CHAPTER NINE

Biological fitness landscapes by deep mutational scanning

Jacob D. Mehlhoff and Marc Ostermeier*
Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD, United States
*Corresponding author: e-mail address: oster@jhu.edu

Contents

Abstract

Knowledge of the distribution of fitness effects (DFE) of mutations is critical to the understanding of protein evolution. Here, we describe methods for large-scale, systematic measurements of the DFE using growth competition and deep mutational scanning. We discuss techniques for producing comprehensive libraries of gene variants as well as provide necessary considerations for designing these experiments. Using these methods, we have constructed libraries containing over 18,000 variants, measured fitness effects of these mutations by deep mutational scanning, and verified the presence of fitness effects in individual variants. Our methods provide a high-throughput protocol for measuring biological fitness effects of mutations and the dependence of fitness effects on the environment.

Methods in Enzymology, Volume 643
ISSN 0076-6879
https://doi.org/10.1016/bs.mie.2020.04.023

1. Introduction

A protein's amino acid sequence evolves under the presence of mutation, genetic drift, and selection. One key determinant for how proteins evolve is the distribution of fitness effects (DFE) of mutations, which describes both the frequency and magnitude of fitness effects. The frequency of beneficial mutations impacts the rate of adaptive evolution. In addition, the distribution of neutral and deleterious mutations plays a crucial role in determining the rate of protein evolution (how fast protein sequences change), which is of central importance for the reconstruction of evolutionary history and mechanisms.

With the advent of deep sequencing technology, it has become possible to measure fitness effects in a high-throughput and comprehensive manner through Deep Mutational Scanning (DMS) experiments. DMS studies subject large-scale libraries of mutations to an enrichment process in which the enrichment ratio is dependent on the effect of the mutation. Over the course of the last decade, DMS has been harnessed as a powerful technique for measuring the fitness effects of mutations on a large-scale (Boucher, Bolon, & Tawfik, 2016; Canale, Cote-Hammarlof, Flynn, & Bolon, 2018; Fowler & Fields, 2014; Kowalsky et al., 2015). The design of DMS experiments determines what the enrichment basis and, thus, what the calculated fitness metric measures. Most DMS studies have focused on the proteins themselves. These studies use the term "fitness" to refer to a measure of the protein's property (e.g., catalytic activity or ligand affinity) or a phenotype it confers (e.g., antibiotic resistance) relative to some wildtype reference. Often, the selection pressure is artificial. For example, mutational effects on ligand affinity can be observed using cell surface display and fluorescence-activated cell sorting (Whitehead et al., 2012) and mutational effects on enzyme activity can be observed using engineered cells whose growth rate in a particular environment depends on enzyme activity (Wrenbeck, Azouz, & Whitehead, 2017). Although such studies provide valuable information on a component of fitness, they do not accurately capture the DFE upon which evolution acts. There are multiple reasons for this issue. (a) If the protein is characterized outside its native environment, the measure may not accurately reflect the protein's properties in its native environment. (b) A protein may have secondary functions other than the one being characterized. (c) A variety of mechanisms can buffer fitness effects from mutational effects on the protein's physicochemical properties.

These include a protein's stability buffer (Tokuriki & Tawfik, 2009), the action of chaperones to maintain properly folded proteins, cells may have an excess of the protein's activity (more than enough to maintain fitness) (Hartl, Dykhuizen, & Dean, 1985), and the fact that cells are complex with multiple levels of organization between DNA sequence and organismal fitness. (d) A mutation may affect fitness through mechanisms other than effects on its ability to perform its physiological function. We recently characterized such "collateral" fitness effects for TEM-1 β-lactamase (Mehlhoff et al., 2019).

Growth competition experiments with the gene in its native environment should be used if the focus of the experiment is on organismal fitness and the DFE upon which evolution acts (e.g., Lind, Arvidsson, Berg, & Andersson, 2016; Mehlhoff et al., 2019; Melamed, Young, Gamble, Miller, & Fields, 2013; Noda-Garcia et al., 2019; Roscoe, Thayer, Zeldovich, Fushman, & Bolon, 2013). Here, we present our method for constructing comprehensive libraries of gene variants and measuring the fitness effects of these libraries through growth competition and deep mutational scanning. Our method consists of a single-flask growth competition experiment with quantification of allele frequency by deep sequencing at select timepoints during exponential growth. These experiments measure the rate at which allele frequencies change within the population. Shifts in allele frequency are one of the closest estimates for organismal fitness as they represent the effect of mutations on the cellular exponential growth rate. However, such measures do not capture the fitness effects of mutation on lag phase, stationary phase, and recovery from long-term dormancy. They also only capture the fitness effects in the experiment's environment. Fitness across a range of environments will govern a protein's evolution (Noda-Garcia et al., 2019).

2. General considerations

One consideration in the experimental design is whether to leave the gene on the chromosome under its native promoter. Mutations for DMS can be introduced into genes in their native chromosomal environment by integrating libraries of mutations into the selected strain through λ Red recombineering (Lind et al., 2016; Lind, Berg, & Andersson, 2010; Lundin, Tang, Guy, Nasvall, & Andersson, 2017) or other methods (Noda-Garcia et al., 2019). The advantages of doing so are that it ensures the gene is expressed at native levels, expression level responds to any native

regulation the mutations might affect, and cell-to-cell gene copy number does not have the potential to vary like it might when incorporated into a plasmid. A disadvantage is that the alleles will be under selective pressure during library construction and growth that precedes the growth competition experiment (assuming the gene's promoter cannot be turned off during this time). This may limit identification of strongly deleterious mutations because they are not observed in the sequencing or the fitness effects cannot be accurately quantified because the mutation's frequency is too low.

We prefer to place the gene on a low-copy plasmid (to facilitate large-scale library construction) under an inducible promoter to repress gene expression levels until the growth competition experiment. Doing so helps to prevent the loss of deleterious variants during library construction or propagation of cells preceding growth competition. We placed our gene of study under the IPTG-inducible tac promoter on a plasmid with a *p15A* origin. The *tac* promoter can be strongly repressed in the absence of IPTG by using a strain which overexpresses LacI, as is the case for the strain NEB 5-alpha $LacI^q$. Strong repression may not be possible for essential genes. Conversely, too high an expression level may lead to the masking of fitness effects due to protein overabundance. Expression levels should ideally be maintained at a level that allows for observation of fitness effects without expression of the gene itself causing significant fitness or phenotypic effects. Expression at native levels ensures that the measured fitness effects are most relevant to the evolution of that gene.

A variety of methods for library construction exist and one's choice depends on the type of library desired. A common strategy is comprehensive site-saturation libraries, which contain all possible single amino acid substitutions. Our lab developed PFunkel, a single-pot oligo-directed mutagenesis method, specifically for constructing such libraries (Firnberg & Ostermeier, 2012). The method works best on a single-stranded DNA template prepared using helper phage. This requires an f1 phage origin on the plasmid. The Whitehead group developed an alternate method for preparing the single-stranded template for PFunkel mutagenesis (Wrenbeck et al., 2016). This method, called nicking mutagenesis, avoids the need for an f1 origin or the use of phage. Inverse PCR (Jain & Varadarajan, 2014) is an alternative to PFunkel and nicking mutagenesis and is described in this chapter. Although the method is more labor-intensive and requires twice as many primers, it is simpler (it requires fewer manipulations of the DNA) and offers some other advantages. Success of each mutagenesis reaction can be evaluated at every position by electrophoresing the inverse PCR

reaction on an agarose gel. The PCR reactions for problematic positions can be optimized by standard approaches including primer redesign. Reactions can be combined at different volume ratios based on the gel intensities or fluorescent nucleic acid stain assays to better balance the frequency of mutations across positions.

Sub-libraries for different regions of the gene are often used due to read-length limitations of deep sequencing methods. The number of sub-libraries created depends on the length of the gene and the deep sequencing method to be used. PacBio allows sequencing of the entire gene but is considerably more expensive per sequencing read. More often, the gene is divided into sub-libraries in which mutagenesis is limited to a region that is at or less than the length of the sequencing read. Such libraries typically undergo separate growth competition experiments. One artifact of limited sequencing read-length is that fitness effects can arise from unintended mutations that are unobserved because they lie outside the sequencing read (either in other sub-regions of the gene or elsewhere on the plasmid, such as in the promoter or plasmid origin). Such mutations may occur spontaneously during growth but more likely arise during library creation. Sub-cloning the library (and especially just the mutagenized region) into a fresh vector after creation reduces the frequency of these spurious effects.

3. Methods

Some elements of this protocol are specific to the experimental system we used in a recent study of the fitness effects of mutations in TEM-1 β-lactamase in the absence of β-lactam antibiotics (Mehlhoff et al., 2019). Most of these derive from the fact that we used a plasmid encoding spectinomycin resistance in which our gene of interest (TEM-1 β-lactamase) was under the control of the IPTG-inducible *tac* promoter. This plasmid contained a unique *Sph*I site outside of *TEM-1*, which was used to linearize the vector to prepare the DNA for deep sequencing.

3.1 Primer design for inverse PCR

Pairs of forward and reverse primers need to be designed for every codon in the gene in order to form a comprehensive library of single-codon substitutions. One of the primers in a pair contains the degenerate codon at the 5′ end. N base pair degeneracy in this codon allows for 64 possible codon combinations and is used if one is interested in codon-specific effects. Data for synonymous codons can be compared to identify potential fitness effect

artifacts from unintended mutations outside the sequence region. However, synonymous codons can have different fitness effects (Faber, Wrenbeck, Azouz, Steiner, & Whitehead, 2019; Firnberg, Labonte, Gray, & Ostermeier, 2014). Alternatively, NNK, NNS, or NNB libraries can be used to reduce the frequency of stop codons while keeping all 20 amino acids. It is best to design the primers to have similar melting temperatures as PCR reactions are performed in parallel in a 96-well PCR plate. Examples of forward and reverse primers for the thirtieth codon in *TEM-1* (K30) are shown below.

Forward primer:

5′-(N:25252525)(N)(N) GTA AAA GAT GCT GAA GAT CAG TTG GG-3′.

Reverse primer:

5′-CAC CAG CGT TTC TGG GTG A-3′.

Oligonucleotides containing degenerate bases should ideally be ordered as hand-mixed, a more expensive option that is designed to produce an equimolar ratio of A:C:T:G at the variable N positions. We ordered primers at a concentration of 100 μM in IDTE Buffer (pH 8.0) then diluted the primers to 10 μM as a working solution.

3.2 Library construction by inverse PCR

Inverse PCR reactions are performed in separate wells of 96-well PCR plates for each substitution site (Fig. 1). Reactions are electrophoresed on a gel to confirm successful amplification of the full-length plasmid. The successful reactions are later pooled.

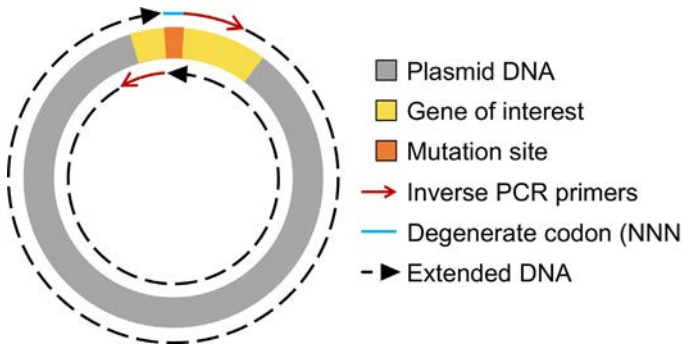

Fig. 1 Schematic depiction of multiplexed inverse PCR using degenerate primers with a 5′ NNN at the mutation site.

Table 1 Thermocycling conditions for using inverse PCR to make single-codon substitutions.

Step	Temperature	Time
1	95 °C	2 min
2	95 °C	30 s
3	58 °C	30 s
4	65 °C	2 min[a]
Repeat steps 2–4 for 30 cycles		
5	65 °C	5 min
6	4 °C	∞

[a]The extension time corresponds to that for a ~4 kb plasmid and should be adjusted based on the size of the plasmid. At least 30 s of extension time should be used per kb of PCR product.

1. Make a master mix consisting of 1 μL of 50 ng/μL template plasmid:19 μL of H_2O:25 μL of Phusion High-Fidelity PCR Master Mix with HF Buffer (NEB) for each reaction. Add 45 μL of the master mix to each well of a 96-well PCR plate.
2. Add 2.5 μL of 10 μM forward primer and 2.5 μL of the corresponding 10 μM reverse primer to each well.
3. Perform inverse PCR using the thermocycling conditions shown in Table 1.
4. Following PCR, combine 3 μL of the PCR product with 1 μL 6× Purple Loading Dye and 2 μL H_2O.
5. Prepare a TAE 1.0% agarose gel containing 0.5 μg/mL ethidium bromide.
6. Load the samples into the gel and electrophorese at 110 V for approximately 40 min. The time and voltage may need to be adjusted depending on the size of the PCR product.
7. Visualize bands under UV light and image.
8. For any samples which do not have a prominent band at the same size as the template, repeat steps 1–7. A change in annealing temperature may be necessary. The extension time and amount of starting template added to each well can also be altered if necessary.

3.3 Purification of PCR products

1. Prepare a 1.0% agarose gel containing 0.5 μg/mL ethidium bromide using a wide comb which can ideally hold 150 μL or more in each well.

2. For each sub-library, pool all the successful PCR reactions together at the desired ratio. Sub-libraries are used when the sequencing read-length is less than the gene's length (see General Considerations).
3. Add 6× Purple Loading Dye such that it makes up 1/6th of the final volume.
4. Load the pooled successful PCR reactions into the well. In a separate well, load linearized template as a control.
5. Run the DNA gel at 110 V for 40 min.
6. Isolate the band that is the same size as the template and extract the desired DNA using a PureLink™ Quick Gel Extraction Kit (Invitrogen) adhering to the manufacturer's instructions.
7. Concentrate the DNA using the DNA Clean & Concentrator Kit (Zymo) according to the manufacturer's instructions.
8. Determine the concentration of the DNA spectrophotometrically using Eq. (1)

$$[\text{dsDNA}] = 50\frac{\mu\text{g}}{\text{mL}} * (A_{260} - A_{320}) * \text{d} \quad (1)$$

where A_{260} and A_{320} are the absorbances at 260 and 320 nm, respectively and d is the dilution factor by which the sample has been diluted if necessary.

3.4 Phosphorylation and ligation of PCR product

1. Add approximately 400 ng of DNA, 2 μL of T4 DNA Ligase Buffer, 1 μL (10 units) of T4 PNK (NEB), and ddH_2O to a total volume of 20 μL. Pipette the solution up and down to mix and then briefly spin down in a microcentrifuge.
2. Incubate the sample at 37 °C for 1 h.
3. Use the DNA Clean & Concentrator Kit (Zymo) to purify the phosphorylated DNA. Elute the phosphorylated DNA using 15 μL of nuclease-free water.
4. Add 2 μL of T4 DNA Ligase Buffer, 1 μL (400 Cohesive End Units) of T4 DNA Ligase (NEB), and nuclease-free water to a final volume of 20 μL. Pipette the solution up and down to mix and then briefly spin down in a microcentrifuge.
5. Incubate at room temperature for 1 h.
6. Store the prepared DNA at −20 °C. One μL will be used per transformation.

3.5 Transformation of library into strain for growth competition

The ligation mixture is then transformed into the desired strain. If the desired strain transforms poorly, the library can be first transformed into a high competency strain to maximize library size. Subsequently, plasmid DNA can be prepared from these transformants and transformed into the desired strain. We have used NEB 5-alpha LacI^q cells (*F′ proA+B+ lacI^q Δ(lacZ)M15 zzf::Tn10 (TetR) / fhuA2Δ(argFlacZ) U169 phoA glnV44 Φ80Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*). We elected to use this particular derivative of DH5α cells for their high transformation efficiency, the lack of functional *endA* and *recA* genes, and the presence of the *lacI^q* mutation. LacI^q denotes a mutation within the *lacI* promoter which leads to increased transcription of *lacI*. Overexpression of LacI causes the *tac* promoter to be strongly repressed in the absence of IPTG. We also use 2% w/v glucose in the media to help repress the promoter. We chose this experimental design to guard against losing deleterious alleles during library creation and propagation that preceded the growth competition experiment.

The desired number of transformants for adequate coverage of the library depends on the degeneracy of the target library and can be estimated (Bosley & Ostermeier, 2005). Under the simplifying assumption that each library member is equally frequent, one needs ~4.6 times more transformants than the number of variants to ensure that a particular variant has a 99% probability of being present in the library. To have a 99% probability that a library is complete requires the number of transformants be at least 10- to 25-fold higher than the number of variants (the precise number depends on the number of intended variants). Multiple transformations can be performed to reach the target number of transformants.

1. Make LB-agar plates containing 2% w/v glucose and 50 μg/mL spectinomycin. We used 2% glucose to help repress expression, but the addition of this high level of glucose also helps recovery of the library from the plate (the colonies lift off the plate more readily). We recommend making sets of plates at two different sizes. One set of plates should be made using 200 mL of agar in 245 mm × 245 mm square sterile dishes. These plates will be used for collecting the transformants that will constitute the library. A smaller set of plates should be made using 20 mL of agar in 100 mm × 15 mm round Petri dishes to be used in estimating the total number of transformants on the large plates.

2. Transform the ligated plasmid into NEB 5-alpha LacI^q cells following the manufacturer's high efficiency transformation protocol. In brief, thaw the competent cells on ice for 10 min, add 100 ng of plasmid DNA, and gently flick the tube to mix. Allow the mixture to sit on ice for 30 min before performing a heat shock at 42 °C for 30 s and placing back on ice for 5 min. Add 950 μL of room temperature SOC media and incubate at 37 °C with vigorous (250 rpm) shaking for 1 h.
3. Prepare a 10-fold dilution by mixing 15 μL of the cells into 135 μL of room temperature SOC media.
4. Spread 100 μL of the undiluted and 100 μL 10-fold diluted cells onto separate 20 mL agar plates.
5. Spread the remaining 885 μL of the undiluted cells onto a 200 mL agar plate.
6. Incubate the plates at 37 °C for approximately 16 h.
7. Count the number of colonies on the 20 mL agar plates. These colony counts can be used to estimate the number of transformants on the 200 mL agar plate by accounting for the difference in volume of cell suspension plated and any dilution before plating.
8. Select 10 colonies from the 20 mL agar plate to grow cultures from in 10 mL of growth media containing 2% w/v glucose and 50 μg/mL spectinomycin. Purify plasmid from the cultures after they have been incubating overnight using a Qiagen Plasmid Miniprep Kit. Plasmid can be submitted for Sanger sequencing in order to verify the presence of single-codon substitutions.
9. Collect the cells from the 200 mL agar plate by first adding a mix of 7 mL of growth media and 3 mL of 50% glycerol and scraping the colonies off the surface of the agar with an L-shaped spreader. Then add an additional mix of 3.5 mL growth media and 1.5 mL 50% glycerol and recover the remaining cells on the plate.
10. Combine the cell suspensions and centrifuge for 4 min at 4500 × *g*. Draw off supernatant until the volume of the supernatant and pellet are about equal. Resuspend the pellet in the remaining supernatant and aliquot into tubes for −80 °C storage.
11. Repeat the transformation steps until enough transformants have been collected to ensure high coverage of the library (Bosley & Ostermeier, 2005). The frozen aliquots will later be combined into a single library during the pre-growth competition growth phase as described in the subsequent section.

3.6 Growth competition

The choice of the growth conditions (e.g., media, media volume, vessel, temperature, batch vs chemostat) is up to the researcher. The fitness measurements will be specific to that environment. Since libraries are typically stored as frozen stocks, it is important to grow the cells for enough time before the growth competition to allow the cells to recover. Our protocol has an extended growth period prior to the start of the growth competition experiment to allow all cells to completely recover from being frozen and "erase" any difference that might occur because some cells or sub-libraries have been frozen for longer periods of time. We observed evidence of such an effect in our recent study (Mehlhoff et al., 2019). A separate stock of cells with the wildtype allele might also be used. Although library construction techniques typically create wildtype alleles, the advantage of using a separate stock of wildtype allele is that you can spike it in the library at a desired higher frequency to better allow for calculation of the frequency of the wildtype allele. The wildtype frequency is very important, since it is the reference for all fitness values.

If an inducible promoter is used, one must choose when to induce. Induction at the beginning of the growth competition best guards against loss of allele diversity prior to the experiment. However, it will take time for the protein to reach steady-state levels in the cell; thus, the mutation's effect will change as a function of time until steady state is reached. The resulting fitness effect measured is based on the mean growth rate over the growth competition. In our recent experiments using the *tac* promoter (Mehlhoff et al., 2019), fitness effects did not fully manifest until 2–3 h of post-induction growth in LB at 37 °C.

We used the following protocol in a recent experiment in which we induced with IPTG at the start of the growth competition and grew the cells for 10 generations (Mehlhoff et al., 2019). The length of time for the growth experiment is a balance between allowing enough time for small fitness effects to be observed, but not so long such that the magnitude of the effects of deleterious mutations cannot be distinguished. Fig. 2 is a useful guide that shows how allele frequency with deleterious mutations will vary as a function of mutational effect and number of generations. After long periods of growth, comparatively small magnitude fitness effects become detectable. However, there are disadvantages to an increase in the number of generations. The relative allele frequency for highly deleterious mutations will approach zero, eventually making the magnitude of deleterious fitness effects

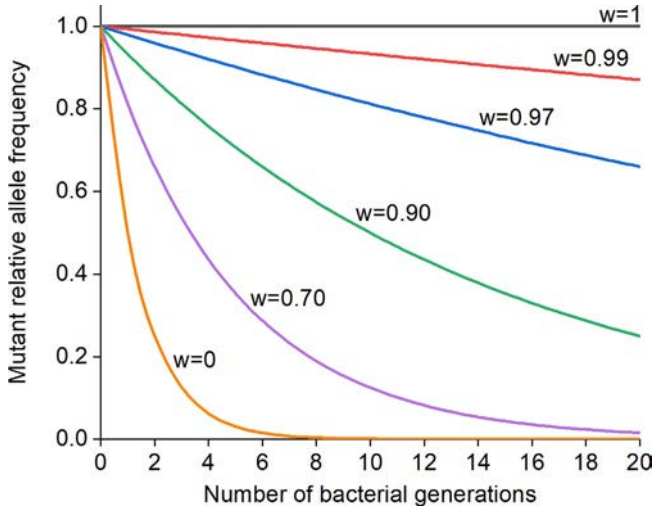

Fig. 2 Mutant relative allele frequency as a function of relative fitness (*w*) and the number of generations of growth. The mutant relative allele frequency is a ratio of the number of times a mutant allele appears compared to the wildtype allele and assumes they are equally as frequent at the start of the growth competition.

indistinguishable from one another. Long periods of growth also increase the risk of background mutations accumulating and affecting allele frequencies.

In our recent experiments, we elected to grow the culture for 10 generations to potentially allow detection of deleterious mutational effects as small as 1%. We chose to dilute the cultures after 5 generations so that growth would remain in exponential phase for all 10 generations (Fig. 3). After a long pre-growth competition growth period, the first phase of the growth competition occurred from a starting optical density (OD) at 600 nm of 0.020 to an OD of 0.640, which corresponds to five generations of growth. We then diluted the culture back to an OD of 0.020 and allowed it to grow to an OD of 0.640 again. The result is 10 generations of growth with the cells remaining in the exponential growth phase for the majority of their growth. All measurements of OD were taken at a wavelength of 600 nm.

In this protocol, samples for deep sequencing analysis are collected at the start and end of growth competition. Similar studies measuring biological fitness through growth competition have used deep sequencing either comparing the relative abundance of mutants across more than two timepoints (Roscoe et al., 2013) or across differences in growth environment (Stiffler, Hekstra, & Ranganathan, 2015).

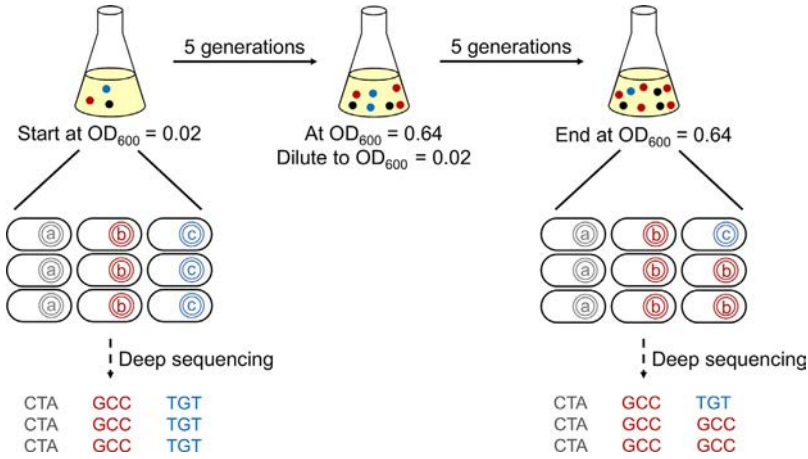

Fig. 3 Schematic of the growth competition experiment and timepoints at which samples are collected for deep sequencing. Cultures were grown for five generations from an OD of 0.020 to 0.640 before being diluted and allowed to grow for an additional five generations. Cells were pelleted and plasmid prepped for deep sequencing immediately before induction and again after 10 generations of induced growth.

1. Dilute frozen library stocks and wildtype cells into separate flasks of 100 mL growth media containing 50 μg/ml spectinomycin and 2% w/v glucose. The volume of cells should be sufficiently large to ensure adequate coverage of the library without notably increasing the turbidity of the media. Incubate the resulting inoculums at 37 °C for approximately 16 h.
2. The next morning, measure the OD of the cultures in triplicate. Average the readings and calculate the dilution to make a 100 mL culture with a final OD of 0.020.
3. Mix wildtype and library cultures at a ratio of 5:95 to make a 100 mL culture.
4. Incubate the flask at 37 °C with 250 rpm shaking until the OD is around 0.5.
5. Collect 10 mL of the co-culture in a centrifuge tube on ice. As soon as time allows, centrifuge at 4 °C and 4000 × *g* for 10 min. Extract plasmid from the pelleted cells using the Plasmid Miniprep Kit (Qiagen). If the strain being used is *endA*$^+$, be sure to perform the extra wash step as indicated in the instructions. Store the DNA at 4 °C as the initial time point sample.

6. Dilute the remaining culture to an OD of 0.020 in 100 mL of media in a shake flask. The media and flask should be pre-warmed to 37 °C.
7. At time zero, add filter sterilized IPTG to a final concentration of 1 mM. Incubate the flasks at 37 °C with shaking until the OD is approximately 0.640 (5 generations of growth).
8. Dilute the culture to an OD of 0.020 in a pre-warmed flask containing 100 mL media.
9. Incubate at 37 °C with shaking until the OD reaches approximately 0.640.
10. Place 10 mL of the co-culture on ice and immediately centrifuge at 4000 × *g* for 10 min at 4 °C. Isolate and store the 10-generation plasmid as in step 5.

3.7 Attaching Illumina index sequences

The specifications chosen for deep sequencing are up to the researcher as well. We elected to use Illumina MiSeq deep sequencing with 2 × 300 bp reads. Illumina HiSeq can be utilized to generate more reads (~25 million total reads from MiSeq; ~300 million total reads from HiSeq). The 300 bp read-length is the longest read-length currently available for Illumina MiSeq. Selection of the read-length will depend on the length of the gene and determines the number of sub-libraries necessary. Paired end reads allow for higher accuracy in read alignment and the ability to detect DNA rearrangements such as insertion-deletion (indel) variants and inversions. Adapter sequences are designed to anneal to the linearized DNA and amplify it in preparation for Illumina deep sequencing. Using an index sequence allows for the samples to be pooled and run in a single lane of Illumina MiSeq as samples can later be identified by their corresponding index. We used an index sequence of GCCAAT for our time zero sample and an index sequence of CTTGTA to represent the 10-generation sample. Examples of forward and reverse adapters for the first region of TEM-1 (the first 95 codons) are shown in Fig. 4.

1. Mix 50 ng of plasmid collected during the growth experiment with 10 μl 10 × CutSmart Buffer (NEB) and water to a total volume of 97 μL.
2. Add 3 μL (60 units) of *Sph*I-HF (NEB) and incubate for 1 h at 37 °C.
3. Purify the plasmid using the DNA Clean & Concentrator Kit (Zymo).
4. Add 1.25 μL of each of the forward and reverse Illumina adapter sequences, 9 μL of nuclease-free water, and 12.5 μL of Phusion High-Fidelity PCR Master Mix with HF Buffer (NEB) to 1 μL of linearized DNA. The reverse adapter contains the index sequence, so a different

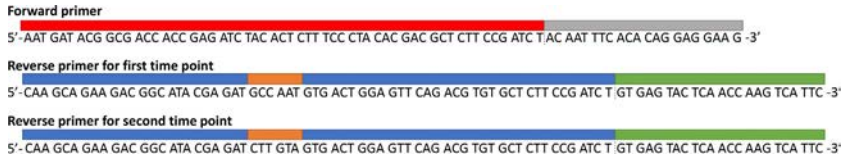

Fig. 4 Forward and reverse primers designed for Illumina deep sequencing of the first sub-library of TEM-1. The forward primer consists of the forward Illumina adapter (red) and forward annealing sequence (gray). The reverse primer consists of the reverse Illumina adapter (blue), index sequence (orange), and reverse annealing sequence (green). The annealing sequence is altered based on the region of the gene to be amplified. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2 Thermocycling conditions for attaching adapters with Illumina index sequences to the linearized DNA.

Step	Temperature	Time
1	98 °C	30 s
2	98 °C	15 s
3	47 °C	15 s
4	72 °C	3 min
Repeat steps 2–4 for 25 cycles		
5	72 °C	5 min
6	4 °C	∞

reverse adapter should be used for the time zero and 10-generation samples in order to identify the samples during demultiplexing.

5. Perform PCR using the thermocycling conditions shown in Table 2.
6. Estimate the concentration of the PCR product using spectrophotometry and Eq. (1).
7. Verify the PCR product runs at the proper size on a 1% agarose gel with 0.5 μg/mL ethidium bromide.
8. Pool the samples together and submit for Illumina MiSeq (2×300 bp reads).

3.8 Deep sequencing analysis of allele frequency

It is necessary to verify that the deep sequencing files being analyzed have not been corrupted during any upload, download, or other file transfer process.

A checksum is a string of characters which acts as a "digital fingerprint." Tracking the checksum of a file is a way to ensure that the contained data has not been damaged during file transfer. Programs like FastQC (Wingett & Andrews, 2018) are another useful tool in performing quality control checks of raw deep sequencing reads. It provides a summary report of sequencing quality and scores as well as identifies any concerning patterns within the raw data.

There are many possible pathways when deciding on tools for processing and analyzing deep sequencing reads. We elected to use PEAR (Zhang, Kobert, Flouri, & Stamatakis, 2014), Trimmomatic (Bolger, Lohse, & Usadel, 2014), and Enrich2 (Rubin et al., 2017). PEAR works to align paired-end reads. Trimmomatic then cuts a specified number of bases from the start and end of the paired-end reads. We adjusted the number of bases needing to be cropped from each sub-library in order to remove bases corresponding to the Illumina adapter sequences as well as bases outside the desired region. Enrich2 then quality filters the reads and counts variants by comparing the read sequence to the input wildtype sequence.

1. Download the Illumina MiSeq reads and verify the checksums in order to ensure the files have not been corrupted during the upload or download process.
2. Inspect the reads for their per base sequence quality using FastQC.
3. Merge the paired-end reads using PEAR. We elected to run PEAR with a minimum assembly length of 200 and a maximum assembly length of 500 based on the expected length of our reads.
4. Crop the paired-reads using Trimmomatic so as to remove the adapter sequences and any base pairs outside the desired region. For TEM-1 and our specific adapter sequences, we used the following settings: Region 1—HEADCROP:24, CROP:285; Region 2—HEADCROP:20, CROP:285; Region 3—HEADCROP:24, CROP:291.
5. Input the paired and cropped reads to Enrich2 for counting of variants. We set Enrich2 to ignore any reads that contained bases with a quality score below 20, bases marked as N, or mutations at more than one codon.

3.9 Fitness calculation for DMS

Our approach to calculating fitness from growth competition experiments is very similar to (and based on) that of Kowalsky et al. (2015) and Rubin et al. (2017). The difference is our approach calculates a fitness based on growth

rate instead of using an enrichment ratio as a proxy for fitness. An additional advantage is that fitness values can be compared between experiments with a different number of generations. A detailed derivation of our equations for calculating fitness and associated statistical measures can be found in our recent study on collateral fitness effects of mutations (Mehlhoff et al., 2019).

We calculated the fitness of each variant using the allele counts tabulated by Enrich2 and the fold increase in the number of cells during the experiment. We assume that the optical density (O) of the cell culture is linearly proportional to the number of cells. The fold increase in the number of cells (r) is calculated by the ratio of optical densities at the start (designated by subscript o) and end (designated by the subscript f) of growth competition with a d-fold dilution midway through the induced phase of growth.

$$r = \frac{O_f d}{O_o} \tag{2}$$

We calculated an enrichment ratio for allele i from the allele counts (c) at the start and end of the growth competition in comparison to the total sequencing counts (designated by the subscript T). A value of 0.5 is added to the counts in order to allow an estimate of the fitness in cases where the initial or final counts of allele i is equal to zero (Rubin et al., 2017).

$$\varepsilon_i = \frac{c_{if} c_{To}}{c_{io} c_{Tf}} \approx \frac{\left(c_{if} + 0.5\right) c_{To}}{\left(c_{io} + 0.5\right) c_{Tf}} \tag{3}$$

Comparing the enrichment scores of allele i and wildtype (i.e., the enrichment ratio) can be used as a measure of fitness (Kowalsky et al., 2015; Rubin et al., 2017). Enrichment scores will vary depending on the number of generations of growth. We prefer to calculate fitness as the growth rate of cells containing allele i relative to the growth rate of cells containing the wildtype allele. To do this, we use the resulting enrichment, as calculated in Eq. (3), along with the fold increase in the number of cells to calculate fitness (w), which is the growth rate (μ) of cells with allele i compared to the growth rate of cells with the wildtype allele (designated by the subscript wt) .

$$w_i = \frac{\mu_i}{\mu_{wt}} = \frac{\ln\left(r\varepsilon_i\right)}{\ln\left(r\varepsilon_{wt}\right)} \tag{4}$$

3.10 Statistical treatment of DMS fitness measurements

There are two intrinsic sources of variance in the fitness measurement: the sequencing counts and the fold increase in total number of cells (r). We assumed that the uncertainty in r is negligible due to it appearing in the numerator and denominator of the fitness calculation along with it being of smaller magnitude than the variance stemming from the counts. Using this assumption, we can calculate the variance in the fitness measurement as

$$\sigma_w{}^2 = w_i{}^2 \left[\frac{\frac{(1-f_{if})}{c_{if}} + \frac{(1-f_{io})}{c_{io}}}{(\ln r\varepsilon_i)^2} + \frac{\frac{(1-f_{wtf})}{c_{wtf}} + \frac{(1-f_{wto})}{c_{wto}}}{(\ln r\varepsilon_{iwt})^2} \right] \tag{5}$$

where $f_i = c_i/c_T$ and $f_{wt} = c_{wt}/c_T$ (Mehlhoff et al., 2019). The confidence interval can then be calculated from the variance in fitness as

$$\pm z^* \frac{\sigma_w}{\sqrt{1}} \tag{6}$$

where $z^* = 2.576$ for the 99% confidence interval.

One can calculate P-values to assess if the fitness of cells containing allele i is different than that of cells with the wildtype allele. This value only depends on the sequencing counts. We test if two proportions are the same using the null hypothesis

$$H_0 : \frac{c_{io}}{c_{wto}} = \frac{c_{if}}{c_{wtf}} \tag{7}$$

and calculate a Z-score

$$Z = \frac{\hat{p}_o - \hat{p}_f}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{c_{wto}} + \frac{1}{c_{wtf}}\right)}} \tag{8}$$

$$\hat{p}_o = \frac{c_{io}}{c_{wto}} \tag{9}$$

$$\hat{p}_f = \frac{c_{if}}{c_{wtf}} \tag{10}$$

$$\hat{p} = \frac{c_{io} + c_{wto}}{c_{if} + c_{wtf}} \tag{11}$$

P-values are calculated by determining the area under the curve for that Z-score for one tail of the normal distribution (using the NORMSDIST(Z)

function in Excel) and then multiplying by 2 for a 2-tailed test (i.e., the fitness might be higher or lower than 1).

The above calculations account for the intrinsic uncertainty in a fitness measure by DMS based on the sequencing counts. They do not measure experiment-to-experiment variability. We recommend biological replicates for DMS experiments as an additional assessment of uncertainty in fitness values. Fitness measures can be presented as a weighted mean of the fitness values. This mean is useful when presenting heat maps of the landscape or when analyzing a set of fitness values as a whole. However, since typically only one or two replicas are performed owing to sequencing cost, the standard error of this weighted mean may not reflect the uncertainty in the fitness value. Fitness effects of individual mutations are best presented as separate values each with their own standard error or confidence interval (see Mehlhoff et al., 2019 for an example).

One final note on uncertainty in fitness measures: an implicit assumption in the DMS analysis is that allele frequency reflects cell frequency. This assumption would be violated if, for example, deleterious mutations caused selection for cells with fewer copies of the plasmid.

3.11 Assays for verifying magnitude of fitness effects

Fitness effects of mutations can be verified in a number of ways. Here, we present a simple monoculture growth assay, in which cells containing the mutant allele and cells containing the wildtype allele are incubated in separate flasks side-by-side and culture growth is monitored by OD. This method works best for larger fitness effects (i.e., more than a few percent). For smaller fitness effects, fluorescent markers can be used to track the relative prevalence of wildtype and mutant alleles in co-cultures (Lind et al., 2010; Lind et al., 2016; Lundin et al., 2017). For this method cells with wildtype and mutant alleles contain different fluorescent proteins. These cells are mixed at a 1:1 ratio, and the fluorescence signal of each marker is monitored over time. Care must be taken to ensure that the two fluorescent proteins do not have different fitness effects of their own. Alternatives for assessing fitness effects of mutation in co-culture include quantifying the frequency of the mutant allele by analysis of Sanger sequencing chromatograms of the isolated plasmid from the co-culture (Mehlhoff et al., 2019)—a method that requires the construction of a standard curve by sequencing samples of known ratios of wildtype and mutant plasmids.

For the monoculture growth assay, one must first make the desired mutant(s) by standard site-directed mutagenesis techniques. Next, cultures containing the wildtype or desired mutant alleles are grown using the same methods as described for growth competition DMS experiments (see Section 3.6) except the cultures are monocultures and not mixtures of wildtype (WT) and mutant alleles. This includes the pre-induction growth periods. After the growth competition starts, growth is monitored by measuring the OD at a desired interval, with the final OD measurement at the end of the experiment being measured in triplicate and averaged (since fitness strongly depends on this value). Fitness is calculated from Eq. (12) by substituting in the initial (designated by subscript *o*) and final (designated by subscript *f*) OD values.

$$w_i = \frac{\mu_i}{\mu_{wt}} = \frac{\left(\ln \frac{O_f d}{O_o}\right)_i}{\left(\ln \frac{O_f d}{O_o}\right)_{wt}} \tag{12}$$

For the initial values, we use the intended starting OD based on the dilution (i.e., an OD of 0.02) instead of a measurement of the starting OD because the uncertainty in such a low OD measurement is too large relative to its value. This monoculture growth assay for measuring fitness assumes that the correlation between optical density (O) and the number of cells is the same for mutant and wildtype cells, which may not be true if the mutations affect cell division, size, or morphology (Mehlhoff et al., 2019).

Alternatively, fitness can be measured using the growth rates as determined by the slope of a $\log_2$(OD) vs. time linear fit during exponential growth once the cells reach steady-state growth. Note that although *E. coli* cells in LB media can grow exponentially to high OD levels (much greater than 1.0), the maximum steady-state growth rate occurs below an OD of about 0.30 (Sezonov, Joseleau-Petit, & D'Ari, 2007).

4. Summary

Our outlined methods provide a protocol for constructing comprehensive libraries of mutations and measuring the associated fitness effects using deep mutational scanning. Characterization of individually constructed mutants by monoculture growth assays or a variety of alternatives allow for the confirmation of the fitness effects. Growth competition experiments and deep mutational scanning can be used in combination as a

powerful tool for high-throughput measurements of fitness effects allowing researchers to study a wide array of gene variants or environmental impacts on biological fitness.

Acknowledgments

We thank Erh-Yeh Tsou for preparing drafts of Figs. 1 and 3. This work was supported by National Science Foundation grants DEB-1353143 and MCB-1817646.

References

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120.

Bosley, A. D., & Ostermeier, M. (2005). Mathematical expressions useful in the construction, description and evaluation of protein libraries. *Biomolecular Engineering*, *22*, 57–61.

Boucher, J. I., Bolon, D. N., & Tawfik, D. (2016). Quantifying and understanding the fitness effects of protein mutations: Laboratory versus nature. *Protein Science*, *25*(7), 1219–1226.

Canale, A. S., Cote-Hammarlof, P. A., Flynn, J. M., & Bolon, D. N. (2018). Evolutionary mechanisms studied through protein fitness landscapes. *Current Opinion in Structural Biology*, *48*, 141–148.

Faber, M. S., Wrenbeck, E. E., Azouz, L. R., Steiner, P. J., & Whitehead, T. A. (2019). Impact of in vivo protein folding probability on local fitness landscapes. *Molecular Biology and Evolution*, *36*(12), 2764–2777.

Firnberg, E., Labonte, J. W., Gray, J. J., & Ostermeier, M. (2014). A comprehensive, high-resolution map of a gene's fitness landscape. *Molecular Biology and Evolution*, *31*(6), 1581–1592.

Firnberg, E., & Ostermeier, M. (2012). PFunkel: Efficient, expansive, user-defined mutagenesis. *PLoS One*, *160*(5), 1–10.

Fowler, D. M., & Fields, S. (2014). Deep mutational scanning: A new style of protein science. *Nature Methods*, *11*(8), 801–807.

Hartl, D. L., Dykhuizen, D. E., & Dean, A. M. (1985). Limits of adaptation: The evolution of selective neutrality. *Genetics*, *111*(3), 655–674.

Jain, P. C., & Varadarajan, R. (2014). A rapid, efficient, and economical inverse polymerase chain reaction-based method for generating a site saturation mutant library. *Analytical Biochemistry*, *449*, 90–98.

Kowalsky, C. A., Klesmith, J. R., Stapleton, J. A., Kelly, V., Reichkitzer, N., & Whitehead, T. A. (2015). High-resolution sequence-function mapping of full-length proteins. *PLoS One*, *10*(3) e0118193.

Lind, P. A., Arvidsson, L., Berg, O. G., & Andersson, D. I. (2016). Variation in mutational robustness between different proteins and the predictability of fitness effects. *Molecular Biology and Evolution*, *34*, 408–418.

Lind, P. A., Berg, O. G., & Andersson, D. I. (2010). Mutational robustness of ribosomal protein genes. *Science*, *330*(6005), 825–827.

Lundin, E., Tang, P. C., Guy, L., Nasvall, J., & Andersson, D. I. (2017). Experimental determination and prediction of the fitness effects of random point mutations in the biosynthetic enzyme HisA. *Molecular Biology and Evolution*, *35*, 704–718.

Mehlhoff, J. D., Stearns, F. W., Rohm, D., Wang, B., Tsou, E. Y., Dutta, N., et al. (2019). Collateral fitness effects of mutations. *BioRxiv*. https://doi.org/10.1101/820068.

Melamed, D., Young, D. L., Gamble, C. E., Miller, C. R., & Fields, S. (2013). Deep mutational scanning of an RRM domain of the Saccharomyces cerevisiae poly(a)-binding protein. *RNA*, *19*(11), 1537–1551.

Noda-Garcia, L., Davidi, D., Korenblum, E., Elazar, A., Putintseva, E., Aharoni, A., et al. (2019). Chance and pleiotropy dominate genetic diversity in complex bacterial environments. *Nature Microbiology*, *4*(7), 1221–1230.

Roscoe, B. P., Thayer, K. M., Zeldovich, K. B., Fushman, D., & Bolon, D. N. (2013). Analyses of the effects of all ubiquitin point mutants on yeast growth rate. *Journal of Molecular Biology*, *425*(8), 1363–1377.

Rubin, A. F., Gelman, H., Lucas, N., Bajjalieh, S. M., Papenfuss, A. T., Speed, T. P., et al. (2017). A statistical framework for analyzing deep mutational scanning data. *Genome Biology*, *18*(1), 150.

Sezonov, G., Joseleau-Petit, D., & D'Ari, R. (2007). Escherichia coli physiology in Luria-Bertani broth. *Journal of Bacteriology*, *189*(23), 8746–8749.

Stiffler, M. A., Hekstra, D. R., & Ranganathan, R. (2015). Evolvability as a function of purifying selection in TEM-1 beta-lactamase. *Cell*, *160*(5), 882–892.

Tokuriki, N., & Tawfik, D. S. (2009). Stability effects of mutations and protein evolvability. *Current Opinion in Structural Biology*, *19*(5), 596–604.

Whitehead, T. A., Chevalier, A., Song, Y., Dreyfus, C., Fleishman, S., De Matos, C., et al. (2012). Optimization of affinity, specificity, and function of designed influenza inhibitors using deep sequencing. *Nature Biotechnology*, *30*(6), 543–548.

Wingett, S. W., & Andrews, S. (2018). FastQ screen: A tool for multi-genome mapping and quality control. *F1000Res*, 7, 1338.

Wrenbeck, E. E., Azouz, L. R., & Whitehead, T. A. (2017). Single-mutation fitness landscapes for an enzyme on multiple substrates reveal specificity is globally encoded. *Nature Communications*, *8*, 15695.

Wrenbeck, E. E., Klesmith, J. R., Stapleton, J. A., Adeniran, A., Tyo, K. E., & Whitehead, T. A. (2016). Plasmid-based one-pot saturation mutagenesis. *Nature Methods*, *13*(11), 928–930.

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: A fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics*, *30*(5), 614–620.