

## Le génome des mycobactéries : étude biologique et interprétation évolutive

Dominique Labie

► La tuberculose est, à l'heure actuelle, la première cause de mortalité dans le monde, et son incidence est en voie de recrudescence, tant dans les pays industrialisés que dans le monde en développement. On peut y voir deux causes majeures, l'apparition de souches résistantes aux traitements, et surtout la synergie dramatique avec toutes les causes de déficit immunitaire, au premier rang desquelles il faut placer l'infection par le VIH. En 1993, l'OMS a déclaré officiellement la tuberculose comme une urgence mondiale. Le développement de nouveaux abords thérapeutiques, et surtout prophylactiques, se présente donc comme une priorité. La détermination de la séquence complète de *Mycobacterium tuberculosis* en 1998 a permis l'ouverture d'un nouveau chapitre dans les travaux, amorcés déjà depuis plusieurs années. Les équipes de l'Institut Pasteur (Paris, France) sont au premier rang d'une recherche qui a cependant été l'objet d'une mobilisation internationale.

Une des caractéristiques de la tuberculose est son caractère universel, le fait qu'elle touche de nombreuses espèces avec des manifestations physiopathologiques extrêmement variées, son ancienneté aussi, puisqu'on a mis en évidence le bacille dans les momies de l'ancienne Égypte comme dans celles du Pérou précolombien, et dans des squelettes du Moyen-Âge, en Hongrie comme en Angleterre. La souche se caractérise aussi par sa croissance très lente sur les milieux de culture, qui explique le caractère chronique de la maladie, une possibilité de latence prolongée dans l'orga-

nisme, et sa réactivation même après plusieurs décennies sous l'influence d'une baisse de l'immunité. Le vaccin BCG (bacille de Calmette et Guérin) universellement utilisé provient de l'isolement en 1927 d'une souche *M. bovis* BCG, dérivée de *M. bovis*, atténuée par des passages en série sur milieux de culture. Son efficacité s'avère cependant très inégale, de 0 à 80 % selon les pays. Il a été longtemps aussi communément admis que la souche virulente chez l'homme était dérivée d'une souche bovine après domestication du bétail.

Déjà, en 1996, la souche de référence de *M. tuberculosis* avait été comparée par des chercheurs scandinaves à une souche virulente de *M. bovis* et à la souche atténuée de *M. bovis* BCG [1]. Les auteurs avaient mis en évidence la présence, dans les souches virulentes *M. tuberculosis* et *M. bovis*, d'une protéine antigénique ESAT-6, absente par ailleurs de *M. bovis* BCG. D'autres auteurs, à Seattle (USA), par une technique d'hybridation génomique soustractive, avaient montré trois différences entre *M. bovis* et *M. bovis* BCG, délétions retrouvées uniquement dans cette dernière, et désignées RD1, RD2 et RD3 [2]. Les chercheurs de l'Institut Pasteur, enfin, avaient démontré en comparant les cartes de restriction l'existence de différences entre *M. tuberculosis* et *M. bovis* [3]. Mais c'est en 1998 que ces derniers, en collaboration avec les équipes du Sanger Centre en Grande-Bretagne, ont réussi, par l'emploi de clones BAC (*bacterial artificial chromosome*), à établir en quasi-totalité la séquence génomique

Institut Cochin,  
Département de génétique,  
développement et  
pathologie moléculaires,  
InsERM U.567,  
24, rue du Faubourg  
Saint-Jacques, 75014 Paris,  
France.

labie@cochin.inserm.fr

de ≈ 4,4 Mb de la souche virulente de référence H73Rv de *M. tuberculosis* isolée depuis 1905 [4]. À partir de là, une étude comparative des différents génomes et une étude *in silico* permettaient une interprétation biologique qui a fait l'objet de la publication majeure parue dans *Nature* la même année [5]. Le génome d'environ 4000 gènes se caractérise par une proportion particulièrement élevée de guanines et de cytosines dont la composition en acides aminés de la souche, statistiquement biaisée (glycocolle, alanine, proline, arginine et tryptophane étant majoritaires), est un reflet. Cette répartition expliquait que l'enveloppe de la bactérie s'avère exceptionnellement riche en lipides, glycolipides et polysaccharides. On constatait aussi qu'une fraction importante des séquences codantes étaient impliquées dans la synthèse d'enzymes de la lipogenèse et de la lipolyse, ainsi que dans celle de deux familles protéiques d'un type nouveau, riches en glycocolle, et dont la structure répétitive faite de motifs Pro-Glu (PE) et Pro-Pro-Glu (PPE) se présentait comme une source possible de variation antigénique. Des enzymes, souvent associées à une virulence bactérienne, en particulier des phospholipases, étaient également caractérisées.

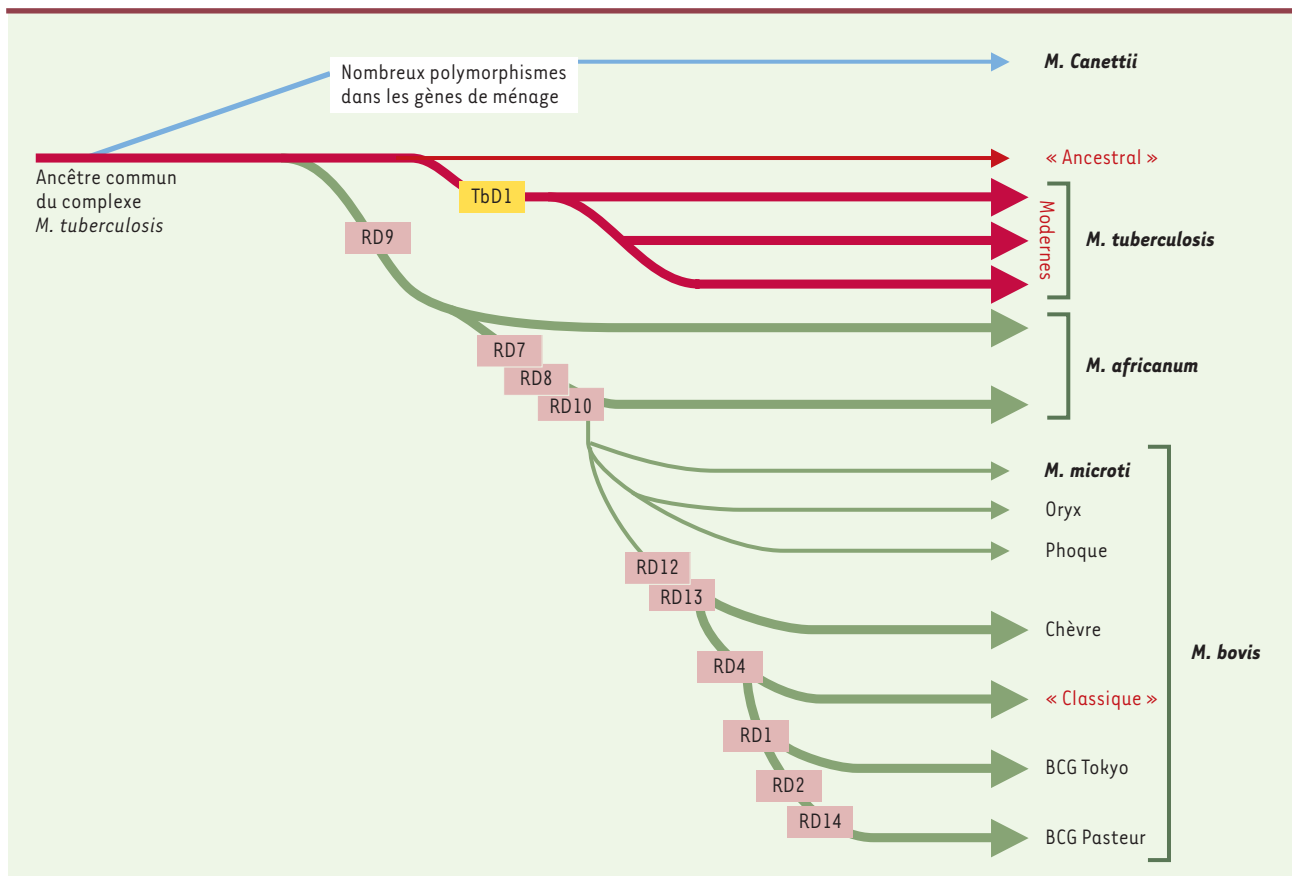
La comparaison des différentes souches, dont la structure génomique avait été déterminée par la même technique de BAC, a fourni un outil à l'équipe de l'Institut Pasteur et aux chercheurs du Centre d'Études du Polymorphisme



Humain (CEPH, Paris, France), leur permettant d'identifier de façon précise les zones variables [6]. Par comparaison avec *M. tuberculosis*, on observe chez *M. bovis BCG* la présence de dix délétions; trois d'entre elles sont les mêmes que celles qui avaient été décrites chez *M. bovis* [2], les sept autres ont été désignées RD4 à RD10. D'autres souches ont été aussi examinées: *M. africanum* s'apparente, de ce point de vue, à *M. tuberculosis* plus qu'à *M. bovis*, alors que *M. microti* a une structure intermédiaire. Les régions délétées semblent coder pour des protéines particulièrement intéressantes. On y trouve entre autres l'invasine Mce, protéine de la mycobactérie qui envahit le macrophage, plusieurs phospholipases, Plc A, B, C et D, sans doute des facteurs de virulence, plusieurs protéines PE, PPE et ESAT-6 vraisemblablement antigéniques

et dont on pourrait envisager l'usage dans une perspective vaccinale, des protéines de type époxyde hydroxylase. On a aussi montré la présence de quelques séquences d'insertion dont l'une, IS6110, interrompt la séquence *plcD*, et de délétions présentes chez *M. tuberculosis* par comparaison à *M. bovis*, désignées RvD1 et RvD2, démontrant donc que l'évolution ne s'est pas faite à sens unique, mais qu'à partir d'un progéniteur commun il y a eu adaptation du bacille à la niche spécifique que lui a fournie son hôte. Le travail le plus récent, enfin, est le fruit d'une collaboration internationale entre chercheurs français, anglais, néerlandais, américains et espagnols [7]. L'étude phylogénique a porté sur 100 souches différentes, toutes apparentées au complexe des mêmes mycobactéries, dont 46 *M. tuberculosis*

incluant les souches provenant des grandes épidémies modernes, mais dont l'origine géographique, l'organisme hôte, humain ou animal, les manifestations phénotypiques étaient extrêmement variés. Un spectre de 20 régions variables, encore accru par rapport aux études précédentes, a été analysé par des approches de génomique comparative en utilisant des techniques d'électrophorèse en champ pulsé, et des analyses *in silico*. Les auteurs ont constaté que la presque totalité des régions RD, sièges de délétions, étaient présentes dans toutes les souches de *M. tuberculosis* quelle que soit leur origine géographique, témoignant ainsi d'un haut degré de conservation. La présence des délétions RvD était, en revanche, inconstante. Une nouvelle délétion était spécifiquement trouvée dans 87 % des souches de *M. tuberculosis*, en particu-



**Figure 1. Schéma proposé pour l'évolution des bacilles tuberculeux, illustrant les pertes successives d'ADN dans certains lignages.** La base du raisonnement est la présence ou l'absence de régions délétées, ainsi que les polymorphismes de certaines séquences (non indiqués) (d'après [7]).



lier dans les souches provenant des dernières grandes épidémies. L'étude *in silico* d'autres génomes montrait à cet emplacement deux gènes, *mmpS6*, absent et *mmL6*, tronqué chez *M. tuberculosis*. Cette délétion spécifique a été désignée TbD1 (*M. tuberculosis specific deletion 1*). Sa taille paraît parfaitement conservée entre toutes les souches d'origines géographiques multiples, faisant penser qu'il s'agit d'un événement très ancien, datant d'un progéniteur.

Les autres souches analysées permettaient une approche phylogénétique. *M. canettii*, souche africaine exceptionnelle, présente un nombre important de polymorphismes suggérant une divergence très ancienne. Globalement, le génome de *M. bovis* est plus petit que celui de *M. tuberculosis*, les génomes de *M. africanum* et de *M. microti* se situant à des positions intermédiaires. Il semble dans ces conditions plausible que *M. bovis* soit l'élément final d'un lignage divergent à partir du progéniteur dont se seraient successivement séparés *M. africanum* et *M. microti* (Figure 1). Des pertes successives d'ADN auraient permis des expansions clonales et des adaptations successives à de nouveaux hôtes. On peut aussi s'interroger sur la pathogénicité du progéniteur ancestral, apparemment très voisin de *M. tuberculosis*. Le fait que les souches animales qui se rapprochent le plus des souches humaines soient également pathogènes est en faveur de cette hypothèse. Il est probable que ces souches anciennes, d'origine africaine ou indienne, étaient à l'origine d'une pathologie endémique; les souches modernes, délétées de la région TbD1, semblent en revanche à l'origine des grandes épidémies et de la diffusion mondiale.

La conservation tout à fait remarquable des gènes ubiquitaires et de la séquence de jonction en TbD1 dans les souches de *M. tuberculosis* rend très vraisemblable l'hypothèse d'un goulot d'étranglement majeur il y a 15 000 à 20 000 ans. Toutes les souches modernes porteuses de la

délétion seraient issues d'un clone unique qui aurait diffusé à travers le monde. La délétion TbD1, dans la mesure où on peut la dater en examinant des ADN de squelettes anciens, se serait produite avant le XVIII<sup>e</sup> siècle, les cas ancestraux prouvés de tuberculose pouvant être dus aux quelques souches de *M. tuberculosis* qui ne présentent pas la délétion TbD1. L'extrême variabilité des hôtes de l'autre branche de *Mycobacterium* (*M. africanum*, *M. microti*, *M. bovis*) rend très vraisemblable qu'il y ait eu là aussi un goulot d'étranglement. Ces souches ont atteint l'homme en Afrique, les phoques en Argentine, les chèvres en Espagne, les blaireaux en Angleterre, les campagnols dans les Orcades. On imagine difficilement sinon la diffusion et l'adaptation d'une même souche à des hôtes aussi variés. Plus de précision et l'analyse d'autres souches sont sans doute encore nécessaires. La vision traditionnelle d'une souche humaine dérivée de la souche bovine est en tous cas tout à fait controversée. ♦

### **Mycobacteria: building an evolutionary scheme from genome comparison**

## RÉFÉRENCES

1. Harboe M, Oettinger T, Gotten Wiker H, Rosenkrands I, Andersen P. Evidence for occurrence of the ESAT-6 protein in *Mycobacterium tuberculosis* and virulent *Mycobacterium bovis* and for its absence in *Mycobacterium bovis* BCG. *Infect Immun* 1996; 64: 16-22.
2. Mahairas GG, Sabo PJ, Hickey MJ, Singh DC, Stover CK. Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. *J Bacteriol* 1996; 178: 1274-82.
3. Philipp WJ, Nair S, Guglielmi G, Laganderie M, Gicquel B, Cole ST. Physical mapping of *Mycobacterium bovis* BCG Pasteur reveals differences from the genome map of *Mycobacterium tuberculosis* H37Rv and from *M. bovis*. *Microbiology* 1996; 142: 3135-45.
4. Brosch R, Gordon SV, Billault A, et al. Use of a *Mycobacterium tuberculosis* H37Rv bacterial artificial chromosome library for genome mapping, sequencing, and comparative genomics. *Infect Immun* 1998; 66: 2221-9.
5. Cole ST, Brosch R, Parkhill J, et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998; 393: 537-44.
6. Gordon SV, Brosch R, Billault A, Garnier T, Eiglmeier K, Cole ST. Identification of variable regions in the genome of tubercle bacilli using bacterial artificial chromosome arrays. *Mol Microbiol* 1999; 32: 643-55.
7. Brosch R, Gordon SV, Marmiesse M, et al. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci USA* 2002; 99: 3684-9.