JOURNAL OF CLINICAL MICROBIOLOGY, Aug. 2004, p. 3594–3599 Vol. 42, No. 8
0095-1137/04/$08.00+0 DOI: 10.1128/JCM.42.8.3594–3599.2004

Genomic Analysis Distinguishes *Mycobacterium africanum*

Serge Mostowy,[1] Anthony Onipede,[2] Sebastien Gagneux,[2] Stefan Niemann,[3] Kristin Kremer,[4] Edward P. Desmond,[5] Midori Kato-Maeda,[2] and Marcel Behr[1]*

McGill University Health Centre, Montreal, Canada H3G 1A4[1]; Division of Infectious Diseases and Geographic Medicine, Department of Medicine, Stanford University, Stanford, California 94305[2]; Forschungszentrum Borstel, National Reference Center for Mycobacteria, D-23845 Borstel, Germany[3]; Mycobacteria Reference Department, Diagnostic Laboratory for Infectious Diseases and Perinatal Screening, National Institute of Public Health and the Environment, 3720 BA Bilthoven, The Netherlands[4]; and Mycobacteriology and Mycology Section, Microbial Diseases Laboratory, California Department of Health Services, Richmond, California 94804[5]

Received 22 March 2004/Returned for modification 28 April 2004/Accepted 7 May 2004

***Mycobacterium africanum* is thought to comprise a unique species within the *Mycobacterium tuberculosis* complex. *M. africanum* has traditionally been identified by phenotypic criteria, occupying an intermediate position between *M. tuberculosis* and *M. bovis* according to biochemical characteristics. Although *M. africanum* isolates present near-identical sequence homology to other species of the *M. tuberculosis* complex, several studies have uncovered large genomic regions variably deleted from certain *M. africanum* isolates. To further investigate the genomic characteristics of organisms characterized as *M. africanum*, the DNA content of 12 isolates was interrogated by using Affymetrix GeneChip. Analysis revealed genomic regions of *M. tuberculosis* deleted from all isolates of putative diagnostic and biological consequence. The distribution of deleted sequences suggests that *M. africanum* subtype II isolates are situated among strains of "modern" *M. tuberculosis*. In contrast, other *M. africanum* isolates (subtype I) constitute two distinct evolutionary branches within the *M. tuberculosis* complex. To test for an association between deleted sequences and biochemical attributes used for speciation, a phenotypically diverse panel of "*M. africanum*-like" isolates from Guinea-Bissau was tested for these deletions. These isolates clustered together within one of the *M. africanum* subtype I branches, irrespective of phenotype. These results indicate that convergent biochemical profiles can be independently obtained for *M. tuberculosis* complex members, challenging the traditional approach to *M. tuberculosis* complex speciation. Furthermore, the genomic results suggest a rational framework for defining *M. africanum* and provide tools to accurately assess its prevalence in clinical specimens.**

Species of the *Mycobacterium tuberculosis* complex traditionally consist of *Mycobacterium tuberculosis, Mycobacterium africanum, Mycobacterium microti,* and *Mycobacterium bovis*. Genomically, species of the *M. tuberculosis* complex present near-identical sequence homology (15, 16, 40) but are most notably differentiated by large sequence polymorphisms (6, 30, 31, 34). Variants of these original species have since been identified via characteristic phenotypic or genotypic attributes, challenging taxonomic classification within the complex (30). As a result, other members of the *M. tuberculosis* complex have emerged, including "*Mycobacterium canettii*" (35, 44), *Mycobacterium caprae* (1, 2), *Mycobacterium pinnipedii* (10, 12), and the dassie bacillus (11, 30).

In contrast to *M. tuberculosis* and *M. bovis*, for which phenotype-based speciation is reliant upon concordant results for a number of biochemical tests, *M. africanum* isolates can present the entire spectrum of biochemical properties employed to discriminate among the *M. tuberculosis* complex. As a result, "*M. tuberculosis*-like" and "*M. bovis*-like" forms of *M. africanum* have been reported, potentially confounding laboratory identification and complicating attempts to accurately determine the true prevalence of *M. africanum* disease (13). Typically, *M. africanum* is subdivided by geographic origin and biochemical properties into two major subgroups: *M. africanum* subtype I originates from West Africa and exhibits *M. bovis*-like properties, while *M. africanum* subtype II originates from East Africa and exhibits *M. tuberculosis*-like properties (13). As part of an epidemiological study of tuberculosis in Guinea-Bissau (22), coexisting isolates presenting a complete spectrum of phenotypic variability was observed and subsequently identified as "*M. africanum*-like." These results challenged the notion of a simple dichotomous subclassification between *M. tuberculosis*-like and *M. bovis*-like forms and stimulated further analysis to determine the basis for this phenotypically variable group of organisms (25).

In recent studies using regions of difference (RD) to discriminate members of the *M. tuberculosis* complex, isolates characterized as *M. africanum* have presented three distinct genetic groupings (6, 31, 34, 39). In the first group are organisms in which TbD1 is missing but the regions RD7, RD8, RD9 and RD10 are present. Some have suggested that these organisms should be called *M. tuberculosis* (6, 34, 39). A second group consists of *M. africanum* strains in which RD9 is deleted but RD7, RD8, and RD10 are still present (6, 18, 31). Finally, a third group of *M. africanum* comprises organisms in which RD7, RD8, RD9, and RD10 are all absent (6, 31). Because of the different observed genomic profiles for organisms classified as *M. africanum*, the DNA content of 12 *M. africanum* isolates was interrogated by Affymetrix GeneChip. Characteristic poly-

* Corresponding author. Mailing address: Division of Infectious Diseases and Medical Microbiology, A5-156, Montreal General Hospital, 1650 Cedar Ave., Montreal QC H3G 1A4, Canada. Phone: (514) 934-1934 42815. Fax: (514) 934 8423. E-mail: marcel.behr@mcgill.ca.

TABLE 1. Characteristics of bacterial isolates[a]

Sample	*M. africanum* subtype	Origin	Reference(s)	Spoligotype	Genetic composition			
					*katG*463	*gyrA*95	*pks15/1*	TbD1
K100/RIVM13876	Subtype II	Netherlands	21,26		CTG	ACC	Intact	−
D2169_99	Subtype II	Uganda	33		CGG	ACC	7 bp	−
D2285_99	Subtype II	Uganda	33		CGG	ACC	7 bp	−
D2241_99	Subtype II	Uganda	33		CGG	ACC	7 bp	−
D10409_02	Subtype I	Ghana	S. Niemann		CTG	ACC	Intact	+
CPHL_A	*M. africanum*	South Africa	E. Desmond		CTG	ACC	Intact	+
D10415_02	Subtype I	Ghana	S. Niemann		CTG	ACC	6 bp	+
K6/RIVM17316	Subtype I	Netherlands	21,26		CTG	ACC	6 bp	+
K85/RIVM17902	Subtype I	Netherlands	21,26		CTG	ACC	6 bp	+
ATCC25420	Subtype I	Senegal	7		CTG	ACC	6 bp	+
ATCC35711	Subtype I	Senegal	7		CTG	ACC	6 bp	+
SN68	"*M. africanum*-like"[b]	Guinea-Bissau	22		CTG	ACC	6 bp	+

[a] Sample number, name, origin, and reference(s) for isolates subjected to GeneChip analysis. The final five columns list isolates' genetic characterization according to spoligotype (where black or white squares indicate the presence or absence, respectively, of 43 specific spacer sequences in the direct repeat region), allele combination at *katG* codon 463 and *gyrA* codon 95 (principal genetic group 1 has *katG*463 CTG and *gyrA*95 ACC; principal genetic group 2 has *katG*463 CGG and *gyrA*95 ACC), base pair composition of *pks15/1* (which either has deletions of 6 bp or 7 bp or remained intact), and presence (+) or absence (−) of TbD1.
[b] SN68 represents an *M. africanum*-like isolate presenting an *M. bovis*-like phenotype (22).

morphisms were also investigated across a panel of phenotypically heterogenous *M. africanum*-like isolates from Guinea-Bissau to assess their distribution and to examine for an association with phenotypic classification.

MATERIALS AND METHODS

Bacterial isolates. Twelve isolates of *M. africanum* of diverse geographic origin, representative of both subtypes, were selected for whole-genome analysis. All isolated from a human host, their laboratory classification and country of origin are provided in Table 1. To genetically characterize these isolates, spoligotype (23) and principal genetic grouping (40) are listed as derived in the supporting citations (7, 21, 22, 26, 33; E. Desmond, unpublished data; S. Niemann, unpublished data). For this study, the presence or absence of *M. tuberculosis*-specific deletion of TbD1 (6) and the characteristic microdeletions of *pks15/1* (28), both regions undetectable via GeneChip analysis, were tested for via PCR across all isolates and confirmed by sequencing.

Thirty-six phenotypically variable isolates from Guinea-Bissau were tested for demonstrated genomic deletions; supporting citations describe their origin and characterization (22, 25). Briefly, by using four biochemical tests (resistance to thiophene-2-carboxylic acid hydrazide, niacin production, nitrate reductase test, and pyrazinamidase test), the isolates were assigned to five different phenotypic biovars varying between classical *M. bovis* (biovar 1) and *M. tuberculosis* (biovar 5).

GeneChip analysis. The *M. tuberculosis* GeneChip represents a high-density oligonucleotide array composed of more than 100,000 sequence patterns in the *M. tuberculosis* genome (37), adopted here to interrogate genomes of *M. africanum*. Extraction of DNA was performed according to standard protocols based on lysozyme and proteinase K (43). Eight micrograms of DNA was fragmented, labeled, and hybridized to the GeneChip as previously described (37). Affymetrix Microarray Suite and the DELSCAN algorithm (AbaSci, LLC, San Pablo, Calif. [http://www.abasci.com]) were used to analyze the data and ultimately suggest candidate deleted regions (30, 32, 42).

PCR amplification and sequencing across deletions. Candidate deleted regions were studied with primers designed to amplify regions in *M. tuberculosis* H37Rv flanking the putative deletion. Reference strains of *M. tuberculosis* and/or *M. bovis* BCG were included as DNA controls for each experiment. To reveal variable genomic sequence, PCR amplicons were run on a 2% agarose gel, where amplicons different from the expected base-pair size of H37Rv could be visualized. All products obtained by amplification across a deleted region were subsequently sequenced by dideoxy terminal sequencing at Stanford University or the McGill University and Genome Quebec Innovation Center. Sequence results were compared by BLAST analysis to *M. tuberculosis* H37Rv sequence by using Tuberculist (http://genolist.pasteur.fr/TubercuList/) and to *M. bovis* AF2122/97 sequence by using Bovilist (http://genolist.pasteur.fr/BoviList/) to confirm whether the amplified *M. africanum* DNA represented a specific deletion event. Annotation of open reading frames affected by deletion events was determined using Tuberculist or, where relevant, Bovilist. Nomenclature of deleted regions disclosed in this study, hereafter referred to as "new deletions," follows that in previously published reports (24, 42).

Analysis of deletions. Primers designed to amplify new deletions from *M. africanum* are provided at www.molepi.mcgill.ca/Mafricanum/primers. Regions are considered deleted only if confirmed by PCR and sequence analysis (30, 31). For isolates to which GeneChip analysis was applied, regions were called present if denoted as such via analysis of GeneChip results.

The deletion of RD7, RD8, RD9, and RD10 from isolates of Guinea-Bissau is described elsewhere (25), whereas new deletions among them were sequence confirmed to determine whether the exact same genomic event was being identified. The distribution of phages RD3 (or phiRv1) (27) and RD11 (or phiRv2) (4, 18) was also determined across Guinea-Bissau isolates. Deletion events were subsequently assigned to the previously characterized bacterial isolates.

RESULTS

As expected, the majority of Affymetrix GeneChip probes hybridized as intact for *M. africanum* genomic DNA, whereas weak or absent hybridization signal was produced for regions of *M. tuberculosis* already described as being deleted from certain *M. africanum* isolates (RD7, RD8, RD9, and RD10). Previously described phages RD3 and RD11, known to have been independently deleted throughout the complex and thus uninformative as phylogenetic markers (6, 20, 31), were also observed as variably missing. Apart from these six regions of deletion, genomic interrogation of *M. africanum* did not reveal any deletions with junctions matching those previously described for other *M. tuberculosis* complex members (4–6, 18, 24, 28, 30, 32, 37, 42). A total of 14 new deletions were uncovered, totaling 46,635 bp (Table 2). For six of these regions, in their place are found rearrangements of parts of the deleted region itself or repetitive sequence, such as the *M. tuberculosis* complex-specific insertion element IS*6110*, a signature of the recombination process observed in mycobacteria (14, 29). The

TABLE 2. Description of deletions from *M. africanum*[a]

Deleted sequence	Start	End	Length (bp)	Affected ORF(s)	Rearranged sequence
RD720	76527	84934	8,408	Rv0069c–Rv0075	
RD721	82923	84589	1,667	Rv0074–Rv0075	
RD701	149241	150980	1,740	Rv0124	
RD702	216795	218516	1,722	Rv0186	217874–217522
RD722	1041193	1049812	8,620	Rv0933–Rv0939	
RD711	1501713	1503655	1,943	Rv1333–Rv1336	
RD742	1875726	1881660	5,935	Rv1661	1881507–1881653, 1881484–1881415, 1875861–1875824
RD713	2219419	2223186	3,768	Rv1977–Rv1979c	2,319 bp*
RD743	2235804	2239118	3,315	Rv1992c–Rv1996	
RD724	2265112	2266239	1,128	Rv2018–Rv2019	Insertion element IS*6110*
RD728	2629497	2634359	4,863	Rv2350c–Rv2352c	Insertion element IS*6110*
RD715	2784617	2785969	1,353	Rv2479c–Rv2480c	
RD727	2902567	2904318	1,752	Rv2578c–Rv2579	29024360–29024396, 2910537–2910595
RD735	3781988	3782408	421	Rv3370c	

[a] The start and end points, length, and affected open reading frames (ORFs) are listed for each region deleted (RD) from *M. africanum* relative to *M. tuberculosis* H37Rv. The final column lists any DNA sequence rearranged at the RD junction. The asterisk * indicates that 2,319 bp of sequence from *M. tuberculosis* strain 210 (http://www.tigr.org/), itself absent from the H37Rv referent, is inverted at the junction of RD713. Deletions are ordered in terms of their locations within the H37Rv genome.

distribution of these deletions across isolates is presented in Table 3, providing three genetic groupings of *M. africanum*.

Targeted genetic analysis confirmed that all isolates classified as *M. africanum* subtype II have a deletion of TbD1 and either have *pks15/1* intact or have its 7-bp microdeletion (Table 1). In agreement with previous reports (6, 39), genomic analysis has revealed that all isolates classified as *M. africanum* subtype II have preserved the regions RD7, RD8, RD9, and RD10 intact. Additionally, new *M. africanum* subtype I deletions (see below) appeared intact in these isolates as determined via GeneChip analysis (Table 3). Conversely, each subtype II isolate has uniquely deleted at least two new regions previously not described (Table 2). The only subtype II isolate to have preserved *pks15/1* intact, K100/RIVM13876, has deletions of RD720 and RD728, affecting 8,408 bp (Rv0069c to Rv0075) and 4,863 bp (Rv2350c to Rv2352c), respectively. Of note, RD720 truncates *sdaA* (Rv0069c), predicted to encode L-serine deaminase, whose disruption may impair in vivo growth (8). Isolates from Uganda, D2169_99, D2285_99, and D2241_99, all have deletions of RD724 and RD727, suggesting regional dominance of a clonal genotype. RD724 affects 1,128 bp (Rv2018 to Rv2019), and RD727 truncates *dhaA* (Rv2579), a putative haloalkane dehalogenase, in its 1,752-bp deletion of Rv2578c to Rv2579. Uniquely, D2169_99 has a deletion of RD721, D2285_99 has a deletion of RD722, and D2241_99 has a deletion of RD735, affecting 1,667 bp (Rv0074 to Rv0075), 8,620 bp (Rv0933 to Rv0939), and 421 bp (Rv3370c), respectively. Together, the genomic content of *M. africanum* subtype II isolates suggests a scattered distribution among clades of "modern" *M. tuberculosis* (20).

Two *M. africanum* isolates, D10409_02 and CPHL_A, have retained TbD1 and *pks15/1* intact (Table 1). These same isolates have a deletion of RD9 but not RD7, RD8, and RD10 (Table 3), thereby excluding them from being classified as *M. tuberculosis* according to previous genomic work (6, 39). Furthermore, the deletions of RD711, RD713, and RD743 are common to only these two isolates, affecting 1,943 bp (Rv1333 to Rv1336), 3,768 bp (Rv1977 to Rv1979c), and 3,315 bp (Rv1992c to Rv1996), respectively. The concordant genomic profile shared among these isolates corroborates their exclusive evolutionary position within the *M. tuberculosis* complex (6, 31). Interestingly, RD713 deletes genes previously implicated in the deletion of RD2 from strains of BCG (Rv1978 to Rv1988) acquired from the Pasteur Institute beyond 1931 (4) and *M. pinnipedii* (Rv1978 to Rv1979c) (28). Finally, isolate CPHL_A has a further deletion of RD715, a

TABLE 3. Large sequence polymorphisms among isolates of *M. africanum* subjected to GeneChip analysis[a]

Sample	Presence of region																			
	RD9	RD7	RD8	RD10	RD724	RD727	RD720	RD728	RD721	RD722	RD735	RD711	RD713	RD743	RD715	RD701	RD702	RD742	RD3	RD11
K100/RIVM13876	+	+	+	+	+	+	−	−	+	+	+	+	+	+	+	+	+	+	−	+
D2169_99	+	+	+	+	−	−	+	+	−	+	+	+	+	+	+	+	+	+	−	+
D2285_99	+	+	+	+	−	−	+	+	+	−	+	+	+	+	+	+	+	+	+	+
D2241_99	+	+	+	+	−	−	+	+	+	+	−	+	+	+	+	+	+	+	+	+
D10409_02	−	+	+	+	+	+	+	+	+	+	+	−	−	−	−*	+	+	−	+	+
CPHL_A	−	+	+	+	+	+	+	+	+	+	+	−	−	−	−	+	+	+	+	+
D10415_02	−	−	−	−	+	+	+	+	+	+	+	+	+	+	+	−	−	+	−	−
K6/RIVM17316	−	−	−	−	+	+	+	+	+	+	+	+	+	+	+	−	−	+	−	−
K85/RIVM17902	−	−	−	−	+	+	+	+	+	+	+	+	+	+	+	−	−	+	−	−
SN68	−	−	−	−	+	+	+	+	+	+	+	+	+	+	+	−	−	−	−	−
ATCC 25420	−	−	−	−	+	+	+	+	+	+	+	+	+	+	+	−	−	−	−	−
ATCC 35711	−	−	−	−	+	+	+	+	+	+	+	+	+	+	+	−	−	−	−	−

[a] Distribution of regions (RD) present (+) or absent (−) in *M. africanum* isolates subject to GeneChip interrogation. Isolates are grouped according to the presence or absence of RD9 and then RD7, RD8, and RD10. The asterisk * indicates that although RD715 is putatively deleted from D10409_02, its sequence was not confirmed.

1,353-bp deletion affecting genes encoding transposase sequence IS*6110* (Rv2479c to Rv2480c). GeneChip analysis also suggests the deletion of RD715 from D10409_02, but this could not be confirmed by PCR because DNA was depleted for this isolate.

The final grouping of *M. africanum* was composed of six isolates in which TbD1 is preserved and 6 bp of *pks15/1* have been deleted (Table 1). The four regions RD7, RD8, RD9, and RD10 are consistently missing, making this group of *M. africanum* the furthest removed phylogenetically from *M. tuberculosis* (6, 31). GeneChip analysis revealed two new deletions, RD701 and RD702, common to only these isolates of *M. africanum*. The *M. africanum*-like isolate from Guinea Bissau presented this same genomic profile, suggesting its inclusion in this group of *M. africanum* by genomic criteria. RD701 deletes 1,740 bp, truncating nearly all of Rv0124 (*PE_PGRS2*). RD702 deletes 1,722 bp, affecting Rv0186 (*bglS*), but has inverted *bglS* sequence (217874 to 217522 of the H37Rv genome) in its place. *bglS* is annotated as a probable beta-glucosidase whose orthologue in *Acetobacter xylinum* belongs to a family of glycosyl hydrolases characterized as having a role in cellulose synthesis (41). Finally, three isolates of this group of *M. africanum* have also deleted and rearranged part of the *pks* region involved in lipid metabolism. RD742 is deleted from ATCC 25420, ATCC 35711, and SN68. RD742 deletes 5,935 bp, affecting the 6,381-bp Rv1661 (*pks7*) annotated as a probable polyketide synthase. The presence of *pks7* sequence (1881507 to 1881653, followed by inverted sequence 1881484 to 1881415 and then 1875861 to 1875824) within the deletion junction further distinguishes these isolates from *M. tuberculosis* H37Rv. *pks7*, involved in the synthesis of phthiocerol derivatives (3, 38), has putative relevance for cell envelope organization and virulence (36).

Because SN68, the *M. africanum*-like isolate from Guinea-Bissau, presented the *M. bovis*-like phenotype but revealed a genomic profile identical to that of this third group of *M. africanum*, we tested for these deletions across a phenotypically diverse panel of 36 *M. africanum*-like isolates from Guinea-Bissau. Unexpectedly, these analyses (available at www.molepi.mcgill.ca/Mafricanum/GB) revealed a common deletion profile for all isolates of this family, irrespective of phenotypic heterogeneity. Without exception, isolates are not missing TbD1 but have deletions of 6 bp of *pks15/1*, RD7, RD8, RD9, RD10, RD701, and RD702. Only RD742 was variable, deleted from 28 of 36 isolates, suggesting that this deletion occurred subsequent to RD701 and RD702. The distribution of RD742 could not be correlated to any examined phenotypic attribute.

DISCUSSION

Genome sequences available for *M. tuberculosis* (9, 15), *M. microti* (5), and *M. bovis* (17) have all confirmed large sequence polymorphisms deleted from *M. tuberculosis* as an obvious medium of genomic variability among species of the *M. tuberculosis* complex (6, 31). *M. tuberculosis* complex organisms are identified to the species level based on morphological, biochemical, and other phenotypic characteristics, but variants have distinguished themselves from conventional species via characteristic phenotypic or genotypic attributes, challenging taxonomic classification within the complex (30). Recognizing the imperfections associated with a classification system based solely upon phenotype, for example, convergence or indeterminate assignments, investigators have more recently explored the utility of genomic deletions in *M. tuberculosis* complex taxonomic derivations (21, 34).

In agreement with other studies (6, 21, 34, 39), genomic analysis has not been able to differentiate *M. africanum* subtype II from modern *M. tuberculosis*, suggesting that *M. africanum* subtype II likely represents phenotypically atypical *M. tuberculosis* strains. In contrast, genomic deletions effectively distinguished *M. tuberculosis* from isolates classified as *M. africanum* subtype I and further presented two groups. One genomotype of *M. africanum* lacks the regions RD9, RD711, RD713, RD743, and likely RD715, while the second genomotype is characterized by the absence of RD7, RD8, RD9, RD10, RD701, RD702, and sometimes RD742. The specificity of these new deletions for *M. africanum* is supported by the fact that these regions are intact as determined by sequence analysis of *M. bovis* 2122 (17) and by a GeneChip-based study of isolates of *M. tuberculosis* (24, 42), the dassie bacillus (30), and the entire BCG family (32). The unique genomic profiles of these *M. africanum* groups point to distinct evolutionary branches within the *M. tuberculosis* complex, both derived from an ancestor in which only RD9 was deleted.

M. africanum characteristically exhibits the range of biochemical traits classically employed to differentiate *M. tuberculosis* from *M. bovis* (13, 19), themselves species located towards the *M. tuberculosis* complex phylogenetic poles (6, 31). As a result, the species name of *M. africanum* is unspecific and is applied towards isolates differing in their phenotypic presentation and genomic content. Although *M. africanum*-specific deletions each have putative biological consequence, predicted functions by gene alignment do not obviously address these documented variable biochemical characteristics. Moreover, a group of isolates spanning the biochemical spectrum was found to share the exact same deletion profile, suggesting that phenotypic heterogeneity among *M. africanum* is due to genetic polymorphisms not detectable by GeneChip analysis, such as small deletions, single-nucleotide polymorphisms, or rearrangements. The observation made here that organisms with an *M. africanum* deletion profile can present the complete range of characteristics from the *M. tuberculosis*-like biovar 5 to the *M. bovis*-like biovar 1 (22) has two important implications. First, these findings imply that during the evolution of the *M. tuberculosis* complex, the biochemical profile associated with *M. bovis* has arisen independently in the *M. africanum* branch. This phenotypic convergence suggests that selective pressures encountered along these distinct lineages of the *M. tuberculosis* complex have favored the emergence of this same biochemical profile. Alternatively, there may be no selective pressure for the maintenance of particular characteristics, and their emergence is of no direct evolutionary consequence. Second, from a more practical vantage, these findings challenge the use of these biochemical characteristics for classifying organisms in diagnostic laboratories. In the absence of genetic typing, the true prevalence of *M. africanum* is difficult to ascertain, especially where *M. tuberculosis*-like or *M. bovis*-like isolates are encountered.

Species of the *M. tuberculosis* complex have been traditionally classified according to the host from which they are pri-

marily isolated, including *M. tuberculosis* from humans, *M. microti* from field voles, and *M. bovis* from cattle. Although the host range of *M. africanum*, like that of *M. canettii* and *M. tuberculosis*, is thought to be human, it is generally isolated at a much lower frequency than *M. tuberculosis*. This may represent a geographically restricted human pathogen or an artifact of laboratory classification. To what extent the genomic definition for *M. africanum* influences its virulence and transmissibility among human and/or nonhuman hosts is the subject of ongoing investigation.

ACKNOWLEDGMENTS

We thank Peter Small for input and assistance. Thanks to David Roquis, Pierre LePage, Daniel Vincent, Yannick Fortin, and Rob Sladek at the McGill University and Genome Quebec Innovation Center and Margarita Zalamea at Stanford University for helping with DNA sequencing and GeneChip experimentation. We also acknowledge Tuija Koivula and Gunilla Kallenius at the Swedish Institute for Infectious Disease Control for their donation of Guinea-Bissau isolates.

This work was supported by the Canadian Institutes of Health Research (CIHR) grant number MOP 36054. M.B. is a New Investigator of the CIHR. S.G. is supported by the Swiss National Science Foundation and by the Novartis Foundation. A.O., S.G., and M.K.-M. are supported by Wellcome Trust grant 176W009.

REFERENCES

1. **Aranaz, A., D. Cousins, A. Mateos, and L. Dominguez.** 2003. Elevation of Mycobacterium tuberculosis subsp. caprae Aranaz et al. 1999 to species rank as Mycobacterium caprae comb. nov., sp. nov. Int. J. Syst. Evol. Microbiol. **53:**1785–1789.
2. **Aranaz, A., E. Liebana, E. Gomez-Mampaso, J. C. Galan, D. Cousins, A. Ortega, J. Blazquez, F. Baquero, A. Mateos, G. Suarez, and L. Dominguez.** 1999. Mycobacterium tuberculosis subsp. caprae subsp. nov.: a taxonomic study of a new member of the Mycobacterium tuberculosis complex isolated from goats in Spain. Int. J. Syst. Bacteriol. **49:**1263–1273.
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. **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.
5. **Brodin, P., K. Eiglmeier, M. Marmiesse, A. Billault, T. Garnier, S. Niemann, S. T. Cole, and R. Brosch.** 2002. Bacterial artificial chromosome-based comparative genomic analysis identifies *Mycobacterium microti* as a natural ESAT-6 deletion mutant. Infect. Immun. **70:**5568–5578.
6. **Brosch, R., S. V. Gordon, M. Marmiesse, P. Brodin, C. Buchrieser, K. Eiglmeier, T. Garnier, C. Gutierrez, G. Hewinson, K. Kremer, L. M. Parsons, A. S. Pym, S. Samper, D. Van Soolingen, and S. T. Cole.** 2002. A new evolutionary scenario for the Mycobacterium tuberculosis complex. Proc. Natl. Acad. Sci. USA **99:**3684–3689.
7. **Castets, M., H. Boisvert, F. Grumbach, M. Brunel, and N. Rist.** 1968. Tuberculosis bacilli of the African type: preliminary note. Rev. Tuberc. Pneumol. (Paris) **32:**179–184. (In French.)
8. **Chen, J. M., D. C. Alexander, M. A. Behr, and J. Liu.** 2003. *Mycobacterium bovis* BCG vaccines exhibit defects in alanine and serine catabolism. Infect. Immun. **71:**708–716.
9. **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.
10. **Cousins, D. V., R. Bastida, A. Cataldi, V. Quse, S. Redrobe, S. Dow, P. Duignan, A. Murray, C. Dupont, N. Ahmed, D. M. Collins, W. R. Butler, D. Dawson, D. Rodriguez, J. Loureiro, M. I. Romano, A. Alito, M. Zumarraga, and A. Bernardelli.** 2003. Tuberculosis in seals caused by a novel member of the Mycobacterium tuberculosis complex: Mycobacterium pinnipedii sp. nov. Int. J. Syst. Evol. Microbiol. **53:**1305–1314.
11. **Cousins, D. V., R. L. Peet, W. T. Gaynor, S. N. Williams, and B. L. Gow.** 1994. Tuberculosis in imported hyrax (Procavia capensis) caused by an unusual variant belonging to the Mycobacterium tuberculosis complex. Vet. Microbiol. **42:**135–145.
12. **Cousins, D. V., S. N. Williams, R. Reuter, D. Forshaw, B. Chadwick, D. Coughran, P. Collins, and N. Gales.** 1993. Tuberculosis in wild seals and characterisation of the seal bacillus. Aust. Vet. J. **70:**92–97.
13. **David, H. L., M.-T. Jahan, A. Jumin, J. Grandry, and E. H. Lehman.** 1978. Numerical taxonomy analysis of *Mycobacterium africanum*. Int. J. Syst. Bacteriol. **28:**467–472.
14. **Fang, Z., C. Doig, D. T. Kenna, N. Smittipat, P. Palittapongarnpim, B. Watt, and K. J. Forbes.** 1999. IS*6110*-mediated deletions of wild-type chromosomes of *Mycobacterium tuberculosis*. J. Bacteriol. **181:**1014–1020.
15. **Fleischmann, R. D., D. Alland, J. A. Eisen, L. Carpenter, O. White, J. Peterson, R. DeBoy, R. Dodson, M. Gwinn, D. Haft, E. Hickey, J. F. Kolonay, W. C. Nelson, L. A. Umayam, M. Ermolaeva, S. L. Salzberg, A. Delcher, T. Utterback, J. Weidman, H. Khouri, J. Gill, A. Mikula, W. Bishai, J. W. Jacobs, Jr., J. C. Venter, and C. M. Fraser.** 2002. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. J. Bacteriol. **184:**5479–5490.
16. **Fraser, C. M., J. Eisen, R. D. Fleischmann, K. A. Ketchum, and S. Peterson.** 2000. Comparative genomics and understanding of microbial biology. Emerg. Infect. Dis. **6:**505–512.
17. **Garnier, T., K. Eiglmeier, J. C. Camus, N. Medina, H. Mansoor, M. Pryor, S. Duthoy, S. Grondin, C. Lacroix, C. Monsempe, S. Simon, B. Harris, R. Atkin, J. Doggett, R. Mayes, L. Keating, P. R. Wheeler, J. Parkhill, B. G. Barrell, S. T. Cole, S. V. Gordon, and R. G. Hewinson.** 2003. The complete genome sequence of Mycobacterium bovis. Proc. Natl. Acad. Sci. USA **100:**7877–7882.
18. **Gordon, S. V., R. Brosch, A. Billault, T. Garnier, K. Eiglmeier, and S. T. Cole.** 1999. Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays. Mol. Microbiol. **32:**643–655.
19. **Haas, W. H., G. Bretzel, B. Amthor, K. Schilke, G. Krommes, S. Rusch-Gerdes, V. Sticht-Groh, and H. J. Bremer.** 1997. Comparison of DNA fingerprint patterns of isolates of *Mycobacterium africanum* from east and west Africa. J. Clin. Microbiol. **35:**663–666.
20. **Hirsh, A. E., A. G. Tsolaki, K. DeRiemer, M. W. Feldman, and P. M. Small.** 2004. Stable association between strains of Mycobacterium tuberculosis and their human host populations. Proc. Natl. Acad. Sci. USA **101:**4871–4876.
21. **Huard, R. C., L. C. Oliveira Lazzarini, W. R. Butler, D. Van Soolingen, and J. L. Ho.** 2003. PCR-based method to differentiate the subspecies of the *Mycobacterium tuberculosis* complex on the basis of genomic deletions. J. Clin. Microbiol. **41:**1637–1650.
22. **Kallenius, G., T. Koivula, S. Ghebremichael, S. E. Hoffner, R. Norberg, E. Svensson, F. Dias, B. I. Marklund, and S. B. Svenson.** 1999. Evolution and clonal traits of *Mycobacterium tuberculosis* complex in Guinea-Bissau. J. Clin. Microbiol. **37:**3872–3878.
23. **Kamerbeek, J., L. Schouls, A. Kolk, M. van Agterveld, D. Van Soolingen, S. Kuijper, A. Bunschoten, H. Molhuizen, R. Shaw, M. Goyal, and J. Van Embden.** 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J. Clin. Microbiol. **35:**907–914.
24. **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.
25. **Koivula, T., M. Ekman, T. Leitner, S. Lofdahl, S. Ghebremicahel, S. Mostowy, M. A. Behr, S. B. Svenson, and G. Kallenius.** 2004. Genetic characterization of the Guinea-Bissau family of Mycobacterium tuberculosis complex strains. Microbes Infect. **6:**272–278.
26. **Kremer, K., D. Van Soolingen, R. Frothingham, W. H. Haas, P. W. Hermans, C. Martin, P. Palittapongarnpim, B. B. Plikaytis, L. W. Riley, M. A. Yakrus, J. M. Musser, and J. D. van Embden.** 1999. Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. J. Clin. Microbiol. **37:**2607–2618.
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. **Marmiesse, M., P. Brodin, C. Buchrieser, C. Gutierrez, N. Simoes, V. Vincent, P. Glaser, S. T. Cole, and R. Brosch.** 2004. Macro-array and bioinformatic analyses reveal mycobacterial 'core' genes, variation in the ESAT-6 gene family and new phylogenetic markers for the Mycobacterium tuberculosis complex. Microbiology **150:**483–496.
29. **McFadden, J.** 1996. Recombination in mycobacteria. Mol. Microbiol. **21:**205–211.
30. **Mostowy, S., D. Cousins, and M. A. Behr.** 2004. Genomic interrogation of the dassie bacillus reveals it as a unique RD1 mutant within the *Mycobacterium tuberculosis* complex. J. Bacteriol. **186:**104–109.
31. **Mostowy, S., D. Cousins, J. Brinkman, A. Aranaz, and M. A. Behr.** 2002. Genomic deletions suggest a phylogeny for the Mycobacterium tuberculosis complex. J. Infect. Dis. **186:**74–80.
32. **Mostowy, S., A. G. Tsolaki, P. M. Small, and M. A. Behr.** 2003. The in vitro evolution of BCG vaccines. Vaccine **21:**4270–4274.
33. **Niemann, S., S. Rusch-Gerdes, M. L. Joloba, C. C. Whalen, D. Guwatudde, J. J. Ellner, K. Eisenach, N. Fumokong, J. L. Johnson, T. Aisu, R. D. Mugerwa, A. Okwera, and S. K. Schwander.** 2002. *Mycobacterium africanum*

subtype II is associated with two distinct genotypes and is a major cause of human tuberculosis in Kampala, Uganda. J. Clin. Microbiol. **40:**3398–3405.
34. **Parsons, L. M., R. Brosch, S. T. Cole, A. Somoskovi, A. Loder, G. Bretzel, D. Van Soolingen, Y. M. Hale, and M. Salfinger.** 2002. Rapid and simple approach for identification of *Mycobacterium tuberculosis* complex isolates by PCR-based genomic deletion analysis. J. Clin. Microbiol. **40:**2339–2345.
35. **Pfyffer, G. E., R. Auckenthaler, J. D. van Embden, and D. Van Soolingen.** 1998. Mycobacterium canettii, the smooth variant of M. tuberculosis, isolated from a Swiss patient exposed in Africa. Emerg. Infect. Dis. **4:**631–634.
36. **Rousseau, C., T. D. Sirakova, V. S. Dubey, Y. Bordat, P. E. Kolattukudy, B. Gicquel, and M. Jackson.** 2003. Virulence attenuation of two Mas-like polyketide synthase mutants of Mycobacterium tuberculosis. Microbiology **149:**1837–1847.
37. **Salamon, H., M. Kato-Maeda, P. M. Small, J. Drenkow, and T. R. Gingeras.** 2000. Detection of deleted genomic DNA using a semiautomated computational analysis of GeneChip data. Genome Res. **10:**2044–2054.
38. **Sirakova, T. D., V. S. Dubey, M. H. Cynamon, and P. E. Kolattukudy.** 2003. Attenuation of *Mycobacterium tuberculosis* by disruption of a *mas*-like gene or a chalcone synthase-like gene, which causes deficiency in dimycocerosyl phthiocerol synthesis. J. Bacteriol. **185:**2999–3008.
39. **Sola, C., N. Rastogi, M. C. Gutierrez, V. Vincent, R. Brosch, and L. Parsons.** 2003. Is *Mycobacterium africanum* subtype II (Uganda I and Uganda II) a genetically well-defined subspecies of the *Mycobacterium tuberculosis* complex? J. Clin. Microbiol. **41:**1345–1346.
40. **Sreevatsan, S., X. Pan, K. E. Stockbauer, N. D. Connell, B. N. Kreiswirth, T. S. Whittam, and J. M. Musser.** 1997. Restricted structural gene polymorphism in the Mycobacterium tuberculosis complex indicates evolutionarily recent global dissemination. Proc. Natl. Acad. Sci. USA **94:**9869–9874.
41. **Tajima, K., K. Nakajima, H. Yamashita, T. Shiba, M. Munekata, and M. Takai.** 2001. Cloning and sequencing of the beta-glucosidase gene from Acetobacter xylinum ATCC 23769. DNA Res. **8:**263–269.
42. **Tsolaki, A. G., A. E. Hirsh, K. DeRiemer, J. A. Enciso, M. Z. Wong, M. Hannan, Y. O. Goguet de la Salmoniere, K. Aman, M. Kato-Maeda, and P. M. Small.** 2004. Functional and evolutionary genomics of Mycobacterium tuberculosis: insights from genomic deletions in 100 strains. Proc. Natl. Acad. Sci. USA **101:**4865–4870.
43. **Van Soolingen, D., P. W. Hermans, P. E. de Haas, D. R. Soll, and J. D. van Embden.** 1991. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. J. Clin. Microbiol. **29:**2578–2586.
44. **Van Soolingen, D., T. Hoogenboezem, P. E. de Haas, P. W. Hermans, M. A. Koedam, K. S. Teppema, P. J. Brennan, G. S. Besra, F. Portaels, J. Top, L. M. Schouls, and J. D. van Embden.** 1997. A novel pathogenic taxon of the Mycobacterium tuberculosis complex, Canetti: characterization of an exceptional isolate from Africa. Int. J. Syst. Bacteriol. **47:**1236–1245.