Title

You cannot tell a book by looking at the cover: Cryptic complexity in bacterial evolution

Journal

Biomicrofluidics, 8(5)

ISSN

1932-1058

Authors

Zhang, Qiucen

Bos, Julia

Tarnopolskiy, Grigory

et al.

Publication Date

2014-09-01

DOI

10.1063/1.4894410

Peer reviewed

You cannot tell a book by looking at the cover: Cryptic complexity in bacterial evolution

Qiucen Zhang,[1] Julia Bos,[2] Grigory Tarnopolskiy,[2] James C. Sturm,[3] Hyunsung Kim,[4] Nader Pourmand,[4] and Robert H. Austin[2]
[1]*Department of Physics, University of Illinois, Urbana-Champaign, Urbana, Illinois 61801, USA*
[2]*Department of Physics, Princeton University, Princeton, New Jersey 08544, USA*
[3]*Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544, USA*
[4]*Genome Sequencing Center, University of California, Santa Cruz, California 95064, USA*

(Received 30 June 2014; accepted 20 August 2014; published online 9 September 2014)

Do genetically closely related organisms under identical, but strong selection pressure converge to a common resistant genotype or will they diverge to different genomic solutions? This question gets at the heart of how rough is the fitness landscape in the local vicinity of two closely related strains under stress. We chose a Growth Advantage in Stationary Phase (GASP) *Escherichia coli* strain to address this question because the GASP strain has very similar fitness to the wild-type (WT) strain in the absence of metabolic stress but in the presence of metabolic stress continues to divide and does not enter into stationary phase. We find that under strong antibiotic selection pressure by the fluoroquinolone antibiotic ciprofloxacin in a complex ecology that the GASP strain rapidly evolves in under 20 h missense mutation in *gyrA* only 2 amino acids removed from the WT strain indicating a convergent solution, yet does not evolve the other 3 mutations of the WT strain. Further the GASP strain evolves a prophage e14 excision which completely inhibits biofilm formation in the mutant strain, revealing the hidden complexity of *E. coli* evolution to antibiotics as a function of selection pressure. We conclude that there is a cryptic roughness to fitness landscapes in the absence of stress. © *2014 AIP Publishing LLC*.
[http://dx.doi.org/10.1063/1.4894410]

We use genetically closely related strains: (1) a wild-type strain of bacteria (WT); and (2) a strain adapted to starvation stress (growth advantage in stationary phase (GASP)). Kolter and his colleagues discovered that under prolonged starvation in stationary phase of no apparent increase in cell density that a new strain of *Escherichia coli* cells evolved with the ability to grow slowly despite the stressful environment.[1,2] These mutant GASP cells have the growth advantage in stationary phase phenotype, or GASP state, which is a chain of sequential mutations under starvation stress. The GASP initial state is a *RpoS*819 allele, the *RpoS* gene encodes the σ^s sigma factor, whose role is to regulate the transcription of genes as a cell protection mechanism in response to deteriorating stress conditions such as decreasing nutrient levels, increased density, or changes in pH. Briefly, we inserted the *RpoS*819 allele, which comes from the ZK819 donor strain, through P1 transduction into each of the Green Fluorescent Protein (GFP) GFPμ2 or Red Fluorescent Protein (RFP) mRFP-1 producing wild type strain. See Ref. 3 for details. The GASP mutation which emerges in stationary phase is not a point mutation but rather a 46-bp duplication of the sequence before the stop codon of the *RpoS* gene resulting in attenuated RpoS response,[2] that is, continued growth under stress.

In order to map out the dynamics of evolution of antibiotic resistance for these two strains,[4,5] we used high spatial stress gradients in a environmental fitness landscape to accelerate the rate of evolution[6] in addressing the question of how these two closely related strains react to antibiotic stress, and metapopulations embedded within a spatial gradient to further accelerate the rate of mutation fixation. Both the WT strain and the GASP strain evolve

resistance to cipro in under 10 h in the high-stress gradient landscape, but the pattern of metapopulation occupation of the evolved strains is substantially different between the two strains, suggesting that the evolved genomes are not convergent, nor are the phenotypes convergent under stress. The WT evolved strain only occupies the perimeter of the ecology, presumably because the evolved WT strain has not yet evolved the GASP mutation of the *RpoS* allele. The GASP mutant strain shows growth across the entire device indicating that the initial genomic and phenotypic difference between the initial strains has not converged to a common solution, as seen in Fig. 1.

The differing cryptic phenotypes of the two strains of *E. coli* used here, a WT and a GASP mutants, are a strong function of the stress being imposed on the system. This can be seen at different in different ecologies: (1) open, flat ecologies such as agar gel competition experiments similar to the work of the Nelson group,[7,8] and (2) closed, complex ecologies such as we construct using microfabrication techniques. Nutrients in an open ecology agar plate are uniformly distributed at time zero, but gradually overpopulation conditions the media and the emergent different phenotypes of the two strains become evident. Fig. 2 shows the expansion of initially mixed WT and GASP population at 24 and 48 h. The WT type cells express GFP, while GASP cells express RFP. At 24 h, both the inner colony and the frontier are equally occupied by the WT and GASP cells as shown by the yellow color of the center circle, and the circular shape of the colony is maintained from inoculation, implying equal growth at the

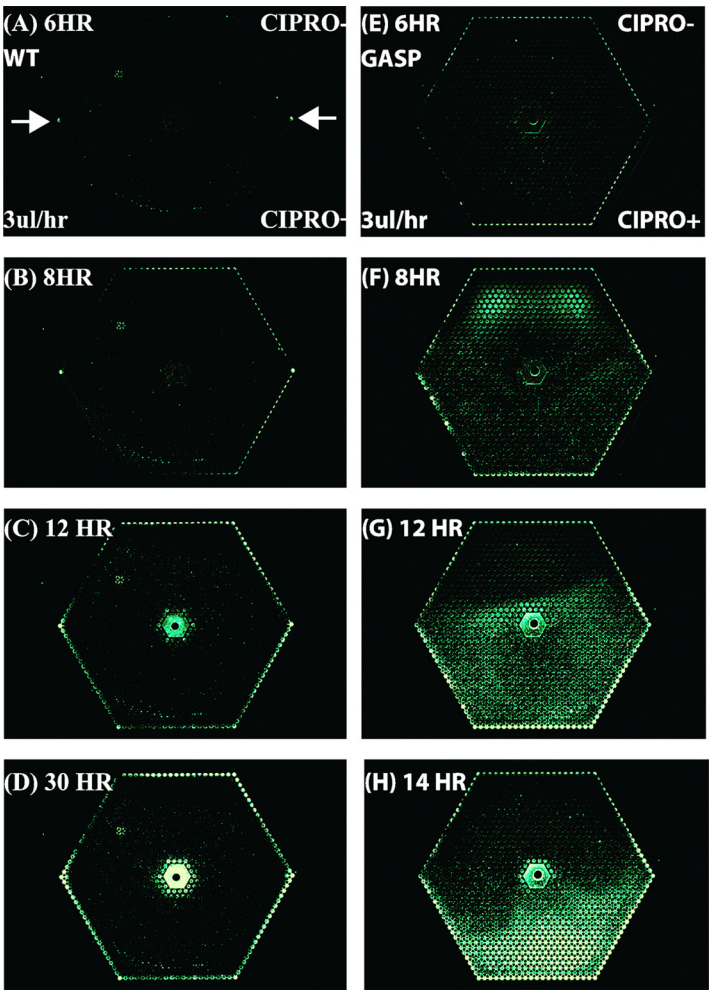

FIG. 1. (a)–(h) Progression of wild type *E. coli* and GASP vs time after inoculation: (a) and (b) +5 h after inoculation, (c) and (d) +10 h after inoculation, and (e) and (f) +15 h after inoculation. Nutrients flow along the bottom channel from right to left, while nutrients +10 μg/ml ciprofloxacin flow along the top channel from left to right. The flow rate is 30 μl/h.

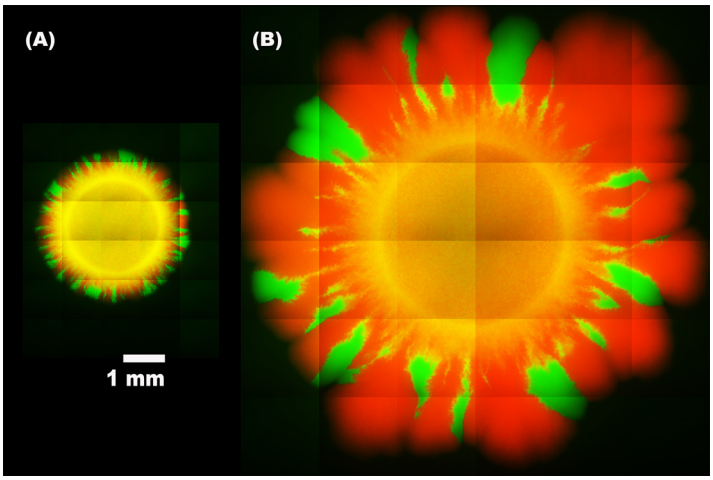

FIG. 2. (a) 24 h and (b) 48 h since inoculation. Segregation of WT (Green) and GASP (Red) population on Lysogenic Broth (LB) agar plate 48 h since inoculation. The inoculation is done by pipetting 5 μl of cell culture on to the center of the plate. The images are taken at 4× magnification and then tiled together to make the montage. GFP and RFP channels are taken at the same intensity of light source and the same exposure time (50 ms).

periphery and equal fitness of the two strains under rich nutrient conditions,[7] see Fig. 2. However, at 48 h, the morphology of the colony dramatically changes. By 48 h, the GASP dominates both the inner colony (turning to more red from yellow) and the expansion front. GASP has a emergent fitness advantage over WT in the developing nutrient-stressed environment, although the WT cells are not completely extinguished (see Fig. 2). However, in a closed boundary with large gradients in nutrients, the mixing of the different phenotypes is much more pronounced, as we have seen in other experiments.[9] Fig. 2 shows the superior fitness of the GASP strain in comparison to the WT strain under stress on an agar plate with no fixed boundaries, while Fig. 3 shows the cohabitation of GASP (red) and WT (green) under much more stringent nutrient availability conditions.

The apparent fitness homogeneity of the genomes of both the WT and GASP strains and the evolved mutants can be seen using the classical technique of counting the number of

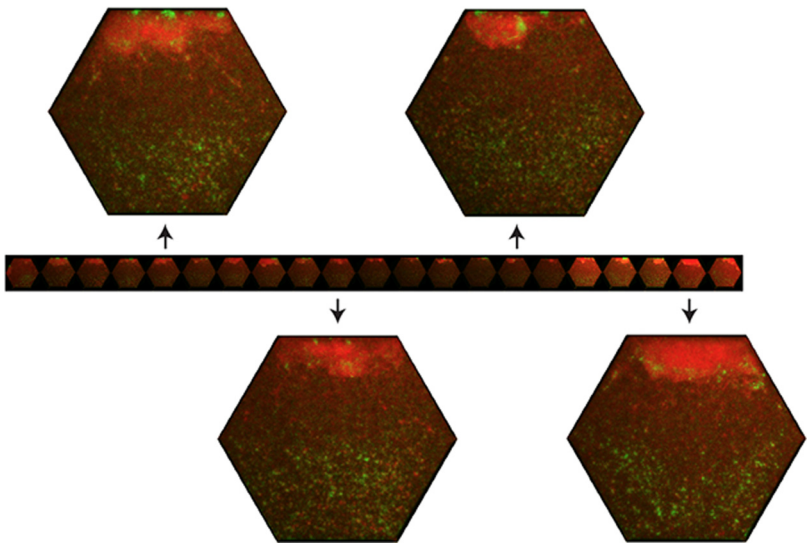

FIG. 3. One of six edges at the periphery of death galaxy chip after 24 h of inoculation. The chip is initially inoculated with a mixture of 107 wild-type and GASP cells at 1:1 ratio in the center hole. It also shows four zoom-in view of every 5 micro-environments along the edge.

colony-forming units (CFUs) as a function of antibiotic concentration, see Fig. 4. We extracted the cells from the array using replica plating at the end of the evolution experiment. WT-R and GASP-R are formally defined as the resistant strains, developed strains by evolution in the chip using, respectively, the WT and GASP *E. coli* as the initial starting population. Both the WT-R and GASP-R strains were cultured in LB 15 ml test tubes for 12 h in varying amounts of cipro, then diluted, and streaked onto agar plates. The fitness landscapes of the two strains had dramatically changed during evolution in the microecology, shown by the number of CFUs that result after exposure of both WT and GASP strains to differ concentrations of streaking on agar plates (see Fig. 4). It was necessary to evaluate fitness via CFU rather than the more typical optical density measurements because of the initial response of *E. coli* to cipro is the formation of long filaments,[10] these filaments give rise to abnormal scattering artifacts in optical density measurements. There are two interesting things to notice from the fitness curves (see Fig. 4): (1) Both the WT-R and the GASP-R mutants have evolved roughly the same resistance to cipro, approximately $\times 15$ the minimum inhibitory concentration of the wild type strains, and (2) the fitness of the WT-R and GASP-R strains is within a factor of 2 of the sensitive strains, so that the increase in resistance came at little cost in terms of fitness in the absence of the antibiotic for either strain, a subject of active concern in microbiology because it makes the resistant strains very difficult to eradicate.[11]

However, whole genome sequencing of the GASP, WT, and evolved resistant strains indicated that there were cryptic aspects to the fitness landscape of the strains. We required *de novo* sequencing because of the 46-bp duplication sequence of the GASP allele. Although *De novo* genome sequencing is considerably more difficult than consensus sequencing,[12] it allowed us to observe any larger scale insertions or deletions in addition to Single Nucleotide Polymorphisms (SNPs), which gives us a more complete picture of the evolved genome of the GASP mutant. Because ciprofloxacin inhibits *gyrA* function by sitting in the active site of the enzyme, it is perhaps not surprising to see mutated amino acid sits very close to ciprofloxacin, which could interfere the binding of ciprofloxacin. Unlike the WT strain, the mutant GASP strain did not evolve the other 3 SNPs seen in the WT resistance experiments,[13] namely, the missense A $\rightarrow$ T in base 3 933 247 in a region coding for the *rbsA* gene and the pair of missense SNPs (1 617 461: A $\rightarrow$ C and 1 617 460: C $\rightarrow$ G) found in the coding sequence for *marR*, encoding a repressor of a regulatory locus controlling multiple antibiotic resistance.

Because we did *de novo* whole gene assembly, it was possible in this experiment to look for larger scale genomic changes than the usual SNP search in the GASP mutant. We also detected the excision of the e14 prophage in 2 of the 3 resistant GASP-R strains. There are 9 cryptic prophages in the *E. coli* genome which are known to confer fitness advantage to differing varieties of stress.[14] The *e*14 prophage is 15 193 bp long and has endpoints (1 195 443 and

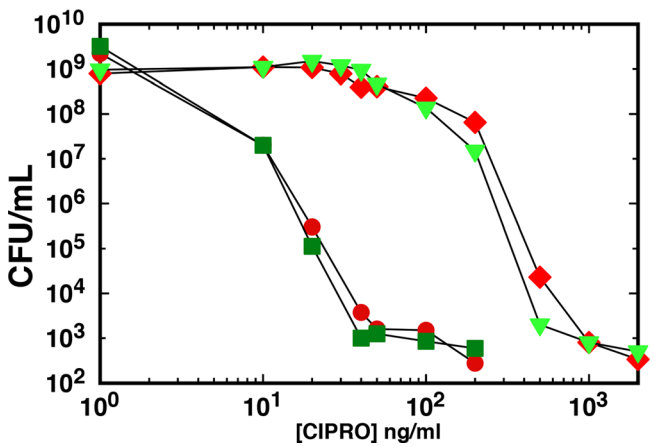

FIG. 4. Colony-forming units as a function of cipro concentration in ng/ml. Solid squares: WT, solid circles: GASP, solid diamond: WT-R, and solid triangle: GASP-R.

1 210 635) on the *E. coli* genome. Knowledge of this excision of the e14 prophage in the GASP mutant allowed us to go back and look at coverage maps of our previous evolved WT cipro resistant strain. Coverage maps showed that the *e*14 prophage was not excised from the WT strain under evolution of cipro resistance. Thus, at least for this set of experiments, the selective excision of the e14 prophage is unique to the GASP-R strain.

We would assume that bacteria have taken advantage of the lysogenic insertion of phage DNA into their genome by using the expression of cryptic phage genes to the bacteria's, not the phage's advantage.[14] The phenotype changes that the e14 excision delivers can be determined by culturing the GASP-R$^-$ strain lacking the e14 prophage and the GASP-R$^+$ which retained e14. There is no apparent difference in resistance to cipro between GASP-R$^-$ and GASP-R$^+$, probably not surprising because they both survived natural selection in our complex ecology. However, there is a profound difference in the ability of the GASP-R$^-$ and GASP-R$^+$ strains to form biofilms. Biofilm formation is a standard way for bacteria to collectively respond to antibiotic stress, and it is known that the e14 excision influences biofilm formation.[14] We assayed for biofilm formation in culture tubes by adding the biofilm stain crystal violet[15] to bacteria cultured for 22 h. Rather astonishingly, the GASP-R$^-$ retains the same ability to form biofilms as the wild-R$^+$ strain (there is no wild-R$^-$ strain seen in our experiments), while the GASP-R$^+$ fails to form biofilms under any cipro concentration, as shown in Fig. 5.

Biofilm formation is not related to the CFU numbers of Fig. 4 and so can be viewed as a cryptic mutation in standard fitness assays, hidden because of the lack of application of stress. However, it may not be cryptic with the microfabricated ecologies used as evolution accelerators used here, because of the profound importance of motility in driving the evolution process. The presence of the e14 excision can be viewed as a "hidden bifurcation" in the fitness landscape, where we imply that underlying the fitness landscape exists a complex set of forcings on the surface by genes that influence many points of the landscape (pleiotropy), as Waddington proposed in his monograph, "The Strategy of the Genes."[16] This then implies that by "fitness landscape," we actually refer to a duality: (1) The fitness landscape generated by the pleiotropy of the genes in a genome in a given environment, and (2) the fitness landscape generated by a genome as a function of changing the multiple parameters of an environment. Mathematically, we can say that simplest level that the local fitness S_{ij} is a second-rank tensor due to the interactions of two genes i and j, of course higher numbers of interacting genes require the construction of higher rank tensors. In this viewpoint, responses of cells to stress which involve multiple interacting genes can be diagonalized into eigenvectors of fitness response, and the net fitness $\langle S \rangle$ can be obtained as the trace Tr S. Because the eigenvector contains the input of several genes, it is then possible that in the process of obtaining a given response to a stress, such as antibiotic resistance, further cryptic phenotypes can be carried along with the desired one. Note that since we did not do single cell sequencing that the fitness matrix elements for a given gene are already averages over many elements, so then the eigen vectors can be viewed as collective values for a community of bacteria.

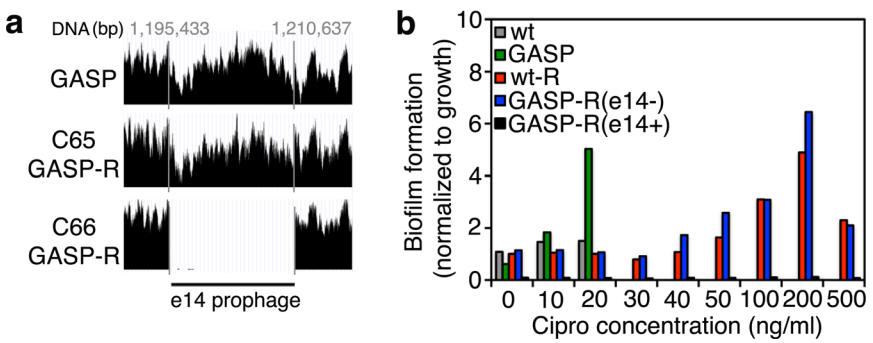

FIG. 5. (a) Coverage map showing e14 prophage deletion in ciprofloxacin resistant GASP. (b) Relative biofilm formation in the different strains. The axis changes scale at 100 ng/ml to accommodate increased resistance.

The combination of utilization of high-gradient fitness landscapes and metapopulations has allowed us to reveal rapidly the multi-path genomic movement of genomes, the diversity of solutions that the plastic genome of bacteria[17] is capable of achieving, given the proper conditions. The genomic changes that the GASP strain evolved are a mixture of convergence and divergence from the WT. There is unsuspected depth to the evolved genome of the bacteria, in that biofilm formation in the evolved strains is now heterogeneous, implying that the bacteria are anticipating necessary evolability in the future.[18] In addition to this surprise, we also note that because the fitness of all the strains that have evolved in our landscape is basically identical, it is probably very difficult for our evolved strains to ever revert back to the WT strain.

This work was performed in part at the Cornell NanoScale Facility supported by the National Science Foundation (Grant No. ECS 03-35765). The research described was supported by Award No. U54CA143803 from the U.S. National Cancer Institute and in part by the National Science Foundation under Grant No. PHYS-1066293 of the Aspen Center for Physics. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute, the National Institutes of Health or the National Science Foundation. We thank James Shapiro for extensive discussions.

[1]E. R. Zinser and R. Kolter, J. Bacteriol. **182**, 4361 (2000).
[2]R. Madan, R. Kolter, and S. Mahadevan, J. Bacteriol. **187**, 7912 (2005).
[3]See supplementary material at http://dx.doi.org/10.1063/1.4894410 for a description of experimental protocols.
[4]S. Wright, in *Proceedings of the 6th International Congress of Genetics* (1932), Vol. 1, p. 356.
[5]Q. C. Zhang and R. H. Austin, Annu. Rev. Condens. Matter Phys. **3**, 363 (2012).
[6]R. Hermsen, J. B. Deris, and T. Hwa, Proc. Natl. Acad. Sci. U.S.A. **109**, 10775 (2012).
[7]O. Hallatschek and D. R. Nelson, Theor. Popul. Biol. **73**, 158 (2008).
[8]K. S. Korolev, M. J. I. Muller, N. Karahan, A. W. Murray, O. Hallatschek, and D. R. Nelson, Phys. Biol. **9**, 026008 (2012).
[9]J. E. Keymer, P. Galajda, G. Lambert, D. Liao, and R. H. Austin, Proc. Natl. Acad. Sci. U.S.A. **105**, 20269 (2008).
[10]S. S. Justice, D. A. Hunstad, P. C. Seed, and S. J. Hultgren, Proc. Natl. Acad. Sci. U.S.A. **103**, 19884 (2006).
[11]D. C. Angst and A. R. Hall, BMC Evol. Biol. **13**, 163 (2013).
[12]C. D. Herring, A. Raghunathan, C. Honisch, T. Patel, M. K. Applebee, A. R. Joyce, T. J. Albert, F. R. Blattner, D. van den Boom, C. R. Cantor, and B. O. Palsson, Nat. Genet. **38**, 1406 (2006).
[13]Q. C. Zhang, G. Lambert, D. Liao, H. Kim, K. Robin, C. K. Tung, N. Pourmand, and R. H. Austin, Science **333**, 1764 (2011).
[14]X. X. Wang, Y. Kim, Q. Ma, S. H. Hong, K. Pokusaeva, J. M. Sturino, and T. K. Wood, Nat. Commun. **1**, 147 (2010).
[15]A. T. A. Jenkins, A. Buckling, M. McGhee, and R. H. Constant, J. R. Soc., Interface **2**, 255 (2005).
[16]C. H. Waddington, *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology* (Allen & Unwin, 1957).
[17]J. A. Shapiro, Phys. Life Rev. **10**, 287 (2013).
[18]J. Sun and M. W. Deem, Phys. Rev. Lett. **99**, 228107 (2007).