JOURNAL OF BACTERIOLOGY, Dec. 2003, p. 6870–6882 Vol. 185, No. 23
0021-9193/03/$08.00+0 DOI: 10.1128/JB.185.23.6870–6882.2003

Subtractive Hybridization Reveals a Type I Polyketide Synthase Locus Specific to *Mycobacterium ulcerans*

Grant A. Jenkin,[1] Timothy P. Stinear,[2] Paul D. R. Johnson,[3] and John K. Davies[1]*

Bacterial Pathogenesis Research Group, Department of Microbiology, Monash University, Clayton,[1] and Department of Infectious Diseases, Austin and Repatriation Medical Centre, Heidelberg,[3] Victoria, Australia, and Unité Génétique Moléculaire Bactérienne, Institut Pasteur, Paris, France[2]

Received 12 May 2003/Accepted 2 September 2003

Mycobacterium ulcerans **causes Buruli ulcer, the third most prevalent mycobacterial infection of immunocompetent humans after tuberculosis and leprosy. Recent work has shown that the production by *M. ulcerans* of mycolactone, a novel polyketide, may partly explain the pathogenesis of Buruli ulcer. To search for the genetic basis of virulence in *M. ulcerans*, we took advantage of the close genetic relationship between *M. ulcerans* and *Mycobacterium marinum* by performing genomic suppressive subtractive hybridization of *M. ulcerans* with *M. marinum*. We identified several DNA fragments specific to *M. ulcerans*, in particular, a type I polyketide synthase locus with a highly repetitive modular arrangement. We postulate that this locus is responsible for the synthesis of mycolactone in *M. ulcerans*.**

The environmental mycobacterium *Mycobacterium ulcerans* is the causative organism of Buruli ulcer (BU), an ulcerative skin disease of humans that is associated with significant morbidity and disability. BU has surpassed leprosy in prevalence in some rural areas of West Africa (2). This combined with the lack of effective preventative strategies and of treatments other than surgical resection has prompted the World Health Organization to recognize BU as an emerging public health problem and to establish of the Global Buruli Ulcer Initiative to coordinate research and public health interventions (44). The characteristic clinical and pathological changes of BU include extensive necrosis of subcutaneous tissue, extracellular location of organisms, and lack of a granulomatous immune response, which differentiate it from diseases caused by other mycobacteria (22, 23). *M. ulcerans* produces a polyketide toxin called mycolactone that appears to be important in pathogenesis (18). The nucleotide sequence of the 16S rRNA gene from *M. ulcerans* is >99.8% identical to that of *Mycobacterium marinum* (31), and multilocus sequencing typing (MLST) confirms this very close genetic relationship (38). In fact, the genetic similarities between *M. ulcerans* and *M. marinum* suggest that *M. ulcerans* is best considered an ecotype of *M. marinum* that has adapted to an as-yet-unidentified environmental niche. Despite these genetic similarities, a number of phenotypic differences exist between these mycobacteria, particularly in growth rate and pigment production (43). The clinical, pathological, and epidemiological features of disease also differ substantially (17), and in particular, mycolactone is not produced by *M. marinum* (18). Further, substantial genetic differences between *M. ulcerans* and *M. marinum* have been documented. Relative binding ratios between *M. ulcerans* and *M. marinum* genomes are surprisingly low, at 37 to 38% (41). Pulsed-field gel electrophoresis reveals a smaller genome size for *M. ulcerans* and substantial macrorestriction fragment polymorphism between the two species (38). Two high-copy-number insertion sequences, IS*2606* and IS*2404*, have also been characterized for *M. ulcerans* and are not present in *M. marinum* (33, 37).

Comparison of mycobacterial genomes suggests that mycobacterial genetic diversity is driven by genetic rearrangements rather than nucleotide variations (10, 13, 14), and insertion sequence elements are a common substrate for such events. We postulated that the observed variations between *M. ulcerans* and *M. marinum* might in part be mediated by the acquisition of DNA by *M. ulcerans* through horizontal gene transfer. The presence of IS*2404* and IS*2606* in *M. ulcerans* provides evidence that such acquisitions have indeed occurred, and the widespread distribution of these elements in the genome (38) would provide a rich substrate for recombination events. To isolate such genomic differences, subtractive hybridization between strains of *M. ulcerans* and *M. marinum* was performed by using the technique of suppressive subtractive hybridization (SSH). This technique was initially developed to isolate differences in cDNA pools, but it has also successfully identified genomic differences between closely related strains of *Helicobacter pylori* (1) and *Aeromonas hydrophilia* (46). Subtractive hybridization between mycobacteria has previously identified genomic differences between *Mycobacterium tuberculosis*, *Mycobacterium bovis*, and *M. bovis* BCG (27); between *Mycobacterium avium* subsp. *avium* and *M. avium* subsp. *paratuberculosis* (40); and between *M. avium* and *Mycobacterium intracellulare* (28). We describe the application of this technique to *M. ulcerans* and the identification of a specific type I polyketide synthase locus.

MATERIALS AND METHODS

Strains and culture media. The strains used in this study are listed in Table 1. The origins of mycobacterial strains and the culture conditions used have been previously described (37, 38). *Escherichia coli* DH5α strains were cultured on Luria-Bertani agar containing selective antibiotics at appropriate concentrations.

Mycobacterial DNA extraction. A combined mechanical lysis method was used for mycobacterial DNA extraction (7, 36). Cells were stripped from the surface of Brown Buckle slopes by aspiration in 0.05% (vol/vol) Tween 80 and then

* Corresponding author. Mailing address: Department of Microbiology, Monash University, Victoria 3800, Australia. Phone: 61 3 9905 4824. Fax: 61 3 9905 4811. E-mail: John.Davies@med.monash.edu.au.

TABLE 1. Bacterial strains used in this study

Strain	Description
E. coli DH5α	For general DNA manipulations
M. ulcerans	
Chant	Victoria, Australia, 1994
Benin 9825	Benin, West Africa, 1994
13822/70	Queensland, Australia 1971
5145	Human isolate, Democratic Republic of Congo, 1976
M. marinum	
NCTC 2275	Environmental isolate; Philadelphia, Pa.; MLST type I[a]
JKD 2396	Human isolate; Victoria, Australia; 1998; MLST type V[a]
99/86	Human isolate; Tasmania, Australia; 1993; MLST type V[a]

[a] MLST type as described by Stinear et al. (38).

resuspended in 1 ml of TE buffer (10 mM Tris-HCl, 1 mM EDTA [pH 8.0], 0.1% Tween 80), containing 10 mg of lysozyme, at 37°C for 2 h with rolling. The cells were pelleted at 17,000 × *g* for 2 min, the supernatant was removed, and the cells were resuspended in 700 μl of lysis buffer (600 μl of 1× TE buffer [pH 8.0], 70 μl of 10% sodium dodecyl sulfate [SDS], 10 μl of proteinase K [20 mg/ml]) and incubated at 50°C for 1 h. Cells were pelleted and then transferred to 2-ml screw-top skirted lysis tubes containing 250 μl of 150-μm-diameter washed beads (Sigma, St. Louis, Mo.), 500 μl of pH 8.0 lysis buffer (9.6% liquid Pyroneg detergent [Diversey], 120 mM sodium acetate), 500 μl of equilibrated phenol (pH 7.0), and 100 μl of chloroform-isoamyl alcohol (24:1) (CIA). Cell lysis was performed in a FP120 FastPrep bead beater (Savant Instruments, Holbrook, N.Y.) at speed 5.0 for 30 s. The tubes were placed on ice and centrifuged for 10 min at 17,000 × *g*. The supernatant was aspirated and reextracted with an equal volume of CIA. Nucleic acid was precipitated with 0.1 volume of 3 M sodium acetate (pH 5.2) and 1.0 volume of isopropanol, and the pellet was washed in 70% alcohol, dried under vacuum, and resuspended in nuclease-free water. RNase (25 μg/ml) treatment of the pellet was followed by repeat phenol CIA extraction and DNA precipitation. This method results in approximately 40 μg of DNA per Brown Buckle slope.

General DNA manipulations. DNA purification was with QIAquick spin columns (Qiagen Inc.) according to the manufacturer's instructions. Plasmid extractions were performed according to the manufacturer's instructions with the Hipure miniprep kit (Roche Diagnostics, Mannheim, Germany). Sequencing was performed with the Prism Big Dye terminator cycle sequencing kit (Applied Biosystems) and analyzed in a PE Applied Biosystems model 373 automated sequencer.

Ligation reactions used T4 DNA ligase (20 U/μl; Promega) in the presence of 0.1 volume of 10× ligase buffer (0.3 M Tris-HCl [pH 7.8], 100 mM $MgCl_2$, 100 mM dithiothreitol, 10 mM ATP) (Promega) for adaptor ligations. For other ligations, 1 U of T4 DNA ligase was used.

Table 2 shows the sequences and derivations of primers. All primers except

TABLE 2. Oligonucleotides used in this study

Oligonucleotide (reference)	Sequence (5′ → 3′)	Description
Adaptor 1 (1)	CTAATACGACTCACTATAGGGCTCGAGCG GCCGCCCGGGCAGGTGGCCCGTCCA	SSH adaptor
Adaptor 2 (1)	CTAATACGACTCACTATAGGGCAGCGTGG TCGCGGCCGAGGTGCCGGCTCCA	SSH adaptor
Primer 1 (P1) (1)	CTAATACGACTCACTATAGGGC	SSH first-round primer
Nested primer 1 (NP1) (1)	TCGAGCGGCCGCCCGGGCAGGT	SSH second-round primer
Nested primer 2 (NP2) (1)	AGCGTGGTCGCGGCCGAGGT	SSH second-round primer
IS2404F (37)	AGCGACCCCAGTGGATTGGT	IS*2404* primers
IS2404R (37)	CGGTGATCAAGCGTTCACGA	
18.1F (this study)	TGATGCGGTCTTTCATCTTG	*pks* β-ketoreductase domain
18.1R (this study)	ATATGCGCCTTGACTTTTGC	
47.1F (this study)	CAGCCAACTGCGCTACTACA	*pks* loading module
47.1R (this study)	GACCACACTGATCCCGTCTC	
74.1F (this study)	GCAACGATTGATGCTTGAAC	*pks* loading module
74.1R (this study)	AGGAGACACGGTTGGCTATG	
4.1F (this study)	GGACGGCAAGATCACTGTC	*pks* acyltransferase domain
4.1R (this study)	GAACTGATGGCGGATGTGTT	
86.2F (this study)	TGTTGACACAAGCGATCACC	*pks* acyltransferase domain
86.2R (this study)	CAGTAACGCTGATGCTGGAA	
51.2F (this study)	ACTAACGGCGACAGAACAGC	*pks* acyl carrier protein domain
51.2R (this study)	GAGTGGACGAAAGGTTGAGG	
54.2F (this study)	TTCCGGATTACGTGATAGGC	*pks* acyltransferase domain
DHF (this study)	GCCGGCAGTCTCTTAACATC	*pks* dehydratase domain
B13F (this study)	CAGACGTCGGGTTAGGAAAA	*mtFabH*/*dpsC* homologue
B13R (this study)	TCTTAGCAAACCCCAGGATG	

those described by Akopyants et al. (1) were designed by using Primer 3 software (34) with default settings.

Southern hybridization analysis. DNA fragments were separated in 2.0% (wt/vol) agarose gels, and transfer of DNA to Hybond N+ nylon membranes (Amersham Pharmacia) was performed in 0.4 N NaOH. For dot blot analysis, PCR products from amplification of inserts were purified and resuspended in 400 μl of 0.1 M NaOH at 37°C. The denatured DNA from each clone was then applied to separate wells of a dot blot apparatus and transferred under vacuum to a Hybond N+ membrane which had been preequilibrated in 2× SSC (20× SSC is 3 M NaCl plus 0.3 M sodium citrate [pH 7.0]). The individual wells and membrane were then washed with 2× SSC, and the membrane was exposed to 1,200 mJ of UV in a Stratalinker (Stratagene). All hybridizations were performed at 65°C after prehybridization in hybridization buffer at 65°C for 4 h. The hybridization buffer contained 1% (wt/vol) skim milk powder dissolved in maleic acid buffer (0.1 M maleic acid and 0.15 M NaCl [pH 7.5])–0.1% (wt/vol) SDS–5× SSC–0.1% (wt/vol) *N*-lauroylsarcosine. Digoxigenin (DIG)-labeled DNA probes were prepared either by PCR incorporation with appropriate primers and templates or by random-primed DNA labeling with random hexamer priming according to the instructions of the manufacturer (Roche Diagnostics). High-stringency washes (0.1× SSC plus 0.1% [wt/vol] SDS at 65°C) were used, and the substrate for detection of anti-DIG–alkaline phosphatase conjugate (Roche Diagnostics) was CDP-Star (Amersham Pharmacia).

SSH. The methodology for SSH was as previously described (1) with some modifications (Fig. 1). An initial SSH experiment used *M. ulcerans* Chant and *M. marinum* NCTC 2275, and a subsequent independent experiment compared *M. ulcerans* Benin 9825 and *M. marinum* JKD 2396. Genomic *M. ulcerans* DNA was digested with *Alu*I (Roche Diagnostics), which produced a smear of fragments <2 kb in size (data not shown). The technique requires ligation of two different adaptors to digested tester (*M. ulcerans*) DNA in two separate pools. Adaptor ligation was determined to be of acceptable efficiency by using comparative PCR. With the adaptor ligation mix as the template, the strength of the PCR product obtained with IS*2404*-specific primers was compared with that of the product obtained with a single specific IS*2404* primer and the adaptor primer P1 to amplify across the adaptor-*Alu*I fragment join.

The two pools were denatured and annealed to an excess of *M. marinum* single-stranded DNA (ssDNA), which had not been ligated to adaptors, initially in their respective pools and then together in a single tube in the presence of more *M. marinum* ssDNA. The hybridization temperature was 65°C and the sodium salt was concentration 50 mM throughout. After hybridization was terminated, nested PCR amplification was performed with the subtracted hybridization mix as a template and primers directed against the outer sequence (first round) and then the inner sequence (second round) of the adaptors. Manual hot-start PCRs were carried out in a standard PCR buffer (Roche Diagnostics) with an Mg^{2+} concentration of 1.5 mM, deoxynucleoside triphosphates at 0.5 mmol, appropriate primers at 200 nM each, and 1 U of *Taq* polymerase (Roche Diagnostics) per reaction. The design of the adaptors permits efficient amplification of only those double-stranded *M. ulcerans* sequences that have a different adaptor at either end. Such sequences are likely to form only when hybridization occurs between *M. ulcerans* ssDNA present in both adaptor pools but absent from the *M. marinum* pool. Other tester sequences (containing a single adaptor or the same adaptor at both ends) are not efficiently amplified, and the nested PCR product will therefore be enriched with *M. ulcerans* specific DNA. In all SSH experiments, an unsubtracted *M. ulcerans* control was included, which consisted of *Alu*I-digested *M. ulcerans* genomic DNA ligated to both adaptors 1 and 2 in the same ligation reaction. Nested PCR amplification of this template was performed as for the subtracted sample.

Titration of the second-round PCR cycle number was performed by Southern hybridization analysis of the subtracted PCR product compared with the unsubtracted control with the whole genomic DIG-labeled driver DNA probe (either *M. marinum* NCTC 2275 or *M. marinum* JKD 2396). Efficient subtraction was determined by the absence of hybridization of the subtracted second-round PCR product compared with the unsubtracted control run for the same number of second-round PCR cycles. The cycle number chosen for further analysis was that at which adequate PCR product was obtained for analysis with little or no hybridization to the *M. marinum* DNA DIG probe.

Analysis of the subtracted second-round PCR product. The subtracted second-round PCR product was amplified for the appropriate cycle number and purified, and the ends were filled in with 10 U each of DNA polymerase I (Promega) and T4 polynucleotide kinase (Promega) in a 100-μl final reaction volume, containing 250 μM deoxynucleoside triphosphates (Promega), 1 mM ATP, and 10 μl of 10× polymerase I buffer (0.5 M Tris [pH 7.5], 0.1 M $MgCl_2$, 10 mM dithiothreitol, 0.5 mg of bovine serum albumin per ml), incubated at 37°C for 1 h. The reaction was terminated by the addition of 1 μl of 0.5 M EDTA (pH 8.0), and the DNA was purified by using QIAquickspin columns (Qiagen) and eluted in 10 mM Tris-HCl (pH 8.5) buffer. The PCR products were ligated into blunt linearized dephosphorylated pUC18 and used to transform supercompetent *E. coli* DH5α by heat shock. Transformants were selected by ampicillin resistance and screened for white colony color on Luria-Bertani agar plates containing 20 μg of X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) per ml. PCR amplification of inserts was performed. Initial results indicated that IS*2404* would be found in many of the clones, and so efforts were made to screen out IS*2404*-containing clones by Southern hybridization analysis of PCR-amplified inserts with an *IS2404* probe. Any clone insert not showing hybridization to this probe was sequenced by using primers to the pUC18 vector adjacent to the multicloning site.

Nucleotide sequence accession numbers. The polyketide synthase gene sequences described here have been deposited in GenBank with accession numbers AY289593, AY289594, AY289595, and AY289596. Sequences for the *M. ulcerans*-specific sequences identified from the subtraction have also been deposited in GenBank, and the accession numbers are given in Table 3.

RESULTS

SSH. The subtracted second-round PCR resulted in a smear of products between 200 and 800 bp, with a predominant bright band at 220 bp (Fig. 2). Comparisons of Southern hybridizations using genomic driver DNA to probe the subtracted second-round PCR product and the unsubtracted control PCR product showed nearly complete subtraction of driver DNA (Fig. 2). A cycle number of 25 was chosen for further analysis. The 220-bp band was shown by sequencing to consist of an *Alu*I fragment of two adjacent IS*2404* genes in the same orientation, separated by a 32-bp intervening sequence. Up to 80% of clones contained this tandem IS*2404 Alu*I fragment as determined by Southern hybridization. This 32-bp intervening sequence has no significant sequence homology to any known DNA sequence in GenBank or in sequenced mycobacterial genomes. Because the SSH second-round PCR amplification did not produce fragments greater than approximately 800 bp, we did not expect to isolate IS*2606* from the subtraction, as it does not contain an *Alu*I site and is 1,404 bp in length. In the first and second subtractions, respectively, 13 and 11 unique clone inserts not containing IS*2404* showed hybridization to *M. ulcerans* but not *M. marinum* genomic probes (Fig. 3), and these are listed in Table 3.

Sequences with significant BLASTX homology to known proteins encoded mainly type I polyketide synthase genes. In the first subtraction experiment, seven sequences showed significant translated homology to polyketide synthases (Table 3). The polyketide synthase genes appeared to be particularly interesting for several reasons. Although type I polyketide synthases are large and prevalent in mycobacteria (13), these sequences seemed overrepresented in comparison with their likely total representation in the genome. In the second SSH experiment, three of the *M. ulcerans*-specific sequences matched areas of *M. ulcerans pks* identified in the first experiment, and another four matched other *pks* sequences. All BLASTX similarities with type I polyketide synthases show expect values of <0.001, except for sequence B15 (E = 0.054). Sequence B15 shows translated amino acid identity of 63% to the *Saccharopolyspora spinosa* polyketide synthase (accession number AAG23264) over its length, and the expect value reflects its short length of 95 bp. Further, the derived amino acid sequences from B13 showed significant homology to the product of the *M. tuberculosis* gene *mtfabH*, which encodes a β-ketoacyl:acyl carrier protein synthase (KASIII) that has a role in

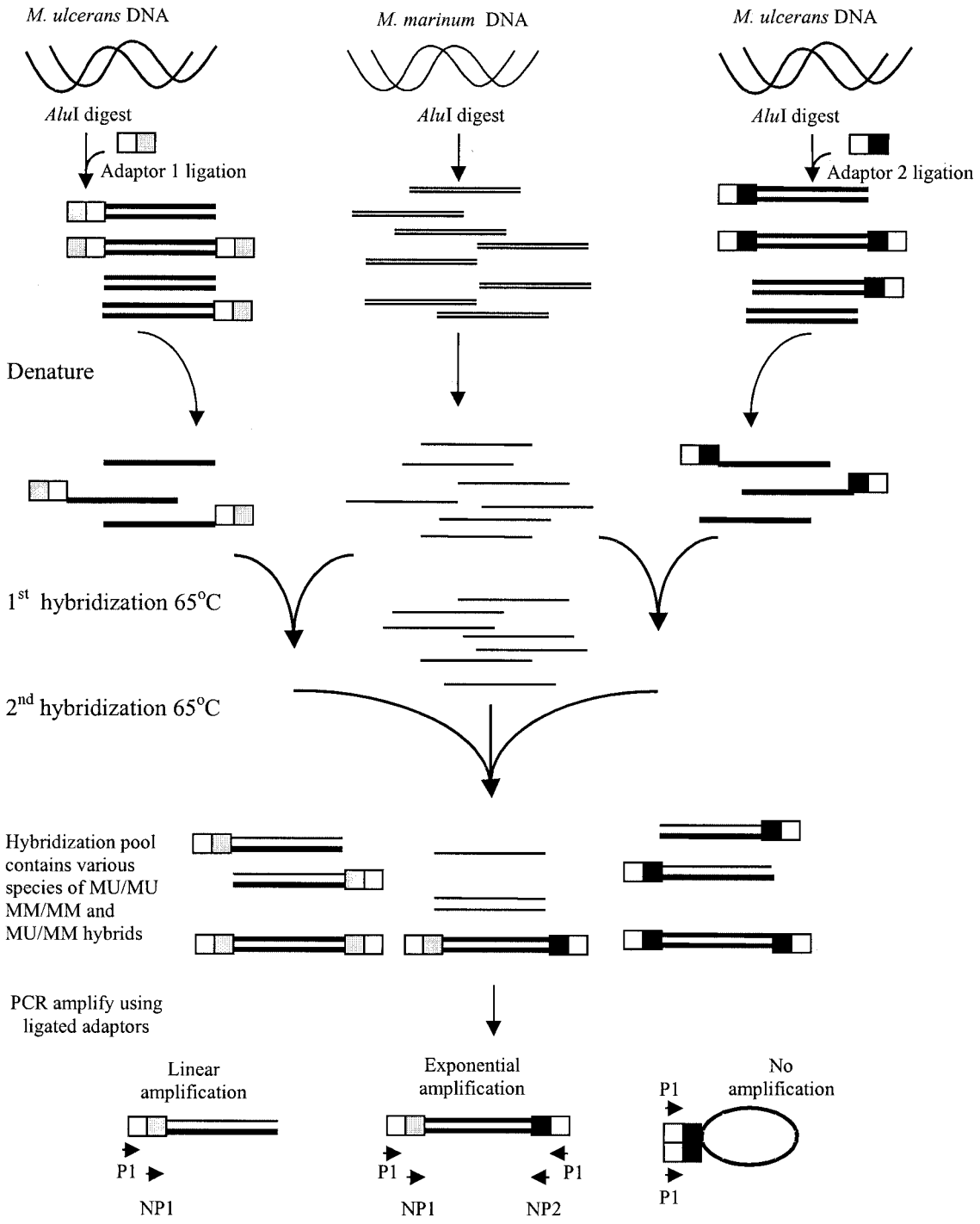

FIG. 1. Schematic diagram of the SSH procedure (1). Thick lines, *M. ulcerans* DNA; thin lines, *M. marinum* DNA. White-and-gray bars, adaptor 1; white-and-black bars, adaptor 2; MU, *M. ulcerans*; MM, *M. marinum*.

mycolic acid synthesis (12), and also to that of the *Streptomyces peucetius* gene *dpsC*, which determines propionyl coenzyme A (propionyl-CoA) starter unit selection for the loading module of the daunorubicin-doxorubicin anthracycline polyketide synthase (32). The results of the two separate SSH experiments, each with different tester and driver species, suggested that the main acquisition of DNA by *M. ulcerans* relative to *M. marinum*, other than the previously described insertion sequence elements, is of polyketide synthase and related genes. We were also interested in these sequences because mycolactone is produced by *M. ulcerans* but not by *M. marinum* and is predicted to require two type I polyketide synthases for its synthesis (21).

With regard to the other sequences found in the subtraction, the derived amino acid sequence from B11 had no significant

TABLE 3. Sequences obtained by subtractive hybridization

Strains used	Clone	Accession no.	Length (bp)	BLASTX match (E value)[a]	Hybridization[b]		PCR[c]	
					M. ulcerans	*M. marinum*	*M. ulcerans*	*M. marinum*
M. ulcerans Chant (tester), *M. marinum* NCTC 2275 (driver)	4.1	AY341915	201	*Streptomyces antibioticus* oleandomycin polyketide synthase, modules 5 and 6 (Q07017) (4e-14)	+++	−	+	−
	18.1	AY341919	469	*Saccharopolyspora spinosa* polyketide synthase (AAG23266) (2e-36)	++	−	+	−
	47.1	AY341913	366	*Streptomyces avermitilis* polyketide synthase (BAB69303) (9e-17)	+	−	+	−
	74.1	AY341914	304	*Streptomyces antibioticus* 8,8a-deoxyoleandolide synthase (AAF82408) (4e-29)	++	−	+	−
	51.2	AY341917	335	*Streptomyces noursei* polyketide synthase (*nysC*) (AAF71776) (3e-22)	+++	−	+	−
	54.2	AY341918	231	*M. tuberculosis* CDC1551 polyketide synthase (MT3018) (1e-12)	++	−		
	86.2	AY341916	260	*Polyangium cellulosum* polyketide synthase (CAD43448) (4e-13)	+++	−	+	−
	4.2	AY341921	568	No significant match (>0.1)	++	−		
	31.2	AY341920	545	No significant match (>0.1)	+++	−		
	89.2	AY341922	208	No significant match (>0.1)	++	−		
	4.3	AY341923	319	No significant match (>0.1)	++	−		
	18.3	AY341924	295	No significant match (>0.1)	++	−		
	19.3	AY341925	545	No significant match (>0.1)	+++	−		
M. ulcerans Benin 9825 (tester), *M. marinum* JKD 2396 (driver)	B3	AY341926	162	*Streptomyces caelestis* polyketide synthase modules 1 and 2 (AAC46024) (6e-14)	+	−		
	B11	AY341927	230	Hypothetical protein from *Magnetococcus* sp. (ZP_00042334) (7e-09)	+++	−		
	B13	AY341928	274	*mtFabH* homologue (AAM88612) (6e-05)	++	−	+	−
	B15	AY341929	95	*S. spinosa* polyketide synthase (AAG23264) (0.054)	ND[d]	ND		
	B21		201	***M. ulcerans pks* (4.1)**	++	−	+	−
	B22	AY341930	193	No significant match (>0.1)	+++	−		
	B60	AY341931	160	*Streptomyces avermitilis* MA-4680 polyketide synthase (BAC68128) (4e-07)	ND	ND		
	B64	AY341932	160	*Polyangium cellulosum* polyketide synthase (CAD43448) (7e-04)	ND	ND		
	B68		260	***M. ulcerans pks* (86.2)**	ND	ND	+	−
	B96	AY341933	192	No significant match (>0.1)	+++	−		
	B99		304	***M. ulcerans pks* (74.1)**	ND	ND	+	−

[a] Highest scoring match in GenBank with the BLASTX algorithm. Matches with Expect (E) values of ≤0.1 are reported. Boldface indicates *pks* sequences already obtained from the first subtraction; hybridization was not repeated for these sequences.

[b] Southern hybridization analysis with whole *M. ulcerans* or *M. marinum* DIG-labeled probe.

[c] +, a product was obtained with *M. ulcerans* or *M. marinum* genomic DNA as the template and primers designed from the respective clone sequence (see Table 2 for sequences).

[d] ND, not determined.

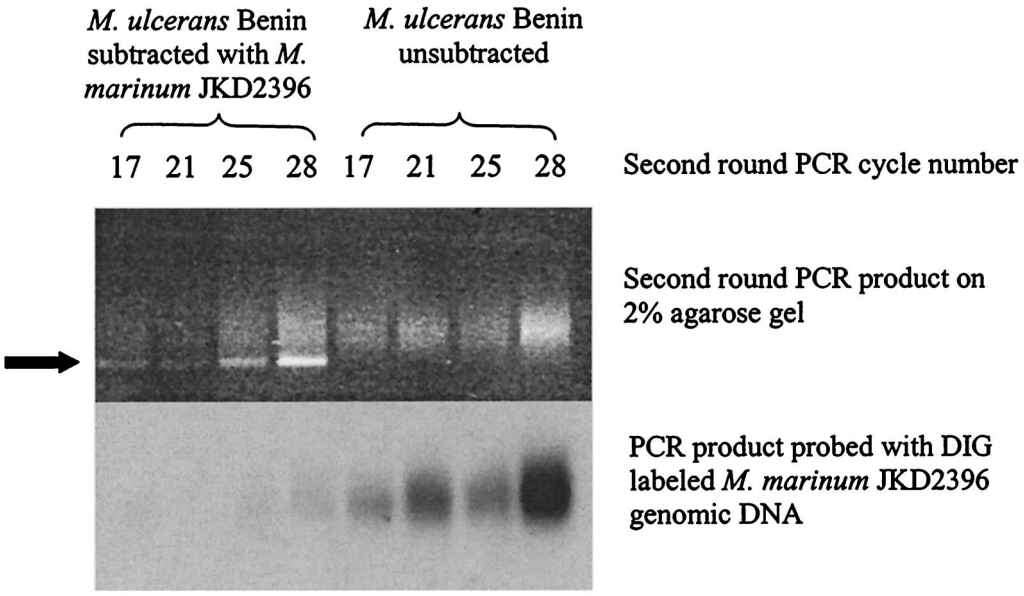

FIG. 2. Titration of second-round PCR product of *M. ulcerans* Benin 9825 with *M. marinum* JKD2396 subtractive hybridization. Southern hybridization with DIG-labeled *M. marinum* JKD2396 genomic DNA is shown. The arrow indicates the position of the 220-bp band mentioned in the text.

homologies to known mycobacterial proteins but did demonstrate 60% similarity to a hypothetical protein from *Magnetococcus* sp. of unknown function. Eight other *M. ulcerans* sequences (4.2, 31.2, 89.2, 4.3, 18.3, 19.3, B22, and B96) demonstrated no significant DNA or derived amino acid sequence similarities and have not been further investigated at this stage.

Seven further sequences identified from the subtractions showed very high sequence homology to known *M. tuberculosis* H37Rv proteins and were also present in *M. marinum* as determined by Southern hybridization. These sequences are not listed in Table 3 and have not been further analyzed.

The polyketide synthase sequences are specific for *M. ulcerans*. PCR analysis (see Table 2 for primers) and Southern hybridization with DIG-labeled probes demonstrated the presence of the identified *pks* sequences in all 18 *M. ulcerans* strains examined from different geographic areas, including Victoria and Queensland in Australia, West Africa, Papua New Guinea, Malaysia, Japan (*Mycobacterium shinshuense* ATCC 33728), Mexico, and Surinam. The *pks* sequences were absent from 18

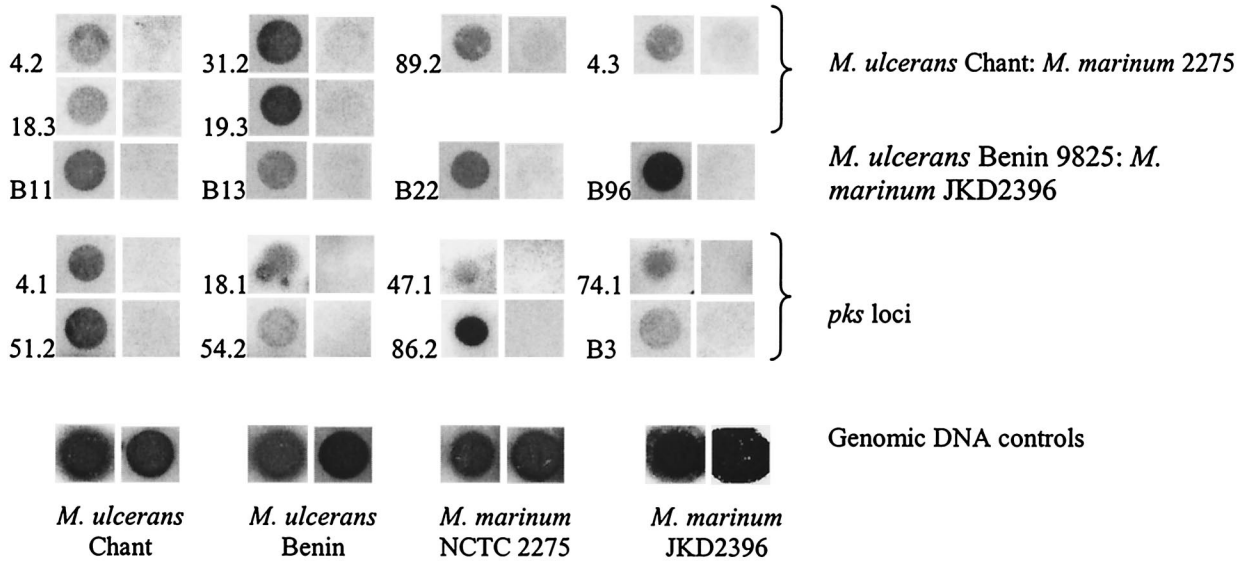

FIG. 3. Dot blot hybridizations of clones identified from the SSH experiments that are specific to *M. ulcerans*. The numbers are the clone numbers (Table 3) shown with the hybridization first to the *M. ulcerans* (Chant or Benin 9825) DIG probe (3-min exposure) and second to the *M. marinum* (NCTC 2275 or JKD 2396) DIG probe (18-min exposure). Control hybridizations with genomic DNA are also shown.

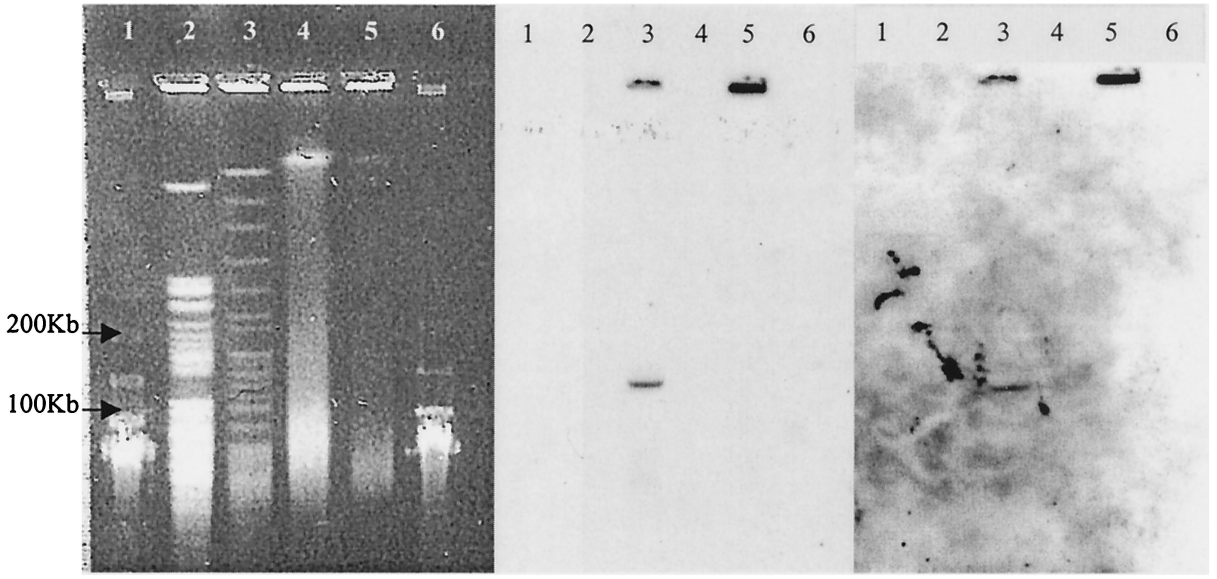

FIG. 4. Pulsed-field gel electrophoresis (left) and Southern hybridization (middle and right) with a DIG-labeled probe to the 74.1 sequence (middle) and the 18.1 sequence (right) of the *M. ulcerans pks* gene. Target DNAs were from *M. ulcerans* 13822/70 (lanes 3 and 5) and *M. marinum* 99/86 (lanes 2 and 4) and were either undigested (lanes 4 and 5) or *Ase*I digested (lanes 2 and 3). Lanes 1 and 6, 50-kb lambda ladder.

M. marinum strains of all five different MLST types described by Stinear et al. (38) and from 10 other mycobacterial species (*M. tuberculosis*, *M. bovis*, *M. smegmatis*, *M. fortuitum*, *M. chelonae*, *M. aurum*, *M. agri*, *M. asiaticum*, *M. triplex*, and *M. avium*) as determined by Southern hybridization (data not shown).

Southern hybridization analysis of an *Ase*I digest of genomic DNA separated by pulsed-field gel electrophoresis demonstrated that the identified *pks* fragments hybridized to the same 135-kb *Ase*I fragment. An example is shown in Fig. 4. They may therefore be from the same type I *pks* loci.

Linkage of subtraction clones by PCR. In view of the findings described above, we performed PCR analysis to determine the relative positions of the identified *pks* sequences. The relationship of the sequences and primers used are indicated in Fig. 5. By using primers 47.1F and 74.1R, a 550-bp product was obtained, and sequencing confirmed clones 74.1 and 47.1 to be adjacent *Alu*I fragments of a *pks* gene (Fig. 5A). Similarly, PCR with primers 4.1F and 86.2R showed these to be nearly adjacent *Alu*I fragments of a separate region of a *pks* gene separated by a 9-bp *Alu*I fragment (Fig. 5B and C). Primers 86.2R and 47.1F produced a product of 2.6 kb (Fig. 5A). Sequencing of this product revealed sequence divergence from the 86.2 subtraction clone upstream of the 86.2R primer binding site, which was subsequently found to match the sequences of clones B60 and B3. Primers 18.1F and 51.2R produced a PCR product of 660 bp, and primers 54.2F and 18.1 R produced a product of 2 kb. Sequencing of these products confirmed sequences 18.1 and 51.2 to be separated by a 163-bp *Alu*I fragment. (Fig. 5D). By using primers B13R and 47.1R, a product of 4 kb was obtained and sequenced. The sequence 31.2 was then found to be derived from an *Alu*I fragment situated between 47.1 and B13 (Fig. 5A).

Screening of an *M. ulcerans* cosmid library for the *pks* locus. A previously constructed cosmid library of *M. ulcerans* Chant genomic DNA was screened for the *pks* locus. Two cosmids (5G10 and 8C6) were confirmed by PCR and Southern hybridization analysis to contain both the sequences 74.1 and 4.1. Sequencing of the termini of the cosmid inserts revealed that were derived from different DNA segments with the arrangements shown in Fig. 5B and C, respectively (see below). Sequence from the cosmid placed B15 immediately upstream of 4.1. No cosmid which contained sequence 54.2, 18.1, or 51.2 was identified.

Southern hybridization demonstrates polymorphism at the *pks* locus. A Southern hybridization analysis of different strains of *M. ulcerans* showed that the 5145 strain had a deletion of a 10-kb *Eco*RI fragment (Fig. 6) which contains sequences that hybridize to clones 18.1 and 4.1. With a probe derived from clone 74.1 and the same *Eco*RI digest, the *M. ulcerans* 5145 and Chant strains showed two bands, but only a single 20-kb band was seen in *M. ulcerans* Benin 9825.

Analysis of the modular organization of the polyketide synthase. The modular organization of type I polyketide synthases is characterized by a repeated arrangement of functional domains grouped into discrete modules. Functional domains were identified within the predicted amino acid sequence from the sequenced regions by using the Pfam HMM server (http://pfam.wustl.edu/) and also by comparison with those identified in other type I polyketide synthases.

The region including the 47.1, 74.1, B3, and B60 subtraction sequences was predicted to encode a loading module β-ketosynthase domain, as can be differentiated from other β-ketosynthase domains by the replacement of the normal active-site cysteine with a glutamine residue (Fig. 7A) (6). The acyl transferase domain of this loading module would be expected to use an acyl group as the initiating carbon by sequence homology of the specificity domain, as described previously (Fig. 7B) (26). In addition, the presence of the KSQ domain and a characteristic arginine at position 117 of the acyl transferase domain

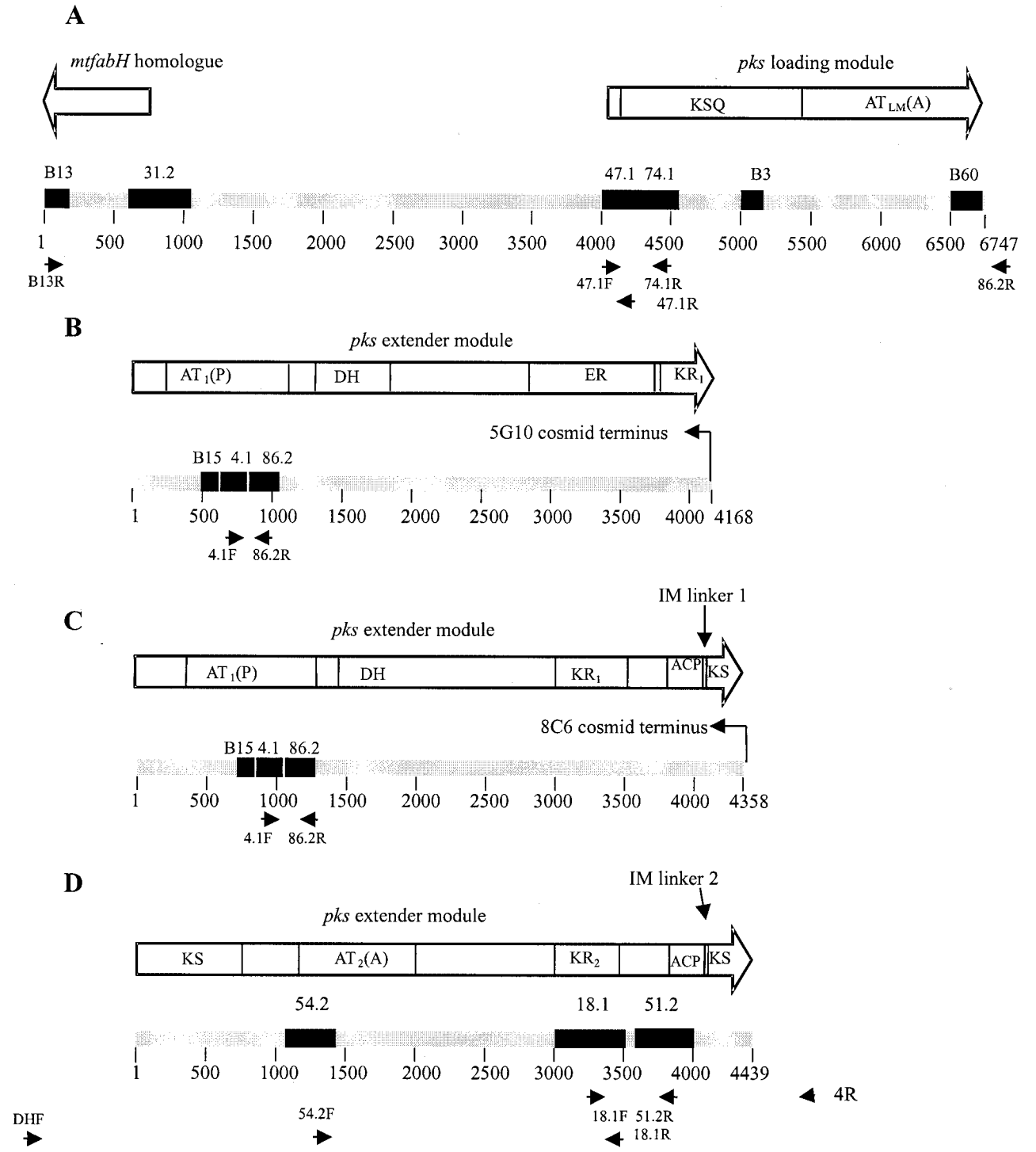

FIG. 5. Diagram of the *M. ulcerans pks* gene fragments sequenced for this study, showing the relative positions of the SSH-derived clones (black boxes). (A) Loading module (AY289593); (B to D) extender modules (AY289594, AY289595, and AY289596, respectively). Displayed above the gene fragments are the functional domains of the predicted translated protein. The primers used to link subtraction clones by PCR are also indicated. KSQ, loading module ketosynthase; KS, β-ketoacyl synthase; AT(A), acyltransferase-accepting acyl group; AT(P), acyltransferase-accepting propionyl group; KR, β-ketoacyl reductase; DH, dehydratase; ER, enoyl reductase; ACP, acyl carrier protein; IM, intermodular linker.

(Fig. 7B) indicates that the initiating group for polyketide synthesis would be a dicarboxylic acid CoA ester (26).

Partial or complete sequences of three extender module domains centered on the identified subtraction clones have also been sequenced (Fig. 5B, C, and D). Two intermodular linkers with a characteristic proline residue (19) were identified in the translated protein sequence (Fig. 5C and D). β-Keto group reduction domains were identified in each of the mod-

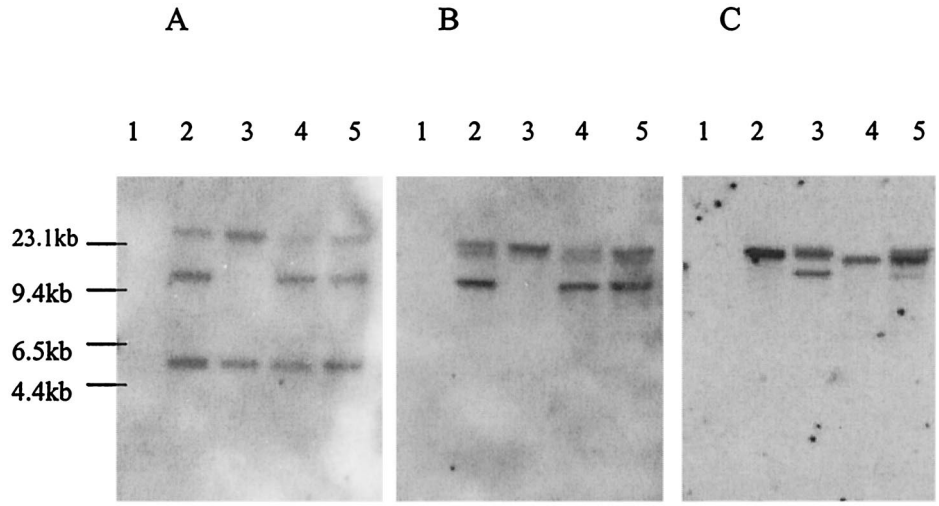

FIG. 6. Southern hybridization of *Eco*RI digests of DNAs from *M. marinum* JKD 2396 (lane 1) and *M. ulcerans* (lane 2, *M. ulcerans* 13822/70; lane 3, *M. ulcerans* 5145; lane 4, *M. ulcerans* Benin 9825; and lane 5, *M. ulcerans* Chant) with DIG-labeled probes to clone 18.1 (A), clone 4.1 (B), and clone 74.1 (C) (Table 3).

ules. The ketoreductase domains showed 36% (KR_1) and 43% (KR_2) identity to the same domains from erythromycin synthase module 1 of *Saccharopolyspora erythraea* (accession no. Q03131). The dehydratase domain and enoyl reductase domains displayed 37 and 50% identity, respectively, to the same domains from OleAII of *Streptomyces antibioticus* (AAF82409). From the differing β-keto processing arrangements of each of the three extender domains and their AT domain carbon preferences, they could fit within the putative modules required for synthesis of either the core or side chain of mycolactone (21). The extender modules sequenced from two separate cosmid termini (Fig. 5B and C) have nearly identical sequences, including an identical acyl transferase domain, except that an enoyl reductase domain is present in module B but absent from module C. An alignment of the amino acid sequences for this region of sequence divergence is shown in Fig. 8.

The previously mentioned *mtfabH*/*dpsC* homologue (B13), as well as the intergenic sequence 31.2, was linked to the region upstream of the *pks* gene (Fig. 5A). The translated protein sequence of this gene fragment demonstrates replacement of an active-site cysteine (as seen in other KASIII homologues, including mtFabH) by serine (Fig. 9). This replacement is seen in DpsC and AknE2 (another anthracycline polyketide synthase-associated gene, from *S. galilaeus*), suggesting that the gene homologue in *M. ulcerans* may have a similar function in determining starter unit selection (32).

Attempts to PCR amplify regions between the modules have been made. A 6-kb product was obtained by using 18.1R and a forward primer to a dehydratase domain (primer DHF [Table 2]), designed from sequence obtained after sequencing of the two cosmid extender modules and already known to be absent from the sequence between 54.2 and 18.1. A product of 2.5 kb was also obtained between 51.2F and 4.1R. However, it is not possible to determine from our data which of the modules shown in Figs. 5B and C are adjacent to extender module D. Partial sequencing of these PCR products provided the sequence upstream of 54.2 and downstream of 51.2 (Fig. 5D).

Because the DNA sequence encoding domains across the various modules has been found to be highly repetitive, it has proved difficult to definitively organize the sequenced modules relative to one another. For example the sequences B64, B68, and B60 have nearly identical sequences but demonstrate single-nucleotide polymorphisms that have introduced different *Alu*I restriction sites resulting in different fragment sizes. Alignments obtained from the cosmid sequence indicate that B60 is from the loading module acyl transferase, while B64 comes from another acyl transferase whose relationships to other domains could not be resolved from our limited sequence. The sequence B68 is identical to 86.2, which suggests that the observed polymorphisms reflect differences across domains rather than strain variations.

It became apparent that multiple copies of the modules were present at the *pks* locus, making attempts to link modules by PCR problematic. Efforts to map the locus by Southern hybridization were also hampered by the absence of an 18.1-containing cosmid from our library. Complete sequencing of this region is ongoing in collaboration with the genome sequencing project at the Institut Pasteur.

The GC contents of the segments of *pks* were calculated as 60 to 62%, which is similar to that seen in whole genomes of other mycobacteria and to the predicted genome GC percentage for *M. ulcerans* (http://genepole.pasteur.fr/Mulc/burulist.html). Codon usage analysis using relative synonymous codon usage calculations for selected amino acids was similar to that seen with averaged coding regions present in the database (http://www.kazusa.or.jp/codon) for other mycobacteria.

DISCUSSION

This study confirms the effectiveness of genomic subtractive hybridization of closely related mycobacteria in isolating inser-

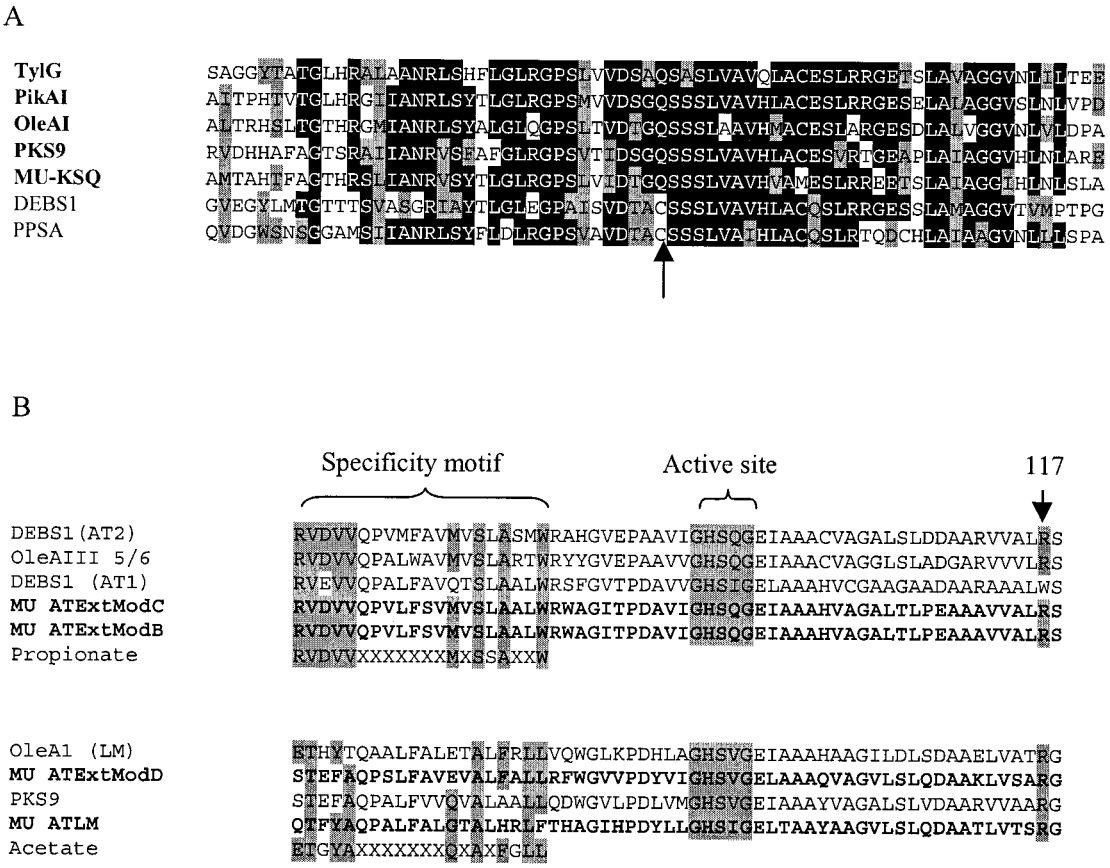

FIG. 7. (A) ClustalW alignment of the amino acid sequences of the β-ketosynthase active sites from type I polyketide synthase loading modules and the *M. ulcerans* KSQ (MU-KSQ) identified in this study. For comparison, the active sites of the β-ketosynthases from the erythromycin synthase module 1 (DEBS1) of *S. erythraea* (accession no. Q03131) and phenolpthiocerol synthesis polyketide synthase (PPSA) from *M. tuberculosis* (accession no. Q10977) are included to demonstrate the replacement of the active-site cysteine by glutamine in β-ketosynthase loading modules (arrow). (B) Amino acid alignment of the acyl transferase (AT) specificity and active sites of the *M. ulcerans* loading module (MU ATLM) and extender modules (MU ATExtModB to -D) with other type I polyketide synthases. The active site, specificity motifs, and position 117 arginine are indicated (26). TylG, tylactone synthase loading module from *Streptomyces fradiae* (accession no. AAB66504); PikAI, pikromycin synthase loading module from *Streptomyces venezuelae* (AAC69329); OleAI and OleAIII, deoxyoleandolide synthase loading modules from *Streptomyces antibioticus* (modules 1/2 [AAF82408] and 5/6 [Q07017], respectively); PKS9, from *M. tuberculosis* (NP_216180).

tions of the tester organism relative to the driver organism. We have identified by SSH a type I polyketide synthase locus which is apparently present in *M. ulcerans* but absent from *M. marinum*. This locus was identified by two independent experiments using different strains of *M. ulcerans* and *M. marinum*, which implies that the major genetic acquisition by *M. ulcerans* (other than the previously identified IS*2404* and IS*2606*) is the identified *pks* locus.

The data are consistent with *M. ulcerans* evolving from *M. marinum* according to the general principles observed in the evolution of other mycobacteria, in that strain variation and speciation events are driven by insertions, deletions, and recombination events rather than by codon mutation. In particular, this has been extensively shown for the *M. tuberculosis* complex (8) in comparisons of *M. bovis* BCG with *M. bovis* and of *M. leprae* with *M. tuberculosis* (9). We were not able to identify DNA fragments that are present in *M. marinum* but deleted from *M. ulcerans* (data not shown), and we have not investigated recombination events within the *M. ulcerans* chromosome relative to *M. marinum*, which are also likely to be important in determining the unique phenotype of *M. ulcerans* and its pathogenesis (38).

Previous mycobacterial genomic subtractions have identified a wide variety of genes. In *M. avium* subsp. *paratuberculosis*, a pathogenicity island containing genes that may have been involved in cell wall polysaccharide modification was identified (40). Twenty-one genes were found in *M. avium* that were absent from *M. intracellulare*, and some of these were postulated to have a role in invasiveness of the intestinal mucosa (28). The *pks* locus that we have identified does not have the characteristics of a pathogenicity island in that the GC content is similar to that of mycobacteria, although we have shown by Southern hybridization analysis that IS*2404* and IS*2606* are both present near the *pks* locus (data not shown).

This SSH technique includes an initial hybridization step that compensates for variation in gene copy number. This was considered important, because the high copy number of IS*2404* in *M. ulcerans* threatened to overwhelm any subtraction, making it difficult to identify single copy number differences between *M. ulcerans* and *M. marinum*. The known presence of

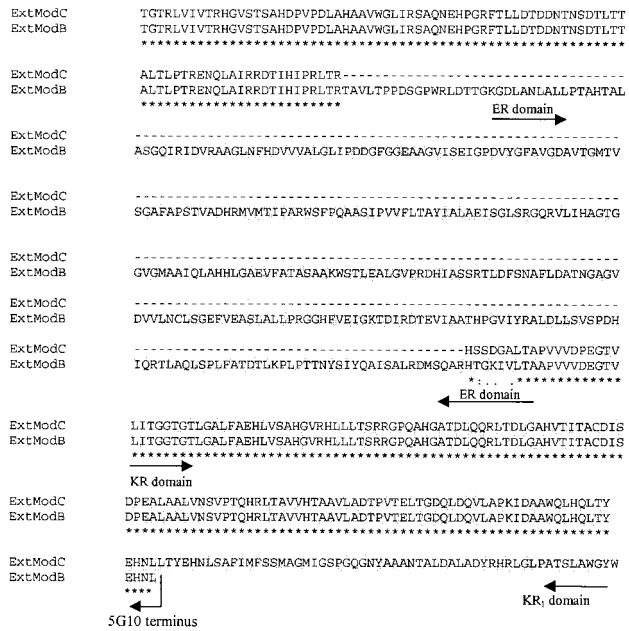

FIG. 8. Alignment of the amino acid sequences for extender modules B and C to show the conservation of sequence between the modules and the presence of the enoyl reductase (ER) domain in module B. The ketoreductase (KR_1) domain is also shown and is truncated in module B at the terminus of the cosmid 5G10 insert.

IS*2404* also acted as an internal positive control to confirm the effectiveness of subtraction. Of the DNA sequences that appear to be specific to *M. ulcerans* but did not have any significant BLAST matches, only one (sequence 4.3) potentially includes a full open reading frame. These may represent specific *M. ulcerans* intergenic sequences, and indeed, the sequence 31.2 was mapped to the upstream region of the *pks* loading module and found to overlap the start of the *mtfabh*/*dpsC* homologue. We have not been able to map any of the other specific sequences to the *pks* locus at this stage. As the genome sequences for *M. ulcerans* and *M. marinum* are assembled, we will attempt to define the specificity and relationships of these sequences. Sequences isolated in the subtractions which appear to be present in *M. marinum* presumably reflect amplifications of rare unsubtracted DNA with significant homologies between driver and tester. This is suggested as the likely explanation, given that no such sequences were obtained from either independent SSH experiment.

Within the *pks* locus, certain apparent hot spots were identified in each experiment independently. This may be a function of *Alu*I fragment size or PCR efficiency related to the template sequence, but these areas may also be the sites of maximal DNA sequence difference compared with other *pks* loci in the *M. marinum* and *M. ulcerans* genomes. Polyketide synthase modules across species share significant amino acid sequence homology in functional domains, and therefore not all areas in a *pks* module may be identified in the SSH because of similar DNA sequences in *M. marinum pks* clusters.

The *M. tuberculosis* genome contains 18 polyketide synthase genes (13), and the potent and varied biological activities of polyketides make them intriguing candidates as virulence determinants in mycobacterial infections. They are known to be involved in mycobactin siderophore synthesis (16) and in synthesis of cell wall lipids that are likely to be involved in pathogenesis (3, 4, 15). Signature-tagged mutagenesis (11) suggested that the *pps* polyketide synthase locus, which is responsible for the synthesis of phthiocerol and phenolphthiocerol, has a role in pathogenesis. Signature-tagged mutagenesis also identified *pks6* (Rv0405) of *M. tuberculosis* and a *pks6* homologue in *M. marinum* as virulence factors (42). *pks2* (Rv 3825c) was recently shown to be involved in synthesis of cell wall sulfolipids (35) and has previously been identified as being upregulated in *M. tuberculosis* after macrophage phagocytosis (20). *M. leprae*, despite marked gene degeneration, still has six functional *pks* genes, implying that they may be essential genes or at least important in the host-pathogen interaction (14).

The structure of the identified *M. ulcerans pks* locus indicates it encodes a type I polyketide synthase protein with significant homology to those from *Streptomyces* spp. and mycobacteria. Given that mycolactone is produced by *M. ulcerans* and not *M. marinum* (18) and that a type I polyketide synthase would be expected to be responsible for its synthesis, we hypothesize that the *M. ulcerans*-specific *pks* encodes a mycolactone synthase. The core 12-member-ring macrolactone of mycolactone may be synthesized by a type I polyketide synthase containing a loading module plus nine extender modules, and the side chain may be synthesized by another polyketide synthase consisting of a loading module plus seven extender modules (Fig. 10). Our finding that the *M. ulcerans*-specific *pks* locus is contained on a 135-kb *Ase*I fragment is consistent with a recent report that the mycolactone polyketide synthase genes are confined to a 100-kb region of DNA (29). Genomic hybridizations suggest that two copies of the loading module are present in the genomes of some *M. ulcerans* strains (Fig. 6), which would be consistent with a model whereby synthesis of both the side chain and central ring of mycolactone is initiated by decarboxylative incorporation of malonyl-CoA resulting in loading of an acyl group on to the first extender module prior to the first extension reaction. The structure of mycolactone suggests that the loading module and first extender unit could be identical in each synthetic locus, as carbons 18 to 20 (macrolactone) and 14′ to 16′ (fatty acid side chain), respectively, have the same structure (Fig. 10). The characteristics of the identified loading module therefore coincide with those anticipated for the mycolactone synthase.

We have evidence of strain variation in the *pks* locus as indicated by deletion of the 10-kb *Eco*RI fragment in *M. ulcerans* 5145. Recombination events have been seen with some other polyketide synthase loci, such as the tylosin locus of *Streptomyces fradiae* (24), and presumably result from the high sequence similarity between equivalent domains of adjacent modules. Heterogeneity in the mycolactones produced by *M. ulcerans* strains has been reported (29), and it is possible that recombination of the *pks* locus genes contributes to this heterogeneity.

Inactivation of the *M. ulcerans pks* loading module would be expected to abolish synthesis of all polyketides from that locus (6, 45). We have attempted to do this by double-crossover allelic exchange with a suicide vector system carrying a portion of the *pks* loading module disrupted by an antibiotic resistance cassette, but we have not been successful (data not shown).

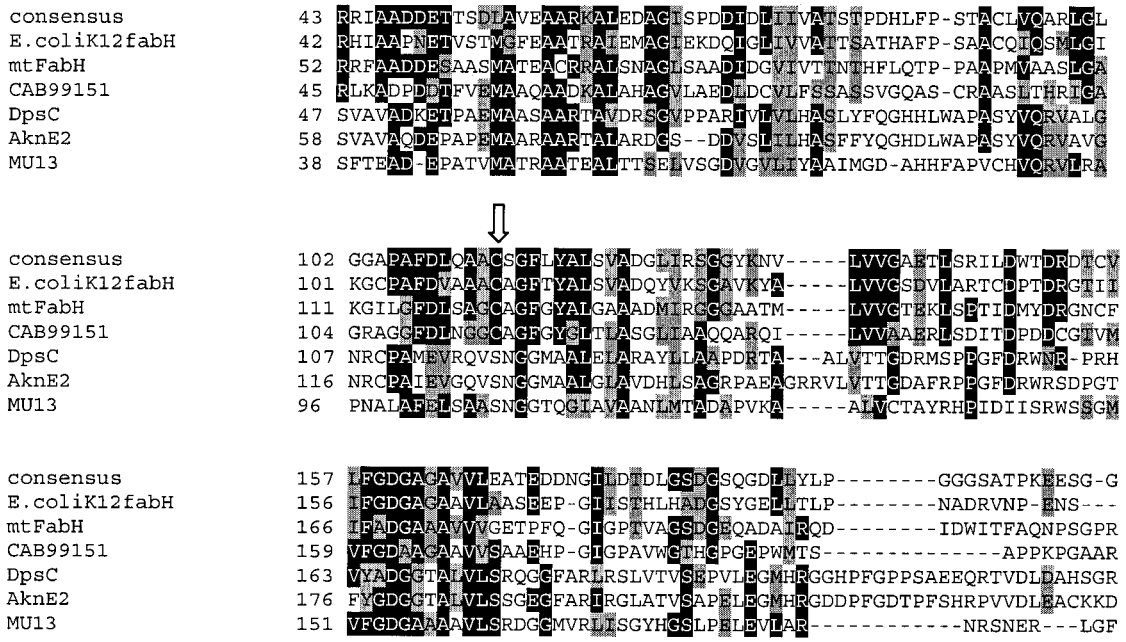

FIG. 9. Partial ClustalW alignment of β-ketoacyl:ACP synthase III (KASIII or FabH) proteins with the *M. ulcerans* homologue identified in this study. The arrow points to the putative active site, which is a cysteine in most family members and the consensus but is a serine in DpsC, AknE2, and the *M. ulcerans* homologue (MU13). The consensus sequence was generated by RBS-BLAST through GenBank alignment of all sequences for protein family COG0332 by using the National Center for Biotechnology Information conserved domain database server. *E. coli* K-12 FabH, accession no. NP_415609; DpsC (from *Streptomyces coelicolor* 2SCG18), CAB99151; AknE2 (from *S. galilaeus*), AAF70109. Accession number AAA87620 is from *Streptomyces* sp. strain C5.

The transformation efficiency achieved by electroporation of *M. ulcerans* under standard conditions for slow-growing mycobacteria (30) with the shuttle vector pMV261 (39) has been only 10^1 to 10^2 transformants/μg of DNA, a level which is likely to be too low to achieve successful homologous recombination in mycobacteria. We are currently undertaking experiments to express the *M. ulcerans pks* locus in *M. marinum* and *M. smegmatis*.

Where genome sequences are available, comparisons can be performed by using microarray or gene chip technology to detect strain differences between isolates of the same or closely related species (5, 25). The ongoing genome sequencing projects for *M. ulcerans* (http://genepole.pasteur.fr/Mulc/burulist.html) and *M. marinum* (http://www.sanger.ac.uk/Projects/M_marinum/) will allow an in silico analysis of the genomic differences between these two organisms, including genetic rearrangements, and should define the limits of difference at the identified *pks* locus between the *M. ulcerans* and *M. marinum* strains used in these projects.

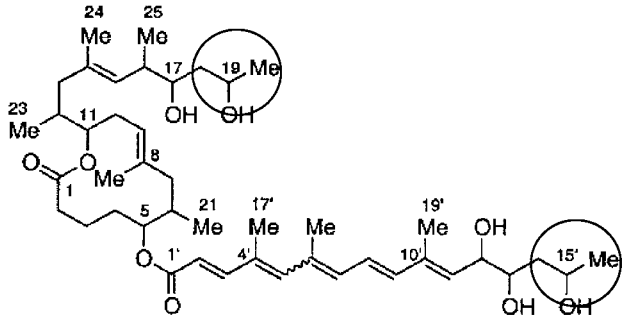

FIG. 10. Structure of mycolactone (21), showing the macrolactone core structure (C1 to C25) and the side chain (C1′ to C19′). The terminal three carbons (C18 to C20 and C14′ to C16′), which are anticipated to be synthesized by the loading module plus the first extender module of the respective polyketide synthases, are circled.

ACKNOWLEDGMENTS

G.J. was financially supported by National Health and Medical Research Council medical postgraduate scholarship 997451.

We thank Frances Oppedisano of the Microbiological Research Unit, Department of Microbiology and Infectious Diseases, Royal Children's Hospital, Parkville, Australia, for providing the *M. ulcerans* Chant strain cosmid library.

REFERENCES

1. **Akopyants, N. S., A. Fradkov, L. Diatchenko, J. E. Hill, P. D. Siebert, S. A. Lukyanov, E. D. Sverdlov, and D. E. Berg.** 1998. PCR-based subtractive hybridisation and differences in gene content among strains of *Helicobacter pylori*. Proc. Natl. Acad. Sci. USA **95:**13108–13113.
2. **Amofah, G., F. Bonsu, C. Tetteh, J. Okrah, K. Asamoa, K. Asiedu, and J. Addy.** 2002. Buruli ulcer in Ghana: results of a national case search. Emerg. Infect. Dis. **8:**167–170.
3. **Azad, A. K., T. D. Sirakova, N. D. Fernandes, and P. E. Kolattukudy.** 1997. Gene knockout reveals a novel gene cluster for the synthesis of a class of cell wall lipids unique to pathogenic mycobacteria. J. Biol. Chem. **272:**16741–16745.
4. **Azad, A. K., T. D. Sirakova, L. M. Rogers, and P. E. Kolattukudy.** 1996. Targeted replacement of the mycocerosic acid synthase gene in *Mycobacterium bovis* BCG produces a mutant that lacks mycosides. Proc. Natl. Acad. Sci. USA **93:**4787–4792.
5. **Behr, M. A., M. A. Wilson, W. P. Gill, H. Salamon, G. K. Schoolnik, S. Rane, and P. M. Small.** 1999. Comparative genomics of BCG vaccines by whole-genome DNA microarray. Science **284:**1520–1523.
6. **Bisang, C., P. F. Long, J. Cortes, J. Westcott, J. Crosby, A. L. Matharu, R. J.**

Cox, T. J. Simpson, J. Staunton, and P. F. Leadlay. 1999. A chain initiation factor common to both modular and aromatic polyketide synthases. Nature **401:**502–505.
7. **Boddinghaus, B., T. Rogall, T. Flohr, H. Blocker, and E. C. Bottger.** 1990. Detection and identification of mycobacteria by amplification of rRNA. J. Clin. Microbiol. **28:**1751–1759.
8. **Brosch, R., S. Gordon, M. Marmiesse, P. Brodin, C. Buchrieser, K. Eiglmeier, T. Garnier, C. Gutierrez, G. Hewinson, K. Kremer, L. Parsons, A. Pym, S. Samper, D. van Soolingen, and S. Cole.** 2002. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. Proc. Natl. Acad. Sci. USA **99:**3684–3689.
9. **Brosch, R., S. V. Gordon, A. Pym, K. Eiglmeier, T. Garnier, and S. T. Cole.** 2000. Comparative genomics of the mycobacteria. Int. J. Med. Microbiol. **290:**143–152.
10. **Brosch, R., A. S. Pym, S. V. Gordon, and S. T. Cole.** 2001. The evolution of mycobacterial pathogenicity: clues from comparative genomics. Trends Microbiol. **9:**452–458.
11. **Camacho, L. R., D. Ensergueix, E. Perez, B. Gicquel, and C. Guilhot.** 1999. Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. Mol. Microbiol. **34:**257–267.
12. **Choi, K. H., L. Kremer, G. S. Besra, and C. O. Rock.** 2000. Identification and substrate specificity of beta-ketoacyl (acyl carrier protein) synthase III (mtFabH) from *Mycobacterium tuberculosis*. J. Biol. Chem. **275:**28201–28207.
13. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, B. G. Barrell, et al.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537–544.
14. **Cole, S. T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honore, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell.** 2001. Massive gene decay in the leprosy bacillus. Nature **409:**1007–1011.
15. **Constant, P., E. Perez, W. Malaga, M.-A. Laneelle, O. Saurel, M. Daffe, and C. Guilhot.** 2002. Role of the *pks15/1* gene in the biosynthesis of phenolglycolipids in the *Mycobacterium tuberculosis* complex. J. Biol. Chem. **277:**38148–38158.
16. **de Voss, J. J., K. Rutter, B. G. Schroeder, H. Su, Y. Zhu, and C. E. Barry III.** 2000. The salicylate-derived mycobactin siderophore of *Mycobacterium tuberculosis* are essential for growth in macrophages. Proc. Natl. Acad. Sci. USA **97:**1252–1257.
17. **Falkinham, J. O.** 1996. Epidemiology of infection by nontuberclous mycobacteria. Clin. Microbiol. Rev. **9:**177–215.
18. **George, K. M., D. Chatterjee, G. Gunawardana, D. Welty, J. Hayman, R. Lee, and P. L. Small.** 1999. Mycolactone: a polyketide toxin from *Mycobacterium ulcerans* required for virulence. Science **283:**854–857.
19. **Gokhale, R. S., S. Y. Tsuji, D. E. Cane, and C. Khosla.** 1999. Dissecting and exploiting intermodular communication in polyketide synthases. Science **284:**482–485.
20. **Graham, J. E., and J. E. Clark-Curtiss.** 1999. Identification of *Mycobacterium tuberculosis* RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS). Proc. Natl. Acad. Sci. USA **96:**11554–11559.
21. **Gunawardana, G., D. Chatterjee, K. M. George, P. Brennan, D. Whittern, and P. L. C. Small.** 1999. Characterization of novel macrolide toxins, mycolactones A and B, from a human pathogen, *Mycobacterium ulcerans*. J. Am. Chem. Soc. **121:**6092–6093.
22. **Hayman, J.** 1985. Clinical features of *Mycobacterium ulcerans* infection. Austr. J. Dermatol. **26:**67–73.
23. **Hayman, J., and A. McQueen.** 1985. The pathology of *Mycobacterium ulcerans* infection. Pathology **17:**594–600.
24. **Hopwood, D. A.** 1997. Genetic contributions to understanding polyketide synthases. Chem. Rev. **97:**2464–2497.
25. **Kato-Maeda, M., J. T. Rhee, T. R. Gingeras, H. Salamon, J. Drenkow, N. Smittipat, and P. M. Small.** 2001. Comparing genomes within the species *Mycobacterium tuberculosis*. Genome Res. **11:**547–554.
26. **Long, P., C. Wilkinson, C. Bisang, J. Cortes, N. Dunster, M. Oliynyl, E. McCormick, H. McArthur, C. Mendez, S. J. A., J. Staunton, and P. Leadley.** 2002. Engineering specificity of starter unit selection by the erythromycin-producing polyketide synthase. Mol. Microbiol. **43:**1215–1225.
27. **Mahairas, G. G., P. J. Sabo, M. J. Hickey, D. C. Singh, and C. K. Stover.** 1996. Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. J. Bacteriol. **178:**1274–1282.
28. **McGarvey, J. A., and L. E. Bermudez.** 2001. Phenotypic and genomic analyses of the *Mycobacterium avium* complex reveal differences in gastrointestinal invasion and genomic composition. Infect. Immun. **69:**7242–7249.
29. **Mve-Obiang, A., R. E. Lee, F. Portaels, and P. L. C. Small.** 2003. Heterogeneity of mycolactones produced by clinical isolates of *Mycobacterium ulcerans*: implications for virulence. Infect. Immun. **71:**774–783.
30. **Parish, T., and N. G. Stoker.** 1998. Electroporation of mycobacteria. Methods Mol. Biol. **101:**129–144.
31. **Portaels, F., P. A. Fonteyene, H. de Beenhouwer, P. de Rijk, A. Guedenon, J. Hayman, and M. W. Meyers.** 1996. Variability in 3′ end of 16S rRNA sequence of *Mycobacterium ulcerans* is related to geographic origin of isolates. J. Clin. Microbiol. **34:**962–965.
32. **Rajgarhia, V. B., N. D. Priestley, and W. R. Strohl.** 2001. The product of *dpsC* confers starter unit fidelity upon the daunorubicin polyketide synthase of *Streptomyces* sp. strain C5. Metab. Eng. **3:**49–63.
33. **Ross, B. C., L. Marino, F. Oppedisano, R. Edwards, R. M. Robins-Browne, and P. D. Johnson.** 1997. Development of a PCR assay for rapid diagnosis of *Mycobacterium ulcerans* infection. J. Clin. Microbiol. **35:**1696–1700.
34. **Rozen, S., and H. J. Skaletsky.** 2000. Primer3 on the WWW for general users and for biologist programmers, p. 365–386. *In* S. Krawetz and S. Misener (ed.), Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, N.J.
35. **Sirakova, T. D., A. K. Thirumala, V. S. Dubey, H. Sprecher, and P. E. Kolattukudy.** 2001. The *Mycobacterium tuberculosis pks2* gene encodes the synthase for the hepta- and octamethyl-branched fatty acids required for sulfolipid synthesis. J. Biol. Chem. **276:**16833–16839.
36. **Stinear, T., J. K. Davies, G. A. Jenkin, J. A. Hayman, F. Oppedisano, and P. D. Johnson.** 2000. Identification of *Mycobacterium ulcerans* in the environment from regions in southeast Australia in which it is endemic with sequence capture-PCR. Appl. Environ. Microbiol. **66:**3206–3213.
37. **Stinear, T., B. C. Ross, J. K. Davies, L. Marino, R. M. Robins-Browne, F. Oppedisano, A. Sievers, and P. D. Johnson.** 1999. Identification and characterization of IS*2404* and IS*2606*: two distinct repeated sequences for detection of *Mycobacterium ulcerans* by PCR. J. Clin. Microbiol. **37:**1018–1023.
38. **Stinear, T. P., G. A. Jenkin, P. D. Johnson, and J. K. Davies.** 2000. Comparative genetic analysis of *Mycobacterium ulcerans* and *Mycobacterium marinum* reveals evidence of recent divergence. J. Bacteriol. **182:**6322–6330.
39. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. P. Young, M. H. Lee, G. F. Hatfull, S. B. Snapper, R. G. Barletta, W. R. Jacobs, Jr., and B. R. Bloom.** 1991. New use of BCG for recombinant vaccines. Nature **351:**451–460.
40. **Tizard, M., T. Bull, D. Nillar, T. Doran, H. Martin, N. Sumar, J. Ford, and J. Hermon-Taylor.** 1998. A low G+C content genetic island in *Mycobacterium avium* subsp. *paratuberculosis* and *M. avium* subsp. *silvaticum* with homologous genes in *Mycobacterium tuberculosis*. Microbiology **144:**3413–3423.
41. **Tonjum, T., D. B. Welty, E. Jantzen, and P. L. Small.** 1998. Differentiation of *Mycobacterium ulcerans*, *M. marinum*, and *M. haemophilum*: mapping of their relationships to *M. tuberculosis* by fatty acid profile analysis, DNA-DNA hybridization, and 16S rRNA gene sequence analysis. J. Clin. Microbiol. **36:**918–925.
42. **Trucksis, M.** 2000. Fishing for mycobacterial virulence genes: a promising animal model. ASM News **66:**668–674.
43. **Wayne, L. G.** 1984. Mycobacterial speciation, p. 25–65. *In* G. P. Kubica and L. G. Wayne (ed.), The mycobacteria: a sourcebook. Dekker, New York, N.Y.
44. **World Health Organization.** 1998. International conference on Buruli ulcer control and research. World Health Organization, Geneva, Switzerland.
45. **Xue, Y., L. Zhao, H. W. Liu, and D. H. Sherman.** 1998. A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: architecture of metabolic diversity. Proc. Natl. Acad. Sci. USA **95:**12111–12116.
46. **Zhang, Y. L., C. T. Ong, and K. Y. Leung.** 2000. Molecular analysis of genetic differences between virulent and avirulent strains of *Aeromonas hydrophila* isolated from diseased fish. Microbiology **146:**999–1009.