Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 11554–11559, September 1999
Microbiology

Identification of *Mycobacterium tuberculosis* RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS)

JAMES E. GRAHAM* AND JOSEPHINE E. CLARK-CURTISS†

Department of Biology, Washington University, St. Louis, MO 63130

Edited by Stanley Falkow, Stanford University, Stanford, CA, and approved August 5, 1999 (received for review May 18, 1999)

ABSTRACT **A widely applicable, positive cDNA selection method was developed to identify RNAs synthesized by *Mycobacterium tuberculosis* in response to phagocytosis by cultured human primary macrophages. cDNAs for *sigE* and *sigH* (alternative sigma factors), *aceA* (isocitrate lyase), *ponA* (class I penicillin-binding protein), *pks2* (polyketide synthase), *uvrA* (UvrABC endonuclease), and *ctpV* (putative cation transporter) were obtained from macrophage-grown bacteria. cDNAs for ORFs Rv3070, Rv3483c, Rv0903c (encoding a putative bacterial two-component transcriptional activator), and Rv0170 of the *mce1* virulence operon also were obtained from phagocytized bacilli. cDNAs for these genomic regions were not obtained from approximately 1,000-fold more bacteria grown in laboratory broth. Methods described here, which have identified *M. tuberculosis* genes expressed in response to host interaction, will allow the study of gene expression in a variety of microorganisms, including expression resulting from interaction with human tissues in natural disease states.**

Tuberculosis remains the leading cause of death from disease due to a single infectious agent among youth and adults worldwide. The tubercle bacillus caused disease in an estimated 6.7 million people in 1998, killing an estimated 2.4 million (1).

Mycobacterium tuberculosis initiates infection after inhalation of droplet nuclei capable of transporting between 1 and 3 bacilli into the alveoli of the lung (2). In encounters between bacteria and resident alveolar macrophages not activated previously by inhaled particulates, mycobacteria often are able to survive phagocytosis and multiply within the macrophage phagosomal compartment. After host–cell lysis, bacteria are phagocytized by other nonactivated, infiltrating peripheral blood-derived mononuclear cells, leading to widespread dissemination and transient bacteremia. Survival and growth of *M. tuberculosis* within mononuclear phagocytes are necessary for establishment of both clinical disease and latent infections, which are resolved only by lengthy antimicrobial therapies.

Identifying *M. tuberculosis* genes expressed in response to phagocytosis by human macrophages will increase our basic understanding of the host–pathogen interaction and identify bacterial factors necessary for survival and growth within the human host. Various approaches have been used to identify genes expressed by, or necessary for the growth of, bacterial pathogens or symbionts within the specific environments of host cells and tissues (3–9). Methods based on RNA analysis, including whole-genome expression monitoring (10), although particularly suited to microbial pathogens for which other genetic analysis methods are lacking (6), have been limited by the nature of prokaryotic mRNA. Short mRNA half-lives, limited message polyadenylation, and a scarcity of starting material, particularly in terms of lower abundance class messages, have greatly impeded identification of relevant host interaction-mediated gene expression by direct examination of bacterial mRNA.

We have addressed each of these inherent limitations in the development of a cDNA selection method in which total microbial and host cell RNA are first extracted directly from infected cells or tissue and then converted *en masse* to cDNA fragments containing defined terminal sequences for amplification by PCR. Microbial cDNAs are then selectively captured by hybridization to biotinylated microbial genomic DNA that has been prehybridized with blocking microbial ribosomal DNA (rDNA) fragments (Fig. 1*A*). To identify microbial genes transcribed in response to host interaction, these captured cDNAs then are hybridized to genomic DNA fragments that have been prehybridized with both rDNA and cDNA similarly prepared from microorganisms cultured *in vitro* (Fig. 1*B*). Bacterial cDNA obtained thereby is enriched in sequences corresponding to genomic regions that are preferentially transcribed during growth in the host. To reduce the level of abundant cDNAs captured and obtain more cDNAs for lower abundance class messages (i.e., "normalize"), denatured cDNA is first allowed to partially renature before hybridization of the remaining single-stranded polynucleotides to limiting amounts of microbial genomic DNA (11). Here we describe the method, designated selective capture of transcribed sequences (SCOTS), and an initial application identifying numerous *M. tuberculosis* mRNAs synthesized during infection of cultured primary human macrophages.

MATERIALS AND METHODS

Mycobacterial Cultures. *M. tuberculosis* H37Rv (ATCC no. 25618) was obtained from the American Type Culture Collection and grown in Middlebrook 7H9 broth (100 ml per cotton-stoppered, 500-ml Erlenmeyer flask) supplemented with OADC enrichment [oleic acid, albumin (bovine, Fraction V), dextrose, catalase (DIFCO)] and 0.05% Tween-80 on a rotary shaker (150 rpm) at 37°C. Aliquots (1 ml) containing logarithmic-phase bacteria were frozen (−70°C) as seed stocks and used to inoculate subsequent 100-ml broth cultures. Bacterial population doubling times were approximately 20 h under these growth conditions.

Culture of Human Macrophages and Experimental Infections. Blood was collected from a healthy purified protein derivative skin test-negative donor, and mononuclear cells were isolated immediately by Ficoll (Pharmacia) gradients according to the manufacturer's instructions. Mononuclear cells were incubated in GIBCO RPMI 1640 medium with 25

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office.
Abbreviation: SCOTS, selective capture of transcribed sequences.
*Present address: Department of Medicine, Division of Infectious Diseases, Vanderbilt University, Nashville, TN 37232.
†To whom reprint requests should be addressed. E-mail: clarkcur@biology.wustl.edu.

mM Hepes/4 mM L-glutamine/MEM nonessential amino acids (RPMI) with 20% autologous serum in Teflon wells for 5 days (2×10^6 cells/ml) before separation of monocytes by attachment to plastic (12). Adherent macrophages were infected after 7 days of further *in vitro* maturation (T. M. Kaufman and L. S. Schlesinger, personal communication). Infections at low multiplicity (0.5 bacteria: 1 macrophage) with a predominantly single-cell suspension (13) of opsonized H37Rv (grown to midlogarithmic phase, as described, and diluted into RPMI with 10% autologous serum) were carried out by incubating cell monolayers (2×10^6 cells) and bacteria for 2 h in RPMI with 10% autologous serum. Infected monolayers were washed with RPMI and maintained in RPMI with 1% autologous serum (which does not support growth of H37Rv) with media changes both 16 and 0.5 h before RNA isolation. Intracellular bacterial populations doubled approximately every 18 h during these infections as determined by colony-forming units and Kinyoun staining.

Isolation of Nucleic Acids. *M. tuberculosis* bacterial genomic DNA was prepared either by organic extraction (14) or differential precipitation (15). Total RNA from broth-grown bacteria or from *M. tuberculosis*-infected cells or tissue was obtained by mechanical disruption with silica-zirconium beads (16) and organic extraction (17). Broth cultures were poured directly over ice and centrifuged at 4°C for 10 min before suspension of bacteria in heated (90°C) guanidinium thiocyanate solution. Infected macrophage monolayers were lysed directly with 3 ml of hot guanidinium solution (17).

cDNA Synthesis, Amplification, and Analysis. RNA isolated from approximately 10^6 infected macrophages (20 μg) or from approximately 10^9 log-phase, broth-grown bacilli (5 μg) was treated with DNase I (Ambion) and converted to first-strand cDNA by random priming with mouse murine leukemia virus reverse transcriptase (Superscript II; GIBCO) according to the manufacturer's instructions. A primer with a defined 5′ sequence and nine random 3′ nucleotides was used for both first- and second-strand synthesis as described by Froussard (18). Different terminal sequences therefore were added to broth (5′-CCTCTGAAGGTTCCAGAATCGATAG-3′) and macrophage-derived cDNAs (5′-GACACTCTCGAGACATCAC-CGGTACC-3′). PCR was carried out for 25 cycles (94°C for 40 s, 57°C for 60 s, and 72°C for 30 s), as described previously (19), with the addition of 0.1% Nonidet P-40 and 1% acetamide.

SCOTS. To block capture of bacterial rRNA-derived cDNA, sonicated, photobiotinylated (CLONTECH protocol K1012–1) H37Rv genomic DNA (0.3 μg) was mixed with 5 μg of sonicated *M. tuberculosis rrnA* DNA (plasmid pYA1504) and heated under oil to 98°C for 3 min in 4 μl of 10 mM EPPS [*N*-(2-hydroxyethyl)piperazine-*N*′-(3-propanesulfonic acid) (Sigma)]/1 mM EDTA. One microliter of 1 M NaCl was added, and fragments were hybridized at 77°C (calculated T_m −20°C for 65% G+C DNA content) for 30 min. Total amplified cDNA (3 μg) from either infected macrophages or broth-grown bacteria then was similarly denatured and allowed to reanneal for 30 min at 77°C to remove abundant sequences before hybridization of remaining single-stranded cDNA to rDNA-blocked genomic DNA for 24 h at 77°C (final volume, 10 μl). The length of 30-min prehybridizations used to normalize cDNA and block genomic DNA was based on theoretical calculations and kinetic data (11, 20). DNA mixtures were then diluted to 500 μl with water and bacterial cDNA–chromosomal DNA hybrids captured by binding to 60 μg of streptavidin-coated beads (Dynal M280) according to the manufacturer's instructions. cDNAs were eluted from DNA bound to washed beads with 100 μl of 0.5 M NaOH/0.1 M NaCl, precipitated, and amplified by PCR as described. To maximize representation of original transcripts in these amplified cDNAs, 10 parallel hybridizations and PCR amplifications (21) were used in the first round of SCOTS for each growth condition. These 10 amplified cDNA preparations then were combined, and single, 3-μg aliquots were denatured and again hybridized to rDNA-blocked genomic DNA as described for subsequent rounds of SCOTS (Fig. 1*A*).

Enrichment for cDNAs Specific to Macrophage-Grown Bacilli. Denatured, biotinylated, sonicated H37Rv genomic DNA (0.3 μg) was prehybridized both with cDNA from broth-grown bacteria normalized by three rounds of SCOTS (3 μg) and rDNA fragments (5 μg) for 30 min at 77°C as described (Fig. 1*B*). cDNA from macrophage-grown bacteria normalized by three rounds of SCOTS (1.5 μg) also was denatured and allowed to reanneal for 30 min at 77° before being combined with the prehybridized genomic DNA fragment mixture. After a further 24-h hybridization at 77°C, cDNAs were eluted from DNA bound to streptavidin-coated beads as described above, and PCR-amplified in triplicate via terminal sequences specific for macrophage-grown *M. tuberculosis* cDNA. Aliquots of pooled amplified cDNA (1.5 μg) were denatured and hybridized to blocked genomic DNA for two subsequent rounds of enrichment (Fig. 1*B*). After three rounds of enrichment by positive selection (Fig. 1*B*), cDNAs from macrophage-grown bacteria were cloned via their terminal-sequence *Kpn*I restriction sites into pBluescript II SK(+) (Stratagene). cDNA clones that hybridized specifically to normalized cDNA from macrophage-grown bacteria but not to cDNA from broth-grown bacteria were sequenced.

RESULTS

Preparation of cDNA from Bacilli Grown *in Vitro* and *in Vivo*. Total RNA and PCR-amplified, random-primed cDNA were prepared from log-phase *M. tuberculosis* H37Rv cultures grown in Middlebrook 7H9 broth and directly from *M. tuberculosis* H37Rv-infected, single-donor human macrophages 48 h after infection as described. Complexity, extent of normalization, and reduction of the ribosomal sequence content of the bacterial cDNAs were monitored by radiolabeling PCR-

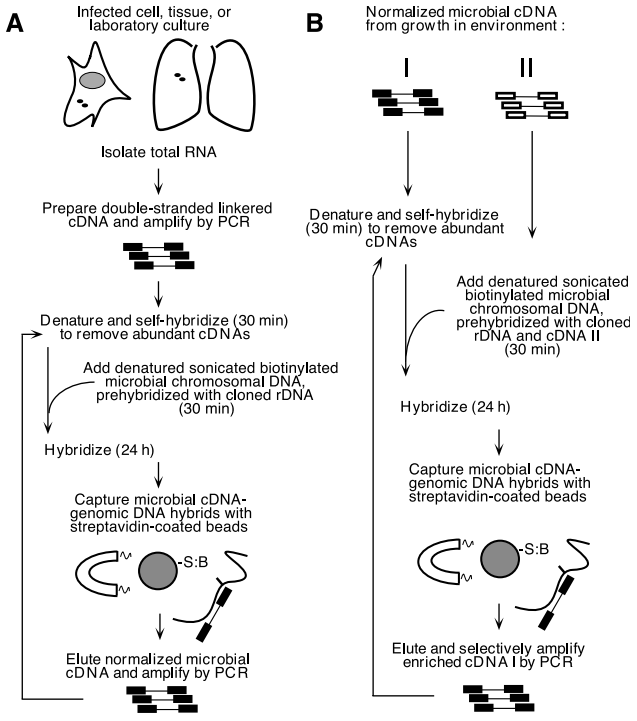

FIG. 1. Identification of microbial genes transcribed *in vivo* by SCOTS. Normalized microbial cDNA is obtained directly from limited samples of infected host cells or tissue (*A*). cDNAs corresponding to genes preferentially expressed in environment I relative to environment II are obtained by differential cDNA hybridization (*B*).

amplified cDNA after each of three rounds of SCOTS for use as probes in Southern hybridizations against H37Rv chromosomal DNA digested with *Pst*I (Fig. 2). Southern hybridizations with restriction endonuclease-digested DNA from 20 randomly chosen cosmids from an *M. tuberculosis* H37Rv genomic library also showed a progressive increase in cDNA normalization and complexity (data not shown). Similar complex bacterial cDNA populations from macrophages 18 and 110 h after infection and from 30 mg of frozen mouse lung tissue [kindly provided by A. M. Cooper and I. M. Orme (Colorado State University, Fort Collins, CO)] containing approximately 10^6 tubercle bacilli also were obtained (data not shown).

Analysis of cDNAs Encoding Mycobacterial Sigma Factors. To establish the limit of our ability to obtain bacterial cDNAs by three rounds of SCOTS, we prepared normalized, radiolabeled cDNA from broth-grown H37Rv for hybridization to a subset of the PCR-generated sigma factor DNA fragments described recently by Manganelli *et al.* (22). These authors determined cDNA levels for 10 H37Rv sigma factors by a novel reverse transcription–PCR assay on total RNA prepared from exponential-phase cultures grown in Middlebrook 7H9 broth. We were able to detect cDNAs hybridizing to DNA fragments encoding major vegetative sigma factors *sigA* and *sigB*, and alternative sigma factors *sigD*, *sigE*, *sigF*, *sigG*, and *sigH*, but not alternative sigma factor *sigC* (as discussed below) in normalized cDNA prepared from log-phase, broth-grown cells (Fig. 3 *B* and *C*). Detection of cDNAs as components of a complex cDNA probe was related to the size of the target fragments used for Southern hybridizations, because larger target fragments containing the same *sigA*, *sigF* (not shown), and *rpoB* DNA regions (Fig. 4 and 5) gave stronger hybridization signals. We therefore detected cDNAs reported to be present at as few as 2.5×10^3 copies per 10 ng of reverse-transcribed total bacterial RNA (*sigG*; ref. 22), when probing small target fragments ranging in size between 81 (*sigH*) and 160 (*sigA*) bp.

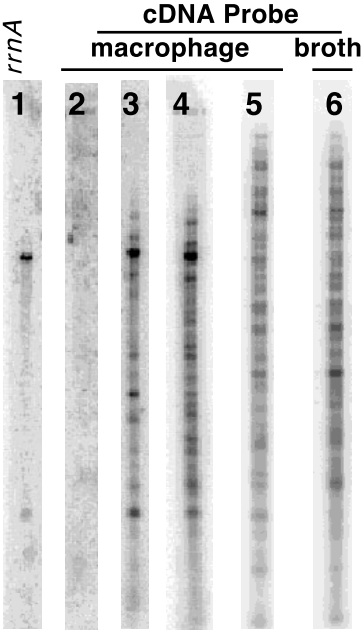

FIG. 2. Bacterial cDNA obtained from intracellular (48-h infection) or broth-grown *M. tuberculosis* H37Rv. cDNAs were prepared as described in the text and used as probes against a Southern blot of H37Rv chromosomal DNA digested with *Pst*I. The probe for lane 1 was the cloned *rrnA* operon of *M. tuberculosis*. Lane 2 was probed with the PCR-amplified total cDNA prepared from infected macrophages as described. Probes for lanes 3, 4, and 5 were bacterial cDNA obtained from infected macrophages by one, two, and three rounds of SCOTS, respectively. Lane 6 was probed with bacterial cDNA obtained by three rounds of SCOTS from a broth culture grown to an OD (A_{600}) of 0.4.

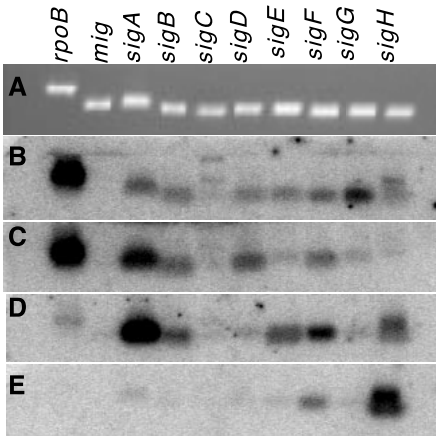

FIG. 3. Identification of sigma factor cDNAs obtained by SCOTS. cDNA probes normalized by three rounds of SCOTS obtained from broth cultures grown to an OD (A_{600}) of 0.15 (*B*) or 0.54 (*C*) or from infected macrophages 18 h (*D*) or 48 h (*E*) after phagocytosis were used in Southern hybridization assays with previously described (22) agarose gel-purified, PCR-amplified DNA fragments (*A*) containing *sigA* (160 bp), *sigB* (106 bp), *sigC* (85 bp), *sigD* (113 bp), *sigE* (115 bp), *sigF* (92 bp), *sigG* (91 bp), *sigH* (81 bp), or *rpoB* (418 bp). *mig* (150 bp) is an *M. avium* DNA fragment not present in H37Rv.

Compared with cDNA prepared from broth-grown cells, cDNA probes prepared by SCOTS from approximately 1,000-fold fewer bacilli growing within human macrophages consistently detected only the major vegetative sigma factors *sigA* and *sigB*, alternative sigma factors *sigE*, *sigF*, and *sigH* 18 h postphagocytosis, and only *sigH* and, to a limited extent, *sigF* and *sigA* 48 h after infection (Fig. 3 *D* and *E*). Differential expression of alternative sigma factors *sigE* and *sigH* in response to macrophage phagocytosis has not been described previously.

Analysis of Intracellular Gene Expression. cDNAs for three genes chosen on the basis of their essential roles in bacterial metabolism (*nrdF*, *rpoB*, and *atpB*) and one previously described as expressed in both broth cultures and during infection of a human monocytic cell line (*fbpB*; ref. 23) all were readily obtained by SCOTS from bacteria grown under different conditions (Figs. 4 and 5; data not shown) as detected by hybridization to target DNA fragments ranging in size from 317 bp (*fbpB*) to 1,753 bp (*rpoB*). After three rounds of

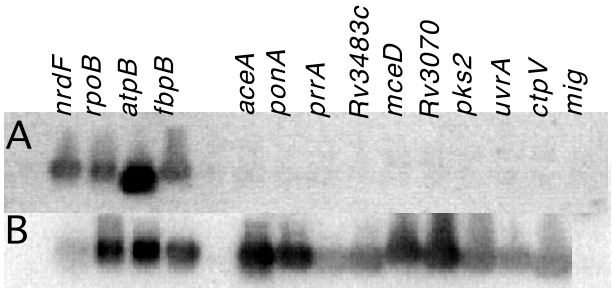

FIG. 4. Transcription of *M. tuberculosis* H37Rv genes during growth in Middlebrook 7H9 broth (*A*) or human macrophages (*B*) as detected by cDNA probes prepared by SCOTS. Bacterial cDNA probes prepared by three rounds of SCOTS (see text) from broth cultures grown to an OD (A_{600}) of 0.4 or from intracellular bacteria 48 h after infection were hybridized to PCR-amplified plasmid fragments containing the indicated H37Rv genes (lanes 1–4) or cloned bacterial cDNA fragments isolated from macrophage-grown bacilli. *mig* (150 bp) is an *M. avium* DNA fragment not present in *M. tuberculosis*. The H37Rv genome coordinates of specific sequences are available on request.

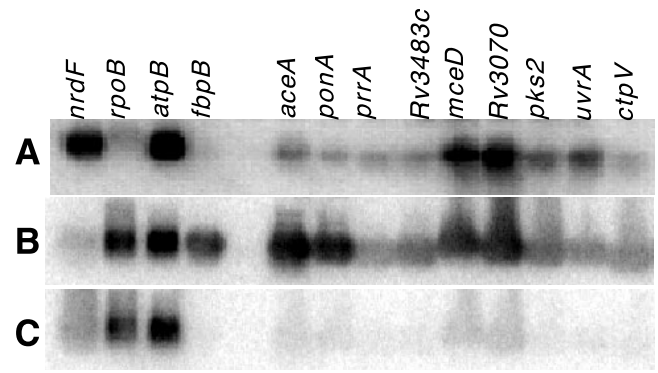

FIG. 5. Modulation of *M. tuberculosis* H37Rv transcript levels during growth in human macrophages as detected by cDNA probes prepared by SCOTS. Bacterial cDNA probes prepared by three rounds of SCOTS (see text) from intracellular bacteria 18 (*A*), 48 (*B*), and 110 h (*C*) after infection were hybridized to PCR-amplified plasmid fragments containing the indicated H37Rv genes (lanes 1–4) or cloned bacterial cDNAs isolated from phagocytized bacilli (48 h).

enrichment by differential hybridization (Fig. 1*B*), nine distinct cDNA clones from macrophage-grown bacilli (48 h) that showed no hybridization to normalized cDNA probes from log-phase, broth-grown bacilli were identified (Fig. 4). cDNA obtained by three rounds of SCOTS from *M. tuberculosis* both 18 and 110 h after macrophage infection also contained cDNA that hybridized to these clones, but at reduced levels relative to control cDNAs at 110 h (Fig. 5). cDNA prepared by three rounds of SCOTS and three rounds of enrichment from an independent, early-log-phase broth culture and 18 h macrophage infection showed a hybridization pattern similar to enriched 48-h macrophage cDNA when used to probe an H37Rv cosmid library (data not shown).

cDNAs obtained only from macrophage-grown bacilli identified several classes of *in vivo* gene expression established by recent studies of other pathogenic bacteria (24–27). These include DNA repair (*uvrA*), nutrient acquisition (*ctpV*, *aceA*), cell-wall metabolism (*ponA*), and previously described virulence factors (*mceD*; ref. 28). In addition, cDNAs encoding other factors potentially relevant to the specific intracellular lifestyle of *M. tuberculosis*, including a previously undescribed, membrane-associated protein (Rv3483c; ref. 29), polyketide synthase (*pks2*), and two-component transcriptional activator (Rv0903c; designated *prrA* for phagocytosis response regulator), were obtained from bacilli grown in human macrophages, but not from approximately 1,000-fold more broth-grown tubercle bacilli. We also were able to obtain cDNAs that hybridized to cloned 350- to 400-bp DNA fragments containing *uvrB*, *uvrD*, *glcB* (malate synthase), and the *prrA* cognate sensor kinase *prrB* from macrophage-grown bacilli 48 h postinfection (data not shown).

DISCUSSION

We have used a novel combination of selective and differential hybridization strategies to identify numerous *M. tuberculosis* genes that appear to be transcribed in response to host–cell interaction. Rather than attempting to maintain proportional representation of mRNAs in PCR-amplified cDNA as done with genomic DNA fragment libraries in a similar semiquantitative approach described recently by Alland *et al.* (30), we have used both limited self-hybridization (11) and hybridization to genomic DNA to increase the number of sequences identified as differentially expressed and to extend our analysis beyond the most abundant bacterial mRNAs. In broth-grown *Escherichia coli*, the most abundant 250 transcripts comprise about 95% of the total mass of mRNA (20). Although there are potential biases in representing total mRNA first as normalized cDNA and then enriched cDNA clones, the use of random priming to create cDNA within a limited size range (200–400 bp) is expected to provide a variety of different contexts for each sequence, thereby reducing potential losses during amplification, normalization, enrichment, and plasmid cloning (21, 31, 32).

SCOTS was able to obtain cDNAs for seven of eight previously characterized sigma factor transcripts from broth-cultured bacteria, with the exception of *sigC* (22). Manganelli *et al*. (22) indicated that SigC is not found to be associated with purified RNA polymerase at levels similar to the major vegetative sigma factors SigA and SigB, suggesting that their detection of *sigC* expression by reverse transcription–PCR (RT-PCR) may have been influenced by the relative efficiency of reverse transcription of template RNAs. In addition to using mouse murine leukemia virus-derived reverse transcriptase (rather than avian myeloblastosis virus polymerase) in our studies, the Middlebrook 7H9 broth we used contained oleic acid as OADC supplement rather than the ADC supplement used by Manganelli *et al.* (22), which may have had some effect on steady-state *sigC* mRNA levels. Sequence analysis of the predicted SigC peptide indicates that it is a member of the extracytoplasmic family of alternative sigma factors (33) rather than the class of major vegetative sigma factors represented by *sigA* and *sigB*, for which cDNAs were readily obtained by SCOTS from logarithmic-phase bacteria. We also did not see any reproducible increase in our ability to obtain *sigF* cDNA once bacteria entered stationary phase (Fig. 2*C*), which is consistent with results obtained by RT-PCR (22) using primers that target the identical *sigF* cDNA region we analyzed and in contrast to previous analysis of stationary-phase expression in *Mycobacterium bovis* bacillus Calmette–Guérin by ribonuclease protection assay (34). Because SCOTS is intended as a screen for identification of mRNAs whose steady-state levels change dramatically under different growth conditions and uses cDNA normalization to extend analysis to less abundant transcripts, we are not able to provide more quantitative data to resolve this issue without further modification of the SCOTS method.

We also were able to reproducibly obtain bacterial cDNAs indicating, first, elevated steady-state *sigE* mRNA levels and, later, *sigH* transcript levels during the course of human macrophage infection. This is, to our knowledge, the first indication of a potential role for alternative sigma factors during survival and growth of *M. tuberculosis* within the human macrophage phagosome. Recently, another bacterial extracytoplasmic sigma factor, SigE, has been reported to have an important role in the regulation of the expression of *Salmonella typhimurium* virulence and antigenicity determinants (35). The extracytoplasmic subfamily of bacterial sigma factors, which are remarkably conserved between Gram-positive and Gram-negative species (36), includes all 11 putative alternative sigma factors identified in the *M. tuberculosis* genome (33). These, indeed, may be key regulatory molecules of bacterial host-adaptive responses, functioning in transduction of extracytoplasmic signals to the bacterial transcription apparatus. Of the 13 putative *M. tuberculosis* sigma factors, SigH has the greatest amount of sequence similarity to SigE of *S. typhimurium* and *Escherichia coli* and, like *E. coli* sigE, is expressed in response to elevated temperature (22).

Our ability to obtain cDNAs for several genes only from *M. tuberculosis* growing within human macrophages indicates what appear to be common challenges faced by a variety of pathogenic bacteria during infection, despite differences in host environments encountered and survival strategies employed. The induction of operons involved in purine, pyrimidine, and amino acid biosynthesis, iron acquisition, and cation transport in phagocytized *S. typhimurium* suggests that phagosomal compartments containing these bacteria are low in nutrients (37). Responses to cation limitation during growth *in*

vivo, as indicated here by cDNA for *M. tuberculosis ctpV* (Fig. 4), which encodes a putative ATP-dependent cation transport protein, previously have been described in both Gram-positive and Gram-negative bacterial pathogens (24, 26). Magnesium ion limitation specifically has been shown to be an environmental signal capable of mediating expression of the *S. typhimurium* PhoP regulon, which consists of numerous genes necessary for growth and survival within host macrophages (38).

Expression of *aceA,* which encodes the *M. tuberculosis* isocitrate lyase, may be important in obtaining carbon from the breakdown of host lipids during growth in the human macrophage phagosome (33). zu Bentrup *et al.* (K. H. zu Bentrup, D. L. Swenson, A. Miczak, and D. Russell, unpublished data) recently have shown AceA to be one of two predominant *Mycobacterium avium* peptides synthesized in response to phagocytosis by mouse macrophages. Immunocrossreactive material to this peptide also was detected in *M. tuberculosis*-infected mouse macrophages and tissues (K. H. zu Bentrup, D. L. Swenson, A. Miczak, and D. Russell, unpublished data). The fact that we did not identify genes of the amino acid and nucleotide biosynthetic pathways differentially expressed in other studies of bacterial *in vivo* gene expression may reflect the lack of these nutrients in Middlebrook 7H9 broth compared with the relatively nutrient-rich culture media used by other investigators (7, 9, 24–26, 39).

A bacterial SOS response, as indicated here by *uvrA* cDNA (Fig. 4), also appears to be a nearly universal response of pathogenic bacteria to growth within the host, regardless of whether bacteria reside within host cells, bloodstream, or tissues (24, 25, 27, 40). Although *M. tuberculosis* may be able to avoid the oxidative burst associated with Fc receptor-mediated phagocytosis (41), our ability to obtain *uvrA* cDNA from bacteria 48 h after infection suggests an SOS response followed phagocytosis by nonactivated human macrophages. Oxidative DNA damage by other host defense mechanisms (42) therefore may remain a challenge to *M. tuberculosis* well after internalization by human macrophages.

Enzymes involved in cell-wall metabolism, particularly amidases and peptidoglycan-modifying enzymes in Gram-positive bacteria, also have been identified as differentially expressed and only necessary for *in vivo* growth in various animal infection models (25–27). *M. tuberculosis ponA* (Fig. 4) likely encodes a class I penicillin-binding protein that is similar to a *Bacillus subtilis* penicillin-binding protein that is only necessary for growth in Mg^{2+}- or Ca^{2+}-limited media (43).

One of the most striking features of the H37Rv genome is the presence of four *mce* gene clusters, which share a core sequence of eight colinear ORFs. Arruda *et al.* (28) identified a 1,535-bp *mce1*-containing DNA fragment that conferred upon *E. coli* strains an ability to invade HeLa cells and increased both uptake by and short-term (24 h) survival within human macrophages. The invasive activity was associated with the upstream Rv0169 (*mce*) sequence and the ability to survive within macrophages with the downstream Rv0170 (*mceD*) sequence. Chubb *et al.* (44) subsequently have shown that the hydrophobic amino-terminal of MceD (Rv0170) indeed is functional in *E. coli*, directing export of a leaderless β-lactamase into the periplasm. If the four *mce* clusters arose by gene amplification, it is likely that the encoded products offer a significant advantage to *M. tuberculosis* in its normal ecological niche (45). Elucidation of a potential role for *mceD* specifically in *M. tuberculosis* pathogenesis as well as the extent of differential transcription within and across the *mce* operons is now being studied.

A significant portion of the *M. tuberculosis* genome is dedicated to polyketide and nonribosomal peptide synthesis (33). Polyketides are a diverse group of microbial secondary metabolites with potent biological activities, acting as virulence factors, toxins, antibiotics, and immunosuppressants. H37Rv *pks2* is 1 of 18 ORFs predicted to encode polyketide synthases. The putative *pks2* gene product is a member of a polyketide subgroup that shows similarity to mycocerosic acid synthase. Some members of this subgroup may contribute to the formation of branched fatty acids in cell-wall-associated phthiocerol dimycocerosate and mycosides (46). Recently, George *et al.* (47) identified a polyketide toxin from *M. ulcerans*, which the authors suggest may represent a family of mycobacterial virulence factors. The role of *pks2* during macrophage infection is unclear; however, polyketide synthase activity in cell-wall metabolism and the cytopathic and immunomodulatory activity of certain polyketides (33, 48) make this an important area for future investigation.

We also obtained cDNAs for two H37Rv ORFs predicted to encode membrane-associated proteins only from phagocytized bacilli. Rv3070 (33) shows reasonable overall similarity (35% identity and 52% similarity over 117 of a total of 122 aa) to the *E. coli* CrcB protein. Proteins with comparable degrees of similarity to Rv3070 are widespread among prokaryotes; their function is largely unknown, beyond a potential role in chromosome condensation (49). Rv3483c contains a predicted amino-terminal transmembrane helix (29) and is unusual in that it shows limited similarity to eukaryotic proteins: neural cell adhesion molecule NCAM30 (27% identity and 33% similarity over 181 aa) and human herpes simplex virus glycosylated capsid protein C (gC-2; 29% identity and 46% similarity over 98 aa). The significance of these potentially important similarities remains to be determined.

Bacteria typically respond to changing environmental conditions through paired membrane protein sensors and DNA-binding response regulators. The *M. tuberculosis* genome contains only 11 such two-component regulatory systems, compared with the similarly sized genomes of *E. coli* and *B. subtilis*, which each have more than 30 (33). Isolation of cDNAs encoding the *prrA* (Rv0903c) response regulator and the *prrB* cognate sensor kinase (Rv0902c) from *M. tuberculosis* growing intracellularly indicates that this two-component response regulator (33) may play an important role in the environmental adaptation of the bacterium to the phagosome of the human macrophage. Identifying relationships between *prrA* expression and that of other genes reported here, as well as determining exactly what genes constitute this novel phagocytosis-induced regulon, will contribute to our understanding of what *M. tuberculosis* does within the host macrophage—that which makes it among the most successful of human pathogens.

Functional analysis of the nine *M. tuberculosis* genes identified by SCOTS is expected to provide further insights into important aspects of the human host–pathogen interaction. The SCOTS procedure can be applied to any microbial pathogen from which total nucleic acids can be obtained, with no requirement for specialized genetic techniques, DNA microarrays, libraries, or species-specific cloning vectors. Unlike other recently described methods of analyzing *in vivo* gene expression (9, 24), SCOTS identifies relevant genes, rather than promoter regions, and is not confounded by polar effects when genes are arranged in polycistronic operons (7). SCOTS also has the advantage of being able to detect *in vivo* gene expression from the small numbers of microbial cells obtained in samples from living tissues in natural disease states, including human biopsies, an application for which no other methods are currently available.

We thank Joan Y. Hou for assistance with human macrophage isolation, Larry Schlesinger for advice on macrophage cultivation, Brian J. Morrow, Jubril O. Hassan, Peter K. Brown, and France Daigle for helpful discussions, and Christos Stathopoulos and Roy Curtiss III for comments on the manuscript. This research was supported by Public Health Service Grants AI35267, AI38672, and AI45244 from the National Institutes of Health.

1. Murray, C. J. & Salomon, J. A. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 13881–13886.
2. Dannenberg, A. M. & Rook, G. A. W. (1994) in *Tuberculosis: Pathogenesis, Protection, and Control*, ed. Bloom, B. R. (ASM Press, Washington, DC), pp. 459–484.
3. Scott-Craig, J. S., Guerinot, M. L. & Chelm, B. K. (1991) *Mol. Gen. Genet.* **228,** 356–360.
4. Abshire, K. Z. & Neidhardt, F. C. (1993) *J. Bacteriol.* **175,** 3734–3743.
5. Mahan, M. J., Slauch, J. M. & Mekalanos, J. J. (1993) *Science* **259,** 686–688.
6. Plum, G. & Clark-Curtiss, J. E. (1994) *Infect. Immun.* **62,** 476–483.
7. Hensel, M., Shea, J. E., Gleeson, C., Jones, M. D., Dalton, E. & Holden, D. W. (1995) *Science* **269,** 400–403.
8. Abu Kwaik, Y. & Pederson, L. L. (1996) *Mol. Microbiol.* **21,** 543–556.
9. Valdivia, R. H. & Falkow, S. (1997) *Science* **277,** 2007–2011.
10. de Saizieu, A., Certa, U., Warrington, J., Gray, C., Keck, W. & Mous, J. (1998) *Nat. Biotechnol.* **16,** 45–48.
11. Ko, M. S. (1990) *Nucleic Acids Res.* **18,** 5705–5711.
12. Schlesinger, L. S. & Horwitz, M. A. (1991) *J. Immunol.* **147,** 1983–1994.
13. Clemens, D. L. & Horwitz, M. A. (1995) *J. Exp. Med.* **181,** 257–270.
14. Clark-Curtiss, J. E., Jacobs, W. R., Docherty, M. A., Ritchie, L. R. & Curtiss, R., III (1985) *J. Bacteriol.* **161,** 1093–1102.
15. van Soolingen, D., de Haas, P. E., Hermans, P. W. & van Embden, J. D. (1994) *Methods Enzymol.* **235,** 196–205.
16. Cheung, A. L., Eberhardt, K. J. & Fischetti, V. A. (1994) *Anal. Biochem.* **222,** 511–514.
17. Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159.
18. Froussard, P. (1992) *Nucleic Acids Res.* **20,** 2900.
19. Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J. (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, San Diego).
20. Hahn, W. E., Pettijohn, D. E. & Van Ness, J. (1977) *Science* **197,** 582–585.
21. Wieland, I., Bolger, G., Asouline, G. & Wigler, M. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 2720–2724.
22. Manganelli, R., Dubnau, E., Tyagi, S., Kramer, F. R. & Smith, I. (1999) *Mol. Microbiol.* **31,** 715–724.
23. Lee, B. Y. & Horwitz, M. A. (1995) *J. Clin. Invest.* **96,** 245–249.
24. Heithoff, D. M., Conner, C. P., Hanna, P. C., Julio, S. M., Hentschel, U. & Mahan, M. J. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 934–939.
25. Mei, J. M., Nourbakhsh, F., Ford, C. W. & Holden, D. W. (1997) *Mol. Microbiol.* **26,** 399–407.
26. Lowe, A. M., Beattie, D. T. & Deresiewicz, R. L. (1998) *Mol. Microbiol.* **27,** 967–976.
27. Polissi, A., Pontiggia, A., Feger, G., Altieri, M., Mottl, H., Ferrari, L. & Simon, D. (1998) *Infect. Immun.* **66,** 5620–5629.
28. Arruda, S., Bomfim, G., Knights, R., Huima-Byron, T. & Riley, L. W. (1993) *Science* **261,** 1454–1457.
29. Sonnhammer, E. L. L., von Heijne, G. & Krogh, A. (1998) in *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology* (AAAI Press, Menlo Park, CA).
30. Alland, D., Kramnik, I., Weisbrod, T. R., Otsubo, L., Cerny, R., Miller, L. P., Jacobs, W. R., Jr., & Bloom, B. R. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 13227–13232.
31. Ko, M. S., Ko, S. B., Takahashi, N., Nishiguchi, K. & Abe, K. (1990) *Nucleic Acids Res.* **18,** 4293–4294.
32. Grothues, D., Cantor, C. R. & Smith, C. L. (1993) *Nucleic Acids Res.* **21,** 1321–1322.
33. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., III, *et al.* (1998) *Nature (London)* **393,** 537–544.
34. DeMaio, J., Zhang, Y., Ko, C., Young, D. B. & Bishai, W. R. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 2790–2794.
35. Humphreys, S., Stevenson, A., Bacon, A., Weinhardt, A. B. & Roberts, M. (1999) *Infect. Immun.* **67,** 1560–1568.
36. Missiakas, D. & Raina, S. (1998) *Mol. Microbiol.* **28,** 1059–1066.
37. Heithoff, D. M., Conner, C. P. & Mahan, M. J. (1997) *Trends Microbiol.* **5,** 509–513.
38. Garcia Vescovi, E., Soncini, F. C. & Groisman, E. A. (1996) *Cell* **84,** 165–174.
39. Coulter, S. N., Schwan, W. R., Ng, E. Y., Langhorne, M. H., Ritchie, H. D., Westbrock-Wadman, S., Hufnagle, W. O., Folger, K. R., Bayer, A. S. & Stover, C. K. (1998) *Mol. Microbiol.* **30,** 393–404.
40. Young, G. M. & Miller, V. L. (1997) *Mol. Microbiol.* **25,** 319–328.
41. Schlesinger, L. S. (1996) *J. Invest. Med.* **44,** 312–323.
42. Fang, F. C. (1997) *J. Clin. Invest.* **99,** 2818–2825.
43. Murray, T., Popham, D. L. & Setlow, P. (1998) *J. Bacteriol.* **180,** 4555–4563.
44. Chubb, A. J., Woodman, Z. L., da Silva Tatley, F. M., Hoffmann, H. J., Scholle, R. R. & Ehlers, M. R. (1998) *Microbiology* **144,** 1619–1629.
45. Romero, D. & Palacios, R. (1997) *Annu. Rev. Genet.* **31,** 91–111.
46. Kolattukudy, P. E., Fernandes, N. D., Azad, A. K., Fitzmaurice, A. M. & Sirakova, T. D. (1997) *Mol. Microbiol.* **24,** 263–270.
47. George, K. M., Chatterjee, D., Gunawardana, G., Welty, D., Hayman, J., Lee, R. & Small, P. L. (1999) *Science* **283,** 854–857.
48. Motamedi, H., Cai, S. J., Shafiee, A. & Elliston, K. O. (1997) *Eur. J. Biochem.* **244,** 74–80.
49. Hu, K. H., Liu, E., Dean, K., Gingras, M., DeGraff, W. & Trun, N. J. (1996) *Genetics* **143,** 1521–1532.