INFECTION AND IMMUNITY, Apr. 2005, p. 2533–2540 Vol. 73, No. 4
0019-9567/05/$08.00+0 doi:10.1128/IAI.73.4.2533–2540.2005

Designer Arrays for Defined Mutant Analysis To Detect Genes Essential for Survival of *Mycobacterium tuberculosis* in Mouse Lungs

Gyanu Lamichhane, Sandeep Tyagi, and William R. Bishai*

Center for Tuberculosis Research, Department of Medicine, Johns Hopkins University, School of Medicine, Baltimore, Maryland

Received 8 September 2004/Returned for modification 4 November 2004/Accepted 22 November 2004

The mechanisms by which *Mycobacterium tuberculosis* elicits disease are complex, involving a large repertoire of bacterial genes that are required for in vivo growth and survival. To identify such genes, we utilized a high-throughput microarray detection method to rapidly screen hundreds of unique, genotypically defined transposon mutants for in vivo survival with a high degree of specificity and sensitivity. Thirty-one *M. tuberculosis* genes were found to be required for in vivo survival in mouse lungs. These genes are involved in a broad range of activities, including metabolism, cell wall functions, and regulation. Our screen included 11 of the 12 known members of the mycobacterial membrane protein (mmpL) family genes, and mutation of 6 of these genes—*mmpL4*, *mmpL5*, *mmpL7*, *mmpL8*, *mmpL10*, and *mmpL11*—severely compromised the ability of the mutants to multiply in mouse lungs. Most of the 31 genes are conserved in other pathogenic mycobacteria, including *M. leprae* and *M. bovis*, suggesting that a core of basic in vivo survival mechanisms may be highly conserved despite the divergent human pathology caused by members of the mycobacterial genus. Of the 31 genes reported here, 17 have not been previously described to be involved in in vivo growth and survival of *M. tuberculosis*.

Tuberculosis is a major global health problem (14) and yet the basic mechanisms by which *Mycobacterium tuberculosis*, in contrast to other related mycobacteria, successfully establishes infection, multiplies, and persists in mammalian lungs are not well understood. Identification of genes essential for growth and survival in vivo would provide insight into the pathogenesis of tuberculosis.

The genomic era has ushered in a number of innovative approaches for identifying genes associated with bacterial pathogenesis within the mammalian host. These have included in vivo expression technology and differential fluorescence induction (21, 31), which are based on identifying genes whose expression is upregulated upon entry of the bacterial pathogen into the mammalian host. Signature-tagged mutagenesis offered a novel, high-throughput, subtractive approach to compete pools of tagged bacterial mutants for survival within the mammalian host seeking mutants which dropped out during the infection process (15). Signature-tagged mutagenesis of *M. tuberculosis* was performed by using IS*1096*-based mutagenesis system (5, 10); however, these analyses were limited by the fact that the insertion sequence (IS) element used for mutagenesis harbored a predilection for insertion into AT-rich sequences, and the target sites were not defined, thus making it suboptimal for targeting a large number of genes. Recently, a highly random *Mariner*-derived transposon was developed for mutagenesis of mycobacterial genomes (1, 23). This transposon, *Himar1*, requires only the sequence TA for insertion and thus may insert in all but 16 of the 4,250 known open reading frames of *M. tuberculosis*. Since *Himar1* has a defined target sequence, it is possible to map all susceptible sites a priori. In addition, it has been demonstrated that this transposon inserts randomly in *M. tuberculosis* strains H37Rv (25) and CDC1551 (20). A procedure known as TraSH (transposon site hybridization) has been developed for the subtractive detection of attenuated transposon mutants based upon microarray detection. Screening large pools of undefined *Himar1* mutants, this technique has identified 194 genes of *M. tuberculosis* required for bacterial survival in mouse spleens (27).

We describe here a high-throughput, high-sensitivity approach for subtractive identification of mutants attenuated for survival in animal tissues which implements previously archived collections of defined *M. tuberculosis Himar1* transposon mutants. Using this method, known as DeADMAn (for designer arrays for defined mutant analysis), we rapidly screened 530 archived *M. tuberculosis Himar1* transposon mutants and identified 31 which are significantly attenuated for 7-week survival in mouse lungs. Seventeen of these genes have not been previously implicated in in vivo growth and survival of *M. tuberculosis.* The DeADMAn methodology is semiquantitative and, since it is based on an archived collection, offers the advantage of retaining each mutant for subsequent study.

MATERIALS AND METHODS

Growth and construction of pools of mutants. A clinical isolate of *M. tuberculosis*, CDC1551 (32) was used as the host strain, and a defined collection of *Himar1* transposon mutants previously archived was used in the present study (20). Each mutant was grown separately in Middlebrook 7H9 liquid medium (Difco) supplemented with 0.2% glycerol, 0.05% Tween 80, 10% ADC (0.5% bovine serum albumin, 0.2% dextrose, and 0.085% NaCl) at 37°C. Mutants that

* Corresponding author. Mailing address: Center for Tuberculosis Research, Department of Medicine, Division of Infectious Diseases, Johns Hopkins School of Medicine, 1503 E. Jefferson St., Rm. 112, Baltimore, MD 21231-1001. Phone: (410) 955-3507. Fax: (410) 614-8173. E-mail: wbishai@jhsph.edu.

harbored a transposon in the region upstream of the terminal 100 bp (or if the gene was < 500 bp in length, within the 5′ 80% of the gene) were considered for screening to decrease the likelihood of selecting distal mutations which may not disrupt gene product function. Six pools (A though F), each containing between 61 and 100 randomly selected mutants, and a pool of 18 deliberately chosen mutants (mmpLS, Table 2) were prepared by combining a pure culture of each mutant grown to approximately the same optical density (0.8 to 1.0). Only mutants whose in vitro growth was similar to that of wild type were used for the present study.

In vivo infection. Seven groups of BALB/c mice were infected with different pools via the tail vein with ~10^6 CFU/mouse. Three mice per group were sacrificed at days 1, 20, and 49 postinfection, and the survival of each mutant in the lungs, spleen, and liver was analyzed. An extra time point of 98 days postinfection was incorporated in the group infected with mmpLS pool. The lungs, spleens, and livers were homogenized in phosphate-buffered saline and plated on Middlebrook 7H10 solid medium (Difco) supplemented with 0.5% glycerol, 10% (vol/vol) oleic acid-albumin-dextrose-catalase (OADC; Becton Dickinson), 0.05% Tween 80, 50 μg of cycloheximide/ml, and 20 μg of kanamycin/ml. More than 10,000 colonies per organ were scraped and pooled, and genomic DNA was prepared as described previously (2).

Probe preparation. A total of 2 μg of genomic DNA prepared from pooled colonies from lung homogenates at each time point were combined, digested with AluI, and ligated to an adaptor, a blunt-ended double-stranded DNA made of the oligonucleotides AdTOP (5′-GTAATACGACTCACTATAGGGCACG CGTGGTCGACGGCCCGGGCTGGT) and AdBTM (5′ P-ACCAGCCCGGG C-NH_2). The junctions of transposon insertion sites were selectively PCR amplified by using a transposon-specific primer HimarSP1 (5′-ACCAATAGGCC GAAATCGGCAAAATCC) and an adaptor-specific primer AP1 (5′-GTAATA CGACTCACTATAGGGCAC) with 28 cycles of 94°C for 30 s, 55°C for 25 s, and 72°C for 45 s. The presence of the 3′ amino group in AdBTM allows only the de novo-synthesized DNA to act as a template for primer AP1 to anneal and extend (28). Next, nested PCR was done with primers HimarSP2 (5′-CCGAGATAGG GTGAGTGTTGTTCCAG) and AP1a (5′-ACTATAGGGCACGCGTGGTCG ACG) with 28 cycles. followed by incorporation of amino-allyl dUTP with the nested primers Himar7F (5′-GGGTCTAGAGACCGGGGACTTATC) and AP1a by using the same PCR cycles as described earlier. The PCR products were then coupled with fluorescent dyes Cy3 or Cy5 for 1 h at room temperature (Atlas Fluorescent Labeling Kit; Clontech). Probes prepared from the input pool (1 day postinfection) and output pools were cohybridized to custom microarrays, incubated, washed, and scanned by using Genepix Axon 4000B. Probe preparation and microarray hybridizations were repeated eight times, which included inverse labeling, for each sample pair.

Microarray analysis. Six different microarray sets, one for each pool and one for pools E and mmpLS, were printed on poly-L-lysine-coated glass slides. The microarrays contain spots of ca. 60 base oligonucleotides corresponding to the sequence at the junction of each transposon insertion. Each oligonucleotide, 50 μM in 3× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), was spotted four times in tandem as internal controls to ensure the consistency of the data. Positive controls consisting of equimolar mixture of each oligonucleotide, and four negative controls representing junctions of mutants not present in the pools were also spotted. The data were analyzed by using a custom developed program by using statistical R language.

RESULTS

DeADMAn offers specific, high-throughput detection of attenuated mutants. A prior study identified the site of *Himar1* transposon insertion in 1,425 mutants of *M. tuberculosis* (20). With a priori knowledge on the genotype of each mutant, we designed 60-base oligonucleotides specific for the genomic DNA adjacent to the distal end of the transposon in 530 of these mutants (representing mutations in 514 unique genes) and used the oligonucleotides to print custom microarrays to probe the mutants (Fig. 1). Randomly sampled mutants from the collection were grown individually in culture, and 530 mutants that grew at a similar rate were selected (a step which ensured equivalent in vitro growth rates of mutants) and combined into pools of 100 or fewer mutants per pool. To test the sensitivity of our arrays, chromosomal DNA was isolated from

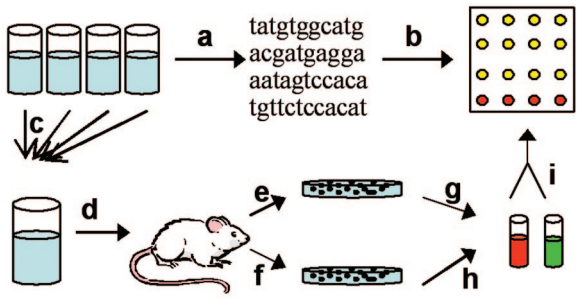

FIG. 1. Schematics of DeADMAn. (Step a) Mutants are grown individually in liquid culture, and the sequence of the transposon insertion junction is determined. (Step b) 60-mer oligonucleotides corresponding to insertion junction sequence of each mutant are printed on a custom microarray. (Step c) The mutants are pooled. (Step d) Mice are infected with pooled mutants to determine in vivo growth. (Steps e and f) The input pool is obtained by harvesting lungs and spleen 1 day postinfection (e) and the output pool is obtained by harvesting lungs and spleen at 20 and 49 days postinfection (f). (Steps g and h) Probes are prepared from the input pool sample (g) and probes are prepared from the output pool sample (h). (Step i) The input and output probes, which are labeled with Cy5 (red) and Cy3 (green) dyes, are cohybridized to custom microarray, and the relative levels of probes corresponding to each mutant are determined. A mutant attenuated for in vivo growth would be absent or underrepresented in the output pool, and this phenotype would be revealed by overabundance of the Cy5-labeled input probe producing red spots. Spots that appear yellow-orange result from the presence and cohybridization of both the input (Cy5) and output (Cy3) probes.

the newly mixed pools, and PCR with a primer specific for the distal end of the *Himar1* transposon was used to generate probes spanning the transposon insertion sites of mutants in the pool. The junctional PCR products were fluorescently labeled and hybridized to the designer oligonucleotide array specific for the pool. These preliminary evaluations revealed that in all instances 100% of the member mutants of each pool were detected by microarray; moreover, negative control spots (60-mers specific for mutants not present in the pools) and positive control spots (a mixture of all 60-mers printed on the array) routinely gave negative and positive signals, respectively, as expected.

Based upon the sensitivity and specificity of DeADMAn in vitro, we proceeded to test the technique in vivo with BALB/c mice. Mice were infected intravenously with ~10^6 CFU of *M. tuberculosis* pools with complexities of 100 or fewer unique mutants. To determine the composition of the surviving members of the mutant pool after in vivo growth, pools were assessed by harvesting lungs, spleens, and livers from groups of mice on days 1, 20, and 49 after challenge. Organ homogenates from these animals were plated on solid medium, and no fewer than 10,000 colonies from these pools were collected and used for the amplification of junctional PCR products representing the various pools. In all instances, junctional PCR products from the in vitro grown pool and from the day 1 lung, spleen, and liver homogenates showed nearly 100% detection of member mutants by DeADMAn, thus corroborating our in vitro sensitivity and specificity data. Mutants that were impaired for in vivo growth or survival were readily detected by the abundance of probe from the input compared to the output time point (Fig. 2), indicating the quantitative nature of the tech-

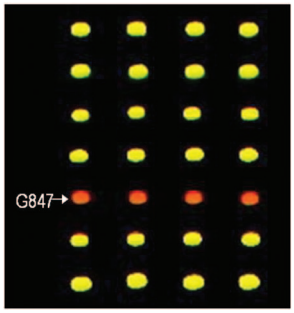

FIG. 2. Growth-attenuated mutants are readily detected by DeADMAn. A subsection of the array for pool C is shown. Probe from the input pool at day 1 postinfection was labeled with Cy5 (red), and the probe from the output pool at day 49 postinfection was labeled with Cy3 (green). The mutant, G847, which has a mutation in the gene *bioB*, is underrepresented in the output pool.

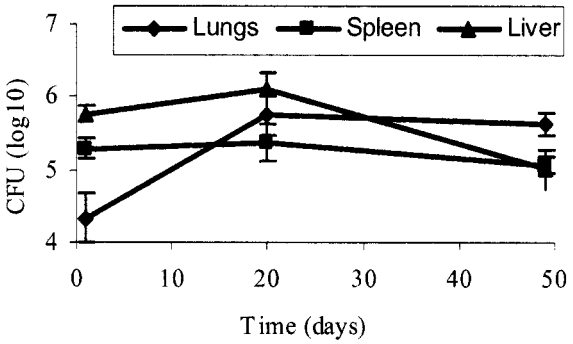

FIG. 3. Growth in vivo of pools of mutants is similar to standard growth of wild-type *M. tuberculosis*. Mice were infected with a pool of 100 different mutants, i.e., pool A. Lungs, spleens, and liver were obtained at days 1, 20, and 49 postinfection, and the number of viable bacilli quantified by plating on solid medium.

nique. Although probes were generated by PCR of the unique junctions of transposon insertion sites, DNA from both input and output pools were amplified under identical conditions to minimize PCR-related artifacts. Indeed, not all pooled amplifications reflected actual quantitative differences between strains when DeADMAn results were analyzed independently by testing with mutant-specific PCR amplifications.. As presented in supplementary data (www.hopkinsmedicine.org/TARGET/GenTools.htm), while 37 mutants gave low output/input ratio, 5 (13.5%) were false positive since the corresponding mutants could be readily detected when tested individually by specific PCR. However, it is important to note that most of the quantitative differences observed in DeADMAn indeed reflect corresponding difference in population size of different mutants in the lungs of mice. We tested in vivo growth kinetics of mutant H28, in which the rel_{Mtb} gene is disrupted by transposon insertion at position +322, in a granuloma model of tuberculosis to test our DeADMAn results that are based on competitive assay in a mixed pool. The mutant was severely attenuated for growth in vivo by 21 days postinfection, whereas there was no difference in growth in vitro compared to the wild type (18).

A small proportion of in vitro nonessential genes are essential for in vivo growth. As assessed by CFU counts in the organs, the mutant pools behaved the same as wild-type *M. tuberculosis*, displaying normal growth with initial rapid multiplication, followed by plateau in mouse lungs, but in contrast showing declining counts in the spleen and the liver (Fig. 3). The 530 unique mutants were grouped in seven different pools and screened for in vivo survival in mouse lungs for 7 weeks by DeADMAn. Of the 530 mutants evaluated (see Table 3 at www.hopkinsmedicine.org/TARGET/GenTools.htm), 32 mutants with transposon insertions in 31 unique genes resulted in in vivo growth attenuation at the 7-week time interval in mouse lungs (Tables 1 and 2). This result suggests ca. 6% of the in vitro nonessential genes to be differentially required for growth in vivo in mouse lungs by *M. tuberculosis* CDC1551. Similar results were obtained for growth in mouse spleens by *M. tuberculosis* H37Rv (27). A total of 94% of the mutants did not show a significant in vivo survival defect. Therefore, these genes are not essential for 7-week in vivo survival in mouse lungs.

The identification of 32 mutants attenuated in vivo in mouse lungs was independently verified by specific PCR amplification of transposon site junctions from mutants that showed growth attenuation by DeADMAn (Fig. 4A and B). A negative control, one of the mutants that was not defective for in vivo survival, was included for each PCR amplification. With exception of mutants H274, H292, H445, H844, and J41, all of the mutants identified by DeADMAn showed concordance with the PCR detection procedure.

A wide range of gene activities are required for growth and multiplication of *M. tuberculosis*. Most of the genes required for 7-week in vivo survival in mouse lungs have reported functions. Of the 31 genes described in Tables 1 and 2, 11 are involved in metabolism (functional groups 1 and 7), 11 in cell wall processes, 2 in regulation, and 1 in adaptation. In addition, five are annotated as conserved hypotheticals and one is annotated as hypothetical since they do not show significant homologies to proteins with known function.

DeADMAn analysis indicates that the rel_{Mtb} gene is required for short-term survival in mouse tissues. Mutant H28, one of the 100 mutants that comprises pool A, with a disruption in the rel_{Mtb} (MT2660/Rv2583c) gene, is significantly attenuated for lung in vivo growth at the 7-week time point (see Fig. 4). rel_{Mtb} encodes GTP pyrophosphokinase which catalyzes the reversible synthesis of (p)ppGpp (3). In *E. coli*, there are two homologues of the *M. tuberculosis rel* gene (rel_{Mtb}), *relA*, and *spoT* that code for a (p)ppGpp synthase and pyrophosphatase, respectively (6). Since the (p)ppGpp molecule is synthesized by RelA in response to low amino acid concentration and elevated levels of uncharged tRNA, conditions characteristic of nutrient starvation, RelA is thought to mediate the cell's response to nutrient deprivation. Recently, an *M. tuberculosis* H37Rv rel_{Mtb} mutant was shown to have a defect in long-term persistence in C57BL/6 mouse lungs after aerosol infection with nearly normal tissue CFU counts at the 5-week time point but a 2-log decline in lung CFU counts at 38 weeks (11). Our results show that when probed at 49 days after intravenous infection of BALB/c mice, the rel_{Mtb} mutant had failed to proliferate and was difficult to detect. Similar observations were made in a granuloma model of tuberculosis in which the rel_{Mtb} mutant was encapsulated within semidiffusible hollow fibers placed in the subcutaneous space of mice (18). These observations in our DeADMAn system suggest that, in addition to

TABLE 1. Mutants that were defective for growth in mouse lungs, sorted in order of gene number

ID[a]	Size (bp)[b]	POI[c]	Gene description[f]	Annotation no.[d]			Identification source or reference[e]
				MT#	Rv#	ML#	
J34	870	373	Oxidoreductase, putative[7]	0106	0097	1992	This study
H496	7,539	364	Peptide synthase (*nrp*)[1]	0110	0101	1996	27; this study
H491	7,539	5881	Peptide synthase (*nrp*)[1]	0110	0101	1996	27; this study
H615	1,023	334	Putative esterase, antigen 85c (*fbpc*)[1]	0137	0129c	2655	This study
SM1	1,239	105	Conserved hypothetical protein[3]	0214	0204c	2618	5, 27; this study
R243	1,512	1221	Hypothetical protein[10]	0350	0336		This study
S49	705	256	Transcriptional regulator, tetR family[9]	0489	0472c	2457	This study
H790	1,317	180	NagC-related protein[9]	0503	0485	2444	This study
H619	1,080	683	ABC transporter (*mkl*)[3]	0684	0655	1892	27; this study
S48	174	43	Membrane protein[3]	0694	0666		This study
R287	1,683	515	Indolepyruvate decarboxylase[7]	0876	0853c	2167	This study
G274	1,182	839	Sugar ABC transporter (*sugC*)[3]	1276	1238	1089	27; this study
G847	1,050	326	Biotin synthase (*bioB*)[7]	1624	1589	1220	27, this study
R252	2,175	1612	4-α-glucanotransferase (*malQ*)[7]	1831	1781c	1545	This study
H587	1,833	331	CbxX/CfgX protein[10]	1847	1798	1536	This study
G723	1,245	745	L-lactate dehydrogenase (*lldD2*)[7]	1921	1872c	2046	This study
G335	531	350	Conserved hypothetical protein[10]	2198	2140c	1289	This study
S63	1,563	862	Proteinase, putative[3]	2282	2224c	1633	13, 27; this study
G504	477	352	Conserved hypothetical protein[10]	2299	2239c	1649	This study
SM3	1,047	898	Hypothetical protein[10]	2496	2423		This study
H28	2,373	322	rel_{Mtb}[7]	2660	2583c	0491	11; this study
G566	1,752	954	Acyl-coenzyme A synthase (*fadD26*)[1]	2999	2930	2358	5; this study
H741	696	291	Hypothetical protein[10]	3306	3210c	0812	27; this study
G522	894	546	Thiosulfate sulfurtransferase (*sseA*)[7]	3382	3283	0728	This study
R237	1,620	601	Chaperonin 69 kDa (*groEL1*)[0]	3526	3417c	0381	This study
G534	420	329	Aspartate-1-decarboxylase (*panD*)[7]	3706.1	3601c	0231	24; this study

[a] ID, identification labels for mutants.
[b] The sizes of the genes described.
[c] POI, point of insertion. This refers to the nucleotide location, from the translation start site, where the transposon has been inserted.
[d] MT#, a unique gene annotation number assigned for each gene by the TIGR database for *M. tuberculosis* CDC1551; Rv#, a unique gene annotation number assigned for each gene by the Sanger database for *M. tuberculosis* H37Rv; ML#, a unique gene annotation number assigned for each gene by the Sanger database for *M. leprae*.
[e] That is, the source reporting the gene to be involved in attenuation.
[f] The superscript numbers in this column indicate the functional classifications of genes by the Sanger database: 0, adaptation; 1, lipid metabolism; 3, cell wall process; 7, intermediary metabolism and respiration; 9, regulator; 10, conserved hypothetical proteins.

having a role in long-term persistence, the rel_{Mtb} gene may also be important for short-term survival in certain in vivo situations.

Similarly, mutation of the gene *panD*, which codes for aspartate-1-decarboxylase, produced a strain that was attenuated in our screen. A previous report also demonstrated that the mutant lacking genes *panC* and *panD* was highly defective for in vivo growth (24). Thus, in addition to identifying genes not previously reported to be associated with pathogenesis, the DeADMAn approach readily identified two genes, rel_{Mtb} and *panD*, which have previously been shown to be required for full virulence of *M. tuberculosis*.

Many members of the mycobacterial membrane large and small protein family (mmpL/S) are important for growth in vivo. A family of genes annotated as mycobacterial membrane protein cluster consists of 17 members; 12 genes that encode large (*mmpL1* through *mmpL12*) proteins and 5 that encode small ones (*mmpS1* through *mmpS5*) (7). These *mmpL* genes belong to the resistance-nodulation-division protein superfamily (30), but several domains in the MmpL proteins are unique to mycobacteria. Proteins of the resistance-nodulation-division family are involved in the transport of a range of substrates across the membrane by using the transmembrane proton gradient. The *mmpL* genes, except for *mmpL6*, are ca. 3 kb in size and encode proteins that have between 10 and 13 transmembrane regions, as predicted by TMPRED (http://www.ch.embnet.org/software/TMPRED_form.html). Although it is clear that the *mmpL* genes are not essential for growth in vitro, the same is not true for survival in vivo in the mouse. MmpL7 is involved in the transport of phthiocerol dimycocerosates, and a mutant lacking this gene activity was severely attenuated for growth in the lungs but not in the spleens and liver (4, 10). MmpL8 is involved in biosynthesis of sulfolipid-1 and may be required for growth and persistence in vivo (9, 12). In addition, in vivo growth of *mmpL2*, *mmpL4*, and *mmpL7* mutants were observed to be compromised when signature-tagged transposon mutants of *M. tuberculosis* Mt103 were screened (5). The effect of mutations in other *mmpL* and *mmpS* genes for in vivo survival is not known. We tested a pool of 18 mutants with mutations in *mmpL* and *mmpS* genes, which includes all family members except for *mmpL3* and *mmpS1*, *mmpS4*, and *mmpS5* (Table 2). In addition, an attenuated mutant, H28 (the rel_{Mtb} mutant), was included as a positive control, and a nonattenuated mutant, H289, with a disruption in a conserved hypothetical gene (MT2749/Rv2675c), was included as a negative control. This pool of mmpL/S mutants showed normal in vivo growth in the lungs, spleen, and liver over a time course of 98 days postinfection (data not shown). Mutants with disruptions in *mmpL4*, *mmpL5*, *mmpL7*, *mmpL8*, *mmpL10*, and *mmpL11* showed significant attenuation for growth in the lungs, whereas the *mmpL1*, *mmpL2*, *mmpL6*, *mmpL9*, and *mmpL12* mutants

TABLE 2. Mutants that comprise pool mmpL/S

ID[a]	Size (bp)[b]	POI[c]	Gene description	Annotation no.[d]			Attenuation[f]	Identification source or reference[e]
				MT#	Rv#	ML#		
G532	2,877	1222	mmpL1	0412	0402c		N	
E168	2,907	458	mmpL2	0528	0507		N	
H446	2,907	2203	mmpL2	0528	0507		N	
G168	2,904	1217	mmpL4	0466	0450c	2378	Y	5; this study
R294	2,895	899	mmpL5	0705	0676c		Y	This study
G868	2,895	1773	mmpL5	0705	0676c		N	
H293	1,194	921	mmpL6	1608	1557		N	
G263	2,763	630	mmpL7	3012	2942	0137	Y	5, 10, 27; this study
R285	3,270	53	mmpL8	3931	3823c		Y	9; this study
G476	2,889	29	mmpL9	2402	2339		N	
R61	3,009	2828	mmpL10	1220	1183	1231	Y	27; this study
G46	2,901	560	mmpL11	0212	0202c	2617	Y	This study
G107	3,441	349	mmpL12	1573	1522c		N	
H353	3,441	772	mmpL12	1573	1522		N	
E262	444	80	mmpS2	0527	0506		N	
L6	900	402	mmpS3	2254	2198	0877	NA	
H28	2,373	322	rel_{Mtb}	2660	2583c	0491	Y	
H289	753	565	Conserved hypothetical protein	2749	2675c	1046	N	

[a] ID, identification labels for mutants.
[b] The sizes of the genes described.
[c] POI, point of insertion. This refers the to nucleotide location, from the translation start site, where the transposon has been inserted.
[d] MT#, a unique gene annotation number assigned for each gene by the TIGR database for *M. tuberculosis* CDC1551; Rv#, a unique gene annotation number assigned for each gene by the Sanger database for *M. tuberculosis* H37Rv; ML#, a unique gene annotation number assigned for each gene by the Sanger database for *M. leprae*.
[e] That is, the source reporting the gene to be involved in attenuation.
[f] Y, attenuated; N, not attenuated; NA, data not available.

did not show an in vivo survival defect (Table 2 and Fig. 4). Only the *mmpL4* mutant showed a survival defect in the spleen, whereas other mutants were not affected in this organ. Interestingly, the degree of in vivo growth attenuation due to mutations in these is not the same. As may be seen in Fig. 4C, the *mmpL4* and *mmpL10* mutants were highly attenuated compared to the *mmpL7* and *mmpL11* mutants. Some of the *mmpL* mutants were included in pools other than mmpL/S pool (see data at www.hopkinsmedicine.org/TARGET/GenTools.htm) as controls for variation in pool complexity to test for their in vivo growth kinetics. Both mutants, G815 (from pool D) and R285 (from pool mmpL/S), with transposon insertion in *mmpL8* at positions +68 and +53, respectively, showed attenuation. Only mutant R285 is reported in Table 2, since the table includes information for all of the mutants that comprise the mmpL/S pool only. The growth data of mutant G815 is available online(see supplementary table at www.hopkinsmedicine.org/TARGET/GenTools.htm) and is not in Table 1. Mutants R294 and G868 have transposon insertion in the same gene, *mmpL5*, which is 2,895 bp long, at positions +899 and +1773 from the translation start site (Table 2). Only R294 showed growth attenuation, suggesting that the disruption was 5′ enough to interfere with the gene activity. The region between +899 and +1773, perhaps, encodes the essential activity since mutations at the two sites produce different outcomes. The segment encoded by the region located 3′ to position +1773 may be dispensable to the activity of mmpL5 protein, but this hypothesis remains to be tested. Interestingly, disruption in the 3′ region of the gene *mmpL10* (+2828) still abolishes its activity since this mutant showed severe in vivo growth defect. This observation suggests that the C terminus of *mmpL10* is essential for its activity.

Most of the 31 in vivo attenuating mutations occur in genes conserved in other species of mycobacteria. All of the *mmpL* genes for which homologues exist in *M. leprae* are required for 7-week in vivo survival by *M. tuberculosis* in mice (Table 2). None of these four *mmpL* homologues are pseudogenes in *M. leprae* despite the fact that 27% of the genes of this organism appear to be decayed, nonfunctional gene remnants (8). *M. bovis* possesses homologues of not only the six *mmpL* genes that are important for *M. tuberculosis* mouse survival, but also for all of the other *mmpL* and *mmpS* genes present in *M. tuberculosis*. Of the 25 non-*mmpL*/*S* genes that are required for *M. tuberculosis* in vivo survival, 22 have homologues in *M. leprae*, a disproportionately high fraction since the *M. leprae* genome possesses homologues to only ca. 40% of the *M. tuberculosis* genetic complement (8). Of these 22 *M. leprae* homologues, 3 (*ML2444*, *ML1545*, and *ML0812*) are predicted to be transcriptionally silent pseudogenes. Although *M. leprae* is different from *M. tuberculosis* in its preferred growth niche, doubling time, and pathogenesis, it appears that the two species have common requirements for in vivo survival since the majority of the genes required for mouse survival of *M. tuberculosis* are also conserved in *M. leprae*. However, *M. bovis* possesses homologues of all of the aforementioned *M. tuberculosis* genes that are important for growth in vivo. This observation is concordant with the fact that both species of mycobacteria share the same growth niche.

DISCUSSION

The genome sequence and development of numerous genetic tools for mutating *M. tuberculosis* has made the develop-

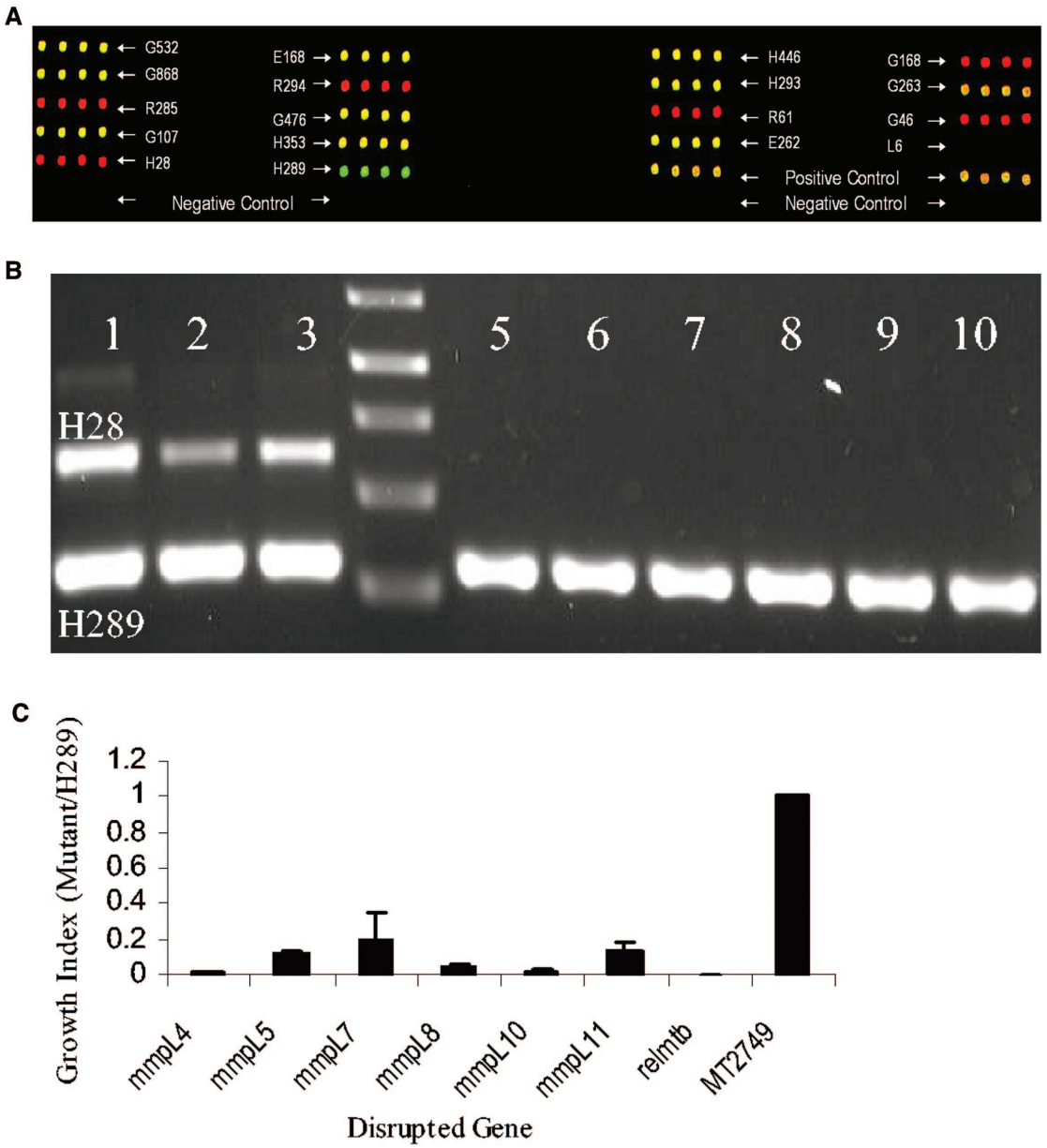

FIG. 4. (A) Microarray with input probes labeled with Cy5 (red) and output probes labeled with Cy3 (green) that were prepared by using colonies obtained from the lungs of mice. Each oligonucleotide represents a junction sequence for a mutant and is printed four times. The mutants that were underrepresented in the output pool (attenuated for in vivo survival) have proportionately more Cy5 probe and produce red spots on the array. There are four controls on this array: (i) H28 is a mutant from a different pool, pool A, that was observed to be attenuated and was used in this pool as a control; (ii) H289 is another mutant from pool A that was not attenuated; (iii) spots labeled as positive controls contain a mixture of all of the oligonucleotides printed on the arrays and thus is an internal positive control; and (iv) spots labeled as negative controls are oligonucleotides representing four different mutants from a different pool and thus serve as internal controls for background noise. (B) Agarose gel electrophoresis showing PCR amplicons of junctions of H28 and H289 (control). Lanes 1, 2, and 3 represent mutants recovered from lungs of three mice sacrificed at 1 day postinfection (input pool). Similarly, lanes 5, 6 and 7, and 8, 9, and 10 represent mutants recovered from three mice sacrificed at 49 and 98 days postinfection, respectively. Lane 4 is 1-kb plus marker (Invitrogen). (C) Six *mmpL* mutants and a control, H28, are compared against H289 for growth in vivo. Mutants H28 and H289 were used as controls based on their phenotypes from an earlier experiment. H28 is one of the 100 mutants of pool A and was observed to be attenuated. H289 is another mutant from pool A that was not attenuated for growth in vivo. Growth index refers to the population of each mutant, as measured from signals from the DeADMAn arrays, in the lungs of mice relative to the population of the positive control strain H289. The errors bars represent the standard deviation.

ment of a comprehensive set of mutants of nonessential genes a realistic goal. Single mutant animal infections to detect virulence defects, although a commonly used tool, is labor, time, and resource intensive and is unlikely to comprise a practical approach for evaluation of the ca. 4,000 genes of *M. tuberculosis*. We sought a way to increase the throughput of mutants tested for animal survival in a manner that was sensitive, specific, and exploited the availability of archived mutant collections.

In previous studies, a procedure called TraSH (transposon site hybridization) was developed for subtractive detection *Himar1* transposon insertion mutants of *M. tuberculosis* attenuated for in vivo growth in mouse spleens (26, 27). Using pools of undefined *Himar1* mutants, TraSH identified 194 genes required for *M. tuberculosis* survival in mouse spleens at 8 weeks. With the TraSH procedure, whole genome spotted microarrays with low representational density (spotted target sequence per gene $\ll$ length of the gene) are used to detect the dropout of attenuated mutants from complex input pools comprising tens of thousands of mutants. Because the arrayed DNA is not fully representational of the full length of each gene, only a fraction of the junctional tags from transposon mutants in the pool share complementarity with the array, leading to incomplete detection specificity and sensitivity. Moreover, the TraSH approach is nonarchival, that is, mutants of interest are part of a complex, mixed pool and must be isolated or recreated for further study. Although it requires an initial investment in the characterization and deposition of defined transposon mutants, the DeADMAn approach offers specific information on the in vivo survival phenotype of each evaluated mutant, is semiquantitative, and retains each mutant for subsequent study.

Using the mouse model of tuberculosis, only a limited number of bacilli, between 10^4 and 10^5 at most, may be delivered to the lungs via tail vein inoculation. Infection via the aerosol route would closely mimic the natural mode of infection but it is not suitable to consistently deliver the large number of different mutants required for high-throughput screening since only a few thousand bacilli can be deposited in the lungs of mice by this method. Although higher inocula may be delivered to the spleen, this organ is not the natural site of tuberculosis infection and may be less relevant to the pathogenesis of the disease. In addition, the requirements for virulence in the two organs are not the same; for example, the *mmpL7* gene is required for growth only in the lungs and not in the spleen or the liver (10). As discussed in this report, DeADMAn revealed that only the *mmpL4* mutant showed growth defect in the spleen, whereas six different *mmpL* mutants were attenuated in the lungs (data not shown).

In view of the numeric limitations on the number of bacilli that may be reproducibly deposited in the mouse lungs, we chose pool complexities of $\leq$100 mutants. From a pilot study, a pool of $\leq$100 mutants appeared to be an optimal compromise between the number of different mutants and the inoculum of each mutant that can be delivered to the lungs on the one hand and the specificity of detecting them on the other. Hensel et al. working with *Salmonella enterica* serovar Typhimurium also observed that a pool containing 192 unique mutants did not give reproducible hybridization results, whereas a pool with 96 mutants did (15). The custom arrays used in the present study were designed specifically for identification and quantification of the mutants. The one-oligonucleotide–one-mutant design of DeADMAn greatly increased specificity compared to the whole-genome arrays and reduced the number of false-positive results.

The mutants that were observed to have in vivo survival defects in the present study have in vitro growth rates similar to those of the parent wild-type strain when cultured in Middlebrook 7H9 and 7H10. Inactivation of some genes showed severe in vivo survival defects which were detectable at 20 days, whereas other genes contributed moderately to in vivo survival, requiring 49 days of infection to discern survival attenuation for those mutants. The data obtained at 20 days postinfection were similar for most mutants and thus not as decisive as that obtained at 49 days (data not shown). The study with mmpL/S pool was extended to 98 days to test the observations made at 49 days postinfection. As discussed above, the rel_{Mtb} mutant showed similar phenotypes at both time points. Similar observations were made for other mutants of the pool. Thus, 49 days postinfection is an appropriate time to screen for attenuated mutants.

The DeADMAn approach may miss mutants defective for in vivo survival that may be rescued by extracellular complementation. Mutants defective in the production of secreted factors, such as siderophores or autoinducers, for example, which may be essential for in vivo survival, might conceivably be rescued by other members of the pool proficient in synthesis of such factors. Another limitation of the DeADMAn approach is that it would be unlikely to detect mutants defective in eliciting tissue pathology though still able to proliferate in vivo. These mutants, known as immunopathology mutants, generate a bacterial burden similar to that of wild type but show diminished pathology in the mouse model of tuberculosis (16). Notably, inactivation of the putative transcriptional regulator *whiB3* and sigma factor H, *sigH*, genes results in such phenotypes (19, 29).

A large number of *M. tuberculosis* genes are arranged in operons and inactivation of the upstream genes may lead to a polar effect if the polycistronic messages are translationally coupled. The genes identified in our study may be in operons, and the in vivo growth defect observed could be due to polar effects on the gene(s) downstream of the mutated gene. In silico analysis showed 11 of the genes identified in Table 1 are likely to be arranged in operons. The genes MT0110 (Rv0101), MT0684 (Rv0655), MT0694 (Rv0666), MT1276 (Rv1238), MT1847 (Rv1798), and MT3706.1 (Rv3601c) seem to be arranged in operons but are located at the distal end. Therefore, disruption of these genes is unlikely to have polar effects. The gene MT0106 (Rv0097) appears to be in an operon consisting of six genes: MT0105 (Rv0096) through MT0110 (Rv0101). Thus, mutation in MT0106 could lead to a polar effect, and a loss of function of the downstream genes may be incorrectly attributed to this gene. Indeed, mutations in MT0110 were observed to result in growth attenuation. Similarly, the genes MT1624 (Rv1589), MT2299 (Rv2239c), MT2999 (Rv2930), and MT3382 (Rv3283) are likely arranged in operons and have at least one gene downstream of their location. Indeed, MT2999 (Rv2930) is part of an operon with 10 genes in *M. tuberculosis* Mt103 (4). Inactivation of the aforementioned genes may have polar effects, in which case these genes would have indirectly contributed to the observed in vivo growth attenuation.

The 31 genes identified in the present study are required for in vivo survival but are dispensable for in vitro growth. Some of the genes reported here, namely, *fadD26*, *mmpL4*, *mmpL7*, *mmpL8*, *panD*, and rel_{Mtb}, have been previously shown to be required for growth in the lungs of mice (5, 9, 11, 24). A subset of the genes on our list were also reported to be necessary for growth in the spleen of mice by TraSH (Tables 1 and 2), whereas some of the genes proposed to lead to in vivo atten-

uation did not show growth phenotype in our analysis (27). Other genes, such as *leuD* and *icl*, that were not studied here have been implicated in the growth and survival of *M. tuberculosis* (16, 17, 22).

The DeADMAn subtractive mutant competition assay is a powerful new tool for high-throughput, yet specific identification of bacterial genes required for in vivo survival in animal tissues. Significant advantages of the technique are that it is archival permitting immediate access to mutants of interest and that it is semiquantitative, allowing mutants of various degrees of attenuation to be identified during the initial screen. Tools of this nature may permit rapid, comprehensive determination of the animal phenotypes of virtually all mutatable genes in a bacterial genome.

ACKNOWLEDGMENTS

We thank Naomi Gauchet for administrative assistance; S. Woolwine for suggesting the acronym DeADMAn; S. Woolwine, R. Tirumalai, and N. Morrison for critical review of the manuscript; Hai Xu for microarray printing; and all members of the Bishai lab for constructive discussions.

This study was supported by National Institutes of Health grants AI36973, AI37856, AI43846, and AI51668 and contract N01-30036.

REFERENCES

1. **Akerley, B. J., E. J. Rubin, A. Camilli, D. J. Lampe, H. M. Robertson, and J. J. Mekalanos.** 1998. Systematic identification of essential genes by in vitro mariner mutagenesis. Proc. Natl. Acad. Sci. USA **95:**8927–8932.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1998. Current protocols and molecular biology. John Wiley & Sons, New York, N.Y.
3. **Avarbock, D., J. Salem, L. S. Li, Z. M. Wang, and H. Rubin.** 1999. Cloning and characterization of a bifunctional RelA/SpoT homologue from *Mycobacterium tuberculosis*. Gene **233:**261–269.
4. **Camacho, L. R., P. Constant, C. Raynaud, M. A. Laneelle, J. A. Triccas, B. Gicquel, M. Daffe, and C. Guilhot.** 2001. Analysis of the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*: evidence that this lipid is involved in the cell wall permeability barrier. J. Biol. Chem. **276:**19845–19854.
5. **Camacho, L. R., D. Ensergueix, E. Perez, B. Gicquel, and C. Guilhot.** 1999. Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. Mol. Microbiol. **34:**257–267.
6. **Cashel, M., D. R. Gentry, V. H. Hernandez, and D. Vinella.** 1996. The stringent response, p. 1458–1496. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. American Society Microbiology, Washington, D.C.
7. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, B. G. Barrell, et al.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537–544.
8. **Cole, S. T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honore, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell.** 2001. Massive gene decay in the leprosy bacillus. Nature **409:**1007–1011.
9. **Converse, S. E., J. D. Mougous, M. D. Leavell, J. A. Leary, C. R. Bertozzi, and J. S. Cox.** 2003. MmpL8 is required for sulfolipid-1 biosynthesis and *Mycobacterium tuberculosis* virulence. Proc. Natl. Acad. Sci. USA **100:**6121–6126.
10. **Cox, J. S., B. Chen, M. McNeil, and W. R. Jacobs, Jr.** 1999. Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. Nature **402:**79–83.
11. **Dahl, J. L., C. N. Kraus, H. I. Boshoff, B. Doan, K. Foley, D. Avarbock, G. Kaplan, V. Mizrahi, H. Rubin, and C. E. Barry III.** 2003. The role of RelMtb-mediated adaptation to stationary phase in long-term persistence of *Mycobacterium tuberculosis* in mice. Proc. Natl. Acad. Sci. USA **100:**10026–10031.
12. **Domenech, P., M. B. Reed, C. S. Dowd, C. Manca, G. Kaplan, and C. E. Barry III.** 2004. The role of MmpL8 in sulfatide biogenesis and virulence of *mycobacterium tuberculosis*. J. Biol. Chem. **279:**21257–21265.
13. **Dubnau, E., P. Fontan, R. Manganelli, S. Soares-Appel, and I. Smith.** 2002. *Mycobacterium tuberculosis* genes induced during infection of human macrophages. Infect. Immun. **70:**2787–2795.
14. **Dye, C., S. Scheele, P. Dolin, V. Pathania, M. C. Raviglione, et al.** 1999. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. JAMA **282:**677–686.
15. **Hensel, M., J. E. Shea, C. Gleeson, M. D. Jones, E. Dalton, and D. W. Holden.** 1995. Simultaneous identification of bacterial virulence genes by negative selection. Science **269:**400–403.
16. **Hingley-Wilson, S. M., V. K. Sambandamurthy, and W. R. Jacobs, Jr.** 2003. Survival perspectives from the world's most successful pathogen, *Mycobacterium tuberculosis*. Nat. Immunol. **4:**949–955.
17. **Hondalus, M. K., S. Bardarov, R. Russell, J. Chan, W. R. Jacobs, Jr., and B. R. Bloom.** 2000. Attenuation of and protection induced by a leucine auxotroph of *Mycobacterium tuberculosis*. Infect. Immun. **68:**2888–2898.
18. **Karakousis, P. C., T. Yoshimatsu, G. Lamichhane, S. C. Woolwine, E. L. Neurmberger, J. Grosset, and W. R. Bishai.** 2004. Dormancy phenotype displayed by extracellular *Mycobacterium tuberculosis* within artificial granulomas in mice. J. Exp. Med. **2000:**647–657.
19. **Kaushal, D., B. G. Schroeder, S. Tyagi, T. Yoshimatsu, C. Scott, C. Ko, L. Carpenter, J. Mehrotra, Y. C. Manabe, R. D. Fleischmann, and W. R. Bishai.** 2002. Reduced immunopathology and mortality despite tissue persistence in a *Mycobacterium tuberculosis* mutant lacking alternative sigma factor, SigH. Proc. Natl. Acad. Sci. USA **99:**8330–8335.
20. **Lamichhane, G., M. Zignol, N. J. Blades, D. E. Geiman, A. Dougherty, J. Grosset, K. W. Broman, and W. R. Bishai.** 2003. A postgenomic method for predicting essential genes at subsaturation levels of mutagenesis: application to *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. USA **100:**7213–7218.
21. **Mahan, M. J., J. M. Slauch, and J. J. Mekalanos.** 1993. Selection of bacterial virulence genes that are specifically induced in host tissues. Science **259:**686–688.
22. **McKinney, J. D., K. Honer zu Bentrup, E. J. Munoz-Elias, A. Miczak, B. Chen, W. T. Chan, D. Swenson, J. C. Sacchettini, W. R. Jacobs, Jr., and D. G. Russell.** 2000. Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. Nature **406:**735–738.
23. **Rubin, E. J., B. J. Akerley, V. N. Novik, D. J. Lampe, R. N. Husson, and J. J. Mekalanos.** 1999. In vivo transposition of mariner-based elements in enteric bacteria and mycobacteria. Proc. Natl. Acad. Sci. USA **96:**1645–1650.
24. **Sambandamurthy, V. K., X. Wang, B. Chen, R. G. Russell, S. Derrick, F. M. Collins, S. L. Morris, and W. R. Jacobs, Jr.** 2002. A pantothenate auxotroph of *Mycobacterium tuberculosis* is highly attenuated and protects mice against tuberculosis. Nat. Med. **8:**1171–1174.
25. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2001. Comprehensive identification of conditionally essential genes in mycobacteria. Proc. Natl. Acad. Sci. USA **98:**12712–12717.
26. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol. Microbiol. **48:**77–84.
27. **Sassetti, C. M., and E. J. Rubin.** 2003. Genetic requirements for mycobacterial survival during infection. Proc. Natl. Acad. Sci. USA **100:**12989–12994.
28. **Siebert, P. D., A. Chenchik, D. E. Kellogg, K. A. Lukyanov, and S. A. Lukyanov.** 1995. An improved PCR method for walking in uncloned genomic DNA. Nucleic Acids Res. **23:**1087–1088.
29. **Steyn, A. J., D. M. Collins, M. K. Hondalus, W. R. Jacobs, Jr., R. P. Kawakami, and B. R. Bloom.** 2002. *Mycobacterium tuberculosis* WhiB3 interacts with RpoV to affect host survival but is dispensable for in vivo growth. Proc. Natl. Acad. Sci. USA **99:**3147–3152.
30. **Tseng, T. T., K. S. Gratwick, J. Kollman, D. Park, D. H. Nies, A. Goffeau, and M. H. Saier, Jr.** 1999. The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. J. Mol. Microbiol. Biotechnol. **1:**107–125.
31. **Valdivia, R. H., and S. Falkow.** 1997. Fluorescence-based isolation of bacterial genes expressed within host cells. Science **277:**2007–2011.
32. **Valway, S. E., M. P. Sanchez, T. F. Shinnick, I. Orme, T. Agerton, D. Hoy, J. S. Jones, H. Westmoreland, and I. M. Onorato.** 1998. An outbreak involving extensive transmission of a virulent strain of *Mycobacterium tuberculosis*. N. Engl. J. Med. **338:**633–639.

Editor: W. A. Petri, Jr.