

Special Issue in Memory of John Maynard Smith

doi:10.1016/j.jtbi.2005.08.036

Crown copyright © 2005 Published by Elsevier Ltd.

## Ecotypes of the *Mycobacterium tuberculosis* complex

**Noel H. Smith<sup>a,\*</sup> , Kristin Kremer<sup>b</sup>, Jacqueline Inwald<sup>a</sup>, James Dale<sup>a</sup>,  
Jeffrey R. Driscoll<sup>c</sup>, Stephen V. Gordon<sup>a</sup>, Dick van Soolingen<sup>b</sup>, R. Glyn  
Hewinson<sup>a</sup> and John Maynard Smith<sup>a</sup>**

<sup>a</sup>TB Research Group, Veterinary Laboratories Agency (VLA), Weybridge, New  
Haw, Addlestone, Surrey KT15 3NB, UK

<sup>b</sup>National Institute of Public Health and the Environment (RIVM), P.O. Box 1, 3720  
BA, Bilthoven, The Netherlands

<sup>c</sup>Wadsworth Center, New York State Department of Health, P.O. Box 22002, New  
Scotland Avenue, Albany, NY 12201-2002, USA

\* Corresponding author. Tel.: +44 1273 873502; fax: +44 1273 678433.

☠ Deceased.

### Abstract

A phylogeny of the *Mycobacterium tuberculosis* complex has recently shown that the animal-adapted strains are found in a single lineage marked by the deletion of chromosomal region 9 (RD9) [Brosch et al., 2002. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. Proc. Natl Acad. Sci. USA 99 (6), 3684–3689]. We have obtained the spoligotype patterns of the RD9 deleted strains used to generate this new evolutionary scenario and we show that the presence of spoligotype spacers 3, 9, 16, 39, and 40–43 is phylogenetically informative in this lineage. We have used the phylogenetically informative spoligotype spacers to screen a database of spoligotype patterns and have identified further members of a group of strains apparently host-adapted to antelopes. The presence of the spoligotype spacers is congruent with the phylogeny generated by chromosomal deletions, suggesting that recombination is rare or absent between strains of this lineage. The phylogenetically informative spacers, in concert with the previously identified single nucleotide mutations and chromosomal deletions, can be used to identify a series of clades in the RD9 deleted lineage each with a separate host preference. Finally, we discuss the application of the ecotype concept to this series of clades and suggest that the *M. tuberculosis* complex may best be described as a series of host-adapted ecotypes.

### 1. Introduction

To further understand the nature and evolution of bacterial species we have investigated the animal-adapted strains of the *Mycobacterium tuberculosis* complex. This group of bacteria are characterized by 99.9% similarity at the nucleotide level and identical 16S rRNA sequences and include the important human pathogen *M. tuberculosis* (Sreevatsan et al., 1997; Brosch et al., 2002). The population structure of this group of organisms is apparently highly clonal, with no recombination of chromosomal sequences between strains (Cole et al., 1998; Gutacker et al., 2002; Hirsh et al., 2004). In the absence of inter-strain recombination once a non-repetitive chromosomal region has been deleted it cannot be replaced and the deletion is a marker of a single cell and all its descendants. This feature of a clonal organism was exploited by Brosch et al.

(Brosch et al., 2002) to describe a new evolutionary scenario for this group which concluded that all of the strains host-adapted to animals formed a nested lineage marked by the absence of a specific chromosomal region (RD9). The RD9 deleted lineage, which excludes the human-adapted pathogen *M. tuberculosis* as well as the more distantly related *M. canettii*, includes strains of an apparently human-adapted species, *M. africanum*, as well as strains adapted to a variety of other mammals including *M. microti* (found in voles, wood mice and shrews), *M. pinnipedii* (found in marine mammals), *M. caprae* (associated with goats) and *M. bovis* (associated with cattle). The lineage also includes strains, such as the Oryx strain, that have not been given a species name but are identified by the host from which they are most frequently recovered (Lomme et al., 1976).

The most common molecular typing method applied to animal-adapted strains of the *M. tuberculosis* complex is spoligotyping (Kamerbeek et al., 1997; Durr et al., 2000). This molecular typing method identifies polymorphism in the presence of spacer units in the direct repeat region in strains of the *M. tuberculosis* complex (Kamerbeek et al., 1997; van der Zanden et al., 1998). The direct repeat region is composed of multiple, virtually identical, 36-bp repeats interspersed with a series of unique spacer sequences of similar size. Spoligotype patterns are polymorphic among isolates due to the absence of one or more spacers and adjacent repeat units. Several studies of this region in strains of *M. tuberculosis* have concluded that the evolutionary trend of this region is primarily by loss of one or several contiguous spacer regions (Groenen et al., 1993; Fang et al., 1998; van Embden et al., 2000). We have investigated the loss of spoligotype spacer units in strains of the RD9 deleted lineage described by Brosch et al. (2002) and we conclude that the loss of some spacer units can be used as phylogenetic markers in a similar manner to the loss of chromosomal regions. In combination with the molecular markers described by Brosch et al. (2002) the spoligotype patterns define a series of groups of bacteria, that we call clades.

## 2. Spoligotypes of the RD9 deleted lineage

We have obtained the spoligotype patterns of all the strains used by Brosch et al. (2002) to define the phylogeny of the RD9 deleted lineage of the *M. tuberculosis* complex. The *M. africanum* isolates were spoligotyped at the Wadsworth Center, caprine and some bovine spoligotype patterns by S. Samper, Spain (Haddad et al., 2001) and all other patterns were obtained from the RIVM database of spoligotype patterns. A representative sample of strains from this lineage and their spoligotype patterns as well as single nucleotide mutations (SNMs) and the chromosomal deletions, that had previously been used to define the phylogeny of the lineage, are shown in Table 1. The full set of spoligotype patterns for the strains used by Brosch et al. (2002) can be obtained from [www.Mbovis.org/pdf/fullpatterns.doc](http://www.Mbovis.org/pdf/fullpatterns.doc). As reported by Brosch et al. (2002) these strains form a series of nested clades marked by phylogenetically informative deletions and SNMs. In Table 1 we have also defined a series of intermediate ancestral strains (anc2-anc7) which describe the hypothetical ancestral strains down this nested lineage (Fig. 1). The caprine clade, for example, were presumably derived from a strain, represented by anc6, that had deleted regions RD12 and RD13 but for all other informative markers was identical to anc5 (Fig. 1). Anc1 represents the most recent common ancestor to all strains of *M. tuberculosis* as well as the strains of the RD9 deleted lineage. This ancestor contains all spoligotype spacer units and all phylogenetically informative RD regions identified by Brosch et al. (2002).

Table 1.

Spoligotype pattern and phylogenetically informative markers of representative strains of the RD9 deleted lineage defined by Brosch et al. (2002)

Representative strain	SNMs <sup>a</sup>			Spoligotype pattern <sup>b</sup>	Chromosomal deletions <sup>c</sup>				Clade
	<i>pncA</i>	<i>OxyR</i>	<i>mmpL</i>		RD9	RD7, 8, 10	RD12, 13	RD4	
<b>anc1</b>	C	G	C	<b>11111111111111111111111111111111111111</b>	+	+	+	+	
<b>anc2</b>	C	G	C	<b>111111101111111111111111111111111101111</b>	-	+	+	+	
IS960090	C	G	C	<b>111111100000111111111111111111111010001111</b>	-	+	+	+	Africanum clade 1
<b>anc3</b>	C	G	C	<b>111111101111111111111111111111111101111</b>	-	-	+	+	
67(131)	C	G	C	<b>1111110001111111111110001111111111111101111</b>	-	-	+	+	Africanum clade 2
<b>anc4</b>	C	G	G	<b>1111111011111101111111111111111111101111</b>	-	-	+	+	
60	C	G	G	<b>1110000001111000010000001111111111001101111</b>	-	-	+	+	Antelope
<b>anc5</b>	C	G	G	<b>1101111101111110111111111111111111100000</b>	-	-	+	+	
40	C	G	G	0001111000000000000000111111111111111100000	-	-	+	+	Seal/vole
<b>anc6</b>	C	G	G	<b>1101111101111110111111111111111111100000</b>	-	-	-	+	
UZ22	C	G	G	01000000000000011111111110100001111100000	-	-	-	+	Caprine
<b>anc7</b>	G	A	G	<b>1101111101111110111111111111111111100000</b>	-	-	-	-	
UZ1	G	A	G	<b>1101111101111110111101111111111111100000</b>	-	-	-	-	Bovine

The full data set is available at [www.Mbovis.org/pdf/fullpatterns.doc](http://www.Mbovis.org/pdf/fullpatterns.doc).

<sup>a</sup> The single nucleotide mutations described are at codon 57 in *pncA*, nucleotide 285 for *oxyR* and at codon 551 in *mmpL*.

<sup>b</sup> The presence of a spacer is indicated by a 1 and its absence by a zero. Phylogenetically informative spacers are shown in bold.

<sup>c</sup> Regions of difference defined by Brosch et al. A dash indicates the region is absent in these strains, a cross indicates it is present.

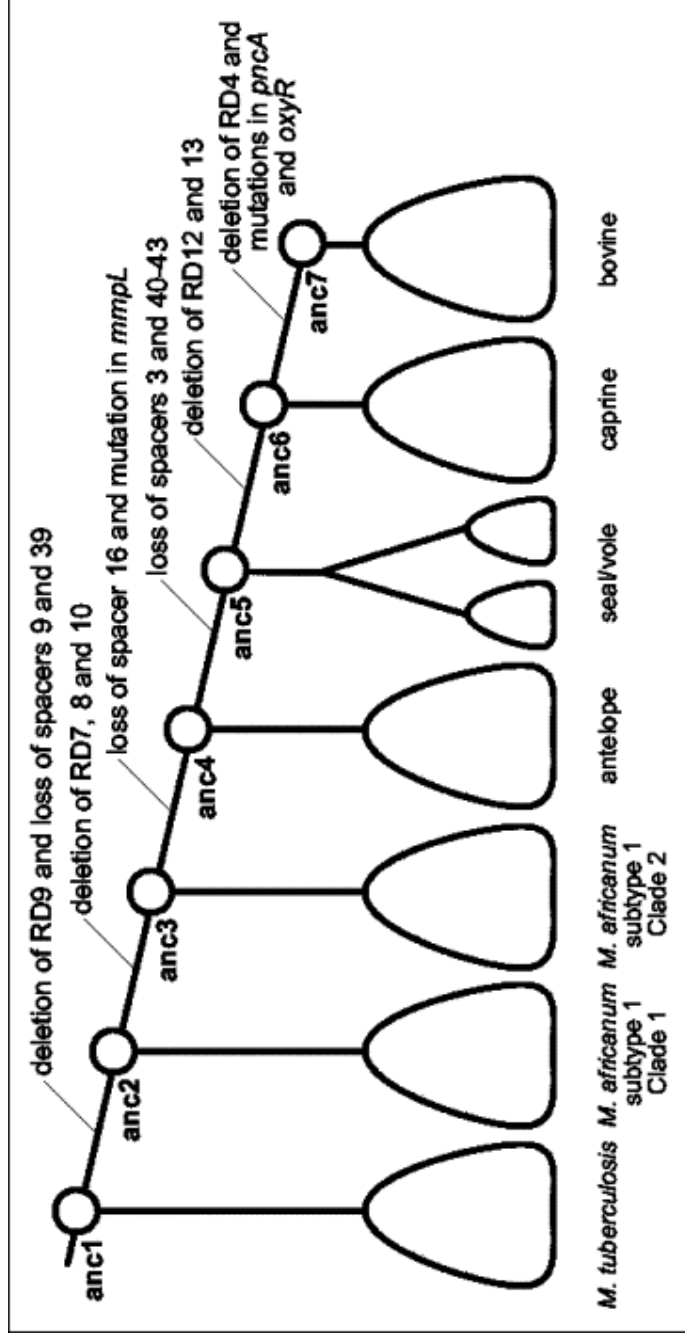


Fig. 1. The phylogeny of the RD9 deleted lineage and *M. tuberculosis* showing the informative deletions (RD deletions), single nucleotide mutations and deleted spoligotype spacers used to define the clades. The hypothetical ancestors, anc1-anc7, are shown as open circles.

The presence of certain spacers in the spoligotype patterns can also be used as phylogenetic markers for this lineage (shown in bold in Table 1). For example, the last four spacers of the standard spoligotype pattern, spacers 40–43, are present in strains derived from anc1 to anc4 but are absent from all descendants of the hypothetical ancestral strain anc5. This suggests that spacers 40–43 were deleted in anc5 and are therefore absent from all strains descended from this hypothetical ancestor. These phylogenetically informative spacers can be identified because they are present in some, or all, isolates of the clades preceding the hypothetical ancestor but absent in all descendants. The phylogenetically informative spacers are spacers 9 and 39 (deleted in anc2), spacer 16 (deleted in anc4) and spacers 3 and 40–43 (deleted in anc5).

### 3. The antelope-associated clade

Although the presence of certain spacers is phylogenetically informative this is not true for the absence of spacers; some spacers have been independently lost in different clades. For example, strains of *M. microti* from the seal/vole clade have lost all but two of the spacer units. Therefore, on its own, the absence of spacers cannot be used to define membership of a clade; it is the presence of certain spacers that is phylogenetically congruent with the deletions and SNMs used by Brosch et al. (2002). It is, however, possible to use the spoligotype pattern as a preliminary screen to identify members of a clade. We have used this approach to identify other strains associated with the oryx strains descended from anc4. A database of 3064 spoligotype patterns of strains of the *M. tuberculosis* complex, obtained from the RIVM, The Netherlands, was screened for spoligotype patterns that contained spacers 3 and 40–43 and lacked spacers 9 and 39. We identified 50 strains with these characteristics; the majority isolated from humans and presumably strains of *M. tuberculosis*. However, nine strains were isolated from artiodactyl hosts rather than humans [gazelle, deer, oryx (2), antelope (2), waterbuck (3)]. Each of these strains had over 19 copies of IS6110 which is unusual for *M. tuberculosis* complex strains isolated from animals (van Soolingen et al., 1994). Isolates were available for five of these strains (two antelope and three waterbuck strains) and chromosomal deletion and SNM analysis showed that they lacked chromosomal regions RD7–10 but retained regions RD4, RD12 and RD13. The presence of a G residue in codon 551 of the *mmpL* gene confirmed that these strains are descendants of anc4 and, along with the two oryx strains identified by Brosch et al. (2002), are members of a group we refer to as the antelope clade. Unlike the other clades shown in Table 1 strains of the antelope clade are not defined by a unique deletion pattern but by a combination of deletion typing, SNMs and spoligotype pattern.

### 4. Discussion

We have shown that the presence of some spoligotype spacers is phylogenetically informative for strains of the RD9 deleted lineage identified by Brosch et al. (2002) and, using the informative spoligotype spacer units, SNMs and chromosomal deletions we have defined a series of nested clades; all of the phylogenetically informative markers are congruent down this lineage. We have used the phylogenetically informative spacers as a screening method to identify further members of a group associated with the previously identified oryx strains and we have named this group the antelope clade.

The phylogenetic congruence shown here between SNMs, chromosomal deletions and the presence of some spoligotype spacers is strong evidence for the absence of inter-strain transfer and recombination of chromosomal DNA between strains of the RD9 deleted lineage. However, this evidence of clonality must be treated with

caution. Strains are most likely to interact with other strains from the same host and, therefore, members of the same clade. The transfer of genetic material between members of the same clade will not distort the relationship between molecular markers described here. Rather, the phylogenetic congruence of chromosomally dispersed molecular markers is evidence that recombination has been rare between members of different clades.

Many of the clades show host association, however, this association is not as strict as with other host-adapted pathogens such as *Salmonella typhi* (Selander et al., 1990). Bovine strains, for example, are most frequently isolated from domestic cattle, however, they are also recovered from other bovids as well as from other mammalian hosts, including humans. The isolation of related strains from several hosts either represents the spill-over of strains into an alternate host or a broadening of host preference.

Strains of the two *M. africanum* sub-type 1 clades were recovered from humans in West Africa (Niemann et al., 2004). However, alternative mammalian hosts for these strains have not been fully investigated and humans may represent a spill-over host for these strains. Furthermore, some of the clades contain sub-groups with different host preferences. Strains of the seal/vole clade (directly descended from anc5 without change in the phylogenetically informative markers) have been isolated from many different hosts broadly falling into two categories, seals and voles. Seal isolates, now known as *M. pinnipedii*, have been found in several species of marine mammals including sea lions and fur seals (Forshaw and Phelps, 1991; Cousins et al., 2003). Vole strains, identified as *M. microti*, have been isolated from a number of small mammals including bank voles, mice and shrews (Wells and Oxon, 1937; Wells, 1946; Kremer et al., 1998; Cavanagh et al., 2002). Isolates from either the seal or the vole sub-groups can readily be distinguished by spoligotype pattern and sub-group specific deletions of chromosomal regions (Ahmed et al., 2003; Frota et al., 2004; Garcia-Pelayo et al., 2004). Therefore, each of the clades may contain, yet to be identified, sub-groups of strains with different host preferences. Remarkably, the phylogeny of the animal-adapted strains of the *M. tuberculosis* complex, as described by Brosch et al. (2002), is a series of nested clades suggesting that new host-adapted groups developed sequentially. In contrast, a star-like phylogeny, seen in other clonal, host-adapted pathogens such as the salmonellae (Selander and Smith, 1990), implies that strains with new host preferences were independently derived from a common source.

The clades we have identified include several mycobacterial species. In general, these species have been identified because they were isolated from the same host and have phenotypic and molecular features in common. For organisms that rarely, if ever, recombine there is no reason to suggest that they will fall into discrete kinds without a mechanism to constrain diversity and stop them diverging without bound (Smith et al., 2000). Recently, Cohan (2002) has suggested a theory-based approach to defining groups of bacteria that have similar properties to eukaryotic species. He suggests that the most important attributes of species are that divergence within the species is limited by some cohesive force, that divergence between different species is irreversible and that different species are ecologically distinct. For bacteria, that do not have their diversity constrained by sexual reproduction, Cohan argues that the cohesive force limiting divergence within a group, which he calls ecotypes rather than species, is periodic selection (selective sweeps) (Cohan, 2002). A new ecotype evolves when a strain adapts to a new niche and becomes immune to periodic selection events in the ancestral population. He suggests that ecotypes manifest

the same important properties held by eukaryotic species (Cohan, 2001 and Cohan, 2002).

We can apply the ecotype concept to strains of the RD9 deleted lineage. Strains of this lineage occupy a series of clades each with a different host preference representing different niches. There are molecular differences between the clades that were present in the most recent common ancestor of the clade and, therefore, present in all descendants; we have used these differences to define the clades shown in Table 1. These phylogenetically informative molecular differences are the result of bottlenecks in diversity presumably during the early origin of each clade and are fixed in the population either as a result of sampling (founder effect or drift) or a result of selective sweeps in the founding population as it adapts to a new host. It is yet to be determined if the phylogenetically informative differences used in this study were the loci under selection during these selective sweeps or if they are selectively neutral, or even deleterious and hitchhiked to fixation with highly selected loci elsewhere on the chromosome.

An important feature of ecotypes is that they are immune to the diversity reducing effects of selective sweeps that occur in ecotypes occupying other niches. We suspect that strains of the *Mycobacteria* adapted to one host will be unable to sustain their population in an alternate host with its own host-adapted strain. It has been shown, for example, that vole strains cause little disease in cattle (S. Gordon, unpublished data) and healthy humans (Wells, 1949; Paul, 1961). It is reasonable to argue, therefore, that strains of the vole-adapted clade will not successfully transmit between cattle or between humans and will not out-compete bovine strains in cattle or *M. tuberculosis* in humans. Similar arguments can be made for bovine-associated strains in humans and *M. tuberculosis* in cattle (van Soolingen et al., 1994; van Soolingen, 2001). Furthermore, assuming that this group of organisms is purely clonal, fixed molecular differences between the clades is evidence that selective sweeps have not recently crossed clade boundaries. Such sweeps would homogenize the population in each niche and remove the diversity we have used to identify different clades.

The *M. tuberculosis* complex, therefore, may represent a series of host-adapted groups that conform to the ecotype concept of Cohan (2002). Each of the ecotypes has a distinct host preference, representing its niche, and is marked by fixed molecular differences within each ecotype. The fixed differences are evidence that selective sweeps in one ecotype do not replace the strains of another ecotype. Strains of the human-adapted *M. tuberculosis* would be regarded as a single ecotype of a much larger series that make up the *M. tuberculosis* complex. However, the clades identified here may not themselves represent ecotypes; each clade may contain several ecotypes with different host preference. The seal/vole strains described above, for example, probably represent the division of a clade into different host-adapted ecotypes. Further division within these groups is possible. The vole strains, all regarded as strains of *M. microti*, have been isolated from hosts as disparate as bank voles and shrews and further work is required to determine if these represent different ecotypes.

Although the ecotype concept is useful for identifying groups of bacteria with population genetics properties similar to eukaryotic species it is not without practical problems. Strains of the bovine-adapted ecotype from cattle in South America, for example, may rarely, if ever, be exposed to the selective sweeps that happen in bovine strains from the UK. It would seem inevitable that geographical isolation will fracture each of these host-adapted ecotypes into smaller and smaller geographical units leading to a very cumbersome nomenclature. However, the ecotype concept does suggest a number of new

experimental approaches. *M. bovis* has traditionally been regarded as a generalist, an organism that can infect many hosts with the implication that there exist genetic differences in *M. tuberculosis* that limit this pathogen to its human host. If, however, the *M. tuberculosis* complex is a series of host-adapted ecotypes then it is more relevant to identify the ecotypes and investigate the genetic differences that control adaptation to each host rather than looking for mechanisms that make *M. bovis* a generalist.

### **Acknowledgements**

We should like to thank S. Samper for supplying spoligotype patterns. This work was funded by DEFRA.



## References

- Ahmed et al., 2003 N. Ahmed and M. Alam et al., Genome sequence based, comparative analysis of the fluorescent amplified fragment length polymorphisms FAFLP of tubercle bacilli from seals provides molecular evidence for a new species within the *Mycobacterium tuberculosis* complex, *Infect. Genet. Evol.* 2 (2003) (3), pp. 193–199.
- Brosch et al., 2002 R. Brosch and S.V. Gordon et al., A new evolutionary scenario for the *Mycobacterium tuberculosis* complex, *Proc. Natl Acad. Sci. USA* 99 (2002) (6), pp. 3684–3689.
- Cavanagh et al., 2002 R. Cavanagh and M. Begon et al., *Mycobacterium microti* infection vole tuberculosis in wild rodent populations, *J. Clin. Microbiol.* 40 (2002) (9), pp. 3281–3285.
- Cohan, 2001 F.M. Cohan, Bacterial species and speciation, *Syst. Biol.* 50 (2001) (4), pp. 513–524.
- Cohan, 2002 F.M. Cohan, What are bacterial species?, *Annu. Rev. Microbiol.* 56 (2002), pp. 457–487.
- Cole et al., 1998 S.T. Cole and R. Brosch et al., Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence, *Nature* 393 (1998) (6685), pp. 537–544.
- Cousins et al., 2003 D.V. Cousins and R. Bastida et al., Tuberculosis in seals caused by a novel member of the *Mycobacterium tuberculosis* complex: *Mycobacterium pinnipedii* sp. nov, *Int. J. Syst. Evol. Microbiol.* 53 (2003) (Part), pp. 1305–1314.
- Durr et al., 2000 P.A. Durr and R.S. Clifton-Hadley et al., Molecular epidemiology of bovine tuberculosis. II. Applications of genotyping, *Rev. Sci. Technol.* 19 (2000) (3), pp. 689–701.
- Fang et al., 1998 Z. Fang and N. Morrison et al., IS6110 transposition and evolutionary scenario of the direct repeat locus in a group of closely related *Mycobacterium tuberculosis* strains, *J. Bacteriol.* 180 (1998) (8), pp. 2102–2109.
- Forshaw and Phelps, 1991 D. Forshaw and G.R. Phelps, Tuberculosis in a captive colony of pinnipeds, *J. Wildlife Dis.* 27 (1991) (2), pp. 288–295.
- Frota et al., 2004 C.C. Frota and D.M. Hunt et al., Genome structure in the vole bacillus, *Mycobacterium microti*, a member of the *Mycobacterium tuberculosis* complex with a low virulence for humans, *Microbiology* 150 (2004) (Part), pp. 1519–1527.
- Garcia-Pelayo et al., 2004 M.C. Garcia-Pelayo and K.C. Caimi et al., Microarray analysis of *Mycobacterium microti* reveals deletion of genes encoding PE-PPE proteins and ESAT-6 family antigens, *Tuberc. Edinburgh* 84 (2004) (3–4), pp. 159–166.
- Groenen et al., 1993 P.M. Groenen and A.E. Bunschoten et al., Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method, *Mol. Microbiol.* 10 (1993) (5), pp. 1057–1065.
- Gutacker et al., 2002 M.M. Gutacker and J.C. Smoot et al., Genome-wide analysis of synonymous single nucleotide polymorphisms in *Mycobacterium tuberculosis* complex organisms: resolution of genetic relationships among closely related microbial strains, *Genetics* 162 (2002) (4), pp. 1533–1543.
- Haddad et al., 2001 N. Haddad and A. Ostyn et al., Spoligotype diversity of *Mycobacterium bovis* strains isolated in France from 1979 to 2000, *J. Clin. Microbiol.* 39 (2001) (10), pp. 3623–3632.
- Hirsh et al., 2004 A.E. Hirsh and A.G. Tsolaki et al., Stable association between strains of *Mycobacterium tuberculosis* and their human host populations, *Proc. Natl Acad. Sci. USA* 101 (2004) (14), pp. 4871–4876.
- Kamerbeek et al., 1997 J. Kamerbeek and L. Schouls et al., Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology, *J. Clin. Microbiol.* 35 (1997) (4), pp. 907–914.

Kremer et al., 1998 K. Kremer and D. van Soolingen et al., *Mycobacterium microti*: more widespread than previously thought, *J. Clin. Microbiol.* 36 (1998) (9), pp. 2793–2794.

Lomme et al., 1976 J.R. Lomme and C.O. Thoen et al., *Mycobacterium tuberculosis* infection in two East African oryxes, *J. Am. Vet. Med. Assoc.* 1699 (1976), pp. 912–914.

Niemann et al., 2004 S. Niemann and T. Kubica et al., The species *Mycobacterium africanum* in the light of new molecular markers, *J. Clin. Microbiol.* 42 (2004) (9), pp. 3958–3962.

Paul, 1961 R. Paul, The effects of vole bacillus vaccination of African mine workers in the Northern Rhodesian copper mines, *Br. J. Ind. Med.* 18 (1961), pp. 148–152.

Selander and Smith, 1990 R.K. Selander and N.H. Smith, Molecular population genetics of *Salmonella*, *Rev. Med. Microbiol.* 1 (1990), pp. 219–228.

Selander et al., 1990 R.K. Selander and P. Beltran et al., Evolutionary genetic relationships of clones of *Salmonella* serovars that cause human typhoid and other enteric fevers, *Infect. Immun.* 58 (1990) (7), pp. 2262–2275.

Smith et al., 2000 J.M. Smith and E.J. Feil et al., Population structure and evolutionary dynamics of pathogenic bacteria, *Bioessays* 22 (2000) (12), pp. 1115–1122.

Sreevatsan et al., 1997 S. Sreevatsan and X. Pan et al., Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination, *Proc. Natl Acad. Sci. USA* 94 (1997) (18), pp. 9869–9874.

van der Zanden et al., 1998 A.G. van der Zanden and A.H. Hoentjen et al., Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* complex in paraffin wax embedded tissues and in stained microscopic preparations, *Mol. Pathol.* 51 (1998) (4), pp. 209–214.

van Embden et al., 2000 J.D. van Embden and T. van Gorkom et al., Genetic variation and evolutionary origin of the direct repeat locus of *Mycobacterium tuberculosis* complex bacteria, *J. Bacteriol.* 182 (2000) (9), pp. 2393–2401.

van Soolingen, 2001 D. van Soolingen, Molecular epidemiology of tuberculosis and other mycobacterial infections: main methodologies and achievements, *J. Intern. Med.* 249 (2001) (1), pp. 1–26.

van Soolingen et al., 1994 D. van Soolingen and P.E. de Haas et al., Use of various genetic markers in differentiation of *Mycobacterium bovis* strains from animals and humans and for studying epidemiology of bovine tuberculosis, *J. Clin. Microbiol.* 32 (1994) (10), pp. 2425–2433.

Wells, 1946 A.Q. Wells, The murine type of tubercle bacillus (the vole acid-fast bacillus). Special Report Series in Medicine, Council of London, no. 259, Sir William Dunn School of Pathology, University of Oxford, Oxford, UK (1946).

Wells, 1949 A.Q. Wells, Vaccination with the murine type of tubercle bacillus vole bacillus, *Lancet* 254 (1949), pp. 53–55.

Wells and Oxon, 1937 A.Q. Wells and D.M. Oxon, Tuberculosis in wild voles, *Lancet* (1937), p. 1221.