Published in final edited form as:

Curr Protoc Microbiol. ; 36: 1E.3.1–1E.3.24. doi:10.1002/9780471729259.mc01e03s36.

Genome-Wide Fitness and Genetic Interactions Determined by Tn-seq, a High-Throughput Massively Parallel Sequencing Method for Microorganisms

Tim van Opijnen¹, David W. Lazinski², and Andrew Camilli²

¹Department of Biology, Boston College, Chestnut Hill, Massachusetts

²Department of Molecular Biology and Microbiology, School of Medicine, Tufts University, and Howard Hughes Medical Institute, Boston, Massachusetts

Abstract

The lagging annotation of bacterial genomes and the inherent genetic complexity of many phenotypes is hindering the discovery of new drug targets and the development of new antimicrobial agents and vaccines. This unit presents Tn-seq, a method that has made it possible to quantitatively determine fitness for most genes in a microorganism and to screen for quantitative genetic interactions on a genome-wide scale and in a high-throughput fashion. Tn-seq can thus direct studies on the annotation of genes and untangle complex phenotypes. The method is based on the construction of a saturated transposon insertion library. After library selection, changes in the frequency of each insertion mutant are determined by sequencing flanking regions en masse. These changes are used to calculate each mutant's fitness. The method was originally developed for the Gram-positive bacterium *Streptococcus pneumoniae*, a causative agent of pneumonia and meningitis, but has now been applied to several different microbial species.

Keywords

transposon sequencing; Tn-seq; *Streptococcus pneumoniae*; *Vibrio cholerae*; genome-wide fitness; genetic interactions; transposon mutagenesis; massively parallel sequencing

INTRODUCTION

The Gram-positive bacterium *Streptococcus pneumoniae* is a common occupant of the nasopharynx and the causative agent of a variety of diseases ranging from otitis media to pneumonia, bacteremia, and meningitis. Each year, 1.5 million people succumb to infection with *S. pneumoniae* (Tuomanen et al., 2004). Due to increasing antibiotic resistance across many bacterial species and because the drug pipeline is running on empty, we are rapidly closing in on a pre-antibiotics era in terms of treatment options (Taubes, 2008). Vaccines are a promising future alternative for protecting against bacterial infections. For *S. pneumoniae*, two different polysaccharide-based vaccines are currently available that protect ~60% of vaccinees. These are expensive, however, and the polysaccharide approach is limiting, as the *S. pneumoniae* serotypes against which the vaccines are directed can be replaced by other serotypes that are not covered by the vaccine.

For *S. pneumoniae* as well as many other pathogens, it is key to gain a better understanding of the link between genotype and phenotype not only to understand processes such as colonization, persistence, and virulence, but also to identify new drug or vaccine antigen targets. One problem, however, is the lagging annotation of bacterial genomes. For instance, more than 30% of *S. pneumoniae* genes are of unknown function, and it is reasonable to assume that most genes to which functions have been ascribed are only partially understood. In addition, microorganisms are complex systems, and a phenotype does not simply arise as the product of individual genetic components; rather, a phenotype is a complex trait that results from interacting components.

To aid in this quest, Tn-seq has been developed as a widely applicable high-throughput tool for gene disruption, with great potential for elucidating gene function, linking genotypes to phenotypes, and resolving complex pathways in microorganisms. In *S. pneumoniae*, for which the method was originally developed, Tn-seq works as illustrated in Figure 1E.3.1. A gene disruption library is first constructed by transposing the mini-transposon *magellan6* into bacterial genomic DNA in vitro using the transposase MarC9. The transposed DNA is subsequently transformed back into bacteria, giving a pool of strains each containing a single transposon insertion in its genome. DNA is isolated from a portion of the bacterial pool prior to selection (t_1), and another portion is used to seed a culture for selection (in vitro or in vivo). After selection, bacteria are recovered and DNA is isolated again (t_2). To accomplish sequencing of only the regions flanking the transposon, a single nucleotide was mutated in the transposon's inverted repeats, introducing an *MmeI* restriction site. DNA from both time points is digested with *MmeI*, which cuts 20 bp downstream of its recognition site, leaving a two-base overhang to which an adapter is ligated. PCR is performed using one primer complementary to the adapter and another complementary to the *magellan6* inverted repeat. The resulting PCR product is 120 bp long, with 20 bp of bacterial-specific DNA flanked by Illumina-specific sequences that enable sequencing on the Illumina Genome Analyzer II (GAII) platform. Note that modifications to the Illumina-specific sequences, which allow sequencing on the HiSeq platform, are provided in Tables 1E.3.1 and 1E.3.2. After sequencing, samples are split based on barcode sequences, and reads are mapped to the genome and counted for each insertion, allowing fitness to be calculated (van Opijnen and Camilli, 2013).

Below, Basic Protocol 1 describes the construction of a mini-Mariner transposon library in *S. pneumoniae* by in vitro transposition, and the Support Protocol describes purification of the Mariner transposase. Next, Basic Protocol 2 describes transformation of linear fragments of genomic DNA from the transposon library into *S. pneumoniae*, including preparation of *S. pneumoniae* starter cultures and making the cells competent for transformation. Basic Protocol 3 describes the use of *S. pneumoniae* transposon library starter cultures in selection to determine each gene's fitness in a specific environment, as well as the sample preparation needed for Illumina sequencing.

Importantly, Tn-seq is not limited to *S. pneumoniae* or Gram-positive bacteria, but can be applied to any microorganism for which there is a suitable means of generating a library of marked (e.g., transposon or suicide plasmid integration) mutations (van Opijnen and Camilli, 2013). Alternate Protocol 1 describes modifications for performing Tn-seq in the

Gram-negative bacterium *Vibrio cholerae*, the causative agent of cholera. Finally, Alternate Protocol 2 describes a more general method for preparing transposon junctions for massively parallel sequencing that can be applied to any transposon or insertional element.

CAUTION

S. pneumoniae and *V. cholerae* are Biosafety Level 2 (BSL-2) pathogens. Follow all appropriate guidelines and regulations for the use and handling of pathogenic microorganisms. See *UNIT 1A.1* (Burnett et al., 2009) and other pertinent resources (APPENDIX 1B) for more information. *S. pneumoniae* is a human pathogen, capable of causing a variety of diseases ranging from otitis media to pneumonia, bacteremia, and meningitis. *V. cholerae* is a human pathogen capable of causing severe secretory diarrhea following oral ingestion.

NOTE

S. pneumoniae is grown on blood agar plates or statically in rich liquid medium such as Todd-Hewitt broth supplemented with 15% yeast extract (THY) or brain heart infusion (BHI). Liquid cultures and agar plates are incubated overnight in an incubator at 37°C with a 5% CO₂ atmosphere or in a candle jar. *V. cholerae* is grown on Luria-Bertani (LB) agar plates or in LB broth with aeration, both at 37°C.

BASIC PROTOCOL 1

GENERATION OF A *magellan6* TRANSPOSON LIBRARY IN *S. PNEUMONIAE* BY IN VITRO TRANSPOSITION

This protocol describes the construction of a transposon library in *S. pneumoniae*. Through in vitro transposition mediated by the transposase MarC9, the mini-transposon *magellan6* (a derivative of the Himar1 Mariner transposon), which encodes spectinomycin resistance, is introduced into linear bacterial genomic DNA fragments. *Magellan6* is provided on the plasmid pMagellan6 and transposes by a cut-and-paste mechanism into the genomic DNA. The transposition event leaves single-stranded gaps at each transposon-chromosome junction, which are repaired using T4 DNA polymerase and *E. coli* DNA ligase.

Materials

S. pneumoniae genomic DNA (obtained using a Qiagen Blood & Tissue Kit)

pMagellan6 DNA (available on request; andrew.camilli@tufts.edu)

2× Buffer A (see recipe)

MarC9 mariner transposase (see Support Protocol)

10× Buffer B (see recipe)

10× (10 mg/ml) bovine serum albumin (BSA; New England Biolabs)

2 mM dNTPs (APPENDIX 2A)

3 U/μl T4 DNA polymerase (New England Biolabs)

2.6 mM nicotinic adenine dinucleotide (NAD)

10 U/μl *E. coli* DNA ligase (New England Biolabs)

30° and 75°C heating blocks

12° and 16°C cooling blocks

Additional reagents and equipment for ethanol extraction (Moore and Dowhan, 2002)
and DNA quantitation (Gallagher, 2011)

Carry out in vitro transposition

1. Combine the following:

1 μg *S. pneumoniae* genomic DNA

1 μg pMagellan6 (plasmid)

10 μl 2× Buffer A

Distilled water to 19.5 μl.

2. Add 0.5 μl MarC9 at the appropriate concentration (determined empirically).
3. Incubate 1 hr at 30°C in a heating block.
4. Heat inactivate 10 min at 75°C in a heating block, then transfer to ice for 5 min.
5. Precipitate DNA by adding 1/10 volume of 3 M sodium acetate, pH 5.2, and 2 volumes of 100% ethanol at 4°C and incubating 20 min at 4°C.
6. Microcentrifuge 15 min at maximum speed, 4°C, to form a pellet, then wash pellet with room temperature 70% ethanol and air dry. Dissolve pellet in 12.5 μl water.
7. Measure DNA concentration (Gallagher, 2011) using 1.5 μl of the DNA solution.

A yield in the range of 50 to 150 ng/μl should be obtained.

Repair transposon junctions

- 8 Mix the following:

2 μl 10× Buffer B

2 μl 10× BSA

1 μl 2 mM dNTPs

3 μl H₂O

1 μl 3 U/μl T4 DNA polymerase

11 μl DNA.

- 9 Incubate 20 min at 12°C, then inactivate the reaction for 10 min at 75°C.
- 10 Add 0.2 μl of 2.6 mM NAD and 1 μl of 10 U/μl *E. coli* DNA ligase and incubate overnight at 16°C.

If a very-small-volume pipettor such as a Rainin P2 is not available, 1 μ l of 0.52 mM NAD can be used; however, the final volume of the ligation reaction will be slightly altered.

Ligation products can be stored up to 1 month at -70°C .

SUPPORT PROTOCOL

PURIFICATION OF MARINER TRANSPOSASE

This protocol describes the purification of a maltose binding protein (MBP)–Himar1 Mariner MarC9 transposase fusion protein essentially as described (Lampe et al., 1996). MarC9 is a hyperactive derivative of the transposase (Lampe et al., 1999). The fusion protein is overexpressed using the *E. coli* pMalC9 strain. The fusion protein is purified by affinity chromatography using an amylose resin column and is stored frozen in single-use aliquots.

Materials

E. coli pMalC9 strain (Lampe et al., 1999; available on request from andrew.camilli@tufts.edu)

LB/ampicillin: LB broth (APPENDIX 4A; Stevenson, 2006) supplemented with 100 $\mu\text{g/ml}$ ampicillin

Isopropyl- β -D-thiogalactopyranoside (IPTG)

Column buffer (see recipe)

Amylose resin (New England Biolabs)

Transposase wash buffer (TWB; see recipe)

Transposase elution buffer (TEB; see recipe)

37°C , 5% CO_2 incubator

15-ml centrifuge tube

1.5-ml microcentrifuge tubes

Prepare cell lysate

1. Grow a 5-ml overnight culture of *E. coli* pMalC9 in LB/ampicillin at 37°C with aeration.
2. Inoculate 80 ml LB/ampicillin with 0.8 ml overnight culture and grow at 37°C with aeration until the OD_{600} of the culture reaches ~ 0.5 .
3. Induce protein production by adding IPTG to a final concentration of 0.3 mM. Continue incubating at 37°C with aeration for 2 hr.
4. Harvest cells by centrifuging 10 min at $8000 \times g$, 4°C . Discard supernatant and resuspend cell pellet in 10 ml cold column buffer.
5. Lyse cells by any reliable method, e.g., bead-beating or a French Press.

Failure to recover protein or poor yields often result from a failure to disrupt the cells. Cells disruption can be assessed by examining the cells before and after breakage by phase-contrast microscopy. At least 90% of cells should be disrupted before proceeding.

6. Clarify the lysate by centrifuging 10 min at $15,000 \times g$, 4°C .
7. Transfer supernatant to a 15-ml centrifuge tube on ice and discard the pelleted insoluble material.

Prepare chromatography resin

8. Thoroughly resuspend amylose resin and aliquot 500 μl into a 1.5-ml tube.
9. Add 1 ml cold TWB, briefly mix to wash the resin, and microcentrifuge 30 sec at $10,000 \times g$, discarding the supernatant. Repeat for a total of three washes. After the final wash, resuspend the resin in 1 ml cold TWB.

Purify MBP-transposase

10. Transfer resin to the clarified cell lysate on ice, then incubate at 4°C on a rocking platform for 1 hr to bind the MBP-transposase protein to the resin.
11. Centrifuge 5 min at $10,000 \times g$, 4°C . Carefully remove the supernatant and discard.
12. Resuspend pellet in 2 ml cold TWB and transfer to two 1.5-ml microcentrifuge tubes.
13. Wash resin a total of four times as in step 9 using cold TWB. Remove the final supernatant from each tube.
14. Elute bound protein by resuspending each resin pellet in 200 μl cold TEB. Incubate 5 min on ice, vortexing gently every 1 to 2 min.
15. Centrifuge 2 min at $10,000 \times g$, 4°C . Remove supernatants and transfer to a single clean 1.5-ml tube.
16. Repeat centrifugation to pellet any residual resin. Transfer supernatant to a clean tube on ice.
17. Dispense $\sim 10\text{-}\mu\text{l}$ aliquots of purified transposase in microcentrifuge tubes or smaller PCR tubes and freeze at -70°C .

The typical concentration of MBP-transposase after this procedure is $\sim 100 \mu\text{g/ml}$. The purity of the final product can be assessed by analysis on an 8% SDS-polyacrylamide gel. The molecular weight of the MBP-transposase fusion protein is 83 kDa. If yields are low, failure to lyse the cells is typically at fault.

MarC9 must be stored in small aliquots, as thawing more than once lowers transposition efficiency. When stored at -70°C or colder, the purified transposase retains activity for at least 1 year.

BASIC PROTOCOL 2

TRANSFORMATION OF *S. PNEUMONIAE* WITH *magellan6* TRANSPOSON INSERTION LIBRARY

This protocol describes how to transform competent *S. pneumoniae* with linear fragments of genomic DNA containing *magellan6* insertions to yield a transposon insertion library. It begins with preparation of competent cells from starter cultures of *S. pneumoniae*. Upon transformation, the *magellan6* insertions are integrated into the genome by double cross-over homologous recombination. Transformants are selected on medium containing an antibiotic for which the transposon encodes resistance. Because in vitro transposition and subsequent transformation of *S. pneumoniae* are both low-frequency events, the overwhelming majority of transformants harbor single transposon insertions. The resulting library can be used in any environment of interest (in vivo or in vitro) to gain insight into each gene's contribution to fitness (survival and/or growth) in that specific environment.

Although this protocol takes advantage of the natural competency of *S. pneumoniae*, non-competent microbial species for which in vivo transposition methods are available may also be subjected to Tn-seq, as in the example of *V. cholerae* (see Alternate Protocol 1). In addition, although this protocol is specific to *magellan6*, other transposons (e.g., Tn5) or insertion elements (e.g., suicide vector) can be used with Tn-seq. A more general method for purifying the transposon-chromosome junctions for sequencing that can work with any transposon or insertional element is also provided (see Alternate Protocol 2).

Of significant importance in the following protocol is the elimination of the plating of transformed bacterial cells in soft agar layers, which has been the standard method for selecting *S. pneumoniae* transformants. Instead, transformed cells are outgrown for a short period of time to allow expression of the transposon drug marker, and then spread on antibiotic-containing blood agar plates. This results in a dramatic increase in the dynamic range of the experiments, allowing recovery of insertion mutants with a fitness 0.05 (van Opijnen and Camilli, 2012). This is in stark contrast to the use of the original soft agar layering method, where recovered insertion mutants have a fitness 0.5 (van Opijnen et al., 2009).

Materials

S. pneumoniae strain (e.g., D39 or TIGR4)

Sheep's blood agar plates (see recipe), with and without 200 µg/ml spectinomycin
Todd-Hewitt yeast (THY): 30 g/liter Todd-Hewitt base (Becton Dickinson, cat. no. 249240), pH 7.3, with 15% yeast extract

Oxyrase (Oxyrase, cat. no. OB-0010)

THY containing 12% (v/v) glycerol

1 N HCl

20% (w/v) glycine in distilled water

1 N NaOH

8% (w/v) bovine serum albumin (BSA)

1 M CaCl₂

350 ng/μl competence stimulating peptide (CSP; AnaSpec) appropriate to the *S. pneumoniae* strain used (e.g., CSP1 for D39, CSP2 for TIGR4)

Transposon library DNA (see Basic Protocol 1)

THY containing 20 μg/ml spectinomycin

37°C, 5% CO₂ incubator

1.5-ml microcentrifuge tubes

Large (16-mm-diameter) and small (13-mm-diameter) glass culture tubes

Spectrophotometer

37°C heating block

Prepare *S. pneumoniae* starter cultures

1. Streak an *S. pneumoniae* strain on a sheep's blood agar plate (without spectinomycin) and incubate overnight at 37°C with 5% CO₂.
2. Pick a single colony and grow in 250 μl THY and 5 μl Oxyrase in a 1.5-ml microcentrifuge tube for 3 hr.
3. Plate 100 μl on a blood agar plate (without spectinomycin) and incubate overnight at 37°C with 5% CO₂.
4. Resuspend all colonies from the plate's surface in 3 ml THY, transfer to a large culture tube, and bring the volume to 10 ml with THY containing 5 μl/ml Oxyrase. Incubate at 37°C with 5% CO₂ until the OD₆₀₀ is ~0.3 (2 to 3 hr).
5. Centrifuge 8 min at 3000 × *g*, room temperature.
6. Discard the supernatant and resuspend the cell pellet in 5 ml THY containing 12% glycerol.
7. Aliquot into twenty 1.5-ml microcentrifuge tubes and store up to 9 months at −70°C or colder.

Carry out transformation

- 8 Thaw a starter culture and transfer to 7.5 ml THY (without antibiotics or Oxyrase) in a large glass culture tube.
- 9 Add 97.5 μl of 1 N HCl and 18.8 μl of 20% glycine. Allow culture to grow for 1.5 to 2 hr and measure the OD₆₀₀ (should be between 0.03 and 0.1).

It is very important to have fast-growing starter cultures; the faster a starter culture grows, the higher the transformation efficiency becomes. For best results, a starter culture should reach the appropriate OD within 2 hr of seeding the growth medium.

- 10** In a small culture tube, combine volumes of culture and prewarmed THY that will give a final volume of 1 ml at an OD₆₀₀ of ~0.03.
- For example, if the OD₆₀₀ is 0.06, combine 500 µl THY and 500 µl culture.
- 11** Place culture tube in a heating block at 37°C.
- 12** Add the following in order, vortexing briefly and gently after each addition:
- 0.02 volume 1 N NaOH
 - 25 µl 8% BSA
 - 1 µl 1 M CaCl₂.
- The volume of NaOH is relative to the culture volume used in step 10. Following that example, if 500 µl culture was used, 10 µl NaOH should be added.
- Each component should be added separately, as they will precipitate if mixed beforehand.
- 13** Add 1.5 µl of 350 ng/µl CSP, immediately start a timer, and return the culture to the incubator for 15 min (CSP1) or 14 min (CSP2).
- The type of CSP and incubation time depend on the *S. pneumoniae* strain used. Use CSP1 for 15 min with strain D39, and CSP2 for 14 min with TIGR4 (Pozzi et al., 1996; Hsieh et al., 2006).
- 14** At precisely the indicated time (14 or 15 min), add 22 µl transposon library DNA and return to the incubator for 45 min.

Prepare transposon library starter cultures

- 15** Transfer 1 ml transformed culture to a microcentrifuge tube and centrifuge 8 min at 3000 × g, room temperature.
- 16** Discard 900 µl supernatant and resuspend the cell pellet in the remaining 100 µl THY.
- The cell pellet is often not visible.
- 17** Plate on a blood agar plate containing 200 µg/ml spectinomycin and incubate overnight at 37°C with 5% CO₂.
- 18** Resuspend colonies in 3 ml THY, transfer to a culture tube, and bring to 10 ml with THY containing 200 µg/ml spectinomycin and 5 µl/ml Oxyrase. Incubate to an OD₆₀₀ of 0.3 (30 to 60 min).
- 19** Centrifuge 8 min at 3000 × g, room temperature, and discard the supernatant. Resuspend cell pellet in 5 ml THY containing 12% glycerol.
- 20** Aliquot into twenty 1.5-ml microcentrifuge tubes and store up to 1 year at -80°C.

BASIC PROTOCOL 3

TRANSPOSON LIBRARY SELECTION AND SAMPLE PREPARATION FOR ILLUMINA SEQUENCING

This protocol describes how a transposon library starter culture can be used to determine each gene's fitness in a specific environment. In short, a starter culture is thawed and briefly cultured (30 min), a sample is taken for DNA isolation (t_1), and an aliquot is used to seed an environment of interest. After several hours of selective culture (e.g., for growth in a liquid medium), another sample is taken for DNA isolation (t_2). The DNA from both time points is then prepared for Illumina sequencing. Here, THY is used as an example of a selective environment, allowing the determination of genes that are important or dispensable for growth in THY. The environment can be changed to any environment of interest, including infection in a mouse model.

A key step in this protocol is digestion of the DNA with the type IIS restriction enzyme *MmeI* (Fig. 1E.3.2). The inverted repeat sequences present at the left and right ends of *magellan6* were engineered to contain an *MmeI* recognition sequence. *MmeI* cuts 20 bp downstream of its recognition site, leaving a random two-base 3' overhang. This generates DNA fragments of which a subset contain the *magellan6* inverted repeat and 20 bp of flanking DNA. By subsequently ligating on an adapter with a random two-base 5' overhang, DNA fragments with the same spacing between adapter and the *magellan6* inverted repeat are created. These fragments can be amplified by PCR in an unbiased manner and then sequenced en masse.

Materials

Transposon library starter culture (see Basic Protocol 2)

Todd-Hewitt yeast (THY): 30 g/liter TH base (Becton Dickinson, cat. no. 249240), pH 7.3, with 15% yeast extract

Spectinomycin (Sigma-Aldrich)

Oxyrase (Oxyrase, cat. no. OB-0010)

DNeasy Blood & Tissue Kit (Qiagen)

Adapter oligonucleotides (Table 1E.3.1)

1 mM Tris·Cl, pH 8.3 (APPENDIX 2A)

2 U/μl *MmeI* restriction enzyme (New England Biolabs, cat. no. R0637S)

32 mM S-adenosyl methionine (SAM; New England Biolabs)

10× Buffer 4 (New England Biolabs)

Calf intestinal phosphatase (CIP; New England Biolabs, cat. no. M0290S)

400 U/μl T4 DNA ligase and buffer (New England Biolabs)

PFU Ultra polymerase and buffer (Stratagene)

25 mM dNTP mix

10 μ M each PCR primer (Table 1E.3.2): primer P1_M6_MmeI and Gex PCR Primer 2

GelGreen DNA stain (Biotium)

50-bp DNA ladder

QIAquick Gel Extraction Kit (Qiagen cat. no. 28704)

Large (16-mm-diameter) glass culture tubes

Spectrophotometer

37°C, 5% CO₂ incubator

37° and 96°C heating blocks

16°C cooling block

Thermal cycler

UV transilluminator

Bioanalyzer (Agilent)

Additional reagents and equipment for isolation of genomic DNA from bacteria (DNeasy Blood & Tissue Kit; also see Wilson, 1997), phenol/chloroform extraction and ethanol precipitation of DNA (Moore and Dowhan, 2002), microvolume DNA quantitation (Desjardins and Conklin, 2011), and agarose gel electrophoresis (Voytas, 2000)

Perform library selection

1. Thaw a transposon library starter culture and transfer to 4 ml THY in a large glass culture tube. Add spectinomycin to 200 μ g/ml and 20 μ l Oxyrase, then grow for 30 to 60 min and check the OD₆₀₀ (should be ~0.1).
2. Set aside a small sample to be used for enumerating colony-forming units (CFU) in the culture at t_1 .
3. Use another small sample to seed 10 ml fresh THY at an OD of 0.005 and incubate at 37°C with 5% CO₂ to an OD of ~0.6 (selection).
4. During selection, centrifuge the remaining undiluted culture (step 1) for 8 min at 3000 $\times$ g, room temperature, discard the supernatant, and immediately isolate t_1 genomic DNA from the pellet using a DNeasy Blood & Tissue Kit.

It is essential that the genomic DNA be clean, as this increases the efficiency of MmeI restriction digestion.

If desired, the pellet can be stored up to 1 year at –20°C for later DNA isolation.

5. When the selective culture reaches an OD of ~0.6, take a small sample to enumerate CFU at t_2 , then centrifuge the rest of the culture and isolate t_2 genomic DNA.

Prepare adapter

- 6** Prepare the adapter by bringing both oligonucleotides (Table 1E.3.1) to 0.2 mM with 1 mM Tris-Cl, pH 8.3. Mix equal volumes in a microcentrifuge tube and place in a heating block at 96°C for 2 min.

See Figure 1E.3.2 and Tables 1E.3.1 and 1E.3.2 for details on how the adapter and primers combine to result in DNA molecules that can be sequenced on an Illumina GAII or HiSeq platform. Note that the two oligonucleotides that form the adapter contain a 4-nt (GAII) or 6-nt (HiSeq) barcode (index) sequence (Table 1E.3.1). By varying this sequence, multiple samples can be mixed into a single Illumina flow cell lane, which can be separated after sequencing based on the barcodes. For instance, by varying the barcode for t_1 and t_2 , these two samples can be mixed and sequenced in the same flow cell lane.

- 7** Anneal oligonucleotides by moving the part of the heating block that holds the microcentrifuge tubes to the bench and allowing it to cool to room temperature (~10 to 20 min).
- 8** Store annealed adapter at -20°C until use (up to 4 years).

Prepare sample for sequencing

- 9** Digest genomic DNA from t_1 and t_2 with the following mixture for 2.5 hr at 37°C:

3 µg *S. pneumoniae* DNA

3 µl (6 units) *MmeI*

0.44 µl 32 mM SAM

20 µl Buffer 4

H₂O to 200 µl.

- 10** Add 2 µl CIP and incubate 1 hr at 37°C.
- 11** Extract DNA with 24:23:1 (v/v/v) phenol/chloroform/isoamyl alcohol, precipitate with 95% ethanol, wash pellet with 70% ethanol, and air dry on the bench for 10 min (Moore and Dowhan, 2002). Resuspend in 25 µl distilled water.

Extraction and precipitation are essential, as they greatly improve the efficiency of ligation and amplification.

- 12** Ligate adapter by mixing:
- 25 µl extracted DNA
- 1 µl annealed adapter
- 1 µl 400 U/µl T4 DNA ligase
- 3 µl T4 DNA ligase buffer.

Incubate overnight at 16°C.

13 Prepare the following PCR mix to amplify the 120-bp target:

5 µl PFU Ultra buffer
 0.5 µl 25 mM dNTPs
 1 µl 10 µM primer P1_M6_MmeI
 1 µl 10 µM Gex PCR Primer 2
 0.5 µl PFU Ultra polymerase
 40.5 µl H₂O
 1.5 µl ligation mix.

14 Carry out the following PCR program:

1 cycle:	30 sec	95°C	(initial denaturation)
18–22 cycles:	10 sec	95°C	(denaturation)
	25 sec	55°C	(annealing)
	5 sec	72°C	(extension)
1 cycle:	10 min	72°C	(final extension).

The goal is to perform as few cycles as possible in order to remain in the linear range of the PCR.

15 Run all PCR products on a 1.8% agarose gel (Voytas, 2000) using GelGreen DNA stain and a 50-bp DNA ladder.

16 Excise the 120-bp band under low UV light, avoiding primer and primer-dimer bands.

CAUTION: UV light is damaging to eyes and exposed skin. Protective eyewear should be worn at all times while using a UV light source.

17 Purify DNA from the excised gel slice using a QIAquick Gel Extraction Kit and elute in 50 µl distilled water.

18 Determine DNA yield on a Nanodrop. Alternatively, determine yield and assess quality of the DNA on an Agilent Bioanalyzer.

The yield should be ~3 ng/µl and the Bioanalyzer should show a distinct peak in the 120-bp size range.

The t_1 and t_2 samples are ready for sequencing on an Illumina HiSeq or GAII platform according to the manufacturer's protocol. After sequencing, reads are mapped to the genome and the change in the number of reads per insertion over time (from t_1 to t_2) is used to accurately determine fitness for each gene. A detailed description of this method and data analysis can be found elsewhere (van Opijnen et al., 2009).

DNA samples can be stored up to 1 year at –20°C.

ALTERNATE PROTOCOL 1

TRANSPOSON LIBRARY CONSTRUCTION IN *V. CHOLERA* BY CONJUGATION

The protocol describes the construction of a mini-Tn5 (mTn5) transposon library in *V. cholerae* using conjugation (via bacterial mating) and is essentially as reported previously (Merrell et al., 2002). mTn5, which encodes kanamycin resistance, is delivered from an *E. coli* donor strain (Sm10λpir) into *V. cholerae* via conjugal transfer of the suicide plasmid pUTmTn5Km2. After transfer, although the plasmid is incapable of replicating in *V. cholerae*, the Tn5 transposase encoded on the plasmid mediates transposition of mTn5 (also on the plasmid) into the *V. cholerae* genome. Because each mutant arises from an independent mating event, each is considered an independent transposition event. Note that insertions into essential genes or essential intergenic sequences will be absent from the final library, and insertions that slow the growth of *V. cholerae* on LB medium will be under-represented.

One disadvantage of this method of generating a transposon insertion library is that ~20% of the exconjugates will have the suicide plasmid randomly integrated into the recipient genome instead of a true transposon hop. These are identifiable because they contain the *bla* (β-lactamase) antibiotic resistance gene on the plasmid backbone and thus are resistant to ampicillin. All plasmid integration mutants will yield identical transposon-plasmid junctional sequence upon sequencing, and thus can be ignored since this sequence will not map to the bacterial genome and therefore is filtered out bioinformatically.

Materials

Streptomycin-resistant *V. cholerae* recipient strain

E. coli donor strain Sm10λpir carrying pUTmTn5Km2 (Merrell et al., 2002; available on request from andrew.camilli@tufts.edu)

LB broth (APPENDIX 4A; Stevenson, 2006), unsupplemented and supplemented individually with 130 µg/ml kanamycin or 100 µg/ml streptomycin

LB agar plates (APPENDIX 4A; Stevenson, 2006), unsupplemented and supplemented with 100 µg/ml streptomycin plus 130 µg/ml kanamycin

80% (w/v) glycerol

37°C incubator

Glass plating rod

1.5-ml or larger microcentrifuge tubes or cryovials

Additional reagents and equipment for replica plating (Elbing and Brent, 2002)

Carry out transposition by mating

1. Grow donor and recipient strains to late-exponential phase (OD₆₀₀ = 0.5 to 1) at 37°C with aeration in LB broth supplemented with 130 µg/ml kanamycin or 100 µg/ml streptomycin, respectively.

2. Pellet 2 ml of each culture for 2 min at $10,000 \times g$, room temperature, and discard the supernatant.
3. Wash cells once with 2 ml LB to remove any traces of antibiotics.
4. Resuspend each pellet in 0.5 ml LB broth and combine at a 1:1 ratio based on OD_{600} .
5. Spread an appropriate amount of mixture (determined empirically to generate 500 to 1000 transposition mutant colonies per plate) onto 40 LB agar plates without antibiotics. Incubate plates 1 hr at 37°C to allow bacterial mating.

Select and pool transposants

6. Replica plate (Elbing and Brent, 2002) each mating plate onto LB agar supplemented with $100 \mu\text{g/ml}$ streptomycin and $130 \mu\text{g/ml}$ kanamycin to select for transposition mutants. Incubate plates 16 hr at 37°C .

Incubating longer than 16 hr or plating too much mating mixture will increase background on the selection plates. If background is problematic, replica plate the selection plates (after 16 hr growth) onto new antibiotic-containing plates and incubate 16 hr at 37°C .
7. Flood each plate with 2 ml LB broth and resuspend colonies with a glass plating rod. Pool transposition mutants by collecting the fluid from each plate and adding to a 50-ml conical tube. Vortex the entire pooled library to resuspend the cells.

Note that approximately half of the added LB will be absorbed into the agar plates.
8. Add glycerol to a final concentration of 15% (v/v) and vortex again to mix.
9. Dispense 1-ml aliquots of library into microcentrifuge tubes or cryovials and freeze at -70°C or colder for long-term storage (up to 2 years).

A yield in the range of 20,000 to 40,000 transposon insertion mutants should be obtained.

ALTERNATE PROTOCOL 2

GENERAL METHOD FOR Tn-seq SAMPLE PREPARATION FOR ILLUMINA SEQUENCING

This protocol describes how the transposon junctions from a *V. cholerae* mTn5 library are prepared for massively parallel sequencing on the Illumina HiSeq platform. Library selection and determination of gene fitness using Tn-seq data are done essentially as described for the *S. pneumoniae magellan6* transposon library. The method below is a variation of homopolymer tail-mediated ligation PCR (HTML-PCR; Klein et al., 2012; Kamp et al., 2013; Lazinski and Camilli, 2013). Although it is described here for the *V. cholerae* mTn5 library, it can be used with essentially any transposon or insertional element.

Materials

PCR and sequencing primers for mTn5 insertion libraries (Table 1E.3.3)

1 mM Tris·Cl, pH 8.3 (APPENDIX 2A)

V. cholerae transposon library (see Alternate Protocol 1)

DNeasy Blood & Tissue Kit (Qiagen)

GelGreen DNA stain (Biotium)

QIAquick PCR Purification kit (Qiagen)

dCTP

ddCTP

Terminal deoxynucleotidyl transferase (TdT) and 5 × reaction buffer (Promega)

DTR Gel Filtration Cartridge (Performa)

25 mM dNTP mix

Easy-A Cloning Enzyme with 10× reaction buffer (Agilent)

1.5-ml microcentrifuge tubes

Branson high-intensity cuphorn sonifier (Branson)

37° and 75°C heating blocks

Thermal cycler

Additional reagents and equipment for library selection and analysis of fitness (see Basic Protocol 3), microvolume DNA quantitation (Desjardins and Conklin, 2011), and agarose gel electrophoresis (Voytas, 2000)

Prepare oligos and sample DNA

1. Dissolve all oligonucleotides (Table 1E.3.3) to 30 μ M with pure water or 1 mM Tris·Cl, pH 8.3. Store at -20°C until use (up to 4 years).
2. Perform library selection as described, generating pre- and post-selection samples (t_1 and t_2) of genomic DNA (using the DNeasy Blood & Tissue kit) and culture for enumerating CFU (see Basic Protocol 3, steps 1–5).
3. Dilute 1 to 10 μ g genomic DNA in a 50- μ l volume. Mechanically shear by sonicating in 1.5-ml microcentrifuge tubes in a prechilled Branson high-intensity cuphorn sonifier for 2 min at 50% intensity with a 10-sec ON/5-sec OFF duty cycle. Spin briefly to bring contents to the bottom of the tube, then repeat sonication for 1 min.

Alternatively, genomic DNA can be sheared by nebulization or partially digested with DNase I or Fragmentase (both from New England Biolabs).

4. Run 5 μ l on a 2% agarose gel (Voytas, 2000) and stain with GelGreen to confirm that the DNA has been sheared to between 100 and 600 bp. If fragments are too large, perform additional sonication.
5. Remove short DNA fragments (<100 bp) using a QIAquick PCR Purification kit.

Add 3'-poly-dC tails

6. Prepare a small volume (e.g., 20 μ l) containing 9.5 mM dCTP and 0.5 mM ddCTP in pure water.
7. Prepare the reaction mix by adding in order:
 - 1 μ l dCTP/ddCTP mix
 - 4 μ l 5 $\times$ TdT reaction buffer
 - 14.5 μ l sheared DNA
 - 0.5 μ l TdT enzyme.
8. Mix by pipetting three times, then pellet contents by brief microcentrifugation, and allow reaction to proceed 1 hr at 37°C.
9. Heat inactivate 20 min at 75°C.
10. Remove small molecules (e.g., buffer components) by passing the reaction mix through a DTR Gel Filtration Cartridge as follows:
 - a. Pre-spin column for 3 min at 850 $\times$ g in a tabletop microcentrifuge to remove water.
 - b. Move the column to a new collection tube, pipet the reaction mix in the middle of the top of the column, and centrifuge 3 min at 850 $\times$ g.
 - c. Dispose of column and save flowthrough.

Amplify transposon junctions

11. Prepare the following mix for the first nested PCR:
 - 34.2 μ l pure water
 - 5 μ l 3'-poly-dC DNA
 - 3 μ l 30 μ M primer olj376
 - 1 μ l 30 μ M primer olj491
 - 0.8 μ l 25 mM dNTPs
 - 5 μ l 10 $\times$ Easy-A buffer
 - 1 μ l Easy-A Cloning Enzyme.
12. Carry out the following PCR program:
 - 1 cycle: 1 min 95°C (initial denaturation)
 - 25 cycles: 30 sec 95°C (denaturation)

30 sec 58°C (annealing)
 2 min 72°C (extension)
 1 cycle: 2 min 72°C (final extension).

- 13** Prepare the following mix for the second nested PCR (to further amplify transposon junctions and add barcode):

40.7 µl pure water
 0.5 µl from first PCR reaction
 1.0 µl 30 µM primer olj492
 1.0 µl 30 µM barcode primer
 0.8 µl 25 mM dNTPs
 5.0 µl 10× Easy-A buffer
 1.0 µl Easy-A Cloning Enzyme.

- 14** Carry out the following PCR program:

1 cycle: 1 min 95°C (initial denaturation)
 15 cycles: 30 sec 95°C (denaturation)
 30 sec 52°C (annealing)
 2 min 72°C (extension)
 1 cycle: 2 min 72°C (final extension).

- 15** Purify PCR products using a QIAquick PCR Purification kit, eluting with 50 µl.

- 16** Run 5 µl on a 2% agarose gel with GelGreen stain.

The second PCR product should run as a smear of DNA fragments ranging from 140 to >700 bp, with the peak of intensity centered between 250 and 400 bp.

- 17** Determine DNA yield on a Nanodrop. Optionally, assess quality of the DNA on a Bioanalyzer.

The sample concentration should be ~50 to 250 ng/µl.

- 18** Multiplex equivalent amounts for sequencing on the Illumina HiSeq.

Samples are sequenced using the custom sequencing primer HK89 (Table 1E. 3.3). The standard Illumina Multiplexing Index Sequencing Primer should be used to read the barcode. After sequencing, reads are demultiplexed and then mapped to the genome, and gene fitness is determined as described above (see Basic Protocol 3).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2A.

Buffer A, 2×

5 µl 80% (v/v) glycerol
4 µl 20 mM dithiothreitol (DTT)
4 µl 10 mg/ml bovine serum albumin (BSA; New England Biolabs)
8 µl 5× salt solution (see recipe)
Prepare fresh just before use

Buffer B, 10×

500 mM Tris-Cl, pH 7.8 (APPENDIX 2A)
100 mM MgCl₂
10 mM DTT
Store up to 1 year at −20°C

Column buffer

20 mM Tris-Cl, pH 7.4 (APPENDIX 2A)
200 mM NaCl
1 mM EDTA
1× complete EDTA-free protease inhibitor cocktail (Roche, cat. no. 1 836 170)
Prepare fresh just before use

Salt solution, 5×

100 µl 125 mM HEPES, pH 7.9
11.5 µl 5 M NaCl
5.75 µl 1 M MgCl₂
Prepare fresh just before use

Sheep's blood agar plates

42.5 g blood agar base no. 2 (Sigma, cat. no. B1676)
1 liter H₂O
Autoclave with magnetic stir bar
Cool to ~50°C
While stirring on magnetic stir plate, add 50 ml prewarmed (37°C) defibrinated sheep's blood (Northeast Laboratory Services, <http://www.nelabservices.com/>)
Add 200 µg/ml spectinomycin (Sigma-Aldrich), as needed
Pour into sterile 100 × 15-mm plates

Store up to 2 months at 4°C in the dark

Transposase elution buffer (TEB)

20 mM Tris-Cl, pH 7.4 (APPENDIX 2A)

200 mM NaCl

1 mM EDTA

2 mM DTT

10% (v/v) glycerol

10 mM maltose

1× complete EDTA-free protease inhibitor cocktail (Roche, cat. no. 1 836 170)

Prepare fresh just before use

Transposase wash buffer (TWB)

20 mM Tris-Cl, pH 7.4 (APPENDIX 2A)

200 mM NaCl

1 mM EDTA

2 mM DTT

10% (v/v) glycerol

1× complete EDTA-free protease inhibitor cocktail (Roche, cat. no. 1 836 170)

Prepare fresh just before use

COMMENTARY

Background Information

Understanding how a genotype leads to a specific phenotype is key to understanding how an organism functions and survives. For some time now, it has been clear that the link between genotype and phenotype is not a linear one—rather, interactions among and between genetic components and the environment create a complex pathway leading to a specific phenotype. The result is intriguing in that through these interactions a robust organism emerges that is able to withstand many kinds of genetic and environmental disturbances. Such interactions are described in terms of epistasis, a phenomenon whereby the effects of a given gene on a biological trait are masked or enhanced by one or more other genes (Kitano, 2004; Moore, 2005). In the last decade, such dependencies between genes have been extensively explored in *Saccharomyces cerevisiae* by identifying synthetic genetic interactions through systematic knockout of all possible double combinations of genes. This method has identified specific networks of genes responsible for important pathways and processes such as DNA integrity (Ooi et al., 2003; Pan et al., 2006), mRNA export (Hieronymus et al., 2004), drug resistance (Begley et al., 2002; Giaever et al., 2004; Dudley et al., 2005; Parsons et al., 2006), and phosphorylation (Fiedler et al., 2009).

Three approaches have been developed to identify genetic interactions in *S. cerevisiae*: synthetic genetic array (SGA; Tong et al., 2001), diploid-based synthetic lethality analysis on microarrays (dSLAM; Pan et al., 2004), and epistatic miniarray profile (E-MAP). Where SGA and dSLAM are able to map synthetic sick/lethal (SSL) relationships, E-MAP is more sensitive and able to measure a larger range of epistatic interactions, including positive interactions (Schuldiner et al., 2005; Collins et al., 2006, 2007; Roguev et al., 2008; Wilmes et al., 2008). These methods all make use of robotic automation to construct double mutants by mating. A query strain, with a drug-resistance marker replacing a gene of interest, is crossed to an arrayed collection of single-gene-deletion strains of opposite mating type marked with a different selectable marker. After meiosis, sporulation, and selection for double mutants, genetic interactions can be identified by measuring colony size.

Although meiotic assortment is not applicable to bacteria, two analogous methods have been developed for *E. coli*: GIANT coli and eSGA. Both make use of conjugation to drive genetic exchange of a marked query gene mutation from an *E. coli* Hfr donor strain into a genome-wide, high-density-arrayed collection of *E. coli* F⁻ recipient single-gene mutants. Robotic pinning, dual-marker selection, and colony imaging are used to identify colony growth defects in the resulting double mutants. However, the applicability of these methods to a wide range of microorganisms is greatly limited, primarily because arrayed collections of single-gene knockouts exist for only a handful of species. More generally applicable methods to identify genetic interactions in bacteria include transposon site hybridization (TraSH; Joshi et al., 2006) and related genomic footprinting methods (Girgis et al., 2007). TraSH is based on the construction of a transposon library, the generation of probes from outward-reading T7 promoters located at each end of the transposon, and subsequent detection by microarray. However, the accuracy of this and related methods is somewhat limited, and the use of a microarray restricts its use to the strain for which the microarray is developed.

Development of Tn-seq—Tn-seq is a robust and sensitive method that enables the determination of each gene's contribution to fitness in a specific environment through massively parallel sequencing. The approach is based on the assembly of a saturated transposon insertion library. Upon growth of the library under a test condition of interest (in vitro or in vivo), insertion mutants with reduced fitness will decrease in frequency in the population, while other mutants, depending on the effect of the transposon insertion, remain the same or increase in frequency. Changes in frequency are determined by sequencing the transposon-flanking regions en masse on an Illumina massively parallel sequencing platform. Subsequently, fitness of each individual gene knockout in a bacterial genome can be calculated. Due to the quantitative nature of massively parallel sequencing, fitness of a gene knockout can be quantitatively determined and provides a direct measure of the growth rate of the bacterial clone containing it. For instance, a mutant with a fitness of 0.5 translates to a two-fold higher doubling time compared to wild type (van Opijnen et al., 2009; van Opijnen and Camilli, 2013).

The development of Tn-seq has made it possible to quantitatively determine fitness for most genes in a microorganism, and to screen for quantitative genetic interactions on a genome-wide scale and in high-throughput fashion (van Opijnen and Camilli, 2013). In addition, the

application of Tn-seq to *S. pneumoniae* has led to the identification of essential growth genes, conditionally important genes, regulatory relationships through genetic interactions (van Opijnen et al., 2009; van Opijnen and Camilli, 2012), non-coding RNAs (Mann et al., 2012), and genes important for in vivo colonization of the nasopharynx and survival in the lung and blood in mouse models of infection (Mann et al., 2012; van Opijnen and Camilli, 2012).

Because Mariner transposons work in many organisms including bacteria, *magellan6*-based Tn-seq is applicable to a variety of bacterial species. For bacteria that are naturally competent for transformation, such as *S. pneumoniae*, the in vitro transposition method illustrated here can be used. However, many bacteria are not naturally competent, and thus require a method of in vivo transposition, such as conjugal transfer of a transposon delivery vector like the mTn5 method illustrated here. Because *V. cholerae* is naturally competent, either in vitro or in vivo transposition can be used.

One disadvantage of Mariner transposons is that they have insertion site preference for TA dinucleotides. This prevents the use Mariner transposons for interrogating certain regions of the genome, such as small genes or promoters, that might lack TA dinucleotides. However, a number of other transposons without such insertion site specificity can be used, including the Tn5 derivative illustrated here.

Obtaining a robust fitness value and sequencing multiple samples in a single flow cell lane—The application of massively parallel sequencing in Tn-seq makes it possible to determine the exact location of each transposon insertion in a complex transposon insertion strain library and simultaneously follow the quantitative change during selection of each of these insertion strains in the population over time. The quantitative change is then used to determine the fitness effect of each transposon insertion. However, what makes Tn-seq so robust is that a gene's fitness is based on the average effect of multiple independent transposon insertions into a gene. In Figure 1E.3.3A, the fitness effects of multiple independent transposon insertions from four independent libraries and their overall means are plotted for the region covering genes SP0456 to SP0460, illustrating how robust fitness values are obtained for individual genes.

In this light, it is important to consider the number of transposon insertions and samples that are sequenced in a single flow cell lane. When work with Tn-seq was first started, 6×10^6 reads per flow cell lane could generally be obtained on the GAII platform when combining three libraries containing 10,000 insertion strains each in a single lane. This works out, on average, to 200 reads per insertion, which gives good resolution to determine small changes in frequency of an insertion mutant. Now, due partly to experience and largely to improvements in the Illumina sequencing platform, 1.6×10^8 reads are generally obtained on the HiSeq platform, enabling the combination of up to 24 independent libraries each containing 10,000 mutants. In other experiments, libraries of much higher complexity (on the order of 100,000 mutants) can be utilized, resulting in much higher insertion density and thus more detailed interrogation of the genome. Thus, the cost effectiveness of Tn-seq has increased by more than 10-fold in just a few years.

Improvement in library construction—The transformation protocol described in Basic Protocol 2 represents a significant improvement over that in the original publication of Tn-seq (van Opijnen et al., 2009). Originally, transformation of *S. pneumoniae* called for plating the transformants in a soft agar overlay on top of a blood agar plate, and then overlaying that with a layer containing antibiotic to select for transformants. Using this method, embedded colonies appear after overnight growth, each containing a transposon insertion.

Transformants must then be separated from the soft agar, which is accomplished by chopping up the agar, diluting it in THY, and incubating for several hours. We hypothesized that due to out-competition during these early steps of library construction, mutants with fitness values less than 0.5 were not identified, and changes to the protocol that reduce the number of generations prior to selection at t_1 would increase the dynamic range of the method. The hypothesis turned out to be correct and the solution extremely simple. After a short outgrowth period in liquid medium, the transformed cells can be directly spread onto a blood agar plate containing the selective antibiotic without loss of transformation efficiency. After overnight growth, antibiotic-resistant transformants can simply be washed off the surface of the plates.

Genetic interactions and different environments, strains, and species—In addition to its utility in determining the fitness contribution of genes, Tn-seq is well suited for subsequently identifying genome-wide genetic interactions. To identify interactions between a gene of interest (query gene) and the rest of the genome, the query gene is deleted, and then a saturated transposon library is introduced in this query gene deletion background. The fitness of each double mutant is determined by Tn-seq and compared to the expected fitness value, which is the product of the individual fitness values previously determined using Tn-seq in the wild-type background. A double mutant fitness value that deviates from the expected value reveals a genetic interaction (Avery and Wasserman, 1992; Drees et al., 2005; Collins et al., 2007; Fiedler et al., 2009; van Opijnen et al., 2009).

When performed in a systematic manner, it is possible to untangle complex phenotypes and regulatory relationships, and assign genes of unknown function to molecular pathways through interaction-profile association with known genes (Tong et al., 2001; Tong, 2004; St. Onge et al., 2007; Carter et al., 2009; van Opijnen et al., 2009; van Opijnen and Camilli, 2012). Moreover, Tn-seq can be used to assign a role to intergenic regions (van Opijnen and Camilli, 2013) or discover noncoding RNAs (Mann et al., 2012), and it can be easily applied to different environments (e.g., different growth media, chemical or physical stresses, or infection models) as well as different strains of a particular species. The HimarI Mariner transposon, illustrated here, can be used in many other microorganisms, including *E. coli* (Lampe et al., 1999), *Staphylococcus aureus* (Bae et al., 2008), *Haemophilus influenzae* (Akerley et al., 1998), *Mycobacterium tuberculosis* (Sassetti et al., 2003), and *Bacteroidetes thetaiotaomicron* (Goodman et al., 2009). Moreover, other transposons (such as mTn5, illustrated here) or insertional elements (such as suicide plasmids) can be used in a wide variety of microorganisms. This unit illustrates two methods of preparing Tn-chromosomal junctions for sequencing, one specific for the *MmeI* mutant form of the Mariner transposon *magellan6*, and a more general HTML-PCR-based method that can be applied to any transposon or insertional element. The reader is also referred to three similar methods

developed simultaneously with Tn-seq: INseq for *B. thetaiotaomicron* (Goodman et al., 2009), HITS for *H. influenza* (Gawronski et al., 2009), and TraDISH for *Salmonella enterica* serovar Typhi (Langridge et al., 2009).

Critical Parameters and Troubleshooting

The most important point to consider when performing Tn-seq in any microorganism is whether the selection condition will impose a bottleneck that will result in stochastic loss of insertion strains regardless of their fitness. If the bottleneck is severe (>15% loss of library complexity), there will be commensurate noise in the data that will greatly limit the ability to identify short genes, which will have few insertions to begin with, or genes having subtle or intermediate fitness alterations. Instead, only long genes or genes with very severe fitness alterations will be identified. If a severe bottleneck is present, steps should be taken to eliminate or reduce it. For example, if selection is being performed in vitro, the population size being selected on should be increased and/or the selective pressure should be reduced in severity. If selection is being performed in vivo (i.e., in an animal model of infection), the inoculum size should be increased and/or the innate or adaptive immune system should be reduced in strength. Alternatively, libraries of lower complexity may be used, although each library will sample only a subset of the genome. If a bottleneck cannot be avoided, two methods of data analysis can be used. In the first, the size of the bottleneck is estimated by determining the proportion of insertion mutants that disappeared from a set of neutral insertions (a transposon into a neutral site, such as a degenerate gene, should not have an impact on fitness). Subsequently, to correct for the bottleneck, for each gene the same proportion of insertions is removed from the set of insertions that were lost. In the second, all insertions with a gene are computationally pooled, and a simple ratio of output reads divided by input reads is used to calculate fitness for the gene.

An important point to consider when performing Tn-seq in *S. pneumoniae* is whether the presence of a capsule is necessary for the experiments to be successful. If the capsule is of no interest, an acapsular strain is preferred because transformation efficiency is at least an order of magnitude higher, and therefore large transposon libraries are more easily obtained. However, large transposon libraries have been successfully prepared in capsular strains by scaling up the number of transformations.

In Basic Protocol 2, it is very important to have fast-growing starter cultures, as faster-growing starter cultures yield higher transformation efficiencies. For best results, a starter culture should reach the appropriate OD within 2 hr of seeding the growth medium.

In Basic Protocol 3, it is essential that the genomic DNA be clean, as this increases the efficiency of *MmeI* restriction digestion. In addition, it is essential to perform phenol/chloroform extraction, as omitting this step will result in failure of the subsequent reactions.

In Alternate Protocol 1, in vivo transposition is based on conjugation. However, another commonly used method of in vivo transposition not illustrated here is to induce transposition from a replicating plasmid within the organism of study.

In Alternate Protocol 2, the two nested PCR primers (olj491 and olj492) and custom sequencing primer (HK89) are specific to the mTn5 transposon. When using this method on libraries generated with other insertional elements, these primers must be redesigned accordingly.

Anticipated Results

Robust fitness effects for individual genes are obtained because they are measured by averaging across multiple independent transposon insertions in each gene (Fig. 1E.3.3A). Due to the abovementioned improvement in the transformation protocol during library construction, the dynamic range of Tn-seq in *S. pneumoniae* is significantly increased. Figure 1E.3.3B shows a Pearson correlation of two independent libraries (10,000 insertion strains each) constructed using the improved transformation protocol, illustrating that reproducibility is very high ($R^2 = 0.92$) and mutants with fitness <0.5 are observed. In addition, Figure 1E.3.3C shows the distribution of fitness for all genes in the *S. pneumoniae* TIGR4 genome based on four independent libraries (containing 10,000 insertion mutants each), which also confirms that with the improved transformation protocol mutants with fitness <0.5 are identified.

Time Considerations

The complete procedure for performing Tn-seq in *S. pneumoniae*, excluding purifying the Mariner transposase and massively parallel sequencing, can be performed in 8 days. Transposon library construction (Basic Protocol 1) takes 1 day. Transformation of the library (Basic Protocol 2) takes 4 days total, including preparation and freezing of *S. pneumoniae* starter cultures (2 days), carrying out the transformation (1 day), and preparing and freezing the transposon library starter cultures (1 day). Selection and sample preparation (Basic Protocol 3) take 3 days total, including library selection (1 day) and preparing the library for sequencing (2 days). The complete procedure for performing Tn-seq in *V. cholerae*, excluding sequencing, can be performed in 5 days, including 3 days for transposon library construction and selection (Alternate Protocol 1) and 1 to 2 days to prepare the library for sequencing (Alternate Protocol 2). Sequencing can take anywhere from 2 to 4 days, depending on the Illumina massively parallel sequencing platform used.

Acknowledgments

This work was supported by U.S. National Institutes of Health grant AI055058 and the Howard Hughes Medical Institute. T.v.O. was supported by a postdoctoral fellowship from the Netherlands Organization for Scientific Research (Rubicon-NWO) and the Charles H. Hood Foundation. A.C. is an investigator of the Howard Hughes Medical Institute.

Literature Cited

- Akerley BJ, Rubin EJ, Camilli A, Lampe DJ, Robertson HM, Mekalanos JJ. Systematic identification of essential genes by in vitro Mariner mutagenesis. *Proc Natl Acad Sci USA*. 1998; 95:8927–8932. [PubMed: 9671781]
- Avery L, Wasserman S. Ordering gene function: The interpretation of epistasis in regulatory hierarchies. *Trends Genet*. 1992; 8:312–316. [PubMed: 1365397]

- Bae T, Glass EM, Schneewind O, Missiakas D. Generating a collection of insertion mutations in the *Staphylococcus aureus* genome using *bursa aurealis*. *Methods Mol Biol*. 2008; 416:103–116. [PubMed: 18392963]
- Begley TJ, Rosenbach AS, Ideker T, Samson LD. Damage recovery pathways in *Saccharomyces cerevisiae* revealed by genomic phenotyping and interactome mapping. *Mol Cancer Res*. 2002; 1:103–112. [PubMed: 12496357]
- Burnett LC, Lunn G, Coico R. Biosafety: Guidelines for working with pathogenic and infectious microorganisms. *Curr Protoc Microbiol*. 2009; 13:1A.1.1–1A.1.14. [PubMed: 19412909]
- Carter GW, Galas DJ, Galitski T. Maximal extraction of biological information from genetic interaction data. *PLoS Comput Biol*. 2009; 5:e100347.
- Collins S, Schuldiner M, Krogan N, Weissman J. A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. *Genome Biol*. 2006; 7:R63. [PubMed: 16859555]
- Collins S, Miller K, Maas N, Roguev A, Fillingham J, Chu C, Schuldiner M, Gebbia M, Recht J, Shales M, Ding H, Xu H, Han J, Ingvarsdottir K, Cheng B, Andrews B, Boone C, Berger S, Hieter P, Zhang Z, Brown G, Ingles C, Emili A, Allis C, Toczyski D, Weissman J, Greenblatt J, Krogan N. Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature*. 2007; 446:806–810. [PubMed: 17314980]
- Desjardins PR, Conklin DS. Microvolume quantitation of nucleic acids. *Curr Protoc Mol Biol*. 2011; 93:A.3J.1–A.3J.16.
- Drees BL, Thorsson V, Carter GW, Rives AW, Raymond MZ, Avila-Campillo I, Shannon P, Galitski T. Derivation of genetic interaction networks from quantitative phenotype data. *Genome Biol*. 2005; 6:R38. [PubMed: 15833125]
- Dudley AM, Janse DM, Tanay A, Shamir R, Church GM. A global view of pleiotropy and phenotypically derived gene function in yeast. *Mol Syst Biol*. 2005; 1:1–11.
- Elbing K, Brent R. Growth on solid media. *Curr Protoc Mol Biol*. 2002; 59:1.3.1–1.3.6. [PubMed: 18265292]
- Fiedler D, Braberg H, Mehta M, Chechik G, Cagney G, Mukherjee P, Silva AC, Shales M, Collins SR, von Wageningen S, Kemmeren P, Holstege FC, Weissman JS, Keogh M, Koller D, Shokat KM, Krogan NJ. Functional organization of the *S. cerevisiae* phosphorylation network. *Cell*. 2009; 136:952–963. [PubMed: 19269370]
- Gallagher SR. Quantitation of DNA and RNA with absorption and fluorescence spectroscopy. *Curr Protoc Mol Biol*. 2011; 93:A.3D.1–A.3D.14.
- Gawronski JD, Wong SMS, Giannoukos G, Ward DV, Akerley BJ. Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for *Haemophilus* genes required in the lung. *Proc Natl Acad Sci USA*. 2009; 106:16422–16427. [PubMed: 19805314]
- Giaever G, Flaherty P, Kumm J, Proctor M, Nislow C, Jaramillo DF, Chu AM, Jordan MI, Arkin AP, Davis RW. Chemogenomic profiling: Identifying the functional interactions of small molecules in yeast. *Proc Natl Acad Sci USA*. 2004; 101:793–798. [PubMed: 14718668]
- Girgis H, Liu Y, Ryu W, Tavazoie S. A comprehensive genetic characterization of bacterial motility. *PLoS Genet*. 2007; 3:e154.
- Goodman AL, McNulty NP, Zhao Y, Leip D, Mitra RD, Lozupone CA, Knight R, Gordon JI. Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe*. 2009; 6:279–289. [PubMed: 19748469]
- Hieronymus H, Yu M, Silver PS. Genome-wide mRNA surveillance is coupled to mRNA export. *Genes Dev*. 2004; 18:2652–2662. [PubMed: 15489286]
- Hsieh YC, Wang JT, Lee WS, Hsueh PR, Shao PL, Chang LY, Lu CY, Lee CY, Huang FY, Huang LM. Serotype competence and penicillin resistance in *Streptococcus pneumoniae*. *Emerging Infect Dis*. 2006; 12:1709–1714. [PubMed: 17283621]
- Joshi SM, Pandey AK, Capite N, Fortune SM, Rubin EJ, Sassetti CM. Characterization of mycobacterial virulence genes through genetic interaction mapping. *Proc Natl Acad Sci USA*. 2006; 103:11760–11765. [PubMed: 16868085]
- Kamp HD, Patimalla-Dipali B, Lazinski DW, Wallace-Gadsden F, Camilli A. Gene fitness landscapes of *Vibrio cholerae* at important stages of its life cycle. *PLoS Pathogens*. 2013; 9:e1003800. [PubMed: 24385900]

- Kitano H. Biological robustness. *Nat Rev Genet.* 2004; 5:826–837. [PubMed: 15520792]
- Klein BA, Tenorio EL, Lazinski DW, Camilli A, Duncan MJ, Hu LT. Identification of essential genes of the periodontal pathogen *Porphyromonas gingivalis*. *BMC Genomics.* 2012; 13:578. [PubMed: 23114059]
- Langridge GC, Phan MD, Turner DJ, Perkins TT, Parts L, Haase J, Charles I, Maskell DJ, Peters SE, Dougan G, Wain J, Parkhill J, Turner AK. Simultaneous assay of every *Salmonella typhi* gene using one million transposon mutants. *Genome Res.* 2009; 19:2308–2316. [PubMed: 19826075]
- Lampe D, Churchill M, Robertson H. A purified Mariner transposase is sufficient to mediate transposition in vitro. *EMBO J.* 1996; 15:5470–5479. [PubMed: 8895590]
- Lampe DJ, Akerley BJ, Rubin EJ, Mekalanos JJ, Robertson HM. Hyperactive transposase mutants of the HimarI Mariner transposon. *Proc Natl Acad Sci USA.* 1999; 96:11428–11433. [PubMed: 10500193]
- Lazinski DW, Camilli A. Homopolymer tail-mediated ligation PCR: A streamlined and highly efficient method for DNA cloning and library construction. *Biotechniques.* 2013; 54:25–34. [PubMed: 23311318]
- Mann B, van Opijnen T, Wang J, Obert C, Wang YD, Carter R, McGoldrick DJ, Ridout G, Camilli A, Tuomanen EI, Rosch JW. Control of virulence by small RNAs in *Streptococcus pneumoniae*. *PLoS Pathogens.* 2012; 8:e1002788. [PubMed: 22807675]
- Merrell DS, Hava DL, Camilli A. Identification of novel factors involved in colonization and acid tolerance of *Vibrio cholerae*. *Mol Microbiol.* 2002; 43:1471–1491. [PubMed: 11952899]
- Moore D, Dowhan D. Purification and concentration of DNA from aqueous solutions. *Curr Protoc Mol Biol.* 2002; 59:2.1.1–2.1.10.
- Moore JP. A global view of epistasis. *Nat Genet.* 2005; 37:13–14. [PubMed: 15624016]
- Ooi S, Shoemaker D, Boeke J. DNA helicase gene interaction network defined using synthetic lethality analyzed by microarray. *Nat Genet.* 2003; 35:277–286. [PubMed: 14566339]
- Pan X, Yuan DS, Xiang D, Wang X, Sookhai-Mahadeo S, Bader JS, Hieter P, Spencer F, Boeke JD. A robust toolkit for functional profiling of the yeast genome. *Mol Cell.* 2004; 16:487–496. [PubMed: 15525520]
- Pan X, Ye P, Yuan DS, Wang X, Bader JS, Boeke JD. A DNA integrity network in the yeast *Saccharomyces cerevisiae*. *Cell.* 2006; 124:1069–1081. [PubMed: 16487579]
- Parsons A, Lopez A, Givoni I, Williams D, Gray C, Porter J, Chua G, Sopko R, Brost R, Ho C. Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast. *Cell.* 2006; 126:611–625. [PubMed: 16901791]
- Pozzi G, Masala L, Iannelli F, Manganelli R, Havarstein LS, Piccoli L, Simon D, Morrison DA. Competence for genetic transformation in encapsulated strains of *Streptococcus pneumoniae*: Two allelic variants of the peptide pheromone. *J Bacteriol.* 1996; 178:6087–6090. [PubMed: 8830714]
- Roguev A, Bandyopadhyay S, Zofall M, Zhang K, Fischer T, Collins S, Qu H, Shales M, Park HO, Hayles J, Hoe KL, Kim DU, Ideker T, Grewal SI, Weissman J, Krogan N. Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. *Science.* 2008; 322:405–410. [PubMed: 18818364]
- Sassetti C, Boyd DH, Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol.* 2003; 48:77–84. [PubMed: 12657046]
- Schuldiner M, Collins S, Thompson N, Denic V, Bhamidipati A, Punna T, Ihmels J, Andrews B, Boone C, Greenblatt J. Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell.* 2005; 123:507–519. [PubMed: 16269340]
- St Onge RP, Mani R, Oh J, Proctor M, Fung E, Davis R, Nislow C, Roth F, Giaever G. Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. *Nat Genet.* 2007; 39:199–206. [PubMed: 17206143]
- Stevenson B. Common bacterial culture techniques and media. *Curr Protoc Microbiol.* 2006; 00:A.4A.1–A.4A.8.
- Taubes G. The bacteria fight back. *Science.* 2008; 321:356–361. [PubMed: 18635788]
- Tong A. Global mapping of the yeast genetic interaction network. *Science.* 2004; 303:808–813. [PubMed: 14764870]

- Tong AH, Evangelista M, Parsons A, Xu H, Bader GD, Page N, Robinson M, Raghibizadeh S, Hogue CW, Bussey H, Andrews B, Tyers M, Boone C. Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science*. 2001; 294:2364–2368. [PubMed: 11743205]
- Tuomanen, L.; Mitchel, TJ.; Morrison, DA.; Spratt, BG. *The Pneumococcus*. ASM Press; Washington, D.C.: 2004.
- van Opijnen T, Camilli A. A fine scale phenotype-genotype virulence map of a bacterial pathogen. *Genome Res*. 2012; 22:2541–2551. [PubMed: 22826510]
- van Opijnen T, Camilli A. Transposon insertion sequencing: A new tool for systems-level analysis of microorganisms. *Nat Rev Microbiol*. 2013; 11:435–442. [PubMed: 23712350]
- van Opijnen T, Bodi KL, Camilli A. Tn-seq: High-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat Methods*. 2009; 6:767–772. [PubMed: 19767758]
- Voytas D. Agarose gel electrophoresis. *Curr Protoc Mol Biol*. 2000; 51:2.5A.1–2.5A.9.
- Wilmes GM, Bergkessel M, Bandyopadhyay S, Shales M, Braberg H, Cagney G, Collins S, Whitworth GB, Kress TL, Weissman J, Ideker T, Guthrie C, Krogan N. A genetic interaction map of RNA-processing factors reveals links between Sem1/DssI-containing complexes and mRNA export and splicing. *Mol Cell*. 2008; 32:735–746. [PubMed: 19061648]
- Wilson K. Preparation of genomic DNA from bacteria. *Curr Protoc Mol Biol*. 2001; 00:2.4.1–2.4.5.

Figure 1E.3.1.

Flowchart depicting Tn-seq using the mini-transposon *magellan6*, from library construction to massively parallel sequencing of transposon-chromosome junctions. **(A)** A gene-disruption library is constructed by transposing *magellan6* into bacterial genomic DNA in vitro, and subsequently transforming a bacterial population with the transposed DNA. The bacterial pool is cultured under selective conditions (in vitro or in vivo), and DNA from pre- and post-selection cultures (t_1 and t_2) is isolated. **(B)** To sequence regions flanking the *magellan6* insertions, DNA from both time points is digested with *MmeI*, which binds a sequence in the *magellan6* terminal inverted repeats (hashed boxes) but cuts 20 bp downstream, leaving a two-base overhang to which a barcoded adapter is ligated. PCR is performed using one primer complementary to the adapter and another complementary to the inverted repeat. **(C)** The resulting 120-bp PCR product contains 20 bp bacterial-specific DNA and a barcode, flanked by Illumina-specific sequences for sequencing.

Figure 1E.3.2.

Detailed schema of how the *magellan6*-specific primers and adapter combine to result in a 120-bp DNA fragment that can be sequenced on an Illumina Genome Analyzer II (GAII) platform. AP-B_bc-ACAC (AP-B) and AP-A_bc-ACAC (AP-A) are two oligos that make up the adapter. AP-A has a random two-base overhang (NN) that is employed to ligate the adapter to the genomic DNA (*MmeI* leaves a random two-base overhang). Note that to prevent self-ligation, AP-B is slightly longer than AP-A. In addition, the adapter contains a four-base barcode (bc) that enables mixing of different samples in a single flow cell lane. Using primers P1-M6-MmeI and Gex PCR Primer 2, a 120-bp fragment is generated that contains Illumina-specific sequences (black tails with asterisk) necessary for sequencing. The Gex Sequencing primer is then used for sequencing samples loaded on the flow cell. (See Tables 1E.3.1 and 1E.3.2 for oligo details.)

Figure 1E.3.3.

Tn-seq fitness analysis. (A) A gene's fitness is based on the average effect of multiple independent transposon insertions into a gene. On the x axis, the genomic region for *S. pneumoniae* TIGR4 is plotted covering genes SP0456 to SP0460. Black data points depict fitness of a single transposon insertion within the coding region of a gene. A black line depicts the average fitness of a specific gene calculated from the individual insertions into that gene. For instance, the average fitness for genes SP0456, SP0458, and SP0460 is ~1 (same as wild type), whereas insertions into genes SP0457 and SP0459 are clearly

disadvantageous, with average fitness values well below 1. Gray data points depict fitness from insertions into intergenic regions. **(B)** High reproducibility ($R^2 = 0.92$) between independently generated libraries consisting of 10,000 transposon insertions each. **(C)** Fitness for all genes across the *S. pneumoniae* genome.

Table 1E.3.1Adapter Oligonucleotide Sequences for Processing *magellan6* Insertion Libraries

Adapter	Sequence (5' to 3')	Purification	Platform
AP-A_bc-ACAC ^a	GTT CAG AGT TCT ACA GTC CGA CGA TCA <u>CAC</u> NN	PAGE	GAII
AP-B_bc-ACAC	p- <u>GTG</u> T GA TCG TCG GAC TGT AGA ACT CTG AAC CTG TC-p	PAGE	GAII
ADBC-F-INDX1a ^b	ACA CGA CGC TCT TCC GAT CTG GAA CTC CTA <u>CTG AC</u> NN	PAGE	HiSeq
ADBC-F-INDX1b	ACA CGA CGC TCT TCC GAT CTT AGC ACA TGC <u>TGA CT</u> NN	PAGE	HiSeq
ADBC-F-INDX1c	ACA CGA CGC TCT TCC GAT CTC CTTGAGGAT <u>GAC TG</u> NN	PAGE	HiSeq
ADBC-F-INDX1d	ACA CGA CGC TCT TCC GAT CTA TCGTGT ACG <u>ACT GA</u> NN	PAGE	HiSeq
ADBC-R-INDX1a	p- <u>GTC AGT</u> AGG AGT TCC AGA TCG GAA GAG CGT CGT GTA GGG A-p	PAGE	HiSeq
ADBC-R-INDX1b	p- <u>AGT CAG</u> CAT GTG CTA AGA TCG GAA GAG CGT CGT GTA GGG A-p	PAGE	HiSeq
ADBC-R-INDX1c	p- <u>CAG TCA</u> TCC TCA AGG AGA TCG GAA GAG CGT CGT GTA GGG A-p	PAGE	HiSeq
ADBC-R-INDX1d	p- <u>TCA GTC</u> GTA CAC GAT AGA TCG GAA GAG CGT CGT GTA GGG A-p	PAGE	HiSeq

^aThe underlined sequences represent the barcodes (ACAC and its complement, GTGT). The NN two-base random nucleotide overhang in the AP-A sequence allows for ligation to the ends of *Mme*I-digested DNA fragments. The phosphorylated 5'-end of the AP-B sequence allows ligation to dephosphorylated genomic DNA fragments. Note that that latter sequence is five nucleotides longer on the 3' end than the former, which prevents adapter-adapter ligation. The phosphorylated 3' end also prevents formation of adapter-adapter products.

^bDue to several technological differences between the Genome Analyzer II (GAII) and HiSeq platforms, the adapters were redesigned and optimized for the HiSeq platform. The single letter at the end of each adapter name indicates the oligo pairs that are annealed together to form an adapter (e.g., ADBC-F-INDX1a is annealed with ADBC-R-INDX1a). The 9-nt sequence in bold is an initiation sequence that is not used in the analysis. It was incorporated because the HiSeq platform is less accurate in the first cycles (reaching its highest accuracy after the first seven to nine cycles). Additionally, the HiSeq is very sensitive to repetitive sequences in the first sequenced bases. Therefore, these adapters are most optimal when used in blocks of four. For instance, adapters 1a, 1b, 1c, and 1d together have a perfect 1:1:1:1 distribution of each nucleotide in the first nine nucleotides, which greatly improves cluster separation by the HiSeq platform. Directly behind this initiation sequence is the 6-nt barcode (underlined) used for multiplexing. The longer barcode is necessary because of the decreased accuracy of sequencing with the HiSeq platform relative to the Genome Analyzer II.

Table 1E.3.2PCR and Illumina Sequencing Primers for Processing *magellan6* Insertion Libraries

Primer	Sequence (5' to 3')	Purification	Platform
P1_M6_MmeI ^{a,b}	CAA GCA GAA GAC GGC ATA CGA <u>AGA CCG GGG ACT TAT</u> CAT CCA ACC TGT	Standard desalting	GAII, HiSeq
Gex PCR Primer 2 ^{b,c}	AAT GAT ACG GCG ACC ACC GAC AGG TTC AGA GTT CTA CAG TCC GA	Standard desalting	GAII
Gex Sequencing Primer ^d	CGA CAG GTT CAG AGT TCT ACA GTC CGA CGA TC	Standard desalting	GAII
ADPT-Tnseq-PCRPrimer ^e	AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATCT	Standard desalting	HiSeq
Genomic DNA Sequencing Primer ^f	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT	Standard desalting	HiSeq

^aThe underlined sequence is the complement of the Magellan inverted repeat.

^bThe italicized regions are Illumina-specific sequences necessary for annealing to oligos present in the flow cell.

^cUsed in combination with P1_M6_MmeI when using GAII adapters.

^dIdentical to the Illumina small RNA primer.

^eUsed in combination with P1_M6_MmeI when using HiSeq adapters.

^fIdentical to the Illumina genomic DNA sequencing primer.

Table 1E.3.3

PCR and Illumina Sequencing Primers for Processing mTn5 Insertion Libraries

Primer	Sequence (5'to 3')	Purification
olj491 ^a	<u>ACC TGC AGG CAT GCA AGC TTC GGC C</u>	Standard desalting
olj376 ^{b,c}	GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC <u>TGG GGG GGG GGG</u> <u>GGG GG</u>	Standard desalting
olj492 ^{a,c}	<i>AAT GAT ACG GCG ACC ACC</i> GAG ATC TAC ACT CTT <u>TCG CGG CCG CAC TTG</u> <u>TGT ATA AGA GT</u>	HPLC
Barcode primer ^{c,d}	CAA <i>GCA GAA GAC GGC ATA CGA</i> GAT NNN NNN GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC T	HPLC
Custom sequencing primer HK89 ^e	ACA CTC TTT CGC GGC CGC ACT TGT GTA TAA GAG TCA G	Standard desalting

^a Nested primers specific to one end of mTn5; the underlined sequence is complementary to mTn5.

^b The underlined sequence is complementary to the poly-dC tail.

^c The italicized regions are Illumina-specific sequences necessary for annealing to oligos present in the flow cell.

^d NNNNNN represents the reverse complement of the barcode and is varied with each sample.

^e Specific to the mTn5 inverted repeat end sequence.