

Protocol 1

High-Throughput Mutant Screening in *Vibrio cholerae* via Transposon Sequencing

Jacob Bourgeois and Andrew Camilli¹

Department of Molecular Biology and Microbiology, Tufts University, School of Medicine, Boston, Massachusetts 02067, USA

Transposon mutagenesis greatly facilitates the study of gene function in microorganisms ranging from viruses to fungi. Traditionally, one would study individual transposon mutants with interesting phenotypes one mutant at a time. Here, we describe methods for the study of tens of thousands of transposon mutants in parallel in the bacterial pathogen *Vibrio cholerae* using transposon-sequencing. The first section outlines methods for making a saturated transposon mutant library. The second section outlines methods for massively parallel sequencing of the transposon junctions. The third section outlines methods for analyzing the sequence data to calculate the fitness contribution of genes.

MATERIALS

It is essential that you consult the appropriate Material Safety Data Sheets and your institution's Environmental Health and Safety Office for proper handling of equipment and hazardous materials used in this protocol.

RECIPES: Please see the end of this protocol for recipes indicated by <R>. Additional recipes can be found online at <http://cshprotocols.cshlp.org/site/recipes>.

Reagents

Reagents required for making a saturated mini-transposon library (see Steps 1–21)

1000× kanamycin stock <R>

Chitin flakes (MilliporeSigma C9213)

Chloramphenicol stock solution (10 mg/mL) <R>

DNeasy Blood & Tissue Kit (QIAGEN 69504)

Glycerol (80% [v/v]; Fisher Scientific G33)

Isopropyl-β-D-thiogalactoside (IPTG; GoldBio I2481C)

LB agar <R>

Prepare LB agar with 50 μg/mL kanamycin and 2 μg/mL chloramphenicol using 1000× kanamycin stock and chloramphenicol stock solution. The plates may be stored for up to 1 mo at 4°C.

LB (Luria–Bertani) liquid medium <R>

Prepare LB as described in the recipe.

¹Correspondence: andrew.camilli@tufts.edu

Transposon Sequencing Genetic Screening in Bacteria

Prepare LB with freshly added 50 µg/mL kanamycin and 2 µg/mL chloramphenicol using 1000× kanamycin stock and chloramphenicol stock solution.

Prepare LB with freshly added 100 µg/mL kanamycin using 1000× kanamycin stock.

Lysozyme (5%; MilliporeSigma L6876) in AE buffer (QIAGEN)

Warm to 37°C and vortex gently until dissolved. Store at –20°C in 1-mL aliquots. Aliquots are stable for up to 5 yr at –20°C. Aliquots can be freeze–thawed up to four times before becoming inactive.

RNase A (20 µg/mL; Fisher EO0491)

Sea salts (0.7% [w/v]; MilliporeSigma S9883)

Strain AC5540 (*Vibrio cholerae* harboring plasmid pDL1093 [Duncan et al. 2018] in a cryopreserved stock vial)

Reagents required for transposon junction sequencing (see Steps 22–54)

AMPure XP magnetic beads (Beckman Coulter A63880)

Deoxy CTP (dCTP; TriLink BioTechnologies N2511)

Deoxynucleotide (dNTP) Solution Mix (New England Biolabs N0047)

Dideoxy CTP (ddCTP; TriLink BioTechnologies N4005)

DNA oligonucleotides in Table 1

Order primers of 50 nt or less unpurified from Integrated DNA Technologies. However, order those 51 nt and larger to be PAGE-purified from Integrated DNA Technologies.

DNA size markers (ladder; New England Biolabs N3200)

Easy-A Hi-Fi Cloning Enzyme (Agilent Technologies)

Easy-A Hi-Fi Cloning Enzyme, Reaction buffer (10×; Agilent Technologies 600402)

Ethanol (70% [v/v]), freshly prepared

GelGreen DNA Stain (Fisher Scientific 41004)

Gel loading buffer (10×) <R>

Molten agarose <R>

Prepare 1.5% molten agarose.

QIAquick PCR Purification Kit (QIAGEN 28104)

Quick Blunting Kit (New England Biolabs E1201L)

TAE <R>

Terminal deoxynucleotidyl transferase (TdT), 5× Reaction buffer (Promega M828C)

Terminal deoxynucleotidyl transferase (TdT) enzyme (Promega)

TABLE 1. DNA oligonucleotide list

Name	Sequence (5' to 3') ^a	Stock
olj363	GTGTGGGCACTCGACATATGACAAG	30 µM
olj376	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGGGGGGGGGGGGGGGGG	30 µM
olj385	AATGATACGGCGACCACCGAGATCTACACTCTTTGGGGGCCAAAATCATTAGGGGATTCATCAG	30 µM
BC45	CAAGCAGAAGACGGCATAACGAGATTGTTGACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	30 µM
BC46	CAAGCAGAAGACGGCATAACGAGATACCGAACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	30 µM
BC47	CAAGCAGAAGACGGCATAACGAGATCTGACATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	30 µM
BC48	CAAGCAGAAGACGGCATAACGAGATCGGGACGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	30 µM
BC49	CAAGCAGAAGACGGCATAACGAGATGTGCGGACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	30 µM
BC50	CAAGCAGAAGACGGCATAACGAGATCGTTTCACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	30 µM
BC51	CAAGCAGAAGACGGCATAACGAGATAAGGCCACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	30 µM
BC52	CAAGCAGAAGACGGCATAACGAGATACCGAAACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	30 µM
BC53	CAAGCAGAAGACGGCATAACGAGATTACGTACGGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	30 µM
olj386	ACACTCTTGGGGGCCAAAATCATTAGGGGATTCATCAG	30 µM

^aThe underlined sequences in BC45–B53 are the inverse-complement of the index 1 (i7) index sequence determined by sequencing on the Illumina NextSeq instrument.

Section 4

Equipment

Agarose gel casting tray and comb
Bowtie2
Branson Sonifier S-450 Cell Disruptor, high-intensity version (Branson Ultrasonics 101147046)
Computer
Conical tubes (15-mL, VWR 89039-666; 50-mL, VWR 89039-660)
Cryovials (1.8-mL; Nunc 377267)
cutadapt.py
Erlenmeyer flasks (sterile, 50-mL and 250-mL)
Fisherbrand Isotemp Refrigerated/Heated Bath Circulators at 37°C (Fisher Scientific 13-874-180) or beaker of H₂O at 37°C (see Step 12)
Freezer at -70°C
Gel electrophoresis chamber (horizontal) and power supply
Genome browser (e.g., Integrated Genome Browser [IGB]; Broad Institute)
GraphPad Prism (optional; see Step 72)
HopCount with HopCount.py
Image-analysis software (see Step 52)
Incubator (static) at 30°C
Incubator-shaker(s) at 30°C and 40°C
Magnetic rack for 0.2-mL PCR tubes (Fisher Scientific NC1831102)
Microcentrifuge tubes (1.5-mL)
Microcentrifuge with adaptor for PCR tubes
Microsoft Excel
Paper towels
PCR tubes (0.2-mL)
Platinum wire loop, wooden stick, or sterile 1000- μ L pipette tip (see Step 1)
Python3
SAMtools
Text editor
Thermocycler
UV or blue light box (see Step 25.v)
UV spectrophotometer
Vortexer

METHOD

Three procedures are described here: making a saturated mini-transposon library (Steps 1–21), transposon junction sequencing (Steps 22–54), and analyzing transposon-sequencing (Tn-seq) data (Steps 55–83). The first procedure has been optimized for constructing transposon insertion libraries in Gram-negative bacteria, using *V. cholerae* as an example. The remaining procedures are applicable to any species of microbe in which a suitable transposon insertion library has been made. See *Introduction: High-Throughput Mutant Screening via Transposon Sequencing* (Bourgeois and Camilli 2023) and the Method section to determine the best procedures for the experiment.

Making a Saturated Mini-Transposon Library

Grow an In Vivo Transposition Strain under Noninducing Conditions (2 d)

1. Streak strain AC5540 (*V. cholerae* harboring plasmid pDL1093) using a sterile 1000- μ L pipette tip, platinum wire loop, or wooden stick from a cryopreserved stock vial onto an LB agar plate supplemented with 50 μ g/mL kanamycin sulfate and 2 μ g/mL chloramphenicol. Incubate for 30 h at 30°C to obtain single colonies.

Plasmid pDL1093 (Fig. 1) is used to construct transposon insertion libraries via in vivo transposition in Gram-negative bacteria such as V. cholerae and Escherichia coli (Duncan et al. 2018). It harbors a mini-Tn10

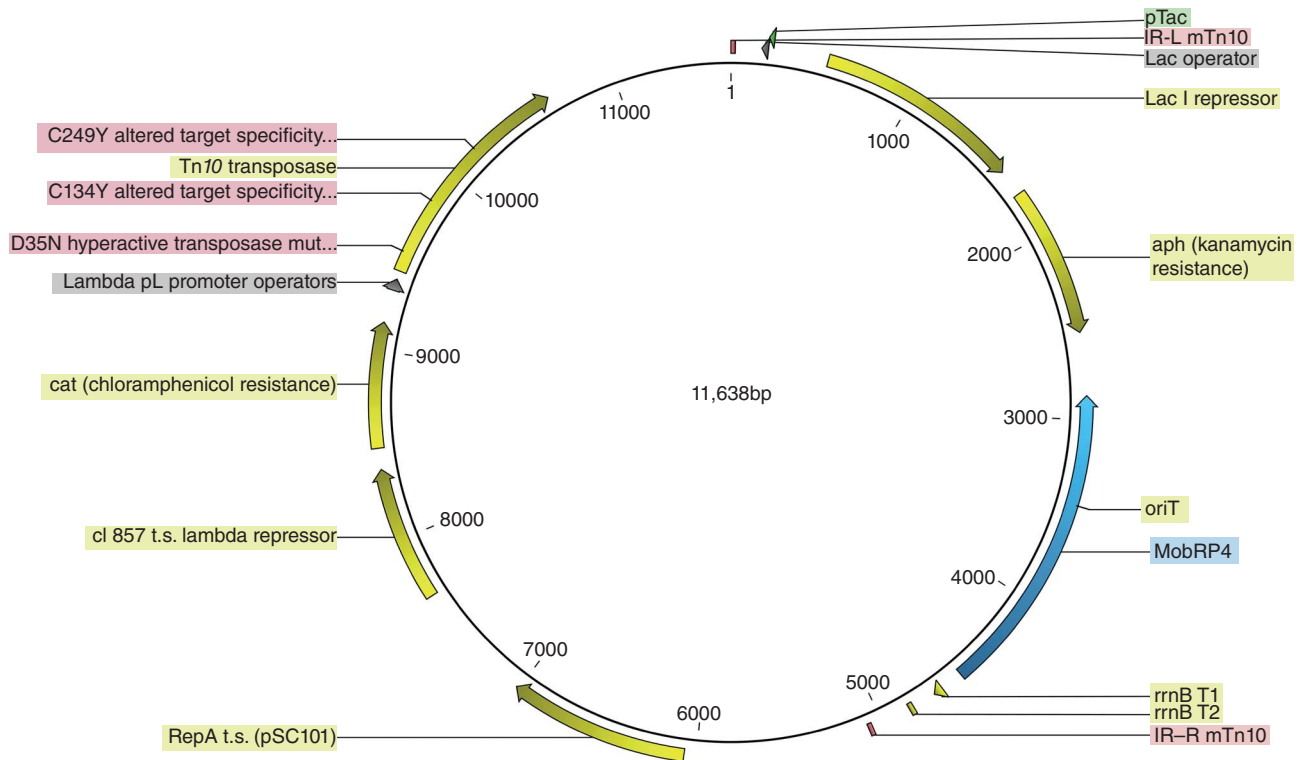


FIGURE 1. Illustration of plasmid pDL1093 used for in vivo transposition in Gram-negative bacteria.

(*mTn10*) containing an *aph* kanamycin-resistance (Km^R) gene, the *lacIq* transcriptional repressor, an outward-reading P_{tac} promoter at the left end, a mobilization region (*mobRP4* and *oriT*), the *oriR6K* origin of replication, and tandem transcriptional terminators *rrnB T1* and *rrnB T2* that prevent any outward-reading transcription from the right end. The mobilization and *oriR6K* features are not used in this protocol. The *mTn10* extends from the inverted repeat-left (*IR-left*) to the inverted repeat-right (*IR-right*), which are the sequences recognized by the *Tn10* transposase (*tnpA*) that is encoded outside the *mTn10*. Also on the plasmid outside the *mTn10* are a *cat* chloramphenicol-resistance (Cm^R) gene and *repA* ts, a temperature-sensitive replication gene. Transcription of *tnpA* is repressed by the phage Lambda temperature-sensitive repressor *cl 857* ts. Induction (derepression) of transposition is triggered by shifting the growth temperature to 40°C, which, at the same time, stops plasmid replication.

2. Pool approximately 20 colonies and use them to inoculate 30 mL of LB broth supplemented with 50 µg/mL kanamycin sulfate and 2 µg/mL chloramphenicol in a 50-mL Erlenmeyer flask. Incubate overnight at 30°C in an incubator-shaker at 200 rpm.

The colonies should be small in size because of the low temperature of growth and presence of dual antibiotics. Pooling many colonies in this way reduces the impact of possible jackpot transposition events on the final library complexity, in which a jackpot event is defined as a spurious transposition during growth of a colony.

Induce Transposition (1 d)

3. Add 100 mL of LB broth supplemented with 100 µg/mL kanamycin sulfate to each of four 250-mL Erlenmeyer flasks.
4. Vigorously swirl the overnight culture for 10 sec to disperse any clumps. Add 20 µL to each of the four flasks containing LB broth supplemented with kanamycin sulfate. Incubate the flasks for 8 h or overnight at 40°C in an incubator-shaker at 200 rpm.

Overnight incubation ensures complete growth. The elevated temperature of 40°C induces transposition.

5. Add 100 mL of LB broth supplemented with 100 µg/mL kanamycin sulfate to each of four new 250-mL Erlenmeyer flasks.

Section 4

6. Vigorously swirl each flask from Step 4 to break up clumps. Add 0.25 mL to the new flasks containing LB broth supplemented with kanamycin sulfate. Incubate the flasks for 6 h or overnight at 40°C in an incubator-shaker at 200 rpm.

Overnight incubation ensures complete growth. This second passage ensures the elimination of cells lacking mTn10 insertions in the genome. Each flask will yield independent transposon insertion mutants and will be combined in the end to make one saturating library.

7. Add 8 mL of 80% glycerol to a 50-mL conical tube.
8. Vigorously swirl each flask. Pool 7.5 mL of each culture into the 50-mL conical tube containing glycerol such that aliquots from all four cultures are mixed together. Vigorously vortex to mix well.

The final concentration of glycerol is 16.8% (v/v).

9. Make 38 1-mL aliquots in cryovials.
10. Freeze the aliquots at -70°C or colder to cryopreserve for future use.

The procedure above typically yields a mTn10 insertion library with a complexity of approximately 100,000 unique insertions. Such a complex library is saturating, meaning that mTn10 insertions can be found in virtually all nonessential genes.

Passage of the mTn10 Insertion Library in the Selection Condition (1 d)

It is important to avoid a bottleneck during the selection, in which a bottleneck is defined as the death of >20% of the bacterial population. Such a bottleneck will negatively impact the results because of the stochastic loss of mTn10 insertion strains irrespective of their fitness. If substantial death of the population cannot be avoided, then the experiment should be scaled up to avoid stochastic loss of mTn10 strains.

11. Prepare eight 250-mL Erlenmeyer flasks of medium, autoclaving before adding chitin or IPTG (filter-sterilized): two flasks containing 100 mL of LB broth, two flasks containing 100 mL of LB broth supplemented with 0.01 mM IPTG, two flasks containing 100 mL of 0.7% sea salts supplemented with 1 g of chitin flakes, and two flasks containing 100 mL of 0.7% sea salts supplemented with 1 g of chitin flakes and 0.01 mM IPTG.
12. Rapidly thaw a cryovial of the mTn10 library in a 37°C H₂O bath or beaker of warm (37°C) H₂O. Invert the tube several times to mix.
13. Add 0.1 mL to each of the eight 250-mL Erlenmeyer flasks. Incubate the four LB broth cultures overnight at 30°C in an incubator-shaker at 200 rpm. Incubate the four chitin cultures for 30 h at 30°C under static conditions (i.e., no shaking). However, swirl the four chitin flasks by hand at the end of the day and first thing the next morning.

The 0.7% sea salts supplemented with chitin mimic estuarine environments that V. cholerae inhabits. The cultures supplemented with IPTG will experience overexpression of genes oriented in the sense direction downstream from the mTn10 P_{tac} promoter, as well as underexpression of genes oriented in the antisense direction. Each culture condition is done in biological duplicate so that a correlation test can be done on the final gene fitness data to examine reproducibility of the results. The results from the chitin selection will be compared to that from the LB broth selection to determine which genes are conditionally important in the chitin growth condition.

14. Vigorously swirl each overnight flask. Transfer 5 mL of each to a 15-mL conical tube containing 0.67 mL of 80% glycerol. Vortex vigorously to mix well.

The final concentration of glycerol is 20% (v/v).

15. Prepare five 1-mL aliquots of each culture in cryovials.
16. Freeze the aliquots at -70°C or colder to cryopreserve.

Purification of Genomic DNA: Lyse Cells (2 h)

The procedure below uses reagents and materials in the DNeasy Blood & Tissue Kit (QIAGEN). Kits from other manufacturers and laboratory procedures are available for purifying DNA from bacteria, such as the Quick-DNA Microprep Kit (Zymo Research D3021).

17. Thaw one cryovial of the input mTn10 library and one cryovial from each of the eight outputs. Invert each tube several times to mix. Dispense 0.5 mL into 1.5-mL microcentrifuge tubes.
18. Pellet the cells at 15,000g for 2 min at room temperature in a microcentrifuge.
19. Remove each supernatant using a 1000- μ L pipette tip. Fully resuspend each cell pellet in 40 μ L of AE buffer (QIAGEN) using a 200- μ L pipette tip.

It is important that each cell pellet is fully resuspended to ensure that the genomic DNA is liberated from all cells in Steps 20 and 21. If the DNA of only a fraction of the cells is liberated, this constitutes a bottleneck, which will negatively impact the results.
20. Add 10 μ L of 5% lysozyme in AE buffer and immediately vortex. Incubate for 30 min at 37°C.
21. Continue to purify genomic DNA with the use of RNase A as described in Steps 17–26 of Section 2, Protocol 1: Isolation and Sequencing of Novel *Vibrio* Species (Camilli 2022).

Transposon Junction Sequencing

Generate a Sequencing-Ready Sample (4 h)

To perform massively parallel sequencing (MPS) of the mTn10–chromosome junctions, user-defined sequences are appended upstream and downstream from each junction using the homopolymer tail-mediated PCR method. Nested PCR primers *olj363* and *olj385* and sequencing primer *olj386* (Table 1) are specific to the mTn10 used in this protocol. These must be replaced with appropriate primers if using another mini-transposon or insertional element.

22. Place 2 μ g of each input and output genomic DNA sample in a microcentrifuge tube and bring the volume to 70 μ L with pure H₂O. Chill to 4°C.
23. Shear the DNA to a size range of 200–800 bp using a high-intensity cup-horn sonicator: 1 min at 50% amplitude with a duty cycle of 5 sec on/5 sec off.
24. Briefly centrifuge at 3000g for 5 sec at room temperature to move the contents to the bottom of the tube, and then repeat the sonication.

The cold H₂O flowing through the sonication chamber is required to prevent heating of the sample and denaturation of the dsDNA fragments. If a high-intensity sonicator is not available, limited digestion with DNase I can be used instead.
25. Perform agarose gel electrophoresis to confirm shearing of the DNA to the desired size range.
 - i. Pour a 1.5% agarose gel in a casting tray. Immediately add 1/10,000 volume GelGreen DNA Stain using a pipette and mix throughout the molten agarose by stirring with a 1000- μ L pipette tip. Place a comb in the gel, and then let the agarose solidify and cool at room temperature.
 - ii. Place the gel in a horizontal electrophoresis chamber with TAE running buffer.
 - iii. Mix 4.5 μ L of PCR products with 0.5 μ L of gel loading buffer. Load the reactions in the agarose gel. Load a DNA ladder in another well.
 - iv. Perform electrophoresis at 5 V/cm at room temperature until the orange dye front reaches the bottom of the gel.
 - v. Photograph the gel on a blue or UV light transilluminator.

A size range >200–800 bp is acceptable so long as the peak amount lies within this range.

The procedure below uses reagents in the Quick Blunt Kit (New England Biolabs) and the QIAquick PCR Cleanup Kit (QIAGEN) and AMPure XP beads (Beckman). Enzymes and kits from other manufacturers are available to accomplish the same tasks, such as the Anza DNA Blunt End Kit (Life Technologies IVGN2404), the DNA Clean & Concentrator Kit (Zymo Research D4029), and the Quick-DNA Magbead Plus Kit (Zymo Research D4081).

26. Move the sheared DNA to PCR tubes. Blunt the ends of the sheared DNA by adding 8.2 μ L of 10 \times Quick Blunting Kit buffer (New England Biolabs), 2 μ L of 10 mM dNTP Solution Mix (New

Section 4

England Biolabs), and 2 μL of Quick Blunt enzyme mix (New England Biolabs) (all included in the kit).

27. Vortex briefly to mix and then centrifuge at 3000g for 5 sec at room temperature.
28. Incubate for 30 min at 25°C followed by heat-inactivation for 20 min at 70°C in a thermocycler with heated lid on.
29. Fully resuspend AMPure XP magnetic beads (Beckman) in the stock container. Add 0.8 volume of AMPure XP beads to the samples.
30. Vortex briefly to mix and then centrifuge at 3000g for 5 sec at room temperature.
31. Incubate for 4 min at room temperature to allow DNA to bind to the beads.
32. Place the tubes on the magnetic rack for 4 min. Remove and discard the supernatant using a 200- μL pipette tip while the tubes remain in the magnetic rack, being careful not to remove any of the beads.

While the DNA is bound to the AMPure XP beads, do not let the beads dry completely, or it will be difficult to recover the DNA.
33. Briefly centrifuge the tubes, place them back on the magnetic rack, and remove any residual liquid using a 20- μL pipette tip.
34. Add 150 μL of freshly made 70% ethanol to each tube. Turn each tube 180° twice, waiting for 1 min in between, while in the magnetic rack so that the beads move through the 70% ethanol twice.
35. Leave the samples for 4 min undisturbed in the magnetic rack, and then remove and discard the supernatant using a 200- μL pipette tip.
36. Centrifuge at 3000g for 5 sec at room temperature, place the tubes back on the magnetic rack, and then remove any remaining liquid using a 20- μL pipette tip.

It is critical to remove all of the 70% ethanol from each tube, as the residual ethanol interferes with recovery of the DNA.
37. Leave the tubes on the magnetic rack with their lids open for 2 min to allow any trace ethanol to evaporate.
38. Add 30 μL of pure H₂O to each tube. Remove the tubes from the magnetic rack and resuspend the beads thoroughly by pipetting. Incubate for 2 min at room temperature.
39. Place each tube back on the magnetic rack for 4 min.
40. Transfer 26 μL of the supernatant to new PCR tubes, leaving the remaining 4 μL behind.

Leaving 4 μL behind is necessary to avoid the beads being taken up into the pipette tip.
41. Mix dCTP and ddCTP to make a solution of 10 mM dCTP and 0.5 mM ddCTP.

Prepare this solution freshly.
42. Add poly(C) tails to the DNA 3' ends by adding 1.8 μL of the 10 mM dCTP/0.5 mM ddCTP mix, 7.17 μL of 5× TdT reaction buffer (Promega), and 0.9 μL of TdT enzyme (Promega).
43. Vortex briefly to mix and then centrifuge at 3000g for 5 sec at room temperature. Incubate for 50 min at 37°C, followed by heat inactivation for 20 min at 70°C in a thermocycler with heated lid on.

We have found that the TdT enzyme from Promega performs better than TdT from other companies. The presence of ddCTP at a 1:20 molar ratio with dCTP serves as a chain terminator of the poly(C) tailing reaction. The resulting poly(C) tail length averages 20 nt.
44. Purify the DNA using AMPure XP beads as described in Steps 29–40.
45. To each 26 μL volume of purified DNA, add 0.7 μL of 30 μM primer olj363, 2 μL of 30 μM primer olj376, 3.4 μL of 10× Easy-A buffer (Agilent Technologies), 1.2 μL of 10 mM dNTPs, and 0.7 μL of Easy-A Hi-Fi Cloning Enzyme (Agilent Technologies).

Section 4

49. Vortex to mix and then centrifuge at 3000g for 5 sec at room temperature.
50. PCR amplify using the following parameters:

Step	Time	Temperature	Cycles
Initial denaturation	1 min	95°C	1
Denaturation	30 sec	95°C	20
Annealing	30 sec	52°C	
Polymerization	2 min	72°C	
Extra extension	2 min	72°C	1
Hold		10°C	1

An example PCR product is illustrated in Figure 2B.

51. Estimate the concentration and size range of the PCR products by agarose gel electrophoresis as described in Step 25.

The PCR products should form a broad smear from ~0.2 to 2 kbp in size. The 400- to 900-bp products will be sequenced; therefore, it is important that the peak amount of product lies in this size range.

52. Pool (multiplex) the input and eight output PCR products at a 5:1:1:1:1:1:1:1:1 ratio, based on the amount of PCR products in the size range of 400–900 bp in the gel photo.

The 5:1 ratio of input to output is used to achieve greater sequencing depth of the input. This is important for determining putatively essential genes, as well as for reducing the number of occurrences in which a transposon insertion strain is present in an output but is absent from the input. One may use image-analysis software to quantify the amount of PCR products in the 400- to 900-bp range, and then use this information to pool the samples to achieve the desired ratio. However, this can be done by eye almost as well.

53. Purify the multiplexed DNA sample using the QIAquick PCR Purification Kit (QIAGEN).

- i. Add five volumes of buffer PB (QIAGEN). Vortex briefly to mix.
- ii. Apply the sample to a QIAquick spin column in a 2-mL collection tube (QIAGEN). Centrifuge at 17,900g for 1 min in a microcentrifuge at room temperature.
- iii. Discard the flowthrough, and then place the QIAquick column back into the same collection tube.
After discarding the flowthrough here and in Step 53.v, one may remove any remaining liquid on the rim of the collection tube by touching it to a clean paper towel.
- iv. Add 0.75 mL of buffer PE (QIAGEN) to the QIAquick column and centrifuge at 17,900g for 1 min at room temperature.
- v. Discard the flowthrough and place the QIAquick column back into the same collection tube. Centrifuge at 17,900g for an additional 1 min at room temperature.
- vi. Place the QIAquick column in a clean microcentrifuge tube.
- vii. In another clean microcentrifuge tube, dilute 30 μ L of buffer EB (QIAGEN) in 70 μ L of dH₂O to make a 0.3 \times working stock. Add 40 μ L of 0.3 \times buffer EB to the center of the QIAquick membrane.
- viii. Centrifuge the column at 17,900g for 1 min at room temperature.
- ix. (Optional): Run the flowthrough through the column a second time for greater recovery of DNA.

The purpose of diluting buffer EB to 0.3 \times is to reduce the concentration of the buffer, so that it is less likely to interfere in subsequent pH- or salt-sensitive reactions or manipulations.

- x. Measure the concentration of the purified DNA by UV spectrophotometry.

The concentration should be in the range of 50–500 ng/ μ L.

TABLE 2. Index 1 (i7) sequences for primers BC45–BC53

Index 1 primer	Index 1 (i7)
BC45	AGTCAACA
BC46	AGTTCCGT
BC47	ATGTCAGA
BC48	CCGTCCCG
BC49	GTCCGCAC
BC50	GTGAAACG
BC51	GTGGCCTT
BC52	GTTTCGGT
BC53	CGTACGTA

MPS (1–2 wk)

54. Submit the multiplexed sample from Step 53 and the custom sequencing primer olj386 to a sequencing facility for MPS on the Illumina NextSeq 550 or similar Illumina instrument, for example, the Tufts University Core Facility (www.tucf.com), which has extensive experience sequencing Tn-seq samples.

- i. Request purification of samples in the size range of 400–900 bp using the Pippin Prep purification system (Sage Science).

If the sequencing facility lacks the Pippin Prep purification system, one may perform gel purification of samples in the size range of 400–900 bp; however, loss of DNA and contamination of the final product with agarose is often problematic.

- ii. Request mid-output, 75-cycle, single-end sequencing using the custom sequencing primer olj386.

Although 75-cycle sequencing is requested, most core facilities running the Illumina NextSeq instrument will generate reads 76 nt in length.

- iii. Provide the sequencing facility with a table of the index 1 (i7) primer index sequences used in Step 48 (see Table 2).

They will use this information to demultiplex the reads and bin them for each sample.

Analyzing Tn-seq Data to Calculate Gene Fitness

The procedure below uses a custom pipeline script called HopCount.py to identify putatively essential genes and conditionally essential or important genes. See Discussion.

Download and Install Python3

55. To determine whether Python3 is already installed on the computer, open Command Prompt (Windows) or Command Terminal (MacOS) and type:

```
$ python3 --version
```

56. Determine whether Python3 is installed.

- Ensure that at least Python 3.6+ is installed for this pipeline.

If Python3 is installed, the output should be the version of Python3 installed.

- If Python3 is not already installed, or out-of-date, Python3 may be installed through one of several methods, including but not limited to (1) directly through python.org (www.python.org/downloads/), (2) through (mini)conda, a language-agnostic package manager (docs.conda.io/en/latest/miniconda.html), and (3) (MacOS) through homebrew, a package manager for MacOS (brew.sh/).

Adapter and Quality Trimming of Reads using *cutadapt.py*

Trim the raw Illumina sequencing reads to eliminate adapter sequences and ambiguous or poor-quality reads. Trimming is done using a publicly available Python script, *cutadapt.py*.

57. Download and install *cutadapt.py* by typing the following into the terminal:

```
$ python3 -m pip install --user --upgrade cutadapt
```

58. Navigate to the directory where the raw sequencing read files are located.

59. Execute *cutadapt.py* with the appropriate parameters to remove 5' and 3' adapter sequences from the sequencing reads in each file by typing the following into the terminal:

```
$ cutadapt -a C{12}N{38} -m 18 -M 51 -j 4 --trim-n -o FILENAME_Trimmed.fastq.gz  
FILENAME.fastq.gz
```

Some of the reads will include the poly(C) tails, added in Step 42, which must be removed. The longest natural poly(C) sequence in the *V. cholerae* genome is 11 nt. Therefore, we trim any sequence including and 3' to poly(C) sequences of 12 nucleotides and greater. We also omit poor-quality reads and very short reads post-trimming to enhance accurate mapping of the *mTn10* insertions.

Parameter	Explanation
-a C{12}N{38}	In a sequence of 12 cytidines in a row, remove everything including and 3' to that motif. The N{38} captures additional nucleotides 3' to the poly(C) motif, up to a 51-nt read. If using shorter or longer sequencing cycle data, adjust the number accordingly.
-m 18 -M 76	Discards reads shorter than 18 bp and longer than 76 bp. If using shorter or longer sequencing cycle data, adjust the -M number accordingly.
-j 4	Uses four logical threads for trimming; speeds execution of the program. Can set it lower if limited by hardware.
--trim-n	Removes ambiguous nucleotides from the ends of reads.
-o FILENAME_Trimmed.fastq.gz	This is the name of the outputted trimmed reads, in fastq.gz file format. Replace "FILENAME" with the desired name.
FILENAME.fastq.gz	Replace "FILENAME" with the name of the raw sequencing read file obtained from the sequencing core facility.

Cutadapt.py should take several minutes to complete. When complete, a panel of information detailing sequence results is displayed. The trimmed sequence read files will appear in the HopCount folder.

Download and Install HopCount

60. Install HopCount using one of the following methods:

- Navigate the browser to the HopCount GitHub page at github.com/camillilab/hopcount. Download the latest release by clicking the "Code" dropdown and then "Download ZIP". Extract the ZIP file on the computer.
- Alternatively, if git is installed on the computer, the repo can be cloned directly by typing into the terminal:

```
$ git clone  
https://github.com/camillilab/hopcount.git
```

Running *HopCount.py* requires that *Bowtie2* and *SAMtools* are installed on the computer.

61. Install brew by entering:

```
$ /bin/bash -c "$(curl -fsSL  
https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
```

62. Install Bowtie2 by entering:

```
$ brew install bowtie2
```

63. Install SAMtools by entering:

```
$ brew install samtools
```

Run HopCount.py

64. Navigate to the HopCount folder. Place the trimmed sequencing read files in the “fastq” folder. Place the appropriate reference GenBank file (in this example, N16961.gbk) into the HopCount folder.

65. Open hopcount.conf using any text editor.

Here the various parameters that HopCount uses in the mapping and analysis of transposon insertions are displayed. A brief description of each parameter is provided above each value. Here we use the following parameter settings:

```
var_hop_threshold = 15  
var_quality = 1  
var_make_wiggle_file = true  
var_keep_sam_file = bam  
var_circular_genome = true (keep as true even if analyzing a linear genome)  
var_percent_cutoff = 0.05
```

66. Open Command Prompt (Windows) or Terminal (MacOS). Navigate to the HopCount folder by typing the following:

```
$ cd PATH_TO_HOPCOUNT_FOLDER
```

PATH_TO_HOPCOUNT_FOLDER is the location of the HopCount folder. For novice users: Command Prompt and Terminal by default initialize in the home directory. cd, or “change directory”, is a command to navigate the file system. If one is unsure of one’s location, type ls (MacOS) or dir (Windows) to provide a current view of files and folders in the current location. From there, type cd ./LOCATION to move along, or cd .. to go backward.

67. Execute the HopCount.py script by entering the following:

```
$ python3 hopcount.py
```

The script should now successfully process the sequencing file, map to the reference genome, tabulate hops, aggregate under genomic regions, and calculate complexity and fold enrichment values for every file provided in “fastq.”

*A number of tab-separated output files have been created under the “output” folder: *_hopcount.tsv, which is a record of each transposon insertion recorded along with its determined genetic context; *_aggregate.tsv, which aggregates the records in hopcount.tsv into each region entry, with DvalGenome calculations and more; *_sense.tsv, which analyzes contributions in DvalGenome regarding same-sense insertions between adjacent genes; *_antisense.tsv, which analyzes contributions in DvalGenome regarding opposite-sense insertions between adjacent genes; .FASTA and .GFF3 files from the reference; the BAM file from Bowtie2 mapping; and WIG files for visualization in a genome browser.*

The aggregate.tsv file contains DvalGenome and additional information that will be used to identify genes important in the selection condition. A number of additional criteria may be used to increase specificity of the results.

Analyze HopCount Results

68. Open the *_aggregate.tsv file with Microsoft Excel.

Section 4

69. Sort the data by the “Sites” column, which is the total number of unique insertion sites per region, from fewest to greatest.

Areas with very few or no insertions may be essential, stochastically lost, or, if analyzing an output sample, may be conditionally essential. Determination requires analysis of these regions in the input sample. Short regions may lack insertions simply because of their small target size.

70. Observe the DvalGenome column.

As described earlier, DvalGenome is a ratio of the observed frequency of hops in a given region divided by the expected number of hops for that region.

71. The expected number of hops is the ratio of the total number of reads in the data set normalized to the region length in question:

$$\text{DvalGenome}(X) = \frac{\text{Number of hops in } X}{\text{Expected hops in } X} = \frac{\text{Number of hops in } X}{\text{Total hops} \times \frac{\text{Length of } X}{\text{Total length}}}$$

72. Compare DvalGenome values with those of a biological replicate of the desired condition. Copy “DvalGenome” values and plot them against each other using Excel (Microsoft), Prism (Graph-Pad), or a similar program.

73. Ensure that the regions are aligned properly. Obtain the Spearman correlation coefficient to determine the strength of reproducibility between the results.

Values of 0.6 and above are well-correlated. A more robust correlation can be done by comparing the subset of regions having a minimal number of unique mTn10 insertions (sites), for example, those with five or more sites. Regions with less than five sites will show greater noise in their DvalGenome values. However, care must be taken to ensure that the regions are aligned properly before performing the correlation.

74. Sort the data by “DvalGenome” values. Identify regions that are significantly under-represented (low DvalGenome values) or over-represented (high DvalGenome values). Highlight genes of interest.

75. Sort the data by locus. Determine whether the genes of interest highlighted in Step 74 are part of operons. If so, determine whether other genes in the operon have similar DvalGenome values.

76. Compare the DvalGenome value of candidate regions to those in the control condition to identify conditionally essential or important regions.

Visualize HopCount Analysis

Use a genome browser, such as IGB (Broad Institute), to visualize the distribution of hops in the genome.

77. Download and install IGB.

Often, genome browsers use an indexed genome file in FASTA format, a GFF3 file that provides annotation information, and an indexed, sorted BAM file from a mapping program. HopCount also generates a WIGGLE file that is a text position histogram representation of the BAM file, which loads quickly in genome browsers. A separate WIG file is generated for all positive-strand, negative-strand, or combined total insertions. We will use the HopCount-generated FASTA, GFF3, and WIG files to visualize insertions around promising candidate genes. Here we describe a brief procedure for IGB:

78. Launch IGB.

79. Load the genome FASTA using Genomes → Load From File....

80. Load the genome GFF3 using File → Load From File....

81. Load WIG file(s) using File → Load from File....

82. Adjust the view as necessary.

- Select individual chromosomes or episomes in the upper-left drop-down box (for example, *V. cholerae* has two circular chromosomes).
- Control base pair resolution with the “plus” and “minus” buttons at the upper right.
- Control location using the scrolling window between the resolution buttons.

83. Observe the distribution of hops in and surrounding genes or operons of interest, that is, those with low or high DvalGenome values, or surrounding putatively essential genes or operons.

DISCUSSION

HopCount.py uses the HopCount program that has been previously used in several studies (Klein et al. 2012; McDonough et al. 2014) as part of the Tufts University Core Facility Galaxy suite (galaxy.med.tufts.edu/). HopCount.py updates and improves the pipeline and allows it to be exported for general distribution and offline use. In this pipeline, first the unique transposon insertions (or “hops”) are detected by mapping sequencing reads to a reference file for *V. cholerae*. Insertion positions (sites) that fall within a coding sequence or intergenic region are aggregated. A fold enrichment value, or DvalGenome, is calculated as the ratio of detected insertions within a given region divided by the number of expected insertions in that region based on the size of the region relative to the genome size. High DvalGenome values indicate that mutants with an insertion in a given region are more abundant than expected in the population, suggesting that disruption of this region is beneficial in the selection condition. Conversely, low DvalGenome values suggest that insertions in that region are detrimental to surviving the selection. The orientation of the mTn10 is considered during tabulation, allowing the user to selectively analyze hops that may influence the expression of nearby genes in the sense or antisense direction in the selection conditions that used IPTG induction of the outward-reading P_{tac} promoter. DvalGenome values between input and output samples are compared for determination of significant genes and intergenic regions in a given selection condition. Finally, DvalGenome values between LB broth and sea H₂O-plus-chitin output samples are compared for determination of conditionally essential or important genes and intergenic regions.

RECIPES

1000× Kanamycin Stock

Kanamycin

Make kanamycin to a final concentration of 50 mg/ml in H₂O.

Filter-sterilize and store in aliquots at –20°C.

Chloramphenicol Stock Solution (10 mg/mL)

1. Prepare the desired volume of stock solution by dissolving chloramphenicol (Sigma-Aldrich C1919) at 10 mg/mL (or 20 mg/mL) in absolute ethanol.
2. Store for up to several months at –20°C.

EDTA

EDTA (ethylenediaminetetraacetic acid)

NaOH

To prepare EDTA at 0.5 M (pH 8.0): Add 186.1 g of disodium EDTA·2H₂O to 800 mL of H₂O. Stir vigorously on a magnetic stirrer. Adjust the pH to 8.0 with NaOH (~20 g of NaOH pellets). Dispense into aliquots and sterilize by autoclaving. The disodium salt of EDTA will not go into solution until the pH of the solution is adjusted to ~8.0 by the addition of NaOH.

Gel Loading Buffer (10×)

Reagent	Quantity for 10 mL
Orange G (Fisher Scientific AC416550100)	100 mg
Tris-EDTA solution (10×, pH 7.4; Fisher Scientific BP2477100)	5 mL
Glycerol (80% v/v)	5 mL

1. Add Orange G to 10× Tris-EDTA solution (pH 7.4).
2. Vortex to dissolve.
3. Add 80% glycerol.
4. Vortex to mix.
5. Store at room temperature.

The buffer is stable for up to 5 yr at room temperature.

LB (Luria–Bertani) Liquid Medium

Reagent	Amount to add
H ₂ O	950 mL
Tryptone	10 g
NaCl	10 g
Yeast extract	5 g

Combine the reagents and shake until the solutes have dissolved. Adjust the pH to 7.0 with 5 N NaOH (~0.2 mL). Adjust the final volume of the solution to 1 L with H₂O. Sterilize by autoclaving for 20 min at 15 psi (1.05 kg/cm²) on liquid cycle.

LB Agar

Agar (20 g/L)
NaCl (10 g/L; Sigma-Aldrich S9625)
Tryptone (10 g/L; BD 211705)
Yeast extract (5 g/L; BD 212750)

Add H₂O to a final volume of 1 L. Adjust the pH to 7.0 with 5 N NaOH. Autoclave. Pour into Petri dishes (~25 mL per 100-mm plate).

Molten Agarose

1. Add either 0.8 or 1.5 g of agarose (Fisher Scientific BP160) to 100 mL of 1× TAE <R> electrophoresis running buffer in a 250-mL glass bottle to make a 0.8% or 1.5% solution, respectively.
2. Microwave until boiling starts, and then carefully swirl the bottle.
3. Continue microwaving at reduced power with occasional swirling until the agarose is fully dissolved.
4. Cool to 60°C before use.

The molten 0.8% agarose can be stored for up to 1 wk at 60°C, but the 1.5% agarose should be used fresh. Single-use aliquots of either can be stored for up to 6 mo at room temperature, and then microwaved to remelt.

TAE

Prepare a 50× stock solution in 1 L of H₂O:

- 242 g of Tris base
- 57.1 mL of acetic acid (glacial)
- 100 mL of 0.5 M EDTA (pH 8.0)

The 1× working solution is 40 mM Tris-acetate/1 mM EDTA.

ACKNOWLEDGMENTS

We thank Cecilia Alejandra Silva Valenzuela, Lauren Shull, Miriam Ramliden, and David Lazinski for expert assistance and modifications of these procedures.

REFERENCES

- Bourgeois J, Camilli A. 2023. High-throughput mutant screening via transposon sequencing. *Cold Spring Harb Protoc* doi:10.1101/pdb.top107867
- Camilli A. 2022. Isolation and sequencing of novel *Vibrio* species. *Cold Spring Harb Protoc* doi:10.1101/pdb.prot108067
- camillilab/hopcount. 2022. github.com/camillilab/hopcount [Accessed July 24, 2022].
- Duncan MC, Forbes JC, Nguyen Y, Shull LM, Gillette RK, Lazinski DW, Ali A, Shanks RMQ, Kadouri DE, Camilli A. 2018. *Vibrio cholerae* motility exerts drag force to impede attack by the bacterial predator *Bdellovibrio bacteriovorus*. *Nat Commun* 9: 4757. doi:10.1038/s41467-018-07245-3
- Galaxy. 2011. galaxy.med.tufts.edu [Accessed July 24, 2022].
- Homebrew. brew.sh [Accessed July 24, 2022].
- Klein BA, Tenorio EL, Lazinski DW, Camilli A, Duncan MJ, Hu LT. 2012. Identification of essential genes of the periodontal pathogen *Porphyromonas gingivalis*. *BMC Genomics* 13: 578. doi:10.1186/1471-2164-13-578
- 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. doi:10.1111/mmi.12557
- Miniconda. 2017. docs.conda.io/en/latest/miniconda.html [Accessed July 24, 2022].
- Python. 2022. www.python.org/downloads/ [Accessed July 24, 2022].
- TUCF. www.tucf.com [Accessed July 24, 2022].