BMC Ecology and Evolution

RESEARCH Open Access

Experimental evolution reveals evolutionary bias and its causes

Haoyuan Wu[1] and Yonghua Wu[1,2*]

Abstract

Background Species generally exhibits evolutionary bias, adapting towards a specific direction rather than others, yet the underlying causes remains unknown.

Results Here, we investigated evolutionary bias and its causes by conducting experimental evolution on *Escherichia coli*. We introduced an *E. coli* strain (lac-), initially unable to utilize lactose due to a frameshift mutation, into two different culture media: one medium (L) containing ample sodium acetate and lactose as carbon sources, and the other medium (G) containing abundant glucose and lactose as carbon sources. After 20 days of experimental evolution, our findings revealed that all L-populations underwent parallel evolution through reverse mutation to utilize lactose (lac+), resulting in a relatively higher fitness gain compared to utilizing sodium acetate. In contrast, all G-populations did not transition towards lactose utilization but instead continued to utilize glucose, which provides a higher fitness gain than utilizing lactose. These results demonstrate that our experimental populations in L and G media respectively exhibit biased evolution towards utilizing different carbon sources, yet all trajectories converge towards higher fitness gains. When lac+ (lactose-eater) and lac- (acetate-eater) were co-cultured in L medium, all lac- individuals were eventually eliminated, while lac + individuals were consistently selected and retained.

Conclusions Our findings indicate that species tend to evolve with a bias towards directions that offer higher fitness gains, partly because high-fitness-gain directions competitively exclude low-fitness-gain directions.

Keywords Evolutionary gravitation, Evolutionary bias, Evolutionary constraint hypothesis, Inter-directional selection, Intradirectional selection

Introduction

Species generally exhibit specific evolutionary directions, often biased towards certain adaptations. For example, giant pandas have evolved specialized adaptations towards bamboo consumption rather than towards utilizing other food resources such as meat. Aphids prefer tender leaves over older ones. The dead leaf butterflies evolve adaptations towards mimicking dead leaves rather than towards mimicking green leaves or branches. Reeds are predominantly found at the edges of water bodies rather than further adapting to deeper water or shorelines. Many trees in arid regions are restricted to shady slopes of mountains rather than sunny slopes. Fossil

*Correspondence:
Yonghua Wu
wuyh442@nenu.edu.cn
[1]School of Life Sciences, Northeast Normal University, 5268 Renmin Street, Changchun 130024, China
[2]Jilin Provincial Key Laboratory of Animal Resource Conservation and Utilization, Northeast Normal University, 2555 Jingyue Street, Changchun 130117, China

studies indicate that this evolutionary bias (such as ecological niche conservatism) in some species has remained stable for millions of years [1]. Within the distribution range of a species, there are often numerous potential ecological niches or evolutionary directions available for adaptive evolution, why does a species adapt evolutionarily biased towards one direction rather than others? What determines the selection of adaptive evolutionary directions for a species? These questions constitute fundamental issues in evolutionary biology [2–4].

We know that species possess evolutionary potential to radiate into different adaptive directions, exemplified by cases such as Darwin's finches [5], Anolis lizards [6], and African cichlids [7]. These diversifications typically occur under substantial geographical isolation. In the absence of such isolation, within-species adaptive divergence (e.g. sympatric speciation) is generally rare [4, 8], and species often evolve collectively in one predominant direction and may not diverge towards other directions. A traditional viewpoint suggests that high levels of gene flow within species homogenize populations and suppress genetic differentiation [4, 9]. Specifically, it is commonly held that genetic swamping and outbreeding depression resulting from high gene flow hinder local adaptation and thereby prevent intraspecific divergence in those directions [2, 3, 10–13]. In addition to the intraspecific factors, a species' inability to adapt in certain directions may also be influenced by abiotic stressors such as severe desert environments or negative interspecific interactions including predation, competition, parasitism, and hybridization with closely related species [2, 14–16]. Nevertheless, the presence of these genetic and ecological negative impacts only explains why a species cannot adapt in some directions, not why it can evolve towards other certain directions, given that such evolutionary trajectories may similarly be affected by those negative factors. Therefore, the mechanisms driving the biases in species evolution remain less clear.

Regarding the evolutionary bias, previous theoretical and experimental studies have shown that in heterogeneous environments, species tend to adapt evolutionarily towards the most productive resource [17–22]. These findings suggest that the evolutionary bias may arise because the fitness gains of adapting in that direction outweigh those of other evolutionary directions [18, 20–22]. We know that a species' environment often presents multiple evolutionary directions (such as utilizing different food resources), each with potentially varying fitness returns. For instance, studies indicate that *Escherichia coli* exhibits higher growth rates when consuming glucose compared to other carbon sources like lactose and sodium acetate [23, 24]. Similarly, one study on Darwin's medium ground finch shows that during drought periods, large birds that can consume large seeds survive better than small birds that feed on small seeds [25], demonstrating differential fitness gains from different food sources. Due to varying fitness returns across different adaptive directions, Darwinian selection acting on mutations in different adaptive directions may tend to preserve the evolutionary direction with higher fitness gains and eliminate those with lower fitness gains. If correct, this may suggest that a species evolves towards a particular direction because it offers relatively higher fitness benefits. The elimination of less advantageous directions may result from genetic and ecological suppressive factors mentioned above [2, 3, 10–16], particularly the gene flow from directions with high fitness gains [17–22], which may facilitate hybridization and competitive exclusion. We hypothesize that evolutionary directions with high fitness gains may evolutionarily suppress those with lower fitness gains, thereby constraining species evolution to biased directions that offer higher fitness gains.

Experimental evolution provides a viable approach to test predictions of relevant evolutionary theories [26, 27]. Model organisms in experimental evolution, such as *Escherichia coli*, are characterized by short generation times, rapid evolution, and ease of cultivation, making them ideal for experimental studies in evolution. Among *E. coli* strains, *E. coli* K-12 GM4792, which is asexual and is characterized by its incapability to use lactose, has been used for experimental evolution [28]. This strain harbors a continuous 212-bp deletion in lactose operon, spanning the lactose operator (O) and promoter (P) region, along with partial sequences of the *lacI* and *lacZ* genes, resulting in a fusion of the *lacI* and *lacZ* genes (Fig. 1). Additionally, a single-base (C) insertion in the *lacI* gene led to a frameshift mutation, inactivating the *lacZ* gene, rendering it unable to utilize lactose (lac-) [28, 29]. Despite this genetic deficiency, when this strain is cultured in a medium with lactose as the sole carbon source, a reverse mutation allowing lactose utilization (lac+) occurs [28, 29]. The lac- and lac+phenotypes are well known to respectively display white and blue clones when plated on LB agar containing IPTG and X-gal (5-bromo-4-chloro-3-indolyl-β-d-galactopyranoside), and this blue-white screening can be used to detect the possible emergence of lac+ [30]. The potential for the re-evolution of lactose utilization in *E. coli* K-12 GM4792 and the convenient detection of lac+through blue-white screening provide an ideal research system for studying possible transitions in evolutionary directions.

In this study, to test the evolutionary constraint hypothesis, we conducted experimental evolution using *E. coli* K-12 GM4792 as a model species. We find that all our experimental populations exhibit evolutionary biases towards utilizing carbon sources that provide higher fitness gains. Our results suggest that the competitive exclusion of low-fitness-gain directions by

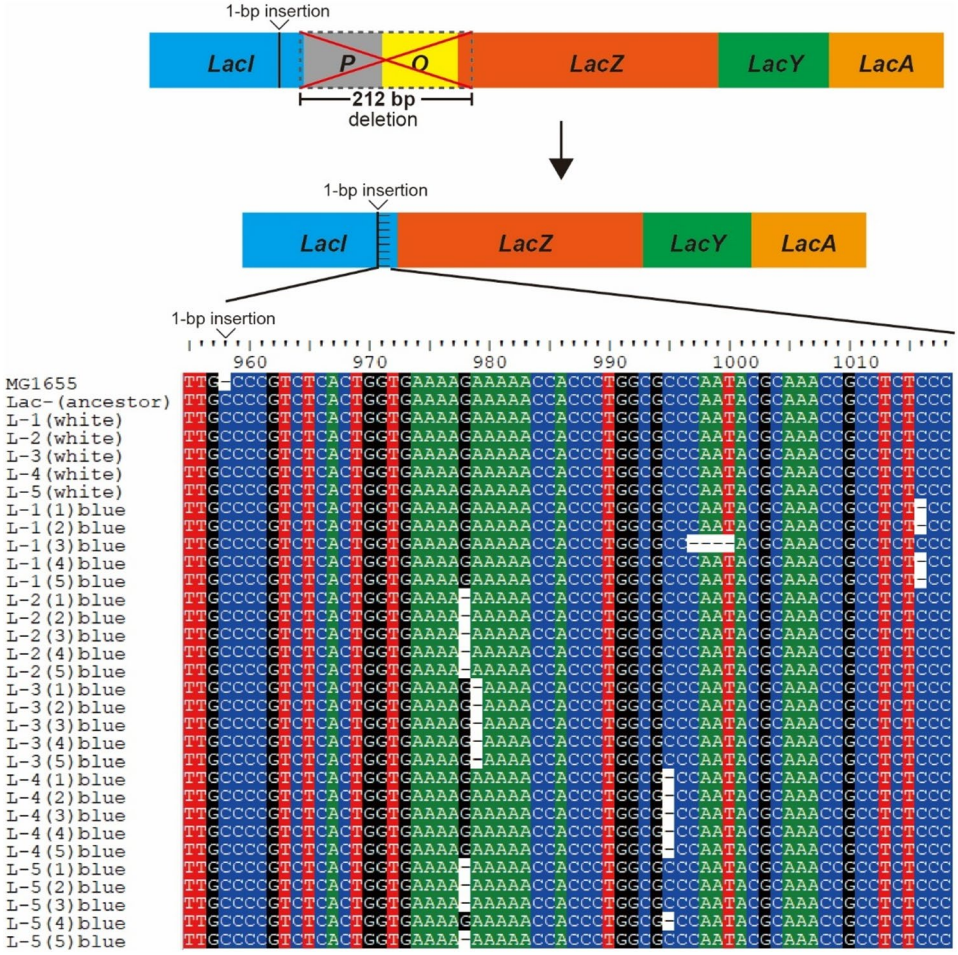

Fig. 1 Structure of the lactose operon and sequence alignment. For lactose operon, the 212 bp deletion and a single base insertion of *E. coli* K-12 GM4792 are illustrated. Regarding the sequences, only the mutation region is displayed. The sequences of all white clones are identical to lac-(ancestor). Conversely, all blue clones carry nucleotide deletions, which are confined to the region (indicated by horizontal stripes) between the single base insertion site 958 of the *lacI* gene and the *lacZ* gene. For convenience, we have also included the corresponding fragment sequences of *E. coli* K-12 MG1655

high-fitness-gain directions plays a crucial role in driving these evolutionary biases.

Results and discussion

Through experimental evolution of *E. coli*, we first investigated the impact of relative fitness gains from different food sources on the formation of evolutionary directions. For this purpose, we utilized a strain of *E. coli* K-12 GM4792, which was initially unable to utilize lactose [28, 29]. We used the *E. coli* K-12 GM4792 strain, which had undergone three consecutive rounds of single-colony cultivation, as the ancestral strain (lac-(ancestor)) for this study. Through blue-white screening and PCR sequencing, our ancestral strain exhibited a white color and a 212-bp deletion, confirming its origin from the lactose-deficient strain of *E. coli* K-12 GM4792 used in this experiment. We cultured lac-(ancestor) in two different M9 culture media (L and G), with the only difference between these two media being the composition of the carbon source. L medium contained abundant sodium acetate and lactose, while G medium contained ample glucose and lactose. Both carbon sources showed substantial remaining amounts after 24 h of cultivation in our experimental strain.

Previous research has indicated that the utilization of different carbon sources can lead to varying growth rates in *E. coli*. Compared to lactose, glucose utilization contributes to a relatively higher growth rate [23], while sodium acetate is considered a low-energy food and even exhibits some toxicity to *E. coli*, exerting a certain inhibitory effect on its growth [24]. The utilization of these three carbon sources results in different growth rates for *E. coli*, providing a possibility for us to examine the impact of relative fitness gains from different food sources on the formation of evolutionary directions.

Experimental evolution and the selection of evolutionary direction

We subjected lac-(ancestor) to experimental evolution in culture media L and G. For each of these two media, we conducted 5 experimental replicates, resulting in the founding of 10 experimental populations (L1 to L5 and G1 to G5). All 10 populations were derived from lac-(ancestor), and they initially shared the same genetic background and inability to utilize lactose.

Since lac-(ancestor) cannot utilize lactose, the initial capability of the 5 L-populations was limited to utilizing only sodium acetate, while the initial capability of the 5 G-populations was restricted to utilizing only glucose. For these 10 populations, we transferred 100 µl into 9.9 ml fresh culture medium in 50 ml plastic culture tubes every 24 h, and incubated them at 37 °C with shaking at 220 rpm. Every 5 days, we performed blue-white screening to detect the appearance of lac+individuals, and then collected 4 ml of each culture for glycerol preservation under −80 °C conditions.

After 25 days of experimental evolution and blue-white screening, our results showed that all five L-populations eventually evolved from lac- to lac+, indicating adaptation toward lactose utilization. In contrast, no blue colonies were detected in any of the 5 G-populations, suggesting that they were less likely to undergo adaptive evolution toward lactose utilization and instead maintained their evolutionary direction (utilizing glucose).

Specifically, for the five L-populations, after conducting blue-white screening on two plates for each population every 5 days, totaling 10,818 colony counts, we observed that by the 5th day of experimental evolution, few blue clones were observed in three L-populations (3 in L2, 1 in L4, and 1 in L5). However, by the 15th day, four populations (L2-5) had a blue colony proportion higher than 95%, with one population, L2, reaching 100%. The proportion of blue colonies in the L1 population remained consistently low during the first 15 days, with no blue colonies detected in the first 10 days, reaching only 1% by the 15th day. However, by the 20th day, the proportion of blue colonies suddenly surged to 100%, similar to the other four populations. An additional 5 days of experimentation revealed that by the 25th day, the proportion of blue colonies in all five L-populations had reached 100%, with no white clones found.

To further confirm the changes in the blue-white colony ratio, after the 25th day of experimental evolution, we conducted another round of blue-white screening on the preserved cultures of L1-5 (Fig.S1). After counting a total of 12,270 colonies, we obtained nearly identical results. By the 20th and 25th days, all five L-populations had completely transformed into blue colonies, with no white clones found. This indicates that the evolved lac+completely dominated the populations, while the lac- strains faced extinction. We combined the results of blue-white colony counts (23,088) for the five L-populations (Table S1), as shown in Fig. 2.

Molecular basis of the parallel evolution of lactose utilization

Our experimental evolution results indicate that after a brief period of approximately 20 days, all five initial L-populations unable to utilize lactose underwent an evolutionary transition towards lactose utilization. To examine the molecular mechanism responsible for their evolution of lactose utilization, we conducted full-length sequencing of the lactose operon in colonies from the 5 L-populations. We sequenced 5 randomly selected blue clones and 1 white clone from each L-population, as well as the ancestral strain, lac-(ancestor) (Table S2), resulting in full-length sequences of the lactose operon spanning 6051 bp. Sequence analysis revealed that all 31 samples examined contained a 212 bp deletion and a single C insertion at position 958 in the *lacI* gene (Fig. 1). This indicates that all the five L-population colonies we tested, including the 25 blue clones capable of utilizing lactose, originated from our ancestral strain, lac-(ancestor), ruling out the possibility of exogenous bacterial contamination. This suggests that the observed lactose utilization mutations in our experimental evolution were acquired during our experimental evolution process.

Further sequence analysis revealed that all of our examined white clones had sequences identical to that of the ancestral strain. However, out of the 25 sequenced blue clones, 24 showed single-base deletions and one exhibited a continuous loss of 4 bases. These deletions were all concentrated within a specific region (mutational region) between the single C insertion site at position 958 in the *lacI* gene and the *lacZ* gene (Fig. 1). These nucleotide deletions appeared to reverse the frameshift mutation of lac-, potentially enabling normal expression of the *lacZ* gene. Specifically, all of our sequenced white colonies with the same sequence as the ancestral strain contained a frameshift mutation caused by a single-base C insertion, resulting in a white appearance on X-gal LB plates, indicating abnormal *lacZ* gene expression. In contrast, all 25 colonies with single-base or continuous 4-base deletions appeared blue on X-gal LB plates, suggesting normal *lacZ* gene expression. Therefore, sequence analysis indicates that the evolution of lactose utilization in our L-populations occurred through base deletions restoring the frameshift mutation of lac-.

To further investigate whether the lactose utilization ability of these 5 L-populations was acquired through parallel evolution, we conducted a comparative analysis of the full-length sequences of the lactose operon from the sequenced samples. The findings revealed that while all blue clones from the 5 L-populations obtained lactose

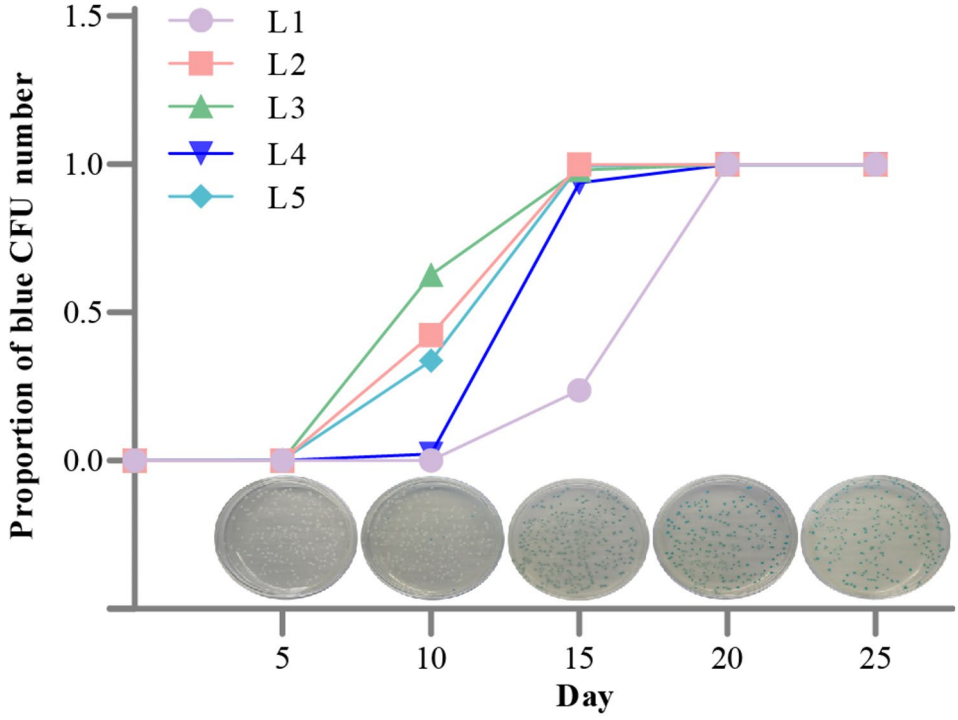

Fig. 2 Changes in the proportion of blue colonies in the 5 L-populations over a 25-day experimental evolution period. Representative plates from day 5 to day 25 are displayed. CFU stands for colony-forming unit

utilization ability through base loss mutations, the specific locations of these deletions varied (Fig. 1). Within these populations, in the L1 population, four blue single clones displayed single-base deletions at position 1016, while another blue single clone exhibited a continuous loss of four bases spanning positions 997 to 1000. In populations L2, L3, and L4, each of the five blue single clones had single-base deletions at positions 978, 979, and 995, respectively. In population L5, four blue single clones had single-base deletions at position 978, and one other blue single clone had a single-base deletion at position 995. The distinct locations of the reverse mutations in these 5 L-populations strongly suggest that these populations independently acquired the ability for lactose utilization through parallel evolution, effectively eliminating the possibility of cross-contamination.

To comprehensively understand the adaptive evolution of the experimental populations, particularly to examine whether parallel evolution also occurred in genes other than those related to lactose utilization, we conducted whole-genome sequencing of the ancestral strain, lac- (ancestor), and 10 evolved experimental populations after 25 days (Tables S3-S5). We performed whole-genome sequencing and comparative analysis of 21 single clones, including one clone from the ancestral strain and two randomly selected clones from each evolved population. The results revealed (Fig. 3, Tables S6-S7) that compared to the ancestral strain, the 10 evolved populations after 25 days showed few mutations, with only 27 SNPs detected in the 20 evolved population clones, averaging 1.35 SNPs per clone. Their evolutionary tree also exhibited very shallow divergence (Fig. S2). Out of these 27 SNPs, 6 occurred in intergenic regions, while the remaining 21 SNPs included 6 synonymous substitutions and 15 non-synonymous substitutions (Table S7). These 15 non-synonymous substitutions were found in different genes across different populations, with only three populations (G2, L2, and L4) sharing the same non-synonymous substitution (A-G) in the gene *staF*, while the rest occurred either in a single population or in two populations (Fig. 3). No G-population shared any non-synonymous substitutions. Only two L-populations (L2 and L4) shared two non-synonymous substitutions (G-A, A-G).

In addition to SNPs, we also analyzed the occurrence of indels and detected a total of 10 indels (Table S8). Interestingly, all 10 indels involved single-base losses and were restricted to the 10 clones from the five L-populations (Fig. 3). Further analysis revealed that these single-base losses all occurred in the mutation region of the *lacI* and *lacZ* fusion genes mentioned above, with the losses in different populations largely occurring at different positions, strongly suggesting parallel evolution of their lactose utilization. To verify the authenticity of these single-base losses, we subsequently performed PCR amplification sequencing of this mutation region in the 10 single clones, confirming the loss of these single bases (Fig. S3).

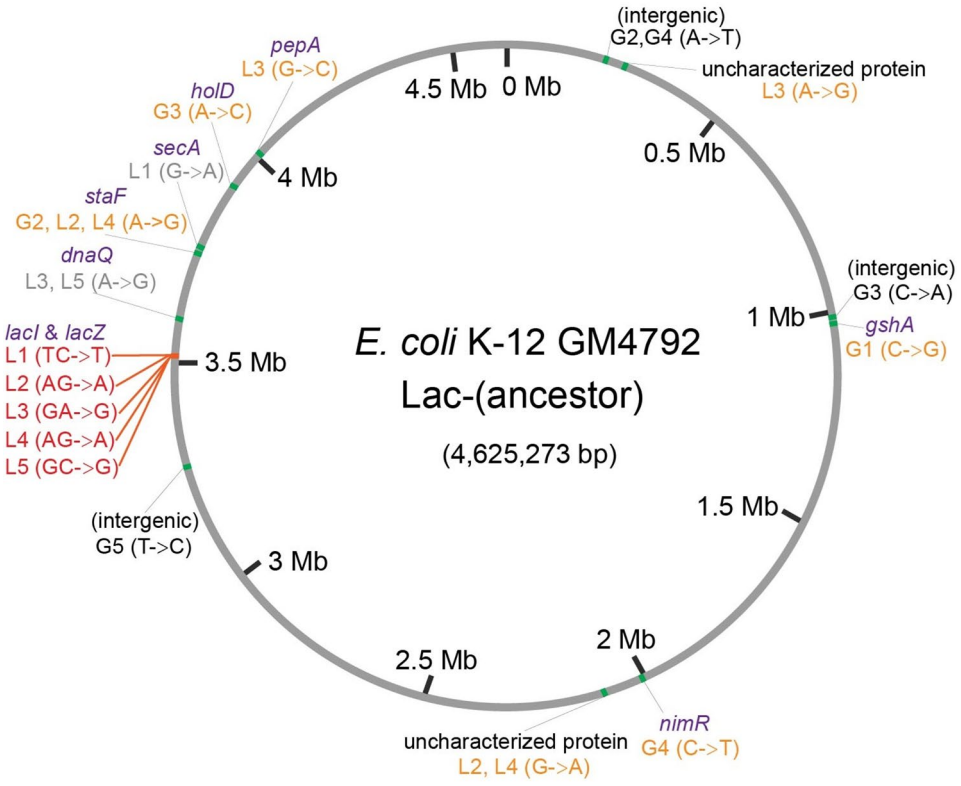

Fig. 3 Mutations found by genome sequencing. The genomes of 25-day evolved populations (L1-5, G1-5) are compared to the ancestral genome (lac-(ancestor)), and SNP (green) and indel (red) are identified. All identified indels are single-base deletions (in red) located exclusively within the mutation region of the fused genes *lacI* and *lacZ* across the five L-populations. Non-synonymous substitutions and synonymous substitutions are indicated in orange and grey for SNPs, respectively. Gene names are highlighted in blue

Therefore, the results of the genome comparison analysis indicate that the *lacI* and *lacZ* fusion genes are the genes that underwent the most significant parallel adaptive evolution in our experimental evolution.

Fitness difference between lac + and lac-

The ancestral strain lac-(ancestor) evolved into lactose-utilizing lac + in L medium, indicating that lac + has a higher relative fitness than lac- in L medium. Conversely, in G medium, no blue colonies were detected, suggesting that lac- may have a higher relative fitness than lac + in G medium. To test this, we randomly selected two blue clones (lac+) and two white clones (lac-) from each L-population to form 10 competitive pairs for competition experiments. Similarly, we randomly selected two white clones (lac-) from each G-population and paired them with two blue clones (lac+) from L-populations to create another 10 competitive pairs. Taking into account the effect of abundance on relative fitness, we adjusted the OD values to set the blue-white colony ratio of the competitive pairs at 10 : 1 and 1 : 10. We conducted 24-hour competitions in both L and G media, under conditions identical to experimental evolution, and performed blue-white screening before and after competition to calculate the density of blue and white colonies.

Table 1 Statistical analysis of the selection rate constant. A one-sample t-test with a two-tailed probability was used, with the null hypothesis stating that the selection rate constant equals zero. The selection rate constants for lac + relative to lac- were calculated in L medium, and for lac- relative to lac + in G medium. Both lac + and lac- were derived from evolved populations. The means (± SD) based on 10 replicate assays are presented

Culture medium (blue: white ratio)	No. of replicates	Selection rate constant(day^{-1})	t	*p*
L (10:1)	10	1.0596 ± 0.60864	5.506	< 0.001
L (1:10)	10	1.5061 ± 0.92026	5.175	0.001
G (10:1)	10	0.3148 ± 0.36055	2.761	0.022
G (1:10)	10	0.2012 ± 0.23699	2.685	0.025

Considering the observed decrease in population density in some populations, we calculated the selection rate constant, r_{ij}, which reflects the difference in the average rate of increase and relative performance of the two competitors [31, 32]. Our results (Table 1) showed that for the 10 competitive pairs in L medium, after 24 h, the density of lac + increased by approximately 1.06 natural logarithms more than the density of lac- (blue : white = 10 : 1, t = 5.51, df = 9, $p < 0.001$), and by about 1.51 natural logarithms more than the density of lac- when using a ratio of 1:10 of blue and white clones (t = 5.18, df = 9,

$p=0.001$). We also calculated the r_{ij} value in G medium, where a difference of 0.31 natural logarithms more lac- than lac+ (blue : white = 10 : 1, t = 2.76, df = 9, $p=0.022$) and 0.20 natural logarithms more lac- than lac+ (blue : white = 1 : 10, t = 2.69, df = 9, $p=0.025$) were found. These results indicate that in L medium, lac+ has a relative fitness advantage over lac-, while in G medium, lac- has a relative fitness advantage over lac+, and their respective advantages are more pronounced when their frequencies are relatively low.

Competitive exclusion of low-fitness-gain direction by high-fitness-gain direction

The experimental results above indicate that in L medium, lac+ exhibits a higher relative fitness than lac-, leading to the evolutionary transition from lac- to lac+. Conversely, in G medium, lac+ shows a lower relative fitness than lac-, and there is no observed evolutionary transition from lac- to lac+. The superior relative fitness of lac+ over lac- in L medium is likely attributed to their differential utilization of carbon sources. To test this, we measured the consumption of carbon sources by lac- and lac+. Our liquid chromatography analysis in L medium revealed that lac+ can metabolize lactose, while lac- was limited to utilizing sodium acetate (Fig. S4). Previous studies have highlighted that sodium acetate serves as a low-energy carbon source that hinders *E. coli* growth, whereas lactose is a relatively favored carbon source for *E. coli* [23, 24]. The capacity of lac+ to metabolize lactose in L medium while lac- can only utilize sodium acetate may contribute to the former's heightened relative fitness (Table 1). Adding to this, we measured population growth of lac+ and lac- in L medium. The results indicated that during the 24-hour growth period, from the second hour to the 24th hour, lac+ demonstrated significantly faster population growth than lac- (averaged t value = 5.13, df = 8, $p<0.05$), with the disparity in population growth between the two gradually widening over time (Fig. 4; Table S9). We also measured the consumption of carbon sources by lac- and lac+ in G medium. Our liquid chromatography results suggested that lac- appears to exclusively utilize glucose, while lac+ exhibited utilization of both glucose and lactose (Fig. S4). The utilization of lactose by lac+ may partially contribute to its diminished relative fitness compared to lac-. Previous studies have indicated that the population growth rate when utilizing lactose is lower than when using glucose [23]. Moreover, the finding that lac+ utilized both glucose and lactose in G medium suggests that carbon catabolite repression—which refers to the reduced use of lactose in the presence of glucose [33, 34], —may have limited effects on the suppression of lactose utilization. This may be partly attributed to the complete deletion of the lactose operator and promoter regions in *E. coli* K-12 GM4792 (Fig. 1) [28]. Therefore, in G medium, the lack of observed evolutionary transition from lac- to lac+ may be more likely due to a lower relative fitness gain of lac+ utilizing lactose compared to lac- using glucose.

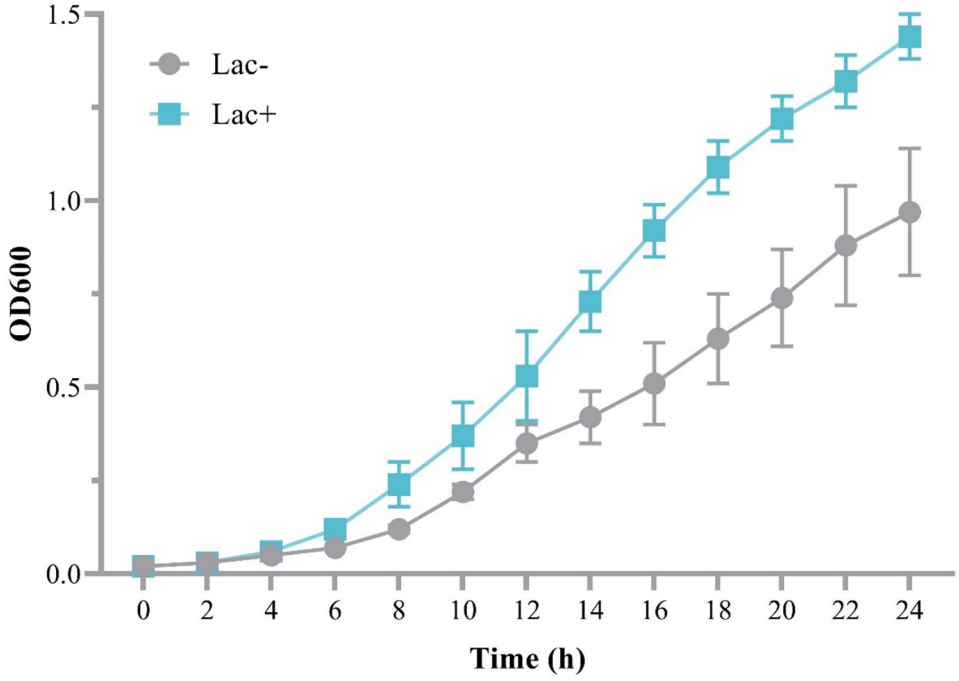

Fig. 4 Growth curve of lac+ and lac- in L medium. Five replicate populations of lac- and lac+ were respectively used to generate the curves. For each time point, the mean and standard deviation of lac+ and lac- based on the five replicate populations are shown separately

Our experimental results show that in the L medium, the lac+, which can consume lactose, exhibited relatively rapid population growth (Fig. 4; Table S9) and higher fitness (Table 1) compared to the lac-, which can only consume sodium acetate. As a result, the lac+ultimately dominated while the lac- went extinct. On the contrary, in the G medium, the lac+mutation, allowing lactose consumption, had lower fitness compared to the lac- (Table 1), which can only consume glucose. Throughout the experimental evolution, we did not observe the appearance of any blue colonies (lac+) in any of the five G-populations. This suggests that the G-populations were less likely to adapt towards lactose utilization, but instead maintained an evolutionary trajectory towards glucose utilization. Considering that the use of different carbon sources represents different adaptive evolutionary directions, these results strongly indicate that the utilization of different carbon sources leads to differences in fitness gains in corresponding evolutionary directions. Mutants in one adaptive direction with high fitness gains will competitively exclude mutations in another adaptive direction with lower fitness gains.

To further validate the phenomenon of competitive exclusion effect, we conducted competitive experiments between lac+and lac- in L medium. From each of the 5 L- populations, we randomly selected 2 blue colonies (lac+) and 2 white colonies (lac-) to create 10 competitive pairs, monitoring the changes in their frequencies over time. To assess the influence of experimental conditions on the outcomes, we carried out four sets of experiments (C, J, L, and H). The C group maintained conditions identical to those of the experimental evolution, while the other three groups experienced modifications: the J group remained static, the L group had reduced sodium acetate levels, and the H group had increased resources apart from the carbon source. Our findings revealed that during a 168-hour competition experiment, the lac- in all four groups eventually became extinct, whereas the lac+became fixed through evolutionary processes (Fig. 5). By the 96th hour, no white colonies were detected in three groups (J, L, and H), and by the 144th hour, the C group also had no white colonies. Across these four groups, the differences in pairwise comparisons at various time points were largely statistically insignificant, except for the L group, where the decrease in white colony frequency was notably significant. Specifically, at the 48th hour, the frequency of white colonies in the L group was significantly lower than that in the C group ($t=2.93$, $df=36$, $p=0.035$) (Table S10). This suggests that the reduction in sodium acetate led to diminished fitness

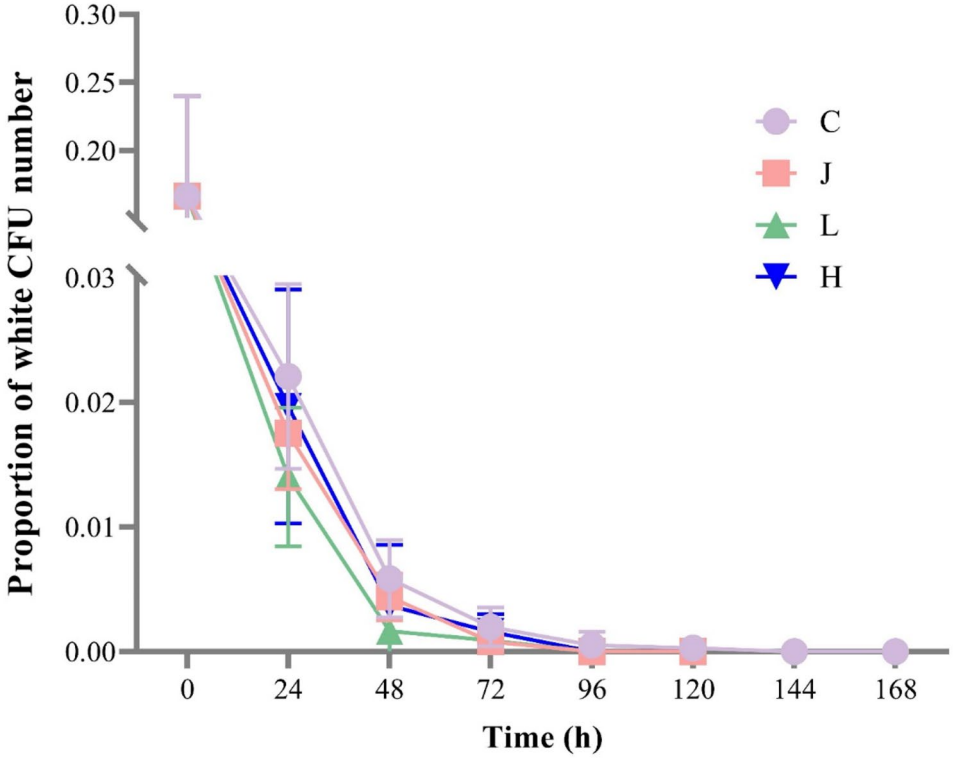

Fig. 5 The changes in the proportion of white colonies over time under four different conditions. C represents L medium with the same culture conditions as the experimental evolution. J represents placing the L medium in a static state. L represents reducing the sodium acetate content in L to 1 g/L, with all other conditions the same as C. H represents tripling the non-carbon source substances in the L medium, with all other conditions the same as C. CFU stands for colony-forming unit. For each time point, the mean and standard deviation based on the 10 replicate subpopulations are shown separately

gains for the lac-, thereby hastening their probability of extinction. These experimental outcomes underscore those mutations favoring a high-fitness direction, such as lac+, competitively exclude mutations favoring a lower-fitness direction, like lac-, affirming the occurrence of Darwinian selection between directions.

In our culture medium, there is differentiation in the utilization of carbon sources between lac+and lac-, with each occupying distinct ecological niches related to carbon sources. Why can't they coexist? The reason may lie in the competition between them despite their differentiation in carbon source utilization. For instance, in the L medium, sodium acetate and lactose serve as private sources for lac- and lac+, respectively. However, other resources such as nitrogen sources, metal elements, and physical space are shared resources where there is no differentiation between the two. Therefore, the inability of both to coexist may stem from their competition for these shared resources. Lac+utilizing lactose may have a relatively higher growth rate than lac- utilizing sodium acetate, potentially giving it a competitive advantage in accessing shared resources and outcompeting lac-. Our findings are consistent with previous studies showing that, in homogeneous environments, competition leads to a rapid decline in diversity [35–39].

A conceptual extension

Darwinian selection is generally considered the primary mechanism driving adaptive evolution [40]. Traditional Darwinian selection emphasizes the selection and elimination of genetic variations within a particular adaptive direction. Darwin's own examples in *On the Origin of Species*—such as faster wolves being more successful hunters or flowers producing more nectar to attract more pollinators—illustrate this selection [41]. The result of this selection leads to enhanced adaptive evolution of species in a particular direction. In this study, our results show that the two mutants, lac+and lac-, utilizing different carbon sources, exhibit different fitness gains, leading to the competitive elimination of the mutant with lower fitness gains by the one with higher gains. This indicates that Darwinian selection can occur between mutations in different adaptive directions, consequently leading to the preservation of adaptive directions with high fitness gains and the elimination of those with lower gains, thereby resulting in biased species evolution, consistent with previous studies [17–22]. This Darwinian selection occurring between different adaptive directions is largely different from traditional Darwinian selection. Therefore, our results indicate the existence of two distinct forms of Darwinian selection: one occurring between different adaptive directions and the other occurring within a particular adaptive direction. For convenience, we refer to these two forms of selection as "inter-directional selection" and "intra-directional selection," respectively. Traditional Darwinian selection, or intra-directional selection, cannot explain why a species adapts toward a particular direction; it only accounts for adaptive evolution in a certain direction. In contrast, inter-directional selection can explain why a species evolves adaptively toward a particular direction while not adapting toward others, thus indicating evolutionary bias. Therefore, distinguishing these two forms of Darwinian selection is fundamentally important for our understanding of how the adaptive direction of species is formed.

In this study, we primarily analyzed the significant role of inter-directional selection in determining the evolutionary direction of species. Inter-directional selection can lead to the elimination of directions with low fitness returns and the preservation of directions with high fitness returns. Our findings indicate that directions with high fitness returns competitively eliminate those with low fitness returns, resulting in their extinction, consistent with previous studies [18, 20–22]. Apart from intraspecific competition mentioned here, previous research suggests that for sexually reproducing organisms, factors such as genetic swamping and outbreeding depression may also contribute to the destruction of local adaptation [2–4, 9–13]. This suggests that these genetic disturbance from directions with high fitness returns may further exacerbate the demise of directions with low fitness returns.

Individuals in directions with high fitness returns potentially have faster growth rates and may contribute to larger populations, and its negative effects (e.g. competition and genetic swamping) on directions with low fitness returns may outweigh any impact in the reverse direction [3]. This asymmetry in negative genetic and ecological effects may ultimately lead to the extinction of directions with low fitness returns, as directions with high fitness returns consistently exert evolutionary suppression over them [18, 20–22]. Consequently, species as a whole can only evolve towards directions with high fitness returns due to this evolutionary inhibition exerted by directions with high fitness returns. Thus, inter-directional selection generates a force that directs species evolution towards directions with high fitness returns. Consistent with our findings, a long-term experimental evolution study reported that when glucose was scarce, *E. coli* underwent evolutionary transition towards the consumption of sodium citrate [42, 43]. Mutants unable to consume sodium citrate went extinct after coexisting with mutants able to consume sodium citrate for approximately 10,000 generations [44, 45], as the latter had relatively higher fitness gains [44]. Similarly, numerous studies have reported that scarcity of one resource can lead to adaptive evolutionary transitions of bacteria towards the utilization of alternative resources [46–51].

Furthermore, previous theoretical and experimental evolutionary research has also shown that in heterogeneous environments, organisms tend to adapt towards the most productive resource [18–22]. All of these studies, along with our findings, suggest that inter-directional selection plays a pivotal role in shaping the adaptive evolution of species towards directions with high fitness gains.

This directional force from inter-directional selection makes each species' evolutionary trajectory appear biased towards a specific direction, devoid of deviation towards other directions, thereby exhibiting evolutionary bias. This aligns with previous findings that in heterogeneous environments, species tend to adapt evolutionarily towards the most productive resources [17–22]. Inter-directional selection could generate an evolutionary force causing each species' evolution to appear attracted to particular directions. Given this force's "attraction" characteristic, it could be termed "evolutionary gravitation," drawing an analogy to gravitational forces in physics. Evolutionary gravitation ensures that a species' evolutionary trajectory always points towards directions with high fitness gain. Unlike the concept of Darwinian selection, the concept of evolutionary gravitation emphasizes the tendency of a species' evolution towards a particular direction. It provides a more intuitive explanation for why species evolve towards specific directions. For example, aphids commonly evolve to prefer eating young and tender leaves while avoiding older leaves, we could say this is due to the stronger evolutionary gravitation of young leaves for aphids, which potentially offer higher fitness returns to them.

Conclusion

Our findings suggest that species generally exhibit an evolutionary bias, tending to adapt towards directions that offer higher fitness gains. The competitive exclusion of directions with lower fitness gains by those with higher fitness gains plays a crucial role in driving this bias. Darwinian selection acting on mutations across different adaptive directions propels species towards paths with high fitness gains, either transitioning from low fitness-gain directions to high ones or maintaining an evolutionary inclination towards high fitness without deviation. Thus, our findings highlight that the evolutionary constraint effects of high-fitness-gain directions on low-fitness-gain directions are a significant factor contributing to evolutionary bias.

Materials and methods

Bacterial strain

The *E. coli* strain used in this experiment is *E. coli* K-12 GM4792. This strain is asexual and its genome has been publicly disclosed, making it suitable for experimental evolution studies [28]. *E. coli* K-12 GM4792 differs from other closely related strains in that it has a continuous 212 bp deletion in its lactose operon region, and its *lacI* and *lacZ* fusion genes has a frameshift mutation caused by a single base insertion, rendering it unable to utilize lactose (lac-) [28]. However, previous studies have shown that when cultured in a medium where lactose is the sole carbon source, this strain undergoes revertant mutations, resulting in the emergence of lactose-utilizing mutants, lac+ [28, 29]. The presence of lac+and lac- can be easily distinguished through blue-white screening [30].

Culture conditions

We retrieved *E. coli* K-12 GM4792 from the −80 °C freezer and allowed it to thaw at room temperature. Subsequently, we conducted three consecutive liquid cultures and streak plate cultures on LB agar, selecting a single clone each time. The culture from the third picking was designated as the ancestral strain (lac-(ancestor)) for the experiment, and we stored multiple backups in glycerol at -80 °C. During blue-white screening, the ancestral strain displayed white colonies, indicating its inability to metabolize lactose. Furthermore, partial gene sequencing of the ancestral strain identified a 212 bp deletion, confirming its consistency with the strain utilized in the experiment.

The ancestral strain was cultured in two distinct media types (L and G) for the experimental evolution, with 5 replicates for each medium, totaling 10 experimental populations (L1-5, G1-5). The composition of 1 L of L medium comprised basic M9 salts (Na_2HPO_4 6.8 g, KH_2PO_4 3.0 g, NaCl 0.5 g, NH_4Cl 1.0 g), supplemented with 2 ml of 1.0 M $MgSO_4$ solution, 0.1 ml of 1.0 M $CaCl_2$ solution, 10 g sodium acetate, and 20 g lactose. The G medium mirrored the L medium but substituted 10 g sodium acetate with 50 g glucose. All reagents were sourced from Sangon Biotech, Shanghai. In the L medium, sodium acetate and lactose functioned as carbon sources, whereas in the G medium, lactose and glucose served as the carbon sources.

Every 24 h, 100 µl of culture from each population was transferred to 9.9 ml of fresh corresponding medium in 50 ml plastic tubes and then placed in an incubator at 37 °C with shaking at 220 rpm. Blue-white screening was conducted on the cultures every 5 days, with a tally of the blue and white colonies (each population was spread on two plates). Subsequently, a 4 ml aliquot of culture mixed with an equal volume of 50% glycerol was stored at -80 °C every 5 days for preservation purposes.

Gene sequencing

We conducted colony PCR amplification and sequencing on various segments of the lactose operon to identify target strains or analyze sequence evolution. For strain identification, we utilized two sets of primers at different

experimental stages. The first primer set was employed to determine if the target bacterial culture harbored a 212 bp deletion of *E. coli* K-12 GM4792. The PCR amplification primers used were LacIF1: TATCCCGCCGT TAACCACCATCAAAC and LaczR1: CTTCCTGTA GCCAGCTTTCATCAACAT. The amplification setup consisted of a 25 μl reaction volume, comprising 12.5 μl of PrimeSTAR Max premix (Takara, Beijing), 0.75 μl of each primer, 1 μl of template (colony), and 10 μl of sterile water. The PCR amplification process involved an initial denaturation at 98 °C for 60 s, followed by 35 cycles of denaturation at 98 °C for 10 s, annealing at 60 °C for 15 s, extension at 72 °C for 5 s, and a final extension at 72 °C for 2 min. Subsequently, the PCR products were detected by 1% agarose gel electrophoresis and were sent for sequencing (ABI3730, GENEWIZ) using the same primers as the PCR amplification. The second primer set was utilized to assess whether the target bacterial strain carried the 212 bp deletion and mutation region of *E. coli* K-12 GM4792. The PCR amplification primers used were LacIF2: CATCTGGTCGCATTGGGTCA and LaczR2: CCAGTTTGAGGGGACGACGACAGT. The amplification system and conditions were akin to the first set, except that the annealing temperature during the 35 cycles was set at 55 °C, and the extension time was 10 s.

We performed full-length sequencing of the lactose operon region. For sample selection, considering the changes in the blue/white clone ratio over time, we chose samples that contained blue and white clones at the same time point for the experiment. Specifically, we selected frozen glycerol stocks from day 15 for populations L1 and L4, and frozen glycerol stocks from day 10 for populations L2, L3, and L5. From each population, we selected one white clone and five blue clones; additionally, we included lac-(ancestor). For these 31 samples, we cultured them overnight and then directly performed colony PCR to amplify the target fragments. The PCR primers used were LacIF: CCATCGAATGGCGCAAAACCTTT C and LacAR: TGCCGGATGCGGCTAATGTAGATC. The PCR amplification system was 25 μl, which included 12.5 μl of PrimeSTAR Max premix (Takara, Beijing), 0.75 μl of each primer, 1 μl of template (colony), and 10 μl of sterilized water. The PCR amplification conditions included an initial denaturation at 98 °C for 60 s, followed by 35 cycles of: denaturation at 98 °C for 10 s, annealing at 55 °C for 15 s, and extension at 72 °C for 60 s, with a final extension at 72 °C for 2 min. After detecting the PCR products with 1% agarose gel, they were sent for sequencing (ABI3730, GENEWIZ) with the relevant sequencing primers (Table S2).

Genome sequencing

We conducted whole-genome sequencing on the ancestral strain used in this study, as well as on single clones from 10 populations that had undergone experimental evolution for 25 days. For the ancestral strain, we performed both second- and third-generation sequencing and de novo assembly of its genome, resulting in a complete genome. For the single clones from the 10 experimental evolution populations, we conducted second-generation genome resequencing.

We conducted overnight plate cultures of the ancestral strain and the 10 populations from day 25 of experimental evolution. The following day, a single clone was randomly selected from the ancestral strain plate, and two single clones were picked from each of the 10 population plates, resulting in a total of 21 samples, which were then placed in G medium for overnight culture. PCR amplification sequencing was performed on these samples to confirm the presence of a 212 bp deletion, indicating that they belonged to the strains used in this study. Subsequently, 5 milliliters of each of the 21 samples were used for DNA extraction using the Bacteria Genomic DNA Extraction Kit (TaKaRa, Beijing) following the provided instructions.

We conducted second-generation genome sequencing on 21 samples using the DNBSEQ platform (BGI, Shenzhen). Initially, we randomly fragmented the genomic DNA of the samples, recovered the necessary DNA fragments through electrophoresis, added adapters for cluster preparation, and finally sequenced them using the MGISEQ-2000 platform. We utilized the SOAPnuke software [52] to filter raw reads (150 bp) and generate high-quality clean reads, which excluded low-quality sequences and adapter sequences.

For our ancestral strain, in addition to second-generation sequencing, we also performed third-generation genome sequencing using the PacBio sequencing platform (BGI, Shenzhen). We processed its DNA into appropriately-sized fragments (20–40 K) using g-TUBE, followed by fragment damage repair and end repair. Adapters with hairpin structures were attached to both ends of the DNA fragments to form a dumbbell structure known as SMRTbell. The annealed smrtbell and the polymerase at the bottom of the ZWM were mixed and used for the final sequencing (PacBio Sequel II). Due to the presence of low-quality sequences and adapter sequences in the original sequencing data, it was necessary to filter out these impurities to obtain reliable subreads. Multiple subreads from a single pore were used to obtain a circular consensus sequencing (CCS) data. We utilized the FalconConsensus [53] and Proovread [54] software to generate corrected reads, which were then used for sequence assembly using the Canu software [55]. Finally, we employed the GATK software [56, 57] for single nucleotide correction using second-generation sequencing short sequences to obtain reliable assembled genome sequences.

SNP, indel analyses and phylogenetic relationships construction

We detected SNPs and indels by aligning the short reads of 20 single clone samples from evolving populations with the complete genome sequence of the ancestral strain, used as a reference. We employed the Burrows-Wheeler Aligner [58, 59] to align the clean data of each sample to the reference genome, resulting in initial alignment results in BAM format. To ensure accurate variant detection, we followed the optimal variant detection analysis workflow recommended by the official GATK website (Genome Analysis Toolkit, GATK). Following alignment, Picard tools (http://broadinstitute.github.io/picard/) were used to remove duplicate reads, and the software GATK [56, 57] was employed for local realignment and base quality recalibration. Evaluation metrics such as sequencing depth, coverage, and alignment rate were statistically analyzed based on the alignment results for each sample.

The genome-level SNP and indel matrix of the 21 samples was then used to construct a neighbor-joining (NJ) tree using MEGA X [60], with the Jukes-Cantor model selected as the best-fit model and bootstrap values set to 1000. Additionally, we reconstructed a maximum likelihood (ML) tree using IQ-TREE [61] based on the genome-level SNP data, selecting K2P + ASC as the best-fit model. The ML tree exhibited high similarity to the NJ tree, and only the NJ tree was presented.

Fitness assay in L culture medium

We compared the relative fitness differences between lac- and lac+in L medium. Considering the changes in blue-white clone ratios over time in L1-5, we selected samples with both blue and white clones at specific time points for the experiment. Specifically, frozen glycerol bacteria saved on day 15 were chosen for L1 and L4, while those saved on day 10 were selected for L2, L3, and L5. After thawing the frozen glycerol bacteria at room temperature, a certain amount of bacterial solution was diluted and subjected to blue-white screening. Two blue single clones and two white single clones were selected from each of L1-5, totaling 20 single clones. Each single clone was then placed in 10 ml of L medium for liquid culture at 37 °C and 220 rpm for 24 h to achieve a similar physiological state. After 24 h, the OD values of these 20 subpopulations were measured three times. Based on the OD values and following a blue: white ratio of 10 : 1 and 1 : 10, two blues and two whites from the same experimental population were paired into competition pairs, resulting in a total of 20 pairs. The paired blue-white bacterial solutions were mixed and prepared as 100 μl bacterial solutions added to 9.9 ml of L medium in 50 ml plastic tubes. Immediately after mixing, bacterial solutions were diluted and plated (3 plates per competition pair), totaling 60 plates. After incubation at 37 °C for 24 h, blue-white screening was performed to count the number of blue and white colonies, calculating the density of blue-white colonies before competition. After 24 h of competition, the same method was used to calculate the density of blue-white colonies after competition. Based on the densities of blue and white colonies before and after competition, the selection rate constant r_{ij} was calculated, which is given by

$$r_{ij} = \frac{\ln\left[N_i(1)/N_i(0)\right] - \ln\left[N_j(1)/N_j(0)\right]}{1\ \mathrm{day}}$$

where $N_i(0)$ and $N_j(0)$ are the initial densities of lac+and lac-, respectively, and $N_i(1)$ and $N_j(1)$ are their densities after 24 h [31, 32].

Fitness assay in G culture medium

We conducted an analysis of relative fitness differences between lac- and lac+in G medium. Since no lac+ (blue clones) were found in the G-populations during the 25-day experimental evolution period, we compared the relative fitness differences between lac- in the G-populations and lac+from the L-populations. Blue-white screening was performed on the frozen glycerol bacteria from the G and L-populations on the 25th day. Two white single clones were selected from each of the five G-populations, and two blue single clones were selected from each L-population, resulting in a total of 20 single clones. These clones were individually placed in 10 ml of G medium and cultured at 37 °C and 220 rpm for 24 h to achieve similar physiological states. After 24 h, the OD values of these 20 subpopulations were measured three times each. We adjusted the OD values to set the blue-white colony ratio of the competitive pairs at 10 : 1 and 1 : 10. We paired 10 lac- and 10 lac+into 10 competition pairs, totaling 20 pairs. Each competition pair was mixed and prepared as 100 μl of bacterial solution added to 9.9 ml of fresh G medium in 50 ml plastic tubes. The tubes were then cultured at 37 °C and 220 rpm for 24 h. Immediately after mixing, bacterial solutions were diluted and plated (3 plates per competition pair), followed by incubation at 37 °C for 24 h and subsequent blue-white screening to count the number of blue and white colonies for density calculation. After 24 h of competition, the density of blue and white colonies was calculated using the same method. The selection rate constant, r_{ij}, was calculated in the same manner as described above, with N_i and N_j representing lac- and lac+, respectively.

Competitive exclusion experiment

To investigate competitive exclusion effect, we conducted competition experiments between lac- and lac+under varied experimental conditions. From each

of the 5 L-populations, two white clones (lac-) and two blue clones (lac+) were selected, resulting in a total of 10 white clones and 10 blue clones. Among these, samples preserved in glycerol on the 15th day were chosen for L1 and L4, while samples preserved on the 10th day were selected for L2, L3, and L5 for the experiment. These 20 clones were cultured in L medium at 37 °C and 220 rpm for 24 h. Subsequently, two blue and two white clones from the same population were paired to create 10 competition pairs (10 replicates). In each pair, lac- and lac+were mixed in a 1 : 1 ratio (50 μl : 50 μl) by volume, generating 4 parallel subpopulations for each competition pair, totaling 40 subpopulations. These four sets were then subjected to competition experiments under four distinct conditions.

The first group (Group C) was cultured under the same conditions as the experimental evolution culture conditions used in this study, with each competition pair placed in 10 ml of L medium at 37 °C and 220 rpm for shaking incubation. The second group (Group J) had similar conditions to Group C, except that the cultures were kept stationary. The third group (Group L) maintained the same conditions as Group C, except that the amount of sodium acetate was reduced from 10 g to 1 g. The fourth group had conditions akin to Group C, but with the mass of other nutrients tripled while keeping the masses of sodium acetate and lactose constant. The final culture volume was increased from 10 ml to 30 ml, with shaking incubation at 37 °C and 220 rpm. Every 24 h for each of these four experimental groups, 100 μl of the culture was sampled and transferred into fresh medium (9.9 ml for Groups C, J, L, and 29.9 ml for Group H) for subsequent generations. For the 40 experimental subpopulations, samples were taken before competition and every 24 h thereafter, and subjected to blue-white screening to count the number of blue and white colonies until no white colonies appeared on two consecutive plating rounds.

Carbon source consumption measurement

We evaluated the consumption of carbon sources by lac- and lac+in L and G media. The frozen glycerol bacteria from the 15th day were used for the experiment. Through blue-white screening, we selected one blue clone (L3-B) and one white clone (L3-W) from the L3 population, and one white clone (G3-W) from the G3 population, making a total of three single clones for the experiment. Initially, these three single clones were cultured overnight in LB liquid medium. The next day, 0.1 ml of overnight cultures of L3-W and L3-B were each added to 9.9 ml of fresh L medium for cultivation, while 0.1 ml of overnight cultures of L3-B and G3-W were added to 9.9 ml of fresh G medium for cultivation, resulting in a total of four experimental populations cultured for 24 h. On the third day, bacterial cultures from these four experimental populations were streaked, single clones were picked again, and placed in their respective 10 ml fresh culture medium for another 24 h of cultivation. This 48-hour physiological conditioning ensured that lac- and lac+reached similar physiological states in their respective culture media. For the formal experiment, after the 48-hour conditioning period, 100 μl of bacterial culture from each of the four experimental populations was added to the corresponding 9.9 ml culture medium, with five replicates for each population, and cultivated at 37 °C and 220 rpm for 24 h. At intervals of 3, 6, 9, 12, and 24 h, we alternated in extracting 1 ml samples from one of the five replicates, stored at -20 °C, and sent for liquid chromatography detection the following day. To ensure result reliability, PCR amplification sequencing was conducted on the three single clone bacterial cultures used in this experiment before starting to confirm a 212 bp deletion, validating the intended strains for the study. Additionally, to prevent potential cross-contamination between lac+and lac-, blue-white screening tests were carried out on the bacterial cultures before and after the 24-hour experiment to verify the absence of any cross-contamination among the four experimental populations. No mixed contamination was found among the total of 6483 colonies observed, eliminating the possibility of cross-contamination and evolutionary shift from lac- to lac+.

For the measurement of acetate content, detection was carried out using a high-performance liquid chromatography system (Agilent 1200) with a Welch Ultimate AQ C18 column (5 μm, 250 mm × 4.6 mm). The samples were diluted 50 times with pure water and directly tested on the machine. The mobile phase used was potassium dihydrogen phosphate: methanol=99 : 1, with a column temperature of 40 °C, flow rate of 0.5 ml/min, detection wavelength of 210 nm, and an injection volume of 10 μl.

For the measurement of lactose and glucose, detection was performed using a high-performance liquid chromatography system (Agilent 1200) equipped with a differential refractive index detector and an Agilent ZORBAX NH2 column (5 μm, 250 mm × 4.6 mm). The sample solution was diluted 20 times with pure water and directly tested on the machine. The mobile phase used was water: acetonitrile=30 : 70, with a column temperature of 40 °C, flow rate of 1 ml/min, and an injection volume of 20 μl.

Growth curve measurement

We measured the growth of lac- and lac+in L medium. One white clone (lac-) and one blue clone (lac+) were selected from each of the five L-populations, resulting in a total of 5 white clones and 5 blue clones for the experiment. For populations L1 and L4, frozen glycerol bacteria saved on the 15th day were chosen, while for populations L2, L3, and L5, frozen glycerol bacteria saved on the 10th

day were used. The 10 clones were each cultured with shaking at 37 °C and 220 rpm in L medium for 24 h to achieve a similar physiological state. The OD values of these 10 clone cultures were measured three times each, and based on the OD values, the initial OD values of the 10 clone cultures were adjusted to be the same. Subsequently, 350 µl of bacterial culture was taken from each of the 10 clone cultures and added to 10 tubes of L medium, making up to 35 ml, and then incubated with shaking at 37 °C and 220 rpm for 24 h. Every 2 h, 1 ml was transferred from each tube to a sterile 1.5 ml centrifuge tube and stored in a refrigerator at 4 °C. Once all samples were collected, the OD600 of each sample was measured sequentially using a UV-visible spectrophotometer (Shimadzu UV-1240) in triplicate, with consistent results for each measurement.

Statistical analyses

We conducted data analysis using the R statistical software. We employed a one-sample t-test to examine the two-tailed probability of potential deviation of the selection rate constant (r_{ij}) from the null hypothesis that the selection rate constant is zero, indicating equal fitness for lac+and lac-. For the growth rate analysis of lac- and lac+in L medium, as well as the competition experiment results of lac- and lac+under four different conditions, we utilized repeated measures ANOVA to assess the statistical significance of corresponding variables at various time points. Prior to the analysis, we performed a normality test on the samples, which indicated a normal or near-normal distribution for all cases. Furthermore, the homogeneity of variance test ($p>0.05$) was conducted, and pairwise comparisons were adjusted using the Holm method for p values.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12862-024-02331-1.

Supplementary Material 1

Acknowledgements
We appreciate Professor Quanguo Zhang from Beijing Normal University for providing us with the experimental strain *E. coli* K-12 GM4792. We express our gratitude to Mengying Liu for her assistance with sterile procedures during bacterial culture. Additionally, we extend our thanks to Hao Tang, Yintian Liu, and Longcheng Fan for their support in the molecular experiments. Special thanks to Chaonan Xu for preparing reagents and consumables, and to Qiang Han for performing component analyses using High-Performance Liquid Chromatography.

Author contributions
Y. W. conceived and designed the research, performed data analyses, and wrote the paper. H.W. conducted the experiment, performed data analyses, and wrote the paper.

Funding
This research was supported by the National Natural Science Foundation of China (grant number 32171604).

Data availability
The genome sequencing data were deposited into the National Center for Biotechnology Information Sequence Read Archive database with BioProject accession number (PRJNA1099467). The full-length sequences of lactose operon were deposited into GenBank with accession numbers (PP693326-56). All other data needed to evaluate the conclusions in the paper are present in the main text or the supplementary materials.

Declarations

Ethical approval
All experimental procedures in this study were conducted in accordance with the Declaration of Laboratory Biosafety Guidance of the Northeast Normal University (approval number NENU-202292).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 7 August 2024 / Accepted: 21 November 2024
Published online: 02 December 2024

References

1. Wiens JJ, Graham CH. Niche conservatism: integrating evolution, ecology, and conservation biology. Annu Rev Ecol Evol Syst. 2005;36:519–39.
2. Bridle JR, Vines TH. Limits to evolution at range margins: when and why does adaptation fail? Trends Ecol Evol. 2007;22:140–7.
3. Kirkpatrick M, Barton NH. Evolution of a species' range. Am Nat. 1997;150:1–23.
4. Mayr E. Animal Species and Evolution. Cambridge: Harvard University Press; 1963.
5. Grant PR, Grant BR. How and Why Species Multiply: the Radiation of Darwin's Finches. Princeton: Princeton University Press; 2007.
6. Losos JB. Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles. Berkeley: University of California Press; 2009.
7. Barlow G. The Cichlid Fishes: Nature's Grand Experiment in Evolution. New York: Basic Books; 2000.
8. Coyne J, Orr H. Speciation. New York: Oxford University Press; 2004.
9. Postma E, van Noordwijk AJ. Gene flow maintains a large genetic difference in clutch size at a small spatial scale. Nature. 2005;433:65–8.
10. Adavoudi R, Pilot M. Consequences of hybridization in mammals: a systematic review. Genes. 2022;13:50.
11. Lenormand T. Gene flow and the limits to natural selection. Trends Ecol Evol. 2002;17:183–9.
12. Moerman F, Fronhofer EA, Wagner A, Altermatt F. Gene swamping alters evolution during range expansions in the protist *Tetrahymena thermophila*. Biol Lett. 2020;16:20200244.
13. Todesco M, et al. Hybridization and extinction. Evol Appl. 2016;9:892–908.
14. Bull C. Ecology of parapatric distributions. Annu Rev Ecol Syst. 1991;22:19–36.
15. Case TJ, Holt RD, McPeek MA, Keitt TH. The community context of species' borders: ecological and evolutionary perspectives. Oikos. 2005;108:28–46.
16. Sexton JP, McIntyre PJ, Angert AL, Rice KJ. Evolution and ecology of species range limits. Annu Rev Ecol Evol Syst. 2009;40:415–36.
17. Hall AR, Colegrave N. How does resource supply affect evolutionary diversification? Proc R Soc B: Biol Sci. 2007; 274:73–8.
18. Holt RD, Gaines MS. Analysis of adaptation in heterogeneous landscapes: implications for the evolution of fundamental niches. Evol Ecol. 1992;6:433–47.
19. Jasmin JN, Kassen R. On the experimental evolution of specialization and diversity in heterogeneous environments. Ecol Lett. 2007;10:272–81.
20. Holt RD. On the evolutionary ecology of species' ranges. Evol Ecol Res. 2003;5:159–78.

21. Kawecki T. Demography of source—sink populations and the evolution of ecological niches. Evol Ecol. 1995;9:38–44.
22. Kawecki TJ, Barton NH, Fry JD. Mutational collapse of fitness in marginal habitats and the evolution of ecological specialisation. J Evol Biol. 1997;10:407–29.
23. Bren A, et al. Glucose becomes one of the worst carbon sources for *E. coli* on poor nitrogen sources due to suboptimal levels of cAMP. Sci Rep. 2016;6:24834.
24. Seong W, et al. Adaptive laboratory evolution of *Escherichia coli* lacking cellular byproduct formation for enhanced acetate utilization through compensatory ATP consumption. Metab Eng. 2020;62:249–59.
25. Boag PT, Grant PR. Intense natural selection in a population of Darwin's finches (Geospizinae) in the Galapagos. Science. 1981;214:82–5.
26. Kawecki TJ, et al. Experimental evolution. Trends Ecol Evol. 2012;27:547–60.
27. McDonald MJ. Microbial experimental evolution–a proving ground for evolutionary theory and a tool for discovery. EMBO Rep. 2019;20:e46992.
28. Zhang Y-C, et al. Genome sequences of two closely related strains of *Escherichia coli* K-12 GM4792. Stand Genomic Sci. 2015;10:125.
29. Foster PL, Trimarchi JM. Adaptive reversion of an episomal frameshift mutation in *Escherichia coli* requires conjugal functions but not actual conjugation. Proc Natl Acad Sci U S A. 1995;92:5487–90.
30. Zhao L, Zhang Q-G, Zhang D-Y. Evolution alters ecological mechanisms of coexistence in experimental microcosms. Funct Ecol. 2016;30:1440–6.
31. Lenski RE, Rose MR, Simpson SC, Tadler SC. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. Annu Rev Ecol Evol Syst. 1991;138:1315–41.
32. Travisano M, Lenski RE. Long-term experimental evolution in *Escherichia coli*. IV. Targets of selection and the specificity of adaptation. Genetics. 1996;143:15–26.
33. Görke B, Stülke J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nat Rev Microbiol. 2008;6:613–24.
34. Aidelberg G, et al. Hierarchy of non-glucose sugars in *Escherichia coli*. BMC Syst Biol. 2014;8:133.
35. Frenkel EM et al. Crowded growth leads to the spontaneous evolution of semistable coexistence in laboratory yeast populations. Proc Natl Acad Sci U S A. 2015; 112:11306–11311.
36. Habets MG, Rozen DE, Hoekstra RF, De Visser JAG. The effect of population structure on the adaptive radiation of microbial populations evolving in spatially structured environments. Ecol Lett. 2006;9:1041–8.
37. Korona R, Nakatsu CH, Forney LJ, Lenski RE. Evidence for multiple adaptive peaks from populations of bacteria evolving in a structured habitat. Proc Natl Acad Sci U S A. 1994;91:9037–41.
38. Rainey PB, Travisano M. Adaptive radiation in a heterogeneous environment. Nature. 1998;394:69–72.
39. Traverse CC, Mayo-Smith LM, Poltak SR, Cooper VS. Tangled bank of experimentally evolved Burkholderia biofilms reflects selection during chronic infections. Proc Natl Acad Sci U S A. 2013;110:E250–9.
40. Futuyma DJ. Evolution. Sunderland (Massachusetts): Sinauer Associates; 2013.
41. Darwin C. On the Origin of Species by Means of Natural Selection. London: John Murray; 1859.
42. Blount ZD, Barrick JE, Davidson CJ, Lenski RE. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. Nature. 2012;489:513–8.
43. Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. Proc Natl Acad Sci U S A. 2008;105:7899–906.
44. Leon D, D'Alton S, Quandt EM, Barrick JE. Innovation in an *E. coli* evolution experiment is contingent on maintaining adaptive potential until competition subsides. PLoS Genet. 2018;14:e1007348.
45. Turner CB, Blount ZD, Lenski RE. Replaying evolution to test the cause of extinction of one ecotype in an experimentally evolved population. PLoS ONE. 2015;10:e0142050.
46. Friesen ML, Saxer G, Travisano M, Doebeli M. Experimental evidence for sympatric ecological diversification due to frequency-dependent competition in *Escherichia coli*. Evolution. 2004;58:245–60.
47. Herron MD, Doebeli M. Parallel evolutionary dynamics of adaptive diversification in *Escherichia coli*. PLoS Biol. 2013;11:e1001490.
48. Lee D-H, Palsson BØ. Adaptive evolution of *Escherichia coli* K-12 MG1655 during growth on a nonnative carbon source, L-1, 2-propanediol. Appl Environ Microbiol. 2010;76:4158–68.
49. Meyer JR, et al. Repeatability and contingency in the evolution of a key innovation in phage lambda. Science. 2012;335:428–32.
50. Rosenzweig RF, Sharp R, Treves DS, Adams J. Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli*. Genetics. 1994;137:903–17.
51. Van Hofwegen DJ, Hovde CJ, Minnich SA. Rapid evolution of citrate utilization by *Escherichia coli* by direct selection requires *citT* and *dctA*. J Bacteriol. 2016;198:1022–34.
52. Chen Y, et al. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. Gigascience. 2018;7:gix120.
53. Yang A, Troup M, Lin P, Ho JW. Falco: a quick and flexible single-cell RNA-seq processing framework on the cloud. Bioinformatics. 2017;33:767–9.
54. Hackl T, Hedrich R, Schultz J, Förster F. Proovread: large-scale high-accuracy PacBio correction through iterative short read consensus. Bioinformatics. 2014;30:3004–11.
55. Koren S, et al. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 2017;27:722–36.
56. DePristo MA, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43:491–8.
57. McKenna A, et al. The genome analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297–303.
58. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009;25:1754–60.
59. Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics. 2010;26:589–95.
60. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35:1547–9.
61. Nguyen L-T, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–74.

Publisher's note