

# UC Irvine

## UC Irvine Previously Published Works

### Title

Analysis of Pools of Targeted Salmonella Deletion Mutants Identifies Novel Genes Affecting Fitness during Competitive Infection in Mice

### Permalink

<https://escholarship.org/uc/item/3p96h29p>

### Journal

PLOS Pathogens, 5(7)

### ISSN

1553-7366

### Authors

Santiviago, Carlos A  
Reynolds, M Megan  
Porwollik, Steffen  
[et al.](#)

### Publication Date

2009-07-01

### DOI

10.1371/journal.ppat.1000477

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

# Analysis of Pools of Targeted *Salmonella* Deletion Mutants Identifies Novel Genes Affecting Fitness during Competitive Infection in Mice

Carlos A. Santiviago<sup>1</sup><sup>¶</sup>, M. Megan Reynolds<sup>2</sup><sup>¶</sup>, Steffen Porwollik<sup>1</sup>, Sang-Ho Choi<sup>1</sup><sup>¶</sup>, Fred Long<sup>1</sup>, Helene L. Andrews-Polymeris<sup>2\*</sup>, Michael McClelland<sup>1\*</sup>

**1** Sidney Kimmel Cancer Center, San Diego, California, United States of America, **2** Department of Microbial and Molecular Pathogenesis, College of Medicine, Texas A&M University System Health Science Center, College Station, Texas, United States of America

## Abstract

Pools of mutants of minimal complexity but maximal coverage of genes of interest facilitate screening for genes under selection in a particular environment. We constructed individual deletion mutants in 1,023 *Salmonella enterica* serovar Typhimurium genes, including almost all genes found in *Salmonella* but not in related genera. All mutations were confirmed simultaneously using a novel amplification strategy to produce labeled RNA from a T7 RNA polymerase promoter, introduced during the construction of each mutant, followed by hybridization of this labeled RNA to a Typhimurium genome tiling array. To demonstrate the ability to identify fitness phenotypes using our pool of mutants, the pool was subjected to selection by intraperitoneal injection into BALB/c mice and subsequent recovery from spleens. Changes in the representation of each mutant were monitored using T7 transcripts hybridized to a novel inexpensive minimal microarray. Among the top 120 statistically significant spleen colonization phenotypes, more than 40 were mutations in genes with no previously known role in this model. Fifteen phenotypes were tested using individual mutants in competitive assays of intraperitoneal infection in mice and eleven were confirmed, including the first two examples of attenuation for sRNA mutants in *Salmonella*. We refer to the method as Array-based analysis of cistrons under selection (ABACUS).

**Citation:** Santiviago CA, Reynolds MM, Porwollik S, Choi S-H, Long F, et al. (2009) Analysis of Pools of Targeted *Salmonella* Deletion Mutants Identifies Novel Genes Affecting Fitness during Competitive Infection in Mice. *PLoS Pathog* 5(7): e1000477. doi:10.1371/journal.ppat.1000477

**Editor:** Brad T. Cookson, University of Washington, United States of America

**Received:** November 3, 2008; **Accepted:** May 15, 2009; **Published:** July 3, 2009

**Copyright:** © 2009 Santiviago et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported in part by NIH grants R21 AI057733, R01 AI034829, R01 AI073971, and R01 AI052237, and the generous support of Mr. Sidney Kimmel. CAS is supported by grant ADI-08/2006 from "Programa Bicentenario de Ciencia y Tecnológica" (CONICYT, Chile) and The World Bank. The funders played no role in the design or conduct of the study, in the collection, analysis, or interpretation of the data, or in the preparation, review, or approval of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: handrews@medicine.tamhsc.edu (corresponding author for animal studies, HLA-P); mclelland.michael@gmail.com (MM)

¶a Current address: Programa de Microbiología y Micología, Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile, Santiago, Chile

¶b Current address: Division of Infectious Diseases, Asian Medical Center, University of Ulsan College of Medicine, Center for Antimicrobial Resistance and Microbial Genetics, University of Ulsan, Seoul, Republic of Korea

¶ These authors contributed equally to this work.

## Introduction

Genetic screening remains one of the most efficient methods to identify genes associated with a phenotype of interest in bacteria. Array-based methods for these screens originated with the transposon-based "signature tagged mutagenesis" (STM) strategy that used unique signature sequences in each transposon to evaluate the relative abundance of individual mutants in pools after selection [1]. STM was later improved by modifying the mutagenizing transposon to include a T7 RNA polymerase promoter ( $P_{T7}$ ) that is used to generate a unique transcript for each mutant from the genomic sequence adjacent to the mutation. This modification makes exogenous unique sequence tags unnecessary. Relative abundance of the input and output  $P_{T7}$  transcripts is monitored using an ORF microarray [2,3,4,5].

Transposon mutagenesis suffers from several drawbacks. First, tens of thousands of random transposon insertion mutants are necessary to ensure that mutations occur in most small genes. There are over 1,100 annotated open reading frames in *Salmonella* that are less than 500 bases in length. In mathematical simulations of pools

of 40,000 random transposon integrations, over 200 of these short genes fail to be disrupted by any transposon, on average (data not shown). Second, this need for a high complexity is a critical limitation of random mutants for genetic screens in environments, including live animals, where the bacterial population may fall to low levels during infection. These 'bottlenecks' may occur at various points during infection – survival of the acidic environment in the stomach, invasion of Peyer's patches and survival in the bloodstream represent some of the processes where the founder population may be very small. Such 'bottlenecks' cause undesirable random loss of mutants and complicate forward genetic screening in such environments. In addition, the polar nature of transposon insertions makes mapping of a phenotype more difficult.

We used the lambda-red recombination method that includes features to minimize polarity [6] to construct targeted deletion mutants in *Salmonella*. We added a  $P_{T7}$  to the cassette inserted during mutagenesis, positioned to produce a gene-specific transcript from the genomic sequence adjacent to the insertion. Thus, our targeted deletion mutants can be pooled for genetic screens. Far fewer specific mutants are needed to ensure

## Author Summary

One strategy to define bacterial genes with a role during infection involves the screening of a pool of random transposon insertion mutants, where each mutant is identifiable by a specific motif or unique transcript. Changes in the survival of each mutant indicate a role of the mutated region during infection. To ensure coverage of most genes of interest, a large number of random transposon mutants would be needed. However, when a pool of bacterial mutants spreads into different sites in an animal host, some mutants are lost at random if the founder population (i.e. the bacteria initially reaching this site, before expansion) is smaller than the number of mutants introduced into the animal. This random loss severely obscures mutants that are truly at a disadvantage. In order to minimize the number of mutants to be screened and thus minimize random loss of mutants from an infecting pool, we generated targeted specific deletions in each gene of interest. Furthermore, we inserted a promoter in each mutant that allows simultaneous monitoring of a pool of these mutants on a novel inexpensive microarray. To demonstrate the utility of our technique, a pool of over 1,000 gene-targeted *Salmonella* mutants was injected into mice and recovered from spleens. We identified mutants that were less fit than wild-type *Salmonella* in this model. Eleven mutants in genes that were not previously known to affect *Salmonella* fitness in intraperitoneal infection were confirmed using the individual mutants in competition with wild-type bacteria. These new phenotypes include the first two examples of attenuation and one of hypervirulence in *Salmonella* due to mutations in small stable RNAs, a class of regulators that bind to other RNAs and proteins.

representation of every gene of interest than is needed when using random transposon mutagenesis. We introduce a novel and inexpensive array designed for monitoring changes in specific deletion mutant representation in the pool after selection. Finally, we demonstrate the utility of these techniques by identifying novel candidate genes encoding proteins and sRNAs deletions of which affect fitness in competitive infection with wild-type *Salmonella*, and we validate eleven of these fitness phenotypes.

## Results

### Genome sequencing

The virulent isolate *Salmonella enterica* serovar Typhimurium ATCC14028 is extensively studied both *in vitro* and *in vivo*. To design primers for generation of our deletion collection, we produced a near complete draft sequence of this genome using the 454 shotgun approach (GenBank accession in process). The ATCC14028 sequence was compared to the completed genome of the 1,000-fold less pathogenic laboratory strain Typhimurium LT2 [7]. As expected, over 95% of the two genomes were orthologous, and the orthologous regions had less than 1% divergence. The ATCC14028 and LT2 genomes differ only by a few hundred single base mutations (including a mutation in *rhoS* that is partly responsible for the attenuation of LT2 [8]), the absence of the two Fels phage in ATCC14028, and other insertions and deletions encompassing less than 40 kb (Sandy Clifton et al., unpublished data).

### Generation of specific gene deletions in ATCC14028

We targeted 1,052 genes for deletion (**Table S1**), primarily genes in *Salmonella* that are not found in *E. coli* [7]. Such genes are

usually in very A+T rich regions [9], and include nearly all of the ~200 genes previously associated with *Salmonella* virulence, including the Type III secretion systems (TTSS) and their known effectors. Targeted deletions were also generated in nearly all of the 100 genes in fimbrial and surface antigen regulons. Finally, we deleted a subset of genes shared by *Salmonella* and *E. coli*, including 44 known and candidate sRNAs, and genes that have known motility, regulatory and pathogenesis functions.

The original vectors for the lambda-red swap strategy, pKD3 (Cm<sup>R</sup>) and pKD4 (Kan<sup>R</sup>) [6], were redesigned to include a T7 RNA polymerase promoter positioned to generate a unique transcript from the *Salmonella* genome directly downstream of each mutant. The construct includes an ATG and ribosome binding site (RBS) to preserve any translation coupling. An outline of our variation on the lambda-red swap strategy is shown in **Figure 1**. The sequences of the redesigned vectors, pCLF3 and pCLF4, are available in GenBank (Accession numbers EU629213 and EU629214, respectively).

The oligos used for mutagenesis were 65mers that contained the same 3' sequence of 20 bases for vector PCR that were used previously [6]. However, they included a 45-base rather than the conventional 35-base gene-specific portion, to allow a secondary function for the oligonucleotides as probes in a microarray (described later). Homology of the oligonucleotides with the genome was positioned so that each swap eliminated the entire gene except for the coding regions of 30 base pairs at the 5' and 3' ends of each gene. These two ends of each gene were preserved to minimize unintended effects on adjacent genes caused by removal of the gene sequence and insertion of the mutagenic cassette. If FLP recombination is used to remove the antibiotic cassette after mutant construction then an open reading frame is generated including ten codons from the 5' end of the original gene, 39 amino acids from the lambda-red mutagenesis referred to as a "scar", and nine amino acids and the stop codon from the original gene. In our constructs, the T7 promoter remains in place in the scar after FLP recombination.

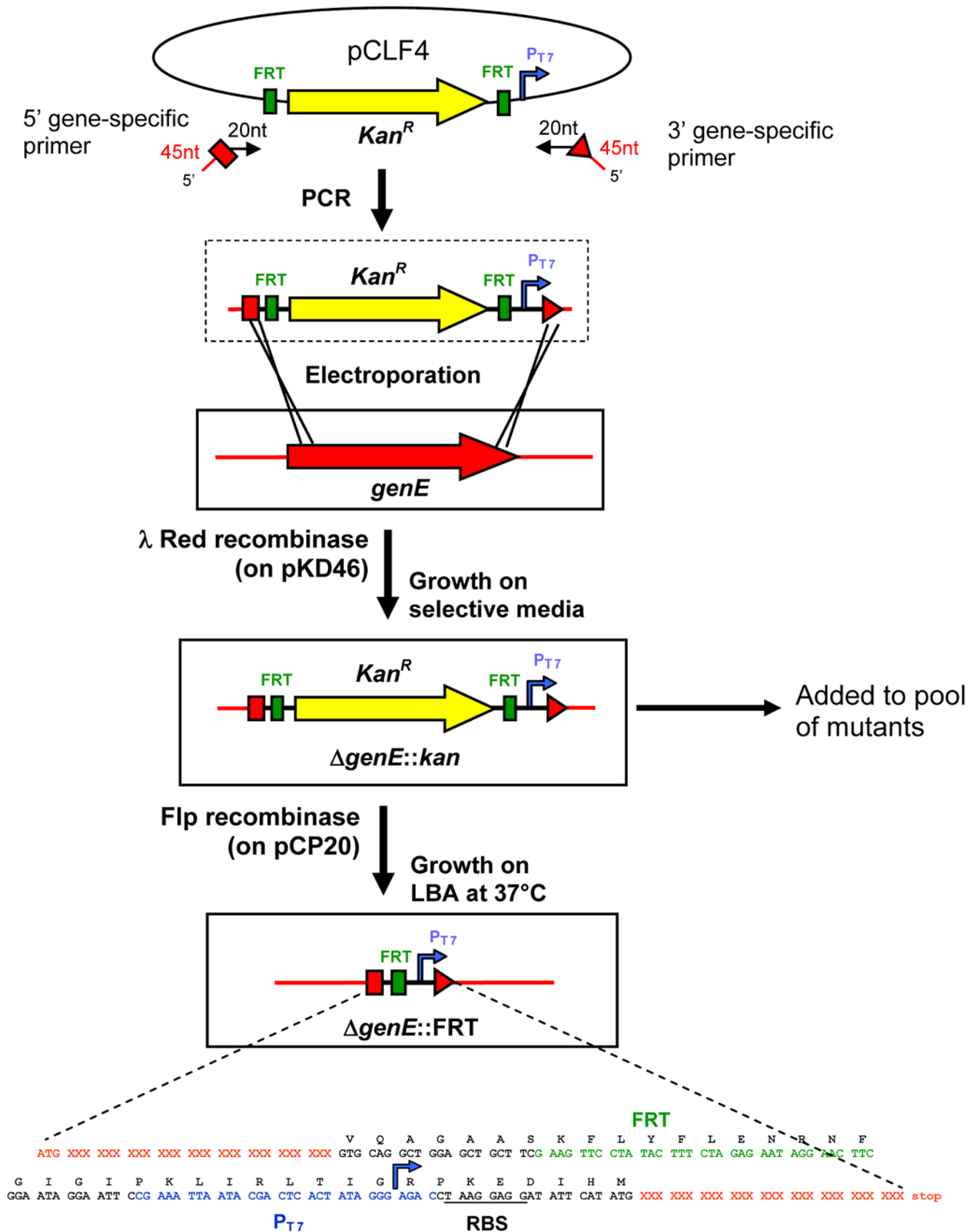
We also increased the level of throughput in the generation of our mutant collections. Lambda-red swap recombination in ATCC14028 was performed with a mixture of two PCR products, one originating from pCLF4 containing a kanamycin resistance cassette (Kan<sup>R</sup>, sense orientation) and one originating from pCLF3 containing a chloramphenicol resistance cassette (Cm<sup>R</sup>, antisense orientation). Each transformation was plated on LB-Kan and LB-Cm, and two transformants from each plate were colony purified and stored (a total of four transformants). These two collections marked with different antibiotic resistance cassettes facilitate the construction of double mutants by transduction.

The *Salmonella* genes targeted for deletion were spot-checked by PCR. Of 1,052 mutants attempted, 1,040 produced Kan<sup>R</sup> or Cm<sup>R</sup> clones, or both of which 1023 were confirmed, below (**Table S1**). To confirm that mutations occurred at the targeted location, two clones obtained for the first 304 Kan<sup>R</sup> mutants and for 231 Cm<sup>R</sup> mutants were checked by PCR using primers to the flanking genomic regions. Only one mutation was incorrect.

The accuracy of FLP recombination to remove the antibiotic resistance marker was verified for twelve transformants (as detailed in the Methods and Materials section). The mutant including flanking regions was amplified by PCR, and both strands of the amplification product were sequenced. All twelve swaps examined were precise, and each recombination event resulted in the intended truncated open reading frame.

### Array-based verification of P<sub>T7</sub> location and activity from a pool of mutants

We used a NimbleGen tiling array to verify the correct insertion and activity of the P<sub>T7</sub> in all of our mutants simultaneously. 1,031

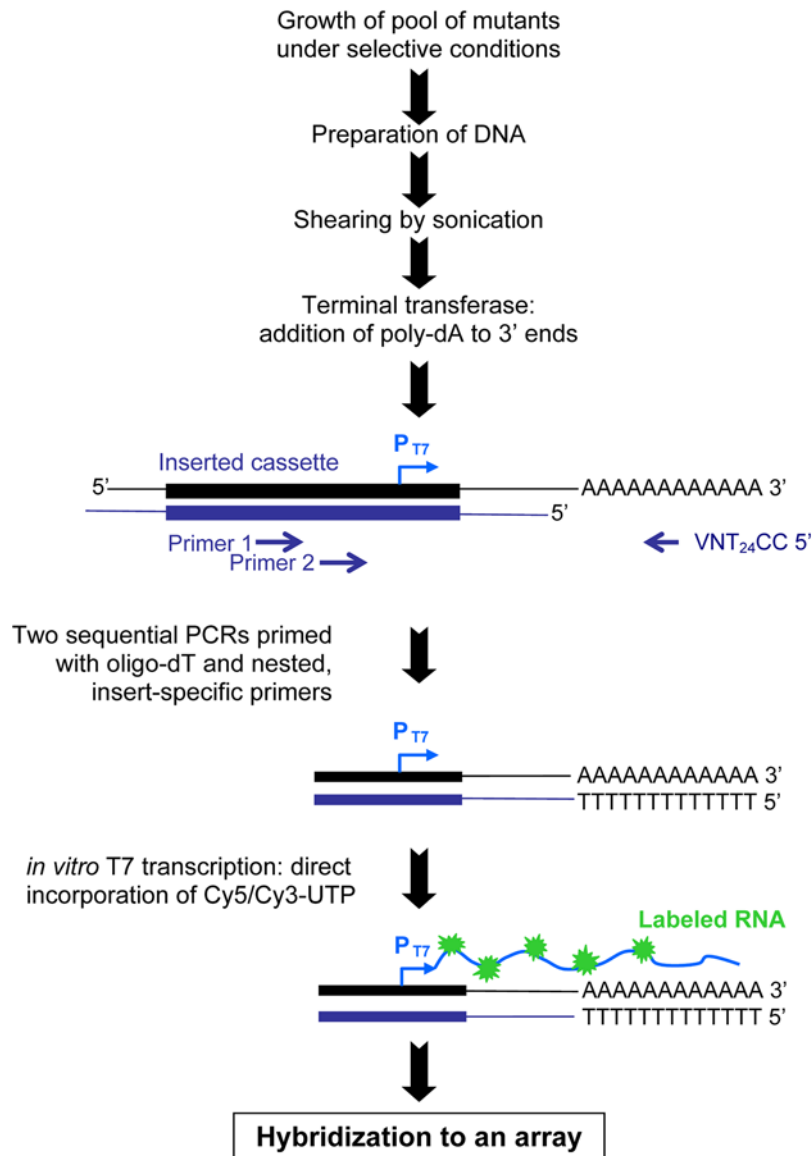


Scar includes 10 codons on the C- and N-terminus of *genE*

**Figure 1. Generation of specific deletions in *S. enterica* serovar Typhimurium ATCC14028.** Our procedure is identical to the Red-swap described in reference [6] with the exception that we re-engineered the original insert to include an in-frame T7 RNA polymerase promoter ( $P_{T7}$ ), and the sequences used for recombination are longer. A gene, identified in the schematic as *genE*, is targeted for deletion. Two 65mer primers (red) are used to amplify the region containing the antibiotic resistance cassette, the  $P_{T7}$ , and the FRT sites (among other elements) from the plasmid pCLF4. The resulting PCR product has 45 base sequences at each end that are homologous to sequences near the 5' and 3' ends of the *targeted ORF*. Transformation of these PCR products into ATCC14028 expressing lambda Red recombinase *in trans* leads to a recombination event resulting in the swap-in of the PCR product, and swap-out of the targeted gene. A ribosomal binding site (RBS) and a downstream ATG start codon near the 3' end of the inserted sequence ensures that a 12 amino acid peptide is made from any RNA that is transcribed in this strand to reduce polar effects. Targeted mutants in many genes were pooled and used for Array-based analysis of cistrons under selection (ABACUS). Finally, antibiotic resistance markers in targeted genes can be removed using the FLP recombinase resulting in a gene encoding a mini-protein of the first ten amino acids of GenE, 39 amino acids from the inserted DNA (called a "scar"), and the last nine amino acids of GenE, and retaining the  $P_{T7}$ . doi:10.1371/journal.ppat.1000477.g001

Kan<sup>R</sup> mutants were pooled by growing each mutant separately to stationary phase in LB, mixing in equal volume, and storing as glycerol stocks at -80°C. The region 3' to each mutation in the pool of 1,031 Kan<sup>R</sup> mutants was labeled by a novel method designed to generate uniform signals from every  $P_{T7}$  (Materials and Methods and **Figure 2**). For the necessary DNA fragmentation prior to amplification and labeling, we replaced the

restriction digestion step used in previous protocols with genome shearing and polyA tailing. To identify active  $P_{T7}$  in the mutants, labeled RNA obtained by T7 *in vitro* transcription from the pool was hybridized to a custom NimbleGen tiling array of 50mer oligonucleotides covering the ATCC14028 genome in overlapping 24 base steps on both strands. Oligos directly adjacent to each functional insert and in the opposite strand from the transcript



**Figure 2. Labeling technique for detection of mutants in a pool by arrays.** doi:10.1371/journal.ppat.1000477.g002

hybridized intensely, with a rapid decrease in signal over a 200 base region (**Figure S1**). Using this method we simultaneously identified 933 inserts from Kan<sup>R</sup> mutants with functional P<sub>T7</sub> sites in the correct location. 48 additional Kan<sup>R</sup> mutants are likely correct but could not be formally confirmed, due to an overlapping transcript from a nearby mutation in another mutant in the pool. 50 Kan<sup>R</sup> mutants (4.8% of all pooled mutants) did not display active transcription at the correct location perhaps due to a mutation in the P<sub>T7</sub>. Four inserts were at an incorrect location.

We repeated these experiments with a pool of 972 Cm<sup>R</sup> mutants and at least 892 of these have a functional P<sub>T7</sub>. 78 Cm<sup>R</sup> mutants were in genes where a Kan<sup>R</sup> clone was not confirmed. Overall, correct Kan<sup>R</sup> or Cm<sup>R</sup> mutants, or both, with functional P<sub>T7</sub> were confirmed in 1,023 genes out of 1052 targeted for deletion. We attempted transduction to construct double mutants of Cm<sup>R</sup> clones in combination with different Kan<sup>R</sup> clones. All ten double mutants attempted were successful (data not shown). In addition, the Cam<sup>R</sup> mutants will also be useful in the future as a pool.

Failure to generate the correct targeted mutation could occur if the mutation was lethal. We did not attempt to delete any genes that are orthologous to genes previously reported to be essential in *E. coli* [10]. However, we did target 38 genes that were previously reported as essential in *Salmonella* [11] and succeeded in obtaining mutants in all but two; *STM1008*, a phage gene that has a close paralogue in the genome, and seems unlikely to be essential in LB, and *STM2087*, encoding an enzyme required for the synthesis of LPS side chains. The reason for the 29 failures to construct a mutant is as yet undetermined.

### Generation and testing of an ‘in house’ microarray for screening

Pools of our mutants can be studied using competitive assays and the representation of each mutant can be determined using microarrays. For such detection, we manufactured an economical “in house” array based on 1,241 65mers derived from the 3′ ends of genes, including all those previously utilized for mutant construction. As noted earlier, these 65mers contained a 45-base gene-specific portion, and oligonucleotides of 45 bases in length produce a more robust RNA hybridization signal [12] than the 35mers that are normally used for lambda-red recombination [6]. As negative controls we included in the array 308 65mer oligos with homology to the 3′ end of mutants that had not yet been generated and were therefore absent from the pool, and 955 65mers designed on the 5′ end of genes (previously used for mutant construction).

One aliquot of the pool of 1,031 Kan<sup>R</sup> mutants was used to test the “in house” array. Labeled RNA probe was prepared from the pool, and hybridized to the array as described (Materials and Methods). The 5′ ends of these RNA molecules consist of conserved 27 bases common to every mutant. A specific 27 base oligonucleotide complementary to this region was included in the hybridization to block cross-hybridization of this portion of the T7 transcripts to a complementary 20 base sequence present in each of the 3′ gene probes on the array (**Figure S2**). Specific signal at least three-fold over the mean signal of the negative controls was detected for 97% (905) of the 933 mutants known to be in the pool. **Table S2** illustrates the specificity and sensitivity of our approach. Only seven (0.7%) of the 5′ oligos included as negative controls in the array hybridized, possibly due to a transcript from a nearby mutant in the complementary strand.

### Genes that affect fitness of *Salmonella* during passage through BALB/c mice

BALB/c mice develop fatal systemic infection from *S. enterica* serotype Typhimurium, as can also occur in humans infected with

*S. enterica* serotype Typhi. We used intraperitoneal infection in BALB/c mice, a well-studied model [1,13,14,15,16], to demonstrate that our mutant screening method could correctly identify both previously observed and novel gene requirements during infection. A group of six BALB/c mice were infected intraperitoneally (IP) and bacteria were recovered from the spleen after euthanasia at 48 hours post-infection. Labeled RNA probe was prepared using aliquots of DNA from the input pool and output pools from each animal as outlined in **Figure 2**, and the labeling protocol was repeated with dyes reversed. Hybridizations were performed using twelve slides, each slide containing the array printed in triplicate. The data were processed using WebArray [17], as described in Materials & Methods, and are presented in **Table S1**, and visualized by genome position in **Figure 3**.

At a threshold of a two-fold change and P<0.0005, 120 mutants showed a change in fitness in IP infection. Among the most unfit were mutants in the TTSS encoded by SPI-2, and associated effector genes, and genes for cell wall biosynthetic enzymes, all known to be important during systemic infection [1,13]. None of these mutants have a change in fitness during growth of the pool in LB for four serial passages to stationary phase with 100-fold dilutions (data not presented). Fifteen mutants without previously known fitness phenotypes during systemic infection in BALB/c mice were selected from the 120 statistically top-ranked candidate mutants. These included mutants in all five candidate sRNA, a supernumerary tRNA (LeuX), and nine randomly picked protein coding genes. These 15 mutants were studied one at a time in competitive assays with wild-type ATCC14028, after P22 transduction of each mutation to a clean genetic background. The phenotypes observed in the pool of mutants were confirmed for eleven mutants, as listed in **Figure 4** and **Table S3**. Four mutants failed to reiterate the phenotype observed in the mutant pool (*sraA*, *rybB*, *STM4529*, and *STM0857*) but this result does not rule out confirmation of a phenotype in a study involving more animals.

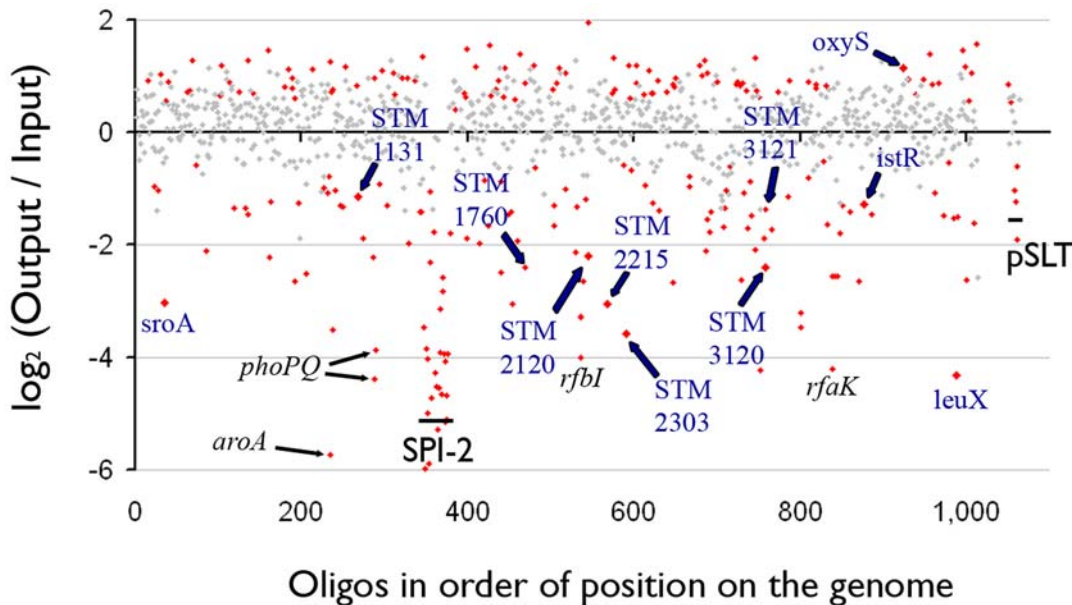
All eleven mutants with confirmed competitive index phenotypes had their Kan<sup>R</sup> cassette removed by FLP recombination to reduce the chance of polar effects on nearby genes. Eight of these (*STM1131*, *STM2303*, *STM3120*, *STM3121*, *sroA*, *istR*, *leuX*, and *oxyS*) were retested in competitive infections in BALB/c mice (**Figure 4A**). The phenotype was confirmed in all eight of these unmarked mutants that should have minimal or no polar effects.

In order to definitively link the deleted genes to the reduced fitness phenotype, two of our unmarked (“FLP-ed”) mutants ( $\Delta$ *STM2303::FRT* and  $\Delta$ *STM3121::FRT*) were re-tested after complementation *in trans*. Intact copies of these genes were cloned into a stable low copy plasmid vector, pWSK29 [18], and the complementing plasmids were transformed into the corresponding unmarked deletion strains. The competitive fitness defect of both mutants was reversed when these strains were complemented with an intact copy of the corresponding gene *in trans* (**Figure 4B**). The presence of the vector alone did not improve the fitness of these two mutants during infection.

## Discussion

Targeted deletion by lambda-red recombination was previously used to construct a library of specific gene deletion mutants for *E. coli* [10]. Here, we use this approach for *Salmonella*, and modify it by adding a T7 RNA polymerase promoter to the insert for the generation of unique transcripts from each mutant. These transcripts are used to identify and measure the relative abundance of each mutant from a pool containing all of our mutants. Furthermore, we introduce several technical innovations to improve the throughput of library construction, and the labeling





**Figure 3. Loss of defined mutants from the input pool after intraperitoneal delivery to six BALB/c mice and recovery from spleen at 48 h post-infection.** Twelve array hybridizations were performed using two each for six mice. WebArray analysis results using quantile normalization for all oligos representing candidate deletions in the pool are shown. Data with  $P < 0.0005$  are depicted in red, those with  $P > 0.0005$  in grey. The X-axis plots the data for each of the mutants in the order in which those mutants occur in the *Salmonella* genome. Mutants confirmed individually in a competitive assay with wild-type are labeled in blue. Examples of important genetic elements are marked in black. doi:10.1371/journal.ppat.1000477.g003

and detection of transcripts from each mutant on an inexpensive, customized microarray. We use a well-studied model, systemic infection of BALB/c mice, to validate our assay and identify novel candidate genes that affect the fitness of *Salmonella* in that infection model.

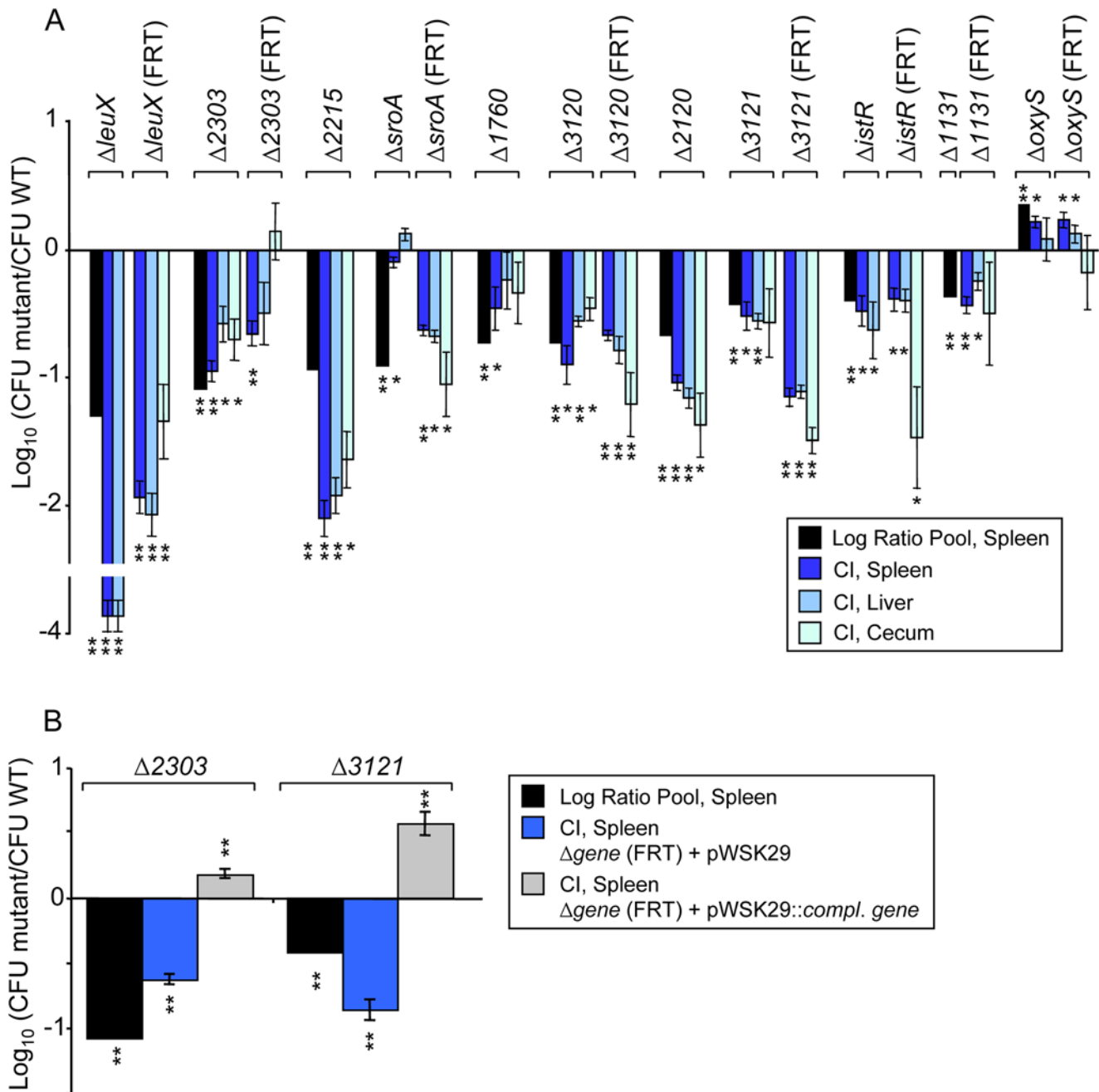
Our targeted deletion strategy has several important advantages over screening of pools of mutants made by random transposon mutagenesis. Screening targeted deletion pools reduces the complexity of the bacterial pool necessary to cover all genes of interest by at least 10-fold over random transposon mutant pools; mathematically, 40,000 random transposon mutants cover about 90% of all non-essential *Salmonella* genes, whereas 4,000 specific mutants in each non-essential gene provides 100% coverage of the targeted genes. This less complex pool is highly advantageous for forward genetic strategies in circumstances where population sizes drop during the selection process, such as in animal models [4,19,20,21,22]. For *Salmonella*, this “bottleneck” is particularly severe in the transition from the intestinal tract to the systemic circulation [23,24] and perhaps in different parts of the gastrointestinal tract where niche conditions can vary considerably. These bottlenecks cause random loss of mutants from the pool if the population falls to numbers close to the complexity of the interrogated pool. For example, if a population falls transiently to 40,000, then a pool of 4,000 should not lose any mutants ( $P < 0.01$  for loss of one mutant). However, 40,000 mutants will lose about 37% of its members at random, because the probability of not picking a mutant that has a frequency of  $1/40,000$  in 40,000 attempts is  $(1 - 1/40,000)^{40,000} = 0.37$ . Furthermore, a pool of lower complexity may be used at a lower infectious dose, allowing a more physiologically relevant infection. The smallest possible pool that includes each mutant of interest is therefore highly desirable.

In any particular experimental system the size of the inoculum needed to avoid bottlenecks will have to be determined empirically. It is possible to determine the optimum size of the inoculum and ascertain whether or not there is a bottleneck by

including a “neutral” mutant carrying a different antibiotic marker at the same CFU as the average mutant CFU in the pool in each experiment. If the representation of this marked strain varies dramatically (either highly over- or under-represented) in the recovered samples in this “fluctuation test”, then the inoculum is of insufficient size to prevent random loss from a pool of that complexity (McClelland et al., unpublished data).

Random transposons can generate pronounced downstream effects that contribute to selection. Our targeted in-frame deletions are engineered to minimize or eliminate these effects. Desirable existing features of the cassette inserted by lambda-red recombination [6], such as an internal ATG start codon at the end of the inserts to allow downstream translational coupling, and the inclusion of FRT sites for the removal of the antibiotic cassette by FLP recombination, remain intact. In our construct, FLP recombination results in the production of an in-frame mini-protein that uses the original start and stop codons of the deleted protein, while retaining the  $P_{T7}$ . The first and last 30 bases of each targeted gene were maintained to reduce the chance that targeted deletions disrupt adjacent genes, overlapping genes, and *cis*-acting sequences thus further minimizing the risk of polar effects. Subsequent analyses should consider that in rare cases the remaining mini-protein may retain some function. Finally, confirmation of mutants identified by genetic screening requires the generation of targeted mutations. A targeted deletion library allows independently generated mutants in the gene of interest to simply be picked from the ordered collection of mutants for further analysis.

Earlier labeling strategies for T7 transcripts from transposon libraries use restriction digestion to fragment the genomic DNA, which has the effect of producing shorter labeled transcripts that are easily mapped near the  $P_{T7}$  [2,3,4,5]. This approach results in non-uniform labeling of transcripts, because those transposons that are not in the ideal proximity to a restriction site are poorly sampled. Our labeling strategy uses randomly sheared ends,



**Figure 4. Competitive Index experiments with individual mutants versus wild-type after intraperitoneal delivery and recovery from the spleen, liver, and cecum.** 10<sup>6</sup> cfu of 933 mutants (pool) or a 1:1 mixture of individual mutants versus wild-type ATCC14028 *ΔphoN::kan* were injected into 5–6 mice and recovered from spleen one or two days post infection. Statistical significance was determined using a Student’s 2-tailed *t* test, and asterisks indicate that normalized output ratios were significantly statistically different from the equivalent ratio in the inoculum. \*\* P value of <0.001, \* P value of <0.05. Strains marked with (FRT) have undergone FLP-mediated recombination to reduce or eliminate polar effects. **A.** Confirmation of mutants, identified by ABACUS, in competitive infections in BALB/c mice. ABACUS prediction for selected mutants (black bars) and confirmation data from competitive infections in BALB/c mice (dark blue bars are from spleen, lighter blue are from liver and cecum) of infected animals. **B.** Complementation of *ΔSTM2303::FRT* and *ΔSTM3121::FRT* mutants *in trans*. The predictions from ABACUS are indicated as black bars. For complementation experiments animals were infected with either a 1:1 mixture of the mutant bearing pWSK29 vs. ATCC14028 *ΔphoN::kan* (negative control, blue bars) or the mutant bearing the corresponding complementing gene on pWSK29 vs. the wild-type ATCC14028 *Δ phoN::kan* (complemented, grey bars). doi:10.1371/journal.ppat.1000477.g004

resulting in more uniform and consistent labeling efficiency for every mutated gene.

The presence of a P<sub>T7</sub> promoter in each targeted mutant allows the presence and the level of each individual mutant to be tracked

when mutants are pooled, using the mutant-specific transcript produced from this promoter. This innovation eliminates the need to introduce a different ‘signature’ sequence in each of the mutants in the collection. The P<sub>T7</sub> facilitates efficient RNA synthesis only



from those DNA fragments that contain P<sub>T7</sub>. Any artifacts that might be generated during the tailing and PCR steps are not transcribed. Our P<sub>T7</sub> procedure generates primarily single stranded labeled RNA, minimizing competition for hybridization to the oligonucleotide array, which might be one reason that the performance of this protocol was consistently superior to direct labeling of PCR products in detection sensitivity and in the standard deviation of measurements, as measured by microarray (data not shown).

The presence of the T7 RNA polymerase promoter in each mutant construct allowed us to check all candidate mutants simultaneously for the location of a functional P<sub>T7</sub>. Using a NimbleGen tiling array of 387,000 oligonucleotides, we confirmed 933 of 1,031 mutants in the Kan<sup>R</sup> pool as present in exactly the correct place, and functional, simultaneously. Similarly, 892 of 972 Cm<sup>R</sup> mutants were confirmed using this technique. Thus, we bypassed much more expensive and laborious traditional PCR confirmation. We also confirm the activity of the P<sub>T7</sub>, which is not possible using PCR.

For routine hybridizations of the pool, we constructed an inexpensive oligonucleotide microarray using the same oligonucleotides that we used to construct the mutants, containing 45 bases of homology to the *Salmonella* genome. This array could also be used to confirm the presence of nearby active P<sub>T7</sub> in a pool. However, unlike the NimbleGen array, this minimal array does not prove the location of the inserted cassettes, or find extraneous insertions that occur elsewhere in the genome. Fortunately, the more expensive NimbleGen arrays need only be used once to determine which mutations are in the correct location and have an active P<sub>T7</sub>. After the library is characterized, the less expensive in-house arrays can be used for further experiments without ambiguity. We expect that multiplexed high-throughput sequencing of experimental samples will replace arrays as the method for assessing changes in the T7 transcripts as soon as prices become competitive.

In order to test our targeted deletion collection and novel protocols, we used systemic infection of BALB/c mice, a model system that has been studied extensively using transposon libraries [1,13,14,15,16]. All the mutants in our collection that were reported to have attenuated growth in these previous transposon assays were confirmed in our assay, including those in SPI-2 and lipopolysaccharide biosynthesis genes. In addition, among the 120 mutants showing the strongest fitness phenotype we identified more than 40 candidates with no previously known phenotype in this model, although many of these new candidate phenotypes are generally more subtle than those seen for SPI-2 (**Table S1, Figure 3**).

We retested fifteen of the novel mutants with at least a 2-fold change in representation between the input and output pool (P<0.0001) in individual competitive infections, and eleven were confirmed in mutants containing the antibiotic cassette. We retested eight of these mutants (*STM1131*, *STM2303*, *STM3120*, *STM3121*, *sroA*, *istR*, *leuX* and *oxyS*) after removal of the antibiotic resistance cassette. The reduced competitive fitness seen in the mutants containing an antibiotic resistance cassette was also seen in all eight of these mutants after removal of this cassette (**Figure 4A** and **Table S3**). We complemented two of these novel mutants, *STM2303* (*pmrM*) and *STM3121* (a LysR-type transcriptional regulator), with the corresponding intact open reading frames *in trans*. The complemented mutants were found in the spleen in numbers similar or greater than the wild-type (**Figure 4B**). These findings directly and unambiguously link the genes we complemented here to the observed defect in fitness of the corresponding mutant during competitive infection with the wild-type. Thus, most of the remaining 25 mutants of the original top 40 novel candidates that we identified may have a phenotype

in this BALB/c acute systemic infection model. There are likely to be further phenotypes yet to be discovered among the many genes that were not mutated for the experiments performed here.

Using our method, the overall false negative rate (failure to observe phenotypes for mutants already known to be under strong selection in this model and measured on the array) was zero. Among 120 mutants showing a phenotype with P<0.0001 and a greater than two-fold change, more than half are already known to have a phenotype and among the new candidates, eleven out of fifteen were confirmed. Thus, the false discovery rate is very low. However, the false discovery rate for new mutants in a system that is already well studied is invariably high. When most genes with a role in a process are known, only a few phenotypes remain to be discovered among the thousands of genes that have no phenotype in that process. The future rate of discovery in the intraperitoneal model in BALB/c mice is unlikely to be better than the rate of 11 out of 15 that we experienced. The rate of confirmation of mutants that show a weaker statistical confidence and lower fold-difference in fitness is likely to be worse. In addition, those mutants with subtle phenotypes in an animal will require a larger number of biological replicates in order to establish the phenotype with confidence. Other phenotypes that will be missed by using a pool of mutants are any of those rescued, *in trans*, by the presence of wild-type bacteria.

Of the mutants we confirm in individual competitive infections with wild-type, four were in small RNA molecules. Mutations in *leuX*, encoding a minor tRNA-Leu, were previously known to reduce the expression of Type I fimbriae, and reduce bladder epithelial invasion and intracellular proliferation of Uropathogenic *E. coli* [25,26]. The attenuation of a *leuX* deletion during *Salmonella* infection was previously unknown. Mutants in the regulatory sRNAs *IstR*, *OxyS* and *SroA* showed small, but reproducible, phenotypes in our *in vivo* experiment, and all have orthologs in *E. coli*. In *E. coli*, *IstR* inhibits the translation of *tixAB*, an SOS-induced toxic peptide, and may arrest growth allowing DNA repair [27,28,29,30]. *SroA* is a known regulator of the *thiBPQ* operon in *E. coli*, that encodes an ABC transport system for the import of thiamine and thiamine pyrophosphate (TPP) into the cell [31]. TPP is an essential cofactor for key enzymes in carbon metabolism.

The *oxyS* mutant was one of the few examples of a mutant that displayed an increase in fitness *in vivo* in our screen. *OxyS* is a member of the OxyR regulon expressed during oxidative stress. This sRNA is a pleiotropic regulator of about 40 genes in *E. coli* including *spoS*, and an anti-mutator that may inhibit alternate stress adaptation pathways when OxyR is activated [30,32]. Our observations for *OxyS*, *SroA*, and *IstR*, are the first examples of phenotypes for small non-coding RNA mutants during *Salmonella* infection, other than the tmRNA system that targets proteins for degradation [33]. The targets of these sRNAs in *Salmonella* are not fully known, and given, for example, the recently reported variable composition of the PhoP regulon across related organisms [34], may differ significantly from their targets in related organisms.

Of the seven remaining confirmed mutant phenotypes, four are in genes that encode putative membrane proteins (*STM1131*, *STM2120*, *STM2215* and *STM2303*), and a fifth gene (*STM1760*) encodes a putative secreted protein. *STM1131* encodes a member of the KdgM superfamily (oligogalacturonate-specific porin). A paralog of this gene, *STM4016* (*yshA*, 26% identity, 40% similarity) is involved in O-Antigen capsule production and environmental persistence [35]. Mutants in a second membrane protein, *STM2120* (*asmA*), are highly sensitive to gastric contents of swine, but otherwise little is known of the function of this gene in *Salmonellae* [36]. Mutants in the *E. coli* ortholog of *asmA* are extragenic suppressors of outer membrane protein assembly and have less LPS on the bacterial surface [37]. A third gene encoding

a predicted membrane protein, *STM2215*, is one of at least eight proteins encoded by Typhimurium that contain cyclic-di-GMP phosphodiesterase, or EAL domains (named after the amino acids in the conserved domain). Unlike several other EAL-containing proteins, mutations in *STM2215* do not influence a multicellular behavior known as the rdar morphotype [38]. Our work reports the first phenotype of any kind for *STM2215*.

Finally, *STM2303* (*pmrM*; *phgE3*) is the terminal gene in the *pmr* operon (polymyxin resistance) and also encodes a predicted membrane protein. This operon is involved in 4-aminoarabinoic acid addition to Lipid A, a modification that confers resistance to polymyxin [39]. Paradoxically, PmrM does not appear to be necessary for this function [39,40], but may be important for resistance of *Salmonella* to high iron concentrations *in vitro* [41,42]. Recent studies show that mutants in *STM2303* have reduced expression of *hilA*, an important regulator of virulence genes in *Salmonella*, and reduced invasion in T84 human colonic epithelial carcinoma cells [43].

Mutants in *STM1760*, encoding a protein containing a classical N-terminal secretion signal sequence, were also confirmed as selected against in our competitive infection assays. The protein encoded by this gene contains multiple Tricopeptide Repeats (TPR) of the Sel-1 superfamily, a motif that mediates protein-protein interactions [44]. Nothing further is reported in the literature about this protein and its potential role during systemic infection by *Salmonella*.

The final genes we confirmed to be important during systemic infection in BALB/c mice, *STM3120* and *STM3121*, are neighboring in the genome. *STM3121* is a LysR family transcriptional regulator neighboring *STM3120* but is transcribed in the opposite direction. Interestingly, expression of *STM3121 in trans* appeared to increase fitness during infection (**Figure 4B**). *STM3121* has not previously been reported to be important during Typhimurium infection, and determination of its regulatory targets may reveal clues as to its role during infection.

*STM3120* and several adjacent genes (cluster *STM3120-STM3117*) are present in serotypes Typhimurium, Enteritidis, and Gallinarum [45] and in some *Yersinia* [46], but are not present in the typhoidal *Salmonellae* or other enterobacterial species studied to date. A second gene in this group, *STM3119*, was also selected against during our ABACUS screen (**Table S1**). Proteomic and transcriptional data indicates that the proteins encoded in this region are highly up-regulated in macrophages [47,48,49]. In serotype Enteritidis, these genes are essential for growth in chicken macrophages and mutants in some of these genes in Gallinarum are defective for systemic colonization after oral infection in chickens [45,50].

Although *STM3120* is a predicted *citE* (citrate lyase) homolog, the other two subunits (*citD* encoding the  $\alpha$ -subunit, and *citF* encoding the  $\gamma$ -subunit) that normally make up bacterial ATP-dependent citrate lyase are not encoded in the region containing this gene. This suggests that *STM3120* may have a different function than a traditional citrate lyase. In *Yersinia pestis*, homologs of *STM3120-3117* are necessary for survival in macrophages that are activated with  $\gamma$ -interferon after infection. These genes appear to function in *Y. pestis* by lowering the level of nitric oxide in macrophages without affecting iNOS levels in  $\gamma$ -interferon post-treated macrophages [46]. It is possible that these genes play a similar role in *Salmonella*, but this hypothesis remains to be investigated.

To summarize, we have designed a collection of novel technologies that permit easy generation and confirmation of specific deletion mutants for genetic screening, both individually and in pools. Use of our collection, designed to minimize the effects of population bottlenecks on screening, will allow relevant

animal models and more relevant doses to be efficiently used to identify *Salmonella* mutants with altered fitness *in vivo*. During our testing of this approach, eleven of such mutants were confirmed in individual competition assays, including three sRNAs and *leuX*. The tools used here are a first step to a more complete description of *Salmonella* genes involved in systemic infection in mice, particularly those genes with milder phenotypes that were difficult to identify with confidence in previous studies. These tools are also of wide applicability to identify genes involved in other aspects of *Salmonella* biology. Finally, similar strategies can be applied in other genetically malleable bacteria.

## Materials and Methods

### Ethics statement regarding the use of animals, choice of species, and numbers of animals

The search for new mechanisms in infection has no *in vitro* surrogate. Mice are natural hosts for *Salmonella* and are also a potential source of human infection. The library and pooling strategies with microarray output (ABACUS) reduces the number of animals required to obtain mutant phenotypes by at least three orders of magnitude compared to screening individual mutants.

### Veterinary care

All animals were housed in AAALAC-accredited animal research facilities at Texas A&M University and were observed daily for assessment of animal health. The facilities are in compliance with the standards set forth in the PHS "Guide for the Care and Use of Laboratory Animals".

### Methods for pain and stress relief

If during the course of these experiments any animal showed anorexia for greater than 24 hr., disinclination to move for 12 hours, loss of normal neurological function as evidenced by locomotory or balance deficits then euthanasia was performed using CO<sub>2</sub> (>70%, inhaled) in accordance with the 2000 AVMA Guide for Euthanasia.

### Strains and standard culture conditions

All strains used in this study are derivatives of *Salmonella enterica* serovar Typhimurium ATCC14028 (Manassas, VA). Strains were routinely cultured in Luria-Bertani (LB) broth and plates, supplemented with 50 mg/l Kanamycin (Kan), 20 mg/l Chloramphenicol (Cm) or 100 mg/l Ampicillin (Amp) where appropriate.

### Construction of plasmids with T7 RNA polymerase promoters

We used a PCR-based strategy to include a T7 promoter in the original pKD3 (Cm<sup>R</sup>) and pKD4 (Kan<sup>R</sup>) vectors for the lambda-red recombination method. Briefly, pKD3 and pKD4 were used as template for independent PCR reactions using primers PT72\_EcoRI (ACTCGAATTCCGAAATTAATACGACTCACTATAG-**GGAGACCTA**GGGAGGATATTCATATG) and FRT3-EcoRI (CATCGAATTCCTATTCCGAAGTTCCCTATTCTAGAAAGTATAGGAACCTCGGCGCGCT) or primers PT72-EcoRI and FRT4-EcoRI (CATCGAATTCCTATTCCGAAGTTCC-TATTCTCTAGAAAGTATAGGAACCTCAGAGCGCT), respectively. The EcoRI sites are underlined and the P<sub>T7</sub> is in bold. Each PCR product was digested with EcoRI, gel purified and self-ligated. Electrocompetent *E. coli* EC100D *pir*-116 (Epicentre), was transformed with aliquots of the ligation, and transformants were selected at 37°C on LB Cm or LB Kan plates. The sequence (both strands) of the resulting plasmids pCLF3 and pCLF4 was

determined by the conventional dye-terminating Sanger method, and deposited under GenBank accession numbers EU629213 and EU629214, respectively.

### Generation of specific deletion mutants

Deletion strains were generated using the lambda-red recombination method [6], with the following modifications, as illustrated in **Figure 1**. Plasmids pCLF3 (Cm<sup>R</sup>) and pCLF4 (Kan<sup>R</sup>) were used as templates to generate unique PCR products for deletion of each gene of interest. The primers used are shown in **Table S1**. PCR amplifications were carried out in 96 well plates using ExTaq polymerase (Takara) in a total reaction volume of 60  $\mu$ l, and a PCR reaction of 30 cycles at an annealing temperature of 55°C. For each gene to be deleted, 30  $\mu$ l of both PCR products bearing different antibiotic markers were combined and purified using the Qiaquick PCR purification kit in 96 well format (Qiagen). Mixed, purified PCR products (10  $\mu$ l) were dialyzed against sterile water for 10 minutes using filters with 0.025  $\mu$ M pore size (Millipore). 2–5  $\mu$ l of mixed, purified PCR products were used for transformation by electroporation (Bio-Rad Gene Pulser) of electrocompetent ATCC14028 expressing lambda-red recombinase, prepared as previously described [6]. Transformations were allowed to recover for 3 hours in LB at 37°C and 150  $\mu$ l of each transformation was plated on LB Kan and LB Cm plates. The remaining transformation was saved and replated when very few colonies were obtained using the shorter grow out period. Two transformants for each gene and each antibiotic resistance marker were twice colony purified, and stored in 30% glycerol at –80°C. The Kan<sup>R</sup> mutants characterized here are available to the community.

### Labeling technique, part I: DNA isolation, sonication and polyadenylation

Our labeling protocol is outlined in **Figure 2**. Genomic DNA was prepared for the input and output pools of ~1000 mutants using the GenElute Bacterial Genomic DNA kit (Sigma). 4  $\mu$ g of genomic DNA was fragmented by sonication using twenty pulses of two seconds in a Branson Sonifier 150 (Branson Ultrasonics Corp., Danbury, CT). PolyA tails were added to fragmented genomic DNA using terminal transferase (TdT) as follows: 1.5  $\mu$ g of DNA fragments were incubated for 30 min at 37°C in a total reaction volume of 50  $\mu$ l containing 40 U TdT (New England BioLabs.), CoCl<sub>2</sub> 0.25 mM, and dATP 0.4 mM. Terminal transferase was subsequently inactivated at 70°C for 10 min and the tailed product was purified using the QIAquick PCR purification kit (Qiagen).

### Labeling technique, part II: PCR amplification and labeling of fragments adjacent the deletion location

Nested PCR was used to amplify the polyA-tailed DNA fragments containing the insert P<sub>T7</sub> and the flanking inserted region. In the first PCR reaction, 50 ng of purified polyA-tailed DNA was used as template for a PCR reaction using primer FRT Out 3-1 (TTCCTATACTTTCTAGAGAA), and a primer designed to anneal to the polyA-tail (CCT<sub>24</sub>VN). The reaction mixture consisted of 1 $\times$  PCR buffer, 0.2 mM of dNTP, 1.5 mM MgCl<sub>2</sub>, 0.05 U Taq polymerase (Promega, WI), and 0.2  $\mu$ M of each forward and reverse primer in a total reaction volume of 25  $\mu$ l. The PCR reaction was performed under the following conditions: initial denaturation at 94°C for 1 min followed by 30 cycles with denaturation at 94°C for 10 s, annealing at 50°C for 10 s, and extension at 72°C for 5 s. The last cycle was followed by a final extension for 3 min at 72°C. In the second amplification step, a nested PCR was performed using 1  $\mu$ l amplified product from the initial PCR in a total volume of 50  $\mu$ l. Internal primer FRT Out 3-2

(TAGGAAC TTCGGAATAGGAA) and primer CCT<sub>24</sub>VN were used under identical cycling conditions as during the initial PCR reaction. PCR products were analyzed on 1% agarose gels.

An aliquot of 8  $\mu$ l of the nested PCR reaction was used directly as template for a 20  $\mu$ l *in vitro* transcription reaction using the AmpliScribe T7 transcription kit (Epicentre), following the manufacturer's protocol with some modifications. The RNA was labeled during the synthesis by including 2  $\mu$ l of 5 mM Cy5- or Cy3-UTP (GE Healthcare) in the *in vitro* transcription reaction. Labeled RNA was treated with RNase-free DNase (Epicentre) and purified with the RNeasy Mini Kit (Qiagen).

### NimbleGen microarray hybridization

Hybridizations were performed according to the manufacturer's protocols (NimbleGen Systems, Madison, WI, <http://www.nimblegen.com/products/lit/lit.html>, [cgh\\_userguide\\_2008\\_05\\_27\[1\].pdf](http://www.nimblegen.com/products/lit/lit.html)) with some modifications. For each hybridization, 4  $\mu$ g of labeled RNA was mixed with alignment oligo, NimbleGen hybridization components and hybridization buffer. The arrays were hybridized at 42°C for 16 hours. Arrays were washed according to the manufacturer's protocol, and scanned using a GenePix 4000B laser scanner (Molecular Devices, Sunnyvale, California) at 5  $\mu$ m resolution. The signal intensities were quantified using NimbleScan software v2.4 (NimbleGen Systems). The data was normalized and analyzed using Webarray [17] and WebarrayDB ([www.webarraydb.org](http://www.webarraydb.org)), with quantile normalization. The moving median intensity of five adjacent probes on the same strand in the genome was calculated and plotted against the genome annotation. Peaks were identified visually (**Figure S1**).

### In-house array hybridization

The 933 65mer 3' oligos and over 1000 control oligos were printed at 10  $\mu$ g/ml in 50% DMSO on Corning UltraGap II slides in triplicate arrays. Oligo sequences are in **Table S1**, and the arrays deposited in GEO as platform GPL5687 and GPL5688. Array hybridization was performed in hybridization chambers (Corning Inc., Corning, NY), following protocols suggested by the manufacturer for hybridizations in formamide buffer for pre-hybridization, hybridization, and post-hybridization washes ([http://www.corning.com/Lifesciences/technical\\_information/techDocs/gaps\\_ii\\_manual\\_protocol\\_5\\_02\\_cls\\_gaps\\_005.pdf](http://www.corning.com/Lifesciences/technical_information/techDocs/gaps_ii_manual_protocol_5_02_cls_gaps_005.pdf)). Immediately before hybridization, 2  $\mu$ g of the labeled probes for input and 2  $\mu$ g of each experimental sample were unified in 36.5  $\mu$ l, mixed with 3.5  $\mu$ l of 100  $\mu$ M 27mer competitor oligo (CATATGAATATCCTCCT-TAGGTCTCCC) complementary to the 27 bases of RNA produced by every mutant (**Figure S2**), and 40  $\mu$ l of 2 $\times$  hybridization buffer (50% formamide, 10 $\times$  SSC [1 $\times$  SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 0.2% sodium dodecyl sulfate) and then denatured at >96°C for 5 min. The probes were hybridized to the *Salmonella* microarray overnight at 42°C in a water bath.

After overnight incubation, slides were washed and scanned using the ScanArray 5000 laser scanner (Packard BioChip Technologies, Billerica, Mass.) with ScanArray 2.1 software. Fluorescence signal intensities were quantified using the QuantArray 2.0 software package (Packard BioChip Technologies). The data were subsequently analyzed using Webarray [17] ([www.webarraydb.org](http://www.webarraydb.org)), with quantile normalization.

### Recombination between FRT sites using FLP recombinase

Antibiotic resistance cassettes were removed from targeted deletion mutants as previously described [6]. Briefly, each mutant

was transformed with the temperature-sensitive plasmid pCP20, that encodes FLP recombinase [51]. Transformants bearing the plasmid were selected at 30°C on LB Ampicillin plates. Several colonies from each transformation were streaked twice and grown at 37°C on LB plates (without antibiotics). Loss of the antibiotic resistance cassettes was determined by patching candidate strains on LB containing the appropriate antibiotics (either Kan or Cm). The concomitant loss of plasmid pCP20 was confirmed by patching the candidate mutants on LB Amp plates.

In our mutant collection, the recombination event leaves ten codons from the 5' end of the original gene, a 39 codon region containing the P<sub>T7</sub> (referred to as a “scar”), and the final ten codons of the original gene (**Figure 1**). Scarred derivatives of mutants  $\Delta STM0854::kan$ ,  $\Delta STM1402::kan$ ,  $\Delta STM2840::kan$ ,  $\Delta STM2884::kan$ ,  $\Delta STM2581::kan$  and  $\Delta STM2581::cam$  were generated to ensure the accuracy of FLP recombination. Two independent examples of each mutant were picked and the presence of each scarred mutant allele was confirmed by PCR amplification using primers flanking mutation, and the subsequent sequence of both strands of the PCR product was obtained by the conventional dye-terminating Sanger method. Non-resistant scarred derivatives of mutants  $\Delta leuX::kan$ ,  $\Delta sroA::kan$ ,  $\Delta istR::kan$ ,  $\Delta oxyS::kan$ ,  $\Delta STM1131::kan$ ,  $\Delta STM2303::kan$ ,  $\Delta STM3120::kan$ , and  $\Delta STM3121::kan$  were generated for further testing in mice.

### Construction of complementing plasmids

*STM2303* was amplified for complementation using the following primers: *STM2303* forward primer 5' GGCGTCTGGTACCAATTCAGTAT 3', and *STM2303* reverse primer 5' TGT-TCTGGATCCGTCGATAGC 3'. The resulting 947 base pair PCR product, in which the only full length ORF is *STM2303*, was purified using agarose gel electrophoresis and extracted from the gel using the Qiaquick gel extraction kit (Qiagen). *STM3121* was amplified for complementation using the following primers: *STM3121* forward primer 5' GCTATTTTCAGGGTACCCGTTTTGGTTCG 3', and *STM3121* reverse primer 5' GCTCCGTTAGCGGATCCCTTTAGACAC 3'. The resulting 1,266 base pair PCR product was purified by agarose gel electrophoresis, and extracted from the gel using the Qiaquick gel extraction kit (Qiagen). The PCR product contains the full-length *STM3121* open reading frame (879 bp) with approximately 296 bp of upstream and 90 bp of downstream sequence.

Purified PCR products were digested with *KpnI* and *BamHI* and ligated into pWSK29 previously cut with the same enzymes and gel purified. *E. coli* DH5 $\alpha$  was transformed with each ligation, and positive clones were selected on LB plates containing Ampicillin. Plasmids bearing the correct insert from each cloning were found by identification of fragments of the proper size after digestion with *PvuI*. pWSK29, pWSK29::*STM2303*, and pWSK29::*STM3121* were transformed into *S. Typhimurium* strain LB5000 (restriction-, modification+) [52]. Plasmid DNA was prepared from LB5000 plasmid-bearing transformants using the Qiaprep spin Miniprep kit (Qiagen), and purified plasmid DNA was used to transform  $\Delta STM2303::FRT$  and  $\Delta STM3121::FRT$  mutants using heat shock. Plasmid-bearing deletion mutants were selected on LB Amp plates, and were twice colony purified prior to use for infections.

### Animal studies in BALB/c mice

The pools of ~1,000 mutants used as inocula were grown overnight at 37°C in LB with aeration and serially diluted in PBS to the proper concentration for inoculation, and the titer of the inoculum was determined by plating on LB containing Kanamycin. A group of 6 BALB/c mice (8–10 week old female) was infected intraperitoneally with  $1 \times 10^6$  CFU of the pool in 100  $\mu$ l

PBS. Mice were monitored twice daily for signs of infection, and were humanely euthanized at 48 hours post infection. Immediately after euthanasia, the spleen was removed and homogenized in 5 ml sterile ice cold PBS. 100  $\mu$ l of the spleen homogenate was used for serial dilution and titer on LB Kan plates. The bacteria in the remaining homogenate were grown in LB to stationary phase and total DNA was extracted.

Individual mutants with an apparent phenotype in the pools were chosen for confirmation in mixed infections with wild-type strain HA431 (ATCC14028 *phaN* Na<sup>R</sup>), to determine the competitive index. Inactivation of *phaN*, encoding alkaline phosphatase, abolishes the ability to cleave 5-bromo-4-chloro-3-indolyl phosphate (XP), but does not reduce the ability of serotype Typhimurium to colonize organs or reduce fecal shedding of this organism in mice during competitive infections [53]. Growth on LB agar plates supplemented with XP thus provided an easy means to distinguish between individual mutants (PhoN<sup>+</sup> blue colonies) and HA431 (PhoN<sup>-</sup> white colonies).

Prior to further analysis, the mutant allele present in each individual mutant to be confirmed was transferred to HA420 (ATCC14028 Na<sup>R</sup>) by P22 transduction. Mutant strains and our wild-type strain HA431 were grown overnight at 37°C with aeration, mixed in a 1:1 ratio, and serially diluted in PBS to the proper concentration for inoculation and titer. Titer was used to determine the exact ratio of strains administered. Groups of 4–6 BALB/c mice (8–10 week old female) were infected intraperitoneally with  $1 \times 10^6$  CFU of mutant and wild-type virulent HA431 in 100  $\mu$ l PBS. Mice were euthanized and spleens, livers, and ceca were recovered as above. Mutant and wild-type organisms were enumerated by serial dilution and plating on LB plates containing 20 mg/L of XP.

### Supporting Information

**Figure S1** NimbleGen tiling array data for T7 RNA polymerase transcripts from a pool of 1031 Kanamycin resistant specific deletion mutants.

Found at: doi:10.1371/journal.ppat.1000477.s001 (3.18 MB PDF)

**Figure S2** Addition of the 27mer competitor oligo results in high specificity of array detection.

Found at: doi:10.1371/journal.ppat.1000477.s002 (0.10 MB PDF)

**Table S1** Mutants constructed and results of selection *in vivo*.

Found at: doi:10.1371/journal.ppat.1000477.s003 (0.57 MB XLS)

**Table S2** Specificity and sensitivity of array-based detection of mutants in a pool.

Found at: doi:10.1371/journal.ppat.1000477.s004 (0.01 MB PDF)

**Table S3** Details of the Competitive Index data.

Found at: doi:10.1371/journal.ppat.1000477.s005 (0.03 MB PDF)

### Acknowledgments

We particularly thank Brian Ahmer for valuable discussions and advice that we consider worthy of being a co-author. We thank Dr. YiPeng Wang for producing Figure S1. We thank Nabil Arrach for helping with the construction of some mutants, Rocio Canals for discussions, and Pui Cheng, Erin Katrube, Andy Bugbee, Christine Shields, Hee-Jeong Yang, and Lydia Bogomolnaya for technical assistance. We thank Jörg Vogel for advice on the positions of candidate sRNA genes.

### Author Contributions

Conceived and designed the experiments: HLAP MM. Performed the experiments: CAS MMR SP SHC HLAP. Analyzed the data: CAS SP FL HLAP MM. Wrote the paper: CAS SP HLAP MM.

## References

- Hensel M, Shea JE, Gleeson C, Jones MD, Dalton E, et al. (1995) Simultaneous identification of bacterial virulence genes by negative selection. *Science* 269: 400–403.
- Sasseti CM, Boyd DH, Rubin EJ (2001) Comprehensive identification of conditionally essential genes in mycobacteria. *Proc Natl Acad Sci U S A* 98: 12712–12717.
- Sasseti CM, Boyd DH, Rubin EJ (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* 48: 77–84.
- Lawley TD, Chan K, Thompson IJ, Kim CC, Govoni GR, et al. (2006) Genome-wide screen for salmonella genes required for long-term systemic infection of the mouse. *PLoS Pathog* 2: e11. doi:10.1371/journal.ppat.0020011.
- Badarinarayana V, Estep PW 3rd, Shendure J, Edwards J, Tavazoie S, et al. (2001) Selection analyses of insertional mutants using subgenic-resolution arrays. *Nat Biotechnol* 19: 1060–1065.
- Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97: 6640–6645.
- McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, et al. (2001) Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* 413: 852–856.
- Wilmes-Riesenberg MR, Foster JW, Curtiss R (1997) An altered *rpoS* allele contributes to the avirulence of *Salmonella typhimurium* LT2. *Infect Immun* 65: 203–210.
- Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, et al. (2006) Selective Silencing of Foreign DNA with Low GC Content by the H-NS Protein in *Salmonella*. *Science*.
- Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, et al. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2: 2006 0008.
- Knuth K, Niesalla H, Hueck CJ, Fuchs TM (2004) Large-scale identification of essential *Salmonella* genes by trapping lethal insertions. *Mol Microbiol* 51: 1729–1744.
- Relogio A, Schwager C, Richter A, Ansorge W, Valcarcel J (2002) Optimization of oligonucleotide-based DNA microarrays. *Nucleic Acids Res* 30: e51.
- Chan K, Kim CC, Falkow S (2005) Microarray-based detection of *Salmonella enterica* serovar Typhimurium transposon mutants that cannot survive in macrophages and mice. *Infect Immun* 73: 5438–5449.
- Gulig PA, Curtiss R 3rd (1988) Cloning and transposon insertion mutagenesis of virulence genes of the 100-kilobase plasmid of *Salmonella typhimurium*. *Infect Immun* 56: 3262–3271.
- Sizemore DR, Fink PS, Ou JT, Baron L, Kopecko DJ, et al. (1991) Tn5 mutagenesis of the *Salmonella typhimurium* 100 kb plasmid: definition of new virulence regions. *Microb Pathog* 10: 493–499.
- Bowe F, Lipps CJ, Tsolis RM, Groisman E, Heffron F, et al. (1998) At least four percent of the *Salmonella typhimurium* genome is required for fatal infection of mice. *Infect Immun* 66: 3372–3377.
- Xia X, McClelland M, Wang Y (2005) WebArray: an online platform for microarray data analysis. *BMC Bioinformatics* 6: 306.
- Wang RF, Kushner SR (1991) Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* 100: 195–199.
- Bohnhoff M, Drake BL, Miller CP (1954) Effect of streptomycin on susceptibility of intestinal tract to experimental *Salmonella* infection. *Proc Soc Exp Biol Med* 86: 132–137.
- Hentges DJ (1970) Enteric pathogen–normal flora interactions. *Am J Clin Nutr* 23: 1451–1456.
- Que JU, Hentges DJ (1985) Effect of streptomycin administration on colonization resistance to *Salmonella typhimurium* in mice. *Infect Immun* 48: 169–174.
- Barthel M, Hapfelmeier S, Quintanilla-Martinez L, Kremer M, Rohde M, et al. (2003) Pretreatment of mice with streptomycin provides a *Salmonella enterica* serovar Typhimurium colitis model that allows analysis of both pathogen and host. *Infect Immun* 71: 2839–2858.
- Meynell GG (1957) The applicability of the hypothesis of independent action to fatal infections in mice given *Salmonella typhimurium* by mouth. *J Gen Microbiol* 16: 396–404.
- Carter PB, Collins FM (1974) The route of enteric infection in normal mice. *J Exp Med* 139: 1189–1203.
- Hamman TJ, Mysorekar IU, Chen SL, Walker JN, Jones JM, et al. (2008) LeuX tRNA-dependent and -independent mechanisms of *Escherichia coli* pathogenesis in acute cystitis. *Mol Microbiol* 67: 116–128.
- Ritter A, Gally DL, Olsen PB, Dobrindt U, Friedrich A, et al. (1997) The *Pai*-associated *leuX* specific tRNA<sup>5(Leu)</sup> affects type 1 fimbriation in pathogenic *Escherichia coli* by control of FimB recombinase expression. *Mol Microbiol* 25: 871–882.
- Vogel J, Argaman L, Wagner EG, Altuvia S (2004) The small RNA *IstR* inhibits synthesis of an SOS-induced toxic peptide. *Curr Biol* 14: 2271–2276.
- Altuvia S (2004) Regulatory small RNAs: the key to coordinating global regulatory circuits. *J Bacteriol* 186: 6679–6680.
- Darfeuille F, Unoson C, Vogel J, Wagner EG (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Mol Cell* 26: 381–392.
- Repoila F, Majdalani N, Gottesman S (2003) Small non-coding RNAs, coordinators of adaptation processes in *Escherichia coli*: the RpoS paradigm. *Mol Microbiol* 48: 855–861.
- Webb E, Claas K, Downs D (1998) *thiBPQ* encodes an ABC transporter required for transport of thiamine and thiamine pyrophosphate in *Salmonella typhimurium*. *J Biol Chem* 273: 8946–8950.
- Altuvia S, Weinstein-Fischer D, Zhang A, Postow L, Storz G (1997) A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. *Cell* 90: 43–53.
- Julio SM, Heithoff DM, Mahan MJ (2000) *ssrA* (tmRNA) plays a role in *Salmonella enterica* serovar Typhimurium pathogenesis. *J Bacteriol* 182: 1558–1563.
- Perez JC, Shin D, Zwir I, Latifi T, Hadley TJ, et al. (2009) Evolution of a bacterial regulon controlling virulence and Mg(2+) homeostasis. *PLoS Genet* 5: e1000428. doi:10.1371/journal.pgen.1000428.
- Gibson DL, White AP, Snyder SD, Martin S, Heiss C, et al. (2006) *Salmonella* produces an O-antigen capsule regulated by *AgfD* and important for environmental persistence. *J Bacteriol* 188: 7722–7730.
- Bearson SM, Bearson BL, Rasmussen MA (2006) Identification of *Salmonella enterica* serovar Typhimurium genes important for survival in the swine gastric environment. *Appl Environ Microbiol* 72: 2829–2836.
- Deng M, Misra R (1996) Examination of *AsmA* and its effect on the assembly of *Escherichia coli* outer membrane proteins. *Mol Microbiol* 21: 605–612.
- Simm R, Lusch A, Kader A, Andersson M, Romling U (2007) Role of EAL-containing proteins in multicellular behavior of *Salmonella enterica* serovar Typhimurium. *J Bacteriol* 189: 3613–3623.
- Gunn JS, Ryan SS, Van Velkinburgh JC, Ernst RK, Miller SI (2000) Genetic and functional analysis of a *PmrA*-*PmrB*-regulated locus necessary for lipopolysaccharide modification, antimicrobial peptide resistance, and oral virulence of *Salmonella enterica* serovar typhimurium. *Infect Immun* 68: 6139–6146.
- Murray SR, Ernst RK, Bermudes D, Miller SI, Low KB (2007) *pmrA*(Con) confers *pmrHFIJKL*-dependent EGTA and polymyxin resistance on *msbB* *Salmonella* by decorating lipid A with phosphoethanolamine. *J Bacteriol* 189: 5161–5169.
- Wosten MM, Kox LF, Chamnongpol S, Soncini FC, Groisman EA (2000) A signal transduction system that responds to extracellular iron. *Cell* 103: 113–125.
- Nishino K, Hsu FF, Turk J, Cromie MJ, Wosten MM, et al. (2006) Identification of the lipopolysaccharide modifications controlled by the *Salmonella PmrA/PmrB* system mediating resistance to Fe(III) and Al(III). *Mol Microbiol* 61: 645–654.
- Van Immerseel F, Eeckhaut V, Boyen F, Pasmans F, Haesebrouck F, et al. (2008) Mutations influencing expression of the *Salmonella enterica* serovar Enteritidis pathogenicity island I key regulator *hilA*. *Antonie Van Leeuwenhoek* 94: 455–461.
- D'Andrea LD, Regan L (2003) TPR proteins: the versatile helix. *Trends Biochem Sci* 28: 655–662.
- Shah DH, Lee MJ, Park JH, Lee SK, et al. (2005) Identification of *Salmonella gallinarum* virulence genes in a chicken infection model using PCR-based signature-tagged mutagenesis. *Microbiology* 151: 3957–3968.
- Pujol C, Grabenstein JP, Perry RD, Bliska JB (2005) Replication of *Yersinia pestis* in interferon  $\gamma$ -activated macrophages requires *ripA*, a gene encoded in the pigmentation locus. *PNAS* 102: 12909–12914.
- Shi L, Adkins JN, Coleman JR, Schepmoes AA, Dohnkova A, et al. (2006) Proteomic analysis of *Salmonella enterica* serovar typhimurium isolated from RAW 264.7 macrophages: identification of a novel protein that contributes to the replication of serovar typhimurium inside macrophages. *J Biol Chem* 281: 29131–29140.
- Hautefort I, Thompson A, Eriksson-Ygberg S, Parker ML, Lucchini S, et al. (2008) During infection of epithelial cells *Salmonella enterica* serovar Typhimurium undergoes a time-dependent transcriptional adaptation that results in simultaneous expression of three type 3 secretion systems. *Cell Microbiol* 10: 958–984.
- Eriksson S, Lucchini S, Thompson A, Rhen M, Hinton JC (2003) Unravelling the biology of macrophage infection by gene expression profiling of intracellular *Salmonella enterica*. *Mol Microbiol* 47: 103–118.
- Zhao Y, Jansen R, Gaastra W, Arkesteijn G, van der Zeijst BAM, et al. (2002) Identification of Genes Affecting *Salmonella enterica* Serovar Enteritidis Infection of Chicken Macrophages. *Infect Immun* 70: 5319–5321.
- Cherepanov PP, Wackernagel W (1995) Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* 158: 9–14.
- Bullas LR, Ryu JI (1983) *Salmonella typhimurium* LT2 strains which are *r<sup>-</sup> m<sup>+</sup>* for all three chromosomally located systems of DNA restriction and modification. *J Bacteriol* 156: 471–474.
- Weening EH, Barker JD, Laarakker MC, Humphries AD, Tsolis RM, et al. (2005) The *Salmonella enterica* serotype Typhimurium *lpf*, *bcf*, *stb*, *stc*, *std*, and *sth* fimbrial operons are required for intestinal persistence in mice. *Infect Immun* 73: 3358–3366.