



HAL
open science

Whole-Genome Analysis of *Mycobacterium avium* subsp. *paratuberculosis* IS900 distribution reveals strain type-specific modalities

Cyril Conde, Marian Price-Carter, Thierry Cochard, Maxime Branger, Karen Stevenson, Richard Whittington, John Bannantine, Franck Biet

► To cite this version:

Cyril Conde, Marian Price-Carter, Thierry Cochard, Maxime Branger, Karen Stevenson, et al.. Whole-Genome Analysis of *Mycobacterium avium* subsp. *paratuberculosis* IS900 distribution reveals strain type-specific modalities. 15th International Association for Paratuberculosis - ICP 2022, Jun 2022, Dublin, Ireland. . hal-04203656

HAL Id: hal-04203656

<https://hal.inrae.fr/hal-04203656>

Submitted on 20 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Whole-Genome Analysis of *Mycobacterium avium* subsp. *paratuberculosis* IS900 Insertion Reveals Strain Type-Specific Modalities

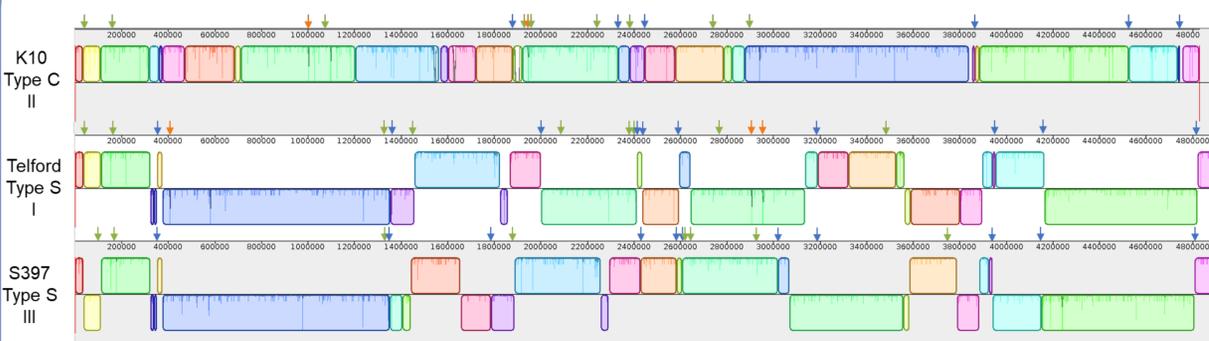
Cyril Condé^a, Marian Price-Carter^b, Thierry Cochard^a, Maxime Branger^a, Karen Stevenson^c, Richard Whittington^d, John P. Bannantine^e and Franck Biet^a
^aISP, INRAE, Université de Tours, UMR 1282, 37380 Nouzilly, France, ^bAgResearch Ltd., Hopkirk Research Institute, Palmerston North, New Zealand, ^cMoredun Research Institute, Penicuik, UK, ^dSchool of Veterinary Science, The University of Sydney, Sydney, Australia, ^eUSDA-ARS, National Animal Disease Center, Ames, USA

Background & Objective

Mycobacterium avium subsp. *paratuberculosis* (Map) is the etiological agent of paratuberculosis or Johne's disease that causes chronic intestinal inflammation in ruminants (1,2). The IS900 insertion sequence, only present in multiple copies in the chromosome of Map, was used as an epidemiological marker and target for qPCR diagnosis but its biological role remains elusive (3,4). During its transposition, the IS may promote a number of important genetic modifications with consequences in the host-pathogen evolution of the strain (5,6). Thanks to the development of sequencing technologies, the number of available Map genomes is increasing rapidly, which makes it easy to analyze the distribution of IS900 in this slow-growing bacterium.

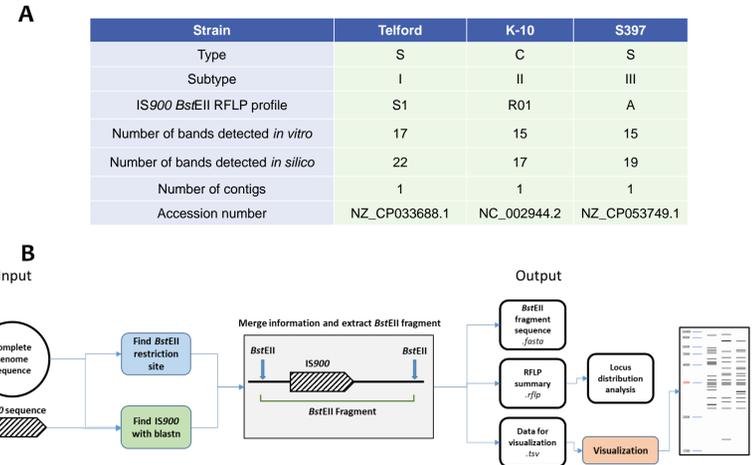
The objective of this study is to use bioinformatics approaches to study the IS900 distribution in the genome of Map strains. In this study we included the genomes of strains belonging to the three genetic lineages that emerged in Map. These lineages are classified as S-type (including subtype I and III) (ovine) and C-type (bovine) strains. As a secondary goal, a tool will be developed to automate the *in silico* restriction fragment length polymorphism (RFLP) analysis using IS900.

Figure 2. Distribution of IS900 copies on Map genomes



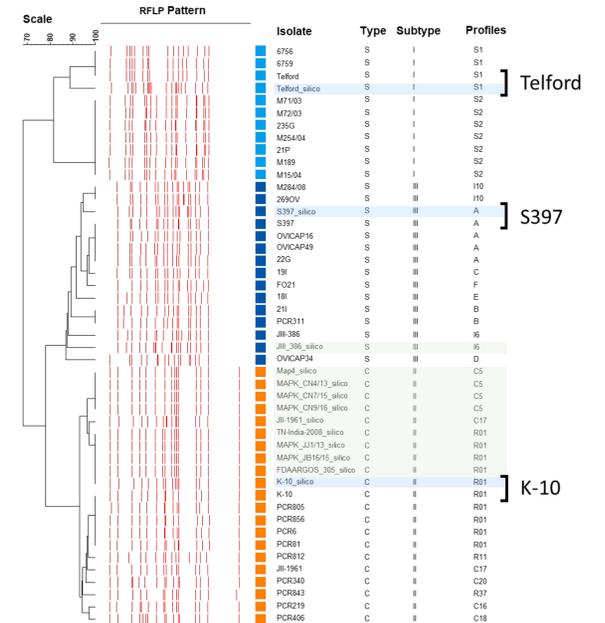
Maui alignment of K-10 (top) (7) with Telford (middle) (8) and S397 (bottom) (9) showing genomic reorganization of the genomes. The colored boxes represent homologous regions present in each genome. Blocks below the centerline indicate regions with inverse orientation. Regions outside the blocks lack homology between the genomes. Within each block there is a similarity profile of the DNA sequences and the white areas indicate sequences specific to a genome. The scale is in base pairs. Orthologous insertions are indicated, by a green arrow, specific insertions are indicated by an orange arrow and conserved loci only in two genomes by a blue arrow.

Figure 1. Materials and methods



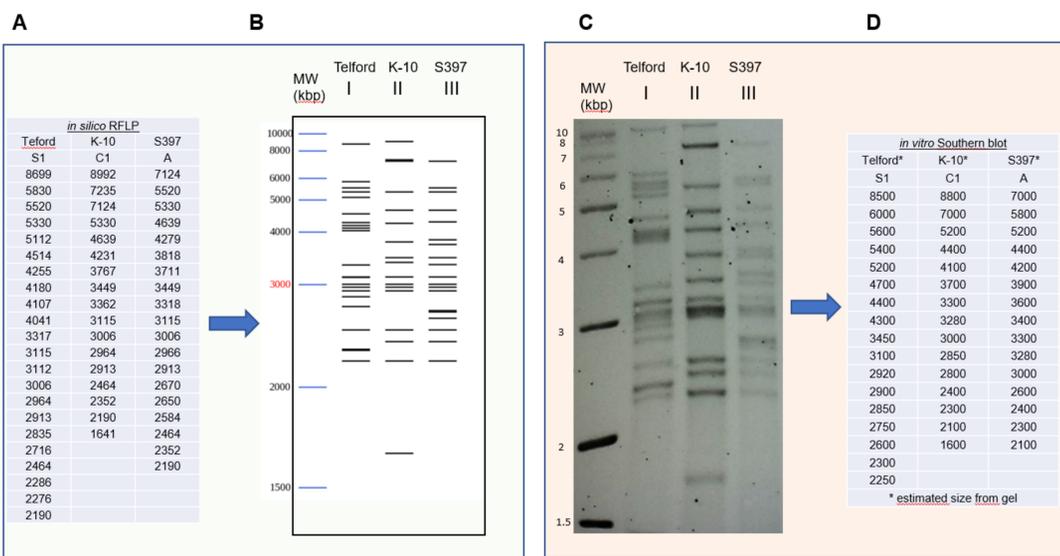
(A) Table 1: details of the strains/genomes used and information on the number of copies of the IS900. (B) Details of steps performed by the IS900 RFLP *in silico* pipeline

Figure 4. Comparison of digital profiles with the existing profiles in the database



UPGMA dendrogram based on IS900 RFLP typing using BstEII (10). Profiles obtained *in silico* can be included and compared to those obtained by classical technic (dark brackets)

Figure 3. IS900 restriction fragment length polymorphism (RFLP) "in vitro" versus "in silico"



(A) An *in silico* RFLP analysis was developed using complete genomes. This automated procedure identified the BstEII restriction sites to obtain the exact size of all DNA fragments carrying a copy of the IS900 sequence. IS900 RFLP profiles were compared using fragment sizes (A) or (B) by digital visualization of the restriction fragments separated according to their size by mimicking their migration in agarose gel. (C) The IS900 RFLP profiles obtained by classical Southern blot method and hybridization to IS900 were used to find the approximate fragment sizes by band analysis software (D)

Results

- The program developed in this study allowed to automatically locate the sequences of the IS900, to know their positions and their number.
- Between 16 to 22 copies of the IS900 were found in the genomes studied.
- Nine insertion site locations were conserved across all genomes studied.
- Analysis of insertion sites reveal that IS900 rarely insert within coding sequences
- The *in silico* IS900 RFLP profiles can be compared by size or by digital visualization mimicking their migration in agarose gel.

Conclusion & Perspectives

- This study provides a new tool that makes it possible to automate the IS900 distribution analysis in Map genome to enrich our knowledge on the dynamics of distribution of this IS for epidemic purpose in link of the evolution within Map species and study the biological implication of the presence of this IS900.
- The profiles and digital profiles can be compared to the many existing profiles in the literature using our dedicated application: <http://mac-inmv.tours.inra.fr>
- More genomes will be analyzed (particularly type S) to determine if there is any correlation between copy number and host lineage.

References : 1. Rathnaiah G, Front Vet Sci 2017 2. Biet F & Boschirolini ML, Res Vet Sci 2014 3. Pavlik I, J Mic Met 1999 4. Bull T, Microbiology 2000 5. Moran NA & Plague GR, 2004 6. Toft & Anderson, 2010 7. Li L, 2005 8. Brauning R, MRA 2019 9. Bannantine JP, MRA 2021 10. Biet F, BMC Microbiology 2012