

## UTILISATION DE LA GÉNOMIQUE DES MYCOBACTÉRIES POUR LA DÉFINITION DE NOUVEAUX VACCINS ET DE CIBLES THÉRAPEUTIQUES

C. DEMANGEL, R. BROSCHE, M. MARMIESSE, S.T. COLE

*Med Trop* 2004; 64 : 221-223

Malgré la multitude de connaissances acquise depuis 120 ans sur *Mycobacterium tuberculosis*, l'agent étiologique de la tuberculose humaine, découverte par Robert Koch en 1881, la tuberculose reste la première cause de mortalité par maladie transmissible en ce début de 21<sup>e</sup> siècle. On estime en effet qu'un tiers de la population mondiale a été infectée par le bacille, et on recense chaque année dans le monde 8 millions de nouveaux cas actifs et trois millions de morts. Depuis les années 80, l'incidence de cette maladie a considérablement augmenté. Cette recrudescence de la tuberculose s'explique d'une part, par l'augmentation de la précarité et de la population dans les pays pauvres, et d'autre part par l'urbanisation croissante, les migrations, l'importante épidémie du sida et l'émergence de souches multirésistantes aux antibiotiques qui posent actuellement de plus en plus de problèmes graves. De plus, l'efficacité du vaccin existant, le bacille de Calmette et Guérin (BCG) est controversée. La protection due au vaccin est en effet souvent insuffisante voire inefficace. Dès lors, en 1993 l'Organisation Mondiale de la Santé (OMS) a déclaré la tuberculose une urgence globale.

La plupart des médicaments antituberculeux utilisés actuellement ont été élaborés dans les années 50. Ils proviennent de criblages de banques de composés chimiques ou résultent d'initiatives fortuites. Parmi eux, la streptomycine découverte en 1943 et produite par *Streptomyces griseus* a été le premier antibiotique antituberculeux. L'observation, faite par hasard en 1945 par Chorine (1) du fait que la nicotine inhibe le développement des mycobactéries chez la souris a conduit à la synthèse puis à l'expérimentation de composés apparentés qui ont abouti à l'élaboration de l'isoniazide, un puissant médicament (2). Par la suite, d'autres médicaments inhibant des mécanismes particuliers de *M. tuberculosis* tel que la biogenèse de la paroi cellulaire caractéristique, ont été mis au point (3).

Pour le développement de nouveaux antibiotiques, l'idéal serait bien évidemment que leur activité bactéricide ait pour cible des fonctions spécifiques des bactéries essentielles à leur survie et non retrouvées chez l'hôte pour éviter tout effet toxique. Pour l'élaboration de tels médicaments,

différentes approches permettent de déterminer quels sont les gènes essentiels de *M. tuberculosis*. Parmi elles, on retrouve les échanges alléliques où les gènes sont inactivés par haploïdie ou diploïdie partielle de l'hôte (4), ou la mutagenèse dirigée (5-7). Néanmoins, la croissance lente des bacilles tuberculeux et la nécessité de travailler dans des laboratoires de haute sécurité (L3) en raison du risque d'infection, rendent les manipulations génétiques de *M. tuberculosis* très longues et difficiles. Une autre alternative pour identifier les gènes supposés essentiels de *M. tuberculosis* est la génomique comparative et fonctionnelle.

### LE GÉNOME DE *M. TUBERCULOSIS*

Le séquençage et l'analyse systématique du génome de la souche H37Rv de *M. tuberculosis* ont été entrepris en collaboration entre l'Institut Pasteur de Paris et le Sanger Centre de Hinxton, Cambridge, en Grande Bretagne. La construction d'une banque ordonnée de BACs (Bacterial Artificial Chromosome) contenant des larges inserts d'ADN de *M. tuberculosis* H37Rv (8) a constitué une étape décisive dans l'accomplissement de ce projet. Ce système BAC a en fait permis de maintenir et de répliquer des fragments d'ADN mycobactérien d'une taille allant jusqu'à 130 kb dans *Escherichia coli*. Le séquençage du génome a révélé que le chromosome de *M. tuberculosis* était circulaire et qu'il contenait 4411 529 paires de bases avec un pourcentage en G+C d'environ 65,6%. De plus, l'analyse bio-informatique a permis de prédire que 3924 gènes codaient des protéines et a mené à l'identification précise des fonctions d'en moyenne 40% des gènes et à certaines connaissances fonctionnelles en ce qui concerne 20% d'autres gènes, tandis qu'aucune information n'a pu être obtenue pour les 40% restant (9).

L'analyse du génome de *M. tuberculosis* a également permis d'identifier plusieurs nouvelles familles de gènes précédemment inconnues, par exemple, les gènes codant les protéines PE et PPE qui occupent presque 10% du génome. Ces protéines sont caractérisées par des motifs de proline-acide glutamique ou proline-proline-acide glutamique typiques, situés dans la partie aminoterminal et par des régions centrales et carboxyterminales répétitives, très riches en glycine pour les PE et asparagine pour les PPE. La fonction de ces protéines est actuellement inconnue, mais leur abondance suggère qu'elles ont un rôle important dans la biologie de *M. tuberculosis* (9, 10). Les protéines PE et PPE sont ainsi

• Travail de l'Unité de Génétique Moléculaire Bactérienne (C.D., R.B., M.M., S.T.C., MD), Institut Pasteur, Paris, France.

• Correspondance : C. DEMANGEL, l'Unité de Génétique Moléculaire Bactérienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France • Page web: <http://www.pasteur.fr/recherche/units/Lgmb/>

• E-mail : demangel@pasteur.fr •

désormais considérées comme des cibles potentielles très intéressantes pour le développement de nouvelles thérapies.

L'analyse du génome a aussi indiqué que 14 gènes *esx* codant pour les protéines de la famille ESAT-6 sont présents dans 11 régions différentes (9, 11). ESAT-6, une petite protéine codée par le gène *esxA* de cette famille, est un antigène protéique apparemment sécrété indépendamment de la voie habituelle de sécrétion faisant intervenir un peptide signal. Cet antigène est fortement reconnu par les lymphocytes T humains et provoque une production massive d'IFN- $\gamma$  (12). Des études récentes ont démontré que certains membres de la famille ESAT-6 seraient impliqués dans le pouvoir pathogène de *M. tuberculosis* (13). Dès lors, cette famille de protéine est considérée très intéressante d'un point de vue préventif (antigène protecteur) et thérapeutique (cible de médicament).

Une autre famille de protéines mycobactériennes, cette fois-ci membranaires (MmpL) d'une taille d'approximativement de 1000 acides aminés et comportant au total 13 membres de type RND (résistance-nodulation-division) a également été identifiée par la génomique. Ces protéines sont caractérisées par des segments transmembranaires typiques et sont impliquées dans le transport de différents substrats comme des lipides à travers la cytomembrane. Cette fonction de transport jouant un rôle crucial dans la virulence de *M. tuberculosis* (6, 7) et pouvant peut-être être impliquée dans la résistance contre certains antibiotiques, pourrait ouvrir de nouvelles voies pour la réalisation de nouveaux médicaments. Ces trois exemples montrent comment la génomique permet de sélectionner des gènes pour des études plus extensives dans un but thérapeutique.

Pour mieux servir les besoins des équipes scientifiques travaillant sur la tuberculose un serveur web interactif a été installé à l'Institut Pasteur (<http://genolist.pasteur.fr/Tuberculist/>); il permet libre l'accès à la séquence complète de *M. tuberculosis* H37Rv, aux gènes et protéines prédits et à leur références bibliographiques.

## GÉNOMIQUE COMPARATIVE DU COMPLEXE *M. TUBERCULOSIS*

*M. tuberculosis*, l'agent de la tuberculose humaine, partage plus de 99,9 % d'identité de son ADN avec les autres membres du complexe de bacille tuberculeux: *Mycobacterium bovis*, l'agent de la tuberculose bovine; *M. bovis* BCG, la souche vaccinale dérivée de *M. bovis*; *Mycobacterium africanum*, un pathogène de l'homme d'origine africaine, *Mycobacterium canetti*, un bacille tuberculeux atypique chez l'homme, et *Mycobacterium microti*, responsable de la tuberculose chez certains rongeurs. Comme le BCG (bacille de Calmette et Guérin), la plupart des souches de *M. microti* sont inoffensives pour l'homme. Ainsi, *M. microti* a été utilisée comme vaccin dans les années soixante en Tchécoslovaquie (14) et a fait l'objet de larges essais cliniques au Royaume-Uni (15). Environ un demi-million d'enfants a été vacciné avec cette souche qui confère une protection et une sécurité équivalente au BCG. Bien que très apparentés, les membres du complexe de bacille tuberculeux

peuvent alors être différenciés sur la base de leur spectre d'hôte, de leurs caractéristiques physiologiques et de leur virulence pour l'homme.

Des études de génomique comparative employant différentes techniques d'hybridation ont montré que plusieurs régions de différence (RD1-RD14), codant environ 140 protéines, sont absentes chez *M. bovis* BCG, la souche vaccinale, par rapport à la souche virulente *M. tuberculosis* H37Rv (8, 16-18). De plus, une étude de la présence ou absence de ces régions chez un plus grand nombre de souches du complexe *M. tuberculosis* a montré que certains d'entre elles étaient également absentes d'autres membres du complexe. Cette étude a permis de définir des relations phylogénétiques entre les différents membres du complexe et de proposer un nouveau schéma de l'évolution des bacilles tuberculeux, remettant en question l'hypothèse actuelle selon laquelle *M. bovis* serait l'ancêtre de *M. tuberculosis* (19). Ces résultats permettent également d'étudier plus profondément les différences génétiques potentiellement impliquées dans le spectre d'hôte des différents membres du complexe. Par exemple, la région RD1 est la seule région à être absente dans les souches vaccinales *M. bovis* BCG et *M. microti* (16, 19, 20), présente chez tous les autres membres du complexe. Il a été confirmé très récemment que les gènes de cette région pourraient être impliqués dans la virulence de *M. tuberculosis*. En effet, lors de la réintroduction de la région RD1 par complémentation dans le BCG et dans *M. microti*, la virulence de ces deux souches vaccinales est partiellement restaurée chez la souris immuno-déprimée. Par contre, la réintroduction de cinq autres régions de différence (RD3, RD4, RD5, RD7, RD9) suspectées impliquées dans la virulence ne semble pas avoir d'effet sur le pouvoir pathogène de ces deux souches (13). Cet exemple montre comment la génomique comparative peut aider à trouver des régions chromosomiques qui sont impliquées dans la virulence de *M. tuberculosis*. Comme RD1 comprend aussi le gène codant l'antigène protéique ESAT-6, il est clair que toutes les souches vaccinales employées à grande échelle dans l'histoire de la vaccination contre la tuberculose n'expriment pas cet antigène immunodominant. L'expression d'ESAT-6 dans le contexte d'un BCG pourrait en augmenter l'efficacité. Cette information est d'une importance majeure pour la mise au point d'un vaccin plus efficace contre toutes les formes de tuberculose.

## GÉNOMIQUE COMPARATIVE MYCOBACTÉRIENNE

Depuis la publication de la séquence du génome de *M. tuberculosis* H37Rv (9) deux autres génomes mycobactériens ont été décrits (21, 22), et la détermination de huit autres mycobactériens est achevée, ou le sera dans un avenir très proche. L'ensemble de ces informations est disponible et régulièrement actualisée sur le site web de l'Institut Pasteur (<http://www.pasteur.fr/recherche/unites/Lgmb/mycogenomics.html>). Grâce à tous ces projets, les mycobactéries se trouvent parmi les bactéries les mieux

caractérisées sur le plan génétique. L'ensemble de ces informations est donc disponible pour identifier de nouvelles cibles potentielles de médicaments antituberculeux. Grâce au jeu de données fourni par les diverses séquences, il est possible de distinguer les gènes spécifiquement mycobactériens de ceux qui sont limités à une espèce mycobactérienne donnée. De plus, les séquences complètes ou partielles des génomes d'environ 200 bactéries et organismes eucaryotes sont également disponibles (par exemple par le site web d'Integrated Genomics Inc. Chicago, <http://wit.integratedgenomics.com/GOLD/>), pour réaliser des recherches de similarité de séquence. Ainsi, un certain nombre de gènes dont la fonction était inconnue au départ ont été retrouvés dans plusieurs bactéries par analogie de séquence et ont été référencés comme «gènes hypothétiques conservés». Il a été démontré que certains d'entre eux jouent des rôles très importants sur le plan biologique et représentent donc des cibles de choix pour la mise au point de nouveaux antibiotiques à large spectre. Ces comparaisons de séquences *in silico* sont importantes pour évaluer si les nouveaux agents thérapeutiques potentiels inhibent bien des fonctions essentielles spécifiquement bactériennes, ce qui limite ainsi l'apparition d'effets secondaires chez l'homme. Il est désormais possible d'effectuer un criblage *in silico* de la séquence génomique humaine afin de s'assurer qu'il n'existe aucune protéine ou gène apparenté chez l'hôte. Des criblages similaires dans les séquences génomiques d'autres pathogènes sont également réalisables pour évaluer le spectre, voire la spécificité de la nouvelle cible. Pour faciliter cette approche des logiciels comme le Diff-Tool ont été développés.

En conclusion, la génomique générale, la génomique comparative et fonctionnelle mycobactérienne ont apporté une quantité considérable d'informations nouvelles qui aide sans aucun doute le développement de nouveaux agents antituberculeux et préventifs.

## RÉFÉRENCES

- 1 - CHORINE V - Action de l'amide nicotinique sur les bacilles du genre *Mycobacterium*. *Compt Rendu Acad Sci* 1945; **220** : 150-152.
- 2 - Fox HH - Synthetic tuberculostatics show promise. *Chem Eng News* 1951; **29** : 3963-3964.
- 3 - PYM AS, COLE ST - Tuberculosis chemotherapy, from conception to genomics. In «WAX R, LEWIS K, SALYERS A, TABER H - Bacterial resistance to antimicrobials : mechanisms, genetics, medical practice and public health». Marcel Dekker ed, New York, 2002, pp 355-403.
- 4 - Pelicic V, Jackson M, Reyrat JM et Coll - Efficient allelic exchange and transposon mutagenesis in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA* 1997; **94** : 10955-10960.
- 5 - BARDAROV S, KRIAKOV J, CARRIERE C et Coll - Conditionally replicating mycobacteriophages: a system for transposon delivery to *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA* 1997; **94** : 10961-10966.
- 6 - CAMACHO LR, ENSERGUEIX D, PEREZ E et Coll - Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. *Mol Microbiol* 1999; **34** : 257-267.
- 7 - COX JS, CHEN B, MCNEIL M, JACOBS WR - Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. *Nature* 1999; **402** : 79-83.
- 8 - BROSCH R, GORDON SV, BILLAULT A et Coll - Use of a *Mycobacterium tuberculosis* H37Rv Bacterial Artificial Chromosome (BAC) library for genome mapping, sequencing and comparative genomics. *Infect Immun* 1998; **66** : 2221-2229.
- 9 - COLE ST, BROSCH R, PARKHILL J et Coll - Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998; **393** : 537-544.
- 10 - BRENNAN MJ, DELOGU G - The PE multi gene family: a «molecular mantra» for mycobacteria. *Trends Microbiol* 2002; **10** : 246-249.
- 11 - TEKAIA F, GORDON SV, GARNIER T et Coll - Analysis of the proteome of *Mycobacterium tuberculosis* *in silico*. *Tubercle and Lung Disease* 1999; **79** : 329-342.
- 12 - SORENSEN AL, NAGAI S, HOUENG G et Coll - Purification and characterization of a low-molecular-mass T-cell antigen secreted by *Mycobacterium tuberculosis*. *Infect Immun* 1995; **63** : 1710-1717.
- 13 - PYM AS, BRODIN P, BROSCH R et Coll - Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. *Mol Microbiol* 2002; **46** : 709-717.
- 14 - SULA L, RADKOVSKY I - Protective effects of *M. microti* vaccine against tuberculosis. *J Hyg Epid Microbiol Immunol* 1976; **20** : 1-6.
- 15 - HART PD, SUTHERLAND I - BCG and vole bacillus vaccines in the prevention of tuberculosis in adolescence and early adult life. *Br Med J* 1977; **2** : 293-295.
- 16 - MAHAIRAS GG, SABO PJ, HICKEY MJ et Coll - Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. *J Bact* 1996; **178** : 1274-1282.
- 17 - GORDON SV, BROSCH R, BILLAULT A et Coll - Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays. *Mol Microbiol* 1999; **32** : 643-656.
- 18 - BEHR MA, WILSON MA, GILL WP et Coll - Comparative genomics of BCG vaccines by whole-genome DNA microarrays. *Science* 1999; **284** : 1520-1523.
- 19 - BROSCH R, GORDON SV, MARMIESSE M et Coll - A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci USA* 2002; **99** : 3684-3689.
- 20 - BRODIN P, EIGLMEIER K, MARMIESSE M et Coll - Bacterial Artificial Chromosome-based comparative genomic analysis identifies *Mycobacterium microti* as a natural ESAT-6 deletion mutant. *Infect Immun* 2002; **70** : 5568-5578.
- 21 - COLE ST, EIGLMEIER K, PARKHILL J et Coll - Massive gene decay in the leprosy bacillus. *Nature* 2001; **409** : 1007-1011.
- 22 - FLEISCHMANN RD, ALLAND D, EISEN JA et Coll - Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J Bacteriol* 2002; **184** : 5479-5490.