doi:10.1038/nature14207

EVOLUTION

Fitness tracking for adapting populations

A method for tracking the descendants of hundreds of thousands of yeast cells in an evolving population reveals that thousands of individuals contribute to early increases in population-wide fitness.

DAVID GRESHAM

Positive selection for genetic variants that benefit an organism in a particular environment, a process called adaptive evolution, affects all species. As such, knowing how frequently beneficial mutations occur, and quantifying the selective advantage they confer — their fitness — has been a long-standing goal for evolutionary biology[1]. In a paper published on *Nature*'s website today, Levy *et al.*[2] describe a method for tracking individual genetic variants in an evolving population, and measuring their fitness and fate as the population adapts to the environment.

Individuals descended from a common progenitor are said to be of the same genetic lineage. Levy and colleagues tracked individual lineages in yeast (*Saccharomyces cerevisiae*) with extremely high resolution by introducing hundreds of thousands of unique, random DNA sequences into individual yeast cells that have otherwise-identical genomes. These sequences, called barcodes, have no impact on the cell, but can be used to distinguish between different individuals by means of DNA sequencing. Individuals that have the same barcode are part of the same lineage, allowing estimation of how many cells in the population are descended from a common ancestor.

After barcoding the yeast cells, the authors studied the population as it underwent adaptive evolution over many generations in a simple environment. In the evolving population, each daughter cell is born through cell division and so is a clone of its mother. Thus, sexual reproduction plays no part in the population's evolutionary dynamics. Although all cells in the population start out with identical genomes (apart from the barcodes), genetic diversity is introduced by random mutations that arise spontaneously when DNA is replicated during cell division. If a mutation is beneficial in the environment, allowing the cell and its descendants to proliferate more rapidly, that lineage will begin to increase in relative abundance in the population. By sequencing the cells' molecular barcodes at different time points throughout the experiment, beneficial lineages can be identified.

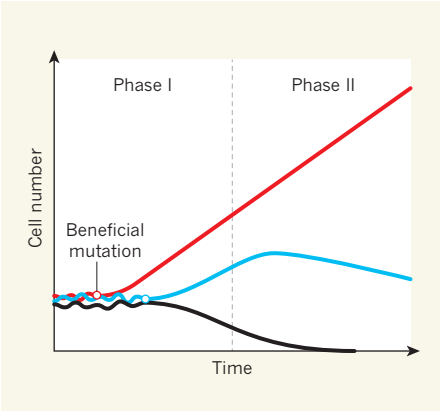

Figure 1 | Get fit or die trying. Levy *et al.*[2] labelled hundreds of thousands of individual yeast cells, and tracked the population as it evolved. Each lineage is initially present in approximately equal numbers. In the first, predictable phase of evolution, thousands of lineages that acquire beneficial mutations (blue, red) expand, increasing the fitness of the population and leading to the decline of lineages that did not acquire beneficial mutations (black). In a second, less predictable phase of evolution, even lineages with beneficial mutations can decline (blue), as those containing mutations that confer an exceptionally high degree of fitness, and that arose early enough, continue to expand (red), further increasing population fitness. (Adapted from Fig. 1a of ref. 2.)

Levy and colleagues used their high-resolution lineage-tracking technique to quantify the fitness of each beneficial lineage, and to determine when the corresponding mutation occurred in the population's history. They found that, in evolving yeast populations containing 70 million cells, about 25,000 lineages showed fitness increases of more than 2% after just over 100 generations. Many of these lineages were present at frequencies lower than 0.001%. This means that there are initially many more competing lineages containing beneficial mutations in evolving populations than previously revealed by whole-population sequencing[3–5].

The aggregate effect of these thousands of beneficial lineages is to push the population fitness higher and higher. In doing so, a process of sequential purging occurs. First, the lineages that did not acquire a beneficial mutation are removed from the population. Then, as population fitness continues to increase, even lineages that contain beneficial mutations are purged once their individual fitness is less than that of the population as a whole.

Levy and colleagues' study shows that there are two distinct phases in the adaptive evolution of a large cell population (Fig. 1). In the first phase, population fitness increases in a predictable manner. This increase is attributable to the cohort of thousands of different lineages with beneficial mutations, and depends on the size of the population and the fitness associated with each mutation. The second phase is less predictable. The ultimate 'winners' must have higher fitness than the overall population and the mutations must have been introduced early enough in the population's history to establish themselves — this phase is unpredictable because such mutations are rare.

The ability to quantify the fitness of each beneficial mutation in a population enabled Levy and co-workers to measure the range of fitnesses conferred by beneficial mutations. Theory predicts[6,7] that the distribution of fitness effects associated with new mutations has a particular mathematical shape, known as an exponential distribution. However, the authors find that this is not the case, at least not in this environment. Instead, they observe a complicated distribution of fitness effects that seems to be composed of a mixture of distributions, which may reflect beneficial mutations in different genes. The nature of the distribution of fitness effects of beneficial mutations is central to understanding and simulating adaptive evolution in future experiments. As such, the ability to empirically measure this distribution with precision provides opportunities to reconcile theory and data.

Despite the power of Levy and co-workers technique, several limitations remain. First, the method does not actually identify the beneficial mutations, a key requirement for understanding the molecular basis of adaptation[8,9]. Second, it tells us about the distribution of fitness effects for beneficial mutations, which are most relevant to the evolution of large asexual populations, but not those for neutral or deleterious mutations, which may be important in populations that are small, sexual or have high mutation rates. Last, and crucially, the method in its current form allows identification of only the earliest stages of adaptive evolution. Once a single lineage has swept to high frequency in the population, its barcode

will be abundant. Loss of barcode diversity limits the ability to detect a second beneficial mutation within these lineages, a problem that could be overcome by somehow regenerating the diversity of barcodes during the course of the experiment.

The ability to track hundreds of thousands of individual lineages in a population is an exciting tool that allows us to address many questions in adaptive evolution. Levy *et al.* performed their experiment using enormous populations, ensuring an ample supply of mutations. However, studying the dynamics of adaptation in much smaller populations would also be informative, and will probably result in less-predictable outcomes in the early stages of adaptation. Furthermore, studying adaptation in different environments and different genetic backgrounds will be crucial for assessing the generality of the results. Application of high-resolution lineage tracking in other organisms may be useful for understanding the evolutionary dynamics of antibiotic resistance in pathogens and the evolution of human tumours. The ability to observe evolution in action with high resolution is certain to reveal unanticipated features of the universal force of adaptive evolution. ■

David Gresham *is at the Center for Genomics and Systems Biology, Department of Biology, New York University, New York, New York 10003, USA.*
e-mail: dgresham@nyu.edu

1. Eyre-Walker, A. & Keightley, P. D. *Nature Rev. Genet.* **8,** 610–618 (2007).
2. Levy, S. F. *et al. Nature* http://dx.doi.org/10.1038/nature14279 (2015).
3. Lang, G. I. *et al. Nature* **500,** 571–574 (2013).
4. Hong, J. & Gresham, D. *PLoS Genet.* **10,** e1004041 (2014).
5. Kvitek, D. J. & Sherlock, G. *PLoS Genet.* **9,** e1003972 (2013).
6. Gillespie, J. H. *Evolution* **38,** 1116–1129 (1984).
7. Orr, H. A. *Genetics* **163,** 1519–1526 (2003).
8. Dean, A. M. & Thornton, J. W. *Nature Rev. Genet.* **8,** 675–688 (2007).
9. Gresham, D. & Hong, J. *FEMS Microbiol. Rev.* http://dx.doi.org/10.1111/1574-6976.12082 (2014).