NIH Public Access
Author Manuscript
Science. Author manuscript; available in PMC 2011 December 21.

Published in final edited form as:
Science. 2011 June 3; 332(6034): 1190–1192. doi:10.1126/science.1203799.

Diminishing returns epistasis among beneficial mutations decelerates adaptation

Hsin-Hung Chou[1,*], **Hsuan-Chao Chiu**[3], **Nigel F. Delaney**[1], **Daniel Segrè**[3,4], and **Christopher J. Marx**[1,2,†]

[1]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

[2]FAS Center for Systems Biology, Harvard University, Cambridge, MA 02138, USA

[3]Graduate Program in Bioinformatics, Boston University, Boston, MA 02215, USA

[4]Department of Biology and Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA

Abstract

Epistasis substantially impacts evolution, in particular the rate of adaptation. We generated combinations of beneficial mutations that arose in a lineage during rapid adaptation of a bacterium whose growth depended upon a newly-introduced metabolic pathway. The proportional selective benefit for three of the four loci consistently decreased when introduced upon more fit backgrounds. These three alleles all reduced morphological defects caused by expression of the foreign pathway. A simple theoretical model segregating the apparent contribution of individual alleles to benefits and costs effectively predicted the interactions between them. These results provide the first evidence that patterns of epistasis may differ for within- and between-gene interactions during adaptation, and that diminishing returns epistasis contributes to the consistent observation of decelerating fitness gains during adaptation.

Epistasis describes genetic interactions in terms of how phenotypic effects of a mutation depend upon other mutations in the genome. If two mutations act upon a given phenotype independently, each would be expected to exert the same proportional effect regardless of whether the other allele was present, although other models can be applied (1–3). Deviations from this null expectation have been used to uncover interacting genes via genetic screens for second mutations that suppress the effect of the first, identify the order of enzymes in biochemical pathways, and unravel systems-level interaction patterns characterized with genome-wide double knockout libraries. One general trend has been that the detrimental effect of a lesion in a pathway (or module) (4) is greater alone than when there is already another deleterious mutation in that process (*i.e.*, antagonistic epistasis). In contrast, lesions

[†]To whom correspondence should be addressed. Email: cmarx@oeb.harvard.edu.
[*]Present address: Institute of Molecular Systems Biology, ETH Zürich, CH-8093 Zürich, Switzerland.

Supporting Online Material
www.sciencemag.org
Materials and Methods
SOM
Text Figs. S1 to S10
Tables S1 to S3
References

Please refer to the complete version of record at http://www.sciencemag.org/. The manuscript may not be reproduced or used in any manner that does not fall within the fair use provisions of the Copyright Act without the prior, written permission of AAAS.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

in parallel pathways producing the same product tend to cause stronger phenotypes (synergistic epistasis) than expected; the extreme case of the latter, termed synthetic lethality, results in a non-viable genotype.

Epistasis between beneficial mutations remains largely unexplored. Previous studies examined epistasis between five amino acid (or promoter) substitutions within an allele of beta-lactamase selected for cefotaxime resistance in *Escherichia coli* (5). By constructing all possible mutation combinations within the beta-lactamase locus, a single-peaked fitness landscape was revealed with numerous cases where the identical mutation increased resistance on some backgrounds but decreased it on others (*i.e.*, sign epistasis). Similar results have been found for cofactor use by isopropylmalate dehydrogenase (6) and for hormone receptors (7). In contrast, few studies have addressed interactions between beneficial mutations in different genes (8, 9).

The distribution of epistatic interactions between mutations may greatly influence evolutionary outcomes —from the maintenance of sexual reproduction to the fixation rate of beneficial alleles— and hence the speed of adaptation itself. The most consistent finding across studies of laboratory-evolved populations has been a rapid deceleration of the rate of fitness increase (10). Theoretical analysis suggests that the observed dynamics of fitness increase and accumulation of substitutions (11) are best described by a class of fitness landscapes with antagonistic interactions between beneficial mutations (12).

We took both experimental and theoretical approaches to investigate potential epistasis in populations that were initiated with an engineered strain of *Methylobacterium extorquens* AM1 (hereafter 'EM'; table S1) and evolved in batch culture with methanol as the sole carbon source (3). In order to grow on methanol, *Methylobacterium* must oxidize formaldehyde into formate. Wildtype *Methylobacterium* (WT) performs this oxidation with a tetrahydromethanopterin-dependent pathway (13). In EM, this native pathway was eliminated and replaced by a non-orthologous, glutathione (GSH)- dependent pathway from *Paracoccus denitrificans* (fig. S1) (14). As a result, the EM strain could grow on methanol, but at a rate three-fold slower than WT (fig. S2). Adaptation in eight replicate populations dependent upon this engineered metabolic function (analogous to natural horizontal gene transfer) resulted in an average fitness increase after 600 generations of 66.8% (fig. S3), as determined by competition assays (3), and was largely carbon substrate-specific (fig. S4).

The genome of an evolved isolate from generation 600 ('EVO') (9) with the highest fitness (W_{EVO}=1.94; table S2) was sequenced to identify the genetic basis of adaptation in that lineage (3). In total, 9 mutations were identified (fig. S5) (3). We found an 11 bp deletion between the two genes that encode the GSH-dependent pathway, *flhA* and *fghA* (*i.e.*, $fghA^{EVO}$), in a plasmid specifically introduced into EM (pCM410, fig. S6). This deletion removed the apparent ribosome binding site for *fghA* and decreased expression of these enzymes by 55 and 73% (3), respectively. This change, however, increased fitness by 14.2% (Fig. 1A), suggesting that production of these enzymes in the EM ancestor was higher than the optimum. In WT, where the GSH pathway is extraneous, a strain with an empty vector had 14.1% fitness advantage relative to when both genes were expressed. It therefore appeared that the primary advantage of the $fghA^{EVO}$ allele was to reduce the costs of protein over-expression, (*e.g.*, energy consumption, ribosome sequestering, protein misfolding). We also identified a SNP in the promoter region of pyridine nucleotide transhydrogenase ($pntAB^{EVO}$), and a 2 bp deletion in the promoter of the most rate-limiting enzyme of GSH biosynthesis, γ-glutamylcysteine synthetase ($gshA^{EVO}$). These gene products have clear linkages to methanol utilization in EM (3). The remaining 6 genetic changes included a large deletion (fig. S7), a synonymous SNP, the loss of a plasmid, two transposon insertions and a 6 bp insertion (3). These latter six are either difficult to genetically reconstruct, were

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

individually neutral under our experimental conditions (15), or were deemed unlikely to greatly contribute to fitness. We thus treated them as a single collective locus, the '*genetic background*' (GB^{EVO}), for the purpose of examining epistasis between beneficial mutations. All identified alleles, when present individually in the ancestral background conferred fitness benefits ranging from 10 to 51% (Fig. 1A).

In order to investigate epistasis between these beneficial mutations, strains with each allelic combination ($2^4 = 16$) were constructed (3) and their fitness values measured (Fig. 1A). The adaptive landscape of this genotypic space contained a single peak; each allele was universally beneficial across genetic backgrounds (*i.e.*, showed no sign epistasis, but the degree of benefit conferred varied (Fig. 1B). Except for $pntAB^{EVO}$, the remaining three alleles exhibited a significant trend of diminishing returns: their selective benefits declined in genetic backgrounds with higher fitness. In contrast, the resistance to cefotaxime conferred by each mutation within the *E. coli* beta-lactamase gene (5) was idiosyncratic in regards to the resistance of the background onto which it was introduced (Fig. 1C).

Interestingly, we found a connection between antagonistic epistasis and a physiological problem caused by protein over-expression in EM. Cells of the EM ancestor showed an increased length and aberrant morphologies relative to WT (Figs. 2, S8), similar to those commonly observed for protein over-expression (16). Reducing expression of the foreign pathway in EM via $fghA^{EVO}$ suppressed cellular abnormalities, while expressing it in WT (where it is redundant) induced similar defects. This confirmed that the morphological defects were caused by over-expression of the foreign pathway. Additionally, the $gshA^{EVO}$ and GB^{EVO} alleles (but not $pntAB^{EVO}$) also individually reduced morphological defects (by ~threefold), and when all the evolved alleles are present together (*e.g.*, the EVO strain) abnormal cells were nearly absent (a finding recapitulated across all eight populations, fig. S2D). These data suggest that part of the benefit conferred by the three alleles whose selective benefit wanes on fitter backgrounds resulted from directly or indirectly decreasing protein over-expression costs.

Epistasis has been often represented as the deviation from a null model in which individual mutations affect the ancestor's fitness ($W_0 = 1.0$) with independent multiplicative factors λ_i (double mutant's fitness, $W_{ij} = \lambda_i\lambda_j W_0$). However, in our system, rather than being captured by a single indivisible phenotype, cell growth seems to depend upon at least one separately measurable component, *i.e.*, the growth burden imposed by expressing the foreign pathway. As stated above, three of the four alleles identified in the EVO strain appear to increase fitness at least partly by reducing this cost. Therefore, in analogy to the contributions to fitness by a single enzyme (17), we developed a mathematical model that partitions fitness into two phenotypes: a 'benefit' component b_0, analogous to a single conglomerate 'enzyme activity' that sets the rate of energy extracted from the substrate to generate biomass; and a cost c_0, encompassing a fixed amount of energy diverted to deal with over-expression of the foreign pathway

(3). Thus the fitness of the ancestral strain can be written, as $W_0 = b_0\text{-}c_0 = 1$. We hypothesize that a new allele *i* could modify the benefit and the cost of the ancestral background by certain multiplicative factors (λ_i and θ_i, respectively), giving rise to a fitness $W_i = \lambda_i b_0\text{-}\theta_i c_0$. A successive allele *j*, on top of the background of mutant *i*, is similarly assumed to act multiplicatively on the benefit and cost components, yielding a fitness $W_{ij} = \lambda_i\lambda_j b_0\text{-}\theta_i\theta_j c_0$.

If we could determine experimentally the values of b_0, c_0, and of λ_i and θ_i for each allele, then the above model should provide predictions for the fitness of any multi-allele strain, computable as $W_{mutant} = \prod_{i \in Alleles} \lambda_i b_0 - \prod_{i \in Alleles} \theta_i c_0$ (3). We estimated these parameters: cost was

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

determined by expressing the foreign pathway in WT, where its metabolic function was fully redundant (c_0=0.141) (17). Setting $W_0 = 1$, results in $b_0 = 1.141$. The lowered cost of expression, θ_i, for each allele was approximated as the decreased relative proportion of morphological defects (table S3) (3). Factors λ_i could then be estimated using the single-allele benefit-cost model.

Without specifying further information, this simple model partitioning fitness into benefit and cost outperformed the standard null model in predicting fitness values of multi-allele combinations ($R^2 = 0.97$ vs. $R^2 = 0.64$, figs. S9, S10) (3). It also recapitulated the antagonistic trend of epistasis among the three alleles affecting cell morphology and correctly predicted the consistent magnitude of benefit from $pntAB^{EVO}$ (Fig. 3). The agreement between our experimental data and model predictions supports our model assumptions and thus our hypothesis as to why diminishing returns epistasis was observed: proportional reductions of a cost became successively less beneficial as the cost itself was alleviated.

Diminishing returns has been predicted (18) but due to very different assumptions. The nonlinearity of fitness increase in these models arises because it is assumed that a given trait is under stabilizing selection for an intermediate optimum, thus explicitly considering fitness as being displaced from a fixed adaptive peak. While the assumption of intermediate optimality holds well for many traits like body weight and length, fitness rises monotonically with increasing growth rate or decreasing protein expression burden. In this study, we considered a higher-level phenotype (growth rate) as the sum of two constituent phenotypes (metabolic rate and protein expression burden) which allowed us to generate a precise expectation for fitness of multi-allele strains without explicitly assuming stabilizing selection. The success of this approach suggests that it may be possible to generalize the idea of expressing higher-level phenotypes (such as fitness) as combinations of multiple underlying traits to provide quantitative predictions of epistasis.

An analogous study (19) of the interactions between beneficial mutations in *E. coli* evolved in minimal glucose medium found similar epistatic trends: four of five new alleles exhibit significant diminishing returns. The fifth such mutation, and a mutation present as a component of our GB^{EVO} allele that is beneficial only in metal poor media (15), showed the opposite trend: an increase in selective advantage with higher background fitness. Thus, across these two distinct model systems 7 of 10 alleles consistently showed antagonism, whereas only two exhibited synergy. This tendency toward diminishing returns between beneficial mutations was predicted from trajectories of fitness increase and substitution rate (12) but had never been tested directly. Furthermore, these results are in stark contrast to the epistatic effects seen among mutations within single proteins, which are varying and unpredictable in their effect with regard to background activity (5, 7). This distinction between results from within- and between-gene epistasis suggests that the underlying causes of epistasis at different physiological scales (*i.e.*, within-gene protein biophysics vs. between-gene physiological networks) lead to categorically distinct, but reproducible trends in genetic interactions which affect both the speed of adaptation and the degree to which possible trajectories are limited.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank S. Kryazhminskiy, D. Weinreich, J. Weitz, C. Wilke, R. Whitaker and members of the Marx and Segrè labs for suggestions and discussions. Supported by NIH grant R01 GM078209, and NSF grant DEB-0845893. The

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

resulting single-end 36-bp read data have been deposited in the NCBI Short Read Archive with accession number SRA030695.1.

References and Notes

1. Phillips PC. Nat Rev Genet. Nov.2008 9:855. [PubMed: 18852697]
2. Cordell HJ. Human Mol Genet. Oct 1.2002 11:2463. [PubMed: 12351582]
3. Materials and methods are available as supporting material on *Science* Online.
4. Segrè D, Deluna A, Church GM, Kishony R. Nat Genet. Jan.2005 37:77. [PubMed: 15592468]
5. Weinreich DM, Delaney NF, Depristo MA, Hartl DL. Science. Apr 7.2006 312:111. [PubMed: 16601193]
6. Lunzer M, Miller SP, Felsheim R, Dean AM. Science. Oct 21.2005 310:499. [PubMed: 16239478]
7. Bridgham JT, Carroll SM, Thornton JW. Science. Apr 7.2006 312:97. [PubMed: 16601189]
8. Sanjuán R, Moya A, Elena SF. Proc Natl Acad Sci U S A. Oct 26.2004 101:15376. [PubMed: 15492220]
9. Applebee MK, Herrgard MJ, Palsson BO. J Bacteriol. Jul.2008 190:5087. [PubMed: 18487343]
10. Elena SF, Lenski RE. Nat Rev Genet. Jun.2003 4:457. [PubMed: 12776215]
11. Barrick JE, et al. Nature. Oct 29.2009 461:1243. [PubMed: 19838166]
12. Kryazhimskiy S, Tkacik G, Plotkin JB. Proc Natl Acad Sci U S A. Nov 3.2009 106:18638. [PubMed: 19858497]
13. Chistoserdova L, Vorholt JA, Thauer RK, Lidstrom ME. Science. Jul 3.1998 281:99. [PubMed: 9651254]
14. Marx CJ, Chistoserdova L, Lidstrom ME. J Bacteriol. Dec.2003 185:7160. [PubMed: 14645276]
15. Chou HH, Berthet J, Marx CJ. PLoS Genet. Sep.2009 5:e1000652. [PubMed: 19763169]
16. Kurland CG, Dong H. Mol Microbiol. Jul.1996 21:1. [PubMed: 8843428]
17. Dekel E, Alon U. Nature. Jul 28.2005 436:588. [PubMed: 16049495]
18. Martin G, Elena SF, Lenormand T. Nat Genet. Apr.2007 39:555. [PubMed: 17369829]
19. Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. (submitted).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

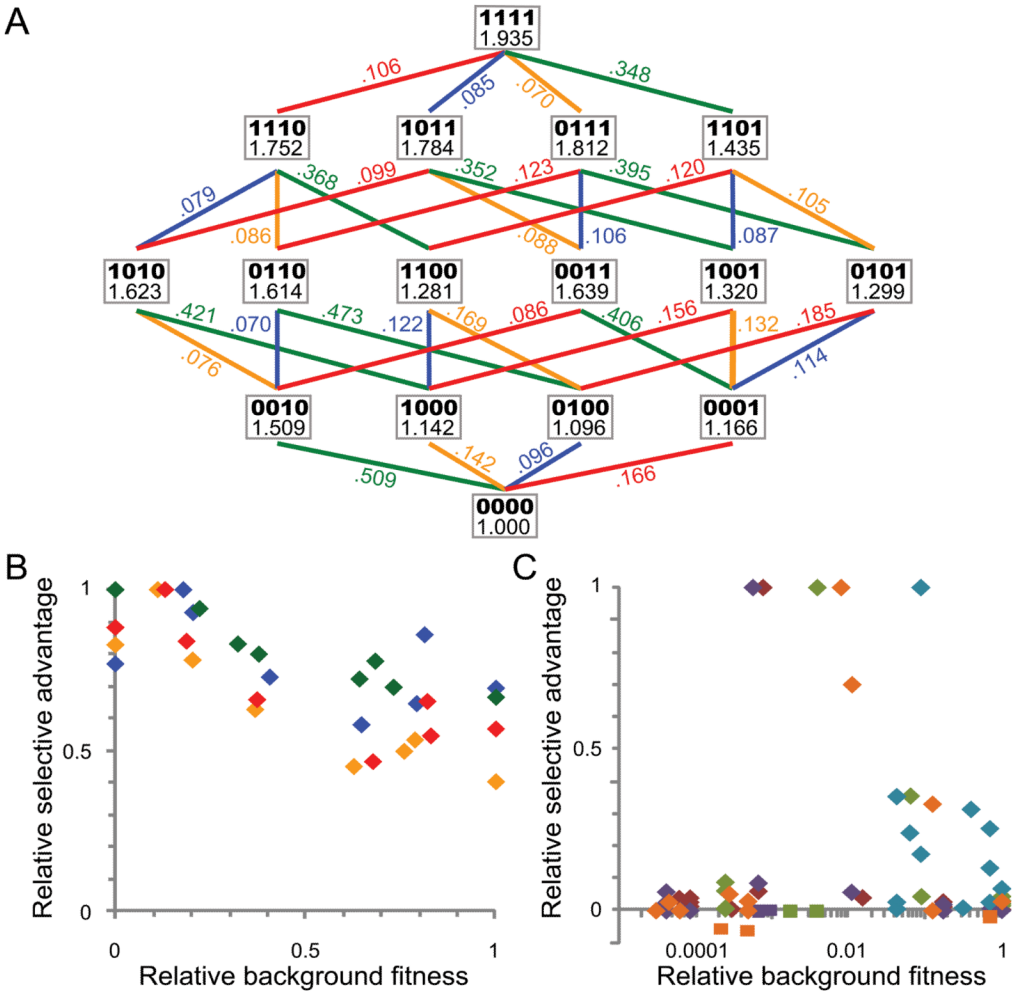

Figure 1.

Mutational network and distinct patterns of epistasis for mutations between and within genes. (**A**) Each node displays the allelic composition (*fghA*, *pntAB*, *gshA*, *GB*) of a given genotype (bold) and its fitness. Ancestral and evolved alleles are indicated by 0 and 1, respectively, leading from the ancestral EM strain (0000) to the evolved EVO isolate (1111). Each edge indicates an allelic replacement (*fghA*, orange; *pntAB*, blue; *gshA*, green; GB, red) and the corresponding selective coefficient. Variation in relative selective effect (normalized to max s_i) of each allele as a function of the fitness of the background was introduced into for: (**B**) between-gene epistasis in *Methylobacterium* (background fitness normalized from EM = 0 to max =1; colors as in A); or (**C**) within-gene epistasis for *E. coli* beta lactamase (background fitness normalized to max MIC = 1; log scale for visualization; squares indicate deleterious effects), data from (5).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

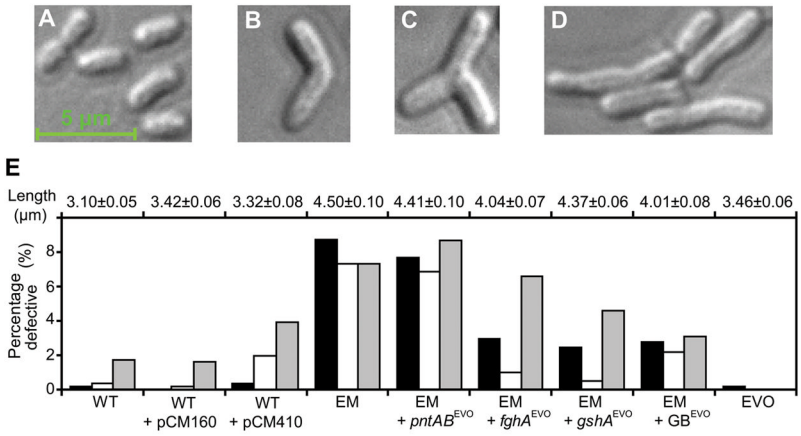

Figure 2.

Morphological aberrations caused by expression of the foreign pathway. Distinct cellular morphologies of (**B**) WT, or EM ancestor showing (**C**) curved, (**D**) branched, or (**E**) elongated cells. (**F**) Mean cell length and proportion of elongated (black), branched (white), and curved (grey) cells for various strains. Plasmid pCM410 expresses the foreign pathway while pCM160 is an empty control plasmid.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

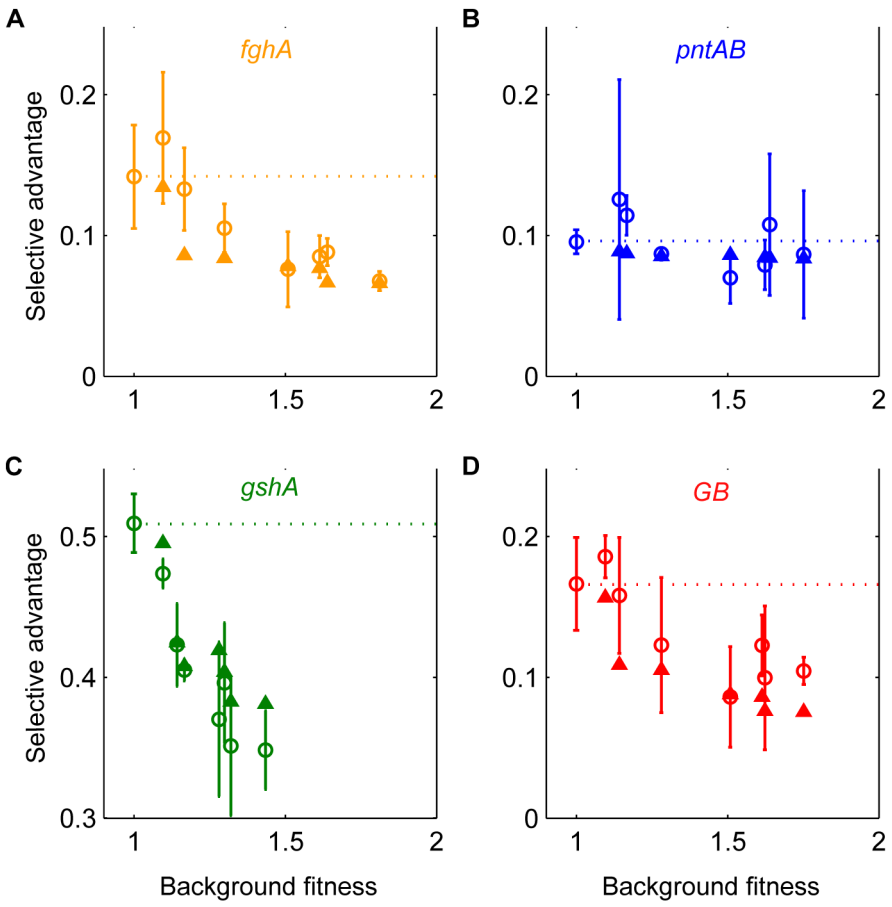

Figure 3.

Antagonistic trend of epistasis detected from the data captured by the benefit-cost model. (**A–D**) Plots of measured (open circles) and predicted (solid triangles) selective coefficients *s* for each of the four evolved alleles, respectively, versus the fitness of the background onto which the allele was introduced. Dashed lines indicate selective advantages for each allele on the ancestral background (*i.e.*, expectation for no epistasis: $W_{ij} = \lambda_i \lambda_j W_o$).