Please note that this is an author-produced PDF of an article accepted for publication following peer review. The definitive publisher-authenticated version is available on the publisher Web site

Molecular Ecology
June 2013, Volume 22, Issue 12, pages 3292–3303
http://dx.doi.org/10.1111/mec.12312
© 2013 John Wiley & Sons Ltd

Archimer
http://archimer.ifremer.fr

The definitive version is available at http://onlinelibrary.wiley.com/

Evolutionary history and genetic parallelism affect correlated responses to evolution

Mickael Le Gac[1, 2, a, *], Tim F. Cooper[3], Stéphane Cruveiller[4], Claudine Médigue[4], Dominique Schneider[1,2]

[1] Laboratoire Adaptation et Pathogénie des Microorganismes, Université Joseph Fourier, Grenoble Cedex 9, France
[2] CNRS UMR 5163, Grenoble Cedex 9, France
[3] Department of Biology and Biochemistry, University of Houston, Houston, TX, USA
[4] CNRS-UMR 8030 & CEA/DSV/IG/Genoscope LABGeM, Evry Cedex, France

[a] Current address : DYNECO/Pelagos, Ifremer, Plouzané, France

*: Corresponding author : Mickaël Le Gac, fax: (+33) 2 98 22 45 48 ; email address : mickael.le.gac@ifremer.fr

Abstract :

We investigated the relationship between genomic and phenotypic evolution among replicate populations of *Escherichia coli* evolved for 1000 generations in four different environments. By resequencing evolved genomes, we identified parallel changes in genes encoding transcription regulators within and between environments. Depending on both the environment and the altered gene, genetic parallelism at the gene level involved mutations that affected identical codons, protein domains or were widely distributed across the gene. Evolved clones were characterized by parallel phenotypic changes in their respective evolution environments but also in the three alternative environments. Phenotypic parallelism was high for clones that evolved in the same environment, even in the absence of genetic parallelism. By contrast, clones that evolved in different environments revealed a higher parallelism in correlated responses when they shared mutated genes. Altogether, this work shows that after an environmental change or the colonization of a new habitat, similar ecological performance might be expected from individuals that share mutated genes or that experienced similar past selective pressures.

Keywords : adaptation ; Epistasis ; experimental evolution ; genome sequencing ; parallelism

38 # Introduction

39 Next-generation sequencing (NGS) technologies are revolutionizing our understanding of the
40 genetics of adaptation. From evolution studies focusing on specific candidate genes,
41 investigations have now switched to the genome scale. Applied to experimental evolution of
42 microorganisms, genome comparisons of ancestral and evolved clones enable the
43 identification of virtually all the mutations associated with experimental adaptation (reviewed
44 in Brockhurst *et al.* 2011; Conrad *et al.* 2011; Dettman *et al.* 2012; Hindré *et al.* 2012). A
45 prominent finding has been that adaptation of independent replicate populations to a given
46 environment often involves genetic parallelism, characterized by mutations that affect
47 identical genes, operons, or functional operational units (Herring *et al.* 2006; Barrick *et al.*
48 2009; Conrad *et al.* 2009; Tenaillon *et al.* 2012). However, the precise mutations altering
49 these genetic loci are often different in replicate bacterial populations (Woods *et al.* 2006;
50 Tenaillon *et al.* 2012).

51 Consideration of the genes that are repeatedly mutated during evolution experiments finds
52 that many are global regulators of gene expression, highlighting the involvement of regulatory
53 network rewiring in adaptation (Philippe *et al.* 2007; Wang *et al.* 2010; Conrad *et al.* 2011;
54 Hindré *et al.* 2012). The same global regulator-encoding genes are often mutated in distinct
55 experimental settings indicating that identical regulatory hubs may provide adaptation to
56 contrasting environments (Conrad *et al.* 2011). Moreover, non-additive, *i.e.* epistatic,
57 interactions between mutations have been shown to be of primary importance in explaining
58 both the evolutionary trajectory of a population and the phenotypes of evolved individuals
59 (Bull *et al*. 2000; Remold & Lenski 2004; Weinreich *et al*. 2005; 2006; Poelwijk *et al*. 2007;
60 Cooper *et al*. 2008; Almahmoud *et al.* 2009; Le Gac & Doebeli 2010). Beneficial mutations in
61 global regulator-encoding genes likely reorganize the architecture of the cellular regulatory
62 networks, potentially contributing to the influence of epistasis (Khan *et al.* 2011; Tenaillon *et*

63 *al.* 2012), and to complex phenotypic changes in environmental conditions distinct from the
64 evolution environment (*i.e.*, correlated responses).
65 In the present study, we investigated the effect of beneficial mutations on correlated responses
66 to evolution. Replicate populations evolving in identical environments are often characterized
67 by having high phenotypic parallelism. The extent to which this parallelism applies to
68 correlated phenotypic responses is, however, less well studied. At one extreme, replicate
69 populations may display parallel correlated responses in many alternative environments. At
70 the other, they may display varied phenotypes under alternative conditions and phenotypic
71 parallelism only in the evolution environment. These phenotypic outcomes have important
72 ecological implications since they dictate how ecologically similar populations will cope with
73 environmental challenges such as the colonization of new habitats/hosts or the modification of
74 ecosystems at both the biotic and abiotic levels. Only few studies using model bacteria have
75 investigated the correlated phenotypic responses to evolution, including some that revealed
76 parallel losses of catabolic abilities and growth phenotypes on alternative resources (Cooper
77 & Lenski 2000; Barrett *et al.* 2005; Fong *et al.* 2005). However, in other studies using the
78 same bacterial species, populations have been found to be more variable when assayed for
79 growth on alternative resources than on the resource on which they first evolved (Travisano *et*
80 *al.* 1995; Travisano & Lenski 1996; MacLean & Bell 2003; Ostrowski *et al.* 2005; 2008).
81 The genetic changes associated with adaptation to a given evolution environment have only
82 rarely been investigated for their effects on correlated responses in alternative environments.
83 Intuitively, one may expect the correlated responses to be more similar for clones sharing
84 mutated genes, but the relationship between genetic parallelism and phenotypic performance
85 in alternative environments may be complex. For example, different mutations affecting a
86 given gene may have similar phenotypic effect in the evolution environment but not
87 necessarily in another one, especially in the case of global regulator-encoding genes.

88 Conversely, even if replicate populations adapt through mutations in different genes, the
89 phenotypic consequences of these mutations could be similar not only in the evolution
90 environment, but also under alternative environmental conditions.
91 Here, we propagated four replicate populations of *Escherichia coli* for 1000 generations in
92 each of four different environments. We sequenced the genome of one evolved clone isolated
93 from each of the 16 populations, allowing us to identify mutations associated with evolution
94 in each environment. We assayed the phenotype of isolated clones in their evolution
95 environment as well as in the three alternative ones to provide a measure of the direct and
96 correlated phenotypic responses to evolution in all four environments. Using this
97 experimental design, we investigated first, the level of genetic parallelism associated with
98 evolution in several environments and second, the respective effect of genetic parallelism and
99 evolutionary history on the correlated responses to evolution. In particular, we tested whether
100 phenotypes in alternative environments (correlated responses to evolution) were more similar
101 for clones sharing mutations in identical genes or genetic loci (genetic parallelism)
102 irrespective of the environment in which they evolved (evolutionary history) or for clones that
103 evolved in the same environment irrespective of genetic parallelism.
104

105 Material and methods

106 *Experimental evolution*
107 *E. coli* B REL606 (Lederberg 1966; Jeong *et al*. 2009) was used as the ancestor to found four
108 replicate populations that were propagated by daily serial transfer in each of four different
109 environments for 1000 generations at 37 °C. Population samples were collected at 100-
110 generation intervals and frozen at -80 °C as glycerol suspensions in the same conditions as the
111 ancestor. All four environments are based on Davis minimal (DM) medium (Lenski *et al*.
112 1991). The first environment, named Ace, comprised 15 mL DM supplemented with 2 g/L

113 sodium acetate trihydrate in 50-mL flasks shaken at 200 rpm. The second, named Gly,
114 comprised 15 mL DM supplemented with 1 g/L glycerol in static Petri dishes. The third,
115 named Glc, comprised 15 mL DM supplemented with 1 g/L D-gluconate in test tubes shaken
116 at 200 rpm. The fourth, named Glu, comprised 600 µL DM supplemented with 1 g/L D-
117 glucose in 1 mL x 96-well plates shaken at 200 rpm. These four environments were chosen in
118 order to apply a combination of parameters impacting bacterial growth and adaptation due to
119 contrasting carbon sources, and homogenization and oxygenation levels. This choice was
120 made *a priori* to allow a wide range of mutational targets, including genes associated with
121 metabolic pathways, such as the catabolism of a specific carbon source, but also regulatory
122 genes that influence a combination of growth aspects such as carbon source consumption and
123 better oxygen use. This strategy should also favor the selection of mutations with large
124 correlated phenotypic effects in alternative environments. Every day (24 +/-2 hours),
125 populations were diluted 300-fold into the same fresh medium, allowing ~8.2 [$\log_2$ (300)]
126 generations per day. Populations therefore experience every day a lag phase with no
127 detectable growth, followed by an exponential phase characterized by the optimal
128 consumption of the available carbon source, and finally a stationary phase when the carbon
129 source has been exhausted (except in the Ace environment, see below), before being diluted
130 again in fresh medium. The bottleneck at each daily transfer was never fewer than ~4×10^6
131 cells in any environment. After 1000 generations, each of the 16 populations was streaked on
132 LB agar plates that were incubated overnight at 37 °C. A single colony was randomly chosen
133 from each plate and frozen at -80 °C as a glycerol suspension.
134
135 *Growth profile assays*
136 Growth profiles were measured first for population samples collected every 200 generations,
137 in triplicate and in their respective evolution environments, and second for the ancestor and

138 individual evolved clones with five-fold replication and in the four different environments.
139 For each experiment, a physiological pre-acclimation was performed in the assay
140 environment, consisting of an overnight culture of bacterial cells followed by a 300-fold
141 dilution and a 24-h incubation. This pre-acclimation step differed slightly for the growth
142 profiles in the Ace environment, in which populations that evolved in the three other
143 environments, as well as the ancestor, grew slowly and were still in exponential phase after 24
144 hours of incubation. In this particular case, the initial overnight cultures were performed in
145 Glu. For all experiments and after a 300-fold dilution, pre-acclimated cultures were incubated
146 in the relevant assay environment.
147 Growth profiles were determined by measuring the optical density for each culture at 600 nm
148 (OD_{600}) at regular intervals during 24 h of incubation. We used the resulting growth curves to
149 calculate the maximum growth rate (μ_{max}) of each evolved culture (populations and individual
150 clones) relative to the ancestor. For the Gly, Glc and Glu environments, maximum growth
151 rates were measured between 0.2 and 0.8 of the maximal ancestral OD_{600}. In the Ace
152 environment, the ancestor was still in exponential phase after 24 h of culture, and growth rates
153 were measured between 0.5 and 1 times the OD_{600} reached by the ancestor following 24 h of
154 growth.
155
156 *Fitness assays*
157 Fitness assays were performed for the populations sampled at 300, 600 and 1000 generations
158 in their respective evolution environments, and for the individual evolved clones in all four
159 environments. All assays were replicated five-fold. Competitions were performed as
160 previously described (Lenski *et al*. 1991). Briefly, all competitors including the REL606
161 ancestor and a marked phenotypic variant called REL607 were pre-acclimated in the assay
162 environment. The ancestor REL606 and all derived evolved clones are unable to use

163 arabinose as a carbon source (Ara$^-$), while REL607 is a REL606 spontaneous revertant that
164 recovered this catabolic ability (Ara$^+$). After pre-acclimation, each evolved sample
165 (population and individual clones) and the REL606 ancestral strain as a control were mixed
166 separately with REL607 at a 1:1 ratio. Mixtures were then diluted 300-fold in fresh medium
167 and incubated for 24 h at 37 °C in the assay environment. At days 0 (when the two
168 competitors are mixed) and 1 (after 24 h of incubation) of each competition experiment, cells
169 were diluted and plated on indicator tetrazolium arabinose (TA) plates, on which Ara$^+$ or Ara$^-$
170 colonies appear pink or red, respectively (Lenski *et al*. 1991). Plates were incubated 24 h at 37
171 °C and each of the competitors was scored. Using the initial and final cell counts we
172 calculated the realized (net) population growth of each competitor, according to the following
173 formula: $G_i = \ln(C_{t1}*300/C_{t0})$, where C_{t0} and C_{t1} are the number of colonies at the beginning
174 and after 24 hours of competition, respectively, and 300 the 300-fold higher dilution factor
175 required for C_{t1} compared to C_{t0}. The fitness of one competitor relative to the other was then
176 calculated as the ratio of their net growth rates during the competition experiment according
177 to the formula: Fitness = G_{Ara^-}/G_{Ara^+}, where G_{Ara^-} and G_{Ara^+} are the realized population
178 growth of the Ara$^-$ ancestor and evolved clones and of the Ara$^+$ REL607 clone, respectively
179 (Lenski *et al.* 1991).
180

181 *Genome sequencing*

182 The genome of each of the 16 clones isolated after 1000 generations of evolution was re-
183 sequenced on the Illumina HiSeq2000 platform (GATC Biotech, Germany) using one lane of
184 single-end 35-bp reads. Barcodes were used for each genome so that reads were clone-
185 specific. Candidate point mutations were identified in comparison to the ancestral genome of
186 REL606 (Jeong *et al*. 2009) using the SNiPer pipeline (Wielgoss *et al*. 2011). The existence
187 of a mutational event was inferred when it was identified in more than 75% of the reads

188 covering a given site and in more than 20 reads. We identified a total of 54 mutations of
189 which 25 were randomly chosen and confirmed by PCR and sequencing of the obtained
190 products. Short read single-end re-sequencing reliably detects SNPs but may miss large indels
191 and mutational events associated with mobile genetic elements (recombination and
192 transposition). To minimize this potential bias, we checked the eight genes that had a
193 mutation in more than one clone, *i.e.* mutated in parallel, for gene size polymorphism by PCR
194 amplification. Using this approach, we identified two additional mutations in the *glpR* gene:
195 one IS*1* insertion 30 bp upstream of its start codon and one 125-bp deletion in its coding
196 region. Re-sequencing the genome of one single evolved clone sampled from each population
197 allowed us to directly link phenotypic changes in different environments to specific
198 combinations of mutations. This approach is not appropriate for analyzing genomic evolution
199 at the level of the entire populations which is beyond the scope of our study and would imply
200 re-sequencing genomes from either multiple evolved clones or mixtures of populations
201 (Barrick & Lenski 2009; Herron & Doebeli 2013).
202
203 *Statistical analyses*
204 As the same global regulator-encoding genes were mutated in more than one environment, we
205 used a generalized linear model (GLM) to test if the correlated responses were more similar
206 for clones that shared mutated genes and/or that evolved in the same environments. To build
207 the model, we measured the variability of the phenotypes (average μ_{max} and fitness values)
208 between clones in their alternative environments. For pairs of clones that evolved in the same
209 environment we considered the three alternative environments, while for those that evolved in
210 two different environments we considered the two shared alternative ones. The variability of
211 the correlated responses was calculated as the mean squares among clones (MSS), as for an
212 ANOVA, but we did not compare the MSS to the mean squares error (MSE) using F statistics.

213 Instead, MSS were ranked within each alternative environment and used as an ordinal
214 response variable in the GLM. By doing this, we could consider as the units of interest the
215 pairs of clones instead of the clones themselves, allowing for the choice of two explanatory
216 variables. The first distinguished the pairs of clones sharing or not a mutated gene (nominal
217 explanatory variable) and the second the pairs of clones that evolved or not in the same
218 environment (nominal explanatory variable). Details of the input data used for the GLM are
219 given in a Dryad file (doi:10.5061/dryad.n2582). Likelihood ratio χ^2 values were used to test
220 for significance of ordinal logistic regressions as implemented in JMP version 3.1.5 (SAS
221 Institute).
222

223 Results

224 *Phenotypic changes in the evolution environments*

225 Four replicate populations of *E. coli* B were propagated for 1000 generations in each of four
226 evolution environments that differed in carbon source, oxygenation and degree of spatial
227 structure. We quantified the phenotypic evolution of each population in its respective
228 evolution environment by measuring its maximum growth rate relative to the ancestor at 200-
229 generation intervals (μ_{max}; Fig. 1a), and its fitness relative to the ancestor at 300, 600 and
230 1000 generations (Fig. 1b). Compared to their common ancestor all populations increased in
231 fitness and all except those evolved in Glu increased their μ_{max} after 1000 generations,
232 indicating that they adapted to their environments (Table S1, Supporting information, Fig. 1).
233 After 1000 generations of evolution, we isolated one evolved clone from each replicate
234 population and measured μ_{max} and fitness in its evolution environment (Table S2, Supporting
235 information, Fig. 2). Phenotypic evolution of the clones and their source populations was
236 correlated (μ_{max}: Spearman Rho = 0.93, $p < 0.0001$; fitness: Spearman Rho = 0.60, $p =$
237 0.0144). However, the clones isolated from the Glu environment tended to display a higher

238 μ_{max} than their entire source populations, indicating that they probably belong to specific sub-
239 lineages that may not be representative of the entire populations. This has however no impact
240 on our results since our main goal here is to relate phenotypes to specific combinations of
241 mutations, a task that could not be achieved by focusing on entire populations. Therefore, all
242 the following analyses focus on comparisons between phenotypes and genomes of clones and
243 *not* populations.
244 Clones sampled from the four environments had quantitative differences in their divergence
245 from the ancestor (Kruskal–Wallis one-way analysis of variance: μ_{max}, $\chi^2 = 12.71$, $p = 0.0053$;
246 fitness, $\chi^2 = 11.59$, $p = 0.0089$), indicating different magnitudes of phenotypic evolution in the
247 four environments (Table S2, Supporting information, Fig. 2).
248
249 *Genome sequencing*
250 Re-sequencing the genomes of the 16 evolved clones revealed a total of 54 mutations
251 compared to the ancestor, with one to six mutations for each clone. These mutations affected a
252 total of 25 genes and included 53 SNPs and one 1-bp insertion (Table 1). Most SNPs (51)
253 occurred within genes, including 49 non-synonymous and only 2 synonymous changes. Only
254 three mutations affected intergenic regions, including two SNPs and the 1-bp insertion. Two
255 additional indels were subsequently identified (see Material and methods), one IS*1* insertion
256 and one 125-bp deletion, both affecting the *glpR* gene (Table 1).
257
258 *Genetic parallelism within and between evolution environments*
259 Defining genetic parallelism as a mutational change occurring in identical genes, operons or
260 functional units (Tenaillon *et al.* 2012) in at least two clones, ~60 % (34/56) of the identified
261 mutations occurred in parallel (Fig. 3, Table 1). The proportion of parallel to total mutations
262 was 7/12, 13/19, 9/16, and 6/9 in the Ace, Gly, Glc, and Glu environments, respectively. This

263 parallelism involved eight genes or operons (Fig. 3, Table 1), including four specific to a
264 single environment (*mreBC* in Ace, *glpR* and *glpK* in Gly, and *lldR* in Glc), and four across
265 multiple environments (*argR*, *spoT*, *rho*, and *nadR*). In all cases genes that changed in parallel
266 within a single environment reflected a greater degree of environmental clustering than
267 expected if the mutations were distributed randomly over the evolved clones (Fisher's exact
268 test, for all genes mutated in parallel $p \leq 0.05$). These data are consistent with the changes in
269 these genes conferring environment-specific adaptations.
270 The four evolved clones sampled from the Ace environment had a total of 12 mutations
271 affecting nine different genes that can be grouped into eight loci. Three of these loci were
272 targeted by mutations in more than one clone: the *mreBC* genetic locus, *mreB* and *mreC* being
273 part of the same operon (Wachi *et al*. 2006), had mutations in two clones, while *rho* and *argR*
274 were mutated in three and two clones, respectively (Fig. 3, Table 1). The four clones sampled
275 from the Gly environment had a total of 19 mutations affecting ten different genes. Seven
276 mutations occurred in *glpR* and *glpK*, which were mutated in four and three clones,
277 respectively, while six occurred in *spoT* and *rho*, which were mutated in four and two clones,
278 respectively (Fig. 3, Table 1). The four evolved clones sampled from the Glc environment had
279 a total of 16 mutations affecting ten different genes, among which three (*argR*, *lldR* and *spoT*)
280 had mutations in three clones (Fig. 3, Table 1). Finally, the four evolved clones sampled from
281 the Glu environment had 9 mutations affecting a total of 5 different genes including *spoT* and
282 *nadR* that were mutated in four and two clones, respectively (Fig. 3, Table 1). Despite the
283 high overall level of genetic parallelism, the number of parallel changes found in individual
284 clones was quite variable (Fig. 3, Table 1), ranging from one (Ace_4, Glc_2 and Glu_2) to
285 four (Gly_3, Glc_4, Glu_1 and Glu_4).
286

287 Four genes were mutated repeatedly in different environments—*argR* (Ace and Glc
288 environments), *spoT* (Gly, Glc and Glu environments), *rho* (Ace and Gly environments), and
289 *nadR* (Gly and Glu). Therefore, evolved clones from all pairs of environments, except
290 Ace/Glu, shared mutated genes, but the combinations of shared mutated genes were
291 environment-specific (Fig. 3, Table 1). The highest level of genetic parallelism was detected
292 for *spoT*, which was modified in 3/4 environments and 11/16 clones, further emphasizing
293 previous observations of repeated *spoT* mutations (see below; Cooper *et al.* 2003; Cooper
294 2007; Ostrowski *et al.* 2008; Woods *et al.* 2011).
295
296 *Functions of mutated genes and distribution of mutations within repeatedly modified genes*
297 We investigated the function of the mutated genes by considering their Gene Ontology
298 categories (Riley *et al*. 2006). Seven (*rpoA*, *rho*, *argR*, *lldR*, *glpR*, *flhD*, and *nadR*) of the 25
299 mutated genes are categorized as transcription regulators, which is more than expected by
300 chance alone (Fisher one tail exact test $p = 0.01$). Although not classified in the GO
301 transcription category, other mutated genes are also involved in transcription regulation: *fabR*
302 encodes a transcription regulator, and *relA* and *spoT* are involved in the metabolism of
303 ppGpp, which reprograms the entire transcription machinery in bacterial cells (Srivatsan &
304 Wang 2008). Therefore, 10 of the 25 genetic loci that were modified during evolution are
305 involved in transcription regulation. Strikingly, six of them (*rho*, *argR*, *lldR*, *glpR*, *nadR*, *and*
306 *spoT*) are among the eight that were mutated in parallel, emphasizing the importance of
307 changes in transcription regulation for adaptation.
308 Two sub-groups can be distinguished among these six genes. The first comprises *glpR* and
309 *lldR*, which are local transcription regulators (Zeng *et al*. 1996; Aguilera *et al*. 2008) affecting
310 the expression of a small number of genes involved in the consumption of a single carbon
311 source (glycerol or lactate, respectively). Parallel changes occurred in these two genes only

312 within specific environments (Gly or Glc, respectively). The second sub-group comprises
313 *argR*, which regulates a large set of target genes (Caldara *et al*. 2006), *spoT* and *rho*, which
314 are involved in global gene regulation (Srivatsan & Wang 2008; Epshtein *et al*. 2010), and
315 *nadR,* which regulates global cell physiology by integrating multiple environmental signals to
316 control the synthesis of NAD, a central co-factor in many bacterial metabolic pathways
317 (Grose *et al*. 2005). Interestingly, it is precisely this second group of transcription regulator-
318 encoding genes that had parallel changes between, as well as within, environments.
319 The genetic parallelism observed at the gene level was associated with mutations affecting
320 different codons largely distributed within these genes, with however two notable exceptions
321 (Table 1). First, specific regions of the same genes were affected by independent mutations,
322 including the 5' end of *argR* and *glpK*. Second, nearby codons were affected by independent
323 non-synonymous mutations in *rho* (codons 322 and 324), and *spoT* (codons 207 and 209). In
324 addition, the exact same codons of *rho* (codons 322 and 324) and *spoT* (codon 393) were
325 repeatedly affected by independent mutations, leading to either the same or different amino-
326 acid changes in the corresponding proteins (Table 1). For both exceptions, similar mutations
327 occurred in more than one environment.
328
329 *Correlated phenotypic responses to evolution in the alternative environments*
330 Next, we investigated how adaptation to a given evolution environment affected μ_{max} and
331 fitness in alternative conditions. The correlated response to selection of each of the 16 evolved
332 clones was compared to the common ancestor in each of the three environments in which it
333 did not evolve (hereafter called alternative environments). As an example, the four evolved
334 clones sampled from the populations that were propagated in Ace were assayed in the
335 alternative environments Gly, Glc and Glu.

336 Growth profiles and competition assays indicated that adaptation to a given environment was
337 often associated with improvements in the alternative environments (Table S2, Supporting
338 information, Fig. 4). Phenotypic evolution in a given environment is constrained by selective
339 pressures that need not to apply in alternative environments. Therefore, we expected the
340 correlated phenotypic responses in the alternative environments to be more variable than the
341 direct response in the evolution environment. Surprisingly, this was not the case (Wilcoxon
342 rank-sum test on the coefficient of variation; μ_{max}: $\chi^2 = 1.9$, $p = 0.17$; fitness: $\chi^2 = 0.2$, $p =$
343 0.69; Table S3, Supporting information).
344 Next, we tested whether genetic parallelism (*i.e.,* whether clones shared or not mutated genes)
345 and evolutionary history (*i.e.,* whether clones evolved or not in the same evolution
346 environment) affected the correlated responses. Correlated responses to evolution may be
347 expected to be more similar for clones with mutations in identical genes, even if they evolved
348 in different environments, as long as the phenotypic effects of individual mutations affecting a
349 given gene are similar and epistatic interactions between mutations in the different genes do
350 not entirely mask the effect of individual mutations. Similarly, correlated responses to
351 evolution may be expected to be more similar for clones that evolved in identical
352 environments, even if they had mutations in different genes, if the evolved mutational
353 pathways resulted in phenotypic parallelism not only in the evolution environment but also
354 under a wider range of environments. We measured the variability of the correlated responses
355 in each alternative environment for all pairs of clones and used a GLM to distinguish the
356 contribution of genetic parallelism and evolutionary history to the correlated responses to
357 evolution. For pairs of clones that evolved in the same evolution environment we considered
358 the three alternative environments, while for pairs of clones that evolved in two different
359 evolution environments we considered the two shared alternative environments. We tested
360 whether sharing or not mutated genes and evolving or not in identical evolution environments

361 influenced the phenotypic variability between clone pairs (Table 2). The interaction between
362 the two factors was significant for both μ_{max} and fitness phenotypes (Table 2), indicating that
363 the impact of genetic parallelism on phenotypic variability was different whether the clones
364 evolved in identical or different evolution environments. The variability of the correlated
365 responses was lower for pairs of clones that evolved in the same environment (Fig. 5).
366 Moreover, clones with different evolutionary history had more similar correlated responses
367 when they shared mutated genes (Fig. 5a, b). However, for clones that evolved in identical
368 evolution environments, the variability of correlated responses was unrelated to the level of
369 genetic parallelism (Fig. 5c, d).
370

371 Discussion

372 We propagated four replicate populations of *E. coli* B in each of four different environments
373 for 1000 generations. During this time all populations adapted to their evolution environment.
374 Adaptation involved a high level of genetic parallelism, and 60 % of the identified mutations
375 affected eight genes that were modified in parallel in at least two populations. Four of these
376 genes, *argR*, *spoT*, *rho*, and *nadR*, were modified repeatedly across environments, although
377 no single gene was mutated in all environments. This result suggests distinct, but overlapping,
378 selective pressures in the different environments. Moreover, the combination of mutated
379 genes was different in the evolved clones sampled from the various environments. Finally, we
380 found that evolution in each environment was associated with improved correlated phenotypic
381 responses in alternative environments and that evolved clones from different evolution
382 environments revealed a higher level of parallel correlated responses when displaying genetic
383 parallelism.
384

385 Evolution was associated with parallel genetic changes in genes specifically involved in the

386 catabolism of the available carbon source (*glpR* and *glpK*) in only one of the four

387 environments (Gly). Adaptive evolution may therefore involve improvements of substrate

388 specific catabolic pathways, but this not a general trend. Consistent with this, other evolution

389 experiments have found mutations in genes specifically involved in the catabolic pathways of

390 the available carbon sources in populations propagated in the presence of glycerol (Herring *et*

391 *al*. 2006) and L-1,2-propanediol (Lee & Palsson 2010), but not in the presence of lactate

392 (Conrad *et al*. 2009) or glucose (Barrick *et al*. 2009). We found that genes involved in global

393 transcription regulation were targets of mutations in all four evolution environments, as

394 observed in most evolution experiments (Cooper *et al*. 2003; Crozat *et al*. 2005; Herring et al.

395 2006; Bantinaki *et al*. 2007; Conrad *et al*. 2009; Le Gac & Doebeli 2010; Maharjan *et al*.

396 2010; Wang *et al*. 2010; Yu *et al*. 2010; Conrad *et al*. 2011; Tenaillon *et al*. 2012), reflecting

397 the importance of restructuring regulatory networks during evolution (Philippe *et al.* 2007;

398 Hindré *et al.* 2012). Even more strikingly, all genes that were mutated in more than one

399 environment encoded global regulators of gene transcription. Six of the eight genetic loci that

400 were repeatedly changed have also been modified by mutations in other evolution

401 experiments, including *glpK, glpR, spoT*, *nadR*, *mreBC* and *rho* (Cooper *et al*. 2003;

402 Raghunathan & Palsson 2003; Herring *et al*. 2006; Woods *et al*. 2006; Ostrowski *et al*. 2008;

403 Woods *et al*. 2011; Tenaillon *et al*. 2012; Herron and Doebeli 2013).

404

405 Genetic parallelism is commonly observed among replicate independent populations that are

406 propagated in similar environments. The nature of the mutations affecting target genes is

407 diverse, from identical nucleotide changes being repeatedly substituted in independent viral

408 populations (Wichman *et al*. 1999) to different types of mutations distributed at various

409 positions within a given gene in replicate bacterial populations (Woods *et al*. 2006; Bantinaki

410 *et al*. 2007; Crozat *et al*. 2010). A recent analysis of 115 replicate populations of *E. coli* B
411 adapted to high temperature revealed few identical mutations among populations, although
412 high genetic parallelism at the gene level was detected (Tenaillon *et al*. 2012). In our study,
413 genetic parallelism at the gene level was associated with different mutations occurring at
414 different positions within genes. For example, *glpR* was affected by two different non-
415 synonymous mutations, a 125-bp deletion in its coding region, and the insertion of an IS*1*
416 element in its promoter region. This profile suggests that the mutations inactivated *glpR*.
417 However, there were two notable exceptions to this pattern. First, different mutations occurred
418 in specific regions of two of the altered genes. All *argR* mutations were localized in the 5'
419 gene region encoding the DNA-binding domain of ArgR (Sunnerhagen *et al*. 1997) and all
420 *glpK* mutations mapped in a ~80-bp region at the 5' end of the gene. Mutations in this *glpK*
421 region have been shown to increase fitness on glycerol-containing medium by reducing the
422 affinity of GlpK for its allosteric inhibitor fructose-1,6-bisphosphate and inhibiting the
423 formation of the GlpK tetramer (Applebee *et al*. 2011). Based on these observations, we
424 hypothesize that mutations in *argR* and *glpK* modulated specific activities of the encoded
425 proteins. Second, we detected an even higher level of specificity for the mutations that
426 occurred in *rho* and *spoT*, which were repeatedly modified both within and between
427 environments. Hence, not only were independently substituted mutations localized to nearby
428 codons (322 and 324 in *rho*, 207 and 209 in *spoT*), but identical codons were changed in
429 clones evolved in different environments, including codons 322 and 324 for *rho* in Ace and
430 Gly and codon 393 for *spoT* in Gly and Glu. In the Gly environment, codon 393 of *spoT* was
431 changed in all four clones, and in 3 out of 4 in Glu. Cross contamination between replicate
432 populations can be excluded since four different SNPs affected codon 393. Such an extreme
433 level of parallelism, extending to the level of the substituted mutations, has also been reported
434 in bacterial evolution experiments for *glpK* (Applebee et al. 2011), *rho* (Tenaillon *et al*.

435 2012), and genes that are the targets for antibiotic resistance (Martinez *et al*. 2011). Therefore,
436 depending on the environment, selection may act either at precise identical codons (in Gly and
437 Glu), leading or not to identical amino-acid changes, or at the level of the entire gene (Glc).
438 In our evolution experiment, *spoT* was the most commonly mutated gene, being modified in
439 3/4 environments and 11/16 clones. It has already been observed as being frequently modified
440 during adaptation to glucose environments. In a commonly-used glucose environment, *spoT*
441 has been affected by a total of 32 non-synonymous mutations, 28 of which involved a unique
442 amino-acid substitution (Cooper *et al.* 2003; Cooper 2007; Ostrowski *et al.* 2008; Woods *et*
443 *al.* 2011). In contrast, in two environments in this study, seven out of eight clones had
444 mutations affecting the exact same *spoT* codon. This suggests that the number of potential
445 beneficial mutations in a given gene may vary from one environment to the other. So far, very
446 little is known about the variability of phenotypic effects associated with different beneficial
447 mutations affecting the same gene. Such studies would undoubtedly increase our
448 comprehension of the relationships between genotypes, phenotypes and fitness.
449
450 We have shown that populations evolving in identical evolution environments may display
451 parallel phenotypic evolution in alternative environments, as was previously described
452 (Cooper & Lenski 2000; Barrett *et al.* 2005; Fong *et al*. 2005). Moreover, we showed that the
453 correlated phenotypic responses were related to both the evolutionary history of the sampled
454 clones, *i.e.* whether or not they evolved in identical evolution environments, and the genetic
455 parallelism, *i.e.* whether or not they shared mutated genes. Our results therefore have two
456 implications. First, individuals that have been subjected to similar historical selective
457 pressures tend to occupy similar ecological niches not only in the environment in which they
458 evolved, but also under a wider range of conditions, even without genetic parallelism. Second,
459 when evolutionary history was different, *i.e.* for clones that evolved in different evolution

460 environments, correlated phenotypic responses were related to the level of genetic parallelism.
461 It therefore suggests that the mutations of the repeatedly modified genes had similar beneficial
462 effects even in different genomic contexts. This finding indicates that although epistatic
463 interactions between mutations are known to considerably alter the effect of individual
464 mutations, they evidently did not entirely mask the individual effects of the mutations under
465 our environmental conditions.
466 In summary, notwithstanding the relatively small number of genes mutated in more than one
467 environment, our results provide a way to experimentally disentangle the relative contribution
468 of evolutionary history and genetic parallelism on the phenotypic performance in naïve
469 environmental conditions. Taken together, our results indicate that populations sharing
470 historical selective pressures could react similarly to the modification of environmental
471 conditions, even if they do not share mutations in specific genes. To a lesser extent, they also
472 indicate that populations with different evolutionary histories may perform more similarly
473 when facing a new environment if they share mutated genes.
474

475 Acknowledgements

476 This research was supported by the French Centre National de la Recherche Scientifique
477 (CNRS), the University Joseph Fourier Grenoble and by a grant from the Agence Nationale
478 de la Recherche (ANR, Program Blanc, Grant ANR-08-BLAN-0283-01 to D. Schneider) and
479 the NSF (DEB0844355 to T. F. Cooper). M. Le Gac thanks the ANR for a post-doctoral
480 fellowship.
481

482 # References

483 Aguilera L, Campos E, Gimenez R, *et al*. (2008) Dual role of LldR in regulation of the
484 *lldPRD* operon, involved in L-lactate metabolism in *Escherichia coli*. *Journal of*
485 *Bacteriology*, **190**, 2997-3005.

486 Almahmoud I, Kay E, Schneider D, Maurin M (2009) Mutational paths towards increased
487 fluoroquinolone-resistance in *Legionella pneumophila*. *Journal of Antimicrobial*
488 *Chemotherapy*, **64**, 284-293.

489 Applebee MK, Joyce AR, Conrad TM, Pettigrew DW, Palsson BO (2011) Functional and
490 metabolic effects of adaptive glycerol kinase (GLPK) mutants in *Escherichia coli*. *Journal*
491 *of Biological Chemistry*, **286**, 23150-23159.

492 Bantinaki E, Kassen R, Knight CG, *et al*. (2007) Adaptive divergence in experimental
493 populations of *Pseudomonas fluorescens*. III. Mutational origins of wrinkly spreader
494 diversity. *Genetics*, **176**, 441-453.

495 Barrett RDH, MacLean RC, Bell G (2005) Experimental evolution of *Pseudomonas*
496 *fluorescens* in simple and complex environments. *American Naturalist*, **166**, 470-480.

497 Barrick JE, Lenski RE (2009) Genome-wide mutational diversity in an evolving population of
498 *Escherichia coli*. *Cold Spring Harbor Symposium Quantitative Biology*, **74**, 119-129.

499 Barrick JE, Yu DS, Yoon SH, *et al*. (2009) Genome evolution and adaptation in a long-term
500 experiment with *Escherichia coli*. *Nature*, **461**, 1243-1247.

501 Brockhurst MA, Colegrave N, Rozen DE (2011) Next-generation sequencing as a tool to
502 study microbial evolution. *Molecular Ecology*, **20**, 972-980.

503 Bull JJ, Badgett MR, Wichman HA (2000) Big-benefit mutations in a bacteriophage inhibited
504 with heat. *Molecular Biology and Evolution*, **17**, 942-950.

505 Caldara M, Charlier D, Cunin R (2006) The arginine regulon of *Escherichia coli*: whole-
506 system transcriptome analysis discovers new genes and provides an integrated view of
507 arginine regulation. *Microbiology*, **152**, 3343-3354.
508 Conrad TM, Joyce AR, Applebee MK, *et al*. (2009) Whole-genome resequencing of
509 *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution in lactate
510 minimal media reveals flexible selection of adaptive mutations. *Genome Biology*, **10**, R118.
511 Conrad TM, Lewis NE, Palsson BO (2011) Microbial laboratory evolution in the era of
512 genome-scale science. *Molecular and Systems Biology*, **7**, 509.
513 Cooper TF, Remold SK, Lenski RE, Schneider D (2008) Expression profiles reveal parallel
514 evolution of epistatic interactions involving the CRP regulon in *Escherichia coli*. *PLoS*
515 *Genetics*, **4(2)**, e35.
516 Cooper TF, Rozen DE, Lenski RE (2003) Parallel changes in gene expression after 20,000
517 generations of evolution in *Escherichia coli*. *Proceedings of the National Academy of*
518 *Sciences*, *USA*, **100**, 1072-1077.
519 Cooper TF (2007) Recombination speeds adaptation by reducing competition between
520 beneficial mutations in populations of *Escherichia coli*. *PLoS Biology*, **5(9)**, e225.
521 Cooper VS, Lenski RE (2000) The population genetics of ecological specialization in
522 evolving *Escherichia coli* populations. *Nature*, **407**, 736-739.
523 Crozat E, Philippe N, Lenski RE, Geiselmann J, Schneider D (2005) Long-term experimental
524 evolution in *Escherichia coli*. XII. DNA topology as a key target of selection. *Genetics*,
525 **169**, 523-532.
526 Crozat E, Winkworth C, Gaffé J, *et al*. (2010) Parallel genetic and phenotypic evolution of
527 DNA superhelicity in experimental populations of *Escherichia coli*. *Molecular Biology and*
528 *Evolution*, **27**, 2113-2128.

529 Dettman JR, Rodrigue N, Melnyk AH, Wong A, Bailey SF, Kassen R (2012) Evolutionary
530 insight from whole-genome sequencing of experimentally evolved microbes. *Molecular*
531 *Ecology*, **21**, 2058-2077.
532 Epshtein V, Dutta D, Wade J, Nudler E (2010) An allosteric mechanism of Rho-dependent
533 transcription termination. *Nature*, **463**, 245-249.
534 Fong SS, Joyce AR, Palsson BO (2005) Parallel adaptive evolution cultures of *Escherichia*
535 *coli* lead to convergent growth phenotypes with different gene expression states. *Genome*
536 *Research*, **15**, 1365-1372.
537 Grose JH, Bergthorsson U, Roth JR (2005) Regulation of NAD biosynthesis by the
538 trifunctional NadR protein of *Salmonella enterica*. *Journal of Bacteriology*, **187**, 2774-
539 2782.
540 Herring CD, Raghunathan A, Honisch C, *et al*. (2006) Comparative genome sequencing of
541 *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nature*
542 *Genetics*, **38**, 1406-1412.
543 Herron MD, Doebeli M (2013) Parallel evolutionary dynamics of adaptive diversification in
544 *Escherichia coli*. *PLoS Biology*, **11**, e1001490.
545 Hindré T, Knibbe C, Beslon G, Schneider D (2012) New insights into bacterial adaptation
546 through *in vivo* and *in silico* experimental evolution. *Nature Reviews Microbiology*, **10**,
547 352-365.
548 Jeong H, Barbe V, Lee CH, *et al*. (2009) Genome sequences of *Escherichia coli* B strains
549 REL606 and BL21(DE3). *Journal of Molecular Biology*, **394**, 644-652.
550 Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF (2011) Negative epistasis between
551 beneficial mutations in an evolving bacterial population. *Science*, **332**, 1193-1196.
552 Lederberg S (1966) Genetics of host-controlled restriction and modification of
553 deoxyribonucleic acid in *Escherichia coli*. *Journal of Bacteriology*, **91**, 1029-1036.

554 Lee DH, Palsson BO (2010) Adaptive evolution of *Escherichia coli* K-12 MG1655 during
555 growth on a nonnative carbon source, L-1,2-propanediol. *Applied and Environmental*
556 *Microbiology*, **76**, 4158-4168.
557 Le Gac M, Doebeli M (2010) Epistasis and frequency dependence influence the fitness of an
558 adaptive mutation in a diversifying lineage. *Molecular Ecology*, **19**, 2430-2438.
559 Lenski RE, Rose MR, Simpson SC, Tadler SC (1991) Long-term experimental evolution in
560 *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *American*
561 *Naturalist*, **138**, 1315-1341.
562 MacLean RC, Bell G (2003) Divergent evolution during an experimental adaptive radiation.
563 *Proceedings of the Royal Society of London series B-Biological Sciences*, **270**, 1645-1650.
564 Maharjan R, Zhou Z, Ren Y, *et al*. (2010) Genomic identification of a novel mutation in *hfq*
565 that provides multiple benefits in evolving glucose-limited populations of *Escherichia coli*.
566 *Journal of Bacteriology*, **192**, 4517-4521.
567 Martinez JL, Baquero F, Andersson DI (2011) Beyond serial passages: new methods for
568 predicting the emergence of resistance to novel antibiotics. *Current Opinion in*
569 *Pharmacology*, **11**, 439-445.
570 Ostrowski EA, Rozen DE, Lenski RE (2005) Pleiotropic effects of beneficial mutations in
571 *Escherichia coli*. *Evolution*, **59**, 2343-2352.
572 Ostrowski EA, Woods RJ, Lenski RE (2008) The genetic basis of parallel and divergent
573 phenotypic responses in evolving populations of *Escherichia coli*. *Proceedings of the Royal*
574 *Society of London series B-Biological Sciences*, **275**, 277-284.
575 Philippe N, Crozat E, Lenski RE, Schneider D (2007) Evolution of global regulatory networks
576 during a long-term experiment with *Escherichia coli*. *BioEssays*, **29**, 846-860.
577 Poelwijk FJ, Kiviet DJ, Weinreich DM, Tans SJ (2007) Empirical fitness landscapes reveal
578 accessible evolutionary paths. *Nature*, **445**, 383-386.

579 Raghunathan A, Palsson BO (2003) Scalable method to determine mutations that occur during
580 adaptive evolution of *Escherichia coli*. *Biotechnology Letters*, **25**, 435-441.
581 Remold SK, Lenski RE (2004) Pervasive joint influence of epistasis and plasticity on
582 mutational effects in *Escherichia coli*. *Nature Genetics*, **36**, 423-426.
583 Riley M, Abe T, Arnaud MB, *et al.* (2006) *Escherichia coli* K-12: a cooperatively developed
584 annotation snapshot—2005. *Nucleic Acids Research*, **34**, 1-9.
585 Srivatsan A, Wang JD (2008) Control of bacterial transcription, translation and replication by
586 (p)ppGpp. *Current Opinion in Microbiology*, **11**, 100-105.
587 Sunnerhagen M, Nilges M, Otting G, Carey J (1997) Solution structure of the DNA-binding
588 domain and model for the complex of multifunctional hexameric arginine repressor with
589 DNA. *Nature Structural Biology*, **4**, 819-826.
590 Tenaillon O, Rodriguez-Verdugo A, Gaut RL, *et al*. (2012) The molecular diversity of
591 adaptive convergence. *Science*, **335**, 457-461.
592 Travisano M, Lenski RE (1996) Long-term experimental evolution in *Escherichia coli*.IV.
593 Targets of selection and the specificity of adaptation. *Genetics*, **143**, 15-26.
594 Travisano M, Vasi F, Lenski RE (1995) Long-term experimental evolution in *Escherichia*
595 *coli*. III. Variation among replicate populations in correlated responses to novel
596 environments. *Evolution*, **49**, 189-200.
597 Wachi M, Osaka K, Kohama T, *et al*. (2006) Transcriptional analysis of the *Escherichia coli*
598 *mreBCD* genes responsible for morphogenesis and chromosome segregation. *Bioscience,*
599 *Biotechnology, and Biochemistry*, **70**, 2712-2719.
600 Wang L, Spira B, Zhou Z, *et al*. (2010) Divergence involving global regulatory gene
601 mutations in an *Escherichia coli* population evolving under phosphate limitation. *Genome*
602 *Biology and Evolution*, **2**, 478-487.

603 Weinreich DM, Delaney NF, Depristo MA, Hartl DL (2006) Darwinian evolution can follow
604 only very few mutational paths to fitter proteins. *Science*, **312**, 111-114.
605 Weinreich DM, Watson RA, Chao L (2005) Perspective: Sign epistasis and genetic constraint
606 on evolutionary trajectories. *Evolution*, **59**, 1165-1174.
607 Wichman HA, Badgett MR, Scott LA, Boulianne CM, Bull JJ (1999) Different trajectories of
608 parallel evolution during viral adaptation. *Science*, **285**, 422-424.
609 Wielgoss S, Barrick JE, Tenaillon O, *et al*. (2011) Mutation rate inferred from synonymous
610 substitutions in a long-term evolution experiment with *Escherichia coli*. *G3: Genes,*
611 *Genomes, Genetics*, **1**, 183-186.
612 Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE (2006) Tests of parallel
613 molecular evolution in a long-term experiment with *Escherichia coli*. *Proceedings of the*
614 *National Academy of Sciences*, *USA*, **103**, 9107-9112.
615 Woods R, Barrick JE, Cooper TF, *et al*. (2011) Second-order selection for evolvability in a
616 large *Escherichia coli* population. *Science*, **331**, 1433-1436.
617 Yu YT, Yuan X, Velicer GJ (2010) Adaptive evolution of an sRNA that controls *Myxococcus*
618 development. *Science*, **328**, 993.
619 Zeng G, Ye S, Larson TJ (1996) Repressor for the sn-glycerol 3-phosphate regulon of
620 *Escherichia coli* K-12: primary structure and identification of the DNA-binding domain.
621 *Journal of Bacteriology*, **178**, 7080-7089.
622

623 **Data Accessibility:**

624 Genome sequencing data: ENA Sequence Read Archive (ERP001471)

625 Fitness, growth, and GLM data: DRYAD data identifier:doi:10.5061/dryad.n2582

626 **Supporting information**

627 Additional supporting information may be found in the online version of this article.

628 **Table S1** Phenotypic traits (μ_{max} and fitness) of the 16 populations during evolutionary time

629 in their respective evolution environments.

630 **Table S2** Phenotypic traits (μ_{max} and fitness) of the evolved clones sampled from each of the

631 16 populations in the four environments.

632 **Table S3** Among-clone coefficient of variation in the four environments for the two

633 phenotypic traits μ_{max} and fitness.

634

635 # Figure Legends

636 **Fig. 1** Phenotypic trajectories of 16 evolving populations during 1000 generations in their

637 respective evolution environments. Average μ_{max} (a) and fitness (b) (+/- SEM) are indicated

638 at, respectively, 200-generation intervals, and 300, 600 and 1000 generations for the replicate

639 populations that evolved in the environments Ace (-·○·-), Gly (—△—), Glc (┈□┈), and Glu

640 (- -◇- -).

641

642 **Fig. 2** Phenotypic traits of the evolved clones sampled from the replicate populations that

643 were propagated in the environments Ace (○), Gly (△), Glc (□), and Glu (◇). The fitness

644 and growth rate (μ_{max}) values were measured for one evolved clone, sampled from each

645 population after 1000 generations, in its respective evolution environment. Values were log-

646 transformed and given in the two-dimensional phenotypic space. Each experiment was

647 performed with five-fold replication and values are given +/-SEM. The phenotypic values for

648 the ancestor are (0;0).

649

650 **Fig. 3** Venn diagrams showing the mutated genes identified in the 16 evolved clones. Each of

651 the 16 ellipses represents one evolved clone sampled from each of the 16 populations. Red,

652 blue, green, and purple ellipses indicate the environments Ace, Gly, Glc, and Glu,

653 respectively. Overlaps indicate mutated genes shared between clones. Names of genes are
654 written in different colours whether they are affected by mutations in at least two different
655 environments (white), or in only one environment either in a single (black) or in different
656 clones in that environment (colour corresponding to the specific environment).
657
658 **Fig. 4** Phenotypic traits of the evolved clones sampled from each replicate population both in
659 their respective evolution environment (○) and in the three alternative ones (●). The μ_{max} (a
660 to d) and fitness (e to h) values (+/- SEM) of each evolved clone are given in each of the Ace,
661 Gly, Glc, and Glu environments (indicated from left to right on the x-axes). The evolution
662 environment from which each set of four clones was sampled is indicated at the top of each
663 panel: Ace (a, e), Gly (b, f), Glc (c, g), and Glu (d, h). Note that y-axis scales vary.
664
665 **Fig. 5** Among-clone phenotypic variability in the alternative environments. Each small hollow
666 symbol (○) corresponds to a comparison between a pair of clones in one of the alternative
667 environments. The y-axis shows the phenotypic variability (mean squares) ranked within each
668 alternative environment, a high rank value indicating that the pair of clones displays very
669 different phenotypes (due to tied ranks, not all of them are seen on the figure). Phenotypic
670 variability is given for μ_{max} (a, c) and fitness (b, d) values for pairs of clones that evolved in
671 different (a, b), and identical (c, d) environments. Median ranks (●) as well as first and third
672 quartiles (------) are indicated.
673

674 Tables

675 **Table 1** Mutations identified in the 16 evolved clones sampled after 1000 generations of
676 evolution in four different environments.

ECB	Gene name*	Ace				Gly				Glc				Glu			
		Ace_1	Ace_2	Ace_ 3	Ace_4	Gly_1	Gly_2	Gly_3	Gly_4	Glc_1	Glc_2	Glc_3	Glc_4	Glu_1	Glu_2	Glu_3	Glu_4
ECB_00549	*entD*						A55T										
ECB_00601	*dacA*										L285Q						
ECB_01027	*ycdT*									P162A							
ECB_01091	*fabF*															A280V	
ECB_01166	*cvrA*		A61S														
ECB_01453	*yddB*											F306L					
ECB_01477	*yneE*														A168T		
ECB_01609	*ydhA*								V98V								
ECB_01863	*flhD*											L21V					
ECB_02323	*cysW*								F191Y								
ECB_02629	*relA*														G318R		
ECB_02686/7	*lysA/lysR*				G/T												
ECB_02832						S138S											
ECB_03097	*argR*		R2L	K45T						A55S		K15N	N60D				
ECB_03109	*mreC*				P230S												
ECB_03110	*mreB*		S10P														
ECB_03146	*rpoA*				E273K												
ECB_03274	*glpR*					IS1	125bp del	R6C	I49N								
ECB_03462	*lldR*									V191A		R244C	W78C				
ECB_03507	*spoT*					P393S	P393A	P393L	P393L		R209H	T442P	R571L	P393T	P393L	P393L	G207D
ECB_03661	*rho*	G324C	D322Y	M219T			D322Y	G324C									
ECB_03811	*glpK*					Q38P		A55S	S32L								
ECB_03848	*fabR*									T30N							
ECB_03885/6	*metA/aceB*			+T													
ECB_04113	*pyrB*										A128V						
ECB_04120/1	*argI/yjgD*	C/A															
ECB_04132	*idnO*									T213P							
ECB_04235	*yjjN*								H62N								
ECB_04266	*nadR*						I211N							Q19*		P228L	

677

678 * The name of the genes carrying mutations are given, together with their ECB numbers

679 (Jeong *et al*. 2009). For intergenic mutations, the name of the two surrounding genes is given.

680

681 **Table 2** Variability of the correlated phenotypic responses for each pair of clones in the

682 alternative environments as a function of genetic parallelism (shared mutated genes or not)

683 and evolutionary history (same evolution environment or not). Likelihood ratio χ^2 tests on

684 ordinal logistic regressions (GLM, Material and methods) were performed with the among-

685 clone variability (Mean square among clones) of the μ_{max} and fitness values (ranked within

686 each alternative environment) taken as ordinal response variables. Evolution in identical

687 environments and level of genetic parallelism among clones were taken as explanatory

688 variables.

689

Source	DF	χ^2	*P*-value
μ_{max}			
Whole model*	3	73.07	<0.0001
Evolutionary history	1	33.88	<0.0001

Genetic parallelism	1	3.55	0.0596
Environment X Parallelism	1	14.61	0.0001
Fitness			
Whole model*	3	24.56	<0.0001
Evolutionary history	1	14.60	0.0001
Genetic parallelism	1	0.08	0.7809
Environment X Parallelism	1	4.60	0.0320

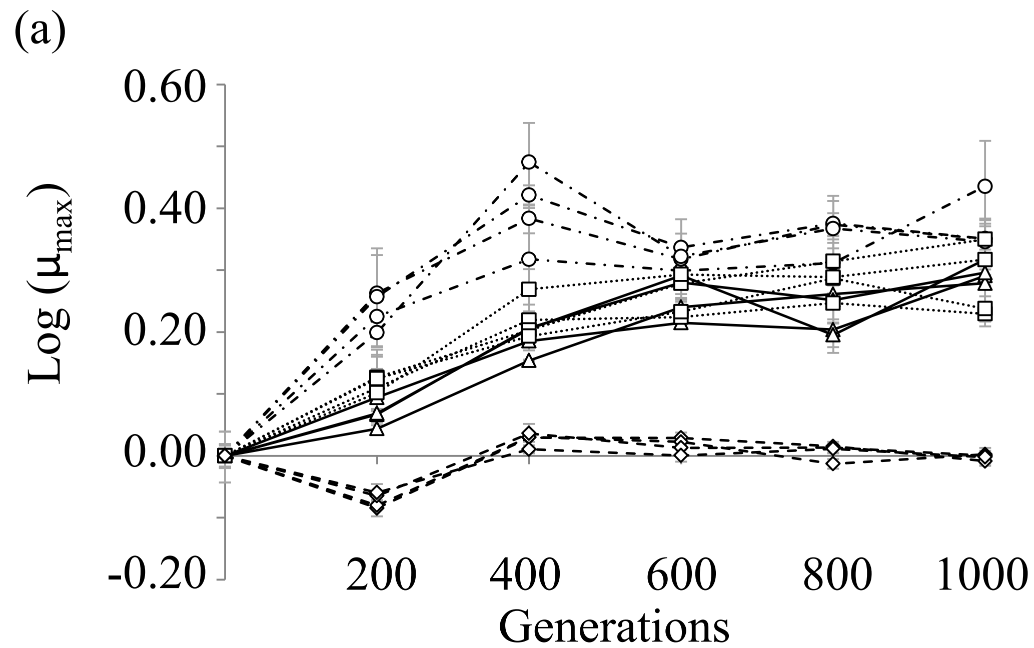
690 * Indicates whether the model based on the two explanatory variables as well as their

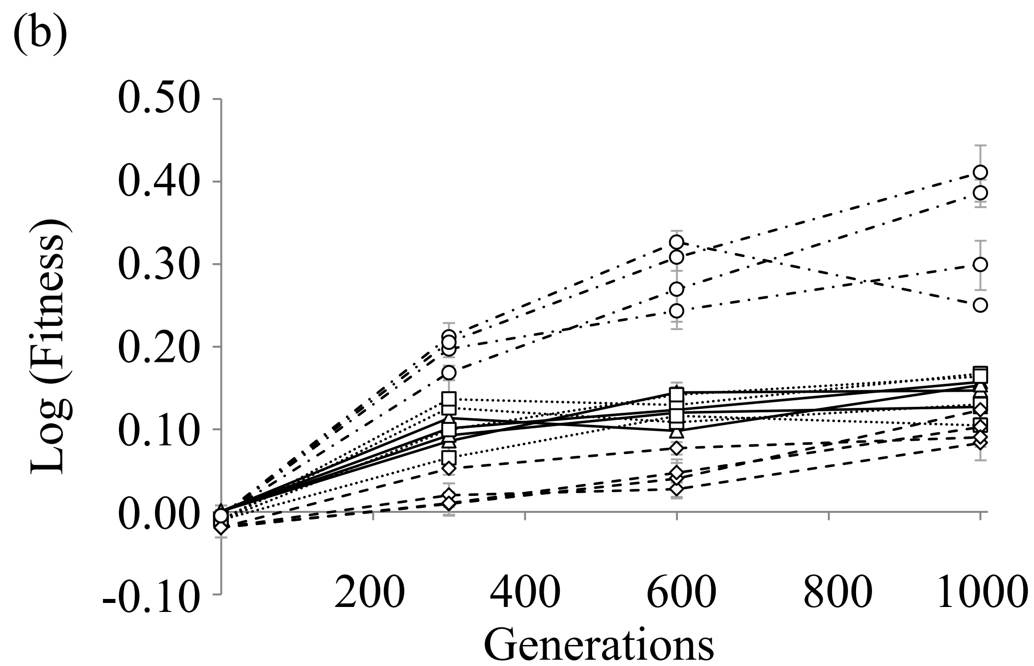
691 interaction significantly explained the among clone variability.

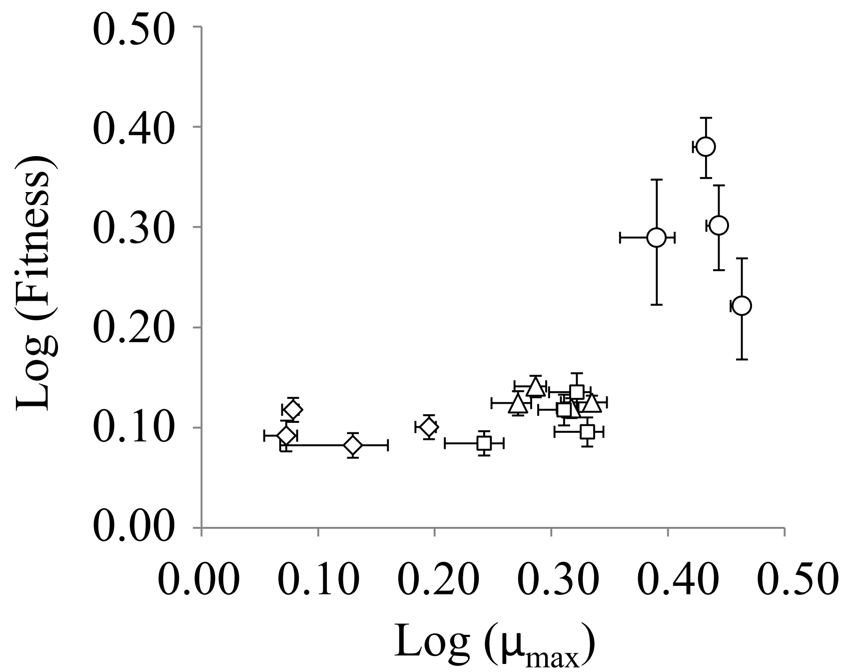
692

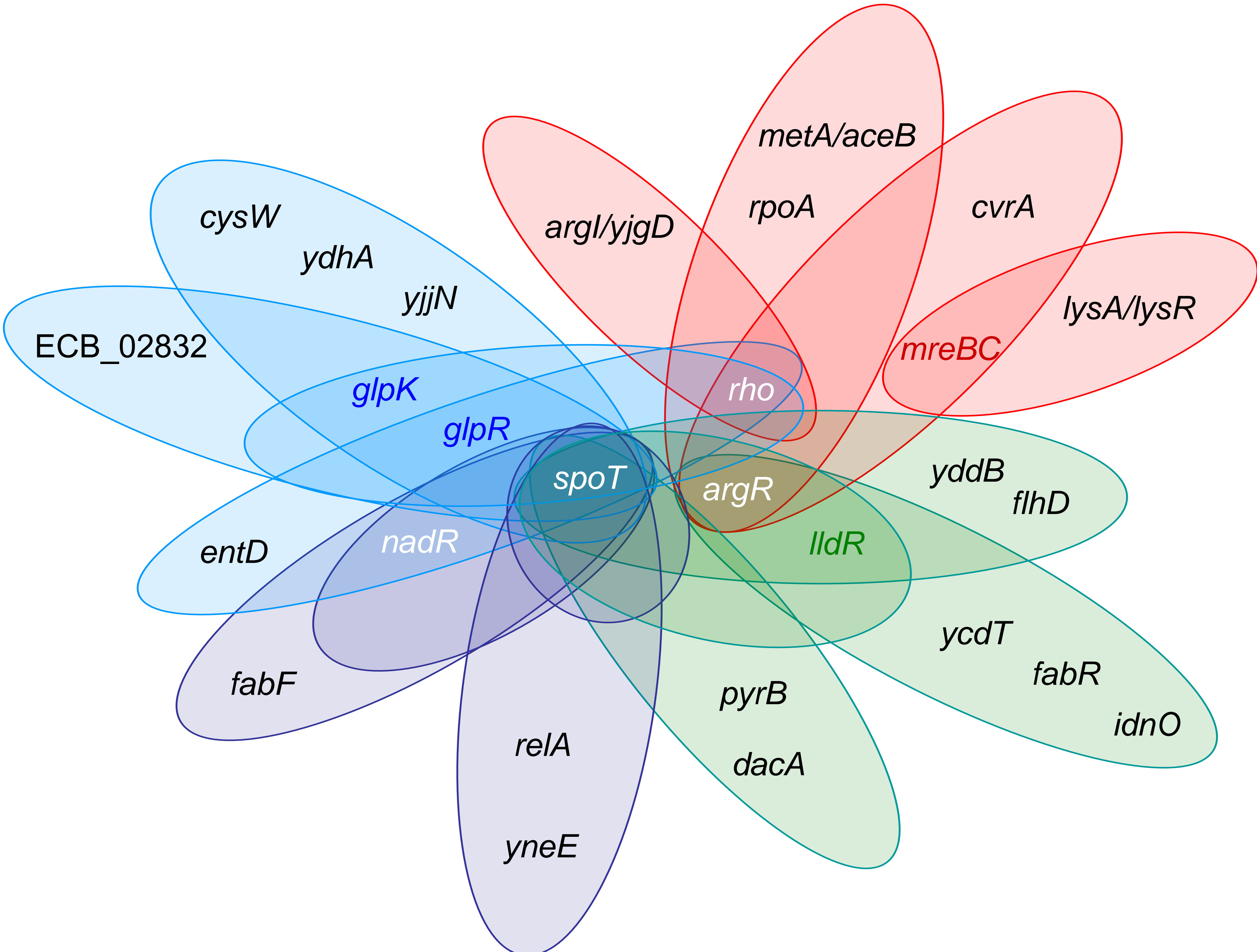
693

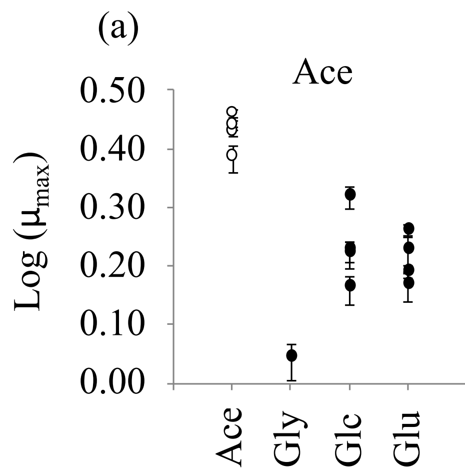
694

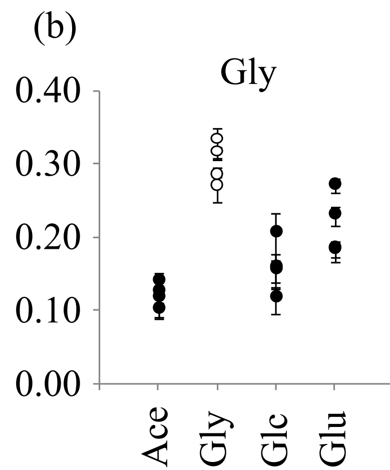

(a)
0.60
0.40
0.20
0.00
-0.20
Log (μmax)
200
400
600
800
1000
Generations

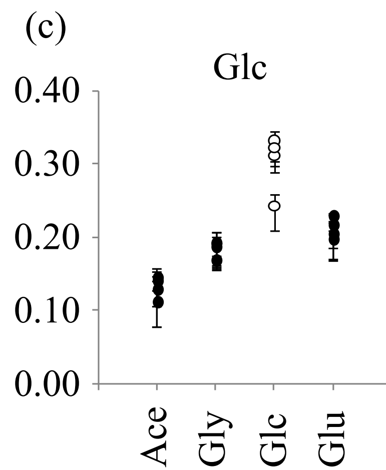

(b)
0.50
0.40
0.30
0.20
0.10
0.00
-0.10
Log (Fitness)
200
400
600
800
1000
Generations

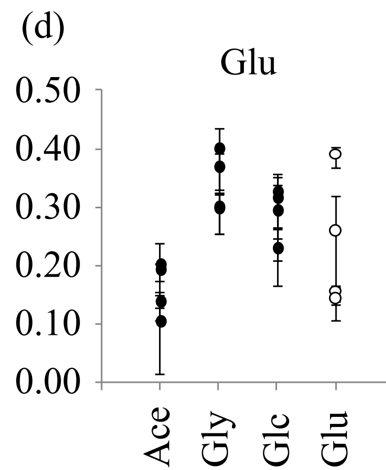

0.50
0.40
0.30
0.20
0.10
0.00
Log (Fitness)
0.00
0.10
0.20
0.30
0.40
0.50
Log (μ_{max})

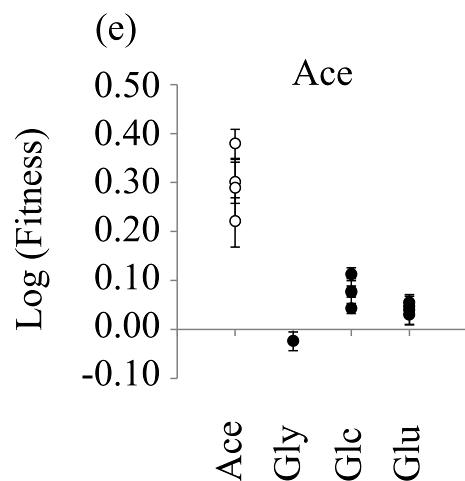

metA/aceB
rpoA
cvrA
argI/yjgD
cysW
ydhA
yjjN
lysA/lysR
ECB_02832
mreBC
glpK
rho
glpR
spoT
argR
yddB
flhD
entD
nadR
lldR
ycdT
fabR
fabF
pyrB
idnO
relA
dacA
yneE

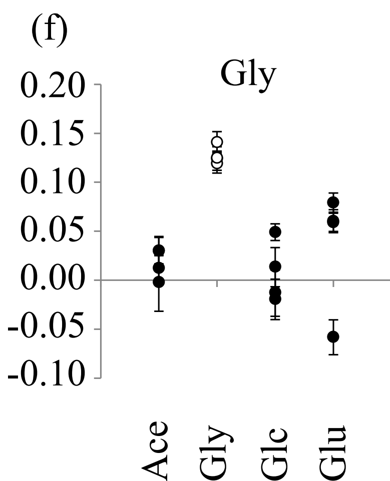

(a)
Ace
Log (μmax)
0.50
0.40
0.30
0.20
0.10
0.00
Ace
Gly
Glc
Glu

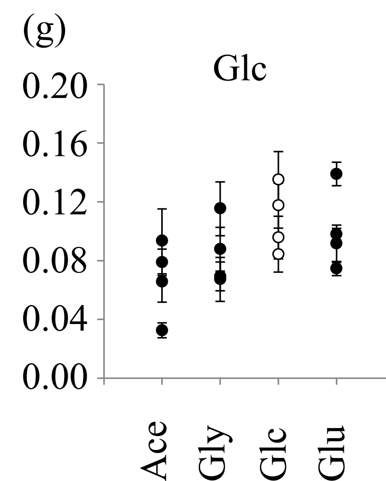

(b)
Gly
0.40
0.30
0.20
0.10
0.00
Ace
Gly
Glc
Glu

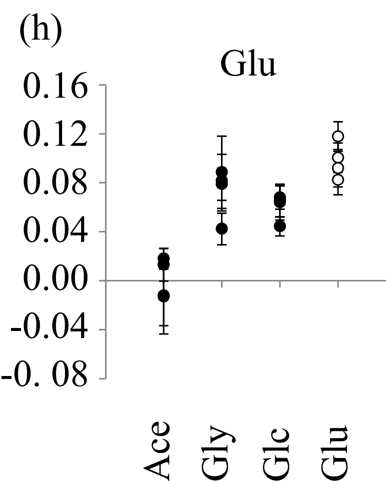

(c)
Glc
0.40
0.30
0.20
0.10
0.00
Ace
Gly
Glc
Glu
Environment

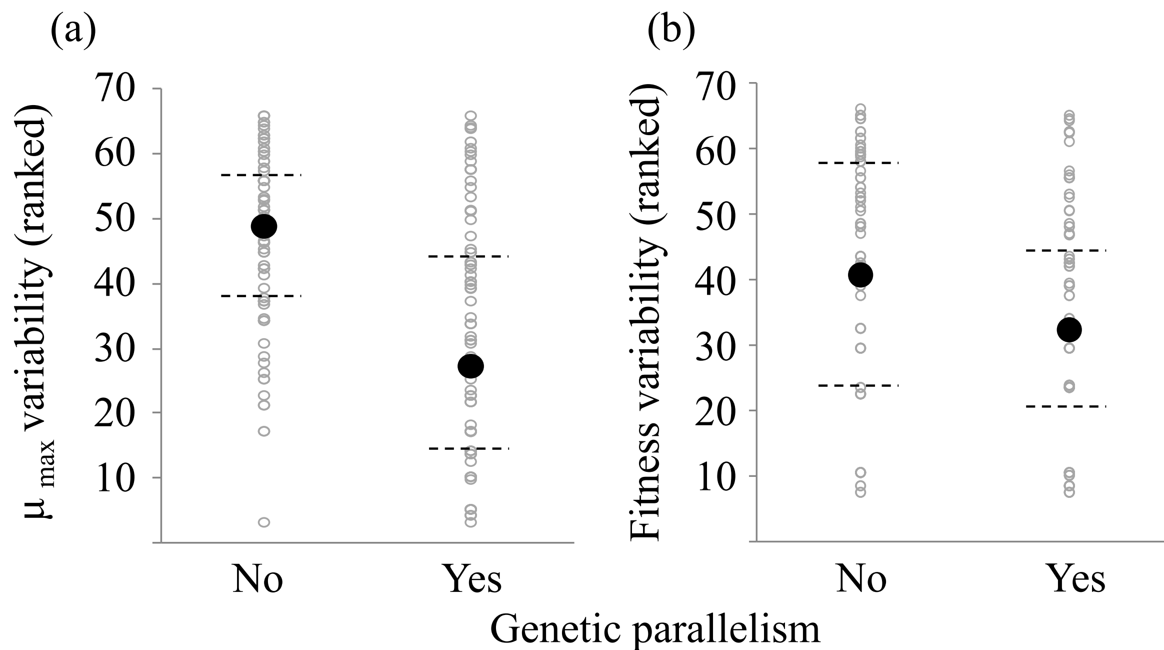
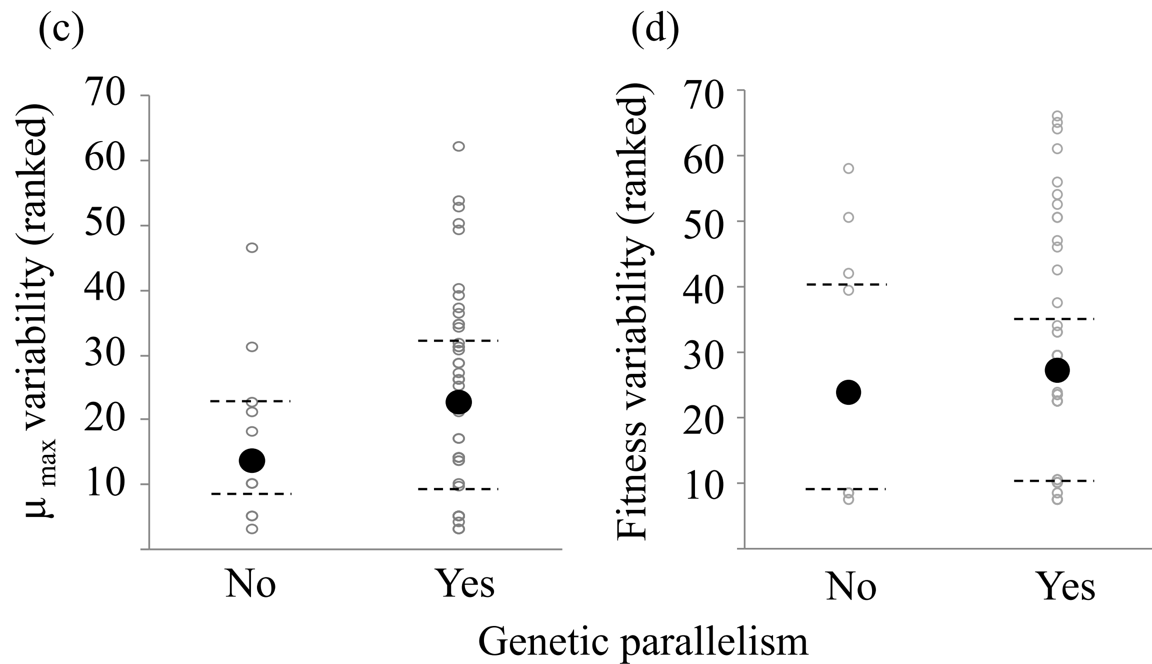

(d)
Glu
0.50
0.40
0.30
0.20
0.10
0.00
Ace
Gly
Glc
Glu


(e)
Ace
Log (Fitness)
0.50
0.40
0.30
0.20
0.10
0.00
-0.10
Ace
Gly
Glc
Glu


(f)
Gly
0.20
0.15
0.10
0.05
0.00
-0.05
-0.10
Ace
Gly
Glc
Glu


(g)
Glc
0.20
0.16
0.12
0.08
0.04
0.00
Ace
Gly
Glc
Glu
Environment


(h)
Glu
0.16
0.12
0.08
0.04
0.00
-0.04
-0. 08
Ace
Gly
Glc
Glu


(a)
(b)
70
60
50
40
30
20
10
μ_{max} variability (ranked)
Fitness variability (ranked)
No
Yes
Genetic parallelism

(c)
(d)
70
60
50
40
30
20
10
μ_{max} variability (ranked)
Fitness variability (ranked)
No
Yes
Genetic parallelism

Supporting information

Evolutionary history and genetic parallelism affect correlated responses to evolution

Mickael Le Gac, Tim Cooper, Stéphane Cruveiller, Claudine Médigue, and Dominique

Schneider

This file includes:

Table S1 Phenotypic traits (μ_{max} and fitness) of the 16 populations during evolutionary time in their respective evolution environments.

Table S2 Phenotypic traits (μ_{max} and fitness) of the evolved clones sampled from each of the 16 populations in the four environments.

Table S3 Among-clone coefficient of variation in the four environments for the two phenotypic traits μ_{max} and fitness.

Table S1 Phenotypic traits (μ_{max} and fitness) of the 16 populations during evolutionary time in their respective evolution environments

Population	μ_{max}*						fitness*			
	0	200	400	600	800	1000	0	300	600	1000
Ace_1	1.00+/-0.09	1.68+/-0.10	2.08+/-0.32	1.99+/-0.09	2.05+/-0.22	2.73+/-0.50	0.97+/-0.06	2.11+/-0.10	1.77+/-0.06	1.39+/-0.05
Ace_2		1.83+/-0.33	2.42+/-0.13	2.08+/-0.21	2.36+/-0.10	2.24+/-0.18		1.94+/-0.05	1.87+/-0.19	2.40+/-0.22
Ace_3		1.81+/-0.30	2.64+/-0.10	2.17+/-0.02	2.37+/-0.21	2.24+/-0.13		2.28+/-0.14	1.73+/-0.14	1.41+/-0.03
Ace_4		1.58+/-0.08	2.98+/-0.47	2.10+/-0.31	2.33+/-0.30	2.22+/-0.18		1.79+/-0.11	1.48+/-0.08	1.38+/-0.07
Gly_1	1.00+/-0.02	1.24+/-0.01	1.53+/-0.08	1.64+/-0.07	1.60+/-0.06	1.95+/-0.21	1.00+/-0.05	1.55+/-0.10	1.95+/-0.15	2.15+/-0.10
Gly_2		1.11+/-0.02	1.43+/-0.01	1.74+/-0.06	1.83+/-0.18	1.90+/-0.13		1.47+/-0.08	1.71+/-0.05	1.81+/-0.06
Gly_3		1.17+/-0.02	1.61+/-0.05	1.91+/-0.03	1.79+/-0.15	1.98+/-0.09		1.51+/-0.05	1.84+/-0.05	1.68+/-0.10
Gly_4		1.17+/-0.01	1.60+/-0.01	1.95+/-0.20	1.57+/-0.07	2.08+/-0.08		1.56+/-0.09	1.61+/-0.09	1.94+/-0.22
Glc_1	1.00+/-0.04	1.29+/-0.17	1.66+/-0.04	1.68+/-0.05	1.76+/-0.02	1.70+/-0.08	0.96+/-0.03	1.63+/-0.12	1.56+/-0.09	1.74+/-0.14
Glc_2		1.34+/-0.04	1.60+/-0.08	1.90+/-0.11	2.06+/-0.18	2.24+/-0.11		1.58+/-0.12	1.69+/-0.05	2.12+/-0.17
Glc_3		1.33+/-0.15	1.56+/-0.04	1.71+/-0.07	1.94+/-0.27	1.73+/-0.08		1.52+/-0.06	1.78+/-0.14	2.12+/-0.00
Glc_4		1.26+/-0.18	1.86+/-0.15	1.96+/-0.14	1.94+/-0.16	2.08+/-0.12		1.32+/-0.05	1.50+/-0.06	1.53+/-0.05
Glu_1	1.00+/-0.04	0.82+/-0.03	1.07+/-0.06	1.07+/-0.02	1.04+/-0.01	0.98+/-0.02	0.91+/-0.06	1.14+/-0.12	1.13+/-0.04	1.39+/-0.14
Glu_2		0.83+/-0.01	1.07+/-0.02	1.05+/-0.01	0.97+/-0.02	1.01+/-0.01		1.03+/-0.06	1.21+/-0.13	1.41+/-0.07
Glu_3		0.86+/-0.00	1.09+/-0.01	1.03+/-0.00	1.03+/-0.02	1.00+/-0.03		1.06+/-0.07	1.21+/-0.03	1.39+/-0.08
Glu_4		0.87+/-0.03	1.02+/-0.02	1.00+/-0.02	1.03+/-0.01	1.00+/-0.01		1.28+/-0.03	1.45+/-0.08	1.52+/-0.02

*μ_{max} and fitness values are presented as log-transformed in Figure 1 but not in Table S1.

Table S2 Phenotypic traits (μ_{max} and fitness) of the evolved clones sampled from each of the 16 populations in the four environments

Environnent	Population	μ_{max} *	fitness*
Ace	Ace_1	2.71+/-0.07	1.36+/-0.06
	Ace_2	2.78+/-0.07	1.27+/-0.04
	Ace_3	2.46+/-0.17	1.31+/-0.06
	Ace_4	2.91+/-0.06	1.34+/-0.08
	Gly_1	Very slow growth**	Very slow growth
	Gly_2	Very slow growth	Very slow growth
	Gly_3	Very slow growth	Very slow growth
	Gly_4	1.12+/-0.11	0.90+/-0.07
	Glc_1	2.10+/-0.12	1.22+/-0.02
	Glc_2	1.70+/-0.09	1.17+/-0.07
	Glc_3	1.68+/-0.11	1.22+/-0.03
	Glc_4	1.47+/-0.11	1.15+/-0.05
	Glu_1	1.56+/-0.05	1.11+/-0.08
	Glu_2	1.70+/-0.16	1.06+/-0.04
	Glu_3	1.49+/-0.11	1.14+/-0.04
	Glu_4	1.84+/-0.06	1.10+/-0.04
Gly	Ace_1	1.39+/-0.06	1.19+/-0.13
	Ace_2	1.27+/-0.04	1.19+/-0.10
	Ace_3	1.34+/-0.06	1.05+/-0.14
	Ace_4	1.32+/-0.08	1.03+/-0.05
	Gly_1	1.87+/-0.10	2.09+/-0.26
	Gly_2	1.93+/-0.08	1.97+/-0.12
	Gly_3	2.08+/-0.06	1.61+/-0.06
	Gly_4	2.16+/-0.13	1.94+/-0.08
	Glc_1	1.45+/-0.10	0.98+/-0.06
	Glc_2	1.61+/-0.19	1.00+/-0.05
	Glc_3	1.44+/-0.06	1.09+/-0.06
	Glc_4	1.32+/-0.07	1.28+/-0.08
	Glu_1	1.71+/-0.07	1.37+/-0.13
	Glu_2	1.54+/-0.05	0.80+/-0.05
	Glu_3	1.87+/-0.05	1.43+/-0.10
	Glu_4	1.53+/-0.06	1.26+/-0.07
Glc	Ace_1	1.29+/-0.10	1.34+/-0.07
	Ace_2	1.34+/-0.15	1.54+/-0.12
	Ace_3	1.39+/-0.02	1.39+/-0.05
	Ace_4	1.38+/-0.11	1.25+/-0.05
	Gly_1	1.56+/-0.10	1.42+/-0.07
	Gly_2	1.54+/-0.09	1.79+/-0.10
	Gly_3	1.47+/-0.04	1.42+/-0.10
	Gly_4	1.55+/-0.11	1.40+/-0.07
	Glc_1	2.05+/-0.10	1.62+/-0.11
	Glc_2	2.14+/-0.14	1.51+/-0.09
	Glc_3	1.75+/-0.13	1.39+/-0.06
	Glc_4	2.10+/-0.11	1.54+/-0.06
	Glu_1	1.64+/- 0.11	1.53+/-0.04
	Glu_2	1.60+/-0.13	1.43+/-0.04
	Glu_3	1.69+/-0.04	1.63+/-0.04
	Glu_4	1.57+/-0.09	1.32+/-0.02
Glu	Ace_1	1.16+/-0.11	0.95+/-0.17
	Ace_2	1.25+/-0.03	1.10+/-0.09
	Ace_3	1.26+/-0.10	1.09+/-0.05
	Ace_4	1.17+/-0.04	0.97+/-0.19
	Gly_1	1.41+/-0.07	2.14+/-0.50
	Gly_2	1.53+/-0.08	1.55+/-0.17

Gly_3	1.59+/-0.13	1.86+/-0.24
Gly_4	1.42+/-0.07	1.86+/-0.18
Glc_1	1.30+/-0.10	1.45+/-0.14
Glc_2	1.41+/-0.14	1.44+/-0.07
Glc_3	1.44+/-0.12	1.46+/-0.13
Glc_4	1.46+/-0.10	1.30+/-0.05
Glu_1	1.20+/-0.03	2.28+/-0.15
Glu_2	1.57+/-0.04	2.12+/-0.20
Glu_3	1.35+/-0.19	2.00+/-0.19
Glu_4	1.18+/-0.05	2.00+/-0.27

The μ_{max} and fitness values are given (+/-SEM) for each evolved clone sampled from each of the 16 evolving populations in each of the four different environments.

[*] μ_{max} values are presented as log-transformed in Figure 2 but not in Table S2.

[**] As no quantitative values could be obtained for these measures, they were not included in the statistical analyses.

Table S3 Among-clone coefficient of variation in the four environments for the two phenotypic traits μ_{max} and fitness

Evolution environment	Alternative environment	Coefficient of variation	
		μ_{max}	fitness
Ace	Ace	0.07	0.15
	Gly	0.04	0.04
	Glc	0.03	0.06
	Glu	0.05	0.04
Gly	Ace	Very slow growth*	
	Gly	0.07	0.02
	Glc	0.03	0.05
	Glu	0.06	0.05
Glc	Ace	0.15	0.07
	Gly	0.08	0.07
	Glc	0.09	0.05
	Glu	0.05	0.02
Glu	Ace	0.09	0.02
	Gly	0.10	0.13
	Glc	0.03	0.06
	Glu	0.13	0.03

*As no quantitative values could be obtained for these measures, they were not included in the statistical analyses.