Visualizing evolution in real-time method for strain engineering

Luis H. Reyes, James Winkler and Katy C. Kao*

Department of Chemical Engineering, Texas A&M University, College Station, TX, USA

INTRODUCTION

The majority of industrially relevant phenotypes in microbial systems involve multiple loci and mechanisms. The identities of these genetic determinants are generally not known, making the rational engineering of strains for these complex phenotypes challenging. Classical strain engineering for these traits generally involve several rounds of random mutagenesis followed by selection. However, with successive rounds of induced mutagenesis, mutations that are deleterious or have negative epistatic effects tend to accumulate by hitchhiking with beneficial alleles. If the desired trait can be coupled with growth, in vitro adaptive evolution can be used to improve the desired phenotype. This process is accomplished by applying a selective pressure so that beneficial mutants (mutants with increased fitness) can be obtained through the process of natural selection. The identities of the mutations residing in adaptive mutants obtained through natural selection or mutagenesis and their subsequent effects on cellular processes must be leveraged for further rational engineering. With advances in genomic tools, the genes and mechanisms involved can now be identified using combinations of whole-genome re-sequencing (Comas et al., 2012; Toprak et al., 2012), transcriptomics (Fitzgerald and Maser, 2001; Paulsen et al., 2001), proteomics (Callister et al., 2008; Bouldis et al., 2010), and metabolomics (Ding et al., 2010; Goodarzi et al., 2010) studies.

In vitro adaptive evolution has been used extensively for the engineering of microbial system for both tolerances to inhibitors (Minny et al., 2011) and for enhanced product formation (Hu and Wood, 2010). The adaptive landscape, also known as fitness landscape, is used to describe the collection of relative fitness effects of each genotype under a specific condition. Detailed molecular characterization of adaptive mutants isolated from in vitro adaptive evolution experiments provides insights into the adaptive landscape for the phenotype of interest.

Characterization of the adaptive landscape will significantly enhance our knowledge on the important parameters underlying complex phenotypes needed for the rational engineering of strains.

In adaptive evolution, clones are typically isolated from the evolving population after an arbitrarily elapsed time or at the end of the experiment. However, since the evolving population is heterogeneous, interclonal competition (clonal interference; Shaver et al., 2002; Kao and Sherlock, 2008) may lead to the extinction of beneficial mutants. Depending on the population structure during the course of evolution, the random isolation of adaptive mutants may fail to identify some adaptive mutations that arise during the course of the evolution. This review will (1) discuss factors that influence population structure and the impact of complex population dynamics on evolutionary engineering and (2) describe a novel evolutionary engineering method called visualizing evolution in real-time (VERT), that was recently developed to help address some of these limitations in traditional evolutionary engineering approach.

ADAPTIVE LANDSCAPE

The idea of an adaptive landscape was first introduced as "surfaces of selective value" by Wright in 1931 (Wright, 1931, 1982, 1988). The adaptive landscape is a multi-dimensional surface representation of the biological fitness of an organism in a particular environment. In an adaptive landscape map for a specific condition, each genotype is correlated with a fitness value (see Figure 1 for an illustration). The resulting landscape can be flat with a single optimum where the evolving population is required to acquire a specific set of mutations, or can be rugged where the accessible optima will depend on the starting point within the landscape. It has been demonstrated that bacteria encounter both types of landscapes in evolution experiments (Orr, 2005; Winnreich et al., 2006;
Reyes et al. Visualizing evolution in real-time

FIGURE 1: Simplified adaptive landscape for two alleles (for one background genotype in one condition). The figure depicts fitness values for beneficial (positive relative fitness values) and deleterious (negative relative fitness values) combinations of alleles.

Gresham et al., 2008). Natural selection usually drives a population to the closest local optimum, but not necessarily the global optimum. Evolving populations tend to be trapped in suboptimal solutions (Lenski et al., 1998) in asexual systems. Thus to reach the global optimum, processes that allow for large “jumps” in the adaptive landscape, such as recombination and horizontal DNA transfer, are necessary to reach new regions of the adaptive landscape in a semi-rational manner. Recombination allows the combination of beneficial mutations with positive synergy and the removal of deleterious mutations acquired in the evolutionary process while horizontal gene transfer allows the acquisition of new functions.

THEORIES GOVERNING POPULATION STRUCTURE DURING ASEXUAL EVOLUTION

Numerous theories have been proposed for the population structure in in vitro adaptive evolution experiments. Several factors, including the selective pressure, size of the population, rate of mutations, frequency of beneficial mutations, and relative fitness of beneficial mutants, are involved in determining the population structure during evolution. In the simplest case, a well-adapted mutant rises in the population, and due to its increased fitness compared to background, the genotype will expand and eventually replace the parental population. This population structure is applicable to situations where the evolution is mutation-limited, the population size is small, and the time between the establishments of successive mutations is much larger than the time it takes for a beneficial mutant to fix in the population (strong positive selection). This theory, called clonal replacement (also called succession-fixation regime or strong-selection weak-mutation regime), implies that only one mutation can become fixed at a time, leading to successive complete selective sweeps (depicted in Figure 2A). The resulting population can be assumed to be homogeneous except during the periods when the beneficial mutant is sweeping through (Desai et al., 2007). However,
FACTORS INFLUENCING POPULATION DYNAMICS

As mentioned above, factors such as mutation rate, relative fitness advantage, population size, and rate of beneficial mutations are important in shaping the population dynamics during evolution. We will briefly discuss each of these factors and how they impact adaptive evolution in different experimental systems. Since the evolution dynamics is dependent on the mutation rate, one would assume higher mutation rate to be advantageous for speeding up evolution by generating mutational diversity. However, an increase in mutation rates does not necessarily accelerate the pace of adaptation (Arjan et al., 1999). While a low mutation rate would result in a slow discovery of beneficial mutations, prolonged exposure to high mutation rate (such as the use of a mutator strain) increases the occurrence and accumulation of deleterious mutations as well as the hitchhiking of apparent “silent” mutations during the course of evolution, increasing the genetic load (Elena and Lenski, 1997; Snigowski et al., 1997; Arjan et al., 1999; Vulic et al., 1999; Lenski et al., 2003) and the fitness advantage conferred by the mutator strains is most likely a result of overcoming a mutation-limited bottleneck during the evolution. Mutagens are often used to increase genetic diversity in evolution experiments. However, since it is not convenient to periodically mutagenize the evolving population, a controllable mutator system can be used, where the expression of mutator alleles can be induced only when needed (Selifonova et al., 2001).

The time it takes a beneficial mutation to become the majority in the population is called the fixation time and is an important factor in determining the population dynamics during evolution. This fixation time depends mainly on two factors, genetic drift and the fitness advantage of the beneficial mutation in comparison with the background, and is inversely proportional to the relative fitness advantage of the beneficial mutant (Lenski et al., 1991). A beneficial mutation with a 10% relative fitness advantage will become the majority of the population after approximately 250 generation in serial batch transfer experiments (Elena and Lenski, 2003) and 100 generations in continuous culture experiments (Gresham et al., 2008). Genetic drift is defined as the probability that a beneficial mutation survives extinction (Joanna, 2011). In *in vitro* adaptive evolution experiments, the main source of drift is genetic bottleneck due to random sampling. This phenomenon takes place when a significant amount of the population suddenly vanishes, as occurs when a fresh batch culture is inoculated from an overnight culture. The survival probability of an allele carrying a beneficial mutation that arose in the culture will depend on its proportion in the culture at the time of transfer and the amount of inoculum transferred; therefore there is a chance that it could be completely lost due to the stochasticity of sampling. In evolution experiments using serial batch transfers, genetic bottlenecks between transfers affect heterogeneity by transferring a small fraction of the population. A reduction of the effect of genetic bottleneck could be achieved by using continuous culture systems such as chemostats or turbidostats (Conrad et al., 2013), where a much smaller genetic bottleneck is present.

VISUALIZING EVOLUTION IN REAL-TIME

As stated above, the population sizes in most *in vitro* evolution experiments are large enough to result in heterogeneous populations due to the effects of clonal interference and multiple mutations. Thus, adaptive evolution experiments can significantly benefit from a more systematic isolation of adaptive mutants and ramping-up schedules for selective pressures. The VERT system was developed (Kao and Sherlock, 2008; Huang et al., 2011) to address these limitations in traditional adaptive evolution experiments. The basis for VERT is the use of isogenic, but differentially labeled (typically with fluorescent proteins) strains to seed the initial evolving population. As a beneficial mutant arises and expands in the population, the colored subpopulation that it belongs is expected to increase in proportion. Using fluorescent activated cell sorting (FACS), the relative proportions of each of the colored subpopulations at each point in time can be measured. Each sustained expansion in the proportion of a colored subpopulation is called an “adaptive event.” Thus, the tracking of the different colored subpopulations can serve as a tool for determining when a fitter mutant arises in the population. The relative subpopulation frequency data collected throughout the course of adaptive evolution represent the history of the population. The observed increase in the relative proportion of a colored subpopulation from consecutive data points is assumed to be the result of the expansion of an adaptive mutant. Therefore, adaptive mutants can be isolated from samples based on the observed expansions and contractions, by sorting out the colored subpopulation that is expected to contain the adaptive mutant of interest. Since experimental data can suffer from noise,
the identification of adaptive events may be challenging. Visual inspection of the data to identify adaptive events is a crude, but relatively effective method (Kao and Sherlock, 2008; Huang et al., 2011). However, since small changes in relative frequencies may be difficult to distinguish from noise, computational methods will provide less biased annotation of adaptive events; our group recently developed a supervised learning method for analysis of VERT data (Winkler and Kao, 2012).

The basic feature of the VERT system, the number of labeled subpopulations, is the aspect that can most readily be manipulated directly by the experimentalist, but is somewhat restricted by the available equipment and properties of the labels themselves. The number of fluorescent markers used represents distinct subpopulations that can be visualized during the course of an evolution experiment. VERT labels must have distinguishable emission spectra and preferably have no significant fitness effect in the condition of interest. Widely used fluorescent proteins such as GFP, YFP, and RFP can be detected on most FACS machines and usually have little effect on the physiology of their host strains. At a minimum, two labeled subpopulations are trivially required to observe population dynamics. Three subpopulations, employing RFP, GFP, and YFP labeled strains, have been used successfully (Kao and Sherlock, 2008; Huang et al., 2011) in fungal systems. Additional subpopulations can be included if suitable equipment is available. Simulated evolution may prove a useful tool for unraveling the connection between adaptive event discovery and initial population diversity.

Visualizing evolution in real-time-based in vitro adaptive evolution experiments can be used in either serial batch transfer or continuous culture systems. Provided that the different fluorescently marked strains show no significant fitness bias, then equal proportions of each strain may be used to seed the population for evolution. Samples are then withdrawn and quantified using FACS every few generations to track the population dynamics. It is typically assumed that the adaptive mutant will expand until a fitter mutant arises in another subpopulation and expands sufficiently to impede its’ expansion. It is further assumed that the generation at which the expanding subpopulation has reached a maximum proportion will contain the largest fraction of the adaptive mutant responsible for the expansion, simplifying the isolation of the adaptive mutant considerably.

In traditional adaptive evolution experiments, selective pressure is generally ramped-up at arbitrarily chosen time intervals. An alternative to this approach, based on using a feedback controller to maintain selective pressure so that the overall population growth rate approaches a user-defined set point, was recently introduced by Toprak et al. (2012). Since the use of VERT allows the users to readily identify when adaptive events occur, it can be used to design a more systematic ramp-up schedule. For example, an increase in selective pressure could be applied when a minimum of 2 adaptive events are observed. The optimal frequency of ramp-up as a function of observed number of adaptive events may differ depending on the adaptive landscape for the phenotype of interest and needs to be investigated.

The isolated adaptive mutants can be further characterized to elucidate the molecular mechanisms of resistance in the selective pressure of interest. Whole-genome re-sequencing, transcriptomics, proteomics, and metabolomics analyses can be used to elucidate the evolutionary trajectories during the process of adaptive evolution. The availability and cost of whole-genome re-sequencing has improved significantly, but in most cases is still more expensive than transcriptome analysis using DNA microarrays. VERT tracks the individual subpopulations, making it easier to distinguish whether genome-wide perturbations observed in the transcriptional regulation found in different isolates arose independently or transitively without whole-genome re-sequencing data (if the isolates come from different colored subpopulations). Since not all the observed perturbations are involved in the complex phenotype of interest, common perturbations observed in independent lineages provide a level of confidence for their involvement. The potential adaptive mechanisms identified can serve as targets for further strain engineering.

The original development of VERT used the yeast Saccharomyces cerevisiae evolving under glucose-limited conditions; a three-colored VERT system was used to seed eight parallel populations (Kao and Sherlock, 2008). The VERT data from one of the populations is shown in Figure 3; generations and subpopulations from which adaptive mutants were isolated are indicated. Detailed genotypic and transcriptomics analyses of the isolated adaptive mutants showed convergence in the perturbation of the protein kinase A regulatory network in independent lineages (Kao and Sherlock, 2008). Subsequent development and application of a two-colored VERT system in E. coli for n-butanol tolerance revealed previously undiscovered resistance mechanisms (Reyes et al., 2012).
CONCLUSION

Understanding the adaptive landscape for the phenotypes of interest in molecular engineering tools is critical. The use of evolutionary engineering has been used extensively to improve strains for complex phenotypes where there is limited knowledge on the associated genetic determinants. Advances in molecular biology tools in recent years have significantly improved our ability to obtain valuable knowledge in the molecular mechanisms involved in the desire phenotypes in isolated adaptive mutants from in vitro evolution experiments. VERT is a recently developed tool for engineering that can help to provide a rough population structure for the evolving population, allowing the systematic isolation of adaptive mutants and ramp-up of selective pressure. Combined with advanced molecular tools, use of VERT in evolutionary engineering can help to gain additional insight regarding the adaptive landscape for complex phenotypes.

ACKNOWLEDGMENTS

The authors would like to thank NSF MCB-1054276 (Kayt Cao), and the Graduate Research Fellowship program (James Winkler) for partial financial support.

REFERENCES

Arai, J. A., Vanessa, M., Zeyl, C. W., Gerrish, P. J., Blanchard, J. L., and Lenski, R. E. (1999). Diminishing returns from mutation supply rate in asexual populations. Science 283, 404–406.

Bachtrog, D., and Gorio, I. (2004). Adaptive evolution of asexual populations under Muller’s ratchet. Evolution 58, 1403–1413.

Barrett, J. E., Yu, D. S., Voyn, S. H., and what, T. K., Schnei- der, D., Lenski, R. E., and Kim, J. F. (2009). Genome evolution and adaptation in a long-term experiment with Escherichia coli. Nature 461, 1243–1247.

Bosek, J., Trout, M., Lands, B. C., Dickman, B., Levy, E. D., Solitary, T. S., and Desjardins, M. (2010). Molecular characterization of the evolution of phagosomes. Mol. Syst. Biol. 6, 423.

Caldwell, J., McCann, L., Burns, J. E., Monnors, M. E., Abrams, K. E., Smith, R. D., Adkins, J. N., and Lipton, M. S. (2008). Comparative central proteome analysis of the core genome concept. PLoS ONE 3, e7942. doi: 10.1371/journal.pone.0007942.

Campos, P. R., and de Oliveira, V. M. (2004). Maternal effects on the clonal interference phenomenon. Evolution 58, 932–937.

Camus, I., Berroil, S., Rostori, A., Rose, G., Malla, B., Kato-Maeda, M., Galle- gos, J., Naiman, S., and Gagnon, S. (2012). Whole-genome sequencing of 100 asexual clones identifies compensatory mutations in RNA polymerase genes. Nat. Genet. 44, 106–110.

Casadevall, M., Lewis, N. E., and Palsson, B. O. (2011). Microbial laboratory evolution in the era of genome-wide science. Annu. Rev. Microbiol. 75, 599–629.

Chiew, M. M., and Murray, A. W. (2005). The speed of evolution and maintenance of variance in asexual populations. Curr. Biol. 17, 385–394.

Drum, M. E., Zhou, X., and Yuan, Y. J. (2010). Metabolome profiling reveals adaptive evolution of Saccharomyces cerevisiae during repeated yeast fermentations. Metabolomics 6, 42–55.

Hana, S. F., and Lenski, R. E. (1997). Long-term experimental evolution in Escherichia coli. VII. Mechanisms maintaining genetic variability within populations. Evolution 51, 1056–1067.

Hana, S. F., and Lenski, R. E. (2003). Evolution experiments with microbial genomics: the dynamics and genetic bases of adaptation. Nat. Rev. Genet. 4, 457–469.

Firginal, J. R., and Musser, J. M. (2001). Evolutionary genomics of pathogenic bacteria. Trends Microbiol. 9, 255–260.

Gerrish, P. (2001). The rhythm of microbial adaptation. Nature 413, 299–302.

Gerrish, P. J., and Lenski, R. E. (1998). The fate of competing beneficial mutations in an asexual population. Genetica 102–103, 127–144.

Goosdori, H., Bennett, B. D., Amini, S., Roses, M. L., Hotten, A. K., Rehav, T., and Tanous, S. (2010). Regulatory and metabolic rewiring during laboratory evolution of ethanol tolerance in E. coli. Mol. Syst. Biol. 6, 178.

Gonzalez, M. M., Tacker, C. M., Jiang, H. T., Pai, D. A., Ward, A., Doebie, C. G., Bost, D., and Dunham, M. J. (2008). The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. PLoS Genet. 4, e1000305. doi: 10.1371/journal.pgen.1000305.

Hu, H., and Wodick, T. E. (2000). An evolved E. coli strain for producing hydrogen and ethanol from glucose. Biotechnol. Biofuels Res. 101, 1055–1058.

Huang, M., McCollum, M., Berman, J., and Kao, K. (2011). Evolutionary dynamics of Cardioac acid ubiquinone in vitro evolution. Environ. Cell 10, 1413–1421.

Joanna, M. (2011). Genetic drift. Curr. Biol. 21, R857–R858.

Kao, K. C., and Sherlock, G. (2008). Molecular characterization of clonal interference during adaptive evolution in asexual populations of Saccharomyces cerevisiae. Nat. Genet. 40, 1499–1504.

Kim, Y., and Stephan, W. (2001). Selective sweeps in the presence of interference among partially linked loci. Genetics 164, 389–396.

Kretz, D. J., and Sherlock, G. (2011). Reciprocal sign opinion between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. PLoS Genet. 7, e1002056. doi: 10.1371/journal.pgen.1002056.

Lenski, R. E., Mengold, J. A., Stragieski, P. D., Travassos, M. A., Vann, B., Gerrish, P. J., and Schmidt, T. M. (1998). Evolution of competitive fitness in experimental populations of E. coli: what makes one genotype a better competitor than another? Antonie van Leeuwenhoek 73, 55–67.

Lenski, R. E., Rose, M. R., Simpson, S. C., and Tadler, S. C. (1991). Long-term experimental evolution in Escherichia coli. I. Adaptation and diversity during 10,000 generations. Am. Nat. 138, 1153–1154.

Lenski, R. E., Winkworth, C. L., and Rife, M. A. (2003). Rates of DNA sequence evolution in experimental populations of Escherichia coli during 20,000 generations. J. Mol. Evol. 56, 496–498.

Mierza, J., Lesniakowska, A., Lin, F., Chen, Y., Zarrat, T., Veloso, A., Xue, X., McConnell, C., Ward, E., Schwartz, D., Bouillard, J. M., Gao, Y., Galari, E., and Lin, X. (2011). Evolution combined with genomic study elucidates genetic bases of isobutanol tolerance in Escherichia coli. Microb. Cell Fact. 10, 18.

Orr, H. A. (2000). The rate of adaptation in asexual populations. Science 287, 509–511.

Paabo, S., Chen, J., Nelson, K. E., and Isaac, M. H. (2001). Comparative genomics of microbial drug efflux systems. J. Microbiol. Biotechnol. 13, 145–150.

Page, L. H., Almario, M. P., Win- delin, J., Ormone, M. M., and Kao, K. C. (2012). Visualizing evolution in real time to determine the molecular mechanisms of r-enthalol tolerance in Escherichia coli. Microb. Eng. (in press).

Polifroni, O., Villes, E., and Schellen- berger, V. (2005). Rapid evolution of novel traits in microorganisms. Appl. Environ. Microbiol. 71, 5645–5649.

Shone, A. C., Dombrowski, P. G., Sweeney, J. Y., Tres, T., Zappaki, B. M., and Stragieski, P. D. (2002). Fitness evolution and the rise of mutants advantageous to genomic Escherichia coli populations. Genetics 162, 557–566.

Stragieski, P. D., Gerrish, P. J., and Lenski, R. E. (1997). Evolution of high mutation rates in experimental populations of E. coli. Nature 387, 705–707.

Terpeki, E., Veres, A., Michal, J. B., Chait, B., Hartl, D. L., and Kishony, R. (2012). Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nat. Genet. 44, 103–105.

Volk, M., Lenski, R. E., and Radman, M. (2005). Mutations, recombination, and inescapable selection of bacteria in the laboratory. Proc. Natl. Acad. Sci. U.S.A. 102, 7348–7351.

Weinreich, D. M., Delaney, N. F., DePristo, M. A., and Hartl, D. L. (2006). Darwinian evolution can follow only very few mutational pathways to fitter proteins. Science 313, 112–115.

Wilke, C. O. (2004). The speed of adaptation in large asexual populations. Genetics 167, 2043–2053.

Winkler, J., and Kao, K. C. (2012). Computational identification of adaptive mutants using the VERT system. J. Bacteriol. 194, 57–63.

Weight, S. (1951). Evolution in Microbial populations. Genetics 36, 907–915.

www.frontierin.org
Wright, S. (1982). Character change, speciation, and the higher taxa. Evolution 36, 427–445.
Wright, S. (1988). Surfaces of selective value revisited. Am. Nat. 131, 115–123.
Yedid, G., and Bell, G. (2001). Microevolution in an electronic microcosm. Am. Nat. 157, 465–487.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 March 2012; accepted: 14 May 2012; published online: 29 May 2012.

Citation: Reyes LH, Winkler J and Kao KC (2012) Visualizing evolution in real-time method for strain engineering. Front. Microbiol. 3:198. doi: 10.3389/fmicb.2012.00198
This article was submitted to Frontiers in Microbiotechnology, Ecotoxicology and Bioremediation, a specialty of Frontiers in Microbiology.

Copyright © 2012 Reyes, Winkler and Kao. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.