PREDICTING EVOLUTION: THEORY AND EXPERIMENTAL PROSPECTS. Philip J. Gerrish[1], Paul D. Sniegowski[2], Kathleen Sprouffske[3] and Tanya Singh[4]. [1]University of New Mexico, [2,4]University of Pennsylvania, [3]University of Zurich.

Introduction: Most work focused on prospective prediction of evolution is based on a fitness landscape approach: if it is known how the myriad possible genotypes map onto fitness, then paths through the landscape can be anticipated. Detailed genotype-fitness data, however, are rarely available, are very costly to obtain, and are subject to change with each change in the organism and/or its environment. Here, we describe prospective prediction of evolution based on real-time sampling of the fitness distribution of an evolving population rather than on prior knowledge of the fitness landscape [1].

Theory. In a large population, mean fitness is predicted to increase over the course of a single generation by an amount equal to the additive variance in fitness. This is Fisher's "fundamental theorem of natural selection"; in discrete time it can be written as:

$$\overline{w}(t+1) = \overline{w}(t) + \sigma_A^2(t)$$

where $\overline{w}$ is mean fitness at time t and $\sigma_A^2(t)$ is the additive variance in fitness at time t. In order to predict how mean fitness will evolve over the course of two generations, we would need to know how fitness variance is expected to change over the course of the first generation. Mathematically speaking, we would need to know the quantity $K_3(t)$ to insert into this equation:

$$\sigma_A^2(t+1) = \sigma_A^2(t) + K_3(t).$$

With this equation in hand, we may now predict how mean fitness will evolve over the course of two generations:

$$\overline{w}(t+2) = \overline{w}(t+1) + \sigma_A^2(t+1) = \left(\overline{w}(t) + \sigma_A^2(t)\right) + \left(\sigma_A^2(t) + K_3(t)\right)$$

If it were known how $K_3(t)$ evolves over the course of the first generation, i.e., if we knew $K_4(t)$ in the equation $K_3(t+1) = K_3(t) + K_4(t)$, then we could, using the same logic, predict how mean fitness would evolve over the course of three generations, and so on. Remarkably, the $K_i(t)$ turn out to be equal to central cumulants in fitness, i.e., $K_i(t) = \kappa_i(t)$; cumulants are standard, measurable statistical properties of the fitness distribution, e.g. $\kappa_2(t) = \sigma_A^2(t)$.

Prospects. In [1], we show that the flux of new mutations affecting fitness may be incorporated into the analysis of fitness evolution described above. With this approach in hand, the near-future course of fitness evolution can, in principle, be predicted despite highly complex underlying dynamics. We have demonstrated the predictive capability of this new framework in simulations (Fig. 1).
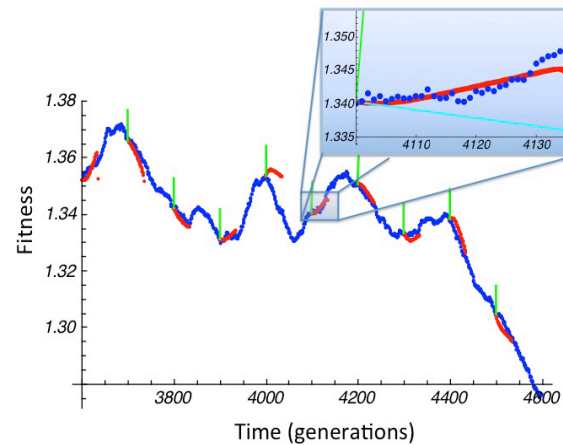

Fig. 1: Visual assessment of the theory's power to predict fitness evolution in a simulated population. The blue curve plots mean fitness of the population over time. Red lines plot predicted mean-fitness trajectories. These predictions are each based on two sets of fitness measurements, each of sample size 100: one set taken 100 generations prior to the start of the red line, and one set taken right at the start of the red line. Green lines indicate predictions of Fisher's theorem, which does not include mutation. The supply of beneficial mutations was limited in these simulations; the downward trend occurs as the beneficial mutations are used up and fitness is degraded by deleterious mutation.

Our framework allows prediction of evolution in real populations based on sampling of the population distribution of fitness at multiple time points. Moreover, we have discovered in recent theoretical work that such sampling of real populations also provides a novel and powerful means of characterizing the distribution of fitness effects (DFE) of new mutations in a population and characterizing its evolution in real time. We discuss ongoing experimental work testing such possibilities in experimental *E. coli* populations.

References:
[1] Gerrish P. J. and Sniegowski P. D. (2012) *J. Roy. Soc. Interface, 9,* 2268–2278.