MBE Advance Access published March 27, 2015

The Effect of Selection Environment on the Probability of Parallel Evolution

Susan F. Bailey,*[‡,1] Nicolas Rodrigue,[1,2] and Rees Kassen[1]

[1]Biology Department and Center for Advanced Research in Environmental Genomics, University of Ottawa, Ottawa, ON, Canada

[2]Department of Biology, Carleton University, Ottawa, ON, Canada

[‡]Present address: Bioinformatics Research Centre, Aarhus University, Aarhus C, Denmark

***Corresponding author:** E-mail: susan.f.bailey@gmail.com.

Associate editor: Miriam Barlow

Abstract

Across the great diversity of life, there are many compelling examples of parallel and convergent evolution—similar evolutionary changes arising in independently evolving populations. Parallel evolution is often taken to be strong evidence of adaptation occurring in populations that are highly constrained in their genetic variation. Theoretical models suggest a few potential factors driving the probability of parallel evolution, but experimental tests are needed. In this study, we quantify the degree of parallel evolution in 15 replicate populations of *Pseudomonas fluorescens* evolved in five different environments that varied in resource type and arrangement. We identified repeat changes across multiple levels of biological organization from phenotype, to gene, to nucleotide, and tested the impact of 1) selection environment, 2) the degree of adaptation, and 3) the degree of heterogeneity in the environment on the degree of parallel evolution at the gene-level. We saw, as expected, that parallel evolution occurred more often between populations evolved in the same environment; however, the extent of parallel evolution varied widely. The degree of adaptation did not significantly explain variation in the extent of parallelism in our system but number of available beneficial mutations correlated negatively with parallel evolution. In addition, degree of parallel evolution was significantly higher in populations evolved in a spatially structured, multiresource environment, suggesting that environmental heterogeneity may be an important factor constraining adaptation. Overall, our results stress the importance of environment in driving parallel evolutionary changes and point to a number of avenues for future work for understanding when evolution is predictable.

Key words: **parallel evolution, experimental evolution, *Pseudomonas fluorescens*, selection environment, environmental heterogeneity.**

Introduction

How repeatable is adaptive evolution? If, as Stephen Jay Gould famously put it, we were to rewind the tape of life and play it again, would the outcome be the same? Gould's answer was no. Evolution, he argued, was the result of too many stochastic events with highly contingent effects to be predictable at all but the crudest scales. A retrospective study of the history of life could make sense of how evolution had occurred and be used to guide thinking for the future, Gould held, but the genetic and mechanistic details underlying adaptive evolution would be nearly impossible to predict.

Despite Gould's skepticism, or perhaps because of it, evolutionary biologists take special note when evolution has been repeated (e.g., Zhang and Kumar 1997; Odeen and Hastad 2003; Harrison et al. 2005; Liu et al. 2010). Repeated evolution of the same phenotype independently in multiple populations, or parallel evolution, seems so unlikely to happen by chance alone that, when it occurs, we usually take it to be strong evidence of adaptation, even in the face of alternative explanations such as chance. Another possibility is mutation rate variation across a genome, which could lead to high levels of parallel evolution if more mutable regions more often contribute variants that are eventually selected during adaptive evolution. These alternatives notwithstanding, compelling examples of repeated evolution probably reflect an unusual set of circumstances: Strong selection acting on genetic systems that are highly constrained in the kind of genetic variation they produce. Our ability to predict when this sort of situation arises or, more generally, the probability of parallel evolution, remains rudimentary at best (but see Conte et al. 2012 for a meta-analysis of the available data from natural populations).

Some progress toward making predictions regarding the probability of parallel evolution can be made, however. As a first step it is important to recognize that the likelihood of observing parallel evolution should depend on the level of the biological hierarchy we are examining. The most repeatable phenotype to evolve during adaptation is fitness itself: Provided there is sufficient genetic variation in the traits associated with fitness and selection is modestly strong, fitness must increase. Thus, the probability of parallel evolution of fitness among multiple independent populations founded from the same genotype or population is necessarily high. However the probability of parallel evolution should decrease when we examine the traits that underlie fitness because there are probably many different combinations of phenotypes, and even more combinations of genotypes, that can generate high fitness. Repeatability should thus be highest at

© The Author 2015. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015

Article

the level of fitness, lower at the level of phenotypes, lower still at the level of the genes themselves, and lowest at the level of individual mutations or nucleotides.

A second step involves making more quantitative predictions about the factors governing the probability of parallel evolution. Genetic models of adaptation like Fisher's geometric model (Fisher 1930) and Gillespie's mutational landscape model (Gillespie 1983, 1991) can be used to make such predictions at the locus or nucleotide level (Orr 2005; Chevin et al. 2010). The central idea is that, among the many possible genetic routes to adaptation, the ones most often followed are those where selection is strong and genetic variation for fitness is abundant. Any factor that modulates the strength of selection or the amount of genetic variation for fitness in that particular environment can impact the probability of parallel evolution.

To see this more clearly, and to articulate more precisely the predictions that can be made, imagine a situation where a large population composed of a single ancestral genotype is situated some distance from a fitness optimum in a single environment. Genetic variation in fitness arises by mutation, and selection leads to a series of successive fixation events resulting in an adaptive walk toward the optimum. Beneficial mutations are more abundant the farther the starting genotype is from the fitness optimum (e.g., Schoustra et al. 2009), meaning that there are likely to be many routes to adaptation available to a mal-adapted population. The probability of parallel evolution should therefore increase as a population adapts because the number of possible adaptive routes available declines as a population approaches a fitness optimum. Orr (2005) has calculated the probability of parallel evolution in the mutational landscape model precisely: Under natural selection, it is $2/(n + 1)$, where n is the number of beneficial mutations available to a given sequence. The available evidence on this relationship is mixed: One study suggests the expected negative relationship between distance to the optimum and the degree of parallelism (Gresham et al. 2008), whereas another observed no evidence of a relationship (Kryazhimskiy et al. 2014).

Now, imagine the population inhabits a spatially heterogeneous environment composed of many different patches with different conditions of growth such that selection is divergent. Strong divergent selection and weak dispersal among patches are expected to lead to the evolution of niche specialization and the maintenance of diversity because beneficial mutations have environment-specific effects, improving fitness in the environment in which they are selected but not in alternative environments. Independently evolved populations from the same selection environment are thus expected to display more parallelism than those from different environments. Indeed, a pattern of higher parallelism within, compared with between, selection environments has been suggested in previous work (Gresham et al. 2008; Spor et al. 2014); however, a formal characterization and comparison of these patterns at the level of genetic changes is lacking.

On the other hand, in a heterogeneous environment where divergent selection is weak relative to dispersal, a single generalist genotype that has high fitness across all patches is expected to evolve because only those beneficial mutations that improve fitness across all patches are likely to be fixed. It seems reasonable to expect that there should be fewer mutations that improve fitness across many environments simultaneously than there are mutations that improve fitness in any single environment, because "generalist" mutations can be selected in both heterogeneous and single environments whereas mutations with environment-specific effects can only be selected in those environments where they actually improve fitness. If so, then we expect the probability of parallel evolution to be higher in populations living in heterogeneous environments where dispersal rate is high enough to drive the evolution of generalists, compared with populations living in more homogeneous environments that tend to drive the evolution of specialists.

Efforts to test these predictions empirically remain limited, in part because many of the theoretical predictions rely on quantities that are difficult to estimate (e.g., distance to the optimum, environmental heterogeneity, as discussed above). For example, Chevin et al. (2010) showed that the degree of pleiotropy can have an important impact on the degree of parallel evolution. Unfortunately this quantity is difficult to estimate in an evolutionary meaningful way (e.g., Rodríguez-Verdugo et al. 2014); however, it is notable that Spor et al. (2014) found that the only gene in which they detected repeat mutations across replicate experimentally evolved yeast populations was one that previous work suggests is highly pleiotropic. Far more common are studies that investigate the genetic basis of striking examples of parallel phenotypic evolution in naturally replicated populations (e.g., Wittkopp et al. 2003; Colosimo et al. 2005; Steiner et al. 2009) or the molecular details underlying a number of traits that appear repeatedly across diverse groups of species (e.g., Yokoyama and Radlwimmer 2001; Copley 2004; Fernald 2006; Castoe et al. 2009). There has been some attempt to quantify the probability of parallel evolution at the locus and nucleotide level in replicated selection experiments with microbes (e.g., Tenaillon et al. 2012; Herron and Doebeli 2013; Kryazhimskiy et al. 2014; Spor et al. 2014). Despite occasionally striking examples of parallel evolution at some loci, on average parallel evolution at the gene level is fairly low, typically in the range of 0–20% or so, and even lower at the nucleotide level (reviewed in Kassen 2014). In phage the situation is quite different, however: Rates of parallel evolution at the nucleotide level are approximately 50% (Wichman and Brown 2010), likely due to the higher mutation rate and more compact genomes of phage relative to bacteria and yeast.

Here, we explore the factors governing the probability of parallel evolution in experimentally evolved populations of bacteria. In a previous paper, we investigated the effect of selection environment on the rate and extent of evolutionary changes in fitness and phenotype in these populations (Bailey and Kassen 2012). We observed clear differences between populations in their evolutionary trajectories, prompting us to further explore the generation-1000 populations using whole-genome sequencing in an

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015

attempt to link the evolved phenotypic differences to underlying genomic differences. More specifically, we examine genomic changes in 15 independently evolved populations of the bacteria *Pseudomonas fluorescens*, all initiated from the same starting genotype, and selected in five different selection environments for approximately 1,000 generations. The selection environments varied in the particular resources available for growth, the number of those resources, and the arrangement of those resources in space. The design of this experiment thus allows us to test the predictions outlined above regarding the change in the probability of parallel evolution as a function of 1) selection environment, 2) the degree of adaptation, and 3) the degree of heterogeneity in the environment. We also examine parallel evolution across multiple levels of biological organization including, in descending order of expected parallelism, fitness, resource phenotype, metabolic phenotype, mutation-bearing genes and nucleotide changes and confirm the expectation that degree of parallelism decreases as one moves down this biological hierarchy from fitness to nucleotide. Taken together, our data thus provide valuable insight into the factors governing the probability of parallel evolution at a range of levels, from fitness down to the nucleotide.

Results

Fifteen populations of *P. fluorescens* SBW25 initiated from the same ancestral genotype were independently evolved for approximately 1,000 generations in five different selection environments, for a total of three replicate populations per selection environment. The three single-resource selection environments contained one type of sugar resource: Either glucose, mannose, or xylose, whereas the two multiresource environments contained all three sugar resources. In the multiresource environments the three sugars were provided in a single, well-mixed patch (MIX), or in three spatially separate patches each containing one of the three sugar types, with random dispersal between patches occurring every 24 h (SPAT). The ancestral *P. fluorescens* SBW25 varied across these five selection environments in how well it was initially adapted—for example, SBW25 reaches cell densities over 10-fold higher in glucose and mannose compared with xylose when grown for 24 h at a 1 M sugar concentration (see Materials and Methods). We refer readers to our previous work for further details regarding the experimental set-up and evolved fitness and phenotype dynamics (Bailey and Kassen 2012). At the end of the experiment each evolved population, as well as the ancestral genotype, was genome sequenced and the resulting sequences compared with the complete *P. fluorescens* SBW25 genome (NC 012660.1; 6,722,540 bp; 6,584 genes). All detected mutations are detailed in supplementary table S1, Supplementary Material online.

Genome Sequencing

Mutational changes were detected in all populations after 1,000 generations of selection, with the number of changes per population ranging from 1 to 6 (mean = 3.4667) for a total of 52 mutational changes across 15 independent populations. The detected mutations consisted of 24 single nucleotide polymorphisms (SNPs), 8 insertions, and 20 deletions. We expect that most, if not all, of these arising mutations had beneficial fitness effects. Our reasoning is that the large population sizes and low mutation rate in our experimental system imply that drift and hitchhiking, respectively, are unlikely to have played a significant role in the evolutionary dynamics. Instead, the evolutionary dynamics should be mainly dictated by selection. Indeed, the majority of mutations detected in our experiment were nonsynonymous (50 of 52), upon which selection is most able to exert influence. These nonsynonymous mutations are thus likely to confer beneficial fitness effects. Interestingly, the two detected synonymous mutations also turned out to have significant beneficial fitness effects, a topic explored in a previous paper (Bailey et al. 2014).

We detected substantial amounts of within-population genetic polymorphism—26 of the 52 detected mutations were present at frequencies less than 1; however, there was no significant difference in the frequency of detected polymorphisms across selection environments (analysis of variance [ANOVA], $F_{4,10} = 0.6557$, $P = 0.6362$). Most insertions and deletions were small and expected to result in loss of function to a single gene, with three notable exceptions—a 2,864-bp deletion (spanning three genes) in a xylose-selected population, a 3,175-bp deletion (spanning two genes) present in two independent MIX-selected populations, and a 35,110-bp deletion (spanning 32 genes) in a glucose-selected population. We did not detect any gene-scale duplication events. The number of mutations per population varied by selection environment (table 1; resampled

Table 1. Number of Genic Mutations Detected in Each Evolved Population.

Selection env.	Rep.	Non-mot.	Mot	All Mutations	Proportion of Mutations that Are Polymorphic
G	1	1	2	3	0.67
	2	2	1	3	0.67
	3	1	–	1	0
M	1	1	1	2	0
	2	1	1	2	0.5
	3	3	1	4	1
X	1	1	3	4	0.75
	2	3	1	4	0.25
	3	2	1	3	0
MIX	1	2	2	4	0.25
	2	3	1	4	0.5
	3	2	1	3	0
SPAT	1	5	1	6	0.5
	2	3	1	5	0.6
	3	3	1	4	1
Totals		34	18	52	0.5

Note.—Mutations are categorized by whether the predicted function of the mutation-bearing gene is motility-related (Mot.) or not (Non-mot.). Frequencies of mutations that were polymorphic are also listed.

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015

ANOVA, $P = 0.0385$), with populations selected in the SPAT environment having the largest number of mutations. These differences appear to be driven by mutations in genes not known to be related to motility (resampled ANOVA on nonmotility-related genes: $P = 0.0361$; resampled ANOVA on motility-related genes: $P = 0.8215$). Details of the observed mutations are summarized in supplementary table S1, Supplementary Material online.

Parallel Evolution as a Function of Biological Hierarchy

Fitness

The relative fitness of all populations increased by 24.6% (±2.3% SE) after 1,000 generations of evolution and this change in fitness did not differ significantly across selection environments (ANOVA, $F_{4,10} = 0.767$, $P = 0.571$; see Bailey and Kassen 2012 for further details). However, the way in which populations gained those fitness improvements differed across selection environments. Figure 1*A* shows how evolved populations differ in their fitness in mannose and xylose (two of the three component sugar resources used in the resource-use fitness assays). Selection environment drove the evolution of different resource-use phenotypes (nested-ANOVA, interaction of selection environment and assay environment [S × A]: $F_{8,140} = 15.192$, $P < 0.001$), with the exception of the X and MIX environments. Within each selection environment, replicate populations did not differ in their resource-use phenotype—including the interaction of population and assay environment (P × A) did not significantly improve the ANOVA model fit (χ^2 test comparing model log-likelihoods: $P = 0.9966$).

Catabolism

The first two principle components of the Biolog growth assay data (explaining 15.8% and 6.53% of the total variation, respectively) were used as a summary of catabolic phenotype. With this phenotypic measure, populations do appear to group loosely by selection environment (fig. 1*B*), but this effect was not significant (nested-ANOVA, S × A: $F_{4,158} = 1.5066$, $P = 0.200$). Catabolic phenotype differed significantly by population, independent of selection environment—including P × A significantly improved model fit (χ^2 test comparing model log-likelihoods, $P < 0.001$).

Gene-Level Changes

At the gene level, there was substantial overlap between populations in the changes that occurred—50% of the 22 mutation-bearing genes were seen in at least two independently evolving populations (fig. 2*A*), and the gene with the greatest number of repeat changes (PFLU4443, likely coding for a flagellar regulatory protein) was mutated in eight independent populations. The mutation-bearing genes for which function could be predicted were distributed across seven functional categories with the greatest proportion of mutation-bearing genes found in the cell motility category—about 25% (fig. 2*B*). Interestingly, those genes that were repeatedly mutated in selection environments that otherwise did not share any carbon resources (e.g., between populations from the mannose and the glucose environments, respectively) were significantly associated with motility ($\chi^2 = 7.0911$, df = 1, $P = 0.0077$). Gene ontology (GO) enrichment analysis showed significant enrichment for genes related to cell motility, and this enrichment was consistent across all selection environments ($P < 0.0001$).

Nucleotide-Level Changes

The majority of the mutational changes detected by whole-genome sequencing were unique (45 of the total 52); however, there were three examples of the same change occurring in multiple independently evolved populations. Those changes were 1) a 11-bp deletion resulting in a frameshift in the *fruR* gene, seen in two SPAT populations; 2) an SNP resulting in a stop gained in the PFLU1301 gene, seen in a MIX population and two SPAT populations; and 3) a 3,175-bp

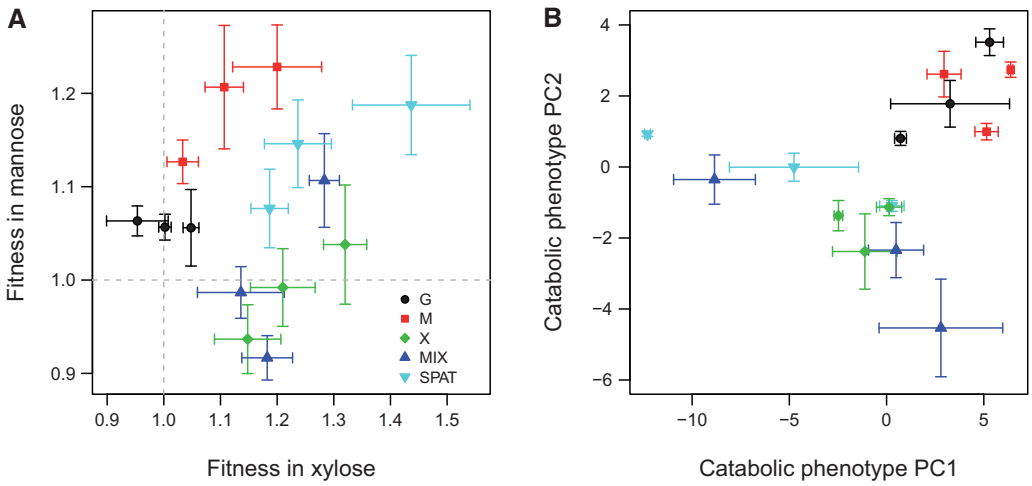

Fig. 1. (*A*) Fitness relative to the ancestor in mannose and xylose (two of the three component resources), for each of the 15 populations. (*B*) Catabolic phenotype (measured as growth on 95 carbon substrates) for each of the 15 populations summarized by their location in the first two dimensions of principle component space. See Bailey and Kassen (2012) for further details on these phenotypic changes. Points and error bars represent mean ± standard error ($N = 6$) for each population.

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015

deletion resulting in the loss of the part of the PFLU2574 and all of the PFLU2575 gene, seen in two MIX populations. Due to the protocol we used to initiate the replicate populations in this experiment (all initiated from the same ancestral population—a clonal population grown up overnight from a single cell), we are unable to distinguish whether these repeat mutations arose de novo in each population independently or were already present at very low frequencies in the founding population. Thus, it is possible that we have overestimated the degree of parallel evolution at the level of the nucleotide in this study; however, note that the frequency of repeat nucleotide-level changes across populations in our experiment is quite similar to that seen in comparable experiments (e.g., Tenaillon et al. 2012).

Using a common measure of similarity (Jaccard index), we compared degree of parallelism across levels of biological hierarchy. Degree of parallelism follows a clear decreasing trend, as expected, as one moves from fitness all the way down to phenotype (fig. 3).

Quantitative Tests of the Probability of Gene-Level Parallel Evolution

Effect of Selection Environment

We estimated the probability of parallel evolution at the gene-level using the Jaccard Index (*J*), which describes the likelihood that the same gene is mutated in two independent populations. Observed values of *J* ranged the full spectrum of possibilities, from 0 to 1, and the distributions in each

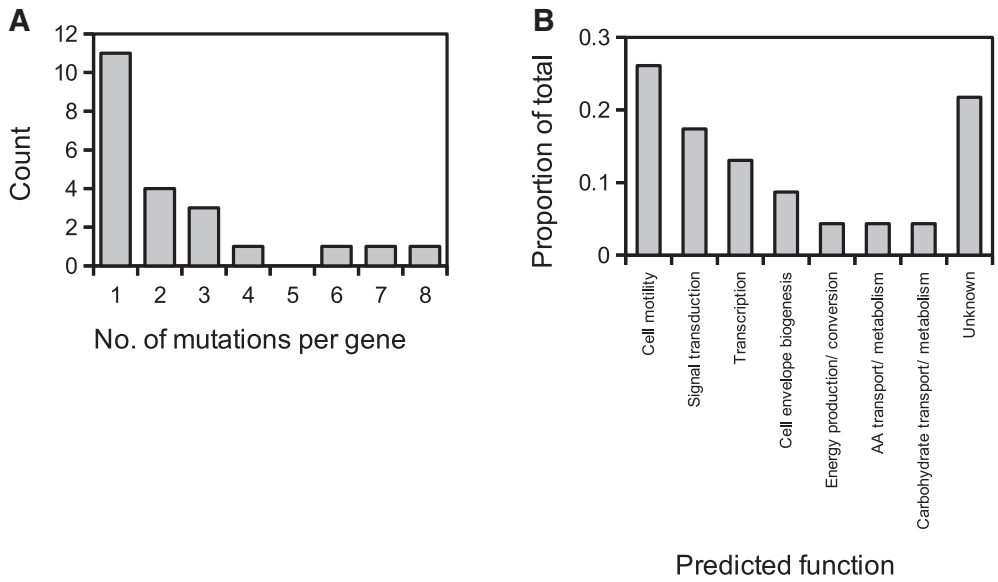

Fig. 2. (*A*) Distribution of mutations across the mutation-bearing genes of the 15 independent populations. (*B*) Proportion of mutation-bearing genes from all 15 populations categorized by predicted function.

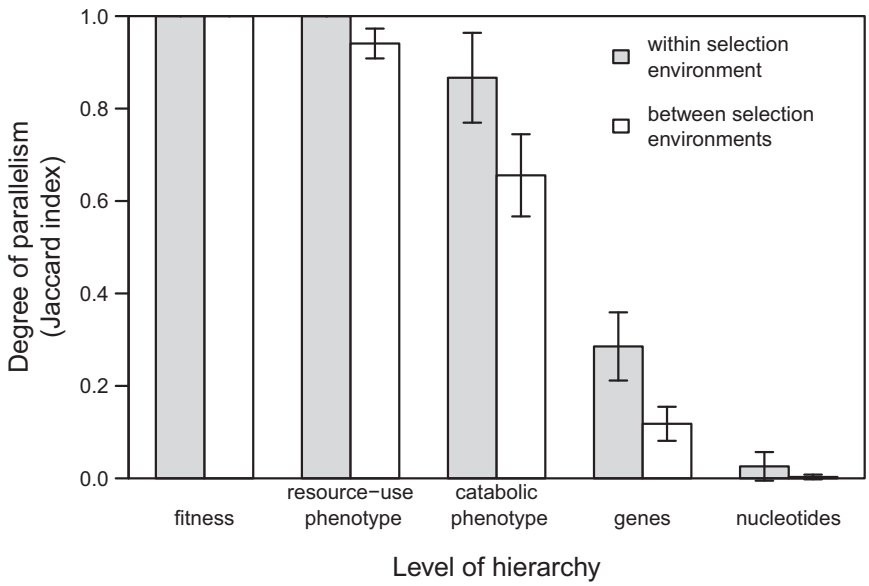

Fig. 3. Degree of parallelism across levels of biological hierarchy. Degree of parallelism is estimated as the mean Jaccard index for populations compared within and between selection environments. Error bars show standard error of all Jaccard index values calculated for a given category.

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015

selection environment appear complex and quite zero-heavy. Figure 4A shows the mean *J* values between populations from different selection environments (black lines and boxes), and between populations evolved within the same selection environment (gray lines and boxes). Mean *J* values that are significantly higher than expected, given the set of all mutation-bearing genes observed, are indicated by solid lines. Some degree of overlap in genetic change was seen between populations in almost all pairs of selection environments (with the exception of the xylose–mannose comparison), and the mannose–SPAT comparison showed significantly more overlap than expected. Although replicate populations evolving within the same selection environment all overlapped to some extent, populations evolving in the mannose and the SPAT environments were significantly more similar than expected. *J* values calculated with only the motility-related genes—those genes expected to have more general, less environment specific, fitness consequences—showed a similar picture (see supplementary fig. S1, Supplementary Material online).

A more general picture of the results comes from examining the average degree of parallel evolution among pairs of populations from the same or different environments. On average, independently evolved populations from the same environment had significantly higher probabilities of parallel evolution compared with those from different environments (resampled ANOVA, main effect of within/between environments; $P = 0.0004$). This effect was variable across treatments, however, and driven largely by the populations from mannose, SPAT, and marginally xylose, environments (mannose $P = 0.0444$, SPAT = 0.0157, xylose $P = 0.0517$; fig. 4B). Independently evolved populations from within the glucose and MIX populations, on the other hand, were no more similar to each other than to populations from other selection environments (glucose $P = 0.4583$, MIX $P = 0.5271$).

Distance to the Optimum

Variation among selection environments in the probability of parallel evolution is not associated with distance to a fitness optimum (fig. 5A; resampled linear regression, $P = 0.9225$). There is, however, a marginally significant negative relationship between the frequency of beneficial mutations in a given selection environment (estimated in an independent assay, see Materials and Methods) and the degree of parallel evolution (resampled linear regression: fig. 5B; $P = 0.0600$), as expected from theory. We note that this marginal significance must be taken with some caution, as the relationship is largely driven by populations evolved in the SPAT environment. When the SPAT populations are removed from the regression analysis, a negative trend remains but the relationship is no longer significant ($P = 0.665$).

Effect of Environmental Heterogeneity

We noted previously that the probability of parallel evolution was markedly higher in the SPAT environment than in most other treatments, consistent with the idea that the combination of divergent selection and gene flow restricts the number of mutations capable of improving fitness relative to a situation involving divergent selection without gene flow. The absence of increased parallelism in the MIX environment, moreover, implies that this result is not a general effect of environmental heterogeneity per se but, rather, a product of the specific properties of the SPAT environment involving sequential rounds of divergent selection followed by gene flow.

Discussion

Parallel evolution is often taken to be compelling evidence of natural selection for a particular phenotypic or genetic

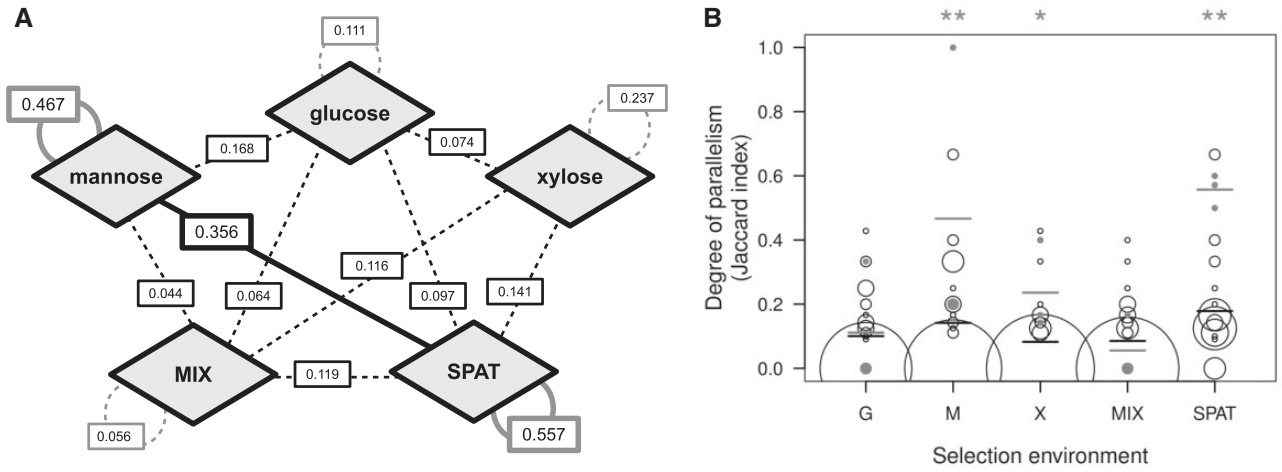

Fig. 4. (*A*) Mean Jaccard index for all mutation-bearing genes, both within (gray) and between (black) selection environments. Solid lines indicate comparisons with significantly higher similarity than expected. (*B*) Jaccard index (J) by selection environment for all population comparisons of the genes bearing mutations. Diameter of each circle is proportional to the number of observations at that value. Open circles indicate between-environment comparisons; black lines indicate the mean between-environment *J* value for each selection environment. Filled gray circles indicate within-environment comparisons; gray lines indicate the mean within-environment *J* value for each selection environment. ** indicates between- and within-environment *J* values are significantly different at $P < 0.01$ and * indicates significance at $P < 0.05$.

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015

"solution" to an adaptive challenge. Although this may often be an appropriate interpretation, it is important to recognize that there must also be appropriate genetic variation available on which selection can act. Any factor that changes the strength of selection or the availability of genetic variation can impact the likelihood of observing parallel evolution. Here we have studied the impact of divergent selection on the probability of parallel evolution by examining a range of phenotypic and genetic endpoints, including whole-genome sequence data, from a previous selection experiment designed to investigate the effects of spatial structure on the dynamics of adaptation.

This study provides insight into the general properties of parallel evolution across levels of the biological hierarchy and, at the same time, allows for more direct, quantitative tests of the theory of parallel evolution. Briefly, we find that parallel evolution is very common at the highest level of biological organization we can study, fitness, and becomes less and less common as one descends down the hierarchy to phenotypes, genes, and nucleotides. In more quantitative terms, we found that the probability of parallel evolution at the level of the gene tends to be higher among replicate populations evolving in the same environment when compared with populations from distinct environments. The probability of parallel evolution can also be highly variable across selection environments, and this variation seems to be related to the supply of beneficial mutations, not distance to a phenotypic optimum. Below we expand on these observations in more depth.

Parallel Evolution and the Hierarchy of Biological Organization

We have documented the extent and variation in evolutionary parallelism in experimentally evolved populations of *P. fluorescens* selected in a range of environments across multiple levels of biological organization, from fitness and phenotype to genotype. The degree of parallelism varied from complete overlap at the level of fitness, where all 15 populations increased fitness by the same amount, to nearly absent at the level of nucleotide changes: Only three repeated nucleotide changes occurred out of a total of 48 unique changes observed. In between, at the level of phenotype and the level of mutation-bearing genes, the degree of parallelism varied substantially among replicate populations in the same environment and across selection environments. These results suggest a pyramid-like relationship between the traits associated with a particular level of organismal organization and the variety of routes to evolution. Fitness, as the trait most important for adaptation, sits at the top of the pyramid and so is likely to most often result in parallel evolution. Below fitness are a wider range of phenotypes, and even more combinations of genotypes, that can contribute to improving fitness, and this is reflected in ever-lower degrees of parallelism. Our work provides a comprehensive view of this pattern of parallelism spanning from nucleotides to fitness, and confirms numerous previous studies showing this same hierarchal pattern for various parts of this nucleotide to fitness spectrum (e.g., Woods et al. 2006; Fox et al. 2011; Tenaillon et al. 2012; Nguyen et al. 2012; Bedhomme et al. 2013).

Phenotypic and Genetic Parallelism Vary among Selection Environments

On average, and consistent with intuition, the degree of parallelism was higher among replicate populations from the same environment and lower when populations are compared across environments at both the phenotypic and genetic level. In terms of resource-use phenotype, we saw that, as expected, populations evolving in different selection environments tended to have different phenotypes, whereas populations evolved in the same selection environment did not differ significantly (fig. 1*A*). An exception to this trend involved the xylose-selected and MIX-selected populations,

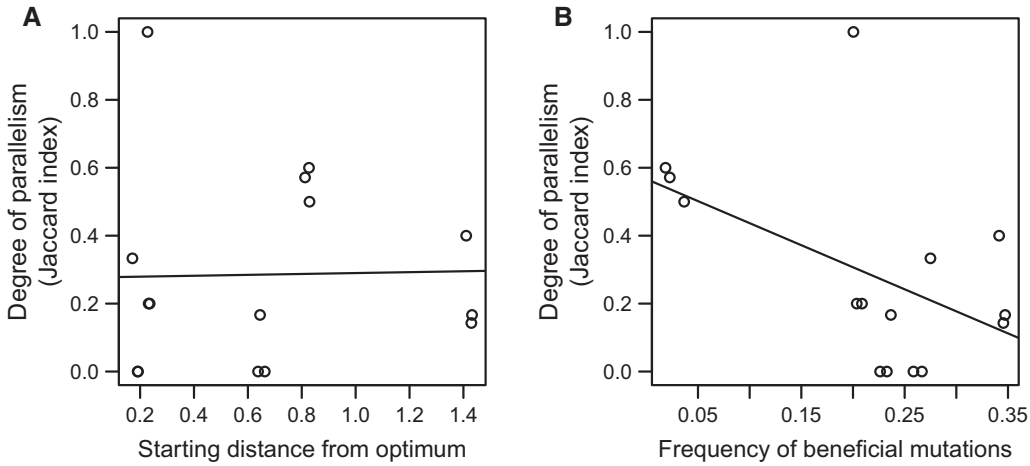

Fig. 5. (*A*) Degree of parallel evolution (calculated for each replicate population as the mean within-environment Jaccard Index) over starting distance from the fitness optimum (estimated in each selection environment as the log difference in 24-h density between ancestral and evolved populations). Line indicates the nonsignificant linear regression fit. (*B*) Degree of parallel evolution versus the frequency of beneficial mutations in a given selection environment estimated from fitness assays of mutants described in Kassen and Bataillon (2006). Line shows the marginally significant linear regression fit ($P = 0.06$). Points are jittered in the *x* axis direction for ease of display.

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015

which both had similar resource-use phenotypes characterized by adaptations specific to xylose (see Bailey and Kassen 2012 for further discussion of this). Notably, however, the genes associated with adaptation to xylose differed between these two environments. For example, all populations selected in xylose evolved what appear to be loss-of-function mutations in PFLU2366—a gene coding for a putative GntR family regulatory protein. However, mutations in PFLU2366 were not a distinguishing feature of populations evolved in the MIX environment. These gene-level differences suggest that the details of the environment are very important in driving selection and that these details may drive evolved differences in phenotype too subtle to detect with our phenotype assays.

We saw a similar pattern, of more pronounced parallelism among populations from the same environment compared with those from distinct environments, at the gene level as well. However it is interesting to note that the degree of parallelism varied among environments, with some displaying markedly higher levels than others. What factors explain this variation is not entirely clear and we discuss each in turn below.

Degree of Adaptation

Recent theoretical work by Chevin et al. (2010) and Orr (2005) suggests that there should be a negative relationship between the degree of adaptation and the probability of parallel evolution. The rationale behind this prediction is straightforward: A mal-adapted population located some distance from a fitness optimum has many more possible beneficial mutations available to it than one located at the optimum itself, as by definition no mutation can be beneficial for a population that resides on a fitness peak. In the selection environments used here, ancestral populations started at a range of distances from their fitness optima, allowing us to test this prediction directly. We found no evidence of such a relationship. However we did find marginal support for the idea that the relative number of beneficial mutations available to selection across environments is negatively related to the degree of parallel evolution, as expected from theory. Notably, there is no significant effect of distance to the optimum on the relative number of available beneficial mutations in our system (linear regression, $F_{1,3} = 0.1346$, $P = 0.7380$). Evidently these two factors, while correlated in analyses of Fisher's geometric model (Chevin et al. 2010), are not strongly linked in our system. Nevertheless this result lends support to the idea that the number of beneficial mutations available to selection is a key parameter determining the likelihood of parallel evolution, consistent with the suggestion by Orr (2005).

Degree of Pleiotropy

Chevin et al. (2010) point out that high degrees of pleiotropy can increase the probability of parallel evolution by effectively restricting the phenotypic dimensions along which selection can act. The degree of pleiotropy is challenging to measure as it requires knowing the "realized" or "effective" number of traits, *m*, that respond to selection and how this varies across environments. Unfortunately our knowledge of the phenotypic traits actually under selection in our system, and how they are pleiotropically related (or not) to each other is extremely limited. Although there is some experimental evidence suggesting that known highly pleiotropic genes may harbor more instances of parallel evolution (Spor et al. 2014), it is much more difficult to compare genomic levels of pleiotropy for populations evolving in one environment versus another, and so a test of this mechanism is beyond the scope of this study.

Environmental Heterogeneity

Environmental heterogeneity generates divergent selection which, depending on the level of gene flow, can lead either to the evolution of specialized types for each environment or a single generalized type that maintains high fitness across a broad range of environments. The latter situation might be expected to lead to high levels of parallel evolution if the supply of mutations that are favorable across a range of conditions is fewer than those that could increase fitness in any of the component environments. Consistent with this expectation, we observed the evolution of a broadly adapted generalist type in the SPAT treatment of the initial experiment and this treatment also showed the highest levels of parallel evolution. Notably, parallel evolution was low in the MIX treatment that, while containing the same range of resources as SPAT, lacked any spatial structure and led to the emergence of xylose-specialists. It is possible that larger population sizes in the SPAT environment (due to the logistics of our experimental setup, see Materials and Methods) could have contributed to its high degree of parallelism as compared with the MIX populations in two possible ways. First, when population sizes are larger rare beneficial mutations with a large fitness effect would be uncovered more often, leading to a higher probability of parallelism. This explanation seems an unlikely one in our experiment, however, because the rate of adaptation in SPAT was actually slower than that in MIX (Bailey and Kassen 2012), contrary to what we would expect when population size limits the supply of beneficial mutations, and so the rate of adaptation. Second, in very large populations, clonal interference can bias the arising mutations to ones with very large effects, again leading to a higher probability of parallel evolution. However, this is also unlikely to have occurred, as clonal interference is also expected to drive the fixation of fewer mutations (Gerrish and Lenski 1998), while in fact we saw the opposite—a greater number of mutations in the SPAT populations compared with the others. We suggest instead that selection for generalists may act as a constraint on the supply of beneficial mutations and so lead to higher levels of parallel evolution. It is further notable that adaptation to the SPAT environment involves more genetic changes than adaptation in the other selection environments in genes unrelated to motility (table 1), suggesting that they are not adaptations to common conditions of culture. Whether these mutations represent the accumulation of a series of mutations that independently increase fitness on each component resource separately or represent adaptations to the fluctuating nature

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015

of the SPAT treatment remains unclear at this point but represents an interesting avenue for future work.

Mutation Rate Heterogeneity

It is also possible that the variation in the degree of parallel evolution seen in this study is influenced by variation in mutation rates across the genome. Mutation rate is known to vary across genomes on many scales—from increasing mutation rates with increasing distance from the origin of replication (Mira and Ochman 2002), to the general mutational bias toward AT (Hershberg and Petrov 2010). Some genes under selection may have higher mutation rates than the other genes under selection in that environment, and that potential bias in mutation rate might lead to biases in the particular mutations that fix. Thus, one would expect populations evolving in environments where mutation rates vary widely across the genes under selection to show a high degree of parallel evolution, whereas populations evolving in environments where mutation rate is more consistent across the genes under selection would be less likely to evolve in parallel. Our lack of knowledge of mutation rate variation in this species, and *Pseudomonads* more broadly, prevents us from testing this prediction directly here. Ongoing work is aimed at characterizing genome-wide variation in mutation rates and selection in *Pseudomonads* using both experimental and theoretical approaches.

General versus Environment-Specific Adaptations

We observed a number of genetic changes in these populations that appear to be general adaptations to well-mixed lab conditions. In particular, those mutations that occur in motility-related genes are commonly predicted to result in loss-of-function of flagella assembly and so are expected to result in fitness improvements common to all conditions in our experiment. Indeed, preliminary work involving allelic replacement of a subset of these mutations lends support to this interpretation (Hinz A, personal communication). Repeated evolution of loss of motility is not surprising for populations selected in well-mixed liquid media where the ability to move is unnecessary, as assembly and use of the complex machinery required for locomotion is likely to be quite costly (Bardy et al. 2003). Thus, these types of mutations might be expected to generate high degrees of parallel evolution across selection environments in our system. Indeed, at the level of phenotype this is certainly the case: In all but one of the populations we identified mutations predicted to be related to motility. Some amount of parallel evolution in this trait also occurred across environments at the gene level—those few gene-level changes that were shared across all three single-resource environments were all genes associated with motility. However, interestingly the availability and arrangement of resources still appears to constrain, to some extent, which motility genes bear mutations in a given population (see supplementary fig. S1, Supplementary Material online). Variation in motility-gene mutations between selection environments must be due to pleiotropic effects of motility genes on resource transport and/or catabolism that we are unable to detect, further underscoring the importance of environmental context on evolutionary processes.

Conclusions

As genome sequencing continues to become faster and cheaper, evolution experiments with large numbers of replicate populations will help to identify more clearly the factors affecting the probability of parallel evolution—a value inherently difficult to characterize simply because it is the outcome of so many stochastic events. Our ability in this study to detect differences between selection environments despite widely varying and complex distributions of parallelism estimates, and with only three replicate populations per environment, is indicative of how important environmental context is for the rates and processes of adaptation. At the most basic level, it appears the availability of beneficial mutations drives the observed differences in parallel evolution in our system as predicted by theory. However in contrast to theory, degree of parallel evolution does not vary with distance to the fitness optimum in our system, and although heterogeneity in the environment does seem to play a role there are certainly other as of yet unidentified mechanisms at work. Indeed, in this experiment and others, the distribution of available beneficial mutations is difficult to characterize and the factors responsible for determining this key distribution are often unclear. Future theoretical, comparative, and experimental work is necessary to further expound the mechanisms driving parallel evolution, and to continue working toward an understanding of how and when evolution is predictable.

Materials and Methods

Selection Experiment

We used a clonal isolate of the bacterium *P. fluorescens* SBW25:lacZ to found 15 independent lines. This strain is isogenic to *P. fluorescens* SBW25 save for the insertion of the lacZ gene, a selectively neutral marker (Zhang and Rainey 2007). Colonies with lacZ are blue on agar plates supplemented with 40 mg/l of 5-bromo-4-chloro-3-indolylbeta-D-galactopyranoside (X-Gal), and can easily be distinguished from the pale yellow colonies of the unmarked strain. All strains and evolving populations were frozen at −80 °C in 16% (v/v) glycerol. Populations were all initiated from a single ancestral population grown up overnight from a single cell and cultured in 24-well plates (Costar, Corning Incorporated), with 2 ml of media in each well, in an orbital shaker (150 rpm) at 28 °C. The culture media consisted of M9 minimal salts (1 g/l NH_4Cl, 3 g/l KH_2PO_4, 0.5 g/l NaCl, and 6.8 g/l Na_2HPO_4) supplemented with 15 mg/l $CaCl_2$, 0.5 g/l $MgSO_4$, and a carbon source. Three monosaccharides were used as carbon sources: Mannose, glucose (both $C_6H_{12}O_6$), and xylose ($C_5H_{10}O_5$). When grown in a 1 M sugar concentration for 24 h (150 rpm at 28 °C), SBW25 reaches cell densities over 10-fold higher in glucose (mean ± SE = $5.33 \pm 1.09 \times 10^8$ colony forming units [CFUs]) and mannose ($5.76 \pm 0.54 \times 10^8$ CFUs) compared with xylose ($1.33 \pm 0.15 \times 10^7$ CFUs). For this reason, concentrations

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015

were adjusted to control for population size resulting in environments containing 9.57 mg/l of mannose and glucose, and 255.22 mg/l of xylose. Single-resource and multiresource environments were constructed containing one sugar type and all three sugar types, respectively. Two types of multiresource environments were created: 1) Fine-grained: Three sugars in a single well (MIX); and 2) coarse-grained: Three sugars available, each in a different well with population mixing and redistribution at each 24-h transfer (SPAT).

In total, the experiment consisted of three replicates in each of five different selection environments. Every 24 h (~6.6 generations), a subsample of each culture was transferred to fresh media. In treatments consisting of a single microwell (the single-resource environments and MIX), 20 µl of culture was transferred. In treatments consisting of multiple microwells (SPAT), each microwell also received 20 µl but this transferred volume was made up of equal parts by volume of samples from each of the microwells making up that selection environment. This transfer regime was repeated for 150 transfers resulting in approximately 1,000 generations of selection. Genome sequence data of the evolved populations confirmed that there was no external- or cross-contamination over the course of the evolution experiment. See Bailey and Kassen (2012) for further methodological details.

Fitness and Phenotype Assays

Fitness of the evolved populations in their selection environment and in the three single resource environments was estimated by competing evolved strains against the ancestral genotype lacking the lacZ neutral marker. Each strain was first acclimated from frozen on the substrate of interest for two growth cycles (48 h), then evolved and ancestor strains were mixed at a 1:1 ratio for a transfer volume of 20 µl. Estimates of the relative frequency of both types were quantified after the first (24 h) and third (72 h) growth cycles by plating on M9 minimal salts + glucose + X-Gal agar. We estimated relative fitness, ω, with the following equation:

$$\omega = (f_{\text{final}}/f_{\text{initial}})\hat{}(1/\text{doublings}), \quad (1)$$

where f_{initial} and f_{final} are the ratios of the frequency of the evolved population to the frequency of ancestral populations before and after competition, respectively, and doublings refers to the number of doublings or generations that occur between the initial and final measurements. Similar results arising from temporally independent replicate assays suggest the absence of both external- and cross-contamination during these assay experiments.

Resource-use phenotype (a three-trait phenotype consisting of a fitness measure in each of the three resources—glucose, mannose, and xylose) was compared across selection environment and population using a mixed effects nested-ANOVA, where population was modeled as a random effect nested within selection environment, and all other variables were modeled as fixed effects. This resulted in the following model, trait value = $P + S + A + P \times A + S \times A$, where P is population, S is selection environment, A is assay environment, and in this case "trait value" is relative fitness of a population in a given assay environment.

The catabolic profiles of the evolved strains were characterized using GN2 Biolog microwell plates containing 95 different carbon sources plus a negative control well. Strains were grown from frozen overnight in 6 ml King's B media (28 °C, shaken at 150 rpm), starved for 2 h (20 µl of each culture in 20 ml M9 minimal salts at 28 °C, 150 rpm), and then 150 µl was transferred into each well of the Biolog plates. Optical density (OD) was measured at the time of inoculation and after 24 h of growth (28 °C, 150 rpm), and growth rate on each carbon substrate was calculated as $r = \ln(\text{OD}_{\text{initial}}) - \ln(\text{OD}_{\text{final}})$. Growth rates were then adjusted relative to the control well by subtracting the maximum growth rate estimate obtained from the control wells (0.09), and setting all negative growth estimates to zero. The adjusted growth estimates were transformed into their principle components (first adjusting and scaling the data, then transforming using the "pr.comp" function in R). We then restricted our analysis to the first two principle components (all other principle components described less than 5% of the total variation) and then compared these summary catabolic phenotypes across population and selection environment using a mixed effects nested-ANOVA as in the resource-phenotype comparison. We note that our estimates of r assume that populations are still in exponential growth phase after 24 h. However, if this is not always the case it may lead to underestimates of growth rate and potentially overestimates of the degree of parallelism at the level of catabolic phenotype. Although we do not have growth curve data to rule out this possibility, we suggest that the assumption likely does hold for the most part, as OD measures from Biolog assays tend to continue to increase when left for second 24-h period of growth (data not shown).

Genome Sequencing and Analysis

For each evolved population and the ancestral SBW25 genotype, genomic DNA was extracted from an overnight culture using the Promega Wizard Genomic DNA Purification kit. 75-bp paired-end Illumina sequencing was performed by the Michael Smith Genome Sciences Centre. Mean coverage across all 15 populations was approximately 60-fold (mean = 60.91; range 43.15–89.88) at a quality score of 20. We performed pair-end mapping of reads on the Pf SBW25 reference genome number NC 012660.1, called and filtered snps/indels, and annotated with respect to the reference genome using the pipeline outlined in the supplementary information of Dettman et al. (2012). We called and filtered snps/indels again using the BRESEQ pipeline (Barrick et al. 2009) to validate the previously identified snps/indels and also to identify larger-scale insertion/deletions that the BRESEQ pipeline is better suited to detect. Mutational changes in the sequence data identified by either pipeline were checked by viewing the mapped reads, and a subset of those mutational changes were then confirmed using targeted sequencing. A number of mutations initially detected in the whole-genome sequencing analysis were

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015

located in intergenic repeat regions. However, these putative mutations were not well supported due to very low coverage (repeat regions are difficult to accurately sequence), and are unlikely to have fitness effects even if they are real. Thus, we removed those putative mutational changes from any further analyses. Because we sequenced population genomic DNA, as opposed to that of single clones, we were able to detect evolved polymorphisms in our populations. We classified variation in reads at a given locus as a true population polymorphism (as opposed to sequencing error) when at least 10% of the reads differed from the others for a given allele.

We classified the mutation-bearing genes into functional groups using predicted Clusters of Orthologous Groups (COGs) classifications (Tatusov et al. 2000) provided in the Pseudomonas Genome Database (Winsor et al. 2011). In particular, we were interested in identifying genes with predicted "cell motility" functions, as mutations in these genes may simply be general adaptations to living in a well-mixed lab environment, whereas we suggest that mutations in genes with non cell motility related functions are more likely to be resource-specific adaptations. We also identified the GO enrichment for all mutations observed in experiment and for mutations observed in each selection environment, using the "Batch-Genes," a tool that is part of GOEAST (Zheng and Wang 2008). We ran this analysis using both the *P. aeruginosa* PAO1 and *P. protegens* Pf-5 GO databases as there was no *P. fluorescens* GO database available for this software. Analyses with the two different GO databases gave identical results.

Gene-Level Parallel Evolution Comparisons

We quantified similarity in genetic changes between pairs of populations using the Jaccard Index *J*. Given two sets, G1 and G2, of the mutation-bearing genes found in populations 1 and 2, respectively,

$$J(G1, G2) = (G1 \cap G2)/(G1 \cup G2). \quad (2)$$

That is, *J* is the number of genes mutated in both populations divided by the total number of genes mutated in population 1 or in population 2. *J* ranges from 0 to 1, with 1 indicating identical genotypes and 0 indicating no shared mutation-bearing genes. This value was calculated for all pairs of evolved populations, both within and between selection environments. Due to the nonindependence of *J* estimates, statistical significance of the mean *J* value for each "comparison set" was estimated using a resampling procedure. By "comparison set" we mean a set of all pairwise population comparisons of a particular type; for example, all mannose-evolved to glucose-evolved population comparisons would make up a single "comparison set." For each comparison set, the populations' mutation-bearing genes were resampled from the set of all mutation-bearing genes identified in the combined 15 population data set. For each population, the number of resampled mutations was set equal to the number of actual mutations in that population. *J* was recalculated for each comparison in the set, and a new mean *J* was calculated. This procedure was repeated 50,000 times for each comparison set, generating distributions of mean *J* values. Using these distributions, we calculated the probability of observing the actual mean *J* values or greater by chance alone (*P* value). Significance was assigned to probabilities less than 0.0033—an experiment-wide *P* value of 0.05, Bonferroni-adjusted for 15 comparison sets. Calculation of *J* and significance testing were performed for all mutation-bearing genes and also for the subset of those mutation-bearing genes that were related to motility-related.

We contrasted estimates of similarity within- and between-selection environments by subtracting the average between-environment *J* values from the average within-environment *J* values for each replicate population. With this metric, a positive value indicates that a population is more similar on average to populations selected in the same selection environment, whereas a value of zero indicates that there is no systematic difference between comparisons of populations selected in the same environment versus different environments. For reasons of nonindependence we again used resampling methods, testing whether the mean differences are significantly greater than zero, and whether the effect differs with selection environment. For this test, the calculated within/between environment differences were randomly reassigned to a selection environment, and *t* statistics were calculated using an ANOVA procedure. This was repeated 50,000 times, generating distributions to which the actual *t* statistics were compared and *P* values were then calculated.

We compared across levels of biological hierarchy using a common measure of parallelism—the Jaccard index. For gene and nucleotide level parallelisms, the Jaccard index was calculated as described above. At the level of fitness, the Jaccard index was simply defined as a 0 or 1 for any given comparison, depending on whether the given pair of populations being compared was significantly different from each other or not (as determined by pairwise *t*-tests using Bonferroni corrected experiment-wide *P* values < 0.05). For the level of resource-use phenotype, Jaccard index could be 0, 1/3, 2/3, or 1 for any given comparison, depending on whether the given pair of populations being compared was significantly different from each other in all, 1, 2, or none of the resources tested (again, determined by pairwise *t*-tests using Bonferroni corrected experiment-wide *P* values > 0.05). At the level of catabolic-phenotype Jaccard index was calculated similarly, but with only two categories of comparison—PC1 and PC2.

We used the Jaccard similarity index measures to look for a relationship between degree of parallel evolution at the gene level and 1) distance to the fitness optimum and 2) the relative number of beneficial mutations available. To estimate distance to the fitness optimum we used the log difference between ancestral and evolved absolute fitness in each selection environment, where absolute fitness was estimated as cell density of a population after 24-h growth. We note that although cell density is just one component of absolute fitness, it is arguably the most important component given our experimental protocol. Having a high cell density, and so large population size, after 24 h is clearly highly beneficial in an experimental setup where populations are bottlenecked to

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015

1% of their size every 24 h. In addition, we note that our measure of absolute fitness assumes that evolved populations have reached a fitness optimum within the time course of the experiment; characterizations of these fitness trajectories suggest that this may in fact be the case (Bailey and Kassen 2012).

To estimate the relative number of beneficial mutations available, we assayed the fitness of a collection of nalidixic acid resistant mutants (described in Kassen and Bataillon 2006) in each selection environment. We then counted the number of mutants that were significantly more fit than the ancestor and divided that number by the total number of mutants tested, to obtain an estimate of the frequency of mutations that were beneficial in a given environment. We stress that this quantity indicates the frequency of beneficial mutations within the set of mutations tested in our assay, so although it can be used as a relative measure across selection environments in our experiment, it is unlikely to be a true estimate of the frequency of beneficial mutations within the set of all possible mutations. A resampled linear regression model was used to quantify and test for the significance of these relationships. All calculations and analyses were performed using R version 3.1.2 (R Development Core Team 2014).

Supplementary Material

Supplementary table S1 and figure S1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

This work was supported by grants from the Natural Sciences and Engineering Research Council (Canada) to S.F.B. and R.K., an Ontario Graduate Scholarship to S.F.B., and a grant from the Canadian Institutes of Health Research to R.K. S.F.B also acknowledges financial support from the European Research Council under the European Union's Seventh Framework Program (FP7/20072013, ERC Grant 311341).

References

Bailey SF, Hinz A, Kassen R. 2014. Adaptive synonymous mutations in an experimentally evolved *Pseudomonas fluorescens* population. *Nat Commun.* 5.

Bailey SF, Kassen R. 2012. Spatial structure of ecological opportunity drives adaptation in a bacterium. *Am Nat.* 180:270–283.

Bardy SL, Ng SY, Jarrell KF. 2003. Prokaryotic motility structures. *Microbiolgy* 149:295–304.

Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461:1243–1247.

Bedhomme S, Lafforgue G, Elena SF. 2013. Genotypic but not phenotypic historical contingency revealed by viral experimental evolution. *BMC Evol Biol.* 13:46.

Castoe TA, de Koning AJ, Kim H-M, Gu W, Noonan BP, Naylor G, Jiang ZJ, Parkinson CL, Pollock DD. 2009. Evidence for an ancient adaptive episode of convergent molecular evolution. *Proc Natl Acad Sci U S A.* 106:8986–8991.

Chevin L-M, Martin G, Lenormand T. 2010. Fisher's model and the genomics of adaptation: restricted pleiotropy, heterogenous mutation, and parallel evolution. *Evolution* 64:3213–3231.

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G Jr, Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307:1928–1933.

Conte GL, Arnegard ME, Peichel CL, Schluter D. 2012. The probability of genetic parallelism and convergence in natural populations. *Proc Biol Sci.* 279:5039–5047.

Copley RR. 2004. Evolutionary convergence of alternative splicing in ion channels. *Trends Genet.* 20:171–176.

Dettman JR, Rodrigue N, Melnyk AH, Wong A, Bailey SF, Kassen R. 2012. Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. *Mol Ecol.* 21:2058–2077.

Fernald RD. 2006. Casting a genetic light on the evolution of eyes. *Science* 313:1914.

Fisher RA. 1930. The genetical theory of natural selection. Oxford: Oxford University Press.

Fox CW, Wagner JD, Cline S, Thomas FA, Messina FJ. 2011. Rapid evolution of lifespan in a novel environment: sex-specific responses and underlying genetic architecture. *Evol Biol.* 38:182–196.

Gerrish PJ, Lenski RE. 1998. The fate of competing beneficial mutations in an asexual population. *Genetica* 102:127–144.

Gillespie JH. 1983. A simple stochastic gene substitution model. *Theor Popul Biol.* 23:202–215.

Gillespie JH. 1991. The causes of molecular evolution. Oxford: Oxford University Press.

Gresham D, Desai MM, Tucker CM, Jenq HT, Pai DA, Ward A, DeSevo CG, Botstein D, Dunham MJ. 2008. The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet.* 4:e1000303.

Harrison CJ, Corley SB, Moylan EC, Alexander DL, Scotland RW, Langdale JA. 2005. Independent recruitment of a conserved developmental mechanism during leaf evolution. *Nature* 434:509–514.

Herron MD, Doebeli M. 2013. Parallel evolutionary dynamics of adaptive diversification in *Escherichia coli*. *PLoS Biol.* 11:e1001490.

Hershberg R, Petrov DA. 2010. Evidence that mutation is universally biased towards AT in bacteria. *PLoS Genet.* 6:e1001115.

Kassen R. 2014. Experimental evolution and the nature of biodiversity. Denver (CO): Roberts and Company Publishers.

Kassen R, Bataillon T. 2006. Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. *Nat Genet.* 38:484–488.

Kryazhimskiy S, Rice DP, Jerison ER, Desai MM. 2014. Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science* 344:1519–1522.

Liu Y, Cotton JA, Shen B, Han X, Rossiter SJ, Zhang S. 2010. Convergent sequence evolution between echolocating bats and dolphins. *Curr Biol.* 20:R53–R54.

Mira A, Ochman H. 2002. Gene location and bacterial sequence divergence. *Mol Biol Evol.* 19:1350–1358.

Nguyen AH, Molineux IJ, Springman R, Bull JJ. 2012. Multiple genetic pathways to similar fitness limits during viral adaptation to a new host. *Evolution* 66:363–374.

Odeen A, Hastad O. 2003. Complex distribution of avian color vision systems revealed by sequencing the SWS1 opsin from total DNA. *Mol Biol Evol.* 20:855–861.

Orr HA. 2005. The probability of parallel evolution. *Evolution* 59:216–220.

R Development Core Team. 2014. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing.

Rodríguez-Verdugo A, Carrillo-Cisneros D, González-González A, Gaut BS, Bennett AF. 2014. Different tradeoffs result from alternate genetic adaptations to a common environment. *Proc Natl Acad Sci U S A.* 111:12121–12126.

Schoustra SE, Bataillon T, Gifford DR, Kassen R. 2009. The properties of adaptive walks in evolving populations of fungus. *PLoS Biol.* 7:e1000250.

Spor A, Kvitek DJ, Nidelet T, Martin J, Legrand J, Dillmann C, Bourgais A, Vienne D, Sherlock G, Sicard D. 2014. Phenotypic and genotypic convergences are influenced by historical contingency and environment in yeast. *Evolution* 68:772–790.

Steiner CC, Römpler H, Boettger LM, Schöneberg T, Hoekstra HE. 2009. The genetic basis of phenotypic convergence in beach mice: similar pigment patterns but different genes. *Mol Biol Evol.* 26:35–45.

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015

Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28:36–33.

Tenaillon O, Rodríguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD, Gaut BS. 2012. The molecular diversity of adaptive convergence. *Science* 335:457–461.

Wichman HA, Brown CJ. 2010. Experimental evolution of viruses: Microviridae as a model system. *Philos T Roy Soc B.* 365:2495–2501.

Winsor GL, Lam DK, Fleming L, Lo R, Whiteside MD, Nancy YY, Hancock RE, Brinkman FS. 2011. Pseudomonas Genome Database: improved comparative analysis and population genomics capability for *Pseudomonas* genomes. *Nucleic Acids Res.* 39:D596–D600.

Wittkopp PJ, Williams BL, Selegue JE, Carroll SB. 2003. *Drosophila* pigmentation evolution: divergent genotypes underlying convergent phenotypes. *Proc Natl Acad Sci U S A.* 100:1808–1813.

Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. 2006. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc Natl Acad Sci U S A.* 103:9107–9112.

Yokoyama S, Radlwimmer FB. 2001. The molecular genetics and evolution of red and green color vision in vertebrates. *Genetics* 158: 1697–1710.

Zhang J, Kumar S. 1997. Detection of convergent and parallel evolution at the amino acid sequence level. *Mol Biol Evol.* 14: 527–536.

Zhang X-X, Rainey PB. 2007. Construction and validation of a neutrally-marked strain of *Pseudomonas fluorescens* SBW25. *J Microbiol Methods.* 71:78–81.

Zhang Q, Wang XJ. 2008. GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. *Nucleic Acids Res.* 36(suppl 2): W358–W363.

Downloaded from http://mbe.oxfordjournals.org/ at UB Frankfurt/Main on April 1, 2015