Supporting Information

Turner et al. 10.1073/pnas.1419677112

SI Materials and Methods

Tn-Seq Illumina Library Preparation from Initial PA14 Mutant Library. To prepare DNA from the PA14 transposon library for Tn-seq analysis transposon mutants were scraped from solid BHI agar into 1 mL 1× buffer A (1) + 0.1% SDS and homogenized in a Mini-Beadbeater (Biospec) in 2-mL vials preloaded with Lysing Matrix B (MP Biomedicals) three to five times for 1 min each, resting on ice in between each pulse. Proteinase K was then added to 1 mg/mL and samples were incubated overnight at 50 °C. Samples were then homogenized once more as above, separate sections from the same wound were pooled, and samples were extracted with an equal volume of 25:24:1 phenol: chloroform:isoamyl alcohol, pH 8.0. DNA was ethanol-precipitated from the aqueous phase and was resuspended in 200–500 μL water after extensive pellet washing with 75% ethanol. Tn-seq sequencing libraries were prepared by a modified version of a published protocol (2). DNA was sheared to ∼400 bp in a Q880R Sonicator (Qsonica). DNA (2.5 μg) was treated with terminal deoxynucleotidyl transferase (TdT; Promega) in a 50-μL reaction with 2.5 μL 9.5 mM dCTP/0.5 mM ddCTP, 10 μL 5× TdT reaction buffer, and 1.25 μL rTdT at 37 °C for 1 h. DNA was then purified with 0.8× volume Agencourt AMPure XP beads (Beckman Coulter, Inc.) and eluted into 30 μL water. Five hundred nanograms of purified, TdT-treated DNA was used as template in a PCR using the primer olj376 (5′-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGGGGGGGGGGGGGGG-3′) and the 5′-biotinylated primer K2PTS-T8OE-1-bio (5′-GGGTTTTCCCAGTCACGACGTTG-3′) and performed with the Expand Long Template PCR System using buffer 2 (Roche) using the following program: 95 °C, 5 min; 10× (94 °C, 30 s; 60 °C, 30 s; 68 °C, 2 min); 68 °C, 10 min. DNA was then purified with 0.8× volume Agencourt AMPure XP beads (Beckman Coulter, Inc.) and eluted into 50 μL water. Biotinylated PCR products were bound to Streptavidin-coupled Dynabeads (Invitrogen) as described (1). The beads were then resuspended in a PCR mix containing the primers K2PTS-T8OE-2 (5′-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNCGTCCAGGACGCTACTTGTG -3′) and BC41 (5′-CAAGCAGAAGACGGCATACGAGATATTATAGTGACTGGAGTTCAGACGTGTG-3′) performed with the Expand Long Template PCR System using buffer 2 (Roche) using the following program: 95 °C, 5 min; 15× (94 °C, 30 s; 58 °C, 30 s; 68 °C, 2 min); 68 °C, 10 min. The final library in the supernatant was removed from the beads on a magnetic stand, purified with 0.8× volume Agencourt AMPure XP beads (Beckman Coulter, Inc.), and eluted into 15 μL water. Libraries were sequenced at the Genome Sequencing and Analysis Facility at the University of Texas at Austin on a HiSEq. 2000 (Illumina) on a 1 × 100 single end run.

Essential Gene Analysis. First, individual insertion sites and the number of reads originating from them were tallied as described (3). When a read could not be unambiguously mapped to one site (e.g., in the case of duplicated genes or regions) it was assigned randomly to one of the matching sites. Thus, duplicated genes cannot be reliably queried with this method. The data analysis method, contained in the Unix script TnSeqEssential.sh and the R script TnSeqDESeqEssential.R was inspired largely by the ESSENTIALS software package (4) and is described in detail below. After removing the 50 most abundant insertion sites from analysis to correct for amplification bias, insertion location versus number of reads data were smoothed using locally weighted LOESS smoothing using a smoothing parameter (α) of 1 to correct for genomic position-dependent effects on apparent mutant abundance. Then, the insertion locations of all sites were randomized to generate an "expected" dataset, and this was repeated 2,000 times (results began to converge after 50 of these Monte Carlo simulations). Then, real and expected data were normalized using estimateSizeFactors() in DESeq version 1.16.0 (5) with default parameters. For gene knockout abundance analysis, a modified annotation was generated with the 3′ 10% of every gene removed (to ignore insertions that may not abolish gene function). Then, the smoothed, normalized number of transposon-derived reads per gene and the number of insertions mapping to each gene were tallied using this modified annotation in Perl. The number of transposon-derived reads was incremented by one for each gene to avoid dividing by zero when comparing to a condition with no mutants detected. Finally, differential mutant abundance between the real data and the expected pseudodata was calculated using a negative binomial test with DESeq, artificially setting normalization factors to 1 (because the data were normalized per insertion) and adjusting the P value for multiple testing correction with the Benjamini–Hochberg method (6). This yielded a typically bimodal-appearing distribution (Fig. 1*A*). To determine whether a gene belonged to the "reduced" or "unchanged" mode, a parameterized bimodal Gaussian mixture model was fit to the $\log_2$-transformed fold change mutant abundance with the R mclust package (7). A gene was declared essential if it met two criteria: (*i*) it was significantly less abundant in the real data than in the expected pseudodata ($P < 0.01$, negative binomial test) and (*ii*) the posterior probability of it's belonging to the reduced mode was less than 0.01. For comparative analyses, only sites that were reliably identified in at least two (PAO1) or three (PA14) of four total replicates were considered for downstream analysis. Tn-seq data were comparatively analyzed as previously described (3), except that the P value was adjusted for multiple testing correction with the Benjamini–Hochberg method (6). All analyses were performed on the Stampede supercomputer at the Texas Advanced Computing Center.

SCFM2 Preparation. SCFM, which contains the average concentrations of ions, free amino acids, glucose, and lactate measured in CF sputum samples, was prepared as previously described (8). To make SCFM2, macromolecules that are present in CF sputum samples were added to SCFM. The concentration of macromolecules in SCFM2 were based on reported measurements of DNA (9–13), lipids (14–16), GlcNAc (17), and mucin (18–20) detected in CF sputum samples. To represent these macromolecules, 600 $\mu g \cdot mL^{-1}$ salmon sperm DNA, 300 μM GlcNAc, 5 $mg \cdot mL^{-1}$ bovine maxillary mucin, and 100 $\mu g \cdot mL^{-1}$ DOPC were added to SCFM to produce SCFM2. Salmon sperm DNA was extracted with phenol:chloroform:isoamylalcohol (25:24:1), pH 8.0, and precipitated with cold ethanol. Once completely dry, the pellet was UV-sterilized for ∼20 min. Bovine maxillary mucin was sterilized under UV for ∼4 h, with occasional mixing using sterile sticks. DOPC was resuspended in chloroform and added to the media before inoculation. To allow chloroform to evaporate, sterile media was shaken at 37 °C for at least 30 min. Sterilized DNA, mucin, and DOPC were stored at −20 °C before use. Step-by-step instructions for SCFM2 preparation are presented in Dataset S8.

1. Goodman AL, Wu M, Gordon JI (2011) Identifying microbial fitness determinants by insertion sequencing using genome-wide transposon mutant libraries. *Nat Protoc* 6(12):1969–1980.
2. Klein BA, et al. (2012) Identification of essential genes of the periodontal pathogen *Porphyromonas gingivalis*. *BMC Genomics* 13:578.
3. Turner KH, Everett J, Trivedi U, Rumbaugh KP, Whiteley M (2014) Requirements for *Pseudomonas aeruginosa* acute burn and chronic surgical wound infection. *PLoS Genet* 10(7):e1004518.
4. Zomer A, Burghout P, Bootsma HJ, Hermans PW, van Hijum SA (2012) ESSENTIALS: Software for rapid analysis of high throughput transposon insertion sequencing data. *PLoS ONE* 7(8):e43012.
5. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106.
6. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - A practical and powerful approach to multiple testing. *J Roy Stat Soc B Met* 57(1):289–300.
7. Fraley C, Raftery AE (1999) MCLUST: Software for model-based cluster analysis. *J Classif* 16(2):297–306.
8. Palmer KL, Aye LM, Whiteley M (2007) Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum. *J Bacteriol* 189(22):8079–8087.
9. Ratjen F, et al. (2005) DNA concentrations in BAL fluid of cystic fibrosis patients with early lung disease: Influence of treatment with dornase alpha. *Pediatr Pulmonol* 39(1):1–4.
10. Kirchner KK, Wagener JS, Khan TZ, Copenhaver SC, Accurso FJ (1996) Increased DNA levels in bronchoalveolar lavage fluid obtained from infants with cystic fibrosis. *Am J Respir Crit Care Med* 154(5):1426–1429.
11. Brandt T, Breitenstein S, von der Hardt H, Tümmler B (1995) DNA concentration and length in sputum of patients with cystic fibrosis during inhalation with recombinant human DNase. *Thorax* 50(8):880–882.
12. Wagener JS, et al.; Pulmozyme Pediatric Broncoscopy Study Group (1998) Aerosol delivery and safety of recombinant human deoxyribonuclease in young children with cystic fibrosis: A bronchoscopic study. *J Pediatr* 133(4):486–491.
13. Henke MO, John G, Germann M, Lindemann H, Rubin BK (2007) MUC5AC and MUC5B mucins increase in cystic fibrosis airway secretions during pulmonary exacerbation. *Am J Respir Crit Care Med* 175(8):816–821.
14. Hull J, South M, Phelan P, Grimwood K (1997) Surfactant composition in infants and young children with cystic fibrosis. *Am J Respir Crit Care Med* 156(1):161–165.
15. Postle AD, et al. (1999) Deficient hydrophilic lung surfactant proteins A and D with normal surfactant phospholipid molecular species in cystic fibrosis. *Am J Respir Cell Mol Biol* 20(1):90–98.
16. Meyer KC, et al. (2000) Function and composition of pulmonary surfactant and surfactant-derived fatty acid profiles are altered in young adults with cystic fibrosis. *Chest* 118(1):164–174.
17. Korgaonkar AK, Whiteley M (2011) *Pseudomonas aeruginosa* enhances production of an antimicrobial in response to N-acetylglucosamine and peptidoglycan. *J Bacteriol* 193(4):909–917.
18. Fung C, et al. (2010) Gene expression of *Pseudomonas aeruginosa* in a mucin-containing synthetic growth medium mimicking cystic fibrosis lung sputum. *J Med Microbiol* 59(Pt 9):1089–1100.
19. Sriramulu DD, Lünsdorf H, Lam JS, Römling U (2005) Microcolony formation: A novel biofilm model of *Pseudomonas aeruginosa* for the cystic fibrosis lung. *J Med Microbiol* 54(Pt 7):667–676.
20. Ghani M, Soothill JS (1997) Ceftazidime, gentamicin, and rifampicin, in combination, kill biofilms of mucoid *Pseudomonas aeruginosa*. *Can J Microbiol* 43(11):999–1004.

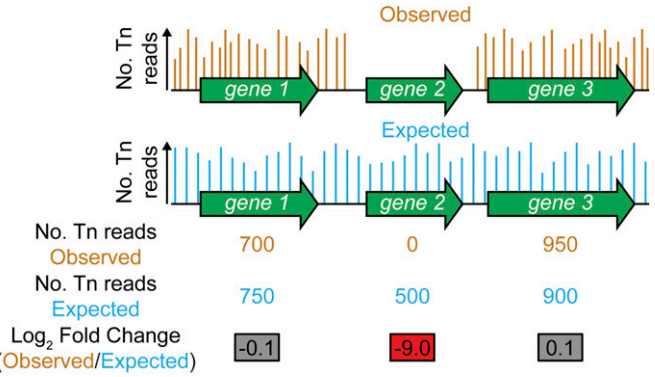

Fig. S1. Explanation of the Monte Carlo method used to determine gene essentiality. In this approach, detected transposon insert sites and their associated read counts per site ("Observed") are randomly distributed across the genome to generate a fitness-neutral pseudodataset ("Expected"). Reads per gene were then tallied for each gene in the genome, and genes with significantly fewer reads observed than expected were declared essential (see *Materials and Methods* for details).

Other Supporting Information Files

Dataset S1 (XLSX)
Dataset S2 (XLSX)
Dataset S3 (XLSX)
Dataset S4 (XLSX)
Dataset S5 (XLSX)
Dataset S6 (XLSX)
Dataset S7 (XLSX)
Dataset S8 (XLSX)
Dataset S9 (XLSX)
Dataset S10 (XLSX)