Identification of host-induced pathogen genes by differential fluorescence induction reporter systems

Dirk Bumann[1,2] & Raphael H Valdivia[3]

[1]Junior Research Group 'Mucosal Infections' OE 9421, Hannover Medical School, Carl-Neuberg-Strasse 1, D-30625 Hannover, Germany. [2]Department of Molecular Biology, Max Planck Institute for Infection Biology, Charitéplatz 1, 10117 Berlin, Germany. [3]Department of Molecular Genetics and Microbiology and Center for Microbial Pathogenesis, Duke University Medical Center, Durham, NC 27710, USA. Correspondence should be addressed to R.H.V. (valdi001@mc.duke.edu).

Published online 5 April 2007; doi:10.1038/nprot.2007.78

The ability to monitor a pathogen's gene expression program in response to the host environment is central to understanding host–microbe interactions. This protocol describes the application of a fluorescence-based promoter trap strategy, termed differential fluorescence induction (DFI), to identify and characterize bacterial genes that are preferentially expressed in infected tissues. In this approach, animals are infected with a library of bacteria expressing random GFP transcriptional gene fusions, and fluorescent bacteria are recovered directly from host tissues using fluorescence-activated cell sorting (FACS). This methodology allows for the identification of bacterial promoters induced in distinct anatomical sites and at different stages of infection. Furthermore, unlike other methodologies, the use of the GFP reporter allows for single cell, temporal and spatial monitoring of pathogen gene expression in infected animals. Library construction, promoter identification and analysis can be done in 4–8 weeks.

INTRODUCTION

Pathogenic bacteria respond to the highly heterogeneous environment of their host by activating and repressing the transcription of a variety of genes. Many of these host-induced genes are required to adapt to new metabolic requirements in host tissues, whereas others encode virulent factors that participate directly in the disruption of host cellular functions and immune responses[1]. Although some 'host-like' conditions can be mimicked *in vitro* (e.g., low iron, acidic pH, antimicrobial peptides and oxidative stress), the host environment remains too complex and dynamic to be accurately modeled in the laboratory[2].

Using specialized gene reporters (e.g., auxotrophy complementation markers, drug resistance cassettes and DNA recombinases), genetic strategies, termed *in vivo* expression technologies (IVET), have been developed that use the host environment to select for active promoters in a classical 'promoter-trap' experimental setting[1,2]. For the most part, bacteria bearing these transcriptionally active promoter fusions are then screened in laboratory media to determine which promoter elements are inactive outside host tissues. Although these approaches have been successful in identifying bacterial genes expressed in infected animals, the selection schemes are biased against the isolation of weaker promoter elements because of the requirement for clonal expansion in tissues under sustained antibiotic or auxotrophy selection. Furthermore, as most of the IVET reporter genes do not provide a readily measurable enzymatic activity that can be quantified *in situ*, they are generally not well suited to quantify bacterial gene expression or to assess the heterogeneity of gene expression in infected animals[3].

Differential fluorescence induction (DFI) is a promoter trap strategy in which a fluorescence-activated cell sorter (FACS) is used to isolate bacteria bearing transcriptionally active *gfp* promoter fusions directly from infected cells or animal tissues[4]. As isolation of single bacterial cells from tissues is based on fluorescence, the selection scheme allows for the unbiased collection of bacteria bearing promoter elements of various strengths. Above all, as bacteria can be analyzed directly by flow cytometry, gene expression levels can be compared between host and nonhost environments with single cell resolution[5].

The identification of differentially expressed genes with promoter trap strategies lacks the parallel processing power of genome-scale approaches such as transcriptional profiling with DNA microarrays. However, DNA microarrays are still of limited use in identifying bacterial genes expressed in animals because of the technical limitation in isolating high quality bacterial RNA from host tissues[6]. In addition, as DNA microarrays measure the average transcriptional response of a bacterial population, the technique cannot resolve the heterogeneity of bacterial gene expression responses in host tissues. Nonetheless, it is theoretically possible that both approaches can be combined using promoter sequences captured after DFI selections as probes for hybridization to DNA microarrays. Such a combination of approaches may prove particularly useful to compare global bacterial promoter activity at distinct anatomical sites, at different stages during infection, in diverse transgenic mouse models or to determine the contribution of mutations in bacterial genes to the regulation of the host-induced bacterial transcriptome.

DFI is not without its pitfalls. For example, as fluorescence intensity of GFP is directly proportional to protein concentration, it lacks the sensitivity of reporter proteins whose signal is amplified by enzymatic activation of a substrate[7]. As a result, most successful DFI applications described to date require the use of multi-copy plasmids to increase the GFP signal artificially[8]. However, expression levels higher than around 200,000 GFP molecules per bacterium should be avoided to prevent nonspecific adverse effects on bacterial fitness[9]. Although the increased copy number and non-chromosomal context can lead to abnormal regulation of promoters, the predicted host-induced gene activation observed with plasmid reporters has been mostly confirmed with single copy gene fusions[10]. In the same way, as GFP is notoriously resistant to degradation, turnover of the GFP is largely dependent on signal dilution by bacterial replication. Therefore, in some instances a reduction in replication rates might be misinterpreted as an

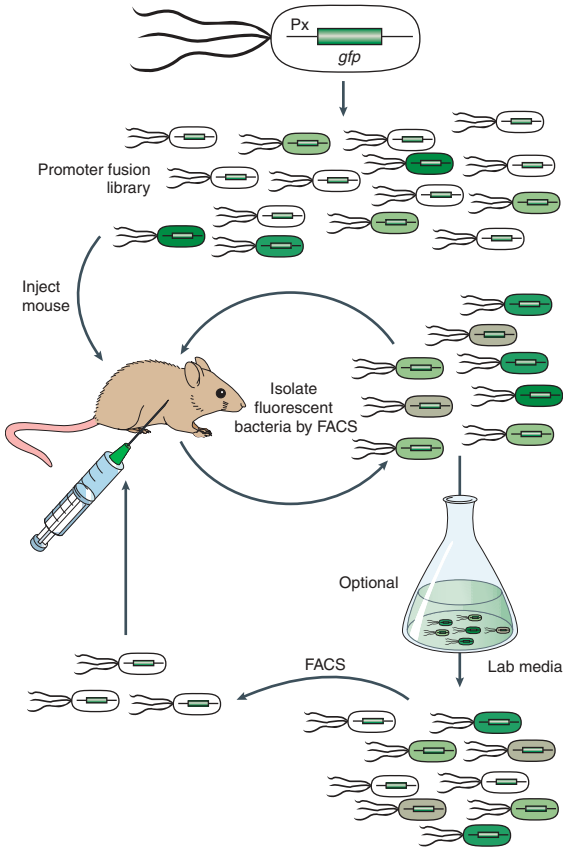

Figure 1 | Identification of bacterial genes expressed in animal tissues. Mice are infected with a pool of *Salmonella enterica* transformed with plasmids containing random *S. enterica* DNA fragments (Px) fused to a promoterless *gfp* gene. Infected tissues are harvested, homogenized and fluorescent bacteria are isolated using fluorescence-activated cell sorting (FACS). Fluorescent bacteria harvested from infected tissues can also be grown in laboratory media (optional—see **Box 1**) and nonfluorescent bacteria (promoters repressed outside host tissues) can be collected using FACS. These nonfluorescent bacteria can then be used to re-infect mice. Fluorescent bacteria recovered from tissues after this second round of infection contain promoters that are preferentially activated in infected animals.

increase in promoter activity. This latter issue has been largely resolved with the use of de-stabilized versions of GFP as gene reporters during infection[11]. A thorough discussion of advantages and disadvantages of GFP as reporter of gene expression can be found elsewhere[12].

Selection rationale

The DFI selection approach is outlined in **Figure 1**. In brief, a plasmid library of random DNA fragments inserted upstream of a promoterless *gfp* gene is transformed into *Salmonella enterica* and the resulting transformants are used to infect mice. Fluorescent bacteria are then isolated directly from infected tissues using FACS and promoters of various strengths are identified by sorting bacterial cells by their fluorescence intensity. Under some circumstances, it may be desirable to define the subset of promoters that are inactive outside the context of a host. An added selection step (**Box 1**) can be included, wherein fluorescent bacteria recovered from infected tissues can be grown in laboratory media and nonfluorescent bacteria can be separated using FACS. These nonfluorescent bacteria are then used to re-infect animals, and fluorescent bacteria are recovered again from tissues. Bacteria recovered after this last selection round contain promoters that are transcriptionally active in host tissues but not in laboratory media.

In this protocol, we illustrate the application of DFI in identifying *S. enterica* serovar Typhimurium genes that are expressed in host tissues. *S. enterica* is a food-borne pathogen that causes ailments ranging from gastroenteritis to typhoid fever[13]. Although the methodology focuses on the identification of *S. enterica* promoters, the same basic flow cytometric-based promoter trap strategy has been applied to a range of Gram positive and negative bacterial pathogens to identify bacterial genes expressed in response to *in vitro* stimuli, tissue culture cells and in infected animals[2,4].

MATERIALS

REAGENTS

- pMW82 (GenBank accession number EF363313, **Fig. 2**), a pBR322-derived (*colE1 bla*) vector carrying a promoter-less unstable *gfp* variant[11] (available upon request from Dirk Bumann) (see REAGENT SETUP) *Note*: A variety of additional GFP-based promoter trap vectors suitable for use in other bacterial species have been described[4]. These vectors can be used as alternatives to pMW82 when different origins of replication, copy number and/or drug resistance markers are required. If desired, the *gfp* genes in these vectors can be destabilized for more accurate reporting of transcriptional activity as outlined elsewhere[9,14]
- Agarose (molecular biology grade; MP Biomedicals)
- Plasmid DNA and gel extraction kits (Qiagen Genomic)

BOX 1 | IDENTIFICATION OF BACTERIAL GENES PREFERENTIALLY EXPRESSED *IN VIVO*

Once *Salmonella enterica* bearing promoters that are transcriptionally active in host tissues has been collected, this complex bacterial pool can be subjected to additional cell sorting cycles after growth *in vitro* to identify promoters that are 'repressed' outside host tissues. As defining *in vitro* conditions that are not encountered in infected animals is highly subjective, it is challenging to determine unequivocally if a gene is only active inside host cells. Nonetheless, one can limit further analysis to promoters that are activated in the host environment but not by host-like stimuli that can be mimicked *in vitro*. For example, a representative sample of the pool of fluorescent bacteria sorted from animal tissues can be grown in tissue culture media supplemented with 10% FBS at 37 °C for 10–20 generations (5–6 h), analyzed using fluorescence activated cell sorter (FACS) and the least fluorescent population (bottom 10%) sorted. This bacterial pool, which is transcriptionally silent in tissue culture media (*in vitro*), is then used to re-infect animals and fluorescent bacteria sorted from the spleens and livers of infected animals as above (Steps 25–28). The most fluorescent bacteria isolated from these sorts contain promoter fusions that are repressed in tissue culture media but active in animal tissues. In this manner, one can focus on the analysis of promoters that are activated by undefined host factors rather than promoters that are simply activated by temperature or serum.

- DNA restriction enzymes (AluI, BamHI; NEB)
- T4 DNA polymerase and T4 DNA polymerase buffer (Roche) (see REAGENT SETUP)
- A 10 mM cocktail of dNTPs (dATP, dCTP, cGTP and dTTP; Roche)
- PCR reagents (*Taq* polymerase and 1× PCR; Promega)
- Oligonucleotides for PCR amplification and sequencing of promoter inserts (5′-GTGATGTCGGCGATATAG-3′ and 5′-TACTCATATGTATATCTC CTTCTTA-3′)
- Ethidium bromide (EtBr; Sigma) **! CAUTION** Carcinogen. Use gloves when handling and disposing EtBr contaminated materials as recommended by Institutional Hazardous Material Disposal guidelines. Wear protective eyewear when excising DNA bands under UV illumination
- Calf-intestinal alkaline phosphatase (CIAP; Roche) (see REAGENT SETUP)
- DNA ligase kit (rapid DNA ligation kit; Roche)
- TE buffer (10 mM Tris–HCl pH 8.0, 1 mM EDTA)
- 1× TBE buffer for DNA electrophoresis (0.089 M Tris, 0.089 M boric acid and 0.002 M EDTA)
- Lennox L liquid media and agar plates (Invitrogen)
- SOC medium for electroporation (Invitrogen)
- Ampicillin (Amp, 100 mg ml^{-1} in sterile water for 1,000× stock; Sigma)
- DMSO (Sigma)
- Triton X-100 (Sigma)
- XL10 electrocompetent *Escherichia coli* (Stratagene) *Note*: XL10 can be substituted with other commercially available, highly electrocompetent *E. coli* strain
- *Salmonella enterica* serovar Typhimurium strain SL1344 (ATCC) **! CAUTION** *S. enterica* is a human pathogen. Standard biosafety level 2 practices and precautions should be in place. Cell sorters potentially generate pathogen-containing aerosols that must be controlled using under-pressure systems equipped with appropriate filters. Consult with the appropriate institutional biosafety review board for the appropriate containment protocols and safety procedures
- 6–12-week-old female BALB/c mice (Charles River) **! CAUTION** All experiments involving mice must be in compliance with institutional and local guidelines for the use of vertebrate animals in research

EQUIPMENT

- High speed cell sorter (e.g., FACSAria; BD Biosciences)
- Bath sonicator (e.g., Elma S30H Elmasonic sonicator)
- Electroporator (e.g., GenePulser; BioRad) and electroporation cuvettes (2-mm gap) (see EQUIPMENT SETUP)
- Thermocycler/PCR machine (see EQUIPMENT SETUP)
- Flow cytometer (see EQUIPMENT SETUP)
- Electrophoresis apparatus

REAGENT SETUP

Restriction digest of vector (pMW82) DNA Digest 10 µg pMW82 with 5 U BamHI for 2 h at 37 °C. BamHI cuts at a single site at the 5′ end of the promoterless *gfp* gene, allowing for the screening of DNA fragments for promoter activity. Resolve the restriction digested DNA on a 0.8% (wt/vol) agarose gel in 1× TBE DNA electrophoresis buffer containing 1 µg ml^{-1} EtBr. In parallel, run 0.5 µg undigested pMW82 as a control in a separate lane. Excise from the agarose gel the linearized vector (size 5129 bp) from trace amounts of undigested vector by monitoring the EtBr-stained DNA bands under low intensity UV illumination. Extract the linearized pMW82 from the agarose plug using a QIAquick gel extraction kit as detailed by the manufacturer.

Fill-in reaction with T4 polymerase To allow cloning of sheared *S. enterica* chromosomal DNA into pMW82, blunt the overhanging BamHI-generated ends by treating the cut vector with 50 U T4 DNA polymerase in a 250 µl reaction volume containing 1× T4 DNA polymerase buffer and 100 µM dNTPs. Incubate for 30 min at 11 °C. Remove enzymes and excess nucleotides from the reaction using the QIAquick PCR purification kit, as detailed by the manufacturer.

CIAP treatment of restricted vector DNA To minimize vector self-ligation, treat the linearized T4 DNA polymerase-treated pMW82 with 1 U CIAP in a 250 µl reaction volume (1× CIAP buffer) for 30 min at 37 °C. Purify the treated vector using QIAquick PCR purification kit as detailed by the manufacturer. *Note:* The vector is ready for use in DNA ligation reactions (at Step 5) and can be stored for approximately 6 months at –20 °C.

EQUIPMENT SETUP

Electroporation of competent cells Adjust electroporator settings to 2.5 kV, 600 Ω, 25 µF for *E. coli* (Step 7) and 1.7 kV, 200 Ω, 25 µF for *S. enterica* (Step 15).

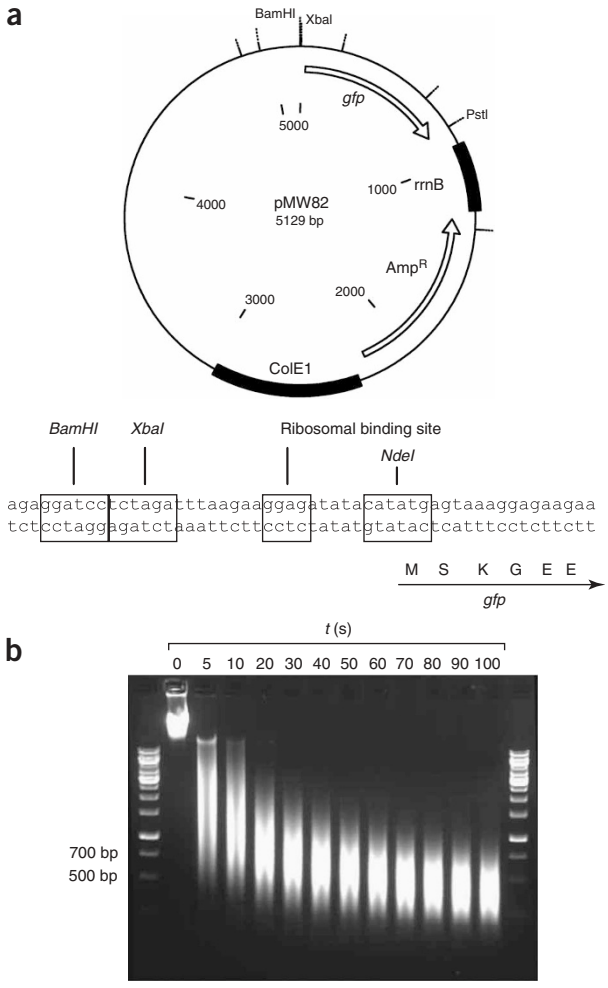

Figure 2 | Construction of promoter trap libraries in *Salmonella enterica*. (**a**) The promoter trap vector pMW82. The promoter trap vector pMW82 is a medium copy plasmid containing a promoterless *gfp* gene, a transcriptional terminator (rrnB) and ampicillin resistance (Amp^R) cassette. The DNA sequence immediately 5′ to *gfp* is shown (bottom), to highlight restriction enzyme recognition sites for cloning of putative promoter elements and the beginning of the gfp open-reading frame (ORF). (**b**) Preparation of *S. enterica* chromosomal DNA for insertion into pMW82. Chromosomal DNA was sheared by sonication for the indicated times (*t*). After each time point, a sample was removed and analyzed by electrophoresis. DNA fragments (500–700 bp range) generated after 50-s sonication were excised from the gel and used for blunt end ligations with pMW82.

Flow cytometer setting Four parameters (forward scatter (FSC), sideward scatter (SSC), uncompensated green (515–545 nm) and orange fluorescence (563–607 nm) after 488 nm excitation) are monitored during sorting. Set photomultipliers (PMTs) for FSC, SSC, FL-1 and FL-2 (FL, fluorescent light; used for 'green' and 'orange' fluorescence detection, respectively) channels on logarithmic scale. Adjust voltage to PMTs to place the bulk of recorded light scattering particles in the center of the FSC/SSC dot plot. Use SSC as the threshold for particle detection. The PMTs of FL-1 and FL-2-channels should be adjusted with *Salmonella* not expressing GFP in such a way that the entire bacterial population is present within the first decade of the logarithmic scale for both fluorescent channels.

PCR settings To amplify promoter captured in pMW82 (Step 28), use the following PCR cycling protocol: 1 cycle 94 °C 5 min; 10 cycles step down (each consecutive cycle is performed at 1 °C lower annealing temperature, i.e., 60 °C, 59 °C), 15 s 94 °C, 30 s annealing (first cycle 60 °C, last cycle 51 °C), 60 s 72 °C elongation; 30 cycles at constant annealing temperature: 15 s 94 °C, 30 s 50 °C, 60 s 72 °C; 1 cycle 7 min 72 °C.

PROCEDURE

Construction of a GFP promoter-trap library in *Salmonella* ● TIMING 6–10 d

1| Isolate genomic DNA from a 10 ml overnight culture of *S. enterica* SL1344 grown in L broth at 37 °C using a Qiagen DNeasy blood & tissue genomic DNA isolation kit according to the manufacturer's description. Run an aliquot of it on a 0.7% agarose gel to ensure that genomic DNA is intact (e.g., no smearing or accumulation of fragments less than 10 kb).
▲ CRITICAL STEP Avoid prolonged vortexing and/or incubation at high temperatures during genomic DNA preparation to minimize shearing and degradation of high MW DNA. For troubleshooting tips, refer to the manufacturer's user manual.

2| Dilute 10 µg genomic DNA in 200 µl TE buffer in a microfuge tube. Set output power in the sonifying water bath to 100% (approximately 320 W) and shear DNA by continuous sonication in 10-s intervals for a total of 30 s.

3| Analyze 4 µl aliquots of sonicated DNA on a 1.5% agarose gel to monitor DNA shearing. Repeat sonication and analysis by gel electrophoresis until most chromosomal DNA consists of 500–700 bp fragments (**Fig. 2b**). This size range should span most bacterial genetic regulatory sequences in intergenic regions and ensure broad coverage of most promoter elements.

4| Run the entire sonicated sample on a preparative agarose gel and excise fragments in the 500–700 bp range. Recover DNA using Qiagen QIAquick gel extraction kit as recommended by the manufacturer.

5| Ligate the sheared genomic DNA with BamHI-digested, T4 polymerase filled-in and CIAP-treated pMW82 prepared as indicated in REAGENT SETUP. Find the optimal insert-to-vector ratios by performing 20 µl test ligations with 50 ng vector and 5, 10, 20 or 40 ng sheared genomic DNA using Roche rapid ligation kit, according to the manufacturer's instructions. Use a vector-only ligation control (no DNA insert) to determine the vector self-ligation background.

6| Purify the ligation mixtures using Qiagen QIAquick PCR purification kit. Elute DNA samples in sterile water.

7| Electroporate 2–5 µl of each purified ligation reaction into 50 µl Stratagene XL10 electrocompetent *E. coli* cells using a Biorad GenePulser (see EQUIPMENT SETUP).
▲ CRITICAL STEP Do not use more than 5 µl ligation mix per electroporation reaction to avoid over-dilution of electrocompetent cells and loss of transformation efficiency.

8| Add 1 ml prewarmed (37 °C) SOC medium immediately after electroporation, transfer to a 10 ml culture tube and incubate for 1 h at 37 °C with shaking at 200 r.p.m.

9| Plate tenfold dilution series (spanning 10^{-3}–10^{-6}) on L-agar plates supplemented with 100 µg ml^{-1} Amp to determine the insert-to-vector ratio DNA ligation reaction that yields the highest number of colonies.
▲ CRITICAL STEP If vector-only ligations control do not yield less than 20% ampicillin resistant (Amp^R) colonies than vector ligations with DNA inserts, repeat DNA ligations with a new batch of linearized and dephosphorylated vector.

10| If necessary, scale up the ligations by setting up multiple reactions at the optimal vector/insert ratio to obtain a minimum of 8×10^4 independent transformants (approximately fivefold genome coverage).

11| Transform purified ligation reactions as detailed in Steps 6–8.

12| Plate transformants on 20 cm diameter L-agar Amp plates at a density of approximately 2×10^4 colonies per plate to minimize crowding, and grow overnight at 37 °C.

13| Scrape colonies from all plates using a wet sterile cotton swab applicator (e.g., Q-Tip^R) and pool cells in a 50 ml centrifuge tube containing 5 ml sterile dH_2O/7% (vol/vol) DMSO. Vortex vigorously and freeze four 250 µl aliquots of the promoter-trap library on dry ice and store at –80 °C. These aliquots represent backups of the promoter library. At any later stage, these aliquots can be thawed and used to inoculate L-broth to re-amplify the plasmid library.
■ PAUSE POINT Plasmid libraries can be kept frozen until needed.
? TROUBLESHOOTING

14| Use the residual suspension (approximately 4 ml) to prepare plasmid DNA using Qiagen Midi plasmid purification kit, according to the manufacturer's instructions. Assess DNA concentration and store the purified plasmid library DNA at –20 °C in 5-µg aliquots at a DNA concentration of 0.1–0.5 mg ml^{-1} in TE buffer.
■ PAUSE POINT Plasmid libraries can be kept frozen until needed. For long-term storage (greater than 6 months), store DNA samples at –80 °C.

15| Prepare electrocompetent *S. enterica* as described by O'Callaghan *et al.*[15]. Transform 0.1–0.2 µg of the promoter library plasmid DNA from Step 14 into *S. enterica* by electroporation (see EQUIPMENT SETUP). To maintain plasmid library diversity, perform

enough electroporations to obtain greater than 10^6 independent Amp^R transformants and plate at a density of less than 2×10^4 colonies per 20 cm plate. Incubate plates overnight at 37 °C. As *E. coli* can be transformed with much higher efficiency (greater than 10^4 fold) than *S. enterica*, *E. coli* should be used as an intermediate host to construct libraries with high diversity and coverage.

16| Scrape colonies from plates with wet cotton swab applicators, resuspend in 20 ml L-broth/7% DMSO (vol/vol) and store the pooled promoter-trap library in 200 µl aliquots at −80 °C.
■ **PAUSE POINT** *S. enterica* libraries can be kept frozen until needed.

Identification of promoters with high activity in infected animals ● TIMING 2–4 weeks
17| Thaw an aliquot of the *S. enterica* library from Step 16 on ice and dilute tenfold in prewarmed L-Amp media. Incubate for 1 h at 37 °C with shaking at 200 r.p.m. to allow bacteria to recover from cryogenesis. For cell sorting controls, also inoculate *S. enterica* transformed with pMW82 (empty vector control) and *S. enterica* constitutively expressing GFP (positive control). Plasmids for the constitutive expression of GFP in *S. enterica* have been described[5,16]. Alternatively, pick a random medium-to-high fluorescent *S. enterica* colony from among the promoter library.

18| Centrifuge bacterial cells at 10,000*g* for 1 min in a microfuge at room temperature (20–25 °C) and wash three times with 1 ml endotoxin-free PBS. Measure OD_{600} for the resuspended culture in a spectrophotometer. For most spectrophotometers an OD_{600} of 1.0 is equivalent to approximately 2×10^9 *S. enterica* per ml. Dilute samples to a density of approximately 10^7 bacteria per ml.

19| Infect BALB/c mice with the washed bacterial suspension by i.p. injection with a 27-gauge needle. To maintain library diversity, an infection dose of 10^6 *S. enterica* cells in 100 µl volume of PBS is recommended. Infect one mouse with the positive (constitutively expressing GFP) and negative (empty vector) *S. enterica* controls.
▲ **CRITICAL STEP** BALB/c mice are naturally sensitive to SL1344 infection and higher infection doses can lead to septic shock. As an alternative to the high infectious doses and/or to increase the infection time to up to 4 d, several mice can be infected in parallel with smaller doses and the bacteria sorted from different animals can be pooled.

20| At 24 h postinfection, kill mice and collect liver and spleens as described in ref. 17
▲ **CRITICAL STEP** The methods of killing used for vertebrate laboratory research animals are regulated by institutional and local guidelines. Consult with your research institution for the appropriate methods for killing.
▲ **CRITICAL STEP** Animals should only be handled by trained personnel to ensure proper care and reduced discomfort to laboratory animals. Consult institutional and local guidelines for proper training.
▲ **CRITICAL** Keep samples on ice throughout Steps 20–22.

21| Homogenize whole tissues in 5 ml ice-cold PBS/0.3% Triton X-100 on sterile 10 ml Petri plates by applying gentle force with the rubber end of a sterile 10-ml syringe plunger. Other mechanical methods of tissue disruption that yield uniform cell suspensions can also be used.

22| Centrifuge the suspension for 5 min at 500*g* and 4 °C to remove tissue debris, unlyzed cells and host cell nuclei.

23| Sort bacteria from infected tissue homogenates with a high-speed FACS (see EQUIPMENT SETUP). Establish cell sorting parameters by comparing the light scattering and fluorescence properties of L-broth grown bacteria (see EQUIPMENT SETUP) with tissue lysates from mice infected with fluorescent and nonfluorescent *S. enterica*. In particular, establish the scatter and fluorescent properties of autofluorescent flavin containing cellular debris (autofluorescent debris) and GFP-expressing *S. enterica* (see **Fig. 3** for details) to set an appropriate green fluorescence threshold to distinguish bacteria from autofluorescent debris. Collect particles with FSC and SSC properties typical for *S. enterica* that have green-to-orange emission ratios typical for GFP (established with *in vitro* cultures of GFP-expressing *S. enterica*) but that do not overlap with the bulk of autofluorescent cellular debris. Trigger sorting events using green fluorescence as a threshold for detection. For improved signal-to-noise ratio in the scatter parameters, remove neutral density filters typically used for eukaryotic cells. To minimize sorting artifacts due to several *Salmonella* cells passing the laser beam simultaneously, adjust flow rate and/or suspension density to maintain total event rates of less than 20,000 particles per second. In the first sorting cycle, collect greater than 10^5 GFP-positive *S. enterica* directly into L-broth Amp to maximize library coverage.
? TROUBLESHOOTING

24| Plate sorted bacteria on L-agar Amp plates and incubate at 37 °C overnight. Pool colonies for cryogenesis as described above in Step 16, and store at −80 °C.
■ **PAUSE POINT** Sorted bacterial libraries can be stored frozen until needed for mouse re-infections.

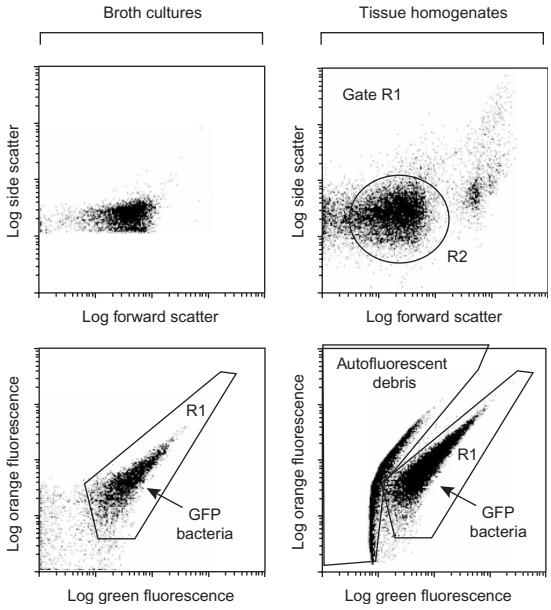

Figure 3 | Gating strategy for FACS of GFP-expressing *Salmonella* from infected mouse tissues. L-broth grown *Salmonella* expressing GFP were analyzed by forward scatter (FSC) and sideward scatter (SSC) (top left) and for fluorescence emission (bottom left) after excitation with a 488 nm laser. Acquired green (515–545 nm) and orange emissions (563–607 nm) in the FL-1 and FL-2 detectors were not compensated. These dot-plot profiles are used to establish appropriate gates (R1) for sorting particles with fluorescent properties similar to GFP (i.e., high 'green', low 'orange') from homogenized tissue lysates. The use of two-color cytometry discriminates against collecting autofluorescent flavin containing cellular debris (bottom right). The FSC and SSC properties of particles with GFP-like fluorescence (Gate R1) from homogenized tissue lysates are shown (top right). Only particles with GFP-like fluorescence properties and low light scattering properties (R2) similar to those observed for bacteria (top left) are sorted. Events with high FSC indicative of large particles or clumps passing the laser beam at the same time are disregarded.

25| The first round of selection is not stringent and thus permits the isolation and amplification of rare *in vivo*-activated promoter fusions. To remove nonproductive promoter fusions from this sorted bacterial population, perform a second round of infection or bacterial sorts by repeating Steps 17–23 with the first passage library obtained in Steps 23 and 24. This time sort approximately 10^4 bacteria with bright GFP-fluorescence to minimize contamination with nonfluorescent bacteria.

To ensure high purity of sorted particles, use flow rates less than 10^4 particles s^{-1} and high abort rates (approximately 25%) in the exclusion of potentially contaminated neighboring droplets. Sort GFP-expressing *Salmonella* directly into dH_2O/7% DMSO to cryopreserve aliquots without plating to minimize redundancy in the second sub-library. This host-induced promoter library can be subjected to further sorting cycles to further refine the type of promoters to be analyzed (**Box 1**).

26| Plate an aliquot of the second sub-library (from Step 25) containing approximately 100 GFP-positive bacteria on L Amp-agar plates and grow overnight at 37 °C.
? TROUBLESHOOTING

27| Next day, resuspend individual colonies in 50 µl dH_2O containing 7% (vol/vol) DMSO.
■ **PAUSE POINT** Suspensions can be stored in 96-well plate at –80 °C until needed.

28| Analyze a 0.5 µl aliquot of each colony suspension by PCR. Set up reactions as instructed in the table below. Run the PCR program detailed in EQUIPMENT SETUP.

Component	Final
Colony suspension	0.5 µl
PCR buffer	1×
$MgCl_2$	2 mM
pMW82-specific forward primer	5 pmol
pMW82-specific reverse primer	5 pmol
dNTPs	4 mM
Taq polymerase	1 U
dH_2O	to 25 µl

29| Determine the heterogeneity of DNA inserts by digesting PCR-amplified inserts with PCR buffer-compatible restriction enzymes that cut frequently (4-base cutters, e.g., AluI). Place half of the PCR in a separate tube, add 5 U of AluI and digest the PCR products for 1 h at 37 °C.

30| Assess restriction fragment length polymorphisms (RFLPs) by running the digested products on a 2% agarose gel. PCR products displaying distinct RFLP patterns can be sequenced directly after excising amplified products from an agarose gel using QIAquick gel purification kit and matched to *S. enterica* genomic sequences[18]. *S. enterica* bearing a unique *gfp* promoter fusion can be tested directly for GFP-expression in cultured macrophages or in infected mouse tissues to confirm expression during infection (e.g., **Fig. 3b**) as previously described[5,16].
? TROUBLESHOOTING

● TIMING

Media and reagent preparation: 1 d (can be prepared in advanced and stored at 4 °C)
Library construction: 2–4 weeks, depending on the use of pausepoints
Animal infections, killing, tissue harvesting and homogenization: Approximately 3–4 d
Cell sorting and analysis: Approximately 2 weeks, depending on the use of pausepoints
Sorted library analysis: Approximately 1–2 weeks, depending on the use of pausepoints
? TROUBLESHOOTING

Steps 1–13

While constructing the GFP promoter-trap libraries, assess library promoter heterogeneity at all stages of construction by monitoring under a fluorescence microscope the levels of GFP expression among colonies of *E. coli* and *S. enterica* grown on L-Amp agar plates. Colonies on the plate viewed under the lowest magnification should display a range of green fluorescent intensities when exposed to blue light. If no fluorescent or very few fluorescent colonies are observed, pick 10–15 random bacterial colonies and determine heterogeneity of inserts by amplifying DNA inserts by PCR followed by RFLP analysis as described in Steps 28–30. If too many plasmids do not contain inserts or the insert sizes are smaller than the 500–700 bp, repeat DNA shearing and ligation steps (Steps 2–9) and reassess heterogeneity of DNA inserts by PCR and RFLP analysis.

Step 23

Some 'sterile' sorter fluids contain autofluorescent bacterial contaminants such as *Pseudomonas* sp. that could overgrow sorted *Salmonella* during subsequent cultivation. To assess potential sterility problems, plate sorted sterile test beads on L-agar. If microbial contamination is a problem, consult with the FACS operator for proper de-contamination steps to take. Alternatively, for moderate to low contamination problems, bacterial growth media can be supplemented with 50 μg ml streptomycin as *S. enterica* SL1344 is naturally resistant to streptomycin.

Steps 23–26

Flow cytometers with cell-sorting capabilities are highly specialized machines that are often operated by a designated trained technician. Discuss the experimental setup and goals with the operator. As these machines are most often used to isolate mammalian cells, the operator may need to make adjustments to the cytometer to ensure proper sorting of bacterial-size particles. If necessary, run test bacterial sorts by pre-mixing GFP-tagged with untagged bacteria at a ratio of 1:100 and determine the sorting efficiency by immediately re-analyzing the sorted population using FACS and by plating sorted cells on L-Amp agar plates. If large discrepancies occur between the number of particles sorted and the recovery of viable of colony forming units, re-test the bacterial sorting parameters as described above and plate a sample of sorted L-broth grown bacteria. Compare the number of sorted particles with plating results. Avoid commercial FACS buffers containing antibiotics that are incompatible with live bacteria sorting.

? TROUBLESHOOTING

Further troubleshooting advice can be found in **Table 1**.

TABLE 1 | Troubleshooting differential fluorescence induction (DFI) selections.

Step	Problem	Possible reason	Solution
30	No enrichment of fluorescent bacteria bearing *in vivo* induced gene fusions	Improper library construction	Test self-ligation frequency of Calf-intestinal alkaline phosphatase (CIAP)-treated pMW82. If there is a high background of self-ligation (greater than 20% of Amp^R colonies), repeat BamHI digestion, and T4 DNA polymerases and CIAP treatments Assess library diversity by visual inspection of colonies in a plate under a fluorescence microscope and/or determine the heterogeneity of plasmids inserts by restriction digest analysis. Reconstruct promoter library if necessary
26		Faulty bacterial sorts	Run test bacterial sorts with L-broth grown cultures. Consult with fluorescence-activated cell sorting (FACS) operator for appropriate changes in flow rates and triggering thresholds
26		Large discrepancies between the number of viable bacteria and sorted particles	Adjust sorting gates to exclude autofluorescent particles (high 'orange' fluorescence) Test viability of sorted cells from L-broth grown cultures. Add osmostabilizing agents (e.g., 10% glycerol) to collection media

ANTICIPATED RESULTS

In a typical sorting experiment, approximately 1% of the input promoter library is recovered from infected spleens in the first round of cell sorting. By the second round of cell sorting, approximately 50% of the clones isolated in the first sort can be recovered from infected tissues. Most (greater than 95%) of the bacteria recovered at this stage bear promoters that are highly expressed in animal tissues but redundancy of promoters captured can be as high as 30% as determined by RFLP analysis (**Fig. 4**). Interestingly, work in *S. enterica* indicates that most of the promoters (approximately 80%) are induced at least threefold higher in animals than in laboratory growth media[11]. If an intermediate negative selection is included, as suggested in **Box 1**, most of the promoters captured will display host-specific transcriptional activity.

DFI analysis in *S. enterica* will reveal promoters for genes encoding proteins involved in stress responses, nutrient scavenging, lipopolysaccharide (LPS) modifications, and a large set of 'effector' proteins and their corresponding transporters (e.g., Type III secretion systems) that are required for *Salmonella* to manipulate the cell biology of their target cells and to down-regulate innate immune responses[11,20]. Interestingly, a large proportion of these identified genes are functionally relevant to infection because mutations in these genes often lead to attenuated virulence[11].

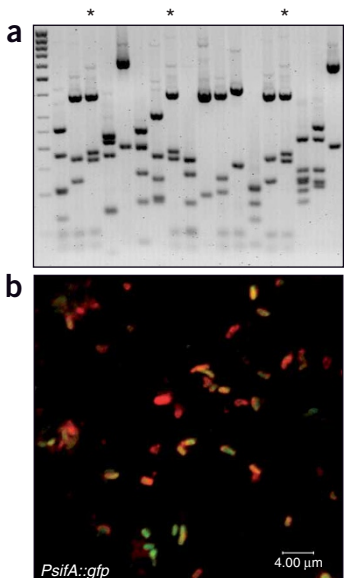

Figure 4 | Analysis of *in vivo* induced bacterial promoters. (**a**) Restriction fragment length polymorphisms (RFLP) analysis promoters captured in MW282 after two-stage cell sorting from infected spleens. Note that 14 independent DNA fragments with promoter activity were obtained from 18 clones. Asterisks indicate examples of DNA inserts with identical RFLP patterns. (**b**) Confocal image of *Salmonella enterica* expressing a *PsifA::gfp* fusion in infected murine Peyer's patches. *S. enterica* bearing *PsifA::gfp* fusions isolated as described above were used to infect mice orally. SifA is an effector protein that regulates membrane dynamics in *Salmonella*-infected cells[21]. Note heterogeneous expression levels. (GFP in green, anti-bacterial lipopolysaccharide (LPS) in red). LPS-positive bacteria without GFP may represent dead bacteria or bacteria that failed to express SifA.

ACKNOWLEDGMENTS We thank Meike Soerensen for excellent technical assistance in developing the procedure, and Claudia Rollenhagen and Nguyen Le Quang for help in preparing the figures. D.B. is supported by the Deutsche Forschungsgemeinschaft, and R.H.V. is supported by grants from the NIH/NIAD, the WhiteHead Foundation and the Pew Charitable Trust. We also acknowledge Stanley Falkow for intellectual input during the development of DFI.

COMPETING INTERESTS STATEMENT The authors declare competing financial interests (see the HTML version of this article for details).

Published online at http://www.natureprotocols.com
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Chiang, S.L., Mekalanos, J.J. & Holden, D.W. *In vivo* genetic analysis of bacterial virulence. *Annu. Rev. Microbiol.* **53**, 129–154 (1999).
2. Rediers, H., Rainey, P.B., Vanderleyden, J. & De Mot, R. Unraveling the secret lives of bacteria: use of *in vivo* expression technology and differential fluorescence induction promoter traps as tools for exploring niche-specific gene expression. *Microbiol. Mol. Biol. Rev.* **69**, 217–261 (2005).
3. Shelburne, S.A. & Musser, J.M. Virulence gene expression *in vivo*. *Curr. Opin. Microbiol.* **7**, 283–289 (2004).
4. Valdivia, R.H. & Ramakrishnan, L. Applications of gene fusions to green fluorescent protein and flow cytometry to the study of bacterial gene expression in host cells. *Methods Enzymol.* **326**, 47–73 (2000).
5. Bumann, D. Examination of *Salmonella* gene expression in an infected mammalian host using the green fluorescent protein and two-colour flow cytometry. *Mol. Microbiol.* **43**, 1269–1283 (2002).
6. Hinton, J.C., Hautefort, I., Eriksson, S., Thompson, A. & Rhen, M. Benefits and pitfalls of using microarrays to monitor bacterial gene expression during infection. *Curr. Opin. Microbiol.* **7**, 277–282 (2004).
7. Tsien, R.Y. The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509–544 (1998).
8. Valdivia, R.H. & Falkow, S. Flow cytometry and bacterial pathogenesis. *Curr. Opin. Microbiol.* **1**, 359–363 (1998).
9. Wendland, M. & Bumann, D. Optimization of GFP levels for analyzing *Salmonella* gene expression during an infection. *FEBS Lett.* **521**, 105–108 (2002).
10. Hautefort, I., Proenca, M.J. & Hinton, J.C. Single-copy green fluorescent protein gene fusions allow accurate measurement of *Salmonella* gene expression *in vitro* and during infection of mammalian cells. *Appl. Environ. Microbiol.* **69**, 7480–7491 (2003).
11. Rollenhagen, C., Sorensen, M., Rizos, K., Hurvitz, R. & Bumann, D. Antigen selection based on expression levels during infection facilitates vaccine development for an intracellular pathogen. *Proc. Natl. Acad. Sci. USA* **101**, 8739–8744 (2004).
12. Chalfie, M. & Kain, S. *Green Fluorescent Protein: Properties, Applications and Protocols,* 2nd Edn. (John Wiley & Sons Inc., New York, 1996).
13. Jones, B.D. & Falkow, S. Salmonellosis: host immune responses and bacterial virulence determinants. *Annu. Rev. Immunol.* **14**, 533–561 (1996).
14. Andersen, J.B. *et al.* New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl. Environ. Microbiol.* **64**, 2240–2246 (1998).
15. O'Callaghan, D. & Charbit, A. High efficiency transformation of *Salmonella typhimurium* and *Salmonella typhi* by electroporation. *Mol. Gen. Genet.* **223**, 156–158 (1990).
16. Valdivia, R.H. & Falkow, S. Bacterial genetics by flow cytometry: rapid isolation of *Salmonella typhimurium* acid-inducible promoters by differential fluorescence induction. *Mol. Microbiol.* **22**, 367–378 (1996).
17. Coligan, J. *et al. Current Protocols in Immunology.* (Wiley Interscience, Hoboken, NJ, 2006).
18. McClelland, M. *et al.* Complete genome sequence of *Salmonella enterica* serovar typhimurium LT2. *Nature* **413**, 852–856 (2001).
19. Becker, D. *et al.* Robust *Salmonella* metabolism limits possibilities for new antimicrobials. *Nature* **440**, 303–307 (2006).
20. Rollenhagen, C. & Bumann, D. *Salmonella enterica* highly expressed genes are disease specific. *Infect. Immun.* **74**, 1649–1660 (2006).
21. Boucrot, E., Henry, T., Borg, J.P., Gorvel, J.P. & Meresse, S. The intracellular fate of *Salmonella* depends on the recruitment of kinesin. *Science* **308**, 1174–1178 (2005).