JOURNAL OF BACTERIOLOGY, Nov. 1993, p. 7282–7289
0021-9193/93/227282-08$02.00/0

Vol. 175, No. 22

The *uraA* Locus and Homologous Recombination in *Mycobacterium bovis* BCG

ANNA ALDOVINI, ROBERT N. HUSSON,† AND RICHARD A. YOUNG*

Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received 29 July 1993/Accepted 20 September 1993

Molecular genetic manipulation of mycobacteria would benefit from the isolation of mycobacterial genes that could serve both as genetic markers and as sequences used to target homologous integration of recombinant DNA into the genome. We isolated the *Mycobacterium bovis* BCG gene encoding orotidine-5′-monophosphate decarboxylase (OMP-DCase) by complementing an *Escherichia coli* mutant defective in this activity. The BCG OMP-DCase gene (*uraA*) and the flanking DNA were sequenced. The predicted BCG OMP-DCase protein sequence is closely related to the *Myxococcus xanthus* OMP-DCase and more distantly related to the other known prokaryotic and eukaryotic OMP-DCases. To investigate whether homologous integration can occur in *M. bovis* BCG, an improved protocol for transformation of BCG was developed and a linear fragment of mycobacterial DNA containing the *uraA* locus, marked with a kanamycin resistance gene, was introduced into BCG cells by electroporation. The kanamycin-resistant BCG transformants all contained vector DNA integrated into the genome. The marked DNA had integrated into the homologous *uraA* locus in approximately 20% of the transformants. These results have implications for understanding the role of mycobacterial genes in disease pathogenesis and for the genetic engineering of improved mycobacterial vaccines.

The World Health Organization estimates that one in three humans is infected with *Mycobacterium tuberculosis*, the etiologic agent of tuberculosis (3, 27). Over the past decade, there has been a recent resurgence in the incidence of tuberculosis in developed countries that has coincided with the AIDS epidemic (25). Because of their impact as major human pathogens and as a result of their profound immunostimulatory properties, mycobacteria have long been intensively studied. In the early 1900s, an attenuated mycobacterium, *M. bovis* bacille Calmette-Guérin (BCG), was isolated for use as a vaccine against tuberculosis (6, 7). Although the efficacy of this vaccine against tuberculosis varied considerably in different trials, and the reasons for its variable efficacy have yet to be resolved, BCG is among the most widely used human vaccines (10, 17).

The recent application of molecular biological technology to the study of mycobacteria has led to the identification of many of the major antigens that are targets of the immune response to infection by mycobacteria (14, 32, 34, 35) and to an improved understanding of the molecular mechanisms involved in resistance to antimycobacterial antibiotics (28, 37). The development of tools that permit molecular genetic manipulation of mycobacteria has also allowed the construction of recombinant BCG vaccine vehicles (1, 8, 11, 12, 16, 18, 23, 24, 26, 31). Genome mapping and sequencing projects are providing valuable information about the *M. tuberculosis* and *M. leprae* genomes that will facilitate further study of the biology of these pathogens (9, 33).

Despite these advances, there are two serious limitations to our ability to manipulate these organisms genetically. (i) Very few mycobacterial genes that can be used as genetic markers have been isolated (8). (ii) Investigators have failed to obtain homologous recombination in slowly growing mycobacteria such as *M. tuberculosis* and *M. bovis* BCG (13, 33), despite being able to accomplish this in the fast-growing species *M. smegmatis* (11). The ability to achieve homologous recombination in slowly growing mycobacteria would provide an important approach to understanding the pathogenesis of tuberculosis and would permit construction of attenuated strains of *M. tuberculosis* for vaccine use.

The gene for orotidine-5′-monophosphate decarboxylase (OMP-DCase) is widely used as a genetic marker. We report here the isolation and characterization of *M. bovis* BCG DNA containing the OMP-DCase gene (*uraA*) and its use to obtain integration of a recombinant plasmid into homologous DNA in the BCG genome.

MATERIALS AND METHODS

Strains and plasmids. The *M. bovis* BCG strain used for DNA isolation and subsequent construction of the recombinant BCG plasmid and λgt11 libraries was the Montreal Strain, ATCC 35735. *M. bovis* BCG was grown in Middlebrook 7H9 medium supplemented with 0.05% Tween 80 as described in Aldovini and Young (1). *E. coli* Y1107 (*pyrF*::Mu *trp*(Am) *lacZ*(Am) *hsdR* m+ su−) was obtained from D. Botstein. Plasmids were propagated in *E. coli* DH5α from Bethesda Research Laboratories. *E. coli* cultures used for plasmid selection were grown in Luria-Bertani broth or agar with 50 μg of ampicillin per ml. Phage M13, used for production of single-stranded DNA, was propagated in *E. coli* JM101 from New England BioLabs. JM101 was grown in YT medium (2). Genomic libraries were generated with pUC19 from Bethesda Research Laboratories. Plasmid pY6002 (11) was the source of the 1.3-kb *Bam*HI DNA fragment containing aminoglycoside phosphotransferase gene *aph*.

Enzymes. The Klenow fragment of *E. coli* DNA polymerase was supplied by Promega. T7 polymerase and *Taq* polymerase (Sequenase and Taquence) were provided by United States Biochemical.

Recombinant DNA library construction. To isolate BCG DNA, cells were harvested by centrifugation, washed, and

* Corresponding author.
† Present address: Division of Infectious Diseases, Children's Hospital, Boston, MA 02115.

resuspended in 50 mM Tris (pH 8.0)–10 mM EDTA–10% sucrose–0.5 mg of lysozyme per ml and incubated at 37°C for 1 h. EDTA was then added to 1%, and the mixture was incubated at room temperature for 15 min. Three phenol-chloroform extractions were performed and followed by RNase treatment, phenol-chloroform extraction, chloroform extraction, and ethanol precipitation. The DNA was then suspended in TE buffer (10 mM Tris [pH 7.5], 1 mM EDTA).

To construct the plasmid library, DNA was subjected to partial digestion with *Sau*3A and DNA fragments of 2 to 6 kb were isolated by agarose gel electrophoresis onto NA45 DEAE-paper and eluted in buffer containing 10 mM Tris HCl, 1 M NaCl, and 1 mM EDTA. The DNA fragments were then phenol-chloroform extracted, ethanol precipitated, and ligated into a *Bam*HI-digested, calf intestinal phosphatase-treated pUC19 plasmid vector. *E. coli* cells were transformed with the ligated mixture, and approximately 4×10^5 recombinants were obtained. Plasmid DNA was obtained from the pool of transformed colonies by an alkaline lysis method.

The λgt11 library was constructed by using a procedure described by Young et al. (36). Briefly, BCG genomic DNA was subjected to random partial digestion with DNase I, *Eco*RI linkers were added to the digestion products, and DNA fragments of 4 to 8 kb were isolated by agarose gel electrophoresis and electroelution. The DNA fragments were then ethanol precipitated and ligated into *Eco*RI-digested λgt11 arms. The ligation mixture was packaged into λ heads, and the packaging mixture was used to infect *E. coli*. Approximately 5×10^6 recombinants were obtained.

Isolation of the BCG OMP-DCase gene by complementation and plasmid DNA manipulation. The BCG recombinant plasmid library was used to transform *E. coli* Y1107. We isolated 21 transformants capable of growing in the absence of uracil and chose 6 for further evaluation by restriction analysis. Plasmid DNA was isolated by alkaline lysis from cells grown in liquid culture, and restriction analysis indicated that all of these plasmids contained the same or very similar insert DNAs. One of these clones (pY6006) was used for further study (see Fig. 1). A 0.6-kb *Bam*HI DNA fragment from pY6006 was used to screen the λgt11 library, leading to the isolation of phage Y3030. This phage carries a 5.6-kb *Eco*RI BCG DNA insert containing the OMP-DCase gene. This insert DNA was subcloned into pGEM7z(f+) to generate pY6011. The 4.4-kb *Sac*I-*Eco*RI fragment of the Y3030 insert was subcloned into pUC19 to generate pY6014. Plasmid pY6015 was derived from pY6014 by replacing *uraA* sequences with the *aph* gene; a 1.15-kb *Hin*cII DNA fragment containing *uraA* sequences was removed by partial *Hin*cII digestion of pY6014 DNA and replaced with a 1.3-kb *Bam*HI fragment containing *aph* from pY6002 that was blunt ended with Klenow.

DNA sequence analysis. The *M. bovis* BCG *uraA* gene was sequenced from the 4.4-kb *Sac*I-*Eco*RI fragment of λgt11 phage Y3030 cloned into M13 in both orientations. The same DNA fragment was subcloned into pUC19 to generate pY6014 for further manipulation. Single-stranded DNA for sequence analysis was prepared from M13 grown in JM101 (30). Both DNA strands were sequenced by the dideoxy method (2). Mycobacterial DNA has a high G+C content, and two different strategies were used to reduce band compression and other artifacts due to high G+C content. A subset of the reactions was carried out with *Taq* polymerase at a high temperature (70°C). In addition, dGTP and dITP were used in independent sequencing reactions (15).

BCG transformation. BCG Pasteur (American Type Culture Collection) was grown in the log phase to an optical density at 600 nm of 0.5 in Middlebrook medium. BCG cells were harvested by centrifugation and washed twice with phosphate-buffered saline (PBS) and suspended in 1 mM MgCl (pH 7.2)–10% sucrose–15% glycerol at a concentration of 10 optical density units at 600 nm per ml. A 0.4-ml volume of BCG cells was mixed with 2 μg of plasmid DNA and electroporated in a 0.2-cm cuvette. Electroporation settings were 2.5 kV of potential and 25 μF of capacitance. After electroporation, cells were resuspended in 10 ml of Middlebrook medium and incubated at 37°C for 2 h before plating on Middlebrook agar containing 20 μg of kanamycin per ml.

Southern blot analysis. Genomic DNAs from BCG strains were isolated as described above, digested with restriction enzymes, subjected to agarose gel electrophoresis in the presence of ethidium bromide, transferred to nitrocellulose, and probed with DNA labelled with ^{32}P by random priming, all by standard procedures (2).

Nucleotide sequence accession number. The nucleotide sequence reported here has been provided to GenBank under accession number U01072.

RESULTS

Isolation of the BCG OMP-DCase gene by genetic complementation. The complementation strategy employed to isolate the BCG OMP-DCase gene was similar to that employed previously to isolate the homologous gene in *M. smegmatis* (11). A recombinant library was constructed in *E. coli* vector pUC19 by using size-selected BCG genomic DNA fragments from a partial *Sau*3A digest. An *E. coli pyrF* mutant strain (Y1107) was transformed with this library, and cells were plated on medium lacking uracil to select for uracil prototrophs and on rich medium containing ampicillin to ascertain the transformation frequency and estimate the fraction of transformants able to complement the *E. coli pyrF* defect. Approximately 0.05% of the cells transformed with the recombinant library became uracil prototrophs. DNA clones were obtained from six colonies able to grow in the absence of uracil, and restriction analysis revealed that these clones contained the same insert DNA. One of these clones, pY6006, was subjected to further study (Fig. 1).

To identify the portion of the 3.5-kb insert DNA of pY6006 that was responsible for complementation, the 1.3-kb *Bam*HI fragment of Tn*903*, which encodes aminoglycoside transferase (*aph*), was inserted into several different sites in pY6006 insert DNA, the resultant plasmids were reintroduced into the *E. coli pyrF* mutant strain, and the ability of the new plasmids to complement the mutant phenotype was assessed as before (Fig. 1). One of the three plasmids with insertion mutations, pY6006B, lost the ability to complement the *pyrF* mutant phenotype, suggesting that sequences necessary for the complementing activity are located in the vicinity of the *Bam*HI site that is disrupted in pY6006B.

Analysis of DNA sequences for the left end of pY6006 insert DNA (as diagrammed in Fig. 1) revealed that the open reading frame of the pUC19 *lacZ* gene in this plasmid continues uninterrupted into an open reading frame for a polypeptide with a sequence similar to that of OMP-DCase proteins. These preliminary data suggested that the left end of the pY6006 insert DNA encoded the amino terminus of the BCG OMP-DCase protein. For later experiments, it was important to have both the OMP-DCase gene and a substantial amount of flanking sequences. To obtain genomic DNA that contains both the OMP-DCase gene and its flanking sequences, the 0.6-kb *Bam*HI DNA fragment from pY6006 was used to probe a λgt11 library of *M. bovis* BCG DNA, as the λgt11 library

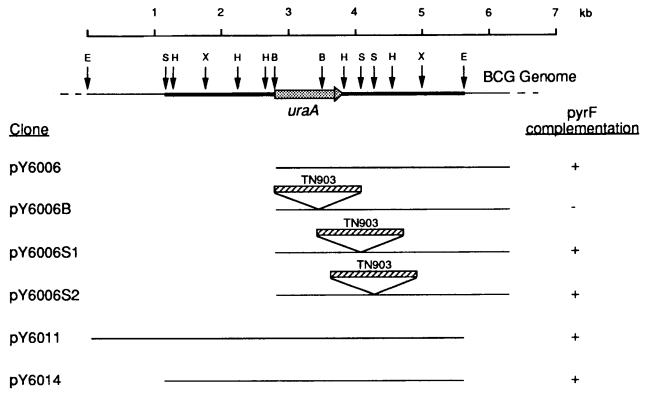

FIG. 1. Structural and functional map of the BCG *uraA* locus. A restriction map of the *uraA* locus and the recombinant insert DNAs for several plasmids that were used to study this region is depicted. The relative positions of the BCG *uraA* gene and the portions of other genes identified through sequence analysis are summarized graphically. Plasmid pY6006 was isolated by complementation of the *E. coli pyrF* mutant, and plasmids pY6006B, pY6006S1, and pY6006S2 were constructed with insertions of the *aph* gene from Tn*903* to obtain clues to the position of the complementing BCG gene. Plasmid pY6011 and pY6014 insert DNAs were obtained from λgt11 clone Y3030 as described in the text. The ability of each of the recombinants to complement the *E. coli pyrF* mutant is indicated. The portion of the *uraA* locus subjected to sequence analysis spans the 4.3 kb between the leftmost *Sac*I site and the *Eco*RI site at the right end of the map. B, *Bam*HI; E, *Eco*RI; H, *Hin*cII; Hin, *Hin*dIII; S, *Sac*I; X, *Xba*I.

contains insert DNA fragments whose size, on average, is larger (4 to 8 kb) than the plasmid library used to obtain pY6006. A lambda clone (Y3030) which contains a 5.6-kb *Eco*RI DNA insert that overlaps that of pY6006 was isolated. The 5.6-kb *Eco*RI DNA fragment and a 4.4-kb *Sac*I-*Eco*RI subfragment were subcloned into plasmid vectors to generate pY6011 and pY6014, respectively (Fig. 1). Both pY6011 and pY6014 were able to complement the defect of *pyrF* mutant *E. coli* Y1107.

Sequences of the BCG OMP-DCase gene and flanking DNA. DNA fragments from phage Y3030 insert DNA were subcloned into M13 vectors and subjected to sequence analysis. The sequences of both DNA strands were determined, and most of the sequence reactions were duplicated with ITP replacing GTP to minimize artifacts due to the G+C-rich nature of mycobacterial DNA. Figure 2 shows the sequences obtained for the BCG OMP-DCase gene (*uraA*) and the flanking DNA. The predicted BCG OMP-DCase protein sequence is 274 amino acids long, similar in size to other OMP-DCase proteins. When the BCG-DCase protein sequence was used to screen the available data bases for similar sequences, the results revealed that the BCG protein is closely related to the *M. xanthus* OMP-DCase (15) and more distantly related to the other known prokaryotic and eukaryotic OMP-DCases. Comparison of the BCG and *M. xanthus* OMP-DCases revealed that 40% of the amino acid residues are identical (Fig. 3). In contrast, only 17% of the residues of the BCG and *E. coli* proteins and 22% of the amino acids of the *M. xanthus* and *E. coli* proteins are identical, although there are a substantial number of conservative amino acid substitutions among these proteins. The relationship of *M. xanthus* OMP-DCase to homologs in other prokaryotes and in eukaryotes was recently described in some detail (15). This comparative sequence analysis revealed four regions which are more highly conserved, and the predicted BCG OMP-DCase also shares this feature with the other homologs. It is interesting that mycobacteria and myxococci both have G+C-rich genomes, but this alone does not account for the degree of sequence conservation between the OMP-DCases from these two prokaryote genera; rather, the two genera appear to be more closely related to one another than either is to the other prokaryotes for which OMP-DCase sequences are available.

Further analysis of the BCG genomic DNA sequences revealed that the 1.7-kb sequence upstream of the OMP-DCase coding sequences contains a single large open reading frame. This open reading frame has no apparent beginning in the cloned DNA fragment, suggesting that it is the coding sequence for the carboxy terminus of a larger protein. A screen of the sequence data base revealed that the 497 amino acid residues of the predicted protein are highly homologous to the carboxyl termini of the large subunit of carbamoyl phosphate synthase. For example, the 497-amino-acid carboxy terminus of the putative *M. bovis* BCG protein was 46% identical to the comparable segment of the *E. coli* carbamoyl phosphate synthase subunit encoded by the *carB* gene (20). Thus, the BCG *carB* gene appears to be located just upstream of *uraA*. This is interesting, because both carbamoyl phosphate synthase and OMP-DCase are involved in pyrimidine biosynthesis. Carbamoyl phosphate synthase catalyzes the first reaction in pyrimidine biosynthesis, production of carbamoyl phosphate, while OMP-DCase catalyzes the last step in the biosynthesis of UMP.

Analysis of BCG DNA sequences downstream of the *uraA* gene revealed a single large open reading frame that continues through the right end of the sequenced DNA fragment. This open reading frame predicts a protein of 501 amino acids. A search of the computer data base revealed that the protein predicted by this open reading frame is similar to previously described proteins from *M. tuberculosis* and *M. leprae* (Fig. 4). The predicted BCG protein is similar to a putative *M. tuberculosis* antigen encoded downstream of the gene for the 65-kDa antigen (22) and to an *M. leprae* antigen that may be an integral membrane protein (29).

Southern analysis with whole genomic DNA revealed a single copy of the *uraA* gene and flanking DNA in the BCG genome (see below). The relative positions of the BCG *uraA* gene and the portions of other genes identified through sequence analysis are summarized graphically in Fig. 1. The position of OMP-DCase sequences is consistent with the genetic analysis described above. The *aph* insertion mutations in plasmid pY6006 that adversely affected complementation of the *E. coli* OMP-DCase mutant occurred within OMP-DCase coding sequences. Conversely, the *aph* insertion mutations that did not affect complementation of the *E. coli* OMP-DCase mutant occurred outside of the BCG OMP-DCase coding sequences.

Introduction of foreign DNA into the BCG genome. Previous attempts to obtain homologous recombination in *M. bovis* BCG have apparently not been successful (13, 32). It is possible that the efficiency of transformation has an influence on the ability to obtain homologous recombination. To maximize the transformation efficiency of BCG, we investigated the effect of adding glycine to the culture medium prior to harvesting cells for electroporation, as the presence of 1.5% glycine can affect the integrity of the cell wall and it seems to improve transformation efficiency in *M. smegmatis* (19). In addition, we compared the efficiency of electroporation of BCG cells in water relative to that in buffer. Autonomously replicating plasmid pYUB12 (23) was used to determine how

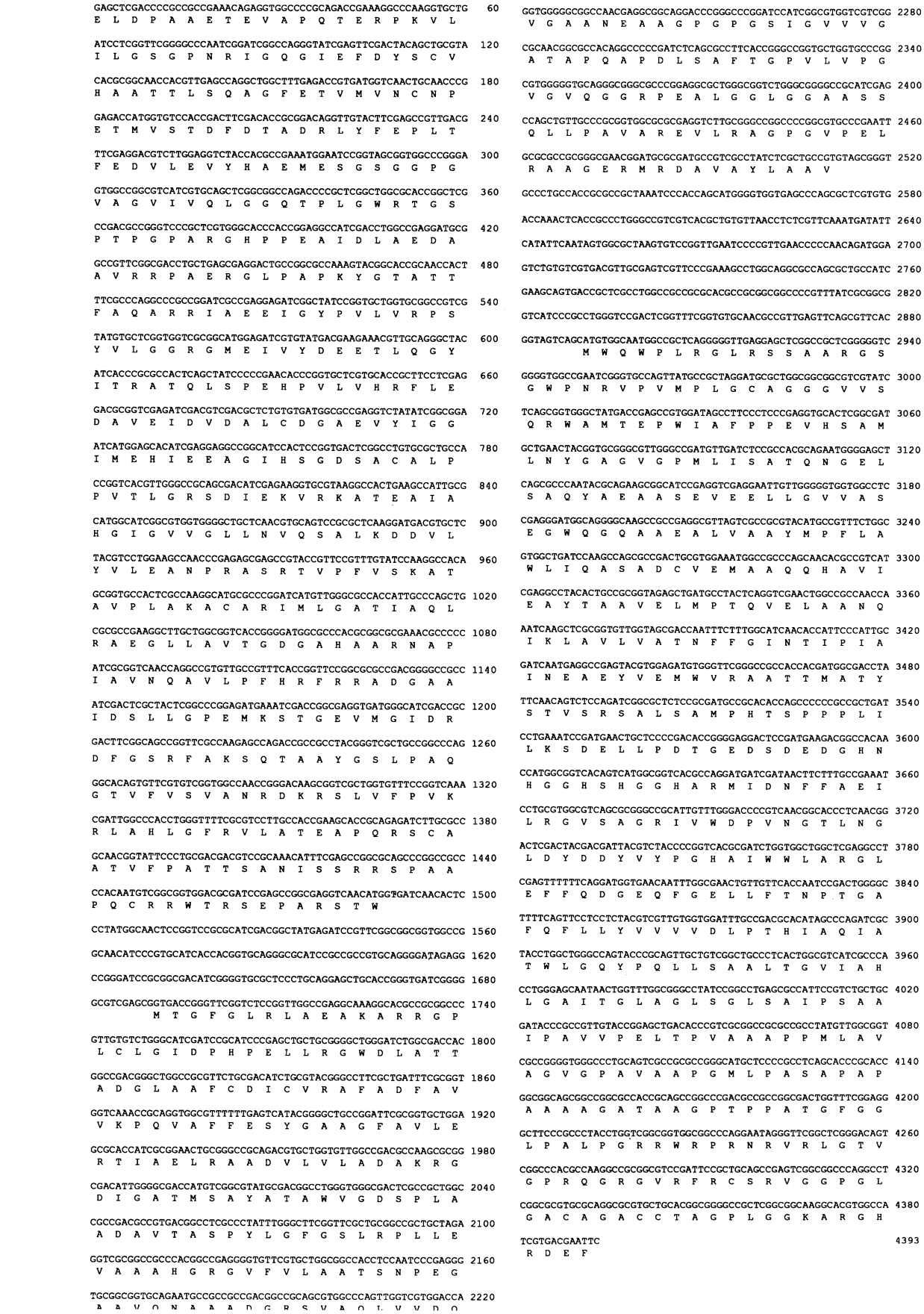

FIG. 2. Sequences of the BCG *uraA* locus and the predicted protein products. The DNA sequence was obtained from the same BCG DNA fragment used to generate pY6014 (Fig. 1).

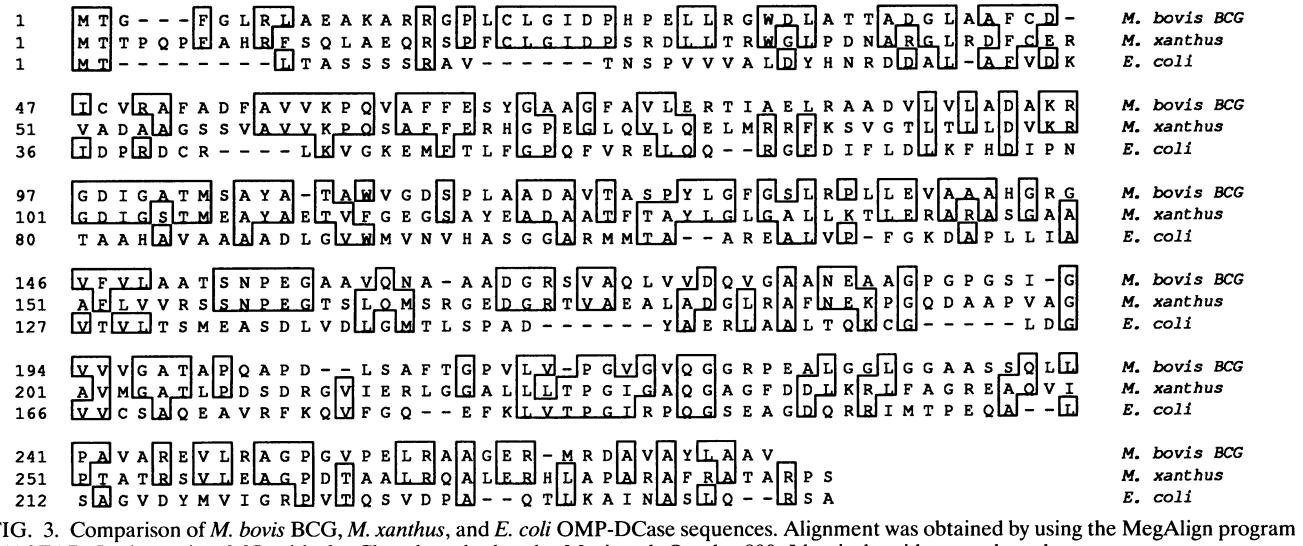

FIG. 3. Comparison of *M. bovis* BCG, *M. xanthus*, and *E. coli* OMP-DCase sequences. Alignment was obtained by using the MegAlign program (DNASTAR, Inc.), version 0.97, with the Clustal method and a Macintosh Quadra 800. Identical residues are boxed.

these variables affect the relative efficiencies of transformation. The results are summarized in Table 1 under experiment 1. Transformation efficiencies were improved substantially by exposing cultures to 1.5% glycine for 24 h prior to harvest and by performing the electroporation in water rather than in buffer.

Experiments with linearized DNA molecules in *Saccharomyces cerevisiae* indicate that the ends of linear DNA molecules

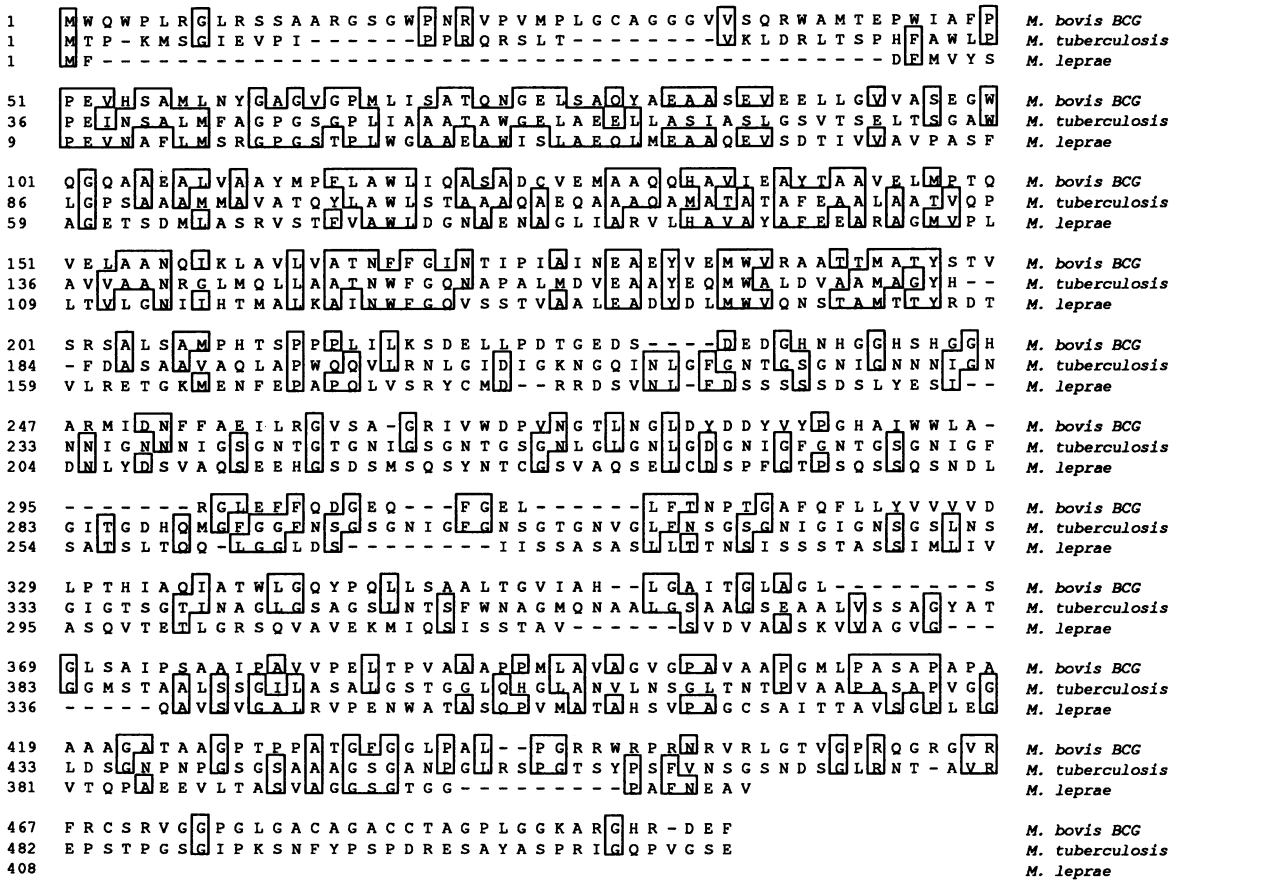

FIG. 4. Comparison of the *M. bovis* BCG protein predicted by the open reading frame downstream of *uraA* with similar proteins from *M. tuberculosis* and *M. leprae*. Alignment was obtained with the MegAlign program (DNASTAR, Inc.), version 0.97, and the Clustal method. Identical residues are boxed.

TABLE 1. BCG transformation efficiencies

Transforming DNA[a]	Glycine treatment[b]	Electroporation medium[c]	No. of transformants/μg of DNA		
			Expt 1	Expt 2	Expt 3
pYUB12	−	Buffer	50		
pYUB12	+	Buffer	250		
pYUB12	−	Water	500		
pYUB12	+	Water	10^4	10^4	10^5
None	+	Water	8	6	35
p6015(I)	−	Buffer		4	
p6015(I)	+	Buffer		22	
p6015(I)	−	Water		39	
p6015(I)	+	Water		98	500

[a] The intact autonomously replicating plasmid pYUB12 was used as a control, and the linear insert DNA of plasmid pY6015 [pY6015(I)] was used as integrating DNA.
[b] Glycine was added to 1.5% to BCG cultures 24 h prior to transformation.
[c] The buffer was 1 mM MgCl (pH 7.2)–10% sucrose–15% glycerol.

are recombinogenic; these ends may facilitate homologous integration by invading genomic DNA at homologous sites to initiate recombination (21). The sequenced 4.4-kb BCG DNA fragment containing *uraA* was used to investigate whether cloned DNA sequences could integrate at the homologous locus in *M. bovis* BCG. To mark the DNA fragment, the OMP-DCase coding sequence was replaced with a kanamycin resistance gene (*aph*) to create pY6015 (Fig. 5A). This left intact approximately 1.5 kb of DNA flanking both ends of *uraA* that could be used to direct homologous integration. The transformation experiment described above for plasmid pYUB12 was repeated with pY6015 insert DNA, and the results are summarized in Table 1 under experiment 2. Again, transformation efficiencies were improved substantially by exposing cultures to 1.5% glycine for 24 h prior to harvest and by performing the electroporation in water rather than in buffer. However, because the transformation efficiencies obtained with the linear DNA were low, we made one additional attempt to improve these efficiencies.

Cultures of *M. bovis* BCG and other slowly growing mycobacteria contain large numbers of cells that are inviable or that have an exceedingly long lag time after plating. Some investigators have suggested that mycobacterial cells have an unusual ability to enter and maintain a dormant state, even when nutrients are available (33). We reasoned that maintenance of BCG cultures in mid-log growth might maximize the fraction of cells that were undergoing DNA synthesis and thus were competent to take up DNA and incorporate it into homologous sites in the genome. A third experiment was performed, in which BCG cultures were diluted approximately 1:4 every 2 days over a 2-month period to ensure persistent log-phase growth before transformation. The results in Table 1 indicate that this approach produced a significant increase in the number of transformants obtained with either the autonomously replicating vector or the linear DNA fragment.

Ten of the BCG colonies obtained in the third experiment were selected for further study after growing to adequate size for picking (24 days after plating). The 10 transformants were colony purified, and DNA was prepared from each. DNA preparations from the wild-type strain and the 10 transformants were digested with a variety of restriction endonucleases, and Southern analysis revealed that the kanamycin-resistant BCG transformants all contained vector DNA integrated into the genome (data not shown). In 2 of the 10 transformants, the transforming DNA had integrated at the homologous locus. Figure 5 shows representative results from a Southern analysis of the wild-type strain and one of the BCG recombinants in which the cloned DNA integrated at the homologous locus.

The results shown in Fig. 5 indicate that a single recombination event led to introduction of the transforming DNA into the *uraA* locus, as there was no replacement of genomic DNA. Two recombination events would have been required to replace the endogenous *uraA* DNA with the transforming DNA.

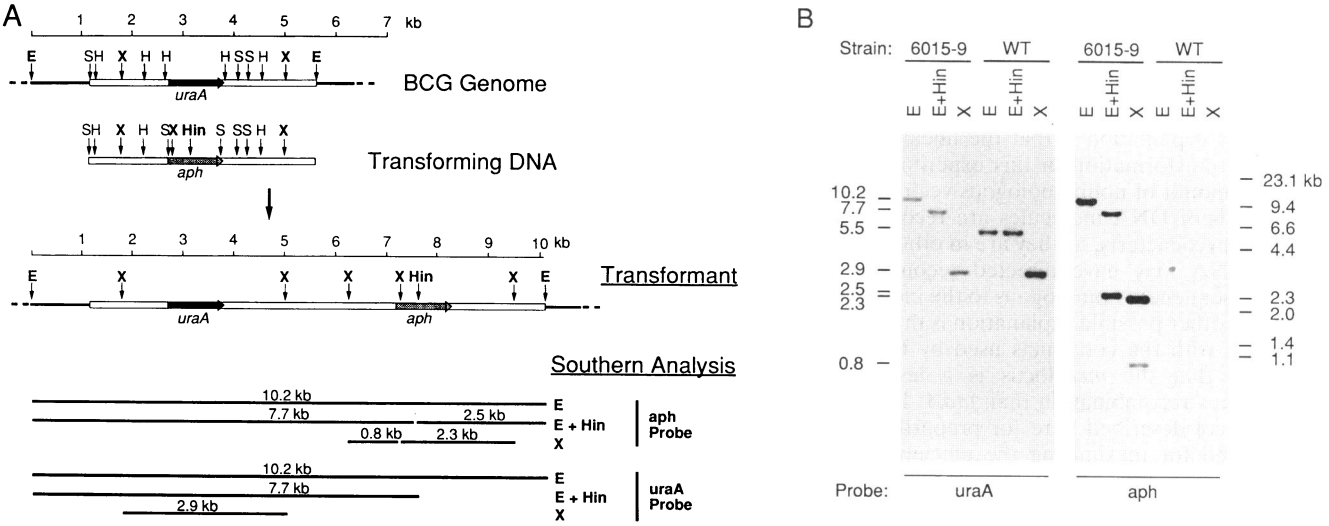

FIG. 5. Integration by homologous recombination in BCG. (A) The diagram depicts the *uraA* locus in wild-type BCG and in a BCG transformant in which the transforming DNA fragment has integrated via homologous recombination. The transforming plasmid was the linear insert DNA of plasmid pY6015, in which the *uraA* gene is replaced by the *aph* gene. A schematic interpretation of the Southern analysis is included. (B) Southern blot analysis. Genomic DNAs isolated from wild-type BCG (WT) and a BCG transformant (6015-9) were digested with restriction enzyme *Eco*RI (E), *Hin*dIII (Hin), or *Xba*I (X), subjected to agarose gel electrophoresis, transferred to nitrocellulose, and probed with labeled restriction fragments containing *uraA* or *aph*. The positions of DNA molecular size markers are indicated to the right, and the apparent size of each of the hybridizing DNA bands is indicated to the left.

The single recombination event seems to have occurred at the *Eco*RI site located at 5.6 kb in Fig. 5A and appears to have involved end-to-end joining of the linear transforming DNA fragment prior to, or during, recombination.

DISCUSSION

We report here attempts to address two limitations to our ability to manipulate the genomes of slowly growing mycobacteria: the lack of mycobacterial genes that can be used as genetic markers and the absence of a method that permits integration of recombinant DNA into homologous sites in the genome. We isolated and sequenced the *M. bovis* BCG *uraA* gene to provide a convenient genetic marker for molecular genetic experiments with *M. bovis* BCG and *M. tuberculosis*. By using a population of BCG cells propagated under conditions that may maximize the fraction of cells undergoing DNA synthesis, we found that a recombinant DNA fragment from the *uraA* locus can integrate into the homologous genomic locus.

Genes encoding OMP-DCase are widely used as genetic markers in diverse fungi, in slime molds, and in prokaryotes (4, 5, 15). These genes are useful because they are nutritional markers and can permit the development of genetic systems that do not rely on drug resistance markers. Such systems may be advantageous where drug resistance is a concern in the manipulation of potential human pathogens. Genes encoding OMP-DCase are also useful as genetic markers because they allow the development of both positive and negative selection schemes (4, 15). The facts that all organisms display the same basic UMP biosynthesis pathway, that all organisms examined have a single OMP-DCase gene, and that this gene can be readily cloned by complementation of *E. coli pyrF* mutants also contributed to our interest in the BCG OMP-DCase gene.

We previously showed that integration of recombinant DNA molecules occurs predominantly at homologous loci in transformed *M. smegmatis* (11). The results described here indicate that integration of recombinant DNA molecules can occur at homologous sequences in transformed *M. bovis* BCG but more frequently occurs at nonhomologous loci. Other investigators have observed only illegitimate recombination events in an attempt to obtain homologous recombination in *M. tuberculosis* and *M. bovis* BCG (13). There are several possibilities why this experiment failed to detect homologous recombination. One likely explanation is that the linearized DNA molecules used for transformation in this experiment contained a substantial amount of nonhomologous vector DNA at their ends; if the ends of DNA molecules are recombinogenic in slowly growing mycobacteria, as they are in other organisms, then the vector DNA may have directed recombination away from genomic sequences homologous to the incoming mycobacterial DNA. Another possible explanation is that integration into the *met* locus with the constructs used by Kalpana et al. (13) is lethal or that the *met* locus is a less favorable site for homologous recombination than *uraA*. It is also possible that the protocol described here for preparing BCG for transformation and for maximizing the efficiency of transformation produces cells that are more competent for homologous recombination.

Although gene replacement was not observed in the experiments described here, the ability to obtain single homologous recombination events should permit targeted gene disruption in slowly growing mycobacteria, as diagrammed in Fig. 6. When the transforming DNA consists of an internal segment of a gene (lacking both ends of the functional gene) and marker DNA has been introduced into the center of this

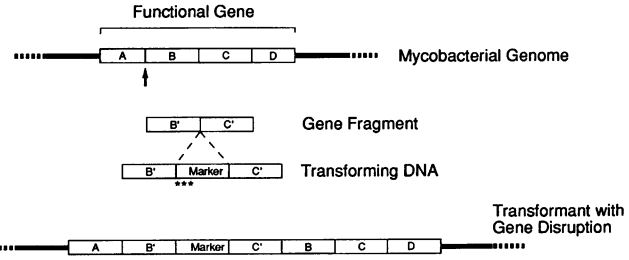

FIG. 6. Gene disruption through targeted integration. The diagram describes how a single recombination event can lead to the disruption of a functional gene. The hypothetical gene has been divided into four arbitrary segments, A, B, C, and D. A cloned internal segment of the gene, represented by B′-C′, can be interrupted with a marker DNA fragment which contains translation termination signals in all three frames (asterisks), producing the transforming DNA shown. Transformation of mycobacteria with the linear transforming fragment can lead to strand invasion at the homologous locus by one of the ends of the transforming DNA. In this hypothetical recombination event, strand invasion has occurred at the junction of segments A and B (thick arrow), the ends of the transforming fragment have come together during the recombination event, and disruption of the endogenous gene has occurred such that incomplete and nonfunctional gene products are expressed.

segment, then homologous recombination can lead to disruption of the endogenous gene such that incomplete and nonfunctional gene products are expressed. Nonetheless, it continues to be important to develop methods that result in gene replacement, as gene loss is a more reliable means of producing null mutations than gene disruption.

The genomes of *M. bovis* and *M. tuberculosis* are very similar. Thus, the cloned *M. bovis* BCG *uraA* locus described here should prove to be a useful tool for genetic manipulation of *M. tuberculosis*. Similarly, the methods used to obtain targeted integration in the *M. bovis* BCG genome should be directly applicable to the achievement of targeted integration in *M. tuberculosis*. The ability to manipulate genomic DNA in slowly growing mycobacteria should provide new approaches to the identification of mycobacterial genes that play a role in disease pathogenesis and construction of attenuated strains of *M. tuberculosis* for vaccine use.

ACKNOWLEDGMENTS

We thank B. James for assistance with BCG DNA sequence analysis, W. Jacobs for plasmid pYUB12, and P. Guilfoyle for helpful discussion.

This work was supported by Public Health Service grants AI23545 and AI26463.

REFERENCES

1. **Aldovini, A., and R. A. Young.** 1991. Humoral and cell-mediated immune responses to live recombinant BCG-HIV vaccines. Nature (London) **351**:479–482.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1987. Current protocols in molecular biology. Green Publishing Associates and Wiley Interscience, New York.
3. **Bloom, B. R., and C. Murray.** 1992. Tuberculosis: commentary on an emergent killer. Science **257**:1055–1067.
4. **Boeke, J., J. Truehart, B. Natsoulis, and G. R. Fink.** 1987. 5-Flouro-orotic acid as a selective agent in yeast molecular genetics. Methods Enzymol. **154**:164–175.
5. **Botstein, D., and R. W. Davis.** 1982. Principles and practice of recombinant DNA research with yeast, p. 607–636. *In* J. Strathern, E. Jones, and J. Broach (ed.), The molecular biology of the yeast Saccharomyces: metabolism and gene expression. Cold Spring

Harbor Laboratory Press, Cold Spring Harbor, N.Y.
6. **Calmette, A., C. Guérin, and B. Weill-Halle.** 1924. Essais de premunition par le BCG contre l'infection tuberculeuse de l'homme et des animaux. Bull. Acad. Natl. Med. (Paris) **91:**787–796.
7. **Collins, F. M.** 1984. Tuberculosis, p. 373–418. *In* R. Germanier (ed.), Bacterial vaccines. Academic Press, Inc., New York.
8. **Donnelly-Wu, M. K., W. R. Jacobs, Jr., and G. F. Hatfull.** 1993. Superinfection immunity of mycobacteriophage L5: applications for genetic transformation of mycobacteria. Mol. Microbiol. **7:**407–417.
9. **Eiglmeier, K., N. Honore, S. A. Woods, B. Caudron, and S. T. Cole.** 1993. Use of an ordered cosmid library to deduce genomic organisation of *Mycobacterium leprae*. Mol. Microbiol. **7:**197–206.
10. **Fine, P. E. M.** 1989. The BCG story: lessons from the past and implications for the future. Rev. Infect. Dis. **11**(Suppl. 2)**:**5353–5359.
11. **Husson, R. N., B. E. James, and R. A. Young.** 1990. Gene replacement and expression of foreign DNA in mycobacteria. J. Bacteriol. **172:**519–524.
12. **Jacobs, W. R., Jr., G. V. Kalpana, J. D. Cirillo, L. Pascopella, S. B. Snapper, R. A. Udani, W. Jones, R. G. Barletta, and B. R. Bloom.** 1991. Genetic systems for mycobacteria. Methods Enzymol. **204:** 537–555.
13. **Kalpana, G. V., B. R. Bloom, and W. R. Jacobs, Jr.** 1991. Insertional mutagenesis and illegitimate recombination in mycobacteria. Proc. Natl. Acad. Sci. USA **88:**5433–5447.
14. **Kaufmann, S. H. E.** 1990. Heat shock proteins and the immune response. Immunol. Today **11:**129–136.
15. **Kimsey, H. H., and D. Kaiser.** 1992. The orotidine-5′-monophosphate decarboxylase gene of *Myxococcus xanthus*. J. Biol. Chem. **267:**819–824.
16. **Lee, M. H., L. Pascopella, W. R. Jacobs, Jr., and G. F. Hatfull.** 1991. Site-specific integration of mycobacteriophage L5: integration-proficient vectors for *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, and bacille Calmette-Guerin. Proc. Natl. Acad. Sci. USA **88:**3111–3115.
17. **Luelmo, F.** 1982. BCG vaccination. Am. Rev. Respir. Dis. **125:**70–72.
18. **Martin, C., J. Timm, J. Rauzier, R. Gomez-Lus, J. Davis, and B. Giquel.** 1990. Transposition of an antibiotic resistance element in mycobacteria. Nature (London) **345:**739–743.
19. **Mizuguchi, Y., and T. Takunaga.** 1968. Spheroplasts of mycobacteria. 2. Infection of phage and its DNA on glycine treated mycobacteria and spheroplasts. Med. Biol. **77:**57.
20. **Nyunoya, H., and C. J. Lusty.** 1983. The carB gene of Escherichia coli: a duplicated gene coding for the large subunit of carbamoylphosphate synthetase. Proc. Natl. Acad. Sci. USA **80:**4629–4633.
21. **Rothstein, R.** 1991. Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. Methods Enzymol. **194:**281–301.
22. **Shinnick, T. M.** 1987. The 65-kilodalton antigen of *Mycobacterium tuberculosis*. J. Bacteriol. **169:**1080–1088.
23. **Snapper, S. B., L. Lugosi, A. Jekkel, R. E. Melton, T. Kieser, B. R. Bloom, and W. R. Jacobs, Jr.** 1988. Lysogeny and transformation in mycobacteria: stable expression of foreign genes. Proc. Natl. Acad. Sci. USA **85:**6987–6991.
24. **Snapper, S. B., R. E. Melton, S. Mustafa, T. Kieser, and W. R. Jacobs, Jr.** 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol. Microbiol. **4:**1911–1919.
25. **Snider, D. E., and W. L. Roper.** 1992. The new tuberculosis. N. Engl. J. Med. **326:**703–705.
26. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennet, G. P. Bansal, J. F. Young, M. H. Lee, G. F. Hatfull, S. B. Snapper, R. G. Barletta, W. R. Jacobs, Jr., and B. R. Bloom.** 1991. New use for BCG recombinant vaccines. Nature (London) **351:**456–460.
27. **Styblo, K.** 1989. Overview and epidemiological assessment of the current global tuberculosis situation with an emphasis on control in developing countries. Rev. Infect. Dis. **11**(Suppl. 2)**:**S339–S346.
28. **Telenti, A., P. Imboden, F. Marchesi, D. Lowrie, S. Cole, M. J. Colston, L. Matter, K. Schopfer, and T. Bodmer.** 1993. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. Lancet **341:**647–650.
29. **Vega-Lopez, F., L. A. Brooks, H. M. Dockrell, K. S. E. De Smet, J. K. Thompson, R. Hussain, and N. G. Stoker.** 1993. Sequence and immunological characterization of a 42-kDa serine-rich antigen from *Mycobacterium leprae*. Infect. Immun. **61:**2145–2153.
30. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153:**3–11.
31. **Winter, N., M. Lagranderie, J. Rauzier, J. Timm, C. Leclerc, B. Guy, M. P. Kieny, M. Gheorghiu, and B. Gicquel.** 1991. Expression of heterologous genes in *Mycobacterium bovis* BCG: induction of a cellular response against HIV-1 Nef protein. Gene **109:**47–54.
32. **Young, D., T. Garbe, R. Lathigra, and C. Abou-Zeid.** 1990. Protein antigens: structure, function and regulation, p. 1–35. *In* J. McFadden (ed.) Molecular biology of mycobacteria. Academic Press Ltd., London.
33. **Young, D. B., and S. T. Cole.** 1993. Leprosy, tuberculosis, and the new genetics. J. Bacteriol. **175:**1–6.
34. **Young, D. B., S. H. Kaufmann, P. W. Hermans, and J. E. Thole.** 1992. Mycobacterial protein antigens: a compilation. Mol. Microbiol. **6:**133–145.
35. **Young, R. A.** 1990. Stress proteins and immunology. Annu. Rev. Immunol. **8:**401–420.
36. **Young, R. A., B. R. Bloom, C. M. Grosskinsky, J. Ivanyi, D. Thomas, and R. W. Davis.** 1985. Dissection of *Mycobacterium tuberculosis* antigens using recombinant DNA. Proc. Natl. Acad. Sci. USA **82:**2583–2587.
37. **Zhang, Y., B. Heym, B. Allen, D. Young, and S. T. Cole.** 1992. The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. Nature (London) **358:**591–593.