaptation is a deterministic and repeatable process, so far as growth and reproduction will generally increase through selection. The genotypic basis of adaptation, however, and therefore the phenotypic traits associated with increased fitness, may not be the same in every case of adaptation to the same environment. In this study, we address three fundamental questions. (i) Do lines selected in the same environment evolve the same phenotype? If not, do they evolve similar phenotypes? (ii) Are phenotypic differences between replicate selection lines the result of selection or drift? (iii) Finally, do replicate lines usually adapt to a given environment through the fixation of the same beneficial mutations?

2. MATERIAL AND METHODS

(a) Selection experiment
A single clone of strain SBW25 of *P. fluorescens* was used to found the four replicate selection lines used in this experiment. The ancestral clone was kept frozen at −80 °C in 50% glycerol (w/v) during the experiment. Populations of *Pseudomonas* were selected on commercially available Biolog GN2 96 well microplates (Biolog, Hayward, CA, USA). Each well on a Biolog plate contains a single agar medium and a tetrazolium dye that acts as an indicator of oxidative carbon metabolism, which correlates with bacterial growth. Ninety-five wells on the plate contain unique carbon compounds and the remaining well is a control that lacks a carbon source. Daily transfers were carried out by using a 96 pin replicator to ‘print’ the populations grown on each selection plate to a fresh selection plate. This resulted in a dilution factor of ca. 2500 per transfer, representing log(2500)/log(2) ~ 11 population doublings. To sustain bacterial populations during the initial stages of adaptation to some of the substrates found on the Biolog plates, it is necessary to supplement the Biolog growth medium with an additional source of nutrients. We used King’s medium B (KB) (glycerol 10 g l⁻¹, proteose peptone 20 g l⁻¹, K₂HPO₄ 1.5 g l⁻¹, MgSO₄ 1.5 g l⁻¹) as an additional source of organic carbon and nitrogen for our cultures. Our cultures were assayed two replicates of each unselected line. The assay procedure was identical to that of the selection experiment, with the exception of an additional 1 day transfer between the reconditioning phase and the growth of cells for the assay.

(b) Assay
After the last transfer, all of our evolved populations (n = 95 × 4 = 380) were frozen at −80 °C in 50% glycerol (w/v). The four replicate selection lines from 23 randomly chosen selection environments and four replicates of the ancestral clone were then reconditioned in KB medium at 28 °C for 1 day. Reconditioned cells were incubated in fresh KB medium at 28 °C for 2 days to give stationary-phase cultures. Stationary-phase cultures were then diluted 150-fold and starved in M9 salts for at least 1.5 h prior to being assayed. We then added 150 µl of starved cells (~1 × 10⁶ cells) from each population to each well on Biolog GN2 plates. The absorbance of the microplate was read immediately at 660 nm using a narrow-beam plate reader (Biotek instruments, Winooski, VT, USA) and again after 24 h ± 5 min of stationary incubation at 28 °C. Scores for each well were calculated as the corrected increase in absorbance (the observed increase minus that of the blank control) over the 24 h period. The direct response to selection was calculated as the difference between the score of an evolved population and the mean score of the four replicates of the ancestral clone for the substrate on which selection was practised. Correlated responses to selection were calculated in the same manner for substrates other than that on which selection was practised.

(c) Mutation accumulation experiment
The clonal isolate that was used to found the selected lines was streaked on KB agar plates to found 20 unselected lines. These lines were propagated by selecting a single colony (the last colony on the streak) and streaking it on a fresh agar plate. Previous studies have shown that this method prevents effective selection and allows mutations, apart from those with highly deleterious effects, to accumulate by drift (Kibota & Lynch 1996; Korona 1999; Zeyl & DeVisser 2001). To avoid selection against small colonies containing deleterious mutations, we took the added precaution of examining the plates under a dissecting microscope. Plates were incubated at 28 °C, giving ca. 20 divisions per line per day. Transfers were carried out daily for 100 days or ca. 2000 bacterial generations. Because the differences in the Biolog profiles between the mutation-accumulation lines and the ancestral clone were predicted to be small, we assayed two replicates of each unselected line. The assay procedure was identical to that of the selection experiment, with the exception of an additional 1 day transfer between the reconditioning phase and the growth of cells for the assay.

(d) Phenotypic-similarity analysis
The phenotypic similarity between a pair of populations was estimated as the Euclidean distance in a multidimensional space in which each Biolog substrate represents one dimension. The Euclidean method estimates the distance between two populations, A and B, as
\[
d_{AB} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2},
\]
where *x* and *y* are the scores of populations *A* and *B*, respectively, on substrate *i*, and the summation is over all substrates. Using this method we computed two distance matrices: the first matrix included all possible pairwise distances between the 92 evolved populations and four replicates of the ancestral clone, and the second matrix included all the possible pairwise distances between the 20 unselected lines. The score of a single unselected line in a given environment was calculated as the mean of the two replicates of the line. Distance matrices can be represented graphically by means of unweighted pair group method with arithmetic mean (UPGMA) phenograms.

3. RESULTS

(a) Estimating the error variance in the Biolog assay
Assaying genetically identical replicates of the ancestral clone provides a means of estimating the error inherent in the Biolog assay. Note that these were ‘true’ replicates: the ancestral clone was independently reconditioned and
Divergent evolution and experimental adaptive radiation  R. C. MacLean and G. Bell 1647

Euclidean distance

Figure 1. UPGMA phenogram constructed from all possible pairwise distances between selection lines, as described in § 2d. Each tip represents a single replicate of a selection line or one of the four replicates of the ancestral clone. The alphanumeric code (e.g. A10) corresponds to the environment in which evolved lines were selected and the number following the dash is the replicate number. The four replicates of the ancestral clone are denoted as ‘base’. Tips that are marked with an asterisk represent replicate selection lines that form sister groups.

Assayed on four different Biolog plates. For each replicate we calculated a mean corrected absorbance across all 95 Biolog substrates. The mean scores for each replicate are very similar (grand mean = 0.53, s.d. = 0.00981), and the coefficient of variation among the mean scores of the replicates is only 1.85%. This shows that the overall mean Biolog scores of replicates of the ancestral clone are essentially identical, but does not exclude the possibility that different replicates produced different scores on different substrates. To address this possibility, we estimated variance components for the four replicates of the ancestral clone by fitting a model that included replicate, assay environment and a replicate × assay-environment interaction. We found that more than 99% of the variance is between assay environments and that both replicate and the interaction term account for less than 1% of the total variance in Biolog scores. We conclude that this assay technique is highly reproducible and that we can have great confidence in a single Biolog assay of a given line.

(b) Replicate selection lines do not converge on a single phenotype

Figure 1 is a UPGMA phenogram that displays phenotypic similarities between the selection lines and the ancestral clone. Each tip corresponds to either a replicate of the ancestral clone, or one of the (23 selection lines × four replicates per line = 92) evolved lines. As expected, the four replicates of the ancestral clone form a distinct phenotypic cluster, and the mean distance between replicates of the ancestral clone, which corresponds to the approximate error in distance measurement, is very small ($D_{\text{error}} = 0.57 \pm 0.03$ s.e.). If evolution is a strictly repeatable process, the four replicates of each selection line should evolve the same phenotype. When placed on a phenogram, these evolved lines should form clusters of four very similar tips in much the same way as the four replicates of the ancestral clone. Our phenogram (figure 1) shows that in no case do the four lines selected in a single environment form a cluster of four tips, although in several cases replicate selection lines form sister tips (marked with asterisks).

(c) Replicate selection lines form weak phenotypic clusters

Although figure 1 clearly demonstrates that replicate selection lines do not evolve the same phenotype, it does not allow us to rule out the possibility that lines selected in a common environment form clusters of similar tips. If this were the case, the mean Euclidean distance between lines selected in the same environment, $D_{\text{self}}$, should be less than the mean distance between these lines and lines selected in other environments, $D_{\text{other}}$ (figure 2). We found

that $D_{self}$ is less than $D_{other}$ in 18 out of 23 selection lines, and that the average difference between these distances is significant under a paired-sample $t$-test ($t_{2,22} = 2.53$, $p = 0.018$). However, both the absolute magnitude of this difference (mean $(D_{other} - D_{self}) = 0.207$) and the relative difference between these distances (mean $(D_{other}/D_{self}) = 1.07$) are small. It is also important to note that values of $D_{self}$ are always much larger (mean = 2.83, range = 1.66–4.73) than the estimate of $D_{error}$ that we obtained from replicate assays of the ancestral clone $(0.57 \pm 0.03$ s.e.). In summary, a quantitative analysis of our distance matrix allows us to draw two conclusions: (i) in no case do the four lines selected in a given environment converge on the same phenotype (i.e. $D_{cell} \gg D_{error}$) and (ii) lines selected in the same environment tend to be more similar to each other than they are to lines selected in other environments (i.e. $D_{cell} < D_{other}$), but the average magnitude of this difference is small.

**d) Unselected lines diverge at a slow rate**

To estimate the contribution of neutral drift to the divergence between replicate selection lines, we propagated 20 unselected lines for 2000 generations. Two lines of evidence demonstrate that mutations accumulated in a neutral manner in these lines as a consequence of drift (R. C. MacLean and G. Bell, unpublished data): (i) some lines fixed heritable and highly conspicuous colony-morphology mutations; and (ii) the mean growth rate of the 20 lines, an important component of fitness, showed a consistent pattern of decay over time. These lines have very similar Biolog profiles and the mean Euclidean distance between the 20 unselected lines after ca. 2000 generations of drift is $1.40 \pm 0.077$ s.e. Because the number of accumulated neutral mutations in an asexual haploid lineage is $U_t$, where $U$ is the genomic mutation rate and $t$ is the number of generations, we estimate the phenotypic divergence between our unselected lines after 1000 generations as $D_{drift} = (1.40 - D_{error})/2 = 0.41$.

**c) Phenotypic divergence between replicate selection lines is caused by correlated responses to selection**

The direct response to selection is a measure of the increase in growth as a result of adaptation: it is the adaptive component of the phenotype. Correlated responses to selection measure changes in non-adaptive components of the phenotype as a result of the pleiotropic effects of adaptive mutations: they can be used to infer the genetic basis of adaptation (Cohan & Hoffman 1989; Gromko et al. 1991). The divergence between replicate selection lines owing to direct and correlated responses to selection can be quantified by the following index (Vasi et al. 1994): $I_R = \langle Var(X) \rangle^{1/2} / X$, where $I_R$ is the divergence associated with a direct response ($I_{DR}$) or a correlated response ($I_{CR}$) to selection, $Var(X)$ is the between-line variance of the response and $X$ is the mean response. The values of $I_{DR}$ are log-normally distributed and in most selection lines (15 out of 23) the value of $I_{DR}$ is less than one (figure 3). This implies that the increase in standard deviation among replicate selection lines for the trait on which they are selected is generally less than the increase in mean performance. The values of $I_{CR}$ tend to be much larger than the values of $I_{DR}$, and the values of $I_R$ expressed in a given selection line are on average 2.74 times larger for each correlated response to selection than for the direct response to selection (figure 3). Each evolved line has a single direct response to selection and 94 correlated responses, so that the divergence between replicate selection lines is almost entirely attributable to correlated responses to selection. Because phenotypic differences between lines imply genetic differences in this system, we interpret this to mean that replicate lines adapt to their common environment to a similar degree, as indicated by the low values of $I_R$ for the direct response to selection, but that the genetic basis of adaptation varies widely between lines, as indicated by the high values of $I_{CR}$ for the correlated response to selection.

Figure 2. Phenotypic clustering of replicate selection lines. Plotted points show the mean distance between replicate selection lines ($D_{cell}$) against the mean distances from these lines to the 88 lines selected in other environments ($D_{other}$). The dashed line is a plot of the equation $y = x$. The dot-dash line is a plot of mean $D_{cell}$ plus the mean distance between replicate measures of the ancestral clone ($D_{error}$).

Figure 3. Standardized variance of direct and correlated responses to selection. Plotted points show the standardized variances of the direct response to selection for replicates of a given line ($I_{DR}$) against the means of the 94 standardized variances of the correlated response to selection (mean $I_{CR}$) ± 1 s.e. One outlier line with $I_{DR} = 52.8$ is not shown on this plot and was excluded from our analysis of $I$. The dashed line is a plot of the equation $y = x$. 

4. DISCUSSION

During the course of the selection experiment, populations of Pseudomonas became adapted to a range of novel habitats provided by Biolog plates (MacLean & Bell 2002). Replicate selection lines evolved very similar traits that are well correlated with fitness, a result that is commonly found when clonal populations of bacteria are allowed to adapt to novel environments (Lenski & Travisano 1994; Korona 1996; Travisano 1997). In agreement with the microevolutionary studies of Lenski & Travisano (1995b), we found that replicate selection lines evolved very different non-adaptive phenotypic traits, such that replicate lines never evolved the same phenotype (see also Yin 1993; Travisano et al. 1995a; Bennett & Lenski 1996; Riley et al. 2001).

Our assay of adaptation and divergence evaluates the overall performance of entire selection lines. The simplest possible interpretation of our results is that each population is dominated by a single genotype, and that adaptive divergence is therefore caused by the emergence of a single dominant genotype in each selection line. This is not necessarily the case, however: we have no information on within-line genetic variation, and it is possible that different consortia of genotypes evolved within some of our lines (Levin 1988; Elena & Lenski 1997; Treves et al. 1998; Rozen & Lenski 2000). The divergence that we have documented would then involve the emergence of different communities in each replicate of a selection line. This interesting possibility can be investigated only by isolating genotypes from each line and testing each genotype on Biolog plates in a more detailed and extensive experiment.

How can we explain the extensive phenotypic divergence between our replicate selection lines? Divergence between replicate selection lines can be caused by two distinct processes: selection and drift. Selection will cause divergence between replicate selection lines if different beneficial mutations arise by chance in different lines (Johnson et al. 1995), or if epistatic interactions occur among beneficial mutations (Mani & Clarke 1990; Korona 1996). The role of drift in between-replicate divergence is somewhat more subtle. In the first place, drift can contribute to divergence through the chance fixation of deleterious mutations in selection lines with small effective population sizes. Although effective population size varied widely between selection lines, the effective population size of our lines was approximately the same as the transfer population size. Given that our replicator transfers ca. 0.1 µl of fluid and that the lowest stationary density of our lines was \( \approx 1 \times 10^7 \) CFU ml\(^{-1} \), we can estimate the minimal effective population size of our lines as \( (1 \times 10^7 \text{ CFU ml}^{-1}) \times (1 \times 10^{-3} \text{ ml}) \approx 1000 \) individuals. This suggests that \( N_e \) (the effective population size) should have been large enough to prevent the chance fixation of deleterious mutations by drift.

Drift can also contribute to divergence when conditionally neutral mutations, which alter the phenotype of a line but not its fitness, accumulate by drift (Lynch 1994). A number of lines of evidence indicate that conditionally neutral mutations did accumulate in the selection lines (MacLean & Bell 2002). To estimate the contribution of neutral drift to divergence between replicate selection lines, we allowed mutations to accumulate in 20 unselected lines that were bottlenecked every 20 generations for 2000 generations. Repeated bottlenecking removes effective selection and allows mutations to accumulate in a neutral manner at a rate \( U \), the genomic mutation rate (Kibota & Lynch 1996; Zeyl & DeVisser 2001). This method will provide an accurate estimate of the rate of neutral divergence in selection lines if selection lines accumulate conditionally neutral mutations at a rate \( U \), as will be the case in a population at equilibrium. Our selection lines were founded from an ancestor that was poorly adapted to life on Biolog plates, and the dramatic response to selection that we observed during the course of the selection experiment demonstrates that our selection lines were not at equilibrium. None the less, the rate of divergence among the unselected lines provides a coarse estimate of the rate of neutral divergence in the selection lines.

The divergence between selection lines can therefore be partitioned as follows. In the first place, measurement error caused a small amount of apparent divergence between any two lines (\( D_{\text{error}} = 0.57 \)). We estimate that neutral drift made a small contribution to between-line divergence, in excess of the divergence attributable to measurement error (\( D_{\text{drift}} = 0.41 \)). Replicate selection lines diverged considerably as a result of adaptation to a common environment (\( D_{\text{adt}} - D_{\text{drift}} - D_{\text{error}} = 1.85 \)). Surprisingly, the divergence attributable to adaptation to completely different substrates expressed by different selection lines was only marginally greater than the divergence attributable to adaptation to a common environment expressed by replicate selection lines (\( D_{\text{other}} - D_{\text{adt}} - D_{\text{drift}} - D_{\text{error}} = 0.21 \)).

Previous studies of adaptive radiation in laboratory microcosms have argued that replicate selection lines tend to form distinct genotypic (Bull et al. 1997) and phenotypic (Dykhuizen & Hartl 1981; Travisano 1997; Cooper & Scott 2001; Riley et al. 2001) clusters. Our system has allowed us to expand the scope of such experiments greatly through the use of a large number of selection and assay environments. The most remarkable feature of our results is that replicate lines do not show a strong tendency to form phenotypic clusters, even though replicate lines evolved similar adaptive components of their phenotypes. The most likely explanation for this result is that different beneficial mutations became common in replicate selection lines.

Comparative studies show that the same niche-specialist phenotypes have evolved independently in adaptive radiations of Anolis lizards (Losos et al. 1998), placental mammals (Madsen et al. 2001), ranid frogs (Bossuyt & Mîlînkovitch 2000) and cichlid fishes (Rüber et al. 1999). Our experiment demonstrates that the convergent adaptive phenotypic evolution shown by isolated populations during an adaptive radiation is often associated with the acquisition of quite different sets of adaptive mutations. This suggests that chance and contingency are likely to be important driving forces in determining the genotypic outcome of adaptive radiation.

We thank D. Schoen, O. Kaltz and three anonymous reviewers for their helpful feedback, and J. Kwan for technical assistance. This work was funded by grants from the Natural Sciences and Engineering Research Council of Canada to G.B. and R.C.M.
REFERENCES


