

Genetic Context Significantly Influences the Maintenance and Evolution of Degenerate Pathways

Eric L. Bruger^{1,2,3,4} , Lon M. Chubiz^{5,6}, José I. Rojas Echenique^{5,7}, Caleb J. Renshaw^{1,2}, Nora Victoria Espericueta^{1,8}, Jeremy A. Draghi⁹, Christopher J. Marx^{1,2,3,4,5,*}

¹Department of Biological Sciences, University of Idaho, Moscow, Idaho, USA

²Institute for Modeling Collaboration and Innovation, University of Idaho, Moscow, Idaho, USA

³Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, Idaho, USA

⁴The BEACON Center for the Study of Evolution in Action, University of Idaho, Moscow, Idaho, USA

⁵Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts, USA

⁶Department of Biology, University of Missouri, St. Louis, Missouri, USA

⁷Department of Molecular Genetics, University of Toronto, Ontario, Canada

⁸Department of Biological Sciences, California State University, Long Beach, California, USA

⁹Department of Biological Sciences, Virginia Institute of Technology, Blacksburg, Virginia, USA

*Corresponding author: E-mail: cmarx@uidaho.edu.

Accepted: 19 April 2021

Abstract

Understanding the evolution of novel physiological traits is highly relevant for expanding the characterization and manipulation of biological systems. Acquisition of new traits can be achieved through horizontal gene transfer (HGT). Here, we investigate drivers that promote or deter the maintenance of HGT-driven degeneracy, occurring when processes accomplish identical functions through nonidentical components. Subsequent evolution can optimize newly acquired functions; for example, beneficial alleles identified in an engineered *Methylobacterium extorquens* strain allowed it to utilize a “Foreign” formaldehyde oxidation pathway substituted for its *Native* pathway for methylotrophic growth. We examined the fitness consequences of interactions between these alleles when they were combined with the *Native* pathway or both (*Dual*) pathways. Unlike the *Foreign* pathway context where they evolved, these alleles were often neutral or deleterious when moved into these alternative genetic backgrounds. However, there were instances where combinations of multiple alleles resulted in higher fitness outcomes than individual allelic substitutions could provide. Importantly, the genetic context accompanying these allelic substitutions significantly altered the fitness landscape, shifting local fitness peaks and restricting the set of accessible evolutionary trajectories. These findings highlight how genetic context can negatively impact the probability of maintaining native and HGT-introduced functions together, making it difficult for degeneracy to evolve. However, in cases where the cost of maintaining degeneracy was mitigated by adding evolved alleles impacting the function of these pathways, we observed rare opportunities for pathway coevolution to occur. Together, our results highlight the importance of genetic context and resulting epistasis in retaining or losing HGT-acquired degenerate functions.

Key words: genomic background, epistasis, degeneracy, methylotrophy, HGT.

Introduction

The means by which novel functions like newly acquired metabolic pathways evolve is an open and highly relevant question for numerous diverse biological systems (Ding et al. 2012; Louca et al. 2018; Shoji 2019). One path to novelty is the

diversification of redundant pathways from a common ancestral function (Hughes 1994; Zhang 2003; Juhas et al. 2009; Näsvall et al. 2012). Alternatively, new functions can be obtained via horizontal gene transfer (HGT). Degeneracy occurs when pathways or processes execute the same

Significance

Although genetic interactions between mutations in evolving lineages of microbes have been investigated, these have not included how evolution impacts the subsequent acquisition of genes via horizontal gene transfer, HGT and thus how these elements interact with existing alleles. We find that the patterns of interactions between evolved alleles appear substantially different when moved into alternative genetic backgrounds. Additionally, these preexisting alleles were found to have strong impacts on the ability of genotypes to maintain and coevolve with functionally redundant genes and pathways, showing that even though they evolved separately, the particular alleles in a genetic background, and the physiological impacts they confer, weigh heavily on whether genes for new horizontally acquired metabolic processes are maintained.

function using different components (such as nonhomologous proteins) in an organism, and in many instances degenerate pathways are acquired by HGT (Tononi et al. 1999; Edelman and Gally 2001; Whitacre 2010; Mason et al. 2015; Rezazadegan and Reidys 2018; Yu and Whalen 2020). Although degeneracy is a prominent feature of biological systems, such as alternative pathways for glucose metabolism (Pfeiffer et al. 2001; Fuhrer et al. 2005), the evolutionary routes to achieving degeneracy are not well understood (Nowak et al. 1997; Lerner et al. 2017; Emamalipour et al. 2020; Hall et al. 2020). In this study, we seek to gain a better grasp of what evolutionary forces impact degeneracy (fig. 1), as it relates to cellular metabolism, by characterizing the factors that drive the maintenance of native and/or heterogenous pathways following an HGT event.

Although it is well understood that barriers to HGT exist—including toxicity, bacterial restriction and defense systems, decreased homologous recombination between diverse species, metabolic costs and mutational load, and phylogenetic barriers—it remains unclear how substantial a role the recipient genetic background plays in this process (Majewski et al. 2000; Kurland 2005; Thomas and Nielsen 2005; Sorek et al. 2007; Oliveira et al. 2017; Hendrickson et al. 2018; Redondo-Salvo et al. 2020). Evidence for HGT-generated degeneracy has been observed in organisms ranging from endosymbionts and pathogens (Xie et al. 2003; de Felipe et al. 2005; Imanian and Keeling 2014; Ghosh and O'Connor 2017). To date, we are unaware of any other studies that examine the interplay between degeneracy and genetic context using combinations of characterized beneficial alleles that emerged during experimental evolution. Here, we examine how genetic context effects selection after HGT through expression of well-characterized, functionally redundant native and novel heterologous pathways within genetic backgrounds harboring alternative sets of alleles (with defined molecular and genetic interactions) that influence the functionality of the new pathway. This approach allows us to address 1) the factors that drive maintenance of a pathway acquired via HGT and 2) how strongly these factors are influenced by other alleles present in a given genetic context, particularly those known to benefit usage of degenerate versions of the pathway (fig. 1).

Metabolism provides a rich source of potential functional redundancy inherent to degeneracy (Edelman and Gally 2001; Sambamoorthy and Raman 2018; Sambamoorthy et al. 2019). Thus, in this study we use methylotrophy as a model system to examine how genetic context shapes the evolution of degenerate metabolic pathways (fig. 2). Methylotrophic organisms have the unique metabolic ability to utilize reduced one-carbon compounds as sole sources of carbon and energy. A direct consequence of methylotrophic metabolism is that all carbon flows through the toxic intermediate formaldehyde. Four described pathways for formaldehyde oxidation exist among bacteria (Chistoserdova 2011). The functional redundancy provided by degeneracy is a common feature among methylotrophs, as many methylotrophs harbor more than one of these pathways. It has been hypothesized that degeneracy within methylotrophic metabolism is selected for, as a means to safeguard against endogenous accumulation of toxic formaldehyde (Marx et al. 2004, 2005; Chistoserdova 2011). Nonetheless, the model methylotroph *Methylorubrum extorquens* AM1

(formerly *Methylobacterium*) encodes a single formaldehyde oxidation pathway that uses the cofactor dephosphotetrahydromethanopterin (referred to throughout as the “Native” or “dH₄MPT” pathway; fig. 2; Chistoserdova et al. 1998). This single-pathway system is well-poised for examining the acquisition of degenerate pathway(s).

During the process of determining the role of the Native dH₄MPT pathway in *M. extorquens*, it was found that this Native pathway could be replaced with an unrelated glutathione (GSH)-dependent bacterial formaldehyde oxidation pathway (Marx et al. 2003). This “Foreign” GSH-dependent pathway is composed of S-hydroxymethyl-glutathione dehydrogenase (Ras et al. 1995) and S-formyl-glutathione hydrolase (Harms et al. 1996), encoded by *flhA* and *fghA*, respectively (fig. 2). An engineered *M. extorquens* strain was constructed, where the dH₄MPT pathway was disrupted by a deletion of *mptG*, encoding β -ribofuranosylaminobenzene 5'-phosphate synthase (Marx et al. 2003), and the two enzymes of the novel (i.e., not having been seen previously by the cell) GSH-dependent formaldehyde oxidation pathway

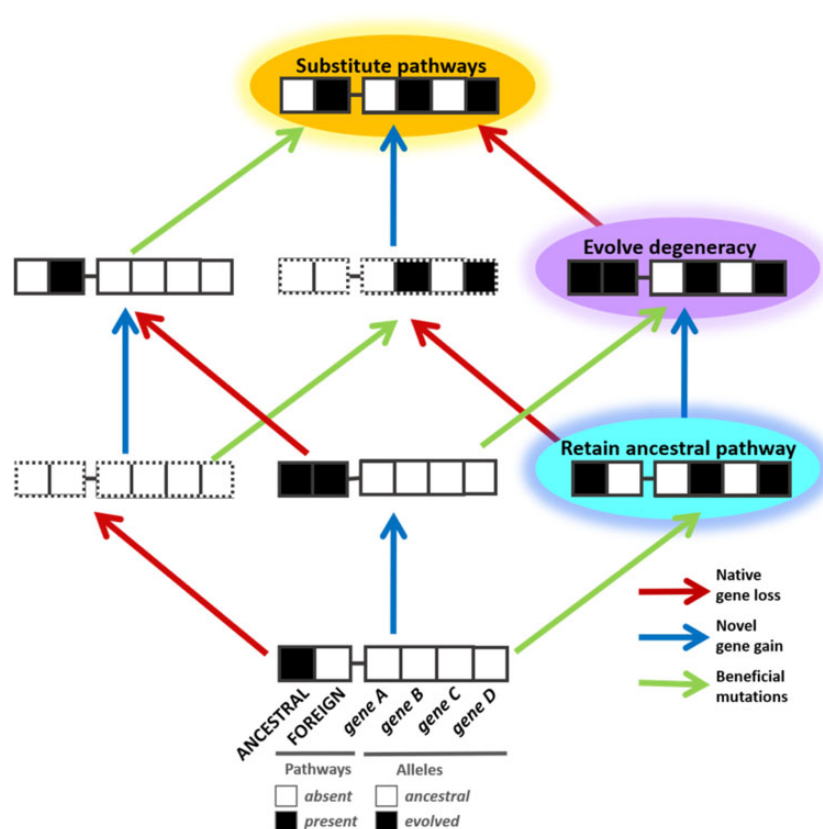


FIG. 1.—Combined impacts of metabolic gene gain, loss, and beneficial mutation on competitive fitness. For a given process, required and redundant genes can be gained and lost, and additional mutations that improve fitness can be acquired and selected. In this diagram, each string of boxes represents a single genotype possessing different combinations of pathways and alleles. The genotypes are separated from their nearest neighbors by single genetic changes such as mutation, gene gain, or gene loss. The left two boxes for each genotype depict the presence or absence of either of two degenerate pathways for a metabolic process of interest. The four boxes to the right for each genotype describe the allelic state of additional genes, which could encode products directly involved in one of the pathways under examination, or be separate and/or located elsewhere in the genome. Epistasis, both in terms of sign and magnitude, can emerge between ancestral, evolved, and newly acquired alleles. Presence or absence of pathways or alleles is indicated by filled or empty boxes, respectively. Boxes bordered by broken lines indicate genotypes that are nonviable under the selective condition examined, due to the absence of any metabolic pathway for the required metabolic process (e.g., methylotrophic metabolism).

from *Paracoccus denitrificans* (Ras et al. 1995) were introduced in trans. Although functionally analogous to the Native dH₄MPT pathway, in terms of converting formaldehyde to formate, the enzymes of the Foreign GSH pathway only use NAD⁺ as a cofactor, generating NADH. In contrast, the enzymes of the dH₄MPT pathway can use either NAD⁺ or NADP⁺, thereby generating NADH or NADPH, respectively (Chistoserdova et al. 1998). As this step is the only central metabolic reaction generating NADPH during growth on methanol, the exclusive use of the Foreign GSH pathway also requires involvement of pyridine nucleotide transhydrogenase (encoded by *pntAB*) to convert NADH (and NADP⁺) to NADPH (and NAD⁺; Carroll and Marx 2013). As PntAB couples hydride transfer from NADH to NADP⁺ with proton translocation from the periplasmic space into the cytosol, the Foreign pathway reduces the proton motive force, making the generation of biomass more energetically costly than for the dH₄MPT pathway (Jackson 2003).

The *M. extorquens* strain with the Foreign GSH pathway alone could grow on methanol, but because it grew approximately 3-fold worse than wild type (WT) (both in terms of rate and yield), experimental evolution was used to select for improved growth (Chou et al. 2011). The following beneficial mutations were identified in a fit sequenced isolate: A) "*fghA*^{evo}," an intergenic mutation alleviating the gene expression burden of both *fghA* and *flhA* (Chou and Marx 2012); B) "*pntAB*^{evo}," a promoter mutation increasing expression of the transhydrogenase and NADPH generation, increasing its availability for biosynthetic reactions (Carroll and Marx 2013); C) "*gshA*^{evo}" (encoding γ -glutamylcysteine synthetase), a promoter mutation predicted to increase cellular GSH pools; and D) mutations at six other loci without obvious direct connection to formaldehyde oxidation (which were treated as a single "allele" designated as "*GB*^{evo}" for "Genetic Background"; Chou et al. 2011). This collection of mutations that there existed multiple avenues to improving

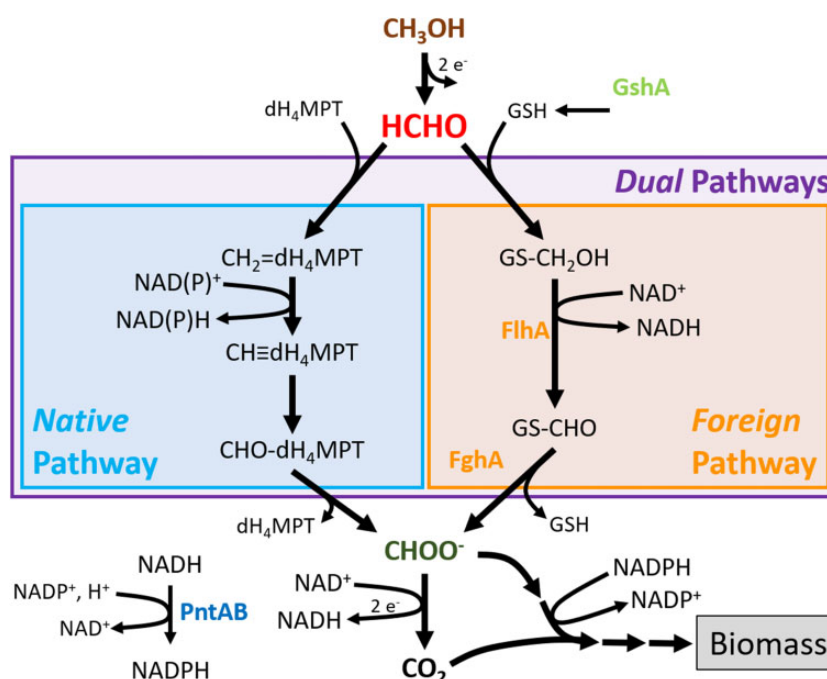


FIG. 2.—Formaldehyde oxidation pathways used in this study. The “Native” pathway for formaldehyde oxidation normally found in *Methylorubrum extorquens* AM1 is the dH₄MPT-dependent branch displayed. Alternatively, genes encoding the “Foreign” degenerate GSH-dependent pathway from *Paracoccus denitrificans* were cloned onto plasmid pCM410 and moved into *M. extorquens*. “Dual” pathway strains have degeneracy, as they possess both of these unrelated pathways. The alternative pathways conjugate the formaldehyde intermediate to dH₄MPT or GSH cofactors, respectively, and oxidize the single carbon through a series of reactions to formate. Formate represents the important metabolic branchpoint between respiration and incorporation into biomass. The Foreign pathway utilizes only the NAD⁺ cofactor, whereas the Native pathway can either NAD⁺ or NADP⁺. Methanol, formaldehyde, and formate are colored with brown, red, and dark green text, respectively, for emphasis. Mutations impacting the products of *fghA/flhA*, *gshA*, and *pntAB* genes are similarly emphasized with colored text.

usage of this pathway in *M. extorquens*. Beneficial mutations affecting *fghA/flhA*, *pntAB*, and *gshA* were identified across multiple populations (Chou and Marx 2012; Carroll and Marx 2013), demonstrating that they were repeated targets for improving the ability of cells to use the *Foreign* pathway. The order of their conferred fitness gains is: *gshA*^{evo} > *fghA*^{evo} > *GB*^{evo} > *pntAB*^{evo} (Chubiz et al. 2012).

Perhaps, the most important consideration in the evolution of degeneracy is whether maintaining the novel function is beneficial, neutral, or deleterious. If the second pathway is immediately an advantage, the emergence of degeneracy is straightforward. In contrast, if the novel pathway is initially neutral or deleterious in the presence of the endogenous pathway, then loss would be expected unless adaptation toward degeneracy can rapidly occur. The *Foreign* GSH pathway was shown to be costly in WT *M. extorquens* that has the *Native* dH₄MPT pathway, due to both plasmid maintenance and enzyme expression costs (estimated previously as fitness costs of 5.5% for *fghA* and 2.7% for *flhA*; Chou and Marx 2012; Chou et al. 2014), but it remains unclear whether these costs might be reduced or even eliminated in the presence of other evolved alleles (such as those listed in the previous paragraph) that

improve pathway function and/or lower plasmid maintenance requirements.

It is particularly of interest whether mutations that evolved to optimize the use of a particular pathway will favor or disfavor the use of alternative or additional degenerate pathways. With a number of known mutations that optimize utilization of the *Foreign* pathway in hand, we could ask whether the same mutations that render a novel-pathway-encoding background more fit also promote degeneracy by facilitating the maintenance of the *Foreign* GSH pathway in a background also encoding the *Native* dH₄MPT pathway. In this study, rather than directly evolve strains with existing degeneracy, we reconstructed strains containing mutations known to improve utilization of the *Foreign* pathway in isolation, placing them in backgrounds containing either the *Native* pathway or both (*Dual*) pathways in combination. This approach allowed us to ask whether preexisting alleles improving one pathway would help or hinder the use of a degenerate pathway for the same metabolic process. We observed that epistatic interactions among both the native alleles and the plasmid-encoded *fghA* and *flhA* genes. Namely, mutations had an altered combinatorial effect on cell fitness (or other traits) than expected from the sum of their effects in

isolation (Moore and Williams 2005; Phillips 2008), and these effects played an important role in the maintenance or loss of the degenerate pathways examined. Importantly, epistasis can be observed as differences in either the strength (magnitude epistasis) and/or directionality (sign epistasis) of the fitness coefficients observed for the strain harboring the pertinent mutations, relative to fitness outcomes for the respective single mutations (Weinreich et al. 2005; Blount et al. 2008, 2012; Chou et al. 2011; Khan et al. 2011; Rokyta et al. 2011; Tokuriki et al. 2012; Kryazhimskiy et al. 2014; Schoustra et al. 2016; Douglas et al. 2017; Wünsche et al. 2017). Interestingly, although none of the alleles tested exhibited sign epistasis in the *Foreign* GSH pathway context where they initially evolved (beneficial both individually and in combination), the same alleles often exhibited sign epistasis in the *Native* and *Dual* pathway contexts. The deleterious epistatic effects observed, which may reflect energetic burdens stemming from expression of these alleles in the genetic backgrounds encoding the dH₄MPT pathway, introduce inaccessible evolutionary trajectories that would constrain the evolutionary routes and endpoints among combinations of these alleles (Chou et al. 2011; Salverda et al. 2011; Chubiz et al. 2012). However, compared with their respective ancestors, several *Dual* strains with alleles that arose during evolution of the *Foreign* pathway alone displayed decreased plasmid loss, and occasionally sustained retention. The increased fitness and corresponding retention arose from synergistic epistasis between a subset of the tested alleles in the *Dual* background. Together, this work highlights the important role of epistatic interactions between pathways and other genomic alleles on retention of newly acquired metabolic pathways, either as sole novel or degenerate functions.

Results

Loss of Degenerate GSH Pathway Was Rapid and Not Dependent upon Mode of Metabolism

In this study, we sought to address which factors lead to the (in)stability of metabolic degeneracy, through the lens of an alternative formaldehyde oxidation pathway introduced into the model methylotroph *M. extorquens*. A formaldehyde oxidation pathway is required for *M. extorquens* to grow on methanol as a carbon source (Chistoserdova 2011). As previously mentioned, substituting the *Native* pathway with the *Foreign* GSH pathway allowed suboptimal methanol growth (Chou et al. 2011). Although the plasmid-encoded GSH pathway is essential for methylotrophic growth when it is the only formaldehyde oxidation pathway present, there was still the possibility of degenerate pathway coevolution. Thus, we hypothesized that the *Foreign* pathway would be less beneficial in the presence of the *Native* pathway under relaxed selection (i.e., *Dual* pathway strains grown in methanol media without kanamycin, used for plasmid

maintenance). To address whether the *Foreign* pathway would be incorporated into the metabolic network of recipient *M. extorquens* genotypes encoding the *Native* dH₄MPT pathway, we measured how well WT maintained the pCM410 plasmid encoding genes for the GSH pathway (genotype 11-0000; fig. 2 and supplementary fig. S1, [Supplementary Material](#) online). Over multiple passages in methylotrophic (methanol)- or nonmethylotrophic (succinate)-based media, we regularly monitored plasmid retention by quantifying viable cell counts on comparable solid media with and without kanamycin (fig. 3A). Through this screen, we observed rapid and uniform loss of the pCM410 plasmid from this basal *Dual* strain, independent of the carbon source included. These results contradicted our expectation that the pathway would be more expendable on the nonmethylotrophic carbon source (succinate). This suggested that even when formaldehyde oxidation was required for growth (methanol), it was beneficial for WT to lose the degenerate pathway, indicating that costs, such as plasmid maintenance and enzyme expression, outweigh any benefits introduction of the *Foreign* GSH pathway confers for dealing with cellular formaldehyde.

Beneficial Alleles Lead to Differential Retention of GSH Pathway

To address whether genetic context could increase pCM410 retention in the presence of the *Native* dH₄MPT pathway, plasmid retention was assessed for several *Dual* pathway genotypes (*gshA*^{evo}, *fghA*^{evo}, *pntAB*^{evo}, and *GB*^{evo}). Each allele was introduced individually into the *Dual* pathway background and the retention frequency of pCM410 was monitored during passage under methylotrophic growth conditions (methanol) without antibiotic selection (fig. 3B). In most cases, strains with evolved alleles showed increased plasmid maintenance compared with the 11-0000 strain (*gshA*^{evo} > *pntAB*^{evo} > *fghA*^{evo} > Ancestor, $P = 4.51\text{e-}06$; fig. 3B, supplementary tables S6 and S9, [Supplementary Material](#) online). This stronger retention of the degenerate pathway suggested that the mutations able to improve growth in the *Foreign* pathway background and linked to increased formaldehyde oxidation by that pathway could also confer benefits (or reduce costs) in the *Dual* pathway context. However, all genotypes tested here demonstrated significant and sustained plasmid loss over a relatively small number of generations. The *GB*^{evo} single mutant strain (11-0001), which harbored mutations that were not directly connected to formaldehyde oxidation by GSH in clearly characterized ways, saw a similar rate of plasmid loss to the 11-0000 strain, suggesting that the role of the mutations composing the *GB*^{evo} allele might enable improved methylotrophic growth when only the *Foreign* GSH pathway is present, but do not seem to influence maintaining the GSH pathway when both GSH and dH₄MPT pathways are available. Further, the

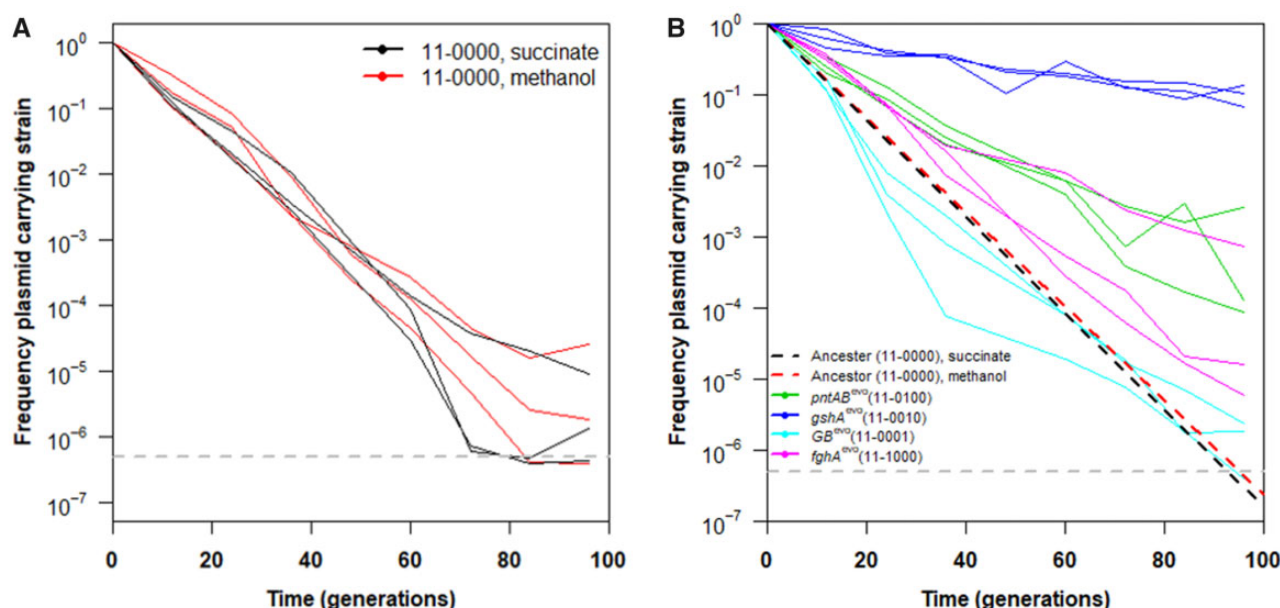


FIG. 3.—(A) Plasmid carriage over time. Maintenance of plasmid pCM410 over time was examined when introduced into WT *Methylobacterium extorquens* AM1 (resulting in strain CM3260, genotype “11-0000”). The experiment was carried out in Hypho media with relaxed selection for the plasmid (no antibiotic present). Regardless of the specific carbon source used, we observed rapid and constant loss of the plasmid over time, suggesting strong fitness costs for its carriage by this genotype. (B) Each strain examined has a separate evolved mutation(s) in its background and was carried out in triplicate. Lines indicate the change of individual lineages and dots indicate the mean of the triplicate cultures for a given strain. Model fits (linear regression) for the rate of loss in the background strain CM3260 grown in both succinate (black) and methanol (red) media are also provided as dashed line for reference.

maintenance of pCM410 was improved by mutations that increase formaldehyde oxidation through that pathway, suggesting that the degenerate pathway could provide fitness benefits beyond harboring the *Native* pathway alone only when flux through the GSH pathway was further genetically optimized. In all, these results raised important questions concerning how maintenance of degenerate pathways might be favored by evolution in cases where the genetic context is arranged optimally for utilization of a newly acquired pathway, whether by expression of that pathway (*fghA*^{evo}), abundance of critical substrates (*gshA*^{evo}), energy contributions of the pathway (*pntAB*^{evo}), or some other factor. In particular, it suggests that shifting GSH abundance via upregulation of *gshA* (*gshA*^{evo} allele) serves as a critical bottleneck to effective usage of the *Foreign* pathway.

Epistatic Interactions among Evolved Alleles Impact the Cost of the Plasmid Encoding the *Foreign* Pathway

As plasmid loss rates differed across the tested genetic backgrounds, we hypothesized that differences in relative fitness between alleles were driving the plasmid retention discrepancies observed. To more directly test the impact of epistatic interactions and resulting fitness of strains on the cost of carrying the degenerate *Foreign* GSH pathway in the *Dual* background, we introduced all eight viable combinations of evolved alleles into the *Native* strain, along with either the

fghA^{anc} or *fghA*^{evo} version of the *Foreign* pathway, and measured the relative fitness values of the resulting *Dual* and *Native* strains against WT (“10-x000,” [supplementary table S6, Supplementary Material](#) online). Comparing fitness values and deriving selection coefficients, we found that the loss of the *Foreign* pathway was nearly universally beneficial in the *Dual* pathway strains, even when the cost of plasmid carriage was decreased by the *fghA*^{evo} allele, which reduces expression of the *fghA* and *flhA* genes from pCM410 (fig. 4; Chou et al. 2011; Chou and Marx 2012). The total carriage cost was generally lower if the plasmid contained the evolved *fghA*^{evo} allele (mean *S* of -0.048 vs. -0.1 , $P = 0.023$), but there were substantial differences across allele backgrounds, with several showing more than a 50% reduction in net plasmid retention costs (fig. 4, [supplementary tables S6 and S9, Supplementary Material](#) online). Interestingly, when the *gshA*^{evo} and *fghA*^{evo} were both present (11-1010), a positive selection coefficient was observed when compared with the corresponding *Native* background strain containing only the *gshA*^{evo} allele (10-x010, fig. 4). This indicated that inclusion of the plasmid encoding the GSH-dependent pathway was beneficial in a background where the *gshA*^{evo} allele presumably increased GSH pools, despite the energetic burden associated with harboring the expression vector for the GSH pathway. This exception aside, the general tendency suggests that the *Native* genotypes are typically more

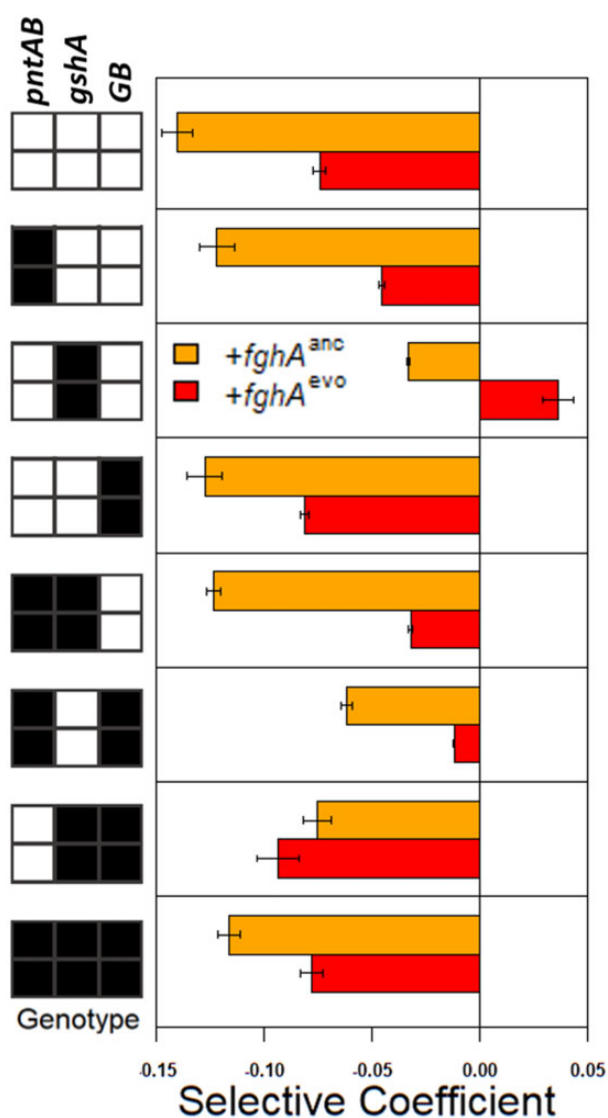


Fig. 4.—Comparison of selection coefficients for inheriting the GSH pathway plasmid by *Native* pathway strains. The bars are color-coded by which focal *fghA* allele is being introduced: *fghA*^{anc} is in orange and *fghA*^{evo} is in red. The specific genotypic background that the mutation is introduced into is indicated in the key on the left. Ancestral alleles (“0”) are coded in white, evolved alleles (“1”) are coded in black. Error bars represent 95% confidence interval estimates.

evolutionarily stable than their *Dual* equivalents and plasmid loss is favorable in most *Dual* backgrounds. Additionally, the selection coefficient values from single allele replacements (using both *fghA*^{anc} and *fghA*^{evo} backgrounds) tracked well with the plasmid retention values, where *gshA*^{evo} > *pntAB*^{evo} > *GB*^{evo} (fig. 4). Overall, these results are consistent with the hypothesis that differences in fitness between *Dual* strains and their *Native* counterparts contributed to the patterns of plasmid loss observed previously ($R^2 = 0.76$, $P = 0.0238$; supplementary table S10, [Supplementary Material](#) online).

Many Altered Epistatic Interactions Are Observed between Alleles Evolved with the *Foreign* Pathway and the dH₄MPT Pathway

Beyond examining genetic interactions impacting retention of the *Foreign* pathway plasmid, we sought to further characterize the epistatic interactions among evolved alleles by testing whether evolved alleles that were beneficial in the *Foreign* pathway context also provided benefits within either the *Native* or *Dual* pathway backgrounds. To comprehensively evaluate patterns of epistasis, fitness values were compared within the *Foreign*, *Native*, and *Dual* pathway backgrounds for strains that represent all combinations of the four characterized alleles viable for methylotrophic growth (40 genotypes in total across all backgrounds; fig. 5, supplementary tables S1 and S3, [Supplementary Material](#) online). To first note, these data confirm that the transition to *Dual* degeneracy is strongly beneficial coming from the *Foreign* background and costly coming from the *Native* background (supplementary fig. S4A, [Supplementary Material](#) online). Additionally, in contrast to the universally beneficial fitness effects observed for evolved alleles in the *Foreign* pathway context, examining the fitness effects of the focal alleles individually in the *Native* and *Dual* pathway backgrounds gave mixed fitness results, indicating

many context-specific epistatic interactions (figs. 5 and 6, and supplementary fig. S4B and supplementary table S5, [Supplementary Material](#) online). Only three of the individual allele replacements showed increased fitness values over the basal *Native* or *Dual* contexts: *gshA*^{evo} into either *Native* or *Dual* backgrounds and *fghA*^{evo} into the *Dual* background. Alternatively, *GB*^{evo} was individually neutral (or slightly beneficial in the transition to 11-0001) in both the *Native* and *Dual* contexts. Finally, although *pntAB*^{evo} enhanced the utilization of the *Foreign* GSH as the sole pathway, *pntAB*^{evo} led to fitness losses in backgrounds (*Native* or *Dual*) able to generate NADPH during methylotrophy by way of the *Native* dH₄MPT pathway. Fitness values calculated from pairwise assays positively correlated well with exponential growth rates when strains grew individually, validating these fitness measures (supplementary fig. S6, [Supplementary Material](#) online; $R^2 = 0.8928$, $P < 0.0001$).

The discrepancies observed in fitness outcomes following single allele replacements within the *Foreign*, *Native*, and *Dual* backgrounds prompted a further investigation into how the tested alleles might also interact with each other. A small number of consistent allelic patterns emerged (figs. 5 and 6). First, the *fghA*^{evo} allele remained beneficial in six of the seven transitions in the *Dual* pathway background. Second, all transitions to the *GB*^{evo} allele in the presence of *pntAB*^{evo} and/or *gshA*^{evo} were beneficial, providing a strong example of synergistic epistasis. This was surprising given that only *gshA*^{evo}, in isolation, provided notable beneficial effects in both *Native* and *Dual*

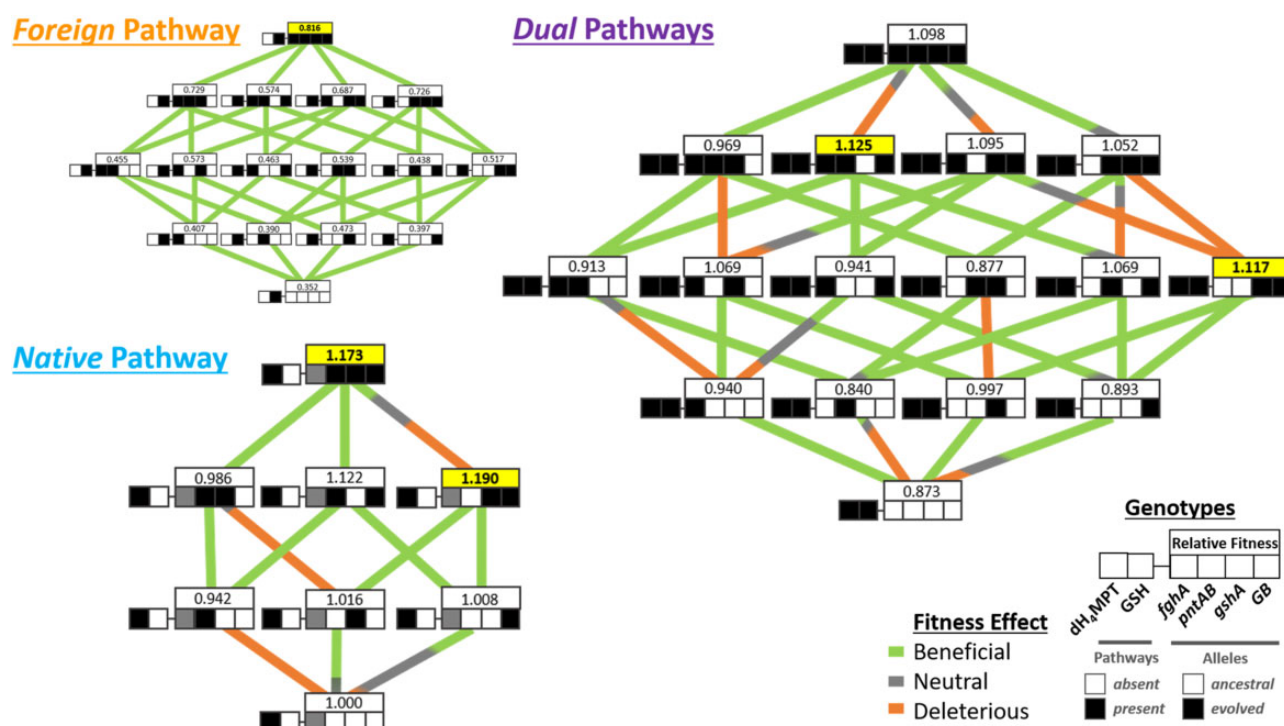


Fig. 5.—Fitness landscape for genotypes of all pathway backgrounds. (A) All *Foreign* strains were competed pairwise with the reference strain CM1232 (genotype “01-0000”). (B) *Dual* and (C) *Native* strains were competed pairwise with the reference strain CM1176 (genotype “10-x000”). For more information on strains, see [supplementary table S1, Supplementary Material](#) online. Three major subsections of the landscape exist: (A) the *Foreign* pathway strains that contain only the GSH pathway, (B) the *Dual* pathway strains that contain both the *Native* pathway and the *Foreign* GSH pathway, and (C) *Native* pathway strains that contain only the dH₄MPT pathway. In the diagram, each row of boxes depicts a single strain genotype, with the genotype of the focal alleles represented by the coloration pattern of the boxes. The leftmost box represents the pathway background of a given genotype and the right four boxes depict the four focal alleles (from left to right: *fghA*, *pntAB*, *gshA*, and *GB*) in either their ancestral (white) or evolved (black) states. The *fghA* box for the *Native* pathway strains is grayed out because the allele is absent in these genotypes. The lines connecting genotypes represent the transitions between mutational neighbors and are color-coded by their net effect based upon probabilities from empirical fitness distributions: beneficial in green, deleterious in orange, and neutral in gray (also see [supplementary fig. S7, Supplementary Material](#) online, for more detailed descriptions).

backgrounds (fig. 6). Third, *gshA*^{evo} was predominantly beneficial or neutral when combined with other alleles, but the pattern of interactions differed slightly between the *Dual* and *Native* contexts. Finally, the *pntAB*^{evo} allele exhibited abundant sign epistasis, where its introduction into various *Native* or *Dual* backgrounds overwhelmingly led to fitness losses (figs. 5 and 6). Interestingly, *pntAB*^{evo} remained exclusively deleterious (4/6) or neutral (2/6) in the backgrounds containing the *gshA*^{evo} allele, and contrastingly beneficial (3/6) or deleterious (3/6) in backgrounds with the *gshA*^{anc} allele. Where *GB*^{evo} was also present in the *gshA*^{anc} backgrounds, the *pntAB*^{evo} introduction increased fitness. These data suggest that the perturbation to nucleotide cofactor pools by introduction of *pntAB*^{evo} is physiologically disadvantageous in backgrounds harboring the dH₄MPT pathway, except in the cases where *GB*^{evo} is already present. It is unclear how *GB*^{evo} might alter the metabolic state of the cell, such that *pntAB*^{evo} introduction into the *Native* or *Dual* backgrounds becomes advantageous. However, regardless of *GB* allelic state, introduction

of *pntAB*^{evo} never allows a fitness advantage in backgrounds where *gshA*^{evo} is encoded, showing that changing both GSH abundance and nucleotide cofactor pools in a background harboring the dH₄MPT pathway is physiologically disadvantageous. On the whole, the set of fitness effects observed in the *Native* and *Dual* backgrounds largely resembled one another and were very divergent (namely, not universally beneficial) from those observed in the *Foreign* background, suggesting that the dH₄MPT pathway is exerting a different and dominant effect upon cellular physiology in comparison with the GSH pathway (fig. 6).

Epistatic Interactions Alter the Observed Fitness Maxima and Accessible Adaptive Trajectories for Evolved Alleles in Altered Pathway Backgrounds

Given that moving evolved alleles that were beneficial in the *Foreign* context into the *Native* or *Dual* contexts exhibited distinct individual and joint epistatic effects, we sought to

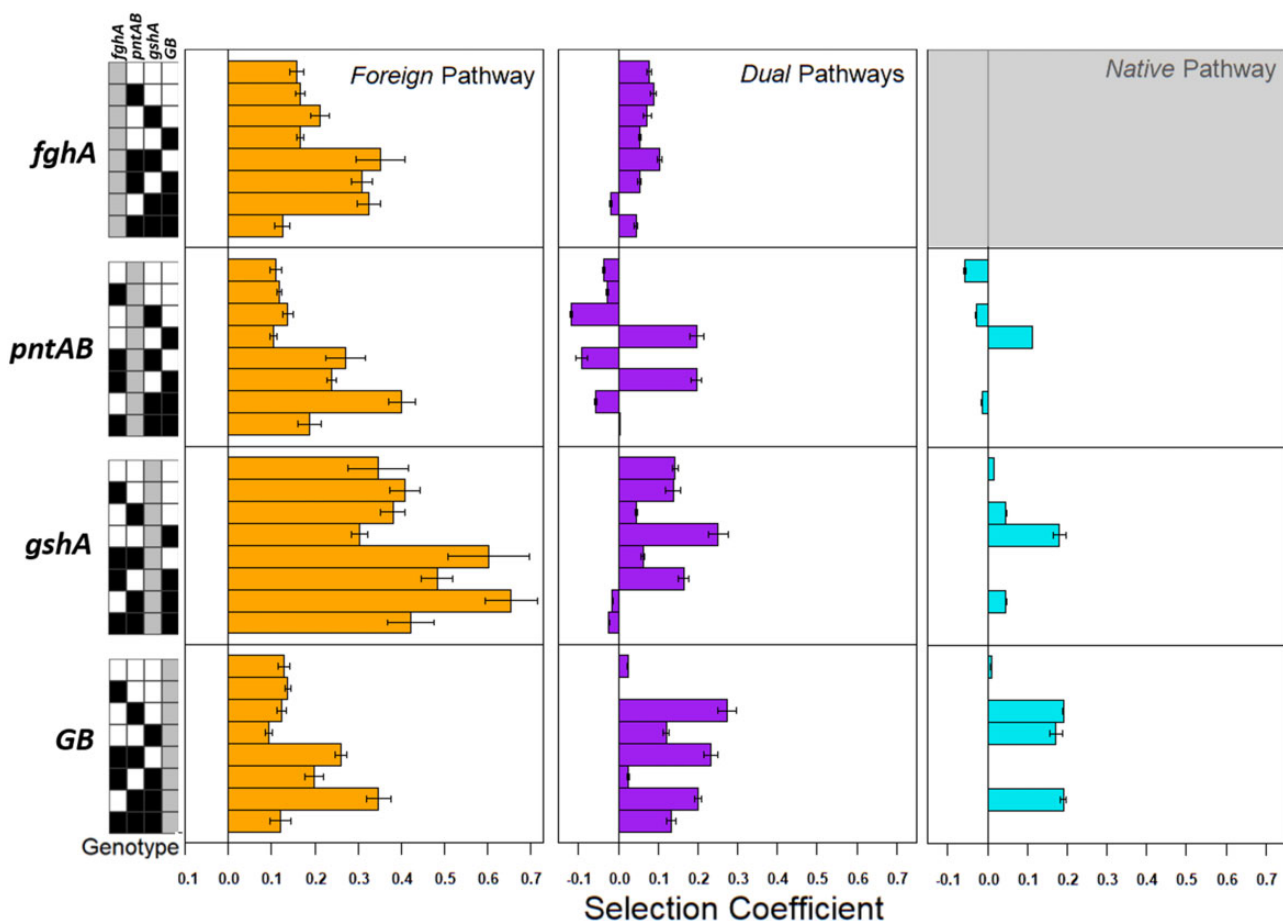


Fig. 6.—Comparison of the selection coefficients of mutational changes in different pathway backgrounds. For each panel, bars to the right of the line are beneficial in effect, whereas those to the left are deleterious. The combinatorial genotype the focal allele is introduced into is depicted by the key on the left—ancestral alleles (“0”) are coded in white, evolved alleles (“1”) are coded in black, and the focal allele being introduced are coded in gray. The *Native* pathway *fghA* section is grayed out because these are inviable genotypes. The bars color intensity is weighted by the likelihood of draw, and those with intermediate draw probabilities (95% confidence intervals overlapping neutral) are lightly shaded. Error bars represent 95% confidence interval estimates.

characterize how these epistatic interactions would alter the evolutionary trajectories for strains encoding these specific mutations. Specifically, we were interested in comparing the likelihood of different orders of mutations occurring in the different pathway backgrounds, given the observed fitness data for all constituent strains. To incorporate measurement uncertainty in the analysis, random draws were simulated from empirical fitness distributions for all tested combinatorial genotypes to assess the likelihood that all mutational changes in a given trajectory were jointly beneficial. In the *Foreign* landscape, all possible trajectories were accessible (i.e., adaptive along the entirety of a series of transitions) as all mutations were universally beneficial (fig. 5A and supplementary fig. S4B and supplementary table S7, [Supplementary Material](#) online). However, not all adaptive trajectories were accessible in the *Native* and *Dual* pathway landscapes (fig. 5 and supplementary table S7, [Supplementary Material](#) online). It was much rarer that either the 10-x111 or 11-1111 genotypes

were adaptively accessible: None of the six trajectories in the *Native* landscape and none of the 24 trajectories available on the *Dual* landscape were accessible in a majority of draws, and most were accessible much more rarely (supplementary table S7, [Supplementary Material](#) online). This stark difference in accessibility between landscapes indicates the existence of alternative local maxima in these landscapes and suggests that alternative mutations would be selected for if evolved under the same conditions. In the *Dual* pathway background, two genotypes (11-0011 and 11-1101; fig. 5, yellow boxes) possessed higher fitness values than the quadruple mutant strain 11-1111, and the overall fitness maxima among this set was occupied by genotype 11-1101. This was intriguing because the genotypic combination producing a peak in the *Dual* landscape differs from the genotype producing a fitness peak on the *Native* landscape (10-x011), suggesting that subtle differences in physiological conditions either favoring or disfavoring these combinations exist among these genotypes.

Particularly, it is likely that excluding the *pntAB*^{evo} allele is more beneficial in the *Native* context where there is no GSH pathway present for it to elicit its originally evolved effect when more generation and cycling of NADH is beneficial. It is also interesting that 11-1101 corresponds to the *Dual* landscape maxima, because it contains all evolved alleles except for the *gshA*^{evo}, which individually provided the largest beneficial fitness effect across all backgrounds. This demonstrated that the upregulation of GSH production is not strictly required for fitness improvement in the *Dual* pathway background.

Though genotypes with all focal evolved mutations were not evolutionarily accessible in the *Native* and *Dual* landscapes, it was possible that alternative accessible maxima existed among the genotypic combinations examined. To completely evaluate all paths, we also considered partial paths, that is, those that did not transition all the way from zero to all of the evolved alleles included in the study. When trajectories which included transitions of only three or four genotypes (rather than five included in complete paths from 0 to four evolved alleles, i.e., 11-0000 to 11-1111) were assessed, 12 of 42 possible trajectories were accessible with over 50% certainty, and three were accessible with 99% certainty (fig. 7 and supplementary table S8, [Supplementary Material](#) online). Comparatively, in the *Foreign* landscape, all 24 of the possible trajectories to 01-1111 were universally beneficial with >99.9% certainty (supplementary table S7, [Supplementary Material](#) online). Conversely, in the *Native* (5/6 full paths accessible in <50% of trials, 3/6 paths accessible in <1% of trials) and *Dual* (24/24 full paths accessible in <50% of trials and 10/24 paths accessible in <1% of trials) landscapes, mutations being universally beneficial were much less likely, regardless of the order acquired ($P < 2e-16$, supplementary table S7, [Supplementary Material](#) online). Thus, although selection appeared to permit any combinatorial adaptive trajectory available to the peak of the *Foreign* landscape, limited paths and alternative peaks exist for these alternative backgrounds due to neutral and deleterious fitness effects that led to inaccessible paths. When we consider both paths within and between *Native* and *Dual* landscapes, we see that many are constrained, and that many of these trajectories would not be favored in selection experiments using strains that have these pathway backgrounds (fig. 7).

Genotype Primed for Degeneracy Reveals Rare Maintenance and Coevolution of *Dual* Pathways

Both direct measurements of plasmid maintenance as well as competition fitness data suggest that the degenerate pathway will simply be lost rapidly if acquired in trans in most *Dual* genotypic backgrounds if antibiotic selection was removed (figs. 4 and 5, supplementary tables S6 and S9, [Supplementary Material](#) online). However, we predicted that some genotypes would potentially be able to maintain

the degenerate GSH pathway under relaxed selection conditions and coevolve with it. To test our prediction and give the adaptation of degeneracy ample opportunity to occur, we propagated 20 populations of the 11-1010 genotype, which was the one example we observed where the introduction of the plasmid with the *Foreign* pathway led to increased fitness (fig. 8). Initially, these populations also experienced rapid loss of the plasmid; however, we saw reversal of plasmid loss around 50 generations into the experiment for one replicate population (fig. 8). At this point, the frequency of the plasmid-carrying strain increased by an order of magnitude and then stabilized at an intermediate frequency (~5%) in the population. The reduced rate of plasmid loss in *gshA*^{evo}-containing lineages (figs. 3B and 8) suggests that they have the capacity to maintain the degenerate pathway long enough for adaptive mutations to arise that maintain the associated degeneracy/novel functions, whereas this is more doubtful in the genotypic backgrounds with more accelerated plasmid loss rates (Silva et al. 2011).

Discussion

Biological organisms are rife with functional redundancy among their traits, and degeneracy is one avenue by which this redundancy can be achieved. We approached the study of degeneracy through the lens of metabolic processes, where it is particularly common. This work sought to address an open question: How is the evolution of a novel degenerate pathway for a given process, often naturally acquired through HGT, impacted in different genetic contexts? In particular, what are the challenges associated with evolving to acquire and retain degenerate pathways that provide analogous functions to the cell? We investigate these questions by substituting or combining alternative formaldehyde oxidation pathways in *M. extorquens*, which, more specifically than generic genetic background changes across the genome, poises cells to exhibit distinct discrete physiological states between different pathway background states. Based on fitness measurements, it appears that combining the *Native* dH₄MPT (WT genome) and *Foreign* GSH pathway (plasmid) would not be favored. Indeed, when the original *Dual* strain was propagated, this background experienced rapid plasmid loss (fig. 3A). Although evolution of the strain engineered with the *Foreign* pathway alone rapidly led to improved growth, it was unclear whether those specific genetic changes would improve or further limit opportunities for coexistence of the two pathways in an evolutionarily accessible way. In alternative (*Native* and *Dual*) pathway backgrounds, those characterized mutations might potentially ameliorate the costs associated with maintaining the *Foreign* pathway (for *Dual* pathway genotypes) or, via tradeoffs, lead to antagonism with the original *Native* pathway. In the latter scenario, each mutation that was beneficial in the context of the *Foreign* pathway alone would be costly when coupled with the

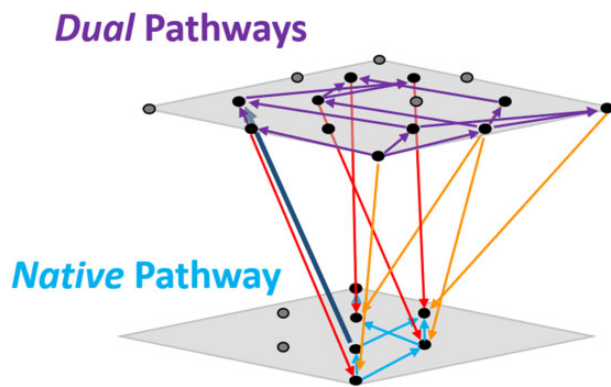


FIG. 7.—Combined conceptual fitness landscape schematic depicting accessible paths within and between (by gene acquisition or loss) *Native* and *Dual* landscapes. Each plane represents one of the sublandscapes, and the dots upon the plane represent the constituent genotypes. Because there are fewer *Native* genotypes viable for methylotrophic growth, there are correspondingly fewer dots present. The dots at the lower corner of each plane correspond to the basal genotypes in the landscape that has no additional evolved focal alleles. Arrows indicate the direction favored by selection. Purple arrows indicate selectively favored transitions among the *Dual* landscape, and cyan arrows indicate selectively favored transitions on the *Native* landscape. Orange and red arrows indicate selection favoring loss of degenerate pathway plasmid (bearing the *fghA*^{anc} and *fghA*^{evo} alleles, respectively), whereas the dark blue arrow indicates selection favoring plasmid retention.

Native pathway. Critical to understanding these adaptive possibilities was to map out the broader picture of epistatic interactions, combining evolved alleles in different metabolic contexts to understand their combined interaction network (figs. 1 and 5). Considering these interactions together gives a broader view of the fitness landscape that illuminates the likelihood of gain or loss of the degenerate pathways under examination.

Data from the plasmid retention experiments provided evidence that the degenerate pathway did not typically provide net benefits beyond harboring the *Native* pathway alone, at least not sufficient to maintain pCM410 at high frequencies, and instead led to substantial fitness costs (Chou et al. 2014). Taken all together, it became clear that the overall tendency for the *Foreign* pathway to be deleterious in the *Dual* pathway context suggests that, in many cases, degeneracy would simply resolve to having just the *Native* pathway (fig. 7). This provides the high-level picture that, on the whole, degeneracy is not greatly favored between these two pathways, at least not with the other alleles examined in this study. Nonetheless, the magnitude of the cost for the *Foreign* pathway was quite flexible, with nearly all allele combinations exhibiting a lower cost than the ancestral plasmid in the basal *Dual* strain (fig. 4). It was unsurprising that the *fghA*^{evo} allele, which has been shown to reduce protein expression costs for both FghA (by 73%) and FlhA (by 55%) (Chou et al. 2011, 2014; Chou and Marx 2012), generally reduced the cost of the degenerate

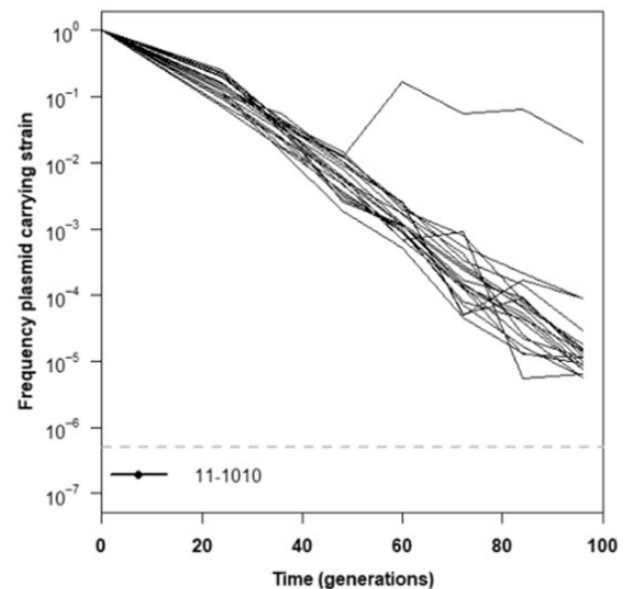


FIG. 8.—Plasmid carriage by genotype 11-1010 over time. Examination of plasmid pCM410 retention/loss over time was examined in strains of *Methylobacterium extorquens* AM1 containing evolved mutations identified in a previous study (Chou et al. 2011). Plasmid carriage by this strain was examined in biological replicate of $N = 20$. Lines indicate the change of each individual lineage. Experiments were carried out in Hypho media with relaxed selection for the plasmid (no antibiotic present). The dashed gray line in each panel indicates the experimental limit of detection for the assay.

Foreign pathway considerably. Beyond this, the precise combinations of alleles in the *Dual* pathway context significantly influenced the fitness outcomes following *Foreign* pathway introduction. Direct tests of plasmid maintenance similarly confirmed that rates of loss for this pathway were strongly dependent on pathway costs, which vary as a direct result of epistatic interactions between the pathways and other evolved alleles present in the genome (figs. 3 and 8, supplementary table S10, [Supplementary Material](#) online). In one case, the evolved plasmid (e.g., *fghA*^{evo}) in a genome with *gshA*^{evo} was even found to impart a benefit (to strain 11-1010; fig. 4). From a total of 20 starting replicates, strain 11-1010 was experimentally evolved to determine whether we could capture the incorporation of the degenerate pathway in real time. Here, a single population exhibited an increased and maintained rate of plasmid carriage through approximately 100 generations, showing that evolution could act to maintain both pathways (figs. 1 and 8; Harrison and Brockhurst 2012). Whereas both the rate of plasmid loss across cells (e.g., lack of segregation) and selection against plasmid-containing cells (Chou et al. 2014; Wein et al. 2019) will contribute to the exponential decline in their frequency in a population, a rise in frequency requires either plasmid exchange, which is absent in this system, or a selective benefit large enough for that plasmid-containing lineage to have a

fitness higher than the average fitness in the population at that time (Lang et al. 2011; Lee and Marx 2013). Given the known prevalence of clonal interference in populations with the *Foreign* pathway (Lee and Marx 2013), at least when it is the sole formaldehyde oxidation pathway in the cell, these dynamics are consistent with that particular lineage emerging through the evolution of beneficial alleles to compete with other high-fitness sublineages also present in the population. Further work will be required to isolate this lineage and determine the genetic basis of the first steps toward adapting toward degeneracy in the *Dual* context. Yet, the majority of populations from this ancestor ultimately lost the plasmid encoding the GSH pathway, speaking to the rarity and difficulty of overcoming the costs imposed by the degenerate pathway encoded on the pCM410 plasmid (fig. 3B).

Whereas the fitness effects of the evolved alleles tested were universally beneficial across genotypes in the original *Foreign* context, we observed rampant pathway-specific epistatic effects when they were placed in the *Dual* and *Native* contexts. None of the evolved alleles had universal selective effects across genotypes tested (figs. 5 and 6). This result contradicted the simple hypothesis that alleles would be universally beneficial in the *Dual* and *Native* contexts if they improve growth on methanol (or other environmental factors in the selective regime) in a manner independent of the pathway(s) allowing methylotrophic growth. Alternatively, they could be universally neutral if they aided the *Foreign* strains in using the GSH pathway in a manner irrelevant to functioning of the *Native* pathway, or universally deleterious if they compromised the function of the *Native* pathway in order to improve use of the *Foreign* pathway (i.e., via tradeoffs between the degenerate pathways). We observed that pathway-specific interaction patterns were prevalent and substantial. Each of the three predominant alleles (*fghA*^{evo}, *pntAB*^{evo}, and *gshA*^{evo}) had their qualitative effect (i.e., beneficial vs. neutral vs. deleterious) altered by the presence of the other alleles, and these patterns tended to be quite consistent between the *Dual* and *Native* pathway contexts. These altered patterns suggest that the presence of the *Native* dH₄MPT shifts the cellular physiology significantly in ways that frequently render these mutations nonbeneficial; this occurs perhaps because although these mutations move metabolite pools closer to that of WT when only the *Foreign* GSH pathway is present, they may disrupt the same pools in strains that already possess the “WT pathway” (fig. 6). Generally speaking, though introducing alleles into different backgrounds than those they evolved in decreases both the likelihood and magnitude of their being beneficial, the occurrence of synergistic epistasis to ameliorate the costs of degeneracy may allow coevolution of multiple functionally redundant pathways in parallel.

The wide range of epistatic effects observed in the *Dual* and *Native* contexts often ran counter to expectations based upon the known physiological connections between them.

The *Foreign* GSH pathway depends upon a distinct C₁ carrier, GSH, rather than dH₄MPT, and requires involvement of transhydrogenase (encoded by *pntAB*) to generate NADPH and recycle NADH to NAD⁺. Correspondingly, it has been observed that the ratio of NADH and NADPH is disturbed in the *Foreign* strain compared with the *Native* WT (Carroll and Marx 2013). Furthermore, GSH is active as a C₁ carrier in its reduced thiol form, which is catalyzed by glutathione reductase and requires NADPH as its source of reducing equivalents. We therefore hypothesized a positive interaction between *gshA*^{evo} and *pntAB*^{evo}, which respectively upregulate expression of the genes contributing to GSH and NAD⁺ production, as high levels of both of these metabolites should enable increased flux through the *Foreign* pathway. Instead, the fitness cost of *pntAB*^{evo} was greater when introduced into backgrounds already containing *gshA*^{evo}, and vice versa ($P=0.005$, unpaired *t*-test). In light of this observation, we hypothesized that increased transhydrogenase activity could be expected to be universally deleterious in the *Native* pathway context due to futile cycling arising from dissipation of the transmembrane proton gradient. However, upon further examination, *pntAB*^{evo} did not provide universal fitness effects in the *Dual* pathway background. Whereas this allele was deleterious as the sole additional allele introduced, in many of the other backgrounds it was neutral, and, in the presence of *GB*^{evo} was even beneficial. The *GB*^{evo} allele is a composite of six mutations, one of which (the large deletion) has been directly demonstrated to be beneficial in the *Native* pathway context of the WT strain during growth on methanol (Lee and Marx 2012). Here we observed that *GB*^{evo} behaved as a neutral allele when introduced alone into *Dual* or *Native* contexts, but in many of the combinations could be strongly beneficial, indicating synergistic epistasis with other evolved alleles (particularly *pntAB*^{evo} or *gshA*^{evo}). At this point, the physiological basis of these many examples of sign epistasis remains open but leads to caution in assuming that the individual effects of mutations in new pathway contexts will hold for combinations of multiple alleles. Additionally, we should question whether known interactions observed in one pathway context will hold when present in an altered metabolic setting.

Given the extent of epistasis observed, would we expect evolution to favor similar beneficial mutations across the contexts of the *Native*, *Foreign*, and *Dual* pathway genotypes? If only individual mutational effects are considered, there was little congruence between these landscapes (fig. 6). Due to numerous sign epistatic interactions, the fitness peaks on each of the *Dual* and *Native* landscapes had many overlapping alleles with the optimal alleles in the *Foreign* context. The frequent neutral or deleterious effects observed in the *Dual* and *Native* landscapes, however, would generate a much more restricted set of paths to these peaks (fig. 5 and supplementary fig. S7, [Supplementary Material](#) online). Interestingly, the characterized trajectory for the *Foreign* pathway population studied (*gshA*^{evo} then *fghA*^{evo} then *pntAB*^{evo};

Chubiz et al. 2012) would not be selectively accessible in either the *Dual* or *Native* pathway contexts.

Our results corroborate the many challenges with adapting to utilize a new, horizontally acquired degenerate metabolic pathway seen in the *Dual* context; however, there are also genetic and physiological factors that might increase the potential for degeneracy to emerge. Many broad-host-range plasmids encode functions to aid their carriage despite imposed fitness costs (Zielenkiewicz and Ceglowski 2001; Feschotte and Pritham 2007; Bahl et al. 2009; Howard-Varona et al. 2017; Legrand et al 2019). Accordingly, foreign pathways introduced directly into the chromosome (such as via recombination) would enjoy several advantages: stable segregation, generally smaller maintenance costs (if any), lower copy number which would reduce gene dosage, and genetic linkage to other potentially beneficial traits. Based on data from this study and previous studies, it can be interpreted that nearly half (~5.8%/14.1%) of the cost associated with pCM410 is associated with maintenance and carriage of the plasmid itself. Consistent with this, integration events of the pCM410 plasmid with host genome were rather common during adaptation of the *Foreign* pathway populations (Chou and Marx 2012; Lee and Marx 2013). However, although it may be more difficult to directly lose a chromosomally encoded pathway, it is still possible we may observe the elimination of degeneracy through selection for knockout mutations impacting one of the degenerate pathways. HGT-acquired pathways with lower expression may also be expected to have corresponding lower initial costs. The plasmid used herein had high, relatively constant level of expression (Marx and Lidstrom 2001; Chou et al 2011, 2014; Chou and Marx 2012). By using the pCM410 plasmid and its lower expression derivative (*fghA^{ev0}*), we were able to show epistatic interactions between beneficial mutations and the pathway(s) present can have tremendous impacts on maintenance costs. Some of these may be compensatory mutations that alleviate plasmid maintenance costs (Bouma and Lenski 1988; Jordt et al. 2020), but others may reduce fitness costs by direct mediation of the physiological effects of the new pathway. Given the abundant evidence for the importance of HGT as an avenue for evolution in microbes (Gogarten and Townsend 2005; Hall et al. 2005, 2017; Boto 2014; Touchon et al. 2017), these hurdles have clearly been overcome in a wide variety of cases. Often this has led to strains with tremendous degeneracy, containing several unrelated versions of functional modules that provide similar benefits to the cell (Chistoserdova 2011; Nayak et al. 2016; von Borzyskowski et al. 2020). In our study, we have shown that this could be potentiated, or even become beneficial, by providing merely one additional genomic mutation. Our results have relevance beyond novel metabolic pathways, as epistasis is likely to impact the retention of redundancy for many diverse biological functions. This work provides evidence that epistatic interactions may be critical drivers in the retention of degenerate

functions achieved over short time periods, even in conditions where heightened selection for the degenerate pathway is seemingly unlikely, by potentiating resulting genotypes for rapid or even immediate incorporation of newly acquired redundant functions. Likewise, natural variation within and between populations, in the light of the results described here, might promote acquisition and/or retention of novel genes via HGT, and should motivate future evolution experiments that focus on the gain of redundant or degenerate functions; like gene duplication, conflicts and redundancies may resolve in diverse ways with large potential consequences for evolutionary innovations.

Materials and Methods

Bacterial Strains

Strains used in this study are all derivatives of *M. extorquens* AM1 and are described in supplementary table S1, [Supplementary Material](#) online. The WT AM1 strain contains a single chromosome as well as four plasmids in its genome (Vuilleumier et al. 2009). The plasmid pCM410 and the evolved derivative pCM410.1145 are low-copy-number, kanamycin-resistant, IncP replicons that express *fghA* and *flhA* under control of the strong *P_{mxsF}* promoter. The strains under analysis for epistatic comparisons contain combinations of evolved alleles previously described, categorized as *fghA*, *pntAB*, *gshA*, and *GB* mutations (Chou et al. 2011). The Genetic Background ("GB") allele is actually a composite of six additional mutations that occurred in the CM1145 population (one SNP on the megaplasmid, loss of plasmid p3META, two IS insertions, one small insertion, and one large deletion) (Chou et al. 2011). One of these IS insertions (Chou et al. 2009) and the large deletion (Lee and Marx 2012) were commonly observed in populations of WT *M. extorquens* evolving in other nutritional environments and were determined to be beneficial for traits unlinked to formaldehyde oxidation. None of the other four loci has an obvious connection to methylotrophy. For further information on these mutations, refer to supplementary table S2, [Supplementary Material](#) online. For convenience, the genotypes used are often referred to by their bitstring representations at the focal alleles under examination, where an allelic state of "0" represents the ancestral ("^{anc}") allele, a "1" represents the evolved ("^{ev0}"), and a "-" represents the absence of an allele/pathway (supplementary fig. S1A and supplementary table S1, [Supplementary Material](#) online). Presence or absence of a given pathway is indicated by "1" or "0," respectively.

Media and Chemicals

All growth experiments were carried out in a modified Hypho growth medium (Vishniac and Santer 1957; Attwood and Harder 1972; Chou et al. 2009). One liter of growth media consisted of 1 ml of a trace metal solution (12.74 g of ethylenediaminetetraacetic acid disodium salt dihydrate, 4.4 g of

ZnSO₄·7H₂O, 1.47 g of CaCl₂·2H₂O, 1.01 g of MnCl₂·4H₂O, 0.22 g of (NH₄)₆Mo₇O₂₄·4H₂O, 0.31 g of CuSO₄·5H₂O, 0.32 g of CoCl₂·6H₂O, and 1.00 g of FeSO₄·7H₂O per liter), 100 ml of phosphate buffer (25.3 g of K₂HPO₄ and 22.5 g of NaH₂PO₄ in 1 l of deionized water), 100 ml of sulfate solution (5 g of (NH₄)₂SO₄ and 0.98 g of MgSO₄ in 1 l of deionized water), 799 ml of deionized water, and the desired carbon source (33.3 mM final phosphate buffer concentration, pH 6.73). Carbon was supplemented to the medium as disodium succinate (3.5 mM, Fisher) or anhydrous methanol (15 mM, Macron Fine Chemicals). These carbon concentrations yield comparable final population sizes between carbon sources (Lee et al. 2009). Additionally, the concentration of methanol used is substantially higher than the reported K_m for methanol dehydrogenase and should not be rate limiting (Anthony and Zatman 1964). The modified trace metal mix (per 100 ml) consisted of 10 ml of 179.5 mM FeSO₄ (5× relative to above recipe), 80 ml of pre-mixed metal mix (12.74 g of ethylenediaminetetraacetic acid disodium salt dihydrate, 4.4 g of ZnSO₄·7H₂O, 1.47 g of CaCl₂·2H₂O, 1.01 g of MnCl₂·4H₂O, 0.22 g of (NH₄)₆Mo₇O₂₄·4H₂O, 0.31 g of CuSO₄·5H₂O, and 0.32 g of CoCl₂·6H₂O in 1 l of deionized water, pH 5), and 10 ml of deionized water.

Competition Fitness Assays

All cultures were grown at 30 °C in 5 ml of media in sealed Balch tubes (22 ml total capacity) that were aerated using roller drums rotating at approximately 60 rpm. All competition experiments were carried out using a minimum of three biological replicates for each competing genotype. Starter cultures for all strains were initiated from single colonies into Hypho media containing 7.5 mM methanol and 1.75 mM succinate for 2 days. The cultures were then diluted 1/64 into acclimation cultures of Hypho media containing 15 mM methanol for 2–4 days (depending on the growth rate of a given strain in methanol). Finally, cultures were mixed and diluted 1/64 into homologous Hypho media containing 15 mM methanol and competed for a duration of 4 days. All competition assays were initiated as a 1:1 volumetric ratio against an *mCherry* fluorescently labeled reference strain (CM1232 or CM1176, see supplementary table S1, [Supplementary Material](#) online) as described previously (Chou and Marx 2012). Frequencies of competitors were determined by passing mixed population samples from the start (F_0) and end (F_1) of the experiment through a Cytotflex S flow cytometer (Beckman Coulter). Approximately 50,000 cells per sample were gated by forward and side scatter, and competitors were differentiated through comparison of fluorescence by excitation at 561 nm and emission through a 610/20-nm BP filter. Where required, additional control strains chromosomally tagged with the Venus fluorescent reporter gene were also included and assessed by excitation at 488 nm and detection of emission with a 525/40-nm BP filter. Malthusian fitness measurements (W) were determined by ratios of log-transformed

changes in relative frequencies over the course of the experiment relative to a common competitor strain (*Foreign* ancestor CM1232, or WT CM1176), by the following formula that assumes an average of 64-fold (2^6) expanded size of mixed populations during competition experiments (Lee et al. 2009; Chou et al. 2011): $W = \log(F_1 * 64 / F_0) / \log((1 - F_1) * 64 / (1 - F_0))$. Competitions against WT utilized the CM1176 strain that lacks the pCM410 plasmid in order to allow competitions against *Native* pathway strains without prior selection for plasmid maintenance and to place WT at the baseline relative fitness of $W = 1$.

To qualify the conditions used in this study, we compared these results against conditions used previously. The correlation between the fitness values reported previously with cells grown in 50-ml flasks and the new data with cultures grown in tubes placed in a roller drum due to the large quantity involved was high ($R^2 = 0.9$, $P = 1.98\text{e-}08$; supplementary fig. S2A, [Supplementary Material](#) online). Directly comparing both sets of strains both grown in roller drums indicated that it was growth conditions and/or the flow cytometry-based fitness assay that generated the remaining differences, not the strains themselves (supplementary fig. S2B, [Supplementary Material](#) online). One reason that this correlation has a slope below unity is due to a confounding effect of cell size and distinguishing fluorescent from nonfluorescent strains via flow cytometry that was discovered in subsequent work that led to a consistent underestimate of fitness values for these strains in the initial study (Chou et al. 2011). Because the *Native* and *Dual* sets of strains had the *Native* dH₄MPT pathway, they were much fitter than any of the *Foreign* pathway strains, even the fittest strain in that context which had all four evolved alleles present (supplementary table S3, [Supplementary Material](#) online). Because of this, a fluorescent version of the WT strain (10-x000) was used for competition assays to provide more precise fitness estimates and the correlation between both competitor strain approaches was reasonably high ($R^2 = 0.68$, $P = 7.85\text{e-}07$; supplementary fig. S3, [Supplementary Material](#) online). However, for dH₄MPT pathway-containing strains, we opted to use the data from competitions against WT because they allowed better resolution between strains and these data had lower levels of variation on the whole, largely as a result of dH₄MPT pathway strains having much higher fitness values in comparison to CM1232 (average coefficient of variation of 1.9% vs. 4.7%).

In order to confirm that our fitness measurements were broadly consistent with prior work, we compared our results with the *Foreign* combinations previously examined (Chou et al. 2011); these strains were regenerated so that all strains considered here shared an identical genetic platform. As reported previously, all permutations of evolved alleles were found to be beneficial in the context of the *Foreign* pathway alone (supplementary tables S3 and S4, [Supplementary Material](#) online). The ordering of selection coefficients of the individual alleles (*gshA*^{evo}, *fgaA*^{evo}, *GB*^{evo}, *pntAB*^{evo}) was likewise maintained (supplementary tables S3 and S4, [Supplementary Material](#)

online). The largest difference between the data sets was that combinations with three or more evolved alleles were found to outperform previous observations (supplementary fig. S2, [Supplementary Material](#) online). Despite this, the overall landscape remained smooth and universally accessible (fig. 5 and supplementary fig. S4B, [Supplementary Material](#) online).

Growth Assays

To assess growth properties of a subset of strains in the study, cell cultures were initiated and passaged identically to competition experiments, going through starter and acclimation cultures prior to actual growth experiments. Cultures were diluted 1/64 from stationary phase cultures and grown in identical Hypho media in sealed Balch tubes (5 ml media in 22 ml total capacity), and growth was monitored by measurement in a Spectronic 200 spectrophotometer. All growth experiments were carried out with at least three biological replicates per genotype. Growth rates were calculated from growth curves using linear regression upon log-transformed absorbance at 600 nm values representing exponential growth of the cultures.

Statistical Comparisons

Model fits were made with linear regression in R (version 3.6.1). Analysis of growth was conducted using GraphPad Prism 8.4.3. Comparison of plasmid retention was conducted via ANOVA on frequency data at 96 generations.

To assess whether a series of substitutions (i.e., an evolutionary trajectory) was selectively favored, we used repeated sampling from empirical distributions of fitness as follows. Sample means and variances of least three empirical fitness measurements per strain were used to parametrize a Gaussian distribution for the fitness of each strain. The probability that a single transition (e.g., 01-0000 → 01-0100) or multiple transitions constituting a trajectory (e.g., 01-0000 → 01-0100 → 01-0110 → 01-0111 → 01-1111) were adaptive (i.e., each step increasing in fitness) could then be calculated by drawing a value for each genotype's fitness from these distributions. This procedure corrects for the fact that steps along a trajectory are not independent, and effectively constitutes a numerical version of a *t*-test for all transitions along a given trajectory. Probabilities of transitions or trajectories being beneficial were then calculated as the fraction of draws in this ensemble for which a given quantitative relationship (e.g., $S_{\text{evolved}} > 1\%$) is true. A probability for the overall likelihood of a transition or trajectory being accessible was determined over the course of those 10^6 samples. Among the *Native* and *Dual* strains, alleles that had slightly negative selection coefficients on average were still likely to generate beneficial draws a significant portion of the time due to uncertainty in the fitness measurements (supplementary fig. S5, [Supplementary Material](#) online). Individual transitions were assigned a designation as beneficial, deleterious, or neutral based on the net probabilities that draws fell into these different classes

(supplementary fig. S1B, [Supplementary Material](#) online). The probability of reaching the genotype in each particular pathway background with evolved alleles at each locus is the proportion of draws in which all mutational steps along that trajectory had $S_{\text{evolved}} > 1\%$ to focus on alleles under sufficient positive selection to allow a higher likelihood of increasing in frequency (Gerrish and Lenski 1998; Patwa and Wahl 2008). Where reported, selection coefficients (*S*) are calculated as previously under a multiplicative model (Lenski et al. 1991; Miller et al. 2018) as $S = W_{\text{evolved}}/W_{\text{ancestor}} - 1$.

Genetic Techniques

Evolved alleles identified in previous studies (Chou et al. 2011) were combined in all possible permutations into *Native* (WT) and *Dual* (WT + plasmid) pathway strains. Chromosomal mutations were introduced by allelic exchange through conjugation as previously described (Marx 2008) and plasmid-based mutations were introduced by transformation of plasmid variants (i.e., pCM410 and the evolved derivative pCM410.1145; Chou et al. 2011).

Plasmid Carriage Experiments

Biological triplicate samples of five genotypes from the *Dual* pathway strain set (11-0000, 11-0001, 11-0010, 11-0100, and 11-1000) were inoculated from plates into liquid media containing 3.5 mM disodium succinate and 50 µg/ml kanamycin to grow cultures up to substantial biomass and ensure cell plasmid carriage. From here, cultures were diluted 1/64 in 5 ml liquid media containing 15 mM methanol and grown for 2 days and then subsequently diluted into new media. Additionally, we passaged genotype 11-0000 in liquid media with 3.5 mM succinate as the sole carbon source and lacking kanamycin as a control where we expected rapid loss of the plasmid. Every 4 days (2 transfers = 12 generations), subsamples were taken from the populations, and a dilution series was completed and plated on media with and without kanamycin to assess any change in frequency of plasmid-containing cells over time. A subsequent plasmid-retention experiment involved 20 replicate populations of the 11-1010 genotype (strain CM3290).

Acknowledgments

This work was supported by funding from the National Science Foundation to C.J.M. (MCB-1714949) and J.A.D. (MCB-1714550), the National Institutes of Health to C.J.M. (R01 GM078209), the National Science Foundation funded BEACON Center for the Study of Evolution in Action (parent award DBI-0939454) to E.L.B. and C.J.M., the University of Idaho Office of Undergraduate Research to C.J.R., and National Science Foundation REU support for N.V.E. (DBI-1757826). We thank J. Bazurto (<https://orcid.org/0000-0001-9012-2260>), A. Borchert (<https://orcid.org/0000-0003-1435-3643>), A. Burmeister (<https://orcid.org/0000-0003-4304-5551>), and

members of the Marx lab for critical reading and commentary on the manuscript.

Author Contributions

Conceptualization: E.L.B., L.M.C., and C.J.M.; Methodology: E.L.B.; Formal Analysis: E.L.B. and J.A.D.; Investigation: E.L.B., L.M.C., J.I.R.E., C.J.R., and N.V.E.; Software: E.L.B. and J.A.D.; Resources: J.A.D. and C.J.M.; Data Curation: E.L.B.; Writing—Original Draft Preparation: E.L.B.; Writing—Review & Editing: L.M.C., J.I.R.E., C.J.R., N.V.E., J.A.D., and C.J.M.; Validation: E.L.B., C.J.R., and N.V.E.; Visualization: E.L.B. and C.J.M.; Supervision: J.A.D. and C.J.M.; Project Administration: J.A.D. and C.J.M.; Funding Acquisition: E.L.B., J.A.D., and C.J.M.

Data Availability

The data underlying this article are available in the article and in its [Supplementary Material](#) online.

Literature Cited

- Anthony C, Zatman LJ. 1964. The microbial oxidation of methanol. 2. The methanol-oxidizing enzyme of *Pseudomonas* sp. M27. *Biochem J*. 92(3):614–621.
- Attwood MM, Harder W. 1972. A rapid and specific enrichment procedure for *Hyphomicrobium* spp. *Antonie Van Leeuwenhoek*. 38(3):369–377.
- Bahl MI, Hansen LH, Sørensen SJ. 2009. Persistence mechanisms of conjugative plasmids. *Methods Mol Biol*. 532:73–102.
- Blount ZD, Barrick JE, Davidson CJ, Lenski RE. 2012. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* 489(7417):513–518.
- Blount ZD, Borland CZ, Lenski RE. 2008. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc Natl Acad Sci U S A*. 105(23):7899–7906.
- Boto L. 2014. Horizontal gene transfer in the acquisition of novel traits by metazoans. *Proc Biol Sci*. 281:1–8.
- Bouma JE, Lenski RE. 1988. Evolution of a bacteria/plasmid association. *Nature* 335(6188):351–352.
- Carroll SM, Marx CJ. 2013. Evolution after introduction of a novel metabolic pathway consistently leads to restoration of wild-type physiology. *PLoS Genet*. 9(4):e1003427.
- Chistoserdova L. 2011. Modularity of methylotrophy, revisited. *Environ Microbiol*. 13(10):2603–2622.
- Chistoserdova L, Vorholt JA, Thauer RK, Lidstrom ME. 1998. C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic Archaea. *Science* 281(5373):99–102.
- Chou H-H, Berthet J, Marx CJ. 2009. Fast growth increases the selective advantage of a mutation arising recurrently during evolution under metal limitation. *PLoS Genet*. 5(9):e1000652.
- Chou H-H, Chiu H-C, Delaney NF, Segrè D, Marx CJ. 2011. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* 332(6034):1190–1192.
- Chou H-H, Delaney NF, Draghi JA, Marx CJ. 2014. Mapping the fitness landscape of gene expression uncovers the cause of antagonism and sign epistasis between adaptive mutations. *PLoS Genet*. 10(2):e1004149.
- Chou H-H, Marx CJ. 2012. Optimization of gene expression through divergent mutational paths. *Cell Rep*. 1(2):133–140.
- Chubiz LM, Lee M-C, Delaney NF, Marx CJ. 2012. FREQ-Seq: a rapid, cost-effective, sequencing-based method to determine allele frequencies directly from mixed populations. *PLoS One* 7(10):e47959.
- de Felipe KS, et al. 2005. Evidence for acquisition of legionella type IV secretion substrates via interdomain horizontal gene transfer. *J Bacteriol*. 187(22):7716–7726.
- Ding Y, Zhou Q, Wang W. 2012. Origins of new genes and evolution of their novel functions. *Annu Rev Ecol Evol Syst*. 43(1):345–363.
- Douglas SM, Chubiz LM, Harcombe WR, Marx CJ. 2017. Identification of the potentiating mutations and synergistic epistasis that enabled the evolution of inter-species cooperation. *PLoS One* 12(5):e0174345.
- Edelman GM, Gally JA. 2001. Degeneracy and complexity in biological systems. *Proc Natl Acad Sci U S A*. 98(24):13763–13768.
- Emamalipour M, et al. 2020. Horizontal gene transfer: from evolutionary flexibility to disease progression. *Front Cell Dev Biol*. 8:1–16.
- Feschotte C, Pritham EJ. 2007. DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet*. 41:331–368.
- Fuhrer T, Fischer E, Sauer U. 2005. Experimental identification and quantification of glucose metabolism in seven bacterial species. *J Bacteriol*. 187(5):1581–1590.
- Gerrish PJ, Lenski RE. 1998. The fate of competing beneficial mutations in an asexual population. *Genetica* 102:127–144.
- Ghosh S, O'Connor TJ. 2017. Beyond paralogs: the multiple layers of redundancy in bacterial pathogenesis. *Front Cell Infect Microbiol*. 7:1–14.
- Gogarten JP, Townsend JP. 2005. Horizontal gene transfer, genome innovation and evolution. *Nat Rev Microbiol*. 3(9):679–687.
- Hall C, Brachat S, Dietrich FS. 2005. Contribution of horizontal gene transfer to the evolution of *Saccharomyces cerevisiae*. *Eukaryot Cell*. 4(6):1102–1115.
- Hall JPJ, Brockhurst MA, Harrison E. 2017. Sampling the mobile gene pool: innovation via horizontal gene transfer in bacteria. *Philos Trans R Soc Lond B Biol Sci*. 372:1–10.
- Hall RJ, Whelan FJ, McInerney JO, Ou Y, Domingo-Sananes MR. 2020. Horizontal gene transfer as a source of conflict and cooperation in prokaryotes. *Front Microbiol*. 11:1–7.
- Harms N, Ras J, Reijnders WN, van Spanning RJ, Stouthamer AH. 1996. S-formylglutathione hydrolase of *Paracoccus denitrificans* is homologous to human esterase D: a universal pathway for formaldehyde detoxification? *J Bacteriol*. 178(21):6296–6299.
- Harrison E, Brockhurst MA. 2012. Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends Microbiol*. 20(6):262–267.
- Hendrickson HL, Barbeau D, Ceschin R, Lawrence JG. 2018. Chromosome architecture constrains horizontal gene transfer in bacteria. *PLoS Genet*. 14(5):e1007421.
- Howard-Varona C, Hargreaves KR, Abedon ST, Sullivan MB. 2017. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *ISME J*. 11(7):1511–1520.
- Hughes AL. 1994. The evolution of functionally novel proteins after gene duplication. *Proc Biol Sci*. 256(1346):119–124.
- Imanian B, Keeling PJ. 2014. Horizontal gene transfer and redundancy of tryptophan biosynthetic enzymes in dinoflagellates. *Genome Biol Evol*. 6(2):333–343.
- Jackson JB. 2003. Proton translocation by transhydrogenase. *FEBS Lett*. 545(1):18–24.
- Jordt H, et al. 2020. Coevolution of host-plasmid pairs facilitates the emergence of novel multidrug resistance. *Nat Ecol Evol*. 4(6):863–869.
- Juhas M, et al. 2009. Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev*. 33(2):376–393.
- Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. 2011. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* 332(6034):1193–1196.
- Kryazhimskiy S, Rice DP, Jerison ER, Desai MM. 2014. Microbial evolution. Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science* 344(6191):1519–1522.

- Kurland CG. 2005. What tangled web: barriers to rampant horizontal gene transfer. *Bioessays* 27(7):741–747.
- Lang GI, Botstein D, Desai MM. 2011. Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics* 188(3):647–661.
- Lee M-C, Chou H-H, Marx CJ. 2009. Asymmetric, bimodal trade-offs during adaptation of *Methylobacterium* to distinct growth substrates. *Evolution* 63(11):2816–2830.
- Lee M-C, Marx CJ. 2012. Repeated, selection-driven genome reduction of accessory genes in experimental populations. *PLoS Genet.* 8(5):e1002651.
- Lee M-C, Marx CJ. 2013. Synchronous waves of failed soft sweeps in the laboratory: remarkably rampant clonal interference of alleles at a single locus. *Genetics* 193(3):943–952.
- Legrand S, et al. 2019. Differential retention of transposable element-derived sequences in outcrossing *Arabidopsis* genomes. *Mob DNA.* 10:1–17.
- Lenski RE, Rose MR, Simpson SC, Tadler SC. 1991. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am Nat.* 138(6):1315–1341.
- Lerner A, Matthias T, Aminov R. 2017. Potential effects of horizontal gene exchange in the human gut. *Front Immunol.* 8:1–14.
- Louca S, et al. 2018. Function and functional redundancy in microbial systems. *Nat Ecol Evol.* 2(6):936–943.
- Majewski J, Zawadzki P, Pickerill P, Cohan FM, Dowson CG. 2000. Barriers to genetic exchange between bacterial species: streptococcus pneumoniae transformation. *J Bacteriol.* 182(4):1016–1023.
- Marx CJ. 2008. Development of a broad-host-range *sacB*-based vector for unmarked allelic exchange. *BMC Res Notes.* 1:1–8.
- Marx CJ, Chistoserdova L, Lidstrom ME. 2003. Formaldehyde-detoxifying role of the tetrahydromethanopterin-linked pathway in *Methylobacterium extorquens* AM1. *J Bacteriol.* 185(24):7160–7168.
- Marx CJ, Lidstrom ME. 2001. Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria. *Microbiology (Reading)* 147(Pt 8):2065–2075.
- Marx CJ, Miller JA, Chistoserdova L, Lidstrom ME. 2004. Multiple formaldehyde oxidation/detoxification pathways in *Burkholderia fungorum* LB400. *J Bacteriol.* 186(7):2173–2178.
- Marx CJ, Van Dien SJ, Lidstrom ME. 2005. Flux analysis uncovers key role of functional redundancy in formaldehyde metabolism. *PLoS Biol.* 3(2):e16.
- Mason PH, Domínguez D JF, Winter B, Grignolio A. 2015. Hidden in plain view: degeneracy in complex systems. *Biosystems* 128:1–8.
- Miller CR, Van Leuven JT, Wichman HA, Joyce P. 2018. Selecting among three basic fitness landscape models: additive, multiplicative and stick-breaking. *Theor Popul Biol.* 122:97–109.
- Moore JH, Williams SM. 2005. Traversing the conceptual divide between biological and statistical epistasis: systems biology and a more modern synthesis. *Bioessays* 27(6):637–646.
- Näsvall J, Sun L, Roth JR, Andersson DI. 2012. Real-time evolution of new genes by innovation, amplification, and divergence. *Science* 338(6105):384–387.
- Nayak DD, Agashe D, Lee M-C, Marx CJ. 2016. Selection maintains apparently degenerate metabolic pathways due to tradeoffs in using methylamine for carbon versus nitrogen. *Curr Biol.* 26(11):1416–1426.
- Nowak MA, Boerlijst MC, Cooke J, Smith JM. 1997. Evolution of genetic redundancy. *Nature* 388(6638):167–171.
- Oliveira PH, Touchon M, Cury J, Rocha EP. 2017. The chromosomal organization of horizontal gene transfer in bacteria. *Nat Commun.* 8:1–11.
- Patwa Z, Wahl LM. 2008. The fixation probability of beneficial mutations. *J R Soc Interface.* 5(28):1279–1289.
- Pfeiffer T, Schuster S, Bonhoeffer S. 2001. Cooperation and competition in the evolution of ATP-producing pathways. *Science* 292(5516):504–507.
- Phillips PC. 2008. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet.* 9(11):855–867.
- Ras J, et al. 1995. Isolation, sequencing, and mutagenesis of the gene encoding NAD- and glutathione-dependent formaldehyde dehydrogenase (GD-FALDH) from *Paracoccus denitrificans*, in which GD-FALDH is essential for methylotrophic growth. *J Bacteriol.* 177(1):247–251.
- Redondo-Salvo S, et al. 2020. Pathways for horizontal gene transfer in bacteria revealed by a global map of their plasmids. *Nat Commun.* 11:1–13.
- Rezazadegan R, Reidys C. 2018. Degeneracy and genetic assimilation in RNA evolution. *BMC Bioinformatics* 19(1):1–17.
- Rokyta DR, et al. 2011. Epistasis between beneficial mutations and the phenotype-to-fitness Map for a ssDNA virus. *PLoS Genet.* 7(6):e1002075.
- Salverda MLM, et al. 2011. Initial mutations direct alternative pathways of protein evolution. *PLoS Genet.* 7(3):e1001321.
- Sambamurthy G, Raman K. 2018. Understanding the evolution of functional redundancy in metabolic networks. *Bioinformatics* 34(17):i981–i987.
- Sambamurthy G, Sinha H, Raman K. 2019. Evolutionary design principles in metabolism. *Proc Biol Sci.* 286(1898):20190098.
- Schoustra S, Hwang S, Krug J, de Visser JAGM. 2016. Diminishing-returns epistasis among random beneficial mutations in a multicellular fungus. *Proc Biol Sci.* 283:1–9.
- Shoji T. 2019. The recruitment model of metabolic evolution: jasmonate-responsive transcription factors and a conceptual model for the evolution of metabolic pathways. *Front Plant Sci.* 10:1–12.
- Silva RF, et al. 2011. Pervasive sign epistasis between conjugative plasmids and drug-resistance chromosomal mutations. *PLoS Genet.* 7(7):e1002181.
- Sorek R, et al. 2007. Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* 318(5855):1449–1452.
- Thomas C, Nielsen K. 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat Rev Microbiol.* 3(9):711–721.
- Tokuriki N, et al. 2012. Diminishing returns and tradeoffs constrain the laboratory optimization of an enzyme. *Nat Commun.* 3(1):1–10.
- Tononi G, Sporns O, Edelman GM. 1999. Measures of degeneracy and redundancy in biological networks. *Proc Natl Acad Sci U S A.* 96(6):3257–3262.
- Touchon M, Moura de Sousa JA, Rocha EP. 2017. Embracing the enemy: the diversification of microbial gene repertoires by phage-mediated horizontal gene transfer. *Curr Opin Microbiol.* 38:66–73.
- Vishniac W, Santer M. 1957. The Thiobacilli. *Bacteriol Rev.* 21(3):195–213.
- von Borzyskowski LS, Bernhardsgrütter I, Erb TJ. 2020. Biochemical unity revisited: microbial central carbon metabolism holds new discoveries, multi-tasking pathways, and redundancies with a reason. *Biol Chem.* 401(12):1429–1441.
- Vuilleumier S, et al. 2009. *Methylobacterium* genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources. *PLoS One* 4(5):e5584.
- Wein T, Hülter NF, Mizrahi I, Dagan T. 2019. Emergence of plasmid stability under non-selective conditions maintains antibiotic resistance. *Nat Commun.* 10(1):13.
- Weinreich DM, Watson RA, Chao L. 2005. Perspective: sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59(6):1165–1174.
- Whitacre JM. 2010. Degeneracy: a link between evolvability, robustness and complexity in biological systems. *Theor Biol Med Model.* 7(1):6.
- Wünsche A, et al. 2017. Diminishing-returns epistasis decreases adaptability along an evolutionary trajectory. *Nat Ecol Evol.* 1(4):61–66.
- Xie G, et al. 2003. Lateral gene transfer and ancient paralogy of operons containing redundant copies of tryptophan-pathway genes in *Xylella* species and in heterocystous cyanobacteria. *Genome Biol.* 4:1–18.

- Yu JJ, Whalen JK. 2020. A new perspective on functional redundancy and phylogenetic niche conservatism in soil microbial communities. *Pedosphere* 30(1):18–24.
- Zhang J. 2003. Evolution by gene duplication: an update. *Trends Ecol Evol.* 18(6):292–298.

- Zielenkiewicz U, Ceglowski P. 2001. Mechanisms of plasmid stable maintenance with special focus on plasmid addiction systems. *Acta Biochim Pol.* 48(4):1003–1023.

Associate editor: George Zhang