JOURNAL OF BACTERIOLOGY, May 2008, p. 3408–3410 Vol. 190, No. 9
0021-9193/08/$08.00+0 doi:10.1128/JB.00009-08

IS*6110*, a *Mycobacterium tuberculosis* Complex-Specific Insertion Sequence, Is Also Present in the Genome of *Mycobacterium smegmatis*, Suggestive of Lateral Gene Transfer among Mycobacterial Species▽†

Abbie Coros,[1] Erin DeConno,[1] and Keith M. Derbyshire[1,2]*

Division of Infectious Disease, Wadsworth Center, New York State Department of Health,[1] and Department of Biomedical Sciences, School of Public Health, University at Albany,[2] Albany, New York 12201-2002

Received 3 January 2008/Accepted 26 February 2008

IS*6110* is an insertion element found exclusively within the members of the *Mycobacterium tuberculosis* complex (MTBC), and because of this exclusivity, it has become an important diagnostic tool in the identification of MTBC species. The restriction of IS*6110* to the MTBC is hypothesized to arise from the inability of these bacteria to exchange DNA. We have identified an IS*6110*-related element in a strain of *Mycobacterium smegmatis*. The presence of IS*6110* indicates that lateral gene transfer has occurred among mycobacterial species, suggesting that the mycobacterial gene pool is larger than previously suspected.

Genetic exchange is thought to be a driving force behind the ability of bacterial species to evolve and adjust to environmental challenges. Lateral gene transfer (LGT) in bacteria is mediated by one of three processes, namely, conjugation, transformation, or transduction; examples of these processes have been described for almost all bacterial species (9, 19). By contrast, the *Mycobacterium tuberculosis* complex (MTBC) species, comprising *M. tuberculosis*, *M. africanum*, *M. bovis*, *M. cannetti*, *M. caprae*, *M. microti*, and *M. pinnipedi* (10, 15), are clonal populations evolved from a single progenitor species that has diversified by the acquisition of spontaneous mutations rather than by LGT. Genome comparisons between these seven species show that they have almost identical 16S rRNA sequences and highly similar genome sequences, and there is no strong evidence for genetic exchange (>99% identity [2, 3, 16]). The lack of clearly documented LGT among members of the MTBC is thought to be a consequence of the organisms' solitary lifestyles within their hosts, preventing their contact with other mycobacterial species, or perhaps even other bacteria. Thus, it has become generally accepted within the scientific community that the MTBC species do not undergo genetic exchange (10, 15, 16).

IS*6110* is an insertion element that is found exclusively within the MTBC; the assumption has been that this restriction is a result of the lack of genetic exchange with other mycobacterial species. A benefit of this exclusivity is that IS*6110* has become an important diagnostic tool in the differentiation of MTBC species from other mycobacteria. Moreover, the element's presence in multiple copies, and at differing locations in the genome, has provided an excellent method by which strains can be genotyped; because of these characteristics, IS*6110* has been used extensively for epidemiological studies (12, 18, 20).

Our studies have focused on DNA transfer between strains of *M. smegmatis*. This work has shown that DNA transfer occurs by a process most similar to conjugation: distinct donor and recipient strains exist and transconjugants are detected only after prolonged cell-cell contact (14, 21, 22). The transfer process is chromosomally encoded and can occur only from a donor to a recipient. The donor and recipient strains are independent isolates of *M. smegmatis* with distinct colony morphologies (13). The genetic basis for donor and recipient ability is unknown. In order to identify genes required for DNA transfer, we have used transposon mutagenesis to isolate transfer-defective mutants in both donor and recipient strains (8; A. Coros, B. Callahan, and K. M. Derbyshire, unpublished data). The transposon mutagenesis screen performed in the recipient strain of *M. smegmatis* (MKD8) identified several insertions into regions of DNA unique to the recipient genome (i.e., the regions are absent from the sequenced mc^{2}155 donor genome; The Institute for Genomic Research), suggesting that they encode recipient-specific functions necessary for transfer (Coros et al., unpublished). We have used a PCR-based chromosome-walking technique to sequence outward from these unique regions into sequences that align with the known *M. smegmatis* genome sequence. Analysis of one of these recipient-specific regions identified an insertion sequence that is very closely related to IS*6110* (3e−105, 67% amino acid [aa] identity; Fig. 1). This is a remarkable finding, because IS*6110*-like elements were thought to be exclusive to the MTBC (12, 17). The presence of this element within the *M. smegmatis* genome is most consistent with the occurrence of a horizontal transfer event, from an MTBC species to *M. smegmatis*; it provides strong supporting evidence for a natural genetic exchange occurring between these two species. It also suggests that the potential shared gene pool of the mycobacteria is much larger than previously thought.

The element (we use "Ms" to differentiate the element from its *M. tuberculosis* counterpart) is not completely intact, because it lacks one inverted repeat (IR); however, the entire transposase gene (encoding 267 aa) and the second IR are highly homologous to ISMt*6110* elements found in the *M.*

* Corresponding author. Mailing address: David Axelrod Institute, NYS Department of Health, P.O. Box 22002, Albany, NY 12201-2002. Phone: (518) 473-6079. Fax: (518) 474-3181. E-mail: keith.derbyshire@wadsworth.org.

† Supplemental material for this article may be found at http://jb.asm.org/.

▽ Published ahead of print on 7 March 2008.

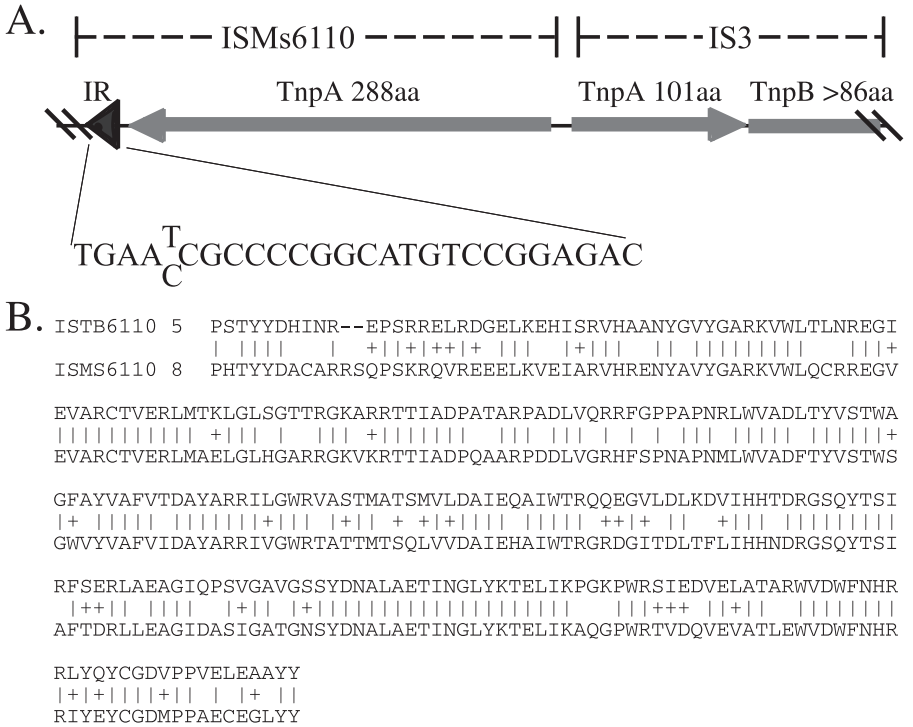

FIG. 1. Schematic representation of the IS*6110* element present in the conjugal recipient strain (MKD8) of *M. smegmatis.* (A) The map shows a 1.44-kb segment of the *M. smegmatis* genome encoding ISMs*6110* and a second, unrelated IS*3* element. The transposases (Tnp) encoded by each open reading frame are indicated above the map, along with their directions of expression and lengths. The IS*3* element is only a partial sequence of 570 bp, but it includes the frameshift sequence that permits expression of a TnpAB fusion protein. ISMs*6110* contains a single *tnpA* gene and a single IR. The sequence of the IR is indicated below the map. The one-nucleotide difference with ISMt*6110* is indicated at position 5, which is a T in ISMs*6110*. (B) The sequences of the transposase proteins for IS*6110* from *M. smegmatis* and *M. tuberculosis* are aligned (67% identity, 79% similarity). The sequences are aligned from aa 5 and aa 8 for *M. tuberculosis* and *M. smegmatis*, respectively.

tuberculosis and *M. bovis* genomes. Immediately adjacent to ISMs*6110* is a second insertion sequence (IS) with weak homologies to IS3-like elements (Fig. 1). It is possible that this IS3-like element inserted into ISMs*6110* and in the process either deleted or displaced the 5′ end of the element and the now-missing IR. Not only is the ISMs*6110* transposase protein highly conserved throughout its length, but also the nucleotide sequence itself contains extensive regions of identity (e.g., 189/248 nucleotide matches, 76% identity) with the *M. tuberculosis* sequences. Most importantly, the transposase-binding site, located immediately 3′ of the transposase gene, is highly conserved (25/26 positions; Fig. 1). This conservation is significant since the IR is a critical *cis*-acting sequence, and its immediate proximity to the gene underscores the functional correlation with the IS*6110* transposase; IS transposases work preferentially in *cis*, and IRs are generally found within a few base pairs of the 3′ end of their cognate transposase gene (4). The one mismatch occurs at the fifth position of the IR; the variations in this region of the IR are not uncommon, because it is a junction between two functional domains of the IR (6, 7, 23).

A Southern analysis was used to determine how many copies of ISMs*6110* were present in the recipient *M. smegmatis* genome. Only one cross-hybridizing band was detected, indicating that only one element was present and that subsequent transposition events had not occurred (see Fig. S1 in the supplemental material). This finding was not entirely unexpected; the element cannot transpose, since it lacks one IR. The chromosomes of other independent isolates of *M. smegmatis* were also probed by Southern analysis for the presence of ISMs*6110* (see Fig. S1 in the supplemental material). ISMs*6110* was not detected in any of these isolates, which included two recipient strains (strains Jucho and Takeuchi) and three donor strains (strains mc^2155, Rabinowitchi, and Nishi) (13). Thus, ISMs*6110* is unique to the recipient strain MKD8; its absence in other *M. smegmatis* isolates is a strong indication that MKD8 acquired the element relatively recently. In addition, we included *M. tuberculosis* genomic DNA as a control in the Southern analysis. No cross-hybridization was detected under the experimental conditions used. It has been shown that an ISMt*6110* probe does not cross-hybridize with ISMs*6110* under standard diagnostic conditions (M. McGarry, personal communication). The lack of cross-hybridization with ISMt*6110* is not surprising given the lack of extended nucleotide identity between the two elements.

IS*6110* has never before been detected outside the MTBC complex. After completion of this work, a bioinformatic survey identified a second IS*6110*-like element in an environmental species of mycobacterium (*Mycobacterium* sp. strain JLS; Mjls_2222, YP_001070499.1) that is being sequenced. This JLS IS*6110* element carries a conserved IR, and its transposase is even more closely related to that of ISMt*6110* (86% aa identity). So, how did the *M. smegmatis* recipient strain acquire a copy of IS*6110*? The most parsimonious explanation is that the segment of DNA was inherited directly from a member of the

MTBC, when the two species co-occupied a single environmental niche, and the sequence subsequently diverged. A second possibility is that the exchange took place through an intermediate species, such as *Mycobacterium* sp. strain JLS. Regardless, the presence of IS*6110* in a non-MTBC bacterium indicates that DNA exchange has occurred between mycobacterial species.

Although plasmid transformation has been described for *M. smegmatis*, exchange of chromosomal DNA by this mechanism has not been demonstrated (1). By contrast, genetic exchange of chromosomal DNA by both conjugation and transduction is known to occur (5, 11). Currently, it is impossible to determine which of the latter processes (if not both) mediated the genetic exchange, but it has not escaped our attention that the MKD8 strain of *M. smegmatis* is a recipient for conjugal DNA transfer. Moreover, many of the genes that we have identified as being required for transfer in *M. smegmatis* are conserved in *M. tuberculosis*. An argument against conjugation is that it requires both parents to have coexisted in a single environmental niche, which might seem unlikely given what is known of the species' normal habitats. By contrast, phage-mediated transfer does not require that the donor and recipient coexist.

Although an IS element has been acquired by the recipient, it is not clear whether the element integrated into the genome via transposition or recombination. As mentioned above, the element lacks one IR, and downstream of the IS are three genes that have only weak homologies with other mycobacterial genes (Coros et al., unpublished), indicating that they too have been acquired from another mycobacterium. It is therefore possible that the entire segment of DNA was inherited by homologous recombination between the upstream IS3 element—of which there are multiple copies in the recipient genome—and genes more distal to the ISMs*6110* element. Regardless of the mechanism of inheritance, the presence of an IS*6110*-like element in *M. smegmatis* implicates the existence of horizontal gene transfer among the mycobacteria and suggests that the transfer event described here occurred in the recent past. Although the present example suggests that the transfer event was from the MTBC to *M. smegmatis*, it is not unreasonable to assume that transfer could occur in the opposite direction, i.e., into or between other members of the MTBC. The only requirement would be that recipient strains of the MTBC exist.

Nucleotide sequence accession number. The ISMs*6110* sequence has been deposited at GenBank (EU366287).

We thank Mike McGarry of the Wadsworth mycobacterial genotyping laboratory for confirming that IS*6110* from *M. tuberculosis* does not cross-hybridize with IS*6110* from *M. smegmatis*. We thank J. Driscoll, V. Escuyer, A Verschoor, and members of the Derbyshire laboratory for comments on the manuscript. The Wadsworth Center Molecular Genetics Core performed DNA sequence analysis.

This work was supported by funding from the NIH (AI 42308) to K.M.D.

REFERENCES

1. **Bhatt, A., H. M. Kieser, R. E. Melton, and T. Kieser.** 2002. Plasmid transfer from *Streptomyces* to *Mycobacterium smegmatis* by spontaneous transformation. Mol. Microbiol. **43:**135–146.
2. **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.
3. **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.
4. **Craig, N. L., R. Craigie, M. Gellert, and A. M. Lambowitz.** 2002. Mobile DNA II. ASM Press, Washington, DC.
5. **Derbyshire, K. M., and S. Bardarov.** 2000. DNA transfer in mycobacteria: conjugation and transduction, p. 93–107. *In* G. F. Hatfull and W. R. Jacobs (ed.), Molecular genetics of mycobacteria. ASM Press, Washington, DC.
6. **Derbyshire, K. M., and N. D. Grindley.** 1992. Binding of the IS903 transposase to its inverted repeat in vitro. EMBO J. **11:**3449–3455.
7. **Derbyshire, K. M., L. Hwang, and N. D. Grindley.** 1987. Genetic analysis of the interaction of the insertion sequence IS903 transposase with its terminal inverted repeats. Proc. Natl. Acad. Sci. USA **84:**8049–8053.
8. **Flint, J. L., J. C. Kowalski, P. K. Karnati, and K. M. Derbyshire.** 2004. The RD1 virulence locus of *Mycobacterium tuberculosis* regulates DNA transfer in *Mycobacterium smegmatis*. Proc. Natl. Acad. Sci. USA **101:**12598–12603.
9. **Frost, L. S., R. Leplae, A. O. Summers, and A. Toussaint.** 2005. Mobile genetic elements: the agents of open source evolution. Nat. Rev. Microbiol. **3:**722–732.
10. **Guttierrez, M. C., S. Brisse, R. Brosch, M. Fabre, B. Omais, M. Marmiesse, P. Supply, and V. Vincent.** 2005. Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*. PLoS Pathog. **1:**1–7.
11. **Lee, S., J. Kriakov, C. Vilcheze, Z. Dai, G. F. Hatfull, and W. R. Jacobs, Jr.** 2004. Bxz1, a new generalized transducing phage for mycobacteria. FEMS Microbiol. Lett. **241:**271–276.
12. **McEvoy, C. R. E., A. A. Falmer, N. C. Gey van Pittius, T. C. Victor, P. D. van Helden, and R. M. Warren.** 2007. The role of IS*6110* in the evolution of *Mycobacterium tuberculosis*. Tuberculosis **87:**393–404.
13. **Mizuguchi, Y., K. Suga, and T. Tokunaga.** 1976. Multiple mating types of *Mycobacterium smegmatis*. Jpn. J. Microbiol. **20:**435–443.
14. **Parsons, L. M., C. S. Jankowski, and K. M. Derbyshire.** 1998. Conjugal transfer of chromosomal DNA in *Mycobacterium smegmatis*. Mol. Microbiol. **28:**571–582.
15. **Smith, N. H., S. V. Gordon, R. de la Rua-Domenech, R. S. Clifton-Hadley, and R. G. Hewinson.** 2006. Bottlenecks and broomsticks: the molecular evolution of *Mycobacterium bovis*. Nat. Rev. Microbiol. **4:**670–681.
16. **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.
17. **Thierry, D., A. Brisson-Noel, V. Vincent-Levy-Frebault, S. Nguyen, J. L. Guesdon, and B. Gicquel.** 1990. Characterization of a *Mycobacterium tuberculosis* insertion sequence, IS*6110*, and its application in diagnosis. J. Clin. Microbiol. **28:**2668–2673.
18. **Thierry, D., M. D. Cave, K. D. Eisenach, J. T. Crawford, J. H. Bates, B. Gicquel, and J. L. Guesdon.** 1990. IS*6110*, an IS-like element of *Mycobacterium tuberculosis* complex. Nucleic Acids Res. **18:**188.
19. **Thomas, C. M., and K. M. Nielsen.** 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat. Rev. Microbiol. **3:**711–721.
20. **van Embden, J. D., M. D. Cave, J. T. Crawford, J. W. Dale, K. D. Eisenach, B. Gicquel, P. Hermans, C. Martin, R. McAdam, T. M. Shinnick, et al.** 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J. Clin. Microbiol. **31:**406–409.
21. **Wang, J., P. K. Karnati, C. M. Takacs, J. C. Kowalski, and K. M. Derbyshire.** 2005. Chromosomal DNA transfer in *Mycobacterium smegmatis* is mechanistically different from classical Hfr chromosomal DNA transfer. Mol. Microbiol. **58:**280–288.
22. **Wang, J., L. M. Parsons, and K. M. Derbyshire.** 2003. Unconventional conjugal DNA transfer in mycobacteria. Nat. Genet. **34:**80–84.
23. **Zerbib, D., P. Prentki, P. Gamas, E. Freund, D. J. Galas, and M. Chandler.** 1990. Functional organization of the ends of IS1: specific binding site for an IS 1-encoded protein. Mol. Microbiol. **4:**1477–1486.