



Published in final edited form as:

*Methods Mol Biol.* 2018 ; 1839: 103–116. doi:10.1007/978-1-4939-8685-9\_10.

## Transposon Sequencing of *Vibrio cholerae* in the Infant Rabbit Model of Cholera

Lauren M. Shull<sup>1</sup>, Andrew Camilli<sup>1</sup>

<sup>1</sup>Tufts University School of Medicine, Boston, USA

### Abstract

Transposon sequencing, or Tn-seq, combines transposon mutagenesis and massively parallel sequencing to allow for rapid and high-throughput identification of genes that play roles in fitness within environments of interest. The bacterial pathogen *Vibrio cholerae* is an excellent candidate for Tn-seq screens due to the availability of a plasmid-based in vivo transposition system and the relative ease with which the cholera disease state can be modeled in animals. This chapter will describe a method for performing Tn-seq screens on *V. cholerae* in the infant rabbit model of cholera.

### Keywords

Transposon sequencing; Tn-seq; *Vibrio cholerae*; Virulence; Cholera; Infant rabbit

## 1 Introduction

Tn-seq is a powerful tool to identify microbial genes that play important roles in a chosen selection condition. A transposon- mutagenized library is constructed such that each individual strain in the library harbors one transposon insertion at a random location in its genome, typically disrupting and inactivating a gene. Libraries of extremely high complexity can be easily generated so that, across the entire library, virtually every gene in the genome has been disrupted at multiple locations. Strains with insertions in essential genes drop out of the library during construction, and thus essential genes are putatively identified by the absence of insertions [1–3]. Genomic DNA from a portion of this starting library is isolated and retained for later PCR amplification and sequencing of the transposon junctions. The library is then subjected to some chosen selection, ideally one that is stringent and reproducible. The surviving bacteria are then recovered from the selection condition, outgrown if necessary, and genomic DNA is collected. The known sequence of the transposon is used as a forward PCR primer binding site, and a different known sequence is added downstream of the transposon by the method described below, providing a reverse PCR primer binding site. The transposon junctions are amplified by PCR, and the end of the transposon is used as a priming site for sequencing of the junctions [4]. The resulting reads are then mapped to the reference genome, identifying the location of each transposon insertion and relative abundance of each mutant in the library [1–3, 5, 6]. Mutants with insertions in genes required for the selection will be underrepresented in the output (or mutants in genes deleterious for the selection will be overrepresented), identifying such genes as candidates for validation and further study [1–3].

Although *V. cholerae* and other bacteria can be mutagenized with many types of transposon (e.g., Tn5 or mariner-Himar1), construction of transposon insertion libraries in *V. cholerae* and other Gram-negative bacteria is facilitated by the availability of an in vivo transposition protocol mediated by the plasmid pDL1098 [5]. On the backbone of this plasmid are antibiotic resistance gene (*cat*; chloramphenicol resistance), mobilization origin of transfer (*oriT*), and a high-temperature-inducible, relaxed target-site specificity Tn10 transposase (*tnpA*) [7]. Also on the plasmid is a mini-Tn10 transposon (mTn10) harboring another antibiotic resistance gene (*aadA*; spectinomycin resistance). A second version of this plasmid was constructed, called pDL1093, which harbors *aphIII* encoding kanamycin resistance within the mTn10. In this system, replication of the plasmid and activity of the transposase are oppositely temperature-sensitive; at low temperatures (less than ~32 °C), the plasmid replicates, and the transposase are repressed, while at high temperatures (greater than ~38 °C), replication of the plasmid ceases, and the transposase gene is expressed. These temperature-sensitive controls allow maintenance of the plasmid without transposition at low temperature and with chloramphenicol selection. A simple shift of an early exponential phase broth culture to a high temperature and subsequent spectinomycin or kanamycin antibiotic selection selects for cells that have undergone a transposon insertion.

A library created in this manner can achieve a complexity of tens of thousands to hundreds of thousands of unique insertion sites in the genome [5]. A library of high complexity is in general desirable: statistical significance and confidence in results is more robust if multiple insertion site mutants of the same gene are analyzed. If only one or two insertion mutants of a particular gene are present in the input population, it is difficult to distinguish if under- or overrepresentation of that mutant in the output population is due to a change in fitness or to random fluctuations in the population. Conversely, if there are many insertions in the same gene and all or almost all behave similarly, it is reasonable to conclude that gene contributes to fitness in the selection.

Another important consideration in addition to library complexity is the bottleneck imposed by the selection. “Bottleneck” refers to the stochastic loss of a subset of the overall population, in contrast to the nonrandom loss of mutants with low fitness in the context of the selection. If the bottleneck is too constrictive, it is impossible to distinguish which mutants were unfit in the context of the selection from those that were lost on account of the bottleneck. This is especially significant during animal infection: while the infant rabbit model for cholera is supportive of highly complex libraries [8], the infant mouse model [9, 10] imposes a bottleneck such that only about  $10^2$ – $10^3$  unique mutants can be recovered from an infection.

Infection efficiency can be improved (and thus the bottleneck effect minimized) in the infant rabbit model by sparing the inoculum from the highly acidic environment of the stomach. Pretreatment of the kits with the proton pump inhibitor Ranitidine reduces stomach acidity, and suspension of the inoculum in a sodium bicarbonate buffer further protects the bacteria from low pH [11].

An inoculum of  $10^8$ – $10^9$  CFU prepared in 500  $\mu$ L buffer is optimal for intragastric infection of infant rabbits. This dose results in cholera symptoms, namely, secretory diarrhea, within

about 12 h after inoculation [11] and is also large enough that even a complex transposon library should be fully represented many times over. Animals are euthanized once symptomatic, at which point about 500  $\mu\text{L}$  of cecal fluid can be recovered. Rabbits have a large cecum, and luminal fluid collects there prior to release in the form of secretory diarrhea [11].

After a low-speed spin to pellet and remove clumps of mucus and host cells, the collected supernatant contains approximately  $10^9$  CFU/mL of planktonic, highly motile *V. cholerae*. This can be rapidly assessed by examining some of the collected supernatant by phase-contrast or dark-field microscopy. Because the number of *V. cholerae* recovered is at the lower limit for efficient genomic DNA isolation, outgrowth of the population is typically required. This is done in a rich broth medium supplemented with spectinomycin or kanamycin to limit growth of other bacteria. This outgrowth also allows for storage of backup cell pellets and cryopreserved glycerol stocks.

Libraries are prepared for sequencing by PCR-amplifying regions adjacent to transposon insertions, which can later be mapped to a reference genome to determine the location and frequency of each insertion [1–3, 5, 6]. Genomic DNA is first sheared by high-intensity sonication to lengths suitable for PCR amplification. Enzymatic methods (e.g., partial DNase I digestion) may also be used to fragment the DNA. While the known transposon sequence can be used as the forward priming site for amplification, the random genomic sequence following the transposon provides no such reverse priming site. To circumvent this, a tail is appended to the 3' ends of the sheared DNA using terminal deoxynucleotidyl transferase (TdT) and a 19:1 mixture of dCTP and the chain terminator ddCTP. This technique, called homopolymer tail-mediated ligation PCR, creates a poly-C tail approximately 20 nt in length as a recognition site for the reverse primer [4].

A second PCR is performed to barcode (index) the samples for Illumina multiplex sequencing using the product from the previous step as a template. Products from the second PCR step are mixed in the desired molar ratio, weighting the input sample most heavily, and this mixture is cleaned up and then ready for sequencing.

Tn-seq data is analyzed by mapping transposon junctions to the reference genome and comparing the number of reads at each insertion site in the input sample to the number in the output samples. Fitness scores are calculated for each insertion site and are then aggregated by gene [1–3]. Bioinformatic processing for Tn-seq analysis is explained in more detail in Subheading 3.4.

## 2 Materials

1. Water-cooled cuphorn sonicator.
2. Phosphate-buffered saline.
3. Luria-Bertani broth, Miller.
4. Antibiotics for selection: kanamycin, chloramphenicol.

5. 300 mM sodium bicarbonate buffer brought to pH 9.0 with 10 N NaOH (made fresh).
6. Ranitidine syrup 15 mg/mL: may require veterinary prescription.
7. Euthasol: requires veterinary prescription.
8. PE50 tubing: about 45 cm per unique inoculum, marked with a black Sharpie at 1, 2, 3, and 4.5 cm from the end, last 1–2 mm “softened” by pinching and pulling with fingernails; plus, one length of about 3 cm for administering Ranitidine.
9. 23 gauge needles and 1 mL syringes: one syringe + needle per unique inoculum, unmarked end of tubing pulled carefully onto needle and pulled to hilt; plus, two extra needles and syringes, one with 3 cm tubing pulled onto needle.
10. SYBR Safe DNA stain.
11. GelGreen DNA stain.
12. Edge Biosystems Performa DTR spin columns.
13. 19:1 dCTP:ddCTP, final concentration of 10 mM.
14. Easy-A cloning enzyme and buffer.
15. 10 mM dNTP mixture.
16. Terminal deoxynucleotidyl transferase (Promega) and buffer.
17. Genomic DNA extraction kit.
18. RNase A, 10 mg/mL stock solution.
19. PCR cleanup kit.
20. *Tn10*-specific PCR primer 1 at 30  $\mu$ M: 5' GTGTGG  
GCACTCGACATATGACAAG 3'.
21. C-tail-specific primer 1 at 30  $\mu$ M: 5' GTGACTGGAGTTCA  
GACGTGTGCTCTTCCGATCTGGGGGGGGGGGGGGGGGGGG 3'.
22. Nested *Tn10*-specific primer 2 at 30  $\mu$ M: this primer uses the *Tn10*-specific primer 1 sequence as a recognition site and introduces the P5 sequence (underlined) required for binding to the Illumina flow cell and cluster generation.  
5' AATGATACGGCGACCACCGAGATCTACACTCTTTGG  
GGGCCAAAATCATTAGGGGATTTCATCAG 3'.
23. C-tail-specific primer 2 at 30  $\mu$ M: this set of primers uses the C-tail primer 1 sequence as a recognition site and introduces the P7 sequence (underlined below) required for binding to the Illumina flow cell and cluster generation, an eight-base index that will be used to demultiplex libraries after sequencing, and a binding site for the Illumina Multiplexing Index Read Sequencing Primer (double underlined). One example of C-tail-specific primer 2 is 5'  
CAAGCAGAAGACGGCATAACGAGATAAAGGAATGTGACTGGAGTTCAG

ACGTGTGCTCTTCCGATCT 3' where the reverse complement of the index, 5' **ATTCCTTT** 3', is bolded.

24. Transposon junction sequencing primer at 30  $\mu$ M: 5' ACACTCTTTGGGGGCCAAAATCATTAGGGGATTCATCAG 3'.
25. Illumina Multiplexing Index Read Sequencing Primer at 30  $\mu$ M: 5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC 3'.

### 3 Methods

#### 3.1 Generation of a Transposon Library

1. Transform the in vivo transposition plasmid pDL1093 or pDL1098 into the target background strain by conjugation, electroporation, or natural transformation. The background strain must be able to grow at 30 °C and 40 °C and be sensitive to chloramphenicol and either kanamycin or spectinomycin. In *V. cholerae* and *Escherichia coli*, the plasmid can be maintained at 30 °C with chloramphenicol selection at 2.5  $\mu$ g/mL. The mTn10 confers either kanamycin resistance (pDL1093) or spectinomycin resistance (pDL1098), both at 100  $\mu$ g/mL. In our experience, selection with kanamycin is more stringent than with spectinomycin in *V. cholerae* and *E. coli*. Thus, the remainder of this protocol deals with pDL1093.
2. Grow the target strain with pDL1093 to stationary phase in Luria-Bertani Miller broth (LB) + 2.5  $\mu$ g/mL chloramphenicol at 30 °C with aeration. Pre-warm 100 mL LB with 100  $\mu$ g/mL kanamycin to 40 °C. Transfer 0.5 mL of the overnight culture to the pre-warmed LB, and grow with shaking at 40 °C overnight. At this high temperature, plasmid replication stops, and transposition is induced. Because transposition is inefficient, many of cells in the overnight culture will lack mTn10 insertions and may or may not still harbor the plasmid. Therefore, a second passage is usually required (next step).
3. Pre-warm an additional 100 mL LB 100  $\mu$ g/mL kanamycin to 40 °C, and transfer 100  $\mu$ L of the overnight culture into this fresh medium. Grow once again to stationary phase at 40 °C (~6–8 h) to reduce the presence of cells that may still harbor the plasmid and/or that lack mTn10 insertions. This culture is the final transposon library. A portion should be used to assay transposition frequency and presence of plasmid (**step 4**, below). In addition, the library should be frozen with 20% glycerol (v/v) at –70 °C in 1 mL aliquots for future use. It is prudent to store many (e.g., 50–100) single-use aliquots. Future amplification of the library to generate additional archived aliquots is not recommended due to skewing of strain representation.
4. To assay transposition efficiency and determine the percentage of cells in the final library that still harbor pDL1093, plate serial dilutions on LB agar plates grown at 30 °C (to enumerate all cells), plates with 100  $\mu$ g/mL kanamycin grown at 40 °C (to enumerate cells with transposon insertions), and plates with both 2.5  $\mu$ g/mL chloramphenicol and 100  $\mu$ g/mL kanamycin grown at 30 °C (to

enumerate cells still harboring the plasmid). A small percentage of cells still harboring the plasmid is tolerable as long as most of the cells in the library contain mTn10 insertions.

### 3.2 Infection of 2- to 3-Day-Old Infant Rabbits and Recovery of Output Library

1. Thaw an aliquot of the *V. cholerae* mTn10 library, and make several dilutions (e.g., 1/20 and 1/100) into 15 mL LB + 100 µg/mL kanamycin, and grow at 37 °C with aeration for several hours to stationary phase. Inoculate with a sufficient volume of the thawed library to ensure representation of the entire library. For a library of 10<sup>5</sup> complexity, i.e., 10<sup>5</sup> unique mTn10 insertion strains, 10<sup>8</sup> CFU or 0.1 mL of the thawed library is sufficient to avoid a bottleneck. Dispose of the rest of the thawed aliquot properly; i.e., do not re-freeze thawed library glycerol stocks (*see* Note 1).
2. 3.5 h before the desired time of inoculation, remove the infant rabbits (kits) from their mother in a container tall enough to prevent them from escaping. Be sure to use a container that allows for airflow. All procedures with the rabbits require Institutional Animal Care and Use Committee approval.
3. Dilute Ranitidine syrup to 1.5 mg/mL with sterile water.
4. Number each kit on the back with a permanent marker and weigh the animal.
5. Using a 1 mL syringe fitted with a 23 gauge needle and about 3 cm of PE50 tubing such that at least 0.5 cm overhangs the end of the needle, administer 5 µg Ranitidine per gram of bodyweight directly into each kit's mouth 3 h before inoculation time (*see* Note 2). This is typically ~0.1 mL.
6. Place kits back in container, and keep at room-temperature (~24 °C) for 3 h, allowing for airflow. Do not place kits back with their dam, as she will often reject and may harm them.
7. Before infecting, prepare the bulk inoculum: pellet the library culture for 2 min at 8000 RCF, remove the supernatant, and resuspend cells to a final concentration of ~10<sup>9</sup> CFU/mL in freshly made sterile 300 mM sodium bicarbonate. Prepare a volume such that each kit can receive 0.5 mL of the inoculum with at least 4 mL extra. Pellet 1 mL of inoculum in duplicate, remove the supernatant, and store the cell pellets at -20 °C for later genomic DNA purification. In addition, temporarily store 1 mL of inoculum in 20% glycerol at -70 °C.
8. Inoculation of kits is ideally done with two people working together. Draw 1 mL of the inoculum into a 1 mL syringe. Place one of the 23 gauge needles with a long section of PE50 tubing onto the syringe (*see* Subheading 2, **step 8**). To account for the void volume in the tubing, place the end of the tube into the remaining inoculum, and depress the plunger just until fluid reaches the end of the tube. Keeping the end of the tubing submerged in the inoculum, draw the plunger back up to the 1 mL mark (*see* Note 3).

9. Intragastrically inoculate each kit: one person will insert the tubing into the mouth and gently feed more and more length down the esophagus, using the markings on the tube as a guide and typically inserting until the last marking is about even with the front teeth (*see* Note 4). There should not be significant resistance when inserting the tubing; if there is, withdraw the tubing, and try once more. Once the end of the tubing is in the stomach, the other person should slowly eject 500  $\mu$ L. If backpressure is felt, the tubing may have kinked and increasing the pressure further could be dangerous in that it may cause the tubing to fly off the syringe and needle. Therefore, if backpressure is felt, withdraw the tubing from the kit and insert once more. After inoculation, gently remove the tubing and inspect for any signs of trauma such as blood in the mouth or blood on the withdrawn tubing. If blood is observed, then the esophagus was likely damaged during the procedure and the animal must be euthanized, since bacteremia is likely.
10. Place animals back in the container; if more than one distinct inoculum is used, group animals by inoculum received to prevent contamination by secondary infection once the kits are symptomatic. Place at room temperature ( $\sim 24$   $^{\circ}$ C) with good airflow. Since the kits will become symptomatic after  $\sim 8$  h, adsorbent bedding material should be placed at the bottom of the container.
11. Begin monitoring kits periodically for cholera symptoms and weight loss. Kits typically become symptomatic about 8–12 h postinoculation, but this varies with *V. cholerae* strain, inoculum size, and status of the kits. Kits should be euthanized after becoming symptomatic and before losing more than 10% of their starting body weight. Kits should also be euthanized if they stop moving freely (paresis) and/or are exhibit abnormal breathing (tachypnea).
12. Euthanize kits via CO<sub>2</sub> asphyxiation until unconscious, followed by intracardiac injection of 200  $\mu$ g Euthasol per gram body weight using a 23 gauge needle and 1 mL syringe.
13. Sanitize the abdomen surface prior to dissection with 70% ethanol. Upon dissection of the abdomen, the cecum should be readily apparent on the animal's right side and swollen with fluid. The small intestine may also be swollen with fluid. Place a sterile container under the cecum (*see* Note 5), which has been gently pulled out to the side of the animal with forceps, and carefully cut open one side of cecum to allow liquid to drain into the container. Collect as much cecal fluid as possible, typically 500–800  $\mu$ L, and transfer to a microcentrifuge tube.
14. Vortex the cecal fluid twice for 5 s to break up aggregates of bacteria. Reserve a small aliquot, 2–10  $\mu$ L, of the cecal fluid prior to any processing, and use this to titer the bacterial burden by serial dilution and plating on LB agar + 100  $\mu$ g/mL kanamycin (*see* Note 6).
15. Unprocessed cecal fluid does not contain enough bacteria to allow adequate genomic DNA recovery for sequencing and must therefore be outgrown. First,



pellet mucus and eukaryotic cells by centrifugation for 2 min at 500 RCF. Transfer all of the supernatant to a new microcentrifuge tube. Outgrow in an appropriate volume of LB with 100 µg/mL kanamycin. After the culture has grown to stationary phase, store cell pellets and glycerol stocks as described above for the inoculum.

### 3.3 Library Preparation for Sequencing

1. Purify genomic DNA from individual cell pellets of each sample, including the inoculum (input) and each animal output. Each cell pellet should contain  $\sim 10^9$  CFU. Include an RNase A step to eliminate contaminating RNA. Measure the DNA concentration by spectrophotometry. Genomic DNA should be at a concentration of at least 50 ng/µL in a minimum volume of 50 µL. Check genomic DNA quality and quantity on a 0.8% agarose gel with SYBR Safe DNA stain.
2. Shear 1 µg of genomic DNA in a 50 µL volume in a PCR tube (or 2 µg in 100 µL in a parabolic-bottom 2 mL microcentrifuge tube) in a water-cooled cuphorn sonicator. Sonicate for 1 min at maximum amplitude, with a 5 s on/5 s off duty cycle. Pellet contents and repeat sonication.
3. Confirm shearing to a range of  $\sim 200$ –800 bp by running a small aliquot of each sample on a 2% agarose gel with GelGreen DNA stain. If necessary, repeat the sonication.
4. Remove small molecules from the sheared DNA by passing through an Edge Biosystems Performa DTR spin column (gel filtration). Quantify the concentration of DNA by spectrophotometry.
5. Append a poly-C tail to the sheared DNA by mixing 0.1 µg sheared DNA, 4 µL TdT reaction buffer, 1 µL 10 mM 19:1 dCTP/ddCTP, 0.5 µL TdT enzyme, and water to a volume of 20 µL. Incubate at 37 °C for 1 h, then heat-inactivate the enzyme at 75 °C for 20 min. Pass through an Edge Biosystems column once more.
6. PCR amplifies and indexes transposon-genomic junctions using 10 µL of the cleaned C-tailing reaction as template. Primers are listed in Table 1 below.
  - a. First PCR:
    - 10 µL C-tailed template.
    - 2 µL 10 mM dNTPs.
    - 5 µL Easy-A Cloning reaction buffer.
    - 1 µL mTn *I/O*-specific primer 1.
    - 3 µL C-tail-specific primer 1.
    - 1 µL Easy-A Cloning enzyme.
    - Water to 50 µL.



Begin the PCR program by heating at 95 °C for 2 min, then 25 cycles of 95 °C for 30 s, 58 °C for 30 s, and 72 °C for 2 min. Finish with another 2 min of extension time, and store at 10 °C or colder.

**b.** Second PCR:

0.5 µL of first PCR.

2 µL 10 mM dNTPs.

5 µL Easy-A Cloning reaction buffer.

1 µL Nested mTn10-specific primer 2.

1 µL C-tail-specific primer 2 (use a different index [“Ng” in Table 1 below] for each sample).

1 µL Easy-A Cloning enzyme.

Water to 50 µL.

Begin the PCR program by heating at 95 °C for 2 min, then 18 cycles of 95 °C for 30 s, 52 °C for 30 s, and 72 °C for 2 min.

Finish with a final extension of 2 min, and store at 10 °C or colder.

7. Visualize 2 and 8 µL of each reaction on a 2% agarose gel with GelGreen DNA stain. Load the 2 µL amounts in adjacent lanes, and the 8 µL amounts in adjacent lanes for later visual comparison. There should be a smear of products from about 200 to 1000 bp. Compare the different lanes by eye (*see* Note 7) to estimate the relative amounts of DNA in each: this information is used below for pooling at roughly equimolar amounts. An example gel analysis of HTML-PCR samples (5 µL was loaded) is shown in Fig. 1.
8. Pool a portion of the second PCR reactions to give an equimolar amount of each output and with the input at four times the molar concentration of each output. This ensures good coverage of the input, which is essential for analysis. Depending on library complexity, up to 24 libraries can be multiplexed in a single Illumina sequencing lane in this manner. Within a single multiplexed mixture, each index primer must be unique. After pooling, temporarily store the remaining second PCR reaction products at –20 °C.
9. Clean the pooled sample with a PCR purification kit to remove residual primers and other small molecules. Quantify the concentration of DNA by spectrophotometry. Also visualize a small aliquot of the mixture on a 2% agarose gel with GelGreen DNA stain to ensure quality. The final DNA sample should be free of primers and primer-dimers, as these will interfere with sequencing. In some instances, gel purification of the 200–800 bp smear of products is required to remove primers and/or primer-dimers. Submit for Illumina single-end 50 cycle sequencing along with an aliquot of the custom transposon junction sequencing primer and information about the Illumina indexes used for multiplexing. You may also want to provide an aliquot of the Illumina

Multiplexing Index Read Sequencing Primer; however, most Illumina sequencing core facilities have this primer.

### 3.4 Data Processing and Analysis

1. Map each file of demultiplexed sample reads to the reference genome using Bowtie. The poly-C tail will appear at the 3' end of short junctional sequences and must be clipped bioinformatically for the reads to map properly.
2. For each sample, normalize the number of readers at each insertion site to the total number of reads from that sample.
3. Calculate fitness scores using the normalized mapped reads. The input sample is T1, and each output sample is T2. This calculation requires an expansion factor  $d$ , which represents the absolute value of the fold change in population size over the experiment, either calculated experimentally or estimated (*see Note 8*). The

fitness calculation is as follows:  $W_i = \frac{\ln\left(N_i(t_2) \times \frac{d}{N_i(t_1)}\right)}{\ln\left((1 - N_i(t_2)) \times \frac{d}{1 - N_i(t_1)}\right)}$  where  $W_i$  is the

fitness value for insertion  $I$ , and  $N_i(t_1)$  and  $N_i(t_2)$  are the frequencies of the insertion at T1 and T2, respectively. Set the cutoff for number of reads per insertion at T1 to 15. This is necessary because the high sequencing error rate of the Illumina platform will generate a low number of reads that map to incorrect positions in the reference genome. In addition, set the program to report “zero” fitness scores, which are insertions that were present in the input sample but absent from the output sample.

4. Normalize fitness scores to known neutral genes. Using a list of 15–30 genes with known neutral fitness (e.g., defective genes, transposases, integron cassette genes but excluding toxin/antitoxin genes), calculate the multiplicative factor required to set the average fitness of all neutral genes in the fitness scores file to 1 (*see Note 8*). If this factor is itself not close to 1, then either the genes in the neutral list are not truly neutral or something went wrong during the selection such as a severe bottleneck. Normalize all the fitness scores by this factor. This step also determines the severity of the bottleneck by calculating the number of insertions in the neutral genes present in the input that are absent from the output.
5. Finally, aggregate insertional fitness scores by gene (*see Note 9*). Randomly exclude a proportion of zero fitness scores equal to the bottleneck calculation in **step 4** to correct for artificially low fitness scores due to the bottleneck. Calculate the mean fitness for each gene based on the fitness of all insertions in that gene. The aggregate fitness scores can be manipulated and sorted as desired, and statistics can be calculated if there are experimental replicates.

## 4 Notes

1. It is a good idea to start cultures at several dilutions to ensure the correct density at time of infection. Make sure the culture volume is sufficiently large for inoculation of animals, storing of cell pellets for genomic DNA isolate, and freezing of one aliquot in glycerol.
2. Place toward the back of the mouth, and the kits will swallow the liquid. Too close to the front of the mouth and they will try to spit it out.
3. The same needle/tubing can be used for several animals; switching to a new tube is necessary only to prevent contamination when using a different inoculum or if the tubing becomes damaged.
4. If the animals are particularly small or large, this length may have to be altered to fit the animal; in this case, estimate the correct length of tubing to insert by holding a (clean) length of PE50 tubing next to the animal and marking as appropriate.
5. We use the lid of a petri plate, which is sterile and shallow enough to place below the cecum. Tilt the lid so that the fluid collects in one pool to facilitate transfer to a microcentrifuge tube.
6. If an estimate of the titer is needed before processing, observe a 5  $\mu\text{L}$  aliquot by microscopy, and compare this to an in vitro- grown culture of a known titer. As a rule of thumb, cecal fluid from symptomatic kits typically contains about  $10^9$  *V. cholerae* per mL.
7. If desired, the molar amounts of DNA in each lane can be quantified by image analysis. Most gel imaging programs have quantification built in; the free software ImageJ also works well. Ensuring that no pixels are oversaturated, compare the DNA quantity in each lane to the known quantity of ladder loaded. Convert from mass to a molar quantity by estimating the peak size in base pairs of the smear and using the following formula:

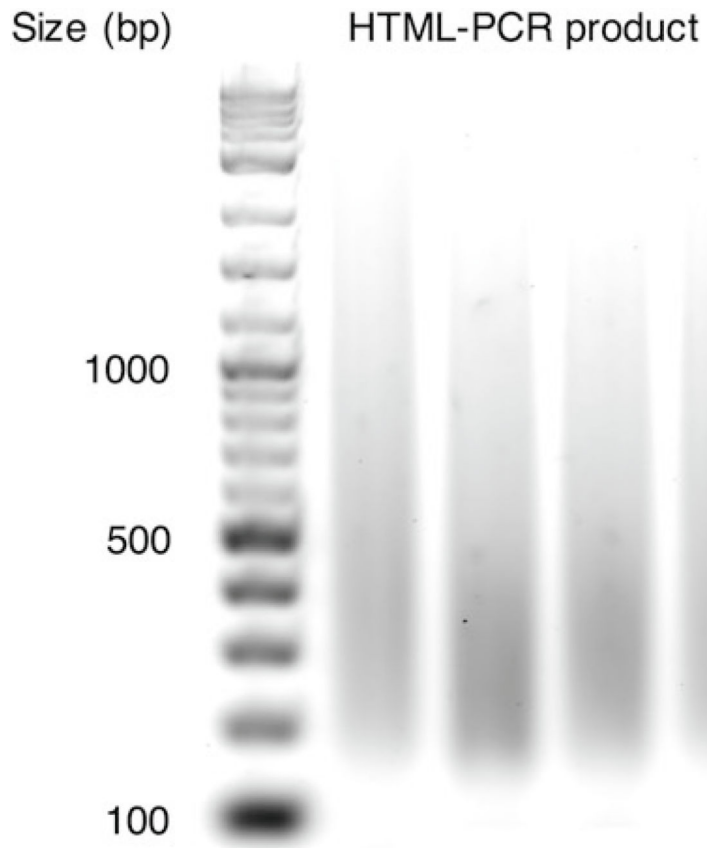
$$\mu\text{g DNA} \times \frac{\text{pmol}}{660 \text{ pg}} \times \frac{10^6 \text{ pg}}{1 \mu\text{g}} \times \frac{1}{\text{Length in base pairs}} = \text{pmol DNA}$$

8. In the case where there is not a neutral gene list available or the expansion factor is unknown, an alternative analysis protocol called HopCount may be used. The method for HopCount analysis is used and described in references 5 and 6.
9. Typically, we exclude insertions from the last 5% of the coding sequence, as these are less likely to truly “knock out” or inactivate the gene they are inserted into.

## References

1. van Opijnen T, Bodi KL, Camilli A (2009) Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat Methods* 6:767–772. 10.1038/nmeth.1377 [PubMed: 19767758]

2. van Opijnen T, Camilli A (2013) Transposon insertion sequencing: a new tool for systems- level analysis of microorganisms. *Nat Rev Microbiol* 11:435–442. 10.1038/nrmicro3033 [PubMed: 23712350]
3. van Opijnen T, Lazinski DW, Camilli A (2014) Genome-wide fitness and genetic interactions determined by Tn-seq, a high-throughput massively parallel sequencing method for microorganisms. *Curr Protoc Mol Biol* 106:7.16.1–7.1624. 10.1002/0471142727.mb0716s106 [PubMed: 24733243]
4. Lazinski DW, Camilli A (2013) Homopolymer tail-mediated ligation PCR: a streamlined and highly efficient method for DNA cloning and library construction. *BioTechniques* 54:25–34. 10.2144/000113981 [PubMed: 23311318]
5. McDonough E, Lazinski DW, Camilli A (2014) Identification of in vivo regulators of the *Vibrio cholerae* xds gene using a high-throughput genetic selection. *Mol Microbiol* 92:302–315. 10.1111/mmi.12557 [PubMed: 24673931]
6. Klein BA, Tenorio EL, Lazinski DW et al. (2012) Identification of essential genes of the periodontal pathogen *Porphyromonas gingivalis*. *BMC Genomics* 13:578 10.1186/1471-2164-13-578 [PubMed: 23114059]
7. Bender J, Kleckner N (1992) Tn10 insertion specificity is strongly dependent upon sequences immediately adjacent to the target-site consensus sequence. *Proc Natl Acad Sci U S A* 89:7996–8000 [PubMed: 1325639]
8. Kamp HD, Patimalla-Dipali B, Lazinski DW et al. (2013) Gene fitness landscapes of *Vibrio cholerae* at important stages of its life cycle. *PLoS Pathog* 9:e1003800 10.1371/journal.ppat.1003800 [PubMed: 24385900]
9. Ujjiye A, Nakatomi M, Utsunomiya A et al. (1968) Experimental cholera in mice: I. First report on the oral infection. *Trop Med* 10:65–71
10. Klose KE (2000) The suckling mouse model of cholera. *Trends Microbiol* 8:189–191. 10.1016/S0966-842X(00)01721-2 [PubMed: 10754579]
11. Ritchie JM, Rui H, Bronson RT, Waldor MK (2010) Back to the future: studying cholera pathogenesis using infant rabbits. *mBio* 1:e00047–10. 10.1128/mBio.00047-10 [PubMed: 20689747]



**Fig. 1.**  
An example of gel analysis of HTML-PCR samples

