Region of Difference 1 in Nontuberculous Mycobacterium Species Adds a Phylogenetic and Taxonomical Character

Jakko van Ingen,1,2* Rina de Zwaan,1 Richard Dekhuijzen,2 Martin Boeree,2 and Dick van Soolingen1

National Mycobacteria Reference Laboratory, Laboratories for Infectious Diseases and Perinatal Screening, National Institute for Public Health and the Environment, Bilthoven, The Netherlands,1 and Department of Pulmonary Diseases, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands2

Received 26 May 2009/Accepted 14 July 2009

The esat-6 and cfp-10 genes are essential for virulence in Mycobacterium tuberculosis. Among nontuberculous mycobacteria, we found these genes only in M. kansasii, M. szulgai, M. marinum, and M. riyadhense, with unique sequences. This adds a phylogenetic and taxonomical characteristic and may represent a virulence factor for nontuberculous mycobacteria.

The 6-kDa early secretory antigenic target (ESAT-6) and 10-kDa culture filtrate protein (CFP-10) of Mycobacterium tuberculosis are potent T-cell antigens (2, 12). The genes encoding ESAT-6 and CFP-10 are situated within “region of difference 1” (RD1) of Mycobacterium tuberculosis but are also present in some nontuberculous mycobacteria (NTM) (2, 12). Deletion of RD1 in M. tuberculosis significantly decreases its virulence in animal models (8), suggesting that genes residing in RD1 are involved in pathogenesis. In M. tuberculosis, the RD1 proteins effect translocation to the cytosol, a mechanism for survival within macrophages (15). Therefore, RD1 in NTM may also play a crucial role in virulence.

To improve our understanding of the pathogenesis of NTM disease and the possible role of RD1 in virulence, we screened a wide diversity of NTM for the presence of RD1.

From our laboratory database, we retrieved isolates of all five Mycobacterium kansasii subtypes based on 16S-23S internal transcribed spacer sequencing (n = 15), M. szulgai (n = 4), M. marinum (n = 4), M. avium (n = 2), M. conspicuum (n = 4), M. genavense (n = 1), M. boehmii (n = 3), M. interjectum (n = 2), M. flavescens (n = 5), M. xenopi (n = 2), M. malmoense (n = 2), M. riyadhense (n = 1), and M. tuberculosis H37Rv. The selection of these strains was based on their phylogenetic relationship with the M. tuberculosis complex in the multigene taxonomical model published by Devulder et al. (3).

To establish the presence of an RD1-like element and sequences of the esat-6 and cfp-10 genes, we used the Esa-12 (CATGACAGAGCAGCAGTG) and Esa-303 (5′-TCAGAAGCCCATTTGCGAGGAC-3′) primers for M. avium and the opBR78 (5′-GTAGCCGGGGATGGCAAGATGGAACCCGATGC C-3′) and opBR103 (5′-TCAGAAGCCCATTTGCGAGGAC AGC-3′) primers for M. marinum (1). The M. smegmatis esat-6 and cfp-10 gene sequences were extracted from the whole-genome sequence in the GenBank database (accession number CP000480). Using these primers, we were able to demonstrate an RD1 for M. tuberculosis H37Rv and for all M. kansasii subtypes, M. szulgai, M. marinum, and M. riyadhense. The PCR was repeatedly negative for isolates of the remaining species, M. avium, M. conspicuum, M. genavense, M. boehmii, M. interjectum, M. flavescens, M. xenopi, and M. malmoense. For these species, we performed Southern blotting and hybridized DNA membranes using the purified M. tuberculosis H37Rv and M. kansasii type I esat-6 amplicons as probes (1). None of the PCR-negative species hybridized with either probe (data not shown).

The presence of RD1, characterized by an esat-6-like gene and an cfp-10-like gene, is a phylogenetic characteristic among the NTM. It is mainly found only among slowly growing NTM species that are phylogenetically related to the M. tuberculosis complex based on the multigene taxonomical model by Devulder et al. (3) and in the more distantly related rapid grower M. smegmatis. Possibly, the presence of RD1 reflects phylogenetic relationships to the M. tuberculosis complex.

Previous authors have recorded the presence of RD1 in M. flavescens (1). We were unable to demonstrate it in four reference strains (ATCC 23008, ATCC 23033, ATCC 23035, and ATCC 23039) and a clinical isolate. Thus, with its presence of an RD1 region, M. smegmatis still stands out among the rapid growing NTM.

Gey van Pittius et al. have demonstrated the presence of the ESX-5 locus, which they assume is a product of duplication of RD1 and its secretion system (ESX-1), in most slow-growing NTM species (5). Our results seem to suggest that in many slow growers, after this duplication, the original ESX-1 either was lost or has undergone extensive mutation, barring hybridization.

All five subtypes of M. kansasii had distinct esat-6 and cfp-10 sequences; type three was subdivided into two separate lineages based on both esat-6 and cfp-10 sequences. Among M. marinum and M. szulgai strains, no difference was noted, and the sequences obtained for M. riyadhense were unique. We aligned the esat-6 and cfp-10 gene sequences separately and concatenated, using the Clustal X software program (14). The resulting topology and tree, inferred by neighbor joining and visualized using the MEGA 4.0 software package (13), were evaluated by bootstrap analyses based on 1,000 resamplings. The resulting trees are shown in Fig. 1A (esat-6), B (cfp-10), and C (esat-6 and cfp-10 concatenated). From these trees, it is

---

* Corresponding author. Mailing address: RIVM, LIS, National Mycobacteria Reference Laboratory, Antonie van Leeuwenhoeklaan 9, P.O. Box 1, 3720 BA Bilthoven, The Netherlands. Phone: 31-30-2742886. Fax: 31-30-2744418. E-mail: jakko.van.ingen@rivm.nl.

† Published ahead of print on 17 July 2009.
obvious that the slow-growing RD1-positive NTM have *esat-6* and *cfp-10* sequences much more closely related to those of *M. tuberculosis* than to those of *M. smegmatis* (Fig. 1A to C).

Thus, the presence of RD1 in these slow growers marks a genetically closely related *Mycobacterium* grouping, and sequencing of RD1 is a tool for (sub)species identification. Therefore, we propose that future introduction of new species phylogenetically related to the RD1-positive grouping should include an investigation of RD1 presence and gene sequences. Deletion of RD1 lowers the virulence of *M. tuberculosis* complex bacteria (8, 9, 11). Although the presence of RD1 may thus be important for virulence, we were not able to detect this genomic region in well-known causative agents of disease in humans, including *M. avium* and *M. malmoense* (7, 12). In *M. smegmatis*, ESAT-6 and CFP-10 secretion has a role in conjugation, rather than translocation to the cytosol, by which *M. tuberculosis* survives within macrophages (4, 15). The remaining NTM species that harbor RD1 are phylogenetically more closely related to *M. tuberculosis* than to *M. smegmatis* (3), which is also expressed in their *esat-6* and *cfp-10* sequences (Fig. 1). Moreover, *M. kansasii*, *M. szulgai*, and *M. marinum* are considered the most pathogenic among the NTM (1, 6, 16). Therefore, RD1 may play a role in virulence of these NTM.

The RD1 sequences differed between *M. kansasii* type 1, an important causative agent of NTM disease, and the other types, which are less involved or not involved in human disease (1). Clinically relevant and nonrelevant *M. szulgai* isolates, determined using the American Thoracic Society diagnostic criteria (6, 16), shared identical sequences. The RD1 presence and gene sequences and thus the protein structure do not provide a complete explanation of the virulences of the different NTM. Presumably, host factors and pathogen factors other than RD1 presence are also important. In vitro infection experiments are necessary to clarify the role of RD1-containing slow-growing NTM.

Demonstration and characterization of RD1 in NTM have gained significance with the advent of the gamma interferon release assays (IGRAs) for the diagnosis of (latent) tuberculosis. These assays measure gamma interferon production and release by patients' T lymphocytes after incubation with the
TABLE 1. GenBank accession numbers of isolates sequenced in this study

<table>
<thead>
<tr>
<th>Species</th>
<th>esat-6</th>
<th>cfp-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. szulgai</td>
<td>EU826486</td>
<td>FJ014490</td>
</tr>
<tr>
<td>M. marinum</td>
<td>EU826487</td>
<td>FJ014491</td>
</tr>
<tr>
<td>M. riyadhense</td>
<td>EU552926</td>
<td>EU552927</td>
</tr>
<tr>
<td>M. kansasii I</td>
<td>EU888292</td>
<td>FJ014492</td>
</tr>
<tr>
<td>M. kansasii II</td>
<td>EU888293</td>
<td>FJ014493</td>
</tr>
<tr>
<td>M. kansasii III</td>
<td>EU888294</td>
<td>FJ014494</td>
</tr>
<tr>
<td>M. kansasii Hb</td>
<td>EU888295</td>
<td>FJ014495</td>
</tr>
<tr>
<td>M. kansasii IV</td>
<td>EU888296</td>
<td>FJ014496</td>
</tr>
<tr>
<td>M. kansasii V</td>
<td>EU888297</td>
<td>FJ014497</td>
</tr>
<tr>
<td>M. tuberculosis H37Rv</td>
<td>FJ014499</td>
<td>FJ014498</td>
</tr>
</tbody>
</table>

ESAT-6 and CFP-10 antigens of *M. tuberculosis* (10). The presence of similar antigens in NTM and thus recognition of these antigens by patients infected by these NTM theoretically lowers the specificity of the IGRAs in diagnosing latent tuberculosis.

In conclusion, an RD1 element, similar to that of *M. tuberculosis*, is present in *M. kansasii*, *M. szulgai*, *M. marinum*, and *M. riyadhense*. The presence of RD1 in general is a phylogenetic and taxonomical characteristic of NTM, which hints at a phylogenetic relationship with the *M. tuberculosis* complex. The RD1 sequence analysis enables distinction to the species containing NTM on functioning of the IGRAs should be the subjects of further studies.

**Nucleotide sequence accession numbers.** All obtained sequences have been deposited in the GenBank database; accession numbers are detailed in Table 1.

**REFERENCES**