nature portfolio

https://doi.org/10.1038/s41596-021-00531-3

Supplementary information

Rapid ordering of barcoded transposon insertion libraries of anaerobic bacteria

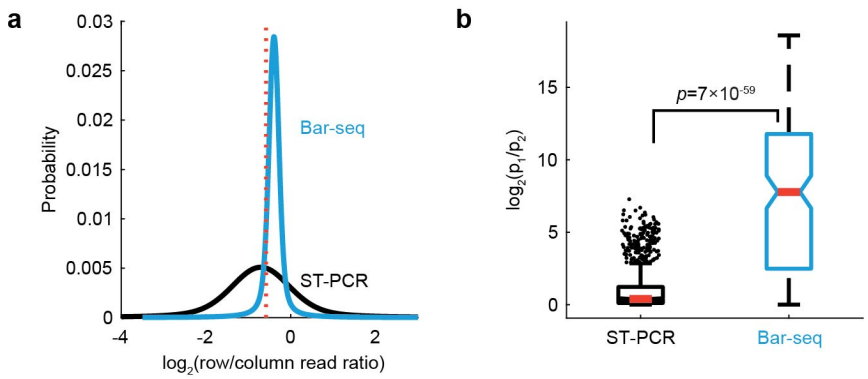
In the format provided by the authors and unedited

Supplementary Information for "Rapid ordering of barcoded transposon insertion libraries of anaerobic bacteria"

Anthony L. Shiver[1], Rebecca Culver[2], Adam M. Deutschbauer[3,4], Kerwyn Casey Huang[1,5,6,*]

[1]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA
[2]Department of Genetics, Stanford University, Stanford, CA 94305, USA
[3]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[4]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA
[5]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA
[6]Chan Zuckerberg Biohub, San Francisco, CA 94158

[*]Correspondence: kchuang@stanford.edu

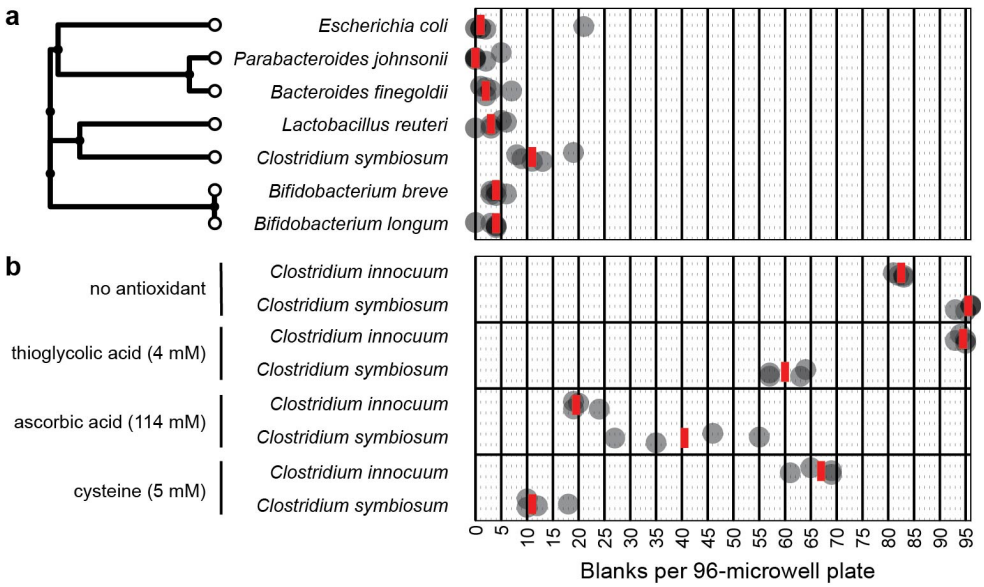


Supplementary Figure 1

The quantitative accuracy of Bar-seq improves probabilistic solutions to barcode locations.

a) The $\log_2$(ratio of row/column reads) should be a Voigt distribution centered around -0.585[2] (red dotted line). The fitted Voigt distributions for Bar-seq (blue, this study) and ST-PCR (black, our re-analysis of previous work[2]) were normalized to an area under the curve of 1 and plotted.

b) The fitted Voigt distributions in (a) were used to predict the likelihood of row and column solutions for all mutant strains that appeared in exactly two row and two column pools. There are two unique solutions to this 2x2 pool inclusion pattern. The log of the ratio of the probabilities for the more (p_1) and less (p_2) likely solution ($\log_2(p_1/p_2)$) is plotted for Bar-seq (blue, n=169) and ST-PCR (black, n=1451). The red dashes are the medians of the distributions, the edges of the boxes are the upper and lower quartiles, whisker length is 1.5 times the inter-quartile range, and the outliers are plotted individually. Bar-seq has a significantly higher log ratio, signifying better predictions.

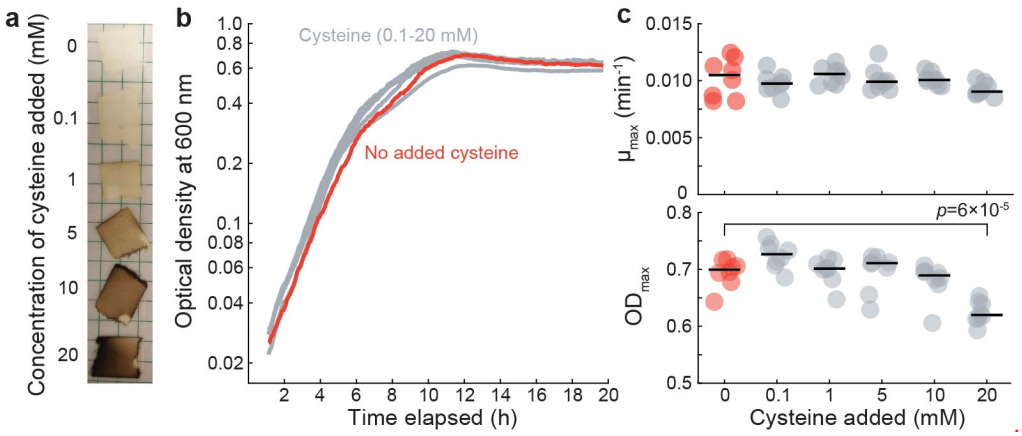


Supplementary Figure 2

Survival of anaerobic species during sorting depends on the presence of antioxidants.

a) Six human gut isolates and a lab strain of *E. coli* survived sorting and regrowth in commercial media. Each strain was grown in a commercial formulation of media with added supplements. *Escherichia coli* BW25113, *Parabacteroides johnsonii* CL02T12C29, *Bacteroides finegoldii* CL09T03C10, and *Lactobacillus reuteri* CF48-3A were grown in Gifu Anaerobic Medium (GAM) supplemented with 1.5 µM hemin, 30 µM menadione, 6 mM arginine, and 1 mM tryptophan. *Clostridium symbiosum* WAL-14163 was grown in Reinforced Clostridial Medium (RCM) supplemented with 1.5 µM hemin, 30 µM menadione, 6 mM arginine, and 1 mM tryptophan. *Bifidobacterium breve* DSM20213 and *Bifidobacterium longum* NCC2705 were grown in de Mann, Rogosa, Sharp (MRS) medium. Each strain was sorted into five 96-microwell plates. Dark circles are the number of blank wells in one 96-microwell plate after outgrowth. Red lines are the median value for the five replicate 96-microwell plates. All strains except *E. coli* and *L. reuteri* are obligate anaerobes[1]. Phylogenetic distances were extracted from the TimeTree database[4].

b) Two highly oxygen-sensitive strains, *Clostridium symbiosum* WAL-14163 and *Clostridium innocuum* 6_1_30, were sorted and grown in media with an antioxidant. For both strains, we used an alternative formulation of RCM with no cysteine and no agar. We then added either no antioxidant, 4 mM thioglycolic acid, 114 mM ascorbic acid, or 5 mM cysteine. Dark circles are the number of blank wells in one 96-microwell plate after outgrowth. Red lines are the median value for the four replicate 96-microwell plates. It is possible to achieve high levels of viability for both of the strains by using ascorbic acid for *C. innocuum* and cysteine for *C. symbiosum*.



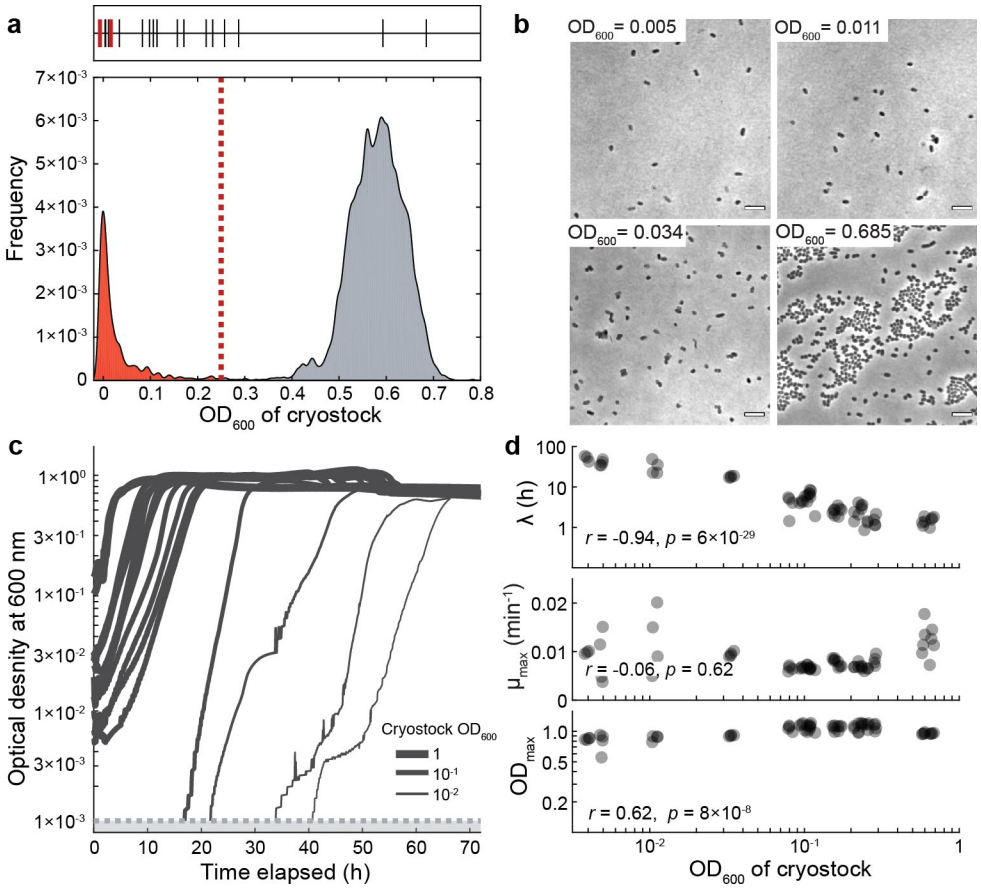
Supplementary Figure 3

***B. thetaiotaomicron* produces hydrogen sulfide from exogenous cysteine.**

a) *B. thetaiotaomicron* produces hydrogen sulfide proportional to the concentration of cysteine added to the medium. The original recipe for BHIS included supplementation with 8 mM cysteine. BHIS media with a range of cysteine concentrations were prepared and inoculated with *B. thetaiotaomicron*. After 24 h of growth, a hydrogen sulfide test strip was held over the culture for 5 s. Reaction of hydrogen sulfide with lead acetate in the strips led to black discoloration. Production of hydrogen sulfide by *B. thetaiotaomicron* in response to added cysteine caused us to investigate the effect of dropping cysteine out of the recipe for BHIS for the sorting procedure.

b) Anaerobic growth of *B. thetaiotaomicron* is not strongly impacted by the addition of cysteine. A 96-well microwell plate was filled with BHIS and increasing levels of cysteine (0-20 mM). The media were inoculated with *B. thetaiotaomicron* VPI-5482 and growth was monitored by reading optical density at 600 nm. The growth curves show the median value for (n=8) independent measurements.

c) The maximum growth rate (μ_{max}) (top) and maximum OD_{600} (OD_{max}) (bottom) were extracted from individual growth curves (n=8). The individual data points (circles) and median (solid line) are plotted for each concentration of cysteine. The omission of cysteine (red circles) did not significantly alter the maximum growth rate or the maximum OD_{600}. The only significant comparison between the no-cysteine condition and the cysteine-added conditions was with 20 mM cysteine for the OD_{max} parameter (multiple hypothesis corrected one-way ANOVA, p=6x10^{-5}).



Supplementary Figure 4

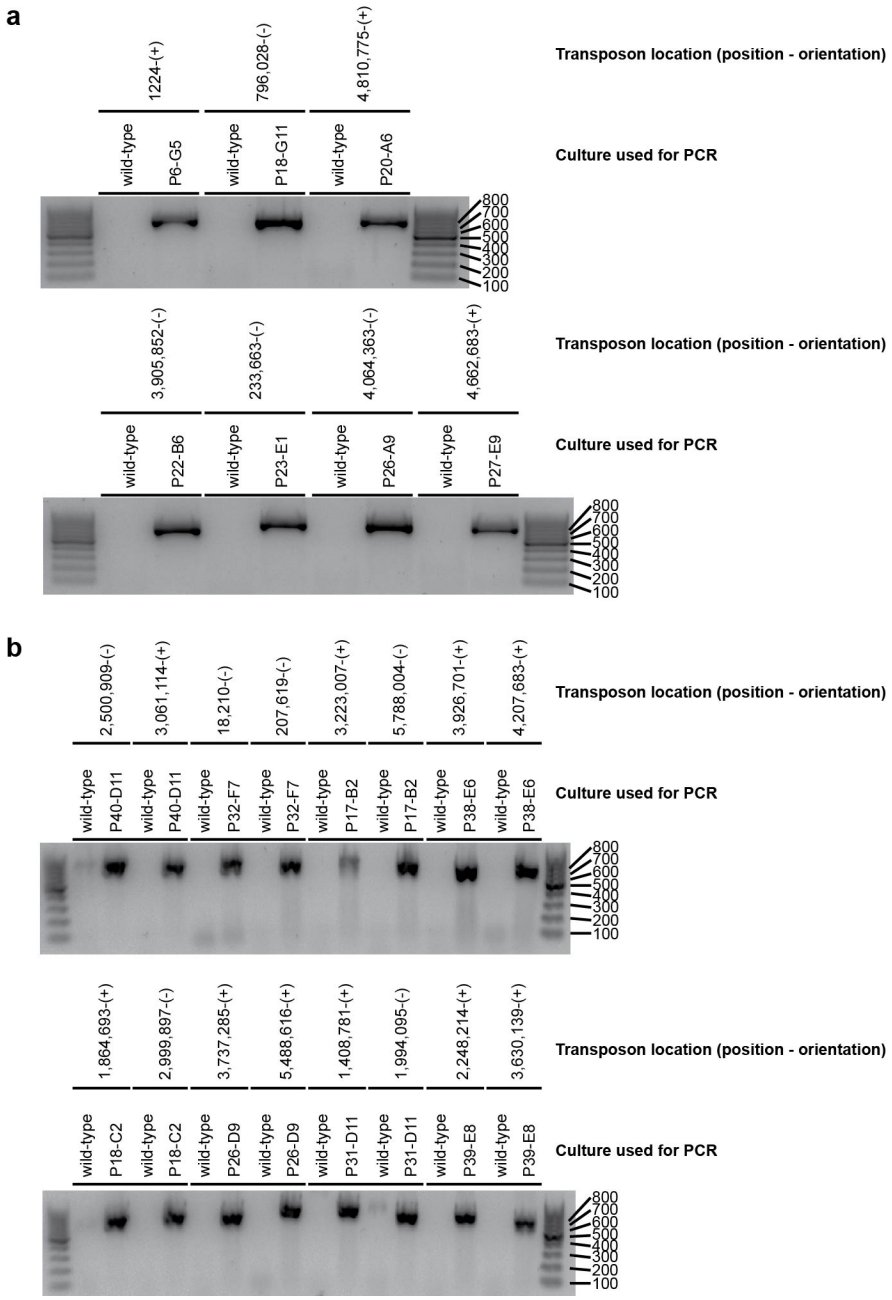
Growth properties of the 40-plate *B. thetaiotaomicron* barcoded transposon insertion library.

a) The distribution of OD_{600} measurements in the cryostocks of the ordered library revealed a surprisingly large population of low OD_{600} wells. A heuristic cutoff of 0.25 (red dotted line) was chosen to discriminate between high (grey distribution) and low OD_{600} (red distribution) in the library. Tick marks at the top of the distribution denote the OD_{600} of wells that were chosen for investigation. Black tick marks record wells in which cells were found by microscopy, while the two red tick marks represent the OD_{600} of wells in which no cells could be found in a 1-µL aliquot.

b) Cells were found using microscopy in most of the low OD_{600} wells in the library, even in wells with OD_{600}<0.01, indicating that many of the low OD_{600} wells are not due to a problem with sorting. Scale bar: 5 µm. The cryostock OD_{600} in the library is shown for each image/well. Images were acquired with a Nikon Ti-E outfitted with a 100X (NA 1.4) objective and an Andor Zyla 5.5 sCMOS camera. Micromanager was used for image acquisition[5].

c) Growth curves after recovery from the cryostock showed that many of the low OD_{600} wells can recover to wild-type densities. The thickness of the line is proportional to the log of the cryostock OD_{600}. All curves represent the mean of (n=4) independent measurements. The horizontal line at an OD_{600} of 0.001 represents the detection limit of the plate reader. The two wells in which cells were not detected did not recover during the 72 h period of monitoring and are not plotted.

d) Growth parameters extracted from the recovery curves (n=4) show that the transposon insertion strains isolated from low OD_{600} wells have maximum OD_{600} values (OD_{max}) and maximum growth rates (μ_{max}) similar to wild-type. As expected, the lag phase (λ) of the cultures was negatively correlated with the cryostock OD_{600} in the library (Pearson's r=-0.94, p=6x10^{-29}).



Supplementary Figure 5

PCR validations of 15 transposon insertion strain assignments within the ordered library.

PCRs were used to confirm that the protocol correctly assigns transposon insertion strains to locations in the ordered library. PCRs were performed with a universal outward facing primer that bound within the Erm cassette and a specific reverse primer that bound to the genome near the insertion point.

a) Seven PCR validations of wells with a single strain with a single transposon insertion. The top label denotes the transposon insertion location that was targeted using a position-(orientation) notation. The bottom label records the culture used as input in the PCR. Either an overnight culture of wild-type *B. theta* VPI-5482 or an aliquot of the cryostock in the ordered library at the position indicated was used as template for the PCR. For positions in the library, P# abbreviates plate number, the letter designates row A-H in the 96-well format, and the number designates column 1-12 in the 96-well format. All seven wells contained the assigned transposon insertion.

b) Eight PCR validations of wells with a single strain with more than one insertion. Two insertion sites were checked for each well. The lanes are labeled as in (a). All eight wells contained both of the tested transposon insertions.

Supplementary Results

Survival of anaerobic species during sorting depends on the presence of antioxidants

To determine whether we could use cell sorting to generate ordered libraries of anaerobic bacteria, we first sought to quantify the ability of bacteria from the gut microbiota, particularly strict anaerobes, to survive the process of cell sorting. We chose two Bacteroidetes (*Bacteroides finegoldii* and *Parabacteroides johnsonii*), two Firmicutes (*Lactobacillus reuteri* and *Clostridium symbiosum*), and two Actinobacteria (*Bifidobacterium breve* and *Bifidobacterium longum*) to test survival of the sorting process; both *Bifidobacterium* species, *C. symbiosum, P. johnsonii*, and *B. finegoldii* are all obligate anaerobes while *L. reuteri* is a facultative anaerobe[1]. As a control, the facultative anaerobe *Escherichia coli* BW25113 was also sorted. Strains were grown and sorted into commercial media formulations of MRS (*B. breve* and *B. longum*), RCM supplemented with hemin, menadione, tryptophan, and arginine (*C. symbiosum*), and GAM supplemented with hemin, menadione, tryptophan, and arginine (*P. johnsonii*, *B. finegoldii*, *L. reuteri*, and *E. coli*).

All plates and cultures spent 1-1.5 h in an aerobic atmosphere while single cells were sorted into five 96-microwell plates for each strain. We counted the number of blank wells after outgrowth in each 96-well plate as a proxy for the number of single cells that did not survive the sorting procedure. We found that even the strict anaerobes survived the aerobic exposure long enough to complete the cell sorting protocol and regrow at high frequency when returned to anaerobic conditions. The fraction of blank wells after sorting varied across species, with *C. symbiosum* having the highest fraction of blank wells (**Supplementary Figure 2a**); nonetheless, at least 89% of *C. symbiosum* cells grew after sorting.

Each of the commercial media formulations we used contained a significant concentration of cysteine, which we hypothesized was acting as an antioxidant to protect the cells during their transient exposure to oxygen. To test this hypothesis and explore the utility of alternative antioxidants, we sorted *C. symbiosum* and *Clostridium innocuum* in formulations of RCM in which cysteine had been either omitted or replaced with another antioxidant. We tested cysteine, thioglycolic acid (a component of GAM), and ascorbic acid for their ability to protect the two species. In the absence of any added antioxidants, the survival of both species was greatly reduced, consistent with their sensitivity to oxygen (**Supplementary Figure 2b**). Both species were protected by the addition of antioxidants. Interestingly, *C. innocuum* preferred ascorbic acid over cysteine while *C. symbiosum* had the opposite preference (**Supplementary Figure 2b**). Thioglycolic acid was the worst performing antioxidant of those tested for both species (**Supplementary Figure 2b**). It remains to be determined if the optimal bacterium-antioxidant pair is idiosyncratic or if there is a general, as yet unidentified, protectant that can be applied across species. Nonetheless, it was straightforward to achieve high survival frequency for both of these strict anaerobes.

B. thetaiotaomicron produces hydrogen sulfide from exogenous cysteine

To test whether *B. thetaiotaomicron* produces hydrogen sulfide from cysteine, we grew cultures of *B. thetaiotaomicron* to saturation in BHI supplemented with hemin (BHIS) and increasing concentrations of added cysteine, from 1-20 mM. We used test strips to detect the presence of hydrogen sulfide in the headspace of the cultures. Reaction of hydrogen sulfide with lead acetate in the test strip led to a black lead sulfide product. Hydrogen sulfide was produced proportionally to the concentration of added cysteine (**Supplementary Figure 3a**).

We then tested growth of *B. thetaiotaomicron* in BHIS as a function of added cysteine and found that cysteine could be left out of the BHIS recipe without adverse effects on anaerobic growth. The growth curve of *B. thetaiotaomicron* grown in BHIS without cysteine was similar to BHIS with cysteine added (**Supplementary Figure 3b**). Calculation of the maximum growth rate and the maximum OD_{600} (optical density at 600 nm) also showed no significant differences between BHIS with or without added cysteine (**Supplementary Figure 3c**).

Limited growth in certain wells is likely due to transient oxygen stress during sorting

After aliquoting and storing the 40-plate library, we measured the OD_{600} after ~60 h of growth. In 14% of the wells, the OD_{600} was <0.25, but in many cases was significantly higher than the expected OD_{600} of a blank well (**Supplementary Figure 4a**). Microscopic examination of the cultures in these low OD_{600} wells revealed that the majority still had detectable culture densities (**Supplementary Figure 4a,b**). Furthermore, the cultures in these low OD_{600} wells could be recovered by outgrowth (**Supplementary Figure 4c**), with a longer lag time as the main feature that distinguished them from wells that exhibited high OD_{600} values during initial outgrowth (**Supplementary Figure 4d**). To test if the low OD_{600} was connected to a biological process disrupted by the transposon insertions, we looked for a correlation between low OD_{600} and the disrupted genes of the transposon insertion strains. Of the 61 genes that had more than one insertion strain isolated as a pure culture in the ordered library, we found that one strain having a low OD_{600} did not make it more likely that the second also had a low OD_{600}; in fact, there was zero overlap between pairwise comparisons (Fisher's exact test: p=0.6, odds ratio=0). We conclude that low OD_{600} wells in the *B. thetaiotaomicron* ordered library are not due to a fitness defect of the mutant that was sorted. Instead, we hypothesize that the transient exposure to oxygen during sorting temporarily slows growth for a fraction of the sorted wells. The presence of these low OD_{600} wells could be a useful indication that post-sort recovery conditions of a strain can be further optimized.

The quantitative accuracy of Bar-seq improves probabilistic solutions to barcode locations

The pool designs used here and elsewhere[2,3] are optimized for the time and effort required to pool the libraries. If a barcode occurs in exactly one of each of the plate, row, and column pools, then for our pool design it has a definite solution for a single location in the library. If a strain occurs more than once in the library, then its barcode will appear in more than one of each of the different pool types and, with our pool design, there will be multiple possible solutions for its locations in the library. To predict the correct locations of these repeated transposon insertion strains, a previous study[2] developed a probabilistic solution that we implemented here.

This approach uses the relative read counts of a barcode across pools to predict the most likely solution out of all alternatives. Briefly, the ratio of reads of the same barcode in two pool types (plate, row, or column) should depend only on the number of strains added to each pool. For example, the ratio of reads for a barcode in its row and column pools should be 8/12. When choosing between alternative combinations of pools, if the ratio of reads for one combination is closer to this ideal ratio than the alternatives, then it is the more likely solution. In reality, the ratio of reads is a distribution centered approximately around this ideal ratio due to both systematic and random noise introduced during the amplification process. A tighter distribution will be more predictive when choosing alternatives.

Because our pool design (plate, row, column) and the pool design from a previous study using ST-PCR (plate-row, plate-column, row, column)[2] share the row and column pool types, we were able to directly compare Bar-seq to ST-PCR according to their ability to distinguish between alternative solutions for indefinite strain locations. We first examined the distribution of row versus column read ratios for definite solutions in the two libraries. The $\log_2$(row/column read ratio) should be a Voigt distribution centered around -0.585 [2]. We found that the Bar-seq method (this study) resulted in a tighter distribution than ST-PCR (reanalyzed data from previous work[2]) (**Supplementary Figure 1a**), indicating that Bar-seq better preserved read count information contained in the genomic DNA pools.

We next examined the subset of barcodes (this study, n=169) or transposon insertion locations (reanalyzed data from previous work[2], n=1451) that were detected in exactly two row and two column pools. There are two unique solutions that could give rise to this pool inclusion pattern. The log ratio of the probabilities of the more (p_1) and less (p_2) likely solutions is a measure of the ability of read ratio information to predict a correct solution for the mutants. The Bar-seq method outperformed ST-PCR (Wilcoxon rank sum test: W=2x10^5, p=7x10^{-59}) (**Supplementary Figure 1b**).

We conclude that Bar-seq is more robust than ST-PCR and more cost-effective than INSeq and Tn-seq, making it the preferred sequencing method for locating mutants within an ordered library.

Supplementary Tables

Supplementary Table 1: Oligos used for Bar-seq PCR amplification in steps 76-79.

Oligo name	Oligo sequence	PAGE purified	Index name	Index sequence	Description
Barseq_P1	AATGATACGGCG ACCACCGAGATC TACACTCTTTCCC TACACGACGCTC TTCCGATCTNNN NNGTCGACCTGC AGCGTACG	yes			Common primer for "TruSeq-indexed" Bar-seq
Barseq_P2 _IT 001	CAAGCAGAAGAC GGCATACGAGAT CGTGATGTGACT GGAGTTCAGACG TGTGCTCTTCCG ATCTGATGTCCA CGAGGTCTCT	yes	IT001	ATCACG	TruSeq-derived indexed Bar-seq primer
Barseq_P2 _IT 002	CAAGCAGAAGAC GGCATACGAGAT ACATCGGTGACT GGAGTTCAGACG TGTGCTCTTCCG ATCTGATGTCCA CGAGGTCTCT	yes	IT002	CGATGT	TruSeq-derived indexed Bar-seq primer
Barseq_P2 _IT 003	CAAGCAGAAGAC GGCATACGAGAT GCCTAAGTGACT GGAGTTCAGACG TGTGCTCTTCCG ATCTGATGTCCA CGAGGTCTCT	yes	IT003	TTAGGC	TruSeq-derived indexed Bar-seq primer
Barseq_P2 _IT 004	CAAGCAGAAGAC GGCATACGAGAT TGGTCAGTGACT GGAGTTCAGACG TGTGCTCTTCCG ATCTGATGTCCA CGAGGTCTCT	yes	IT004	TGACCA	TruSeq-derived indexed Bar-seq primer
Barseq_P2 _IT 005	CAAGCAGAAGAC GGCATACGAGAT CACTGTGTGACT GGAGTTCAGACG TGTGCTCTTCCG ATCTGATGTCCA CGAGGTCTCT	yes	IT005	ACAGTG	TruSeq-derived indexed Bar-seq primer
Barseq_P2 _IT 006	CAAGCAGAAGAC GGCATACGAGAT ATTGGCGTGACT GGAGTTCAGACG TGTGCTCTTCCG ATCTGATGTCCA CGAGGTCTCT	yes	IT006	GCCAAT	TruSeq-derived indexed Bar-seq primer

Barseq_P2_IT 007	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT007	CAGATC	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 008	CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT008	ACTTGA	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 009	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT009	GATCAG	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 010	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT010	TAGCTT	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 011	CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT011	GGCTAC	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 012	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT012	CTTGTA	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 013	CAAGCAGAAGACGGCATACGAGATTTGACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT013	AGTCAA	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 014	CAAGCAGAAGACGGCATACGAGATGGAACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT014	AGTTCC	TruSeq-derived indexed Bar-seq primer

Barseq_P2_IT 015	CAAGCAGAAGACGGCATACGAGATTGACATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT015	ATGTCA	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 016	CAAGCAGAAGACGGCATACGAGATGGACGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT016	CCGTCC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 017	CAAGCAGAAGACGGCATACGAGATCTCTACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT017	GTAGAG	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 018	CAAGCAGAAGACGGCATACGAGATGCGGACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT018	GTCCGC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 019	CAAGCAGAAGACGGCATACGAGATTTTCACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT019	GTGAAA	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 020	CAAGCAGAAGACGGCATACGAGATGGCCACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT020	GTGGCC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 021	CAAGCAGAAGACGGCATACGAGATCGAAACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT021	GTTTCG	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 022	CAAGCAGAAGACGGCATACGAGATCGTACGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT022	CGTACG	TruSeq-derived indexed Barseq primer

Barseq_P2_IT 023	CAAGCAGAAGACGGCATACGAGATCCACTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT023	GAGTGG	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 024	CAAGCAGAAGACGGCATACGAGATGCTACCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT024	GGTAGC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 025	CAAGCAGAAGACGGCATACGAGATATCAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT025	ACTGAT	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 026	CAAGCAGAAGACGGCATACGAGATGCTCATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT026	ATGAGC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 027	CAAGCAGAAGACGGCATACGAGATAGGAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT027	ATTCCT	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 028	CAAGCAGAAGACGGCATACGAGATCTTTTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT028	CAAAAG	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 029	CAAGCAGAAGACGGCATACGAGATTAGTTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT029	CAACTA	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 030	CAAGCAGAAGACGGCATACGAGATCCGGTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT030	CACCGG	TruSeq-derived indexed Barseq primer

Barseq_P2_IT 031	CAAGCAGAAGACGGCATACGAGATATCGTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT031	CACGAT	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 032	CAAGCAGAAGACGGCATACGAGATTGAGTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT032	CACTCA	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 033	CAAGCAGAAGACGGCATACGAGATCGCCTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT033	CAGGCG	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 034	CAAGCAGAAGACGGCATACGAGATGCCATGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT034	CATGGC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 035	CAAGCAGAAGACGGCATACGAGATAAAATGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT035	CATTTT	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 036	CAAGCAGAAGACGGCATACGAGATTGTTGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT036	CCAACA	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 037	CAAGCAGAAGACGGCATACGAGATATTCCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT037	CGGAAT	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 038	CAAGCAGAAGACGGCATACGAGATAGCTAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT038	CTAGCT	TruSeq-derived indexed Barseq primer

Barseq_P2_IT 039	CAAGCAGAAGACGGCATACGAGATGTATAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT039	CTATAC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 040	CAAGCAGAAGACGGCATACGAGATTCTGAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT040	CTCAGA	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 041	CAAGCAGAAGACGGCATACGAGATGTCGTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT041	GACGAC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 042	CAAGCAGAAGACGGCATACGAGATCGATTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT042	TAATCG	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 043	CAAGCAGAAGACGGCATACGAGATGCTGTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT043	TACAGC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 044	CAAGCAGAAGACGGCATACGAGATATTATAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT044	TATAAT	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 045	CAAGCAGAAGACGGCATACGAGATGAATGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT045	TCATTC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 046	CAAGCAGAAGACGGCATACGAGATTCGGGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT046	TCCCGA	TruSeq-derived indexed Barseq primer

Barseq_P2_IT 047	CAAGCAGAAGACGGCATACGAGATCTTCGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT047	TCGAAG	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 048	CAAGCAGAAGACGGCATACGAGATTGCCGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT048	TCGGCA	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 049	CAAGCAGAAGACGGCATACGAGATATGTTTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT049	AAACAT	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 050	CAAGCAGAAGACGGCATACGAGATTGCTTTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT050	AAAGCA	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 051	CAAGCAGAAGACGGCATACGAGATGCATTTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT051	AAATGC	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 052	CAAGCAGAAGACGGCATACGAGATTTTGTTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT052	AACAAA	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 053	CAAGCAGAAGACGGCATACGAGATCAAGTTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT053	AACTTG	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 054	CAAGCAGAAGACGGCATACGAGATAGTCTTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT054	AAGACT	TruSeq-derived indexed Bar-seq primer

Barseq_P2_IT 055	CAAGCAGAAGACGGCATACGAGATTCGCTTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT055	AAGCGA	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 056	CAAGCAGAAGACGGCATACGAGATGTCCTTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT056	AAGGAC	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 057	CAAGCAGAAGACGGCATACGAGATCCTATTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT057	AATAGG	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 058	CAAGCAGAAGACGGCATACGAGATGTTTGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT058	ACAAAC	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 059	CAAGCAGAAGACGGCATACGAGATAGATGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT059	ACATCT	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 060	CAAGCAGAAGACGGCATACGAGATCTGGGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT060	ACCCAG	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 061	CAAGCAGAAGACGGCATACGAGATGCCGGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT061	ACCGGC	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 062	CAAGCAGAAGACGGCATACGAGATTATCGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT062	ACGATA	TruSeq-derived indexed Bar-seq primer

Barseq_P2_IT 063	CAAGCAGAAGACGGCATACGAGATGAGAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT063	ACTCTC	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 064	CAAGCAGAAGACGGCATACGAGATTCTTCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT064	AGAAGA	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 065	CAAGCAGAAGACGGCATACGAGATCTATCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT065	AGATAG	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 066	CAAGCAGAAGACGGCATACGAGATGATGCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT066	AGCATC	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 067	CAAGCAGAAGACGGCATACGAGATAGCGCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT067	AGCGCT	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 068	CAAGCAGAAGACGGCATACGAGATCGGCCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT068	AGGCCG	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 069	CAAGCAGAAGACGGCATACGAGATCCGTATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT069	ATACGG	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 070	CAAGCAGAAGACGGCATACGAGATTAGGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT070	ATCCTA	TruSeq-derived indexed Bar-seq primer

Barseq_P2_IT 071	CAAGCAGAAGACGGCATACGAGATATAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT071	ATCTAT	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 072	CAAGCAGAAGACGGCATACGAGATGCGTGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT072	CCACGC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 073	CAAGCAGAAGACGGCATACGAGATCATGGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT073	CCCATG	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 074	CAAGCAGAAGACGGCATACGAGATTTGCGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT074	CCGCAA	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 075	CAAGCAGAAGACGGCATACGAGATCTAAGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT075	CCTTAG	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 076	CAAGCAGAAGACGGCATACGAGATTTCTCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT076	CGAGAA	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 077	CAAGCAGAAGACGGCATACGAGATCAGCAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT077	CTGCTG	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 078	CAAGCAGAAGACGGCATACGAGATGGTTTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT078	GAAACC	TruSeq-derived indexed Barseq primer

Barseq_P2_IT 079	CAAGCAGAAGACGGCATACGAGATTTATTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT079	GAATAA	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 080	CAAGCAGAAGACGGCATACGAGATTCCGTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT080	GACGGA	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 081	CAAGCAGAAGACGGCATACGAGATTATATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT081	GATATA	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 082	CAAGCAGAAGACGGCATACGAGATAGCATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT082	GATGCT	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 083	CAAGCAGAAGACGGCATACGAGATCCTTGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT083	GCAAGG	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 084	CAAGCAGAAGACGGCATACGAGATAAGTGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT084	GCACTT	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 085	CAAGCAGAAGACGGCATACGAGATTAAGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT085	GCCTTA	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 086	CAAGCAGAAGACGGCATACGAGATTGGAGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT086	GCTCCA	TruSeq-derived indexed Bar-seq primer

Barseq_P2_IT 087	CAAGCAGAAGACGGCATACGAGATTGTGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT087	GGCACA	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 088	CAAGCAGAAGACGGCATACGAGATCAGGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT088	GGCCTG	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 089	CAAGCAGAAGACGGCATACGAGATGGTAGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT089	TCTACC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 090	CAAGCAGAAGACGGCATACGAGATCATTCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT090	TGAATG	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 091	CAAGCAGAAGACGGCATACGAGATATGGCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT091	TGCCAT	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 092	CAAGCAGAAGACGGCATACGAGATCCAGCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT092	TGCTGG	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 093	CAAGCAGAAGACGGCATACGAGATGCGCCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT093	TGGCGC	TruSeq-derived indexed Barseq primer
Barseq_P2_IT 094	CAAGCAGAAGACGGCATACGAGATTTCGAAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT094	TTCGAA	TruSeq-derived indexed Barseq primer

Barseq_P2_IT 095	CAAGCAGAAGACGGCATACGAGATGGAGAAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT095	TTCTCC	TruSeq-derived indexed Bar-seq primer
Barseq_P2_IT 096	CAAGCAGAAGACGGCATACGAGATAAACCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATGTCCACGAGGTCTCT	yes	IT096	AGGTTT	TruSeq-derived indexed Bar-seq primer

Supplementary Table 2: The custom oligos used to make the RB-TnSeq library in steps 87-90. The oligo MOD2_TruSeq is modified to carry 5'-phosphate, as denoted by '/5'P/'. The oligo Mod2_TS_Univ is modified by an internal phosphorothioate bond, as indicated by the '*'. The MOD2_TruSeq and Mod2_TS_Univ oligos are annealed to create a Y-adapter, as specified in Reagent Setup.

Oligo name	Oligo sequence	PAGE purified	Index name	Index sequence	Protocol step(s)	Description
MOD2_TruSeq	/5'P/GATCGGAAGAGCACACGTCTGAACTCCAGTCA	yes			87	Y-adapter for ligation to genomic DNA fragments (1/2)
Mod2_TS_Univ	ACGCTCTTCCGATC*T	yes			87	Y-adapter for ligation to genomic DNA fragments (2/2)
Nspacer_barseq_universal	ATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNGATGTCCACGAGGTCT	yes			88-90	Tn-specific primer for amplification of Tn-seq library
P7_i6	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	yes	i6	GCCAAT	88-90	Adapter-specific primer for amplification of RB-TnSeq library, indexed

Supplementary Table 3: Oligos used to spot check strain assignments in the ordered library. The transposon insertion targeted for the PCR check is annotated with *position-(orientation)*. Every PCR check used Erm-out as the forward primer and a strain-specific reverse primer. These oligos were used to generate the data presented in Supplementary Figure 5.

Oligo name	Oligo sequence	Description
Erm-out	GCAGGTAATACGACTCACTATAG	outward-oriented primer for erm cassette
BtErm-Check-P06G05	GATAAAATTGCATTGCACATTATCAC	1224-(+) transposon insertion, P6-G5
BtErm-Check-P18G11	GAGACACATCCTGAAATCATG	796028-(-) transposon insertion, P18-G11
BtErm-Check-P20A06	TAATGCATGGGAGCTTCATG	4810775-(+) transposon insertion, P20-A6
BtErm-Check-P22B06	TACGGATAGCGTCGTC	3905852-(-) transposon insertion, P22-B6
BtErm-Check-P23E01	TTGGTTCAGCGGTATGTG	233663-(-) transposon insertion, P23-E1
BtErm-Check-P26A09	CAATACTCTCCATTTTCATCTGAC	4064363-(-) transposon insertion, P26-A9
BtErm-Check-P27E09	CTTAGCTAAATACAGTTACCTGG	4662683-(+) transposon insertion, P27-E9
BtErm-Check-P17B02-A	ATTAAAGGATGCATTACGTATATGTC	3223007-(+) transposon insertion, P17-B2
BtErm-Check-P17B02-B	CGTTCAGCTTATAAGCTTCAG	5788004-(-) transposon insertion, P17-B2
BtErm-Check-P18C02-A	CCCGAAACGTGTTTTCAATG	1864693-(+) transposon insertion, P18-C2
BtErm-Check-P18C02-B	GTTCGGGACAAAGAGATTC	2999897-(-) transposon insertion, P18-C2
BtErm-Check-P26D09-A	CAGCCGTATGATTGTCTATG	3737285-(+) transposon insertion, P26-D9
BtErm-Check-P26D09-B	ATCAGTAGGTTCACCTGTAG	5488616-(+) transposon insertion, P26-D9
BtErm-Check-P31D11-A	GATACAGCAAAGCGAATACG	1408781-(+) transposon insertion, P31-D11
BtErm-Check-P31D11-B	TGCTACGACGATCTGTTC	1994095-(-) transposon insertion, P31-D11
BtErm-Check-P32F07-A	AGGTGGTGTATTGCCTG	18210-(-) transposon insertion, P32-F7
BtErm-Check-P32F07-B	CTTCATGCAAAAGCGATTG	207619-(-) transposon insertion, P32-F7
BtErm-Check-P38E06-A	GTACGACAGAGCAAAAATTAGC	3926701-(+) transposon insertion, P38-E6
BtErm-Check-P38E06-B	CAAATAGCCTGTTTTACCCC	4207683-(+) transposon insertion, P38-E6
BtErm-Check-P39E08-A	TAATGTTGTCCTCCTTGCTG	2248214-(+) transposon insertion, P39-E8
BtErm-Check-P39E08-B	GCGTGGAATATTACCACATATTTC	3630139-(+) transposon insertion, P39-E8
BtErm-Check-P40D11-A	AACAGCGAAGGAGACAAG	2500909-(-) transposon insertion, P40-D11
BtErm-Check-P40D11-B	CGACTTCGTTCTAGAAACC	3061114-(+) transposon insertion, P40-D11

Supplementary References

1 Holt J.G., K. N. R., Sneath P.H.A., Staley J.T., Williams S.T. *Bergey's Manual of Determinative Bacteriology*. 9th ed. edn, (Williams and Willkins, 1994).

2 Baym, M., Shaket, L., Anzai, I. A., Adesina, O. & Barstow, B. Rapid construction of a whole-genome transposon insertion collection for Shewanella oneidensis by Knockout Sudoku. *Nature communications* **7**, 13270 (2016).

3 Vandewalle, K., Festjens, N., Plets, E., Vuylsteke, M., Saeys, Y. & Callewaert, N. Characterization of genome-wide ordered sequence-tagged Mycobacterium mutant libraries by Cartesian Pooling-Coordinate Sequencing. *Nature communications* **6**, 7106 (2015).

4 Kumar, S., Stecher, G., Suleski, M. & Hedges, S. B. TimeTree: a resource for timelines, timetrees, and divergence times. *Molecular Biology and Evolution* **34**, 1812-1819 (2017).

5 Edelstein, A. D., Tsuchida, M. A., Amodaj, N., Pinkard, H., Vale, R. D. & Stuurman, N. Advanced methods of microscope control using μManager software. *Journal of biological methods* **1** (2014).