Biochemical and Biophysical Research Communications **258,** 94–101 (1999)
Article ID bbrc.1999.0591, available online at http://www.idealibrary.com on IDEAL®

Search for Genes Potentially Involved in *Mycobacterium tuberculosis* Virulence by mRNA Differential Display

Laura Rindi, Nicoletta Lari, and Carlo Garzelli
Dipartimento di Patologia Sperimentale, Biotecnologie Mediche, Infettivologia ed Epidemiologia, Università di Pisa, I-56127 Pisa, Italy

Received March 22, 1999

An mRNA differential display (DD) assay was developed to compare gene expression between *Mycobacterium tuberculosis* H37Rv and its avirulent mutant H37Ra. The DD protocol made use of an oligo(dT) to prime reverse-transcriptase (RT)-dependent transcription of poly-A tailed mRNAs and a PCR amplification of the RT products by using ten 12-base arbitrary primers in all their pair combinations. This analysis yielded 745 and 708 bands, including 52 and 15 differentially generated bands, in the strains H37Rv and H37Ra, respectively. Six cDNAs that appeared to be expressed in H37Rv, but not in H37Ra, were reamplified and cloned and at least 10 inserts were sequenced for each cloned cDNA. After resolving discrepant results, 6 inserts were found highly homologous to *M. tuberculosis* H37Rv genes. Three of these, *i.e.*, genes Rv2770c, Rv1345, and Rv0288, coding respectively for a member of the PPE protein family, a probable polyketide synthase, and a member of the protein family containing ESAT-6, have been predictively associated to immunological or pathogenetic aspects of *M. tuberculosis* infection; the other genes, *i.e.*, Rv2336, Rv1320c, and Rv2819c, code for proteins with unknown functions. These results show that mRNA DD methodology can represent a potential tool for investigation of *M. tuberculosis* gene expression.

***Key Words: Mycobacterium tuberculosis;* virulence; mRNA differential display.**

The mechanisms by which *Mycobacterium tuberculosis* establishes disease in humans are not well understood, although a number of properties, including the production of cord factor, sulfatides and mycosides, are usually associated with the capacity of virulent *M. tuberculosis* strains to produce progressive disease. The recent advances in the development of molecular genetic methods for mycobacteria has allowed investigation of *M. tuberculosis* virulence factors at the gene level and various strategies have been used to identify mycobacterial virulence genes (1-4). The basic strategy implies the construction of *M. tuberculosis* DNA libraries and their expression in *Escherichia coli* or other non-pathogenic mycobacteria (5); more recently, a strategy based on the disruption of a candidate virulence gene and its reintroduction into an avirulent mutant has been also successfully used (6).

An experimental model for the investigation of *M. tuberculosis* virulence genes is represented by the virulent strain H37Rv and its avirulent mutant H37Ra (7, 8). This model is based on the assumption that the avirulent phenotype must necessarily involve alteration(s) in the expression of genes involved in the virulence of the organism, as compared to the virulent phenotype. By constructing genomic libraries of the virulent strain in a cosmid vector and transforming the avirulent mutant, genomic fragments of *M. tuberculosis* responsible for in vivo growth advantage have been identified (9, 10); moreover, by the use of mRNA subtraction libraries, genes differentially expressed in the virulent strain *M. tuberculosis* H37Rv have been identified and cloned (11).

We have addressed the search for *M. tuberculosis* virulence genes in the H37Rv/H37Ra model by mRNA differential display (DD), a potentially powerful methodology originally developed in eukaryotic systems by Liang and Pardee (12) that makes possible to identify and isolate genes that are over- or under-expressed in one cell type relative to others. So far, the application of DD techniques to prokaryotes has been hampered by the general scepticism on the existence of polyadenylated tracts in bacterial mRNAs that can be targeted to initiate oligo(dT)-primed, reverse transcriptase (RT)-dependent synthesis of complementary DNA (cDNA). Nonetheless, many reports do support the existence of polyadenylated tracts at the 3′-end of bacterial mRNAs. The bacterial 3′adenylation, that seems to play a major role in the regulation of RNA decay (reviewed in 13), differs in the length of poly(A) sequences, which is considerably shorter in bacteria (15-60 nucleotides) compared to eukaryotic cells (80-200 nucleotides), and in the smaller fraction of poly-

0006-291X/99 $30.00

adenylated molecules (1-40%) compared to polyadenylation of most eukaryotic mRNA molecules (reviewed in 14 and 15).

In this paper, on the basis of our recent findings that also in *M. tuberculosis* poly(A)-tailed mRNAs can be targeted for oligo(dT)-primed synthesis of cDNA (16), we describe a DD strategy to study differential gene expression in strains H37Rv and H37Ra of *M. tuberculosis*, similar to that described in the original publication by Liang and Pardee (12), that makes use of an oligo(dT) to prime RT-dependent transcription of mycobacterial mRNAs and a PCR amplification of the RT products by using 12-base arbitrary primers.

MATERIALS AND METHODS

Bacterial strains. *M. tuberculosis* H37Rv and H37Ra, from collection maintained at our Department, grown on Lowenstein-Jensen medium (Difco Laboratories), were used.

RNA preparation. Four to five 10 μl-loopfuls of bacterial cells were collected into a microcentrifuge tube containing 200 μl of TE buffer (0.01 M Tris-HCl, pH 8.0; 0.001 M EDTA). Lysozyme was added to a final concentration of 10 mg/ml and incubated at 37°C overnight. The mixture was incubated in the presence of 1% SDS and at 37°C for 15 min and, after vortexing, proteinase K (200 μg/ml, final concentration) was added and incubated at 42°C for 15 min. RNA was extracted essentially as described by Chomczynsky and Sacchi (17) with minor modifications. Briefly, 500 μl of lysis buffer (4 M guanidinium thiocyanate, 25 mM sodium citrate pH 7.0, 0.5% sarcosyl, 0.1 M 2-mercaptoethanol), 50 μl of 2 M sodium acetate pH 4.0, 500 μl of acid phenol (Sigma, USA), and 100 μl of chloroform/isoamyl alcohol (24/1 v/v) were sequentially added to the mixture with thorough mixing after the addition of each reagent. The suspension was then vigorously shaken for 10 s and cooled on ice for 15 min. After centrifugation for 20 min at 10000 *g* at 4°C, RNA was precipitated by adding 1 volume of isopropanol to the aqueous supernatant. After 1 h at −80°C, the mixture was centrifuged for 20 min at 10000 *g* at 4°C, the pellet was dissolved in 300 μl of lysis buffer, added of 30 μl of sodium acetate, and precipitated by isopropanol as described above. The mixture was centrifuged for 15 min at 10000 *g* at 4°C and the pellet was washed once with 75% ethanol, vacuum-dried and dissolved in diethyl pyrocarbonate (DEPC)-treated water. To remove chromosomal DNA contamination, approximately 50-100 μg of total cellular RNA was incubated for 30 min at 37°C with 20 U RNase inhibitor (Boehringer Mannheim) and 200 U DNase I (Gibco) in 50 mM Tris pH 8.3, 75 mM KCl, 3 mM $MgCl_2$. Another 200 U DNase I were added and incubated as described above. After extraction with phenol/chloroform/isoamylic alcohol, the supernatant was ethanol-precipitated in the presence of 0.1 M sodium acetate and RNA was dissolved in DEPC-treated water. RNA concentration was estimated by spectrophotometer reading (Gene Quant II, Pharmacia Biotech).

mRNA differential display. RNA reverse transcription was carried out as previously reported (16). Briefly, a sample of 1 μg bacterial RNA, dissolved in DEPC-treated water, was heated at 65°C for 15 min and then cooled on ice before cDNA synthesis. Oligo$(dT)_{15}$ (Pharmacia Biotech) was used to initiate the first strand cDNA synthesis, using 200 U M-MLV reverse transcripase (RT) (Gibco) in a buffer containing 50 mM Tris, pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 0.5 mM each dNTP, 20 U RNase inhibitor, and 25 pmoles of each primer for 1 h at 37°C in a 20 μl-volume. The reaction was stopped by incubation at 95°C for 5 min. For amplification, 10 arbitrary 12-mer oligonucleotide primers, coded 1 through 10, the sequences of which are given in Table 1, were used in all combinations to amplify target sequences of the generated cDNAs. PCR was performed in 0.5 ml-microcentrifuge reaction tubes in a final volume of 50 μl containing 10 mM Tris-HCl (pH 8.8), 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X-100, 2-3 μM primers, 100 μM dNTP, 2 U *Taq* polymerase (Dynazyme) and 1 μl of the RT reaction products. The mixture was overlaid with 75 μl mineral oil to prevent evaporation. Hot-start PCR amplification was performed with a OmniGene temperature cycler (Hybaid, U.K.) set for 3-min at 94°C and 2-min at 80°C and for 35 cycles of 1 min at 94°C, 1 min at 50°C and 1 min at 72°C and one final 5-min cycle at 72°C.

TABLE 1

Arbitrary Oligonucleotide Primers Used in This Study[a]

Primer	Sequence
1	GCCAAGCTCCAG
2	GGCGTCATCGAC
3	GCCATCCTCGAC
4	GGCAAGGCACAG
5	GCGGTCATCCTG
6	GCGCTCAAGCTG
7	GGTCTCGCAGTG
8	GGTGTCACCGAC
9	GCGACCAAGGTG
10	GAGGCAGTCGAG

[a] Criteria for arbitrary primer design are given under Results.

Ten-μl aliquots of PCR products were separated by 5% polyacrilamide (29:1), 20 cm-long, 0.75 mm-thick gels in a vertical electrophoresis cell (Protean II, BioRad); sometimes, PCR products were further analyzed by electrophoretic runs on pre-cast 12.5% polyacrilamide, 11 cm-long, 0.5 mm-thick gels (GenePhor, Pharmacia). Molecular mass markers were included in each run. Gels were silver-stained, and then photographed for computer-driven densitometric analysis by GelBase/GelBlot Pro software (UVP Inc., CA). Bands corresponding to amplicons of interest were isolated and re-amplified in a second step PCR amplification with the appropriate pairs of primers under the same conditions used for first-step cDNA amplifications.

Cloning and sequencing procedures. Molecular cloning techniques were performed according to standard protocols (18). In particular, the re-amplification PCR products were purified by QIAquick PCR Purification Kit (Qiagen) and then inserted into the dephosphorylated *EcoR*V site of the plasmid pBluescript SK+ (Stratagene) by the T4 DNA ligase (Pharmacia Biotech) and the ligation product was used to transform *E. coli* JM109 cells. Bacterial clones containing inserts were identified as white colonies from which the plasmid DNA was purified by standard alkaline lysis procedures. At least 10 colonies from each subcloning were chosen and the molecular mass of cloned fragments was checked by restriction of recombinant plasmids with *BamH*I and *Hind*III enzymes.

Sequencing of the inserted DNA fragments of the expected size was carried out with an automated apparatus (ALF*express* DNA sequencer, Pharmacia Biotech) according the T7 DNA polymerase method by the Bluescript T3 and T7 universal sequencing primers. The databases of EMBL and Sanger Centre were used for searching for homologous nucleic acid sequences.

Specific primer-RT-PCR. The sets of primers used for amplification of specific cDNAs generated by oligo$(dT)_{15}$-primed reverse transcription of RNA are shown in Table 2. PCR was performed in 0.5 ml-microcentrifuge reaction tubes in a final volume of 50 μl containing 10 mM Tris-HCl (pH 8.8), 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X-100, 250 ng primers, 200 μM dNTP, 1.25 U of Taq polymerase (Dynazyme) and 2 μl of the RT reaction products. PCR amplification was run for one 3-min cycle at 94°C, 25 cycles of 1 min at 94°C, 1 min at primer annealing temperature and 2 min at 72°C and one final 4-min cycle at 72°C.

TABLE 2

Specific Oligonucleotide Primers Used in This Study[a]

Primer pair[a]	Sequences	PCR product length (bp)
HSP65-U	^{825}CGGTTCGACAAGGGCTACATC^{845}	225
HSP65-L	^{1049}GCGGATCTTGTTGACGACCAG^{1029}	
85-C-U	^{497}GGGCGGCCAATCCAGTTTCTAC^{518}	307
85-C-L	^{803}GTAACCGCCCGAGTCGTTCATC^{782}	
Rv1345-U	^{980}TGCTTGCCGACCGTGTCATC^{999}	175
Rv1345-L	^{1154}CCCAGGTAACCCGCCATCAT^{1135}	
Rv2770c-U	^{977}ATCCAGTGCTCGGTGCTCAG^{996}	192
Rv2770c-L	^{1168}CCAGAAACGGAACGCCTCAG^{1149}	
Rv0288-U	^{76}CAGAGCTTGGGTGCCGAGAT^{95}	160
Rv0288-L	^{235}TGGCTTCATGGGTGCTGGAC^{216}	
Rv2819c-U	^{28}GCACTGACCCATCTCGTAGCA^{48}	158
Rv2819c-L	^{185}AGACCTTTGTCACCGACCAAGA^{164}	
Rv2336-U	^{340}ACCCGAGAAACCGAGGTCAT^{359}	190
Rv2336-L	^{529}CTTCGATCGCGTCATAAAGC^{510}	
Rv1320c-U	^{1189}GAACCGCGGGATAAACAACC^{1208}	209
Rv1320c-L	^{1397}TCATACTGGGCACCTACTGG^{1378}	

[a] Primer codes include the gene name, according to the nomenclature reported by Cole *et al.* (20), and a letter (U or L) that indicates upper and lower primer, respectively. For further details on the gene products and accession numbers, see text and Table 4.

Dot-blot hybridization. RT-PCR product samples were heated to 95°C, chilled on ice and then 1 volume of 20× SSC (3 M NaCl, 0.3 M Na-Citrate, pH7.0) was added. The samples were spotted onto a nylon filter (Hybond N-plus membrane; Amersham) prewet with 10× SSC, in 4-μl aliquots. The membrane was wet in denaturing solution (1.5 M NaCl, 0.5 M NaOH) for 5 min and then transferred to a filter paper pad soaked in neutralizing solution (1.5 M NaCl, 0.5 M Tris-HCl pH7.2, 0.001 M EDTA) for 1 min. After air drying, the DNA was fixed to the nylon filter by placing the blot for 2 min on a Whatmann 3 MM filter soaked in 0.4 M NaOH. After rinsing with 5× SSC, nylon filter was pre-hybridized by incubating in 10-15 ml of hybridization buffer supplied by Amersham ECL Direct System for 15 min at 42°C in roller bottles into a hybridization oven (Hybaid). Filter was then hybridized by adding approximately 200 ng of each probe (see below) and by incubating overnight at temperature not exceeding 42°C in roller bottles at 6 rpm. Filter was then washed twice for 20 min at 42°C with a primary wash buffer (6 M urea, 0.4% SDS, 0.5× SSC) in the roller bottles and then with a secondary wash buffer (2× SSC) for no more than 15 min at room temperature on a shaking platform. Hybridization was then detected on autoradiographic films (Hyperfilm-ECL; Amersham) by using the enhanced chemiluminescence gene detection system (Amersham). Two probing sequences for the genes encoding for the 65 KDa antigen and the 85-C antigen (GenBank accession numbers, M15467 and X57229) were prepared from DNA of *M. tuberculosis* H37Rv by PCR using, respectively, primers HSP65-U and HSP65-L and primers 85-C-U and 85-C-L (see Table 2), as reported above. The 225 and 307 bp amplicons were purified by QIAquick Gel Extraction Kit (Qiagen, Germany) and covalently labeled with horseradish peroxidase (HRP) by glutaraldehyde according to the procedure of ECL Direct System (Amersham).

RESULTS

Selection of arbitrary primers. The 12-base arbitrary primers used in this study were designed according to the following criteria: (a) no stop codon was present in any possible reading frame in both directions; (b) the sequences had minimal tendency, if any, to form internal loops; (c) primers sequences included the most frequently used codons in 10 randomly-chosen *M. tuberculosis* genes and the most frequent codons found in 50 primers for computer-simulated PCRs for 5 randomly-chosen genes of *M. tuberculosis;* (d) the content of GC should be 67% (8 GC for each primer) and either G or C should be at 3′ position.

mRNA differential display in M. tuberculosis H37Rv and H37Ra. Total RNA, extracted from *M. tuberculosis* H37Rv and H37Ra and treated with DNase, was used as a substrate for oligo(dT)-primed, RT-dependent cDNA synthesis. The generated RT products were then PCR-amplified employing the 10 arbitrary primers in all the 55 pair combinations and the reaction products were separated by polyacrylamide gel electrophoresis, silver-stained and then photographed for computer-driven densitometric analysis (a representative example of this analysis is given in Fig. 1). As shown in Table 3, the 55 combinations of the arbitrary primers yielded a total number of 745 and 708 bands for *M. tuberculosis* H37Rv and H37Ra, respectively. A number of differentially generated bands from the RT-PCR products were detected in both mycobacterial strains: in particular, 52 bands were detected in the H37Rv, but not in the H37Ra strain; conversely, 15 differential bands were detected only in the avirulent strain. Six cDNAs that appeared to be expressed in H37Rv, but not in H37Ra, were reampli-

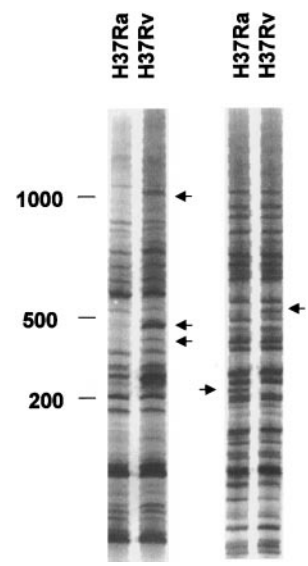

FIG. 1. mRNA differential display. Total RNA from *M. tuberculosis* H37Rv and H37Ra was transcribed by oligo(dT)-primed RT and the generated cDNA were PCR-amplified by arbitrary primers; the reaction products were then separated by polyacrylamide gel electrophoresis and silver-stained. A portion of a 12.5% polyacrylamide gel, obtained with arbitrary primers 7/9 (left panel) and 6/8 (right panel), is shown. Arrowheads indicate differentially amplified products. Molecular mass markers (in bp) are shown on the left.

TABLE 3

Differentially Generated Bands in *M. tuberculosis* H37Rv and H37Ra

	Number of bands		Number and size (bp) of differentially generated bands[a]	
Primer set	H37Rv	H37Ra	H37Rv	H37Ra
1/1	6	6		
1/2	18	18		
1/3	11	9	2 (910, 195)	
1/4	4	4		
1/5	15	12	3 (480, 405, 250)	
1/6	7	6	1 (190)	
1/7	14	12	2 (1080, 270)	
1/6	18	18		
1/9	16	16		
1/10	18	18	1 (560)	1 (410)
2/2	15	14	1 (760)	
2/3	14	12	2 (530, 235)	
2/4	11	11		
2/5	19	17	2 (525, 270)	
2/6	12	10	2 (570, 385)	
2/7	18	17	1 (710)	
2/8	23	24	1 (720)	2 (670, 345)
2/9	23	23		
2/10	30	29	1 (770)	
3/3	10	10		
3/4	5	4	1 (325)	
3/5	13	13		
3/6	16	14	2 (465, 345)	
3/7	14	10	4 (460, 360, 250, 195)	
3/8	18	16	2 (835, 600)	
3/9	13	12	1 (185)	
3/10	0	0		
4/4	10	10		
4/5	4	4		
4/6	3	3		
4/7	6	6		
4/8	10	10		
4/9	11	10	1 (530)	
4/10	10	10		
5/5	15	15		
5/6	16	16		
5/7	8	8		
5/8	16	16		
5/9	12	11	1 (105)	
5/10	11	11		
6/6	14	12	2 (790, 575)	
6/7	9	7	2 (850, 610)	
6/8	26	26	1 (520)	1 (210)
6/9	15	14	2 (2630, 750)	1 (650)
6/10	14	13	3 (675, 565, 360)	2 (500, 270)
7/7	17	16	3 (1330, 1110, 490)	2 (1050, 200)
7/8	18	18		
7/9	23	20	3 (1000, 460, 400)	
7/10	13	12	2 (695, 555)	1 (290)
8/8	15	19	1 (735)	5 (870, 790, 535, 390, 285)
8/9	14	13	1 (370)	
8/10	6	6		
9/9	19	18	1 (365)	
9/10	12	12		
10/10	17	17		
Total	745	708	52	15

[a] Size of bands, expressed in base pairs (bp), is given in parentheses.

fied and subcloned into pBluescript SK+ vector. Since different cDNA species can be contained theoretically within one eluted differentially expressed band (19), at least 10 colonies were selected from each cloned fragment. The corresponding DNA inserts were sequenced and a search for homologous nucleic acid sequences

TABLE 4

Sequence Homology between Cloned cDNAs Detected by mRNA DD and *M. tuberculosis* H37Rv Genes

cDNA		*M. tuberculosis* H37Rv genes			
Code	Length (bp)	Gene name[a] (accession number)	% of homology	Position of homologous sequence	Gene product[b]
2/3v-A	255	Rv2336 (Z83860)	92	315–532	Hypothetical protein
2/3v-B	151	Rv1345 (Z75555)	98	992–1142	Probable polyketide synthase
2/3v-C	198	Rv2770c (AL008967)	100	1–139	PPE protein
3/4v	247	Rv0288 (AL021930)	96	160–291	ESAT-6 family protein
7/10v	77	Rv1320c (Z73902)	95	365–424	Hypothetical protein
8/9v	375	Rv2819c (Z81331)	100	951–1128	Hypothetical protein

[a] Gene names are shown according to the nomenclature reported by Cole *et al.* (20).
[b] See text for references.

was performed in the databases of EMBL and Sanger Centre. Most of the over 60 inserts analyzed showed low or partial sequence homology with *M. tuberculosis* genes or corresponded to rRNA-encoding sequences. Six inserts showed high sequence homology with *M. tuberculosis* H37Rv genes. As summarized in Table 4, these inserts included:

(a) a 198-bp insert 100% homologous in a 139-nucleotide overlap to a sequence of the gene Rv2770c coding for a member of the PPE protein family, which includes 68 glycine-rich proteins with a conserved N-terminal domain of about 180 amino acids and C-terminal segments that vary in sequence and length. The name PPE derives from the motif Pro-Pro-Glu (PPE) found near the N-terminus in most cases (20).

(b) a 151-bp insert 98% homologous to a sequence of the gene Rv1345 coding for protein fadD33, a probable member of the polyketide synthase family that includes mycobacterial enzymes involved in the synthesis of the unusually large number of classes of lipids in these bacteria (21).

(c) a 247-bp insert 96% homologous in a 132-nucleotide overlap to a sequence of gene Rv0288, coding for a member of the protein family containing ESAT-6 (6-kDa early secretory antigenic target) (22, 23).

(d) three inserts of 255, 77, and 375 bp that showed sequence homology, respectively, with sequences of genes Rv2336, Rv1320c and Rv2819c, coding for proteins with unknown functions.

The differential expression of the 6 genes described above was confirmed by RNA dot blot and RT-PCR assays. In particular, 5.0 μg of total RNA preparations from strains H37Rv and H37Ra were spotted onto nylon filters and then probed for the putative differentially expressed transcripts by using peroxidase-labeled DNA sequences prepared from DNA of *M. tuberculosis* H37Rv by PCR using the specific primers listed in Table 2. Hybridization was then detected on autoradiographic films by chemiluminescence emission. A weak signal was obtained for each of the tested transcripts from RNA preparations derived from strain H37Rv and no signal was obtained from H37Ra RNA (data not shown). For the RT-PCR assay, 1 μg of total DNAse-treated RNA from strains H37Rv and H37Ra was transcribed into cDNA by oligo(dT)-primed RT; the reaction products were then probed by PCR employing pairs of specific primers for each of the putative differentially expressed genes (see Table 2). As shown in Fig. 2, a single band of the expected molecular size was detected for each of the tested mRNA transcripts from *M. tuberculosis* H37Rv, but a weak signal was also detected from strain H37Ra. By densitometric analysis, peak areas of the RT-PCR bands for genes Rv1345, Rv2770c, Rv2336, Rv1320c and Rv2819c of *M. tuberculosis* H37Ra ranged from 22 to 35% as compared to H37Rv; the expression of gene Rv0288, coding for a member of the ESAT-6 protein family, was also down-regulated in H37Ra, but to a lower extent (64%). These findings were highly reproducible.

Controls. Control assays were performed at various steps of the mRNA DD procedure. As controls of the successful cDNA synthesis from strains H37Rv and H37Ra, the oligo(dT)-primed RT reaction products of the extracted RNA, together with (*a*) samples of untreated cellular RNA, (*b*) oligo(dT)-primed RT-reaction products of RNase-treated RNA, and (*c*) RT-reaction products of RNA in the absence of oligo(dT)-priming, were probed for cDNAs of two constitutively expressed mycobacterial genes, *i.e.*, the 65 KDa antigen gene (24) and the 85-C antigen gene (25) by a PCR employing specific primers. As positive control, a sample of genomic DNA was PCR-amplified by the two sets of specific primers mentioned above. The expected 225 and 307 bp PCR products were detected when the cDNAs derived from oligo(dT)-primed RT-transcribed RNAs, but not from untranscribed RNA, or from RT-products of RNase-treated, oligo(dT)-primed RNA or from RT-products in the absence of oligo(dT)-priming (data not shown). As already reported (16), these con-

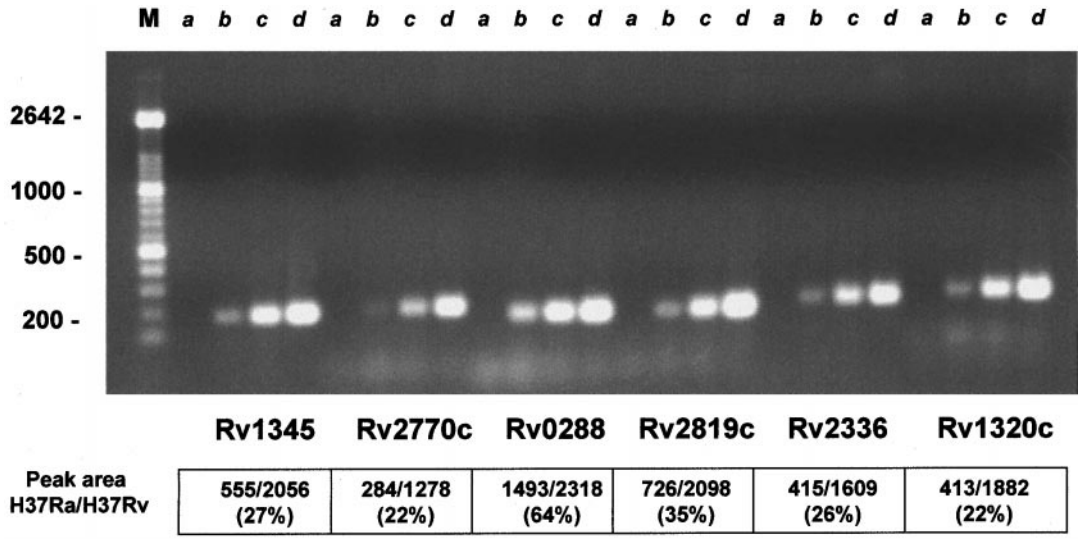

FIG. 2. Gene expression of in *M. tuberculosis* H37Rv and H37Ra assessed by RT-PCR. Total RNA from *M. tuberculosis* H37Rv and H37Ra was transcribed by oligo(dT)-primed RT and the generated cDNA were PCR-amplified by primers specific for the genes differentially expressed in the mRNA DD assay (see text and Table 2 for further details), indicated at the bottom of the panel. The RT-PCR products were separated by 2% agarose gel electrophoresis and stained with ethidium bromide. Lanes: M, molecular marker; *a*, PCR carried out in the absence of cDNA (H_2O); *b*, RT-PCR carried out with RNA preparation from H37Ra; *c*, RT-PCR carried out with RNA preparation from H37Rv; *d*, PCR carried out with *M. tuberculosis* H37Rv DNA. At the bottom of the figure, the peak areas of bands from H37Ra and H37Rv, expressed in gray level values by GelBase Pro software, are reported.

trol data indicate the successful oligo(dT)-primed RT-mediated synthesis of cDNA for mycobacterial mRNAs and, at the same time, rule out contamination of RNA samples with genomic DNA.

Further experiments were run to control whether the battery of the selected arbitrary primers was appropriate for the study of mycobacterial gene expression; in particular, a dot-blot hybridization analysis was performed, in which all the 55 products of oligo(dT)-primed RT-PCRs obtained from H37Rv RNA were spotted onto nylon filters and then probed for the mycobacterial cDNAs of the antigens 65 KDa and 85-C. As shown in Fig. 3, the probe for the 65 KDa antigen gene strongly hybridized with RT-PCR products obtained with primer pair 8-9, and with primer 9 alone (used as up- and down-stream primer) and, also, although weakly, with RT-PCR products obtained with primer pairs 2-9, and 7-9. Probe for the 85-C antigen gene, on the contrary, failed to detect RT-PCR products derived from any arbitrary primer pair combination (not shown). These results indicate that the selected primers are able to detect part of the genes expressed in *M. tuberculosis.*

DISCUSSION

The original mRNA DD protocol uses a two-base anchored oligo-dT primer for RT-dependent cDNA synthesis and the same primer with an arbitrary 10-base oligonucleotide for PCR amplification (12). This approach has a significant limitation as it identifies cDNAs biased to the 3′-ends of the corresponding mRNAs, often within untranslated regions (12, 26). On the other hand, due to the general scepticism on the existence of polyadenylated tracts in bacterial mRNAs that can be targeted to initiate oligo(dT)-primed RT synthesis of cDNA, the very few reported applications of DD techniques to prokaryotes make use of arbitrary primers for cDNA synthesis (27, 28, 29), which may lead to the generation of many false-positive results that need further time-consuming controls. The results reported in our study show that the mRNA DD method, based on the oligo(dT)-primed RT-dependent transcription of poly-A tailed mRNAs and a PCR amplification of the RT products by pairs of 12-base arbitrary primers, can be successfully applied to the investigation of gene expression in *M. tuberculosis.* Moreover, the mRNA DD strategy used in our study allows the cloning of cDNAs complementary to portions of mRNAs other than just their 3′-termini and PCR reactions can be run at a relatively high annealing temperature (*i.e.*, 50°C), which produces clear banding patterns without background smear due to continuous low-stringency priming of oligonucleotides that do not match perfectly.

By our mRNA DD method, 6 genes that appear to be markedly downregulated in the avirulent H37Ra strain have been detected. The detection of these genes stems from the analysis of a limited number of the many differentially generated bands and from the use of a panel of arbitrary primers that, as shown in the control experiments, do not guarantee the detection of the complete *M. tuberculosis* gene repertoire. It has

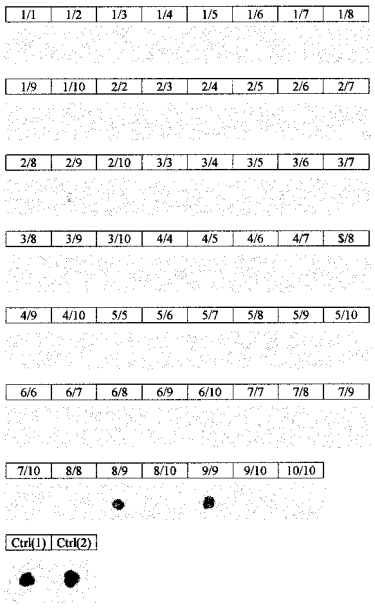

FIG. 3. Dot-blot hybridization. The RT-PCR products, obtained from H37Rv RNA with the 10 arbitrary primers in all the 55 combinations, were spotted onto nylon filters and then hybridized by addition of a peroxidase-labeled probe for the gene of antigen 65 KDa, prepared by PCR with specific primers (see Table 2) from DNA of *M. tuberculosis* H37Rv. Hybridization was then detected on autoradiographic films by the enhanced chemiluminescence gene detection system. Numbers at the top of the strips indicate the abritary primers pairs; positive controls consisted of a spotted RT-PCR-product, obtained from total RNA from *M. tuberculosis* H37Rv transcribed by oligo(dT)-primed RT and PCR-amplified by primers specific for the 65 KDa antigen gene [Ctrl(1)], and the unlabeled 65 KDa antigen gene probe itself [Ctrl(2)].

therefore to be expected that the cloning and sequencing of other differential cDNAs or the screening of gene expression with novel arbitrary primers may lead to detection of more genes that are downregulated in the avirulent H37Ra strain. These considerations, together with the experimental data, suggest that the mutation H37Rv → H37Ra affects the regulatory repertoire of *M. tuberculosis* gene expression, an extensive and complex apparatus that includes many different promotors, 13 putative σ factors and over 100 transcriptional regulatory proteins (20). In agreement with this hypothesis, in fact, Collins and co-workers (30) reported that a mutation in the principal σ factor caused the loss of virulence in a strain of the *M. tuberculosis* complex. It is, however, stricking that 3 of the 6 genes detected by mRNA DD code for proteins that have been predictively associated, although to a different extent, to immunological or pathogenetic aspects of *M. tuberculosis* infection (20), in particular: ESAT-6, a potent T-cell antigen, unique to pathogenic mycobacterial species (22, 23) recognized by a high proportion of protective T cells in animal models (31); the PPE proteins, predictively abundant and antigenically polymorphic, that are thought to be of immunological significance, possibly as a source of antigenic variation or by inhibiting antigen processing (20); and, finally, the polyketide synthase, that is also thought to play a role in *M. tuberculosis* infection or disease, possibly by the synthesis of new lipids and polyketide metabolites (20).

By transfecting and expressing the genes potentially involved in *M. tuberculosis* virulence in other non-pathogenic bacteria, it will now be possible to dissect *M. tuberculosis* virulence properties and to study the candidate virulence factors at a molecular level.

ACKNOWLEDGMENTS

This work was supported by The National Research Program on AIDS (1997) and Grant 50A.0.17 from the Istituto Superiore di Sanità, Italy. We thank Drs. Mauro Pistello and Giulia Freer for careful reading of the manuscript.

REFERENCES

1. Jacobs, W. R., Kaplana, G. V., Cirillo, J. D., Pascopella, L., Snapper, S. B., Udani, R. A., Jones, W., Barletta, R. G., and Bloom, B. R. (1991) *Methods Enzymol.* **204,** 537–555.
2. Hatfull, G. F. (1993) *Trends Microbiol.* **1,** 310–314.
3. Jacobs, W. R., and Bloom, B. R. (1994) *in* Tuberculosis: Pathogenesis, Protection, and Control (Bloom, B. R., Ed.), pp. 253–268, ASM, Washington, DC.
4. Clark-Curtiss, J. E. (1998) *Curr. Top. Microbiol. Immunol.* **225,** 57–79.
5. Arruda, S., Bonfim, G., Knights, R., Huima-Byron, T., and Riley, L. W. (1993) *Science,* **261,** 1454–1457.
6. Berthet, F.-X., Lagranderie, M., Gounon, P., Laurent-Winter, C., Ensergueix, D., Chavarot, P., Thouron, F., Maranghi, E., Pelicic, V., Portnoi, D., Marchal, G., and Gicquel, B. (1998) *Science* **282,** 759–762.
7. Steenken Jr., W., and Gardener, I. U. (1946) *Am. Rev. Tuberc.* **54,** 62–66.
8. Steenken Jr., W., Oatway, W. H., and Petroff, S. A. (1934) *J. Exp. Med.* **60,** 515–540.
9. Pascopella, L., Collins, F. M., Martin, J. M., Jacobs Jr., W. R., and Bloom, B. R. (1993) *Infect. Agents Dis.* **2,** 282–284.
10. Pascopella, L., Collins, F. M., Martin, J. M., Lee, M. H., Hatfull, G. F., Stover, C. K., Bloom, B. R., and Jacobs Jr., W. R. (1994) *Infect. Immun.,* **62,** 1313–1319.
11. Kinger, A. K., and Tyagi, J. S. (1993) *Gene* **131,** 113–117.
12. Liang, P., and Pardee, A. B. (1992) *Science* **257,** 967–971.
13. Cohen, S. N. (1995) *Cell* **80,** 829–832.
14. Sarkar, N. (1996) *Microbiology* **142,** 3125–3133.
15. Sarkar, N. (1997) *Ann. Rev. Biochem.* **66,** 173–197.
16. Rindi, L., Lari, N., Garcia Gil, M., and Garzelli, C. (1998) *Biochem. Biophys. Res. Commun.* **248,** 216–218.
17. Chomczynsky, P., and Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159.
18. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
19. Liang, P., and Pardee, A. B. (1995) *Curr. Op. Immunol.* **7,** 274–280.
20. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E. 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth,

T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Barrell, B. G., et al. (1998) *Nature* **393,** 537–544.

21. Kolattukudy, P. E., Fernandes, N. D., Azad, A. K., Fitzmaurice, A. M., and Sirakova, T. D. (1997) *Mol. Microbiol.* **24,** 263–270.
22. Harboe, M., Oettinger, T., Wiker, H. G., Rosenkrands, I., and Andersen, P. (1996) *Infect. Immun.,* **64,** 16–22.
23. Sørensen, A. L., Nagai, S., Houen, G., Andersen, P., and Andersen, A. B. (1995) *Infect. Immun.* **63,** 1710–1717.
24. Shinnick, T. M. (1987) *J. Bacteriol.* **169,** 1080–1088.
25. Content, J., de la Cuvellerie, A., De Wit, L., Vincent-Levy-Frébault, V., Ooma, J., and De Bruyn, J. (1991) *Infect. Immun.* **59,** 3205–3212.
26. Liang, P., Averboukh, L., and Pardee, A. P. (1993) *Nucleic Acids Res.* **21,** 3269–3275.
27. Fislage, R., Berceanu, M., Humboldt, Y., Wendt, M., and Oberender, H. (1997) *Nucleic Acids Res.* **25,** 1830–1835.
28. Kwaik, Y. A., and Pederson, L. L. (1996) *Mol. Microbiol.* **21,** 543–556.
29. Wong, K. K., and McClelland, M. (1994) *Proc. Natl. Acad. Sci. USA,* **91,** 639–643.
30. Collins, D. M., Kawakami, R. P., de Lisle, G. W., Pascopella, L., Bloom, B. R., and Jacobs Jr., W. R. (1995) *Proc. Natl. Acad. Sci. USA,* **92,** 8036–8040.
31. Pais, T. F., Silva, R. A., Smedegaard, B., Appelberg, R., and Andersen, P. (1998) *Immunology,* **95,** 69–75.