

II.2.1.2 LA RADIOGRAPHIE

La radiographie pulmonaire permet d'établir un bilan initial des lésions thoraciques de forme, de siège et de taille variable. L'aspect des lésions est sans rapport avec l'intensité de la maladie. La spécificité de la radiographie pour la tuberculose pulmonaire varie beaucoup selon différentes études (27% à 81%). Il existe quatre types de lésions : le nodule, le tuberculome, l'infiltrat et la caverne (figure 5).

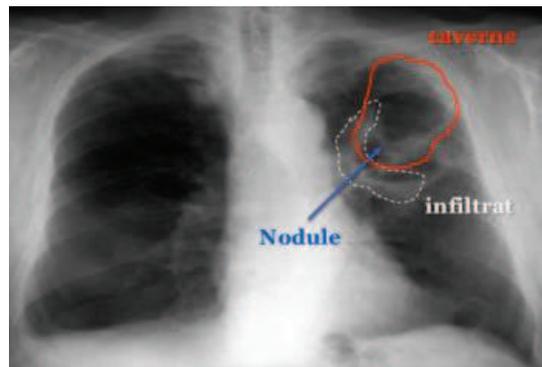


Figure 5: Radiographie d'un tuberculeux pulmonaire

(source : http://www.asnom.org/fr/428_tuberculose.html. Modification sur Adobe Photoshop CS6)

- **L'infiltrat** se traduit comme des lésions débutantes de l'infection. Il se manifeste comme une partie peu opaque, hétérogène et étendue de l'imagerie radiographique.
- **Le nodule** est un granulome de taille variable, isolé ou groupé.
- **Le tuberculome** est un nodule isolé pseudo-tumoral.
- **La caverne** est une région vide due à une perte de substance au sein d'un infiltrat à paroi épaisse.

Dans sa forme pulmonaire, la TB se manifeste par la présence d'infiltrats et de nodules principalement localisés aux sommets des poumons et parfois associés à des cavernes.

II.2.1.3 L'INTRADERMO-REACTION A LA TUBERCULINE

L'intradermo-réaction (IDR) à la tuberculine ou test de Mantoux est un test de diagnostic de la TB basé sur l'observation d'une réaction cutanée d'hypersensibilité retardée après injection intradermique de tuberculine (antigènes mycobactériens composés de mélange de protéines extraites de culture de MTBC et purifiées appelé PPD / « Purified Protein Derivative »). Cette réaction témoigne chez le patient de l'existence d'une immunité à

médiation cellulaire vis-à-vis des mycobactéries, réaction induite soit par une vaccination préalable par le BCG, soit par un contact antérieur avec le BK ou certaines mycobactéries atypiques. Le test IDR est utilisé aussi bien pour le diagnostic de la TB active que pour le diagnostic de la TB latente.

On parle de virage tuberculinique, si sur quelques mois d'intervalle, le résultat de l'IDR initialement négatif (<5 mm) ou faiblement positif (entre 5 et 9 mm) devient positif (>10 mm).

II.2.1.4 AUTRES TESTS IMMUNOLOGIQUES

Des tests immunologiques comme le test QuantiFERON-TB-Gold® et le test TSPOT-TB® ont été développés récemment pour le diagnostic de l'infection tuberculeuse latente (L'Hadj *et al.*, 2006; Lalvani, 2007). Ces tests présentent comme avantage de ne pas être perturbés par une vaccination antérieure au BCG ou par une infection causée par une autre mycobactérie (Al-Orainey, 2009; Bocchino *et al.*, 2009; Richeldi, 2006).

II.2.1.5 LES TESTS MOLECULAIRES

Aujourd'hui, de nombreuses techniques moléculaires adaptées au diagnostic de la TB utilisant l'ADN des mycobactéries sont utilisées. Ces tests sont basés sur l'amplification génique par la classique réaction de PCR (« Polymerase Chain Reaction ») sur des séquences spécifiques des souches *M. tuberculosis* (D'Amato *et al.*, 1995), par la PCR en temps réel comme pour le cas du Xpert MTB/RIF (§ I.6.4.2-c) (Blakemore *et al.*, 2010; Broccolo *et al.*, 2003; Helb *et al.*, 2010), ou aussi par la LAMP-PCR (Loop-mediated isothermal amplification) (Boehme *et al.*, 2007; Cao *et al.*, 2015) permettant la détection de séquences nucléiques spécifiques des bacilles tuberculeux ; cette PCR peut être suivie ou non de l'hybridation des produits d'amplification sur des sondes spécifiques. Cependant, ces techniques présentent, actuellement, une sensibilité insuffisante lorsqu'elles sont appliquées directement sur des échantillons biologiques (Kim *et al.*, 2009). Elles présentent par contre une sensibilité et une spécificité excellentes lors de leur utilisation à partir d'extraits de culture.

II.2.1.6 LA CULTURE

La culture est la méthode de référence pour le diagnostic de la TB. Seule une culture positive à *M. tuberculosis* constitue une preuve formelle du diagnostic d'une TB. Elle est plus sensible que la microscopie. Sa sensibilité varie de 80% à 85% alors que la sensibilité varie de 50% à 80% pour la microscopie. Elle s'effectue sur milieu de culture enrichi, solide (milieu de Löwenstein-Jensen ou milieu de Colestos) où les résultats s'obtiennent entre 4 à 8 semaines ; ou sur milieu liquide (Middlebrook 7H11, 7H10, 7H9, MGIT/Mycobacteria Growth Indicator Tube ou Dubos) où les résultats s'obtiennent plus rapidement (environ 15 jours).

II.3 LES MYCOBACTERIES ET LE COMPLEXE *M. TUBERCULOSIS*

II.3.1 SYSTEMATIQUE

Le genre *Mycobacterium* appartient à la famille des MYCOBACTERIACEAE et à l'ordre des ACTINOMYCETALES (Garrity *et al.*, 2004). Ce genre regroupe actuellement plus de 120 espèces différentes (Shinnick et Good, 1994; Tortoli, 2006), classées en :

- **mycobactéries atypiques** qui regroupent les mycobactéries non tuberculeuses (NTM) connus aussi sous l'appellation de « *Mycobacteria other than tuberculosis* » – MOTT ou encore « Mycobactéries autres que *Mycobacterium tuberculosis* » - MAMT.

- **mycobactéries du complexe *M. tuberculosis* (MTBC)** qui regroupent les agents de la TB humaine ;

- et une distinction particulière faite pour *M. leprae*, et *M. ulcerans*, agents responsables respectivement de la lèpre et de l'ulcère de Buruli.

La systématique des espèces du genre *Mycobacterium* est représentée sur la figure 6.

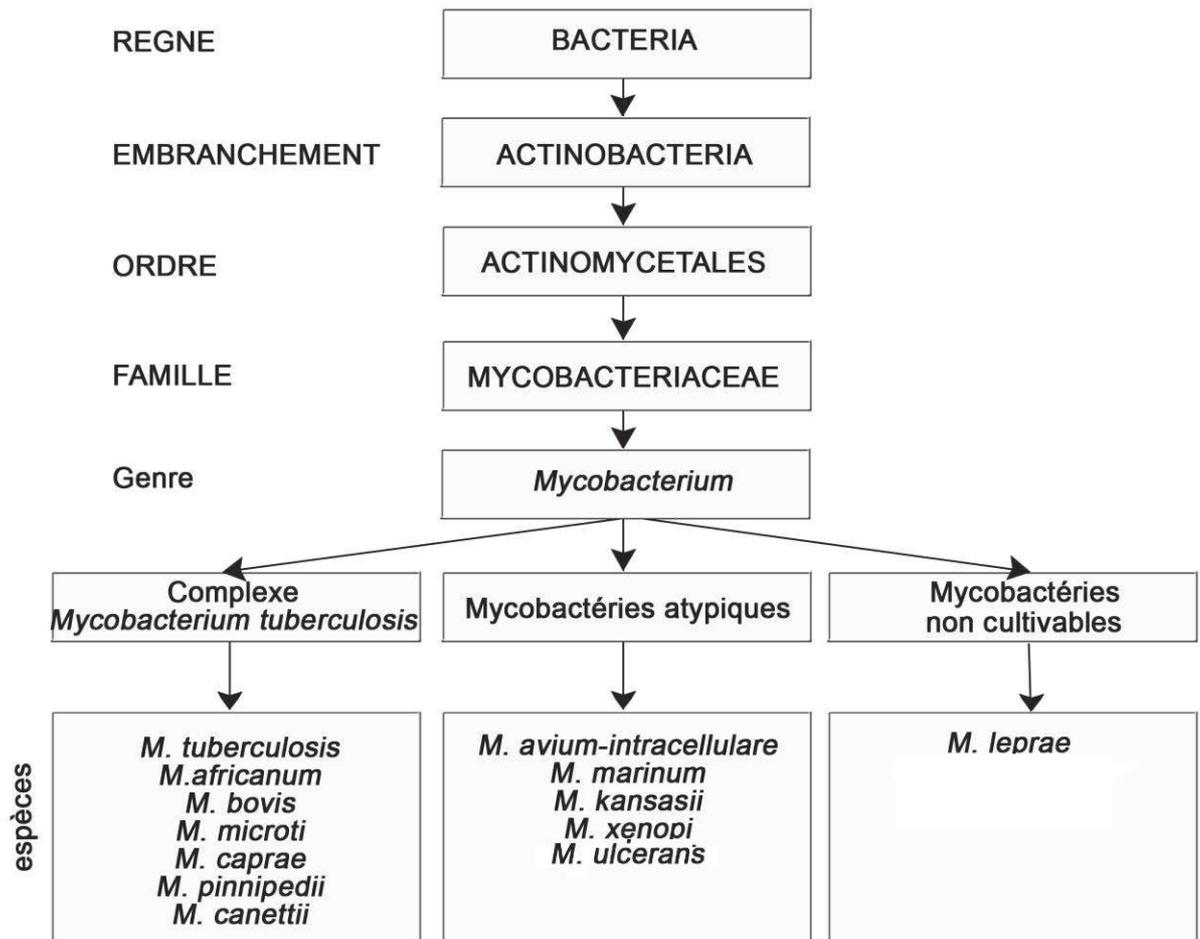


Figure 6: Systématique des espèces du genre *Mycobacterium*

De nouvelles sous-espèces ou de nouvelles espèces sont identifiées régulièrement. Entre 2003 et 2006, plus de 30 nouvelles espèces ont été décrites (Tortoli, 2006). L'arbre phylogénétique de la figure 7 montre la place de MTBC parmi les mycobactéries à croissance lente selon la séquence du gène codant pour l'ARN ribosomal 16S (Tortoli, 2006).

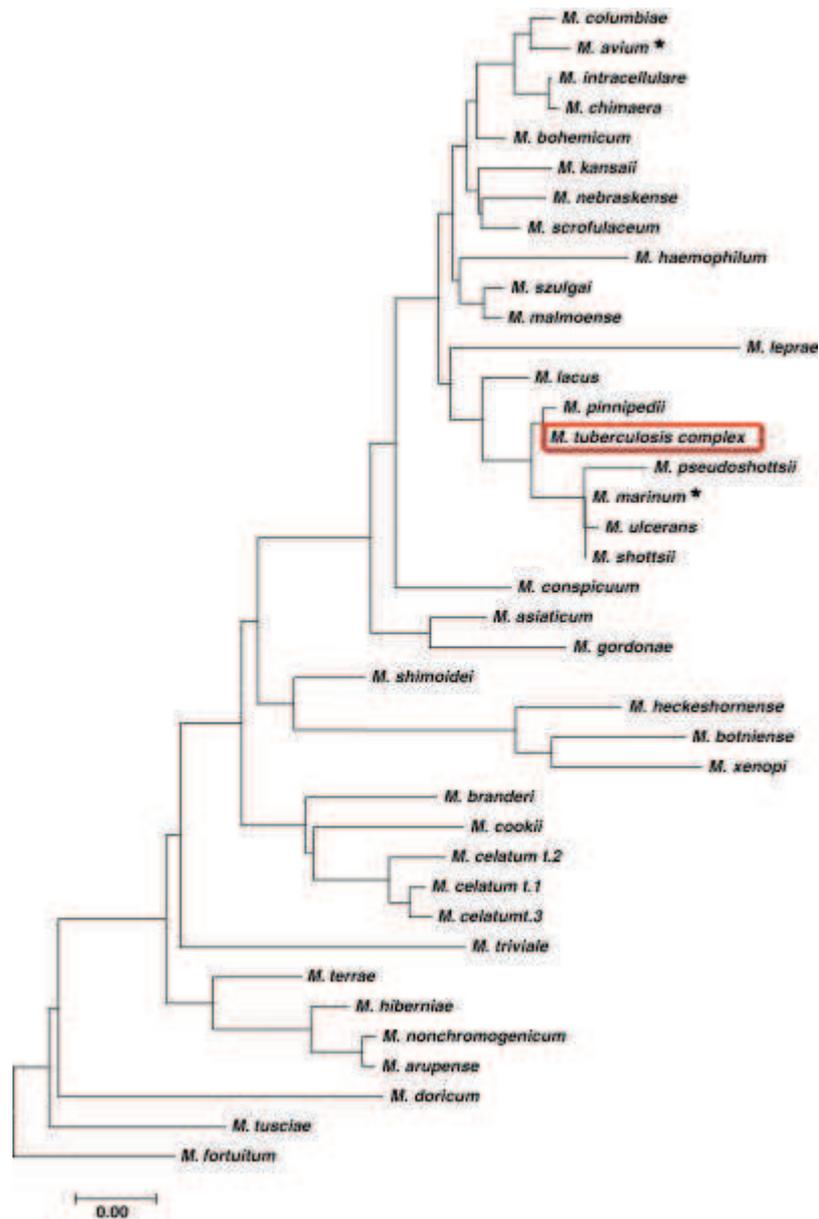


Figure 7: Arbre phylogénétique des mycobactéries à croissance lente selon la séquence codant l'ARNr 16S
M. fortuitum a été utilisé comme « outgroup ».

Les mycobactéries atypiques généralement non pathogènes pour l'homme (MAMT), peuvent être sources d'infections chez les personnes immunodéprimées. Elles se différencient du MTBC par leur moindre virulence et par l'absence de transmission inter-humaine. Ce sont des mycobactéries omniprésentes dans l'environnement. On peut citer parmi ces mycobactéries atypiques : *M. avium-intracellulare*, *M. marinum* (avec un * sur la figure 7) ou encore *M. abscessus* (Bryant *et al.*, 2013).

II.3.2 CARACTERISTIQUES DU GENRE *Mycobacterium*

Les bactéries du genre *Mycobacterium* sont des bactéries aérobies strictes, généralement non sporulantes, non mobiles, et en forme de fin bâtonnet d'environ 4µm de long et 0,4µm de large ou légèrement incurvées et dépourvues de capsule.

Elles se caractérisent par leur acido-alcool-résistance c'est-à-dire leur capacité à résister à la décoloration par les acides et les alcools et leur croissance particulièrement lente avec un temps de dédoublement de 12h à 24h (Harshey et Ramakrishnan, 1977). Ces bactéries se caractérisent également par la composition en acides mycoliques (acides gras à longue chaîne) de leur paroi cellulaire conférant une haute imperméabilité de cette dernière et un faible niveau de fixation de la coloration de Gram qui leur donne la considération par défaut de bactéries Gram positives, ainsi que la résistance naturelle à certains antibiotiques usuels.

La non-sporulation des mycobactéries est admise par la majorité des chercheurs (Traag *et al.*, 2010).

II.3.3 CARACTERISTIQUES DU GENOME DE *M. tuberculosis*

Le génome d'une souche *M. tuberculosis* de référence utilisée jusqu'à nos jours, la souche H37Rv a été séquencé et annoté en 1998 (Cole *et al.*, 1998). Il s'agit d'un génome constitué d'un chromosome unique et circulaire de 4.411.532 paires de bases, comprenant 3995 phases de lecture et formant environ 4000 gènes (figure 8). L'ADN de *M. tuberculosis* présente un taux élevé en guanine et en cytosine (65,6%) sauf dans quelques régions particulières telles que les gènes qui codent pour les protéines transmembranaires. Une grande partie des gènes (6% du génome) semble coder pour des enzymes impliqués dans la synthèse et la dégradation de lipides. Plus de 50% des protéines codées ont des fonctions connues actuellement. L'homologie entre les ADN des différentes sous-espèces du MTBC est très élevée (>99,9%) (Garnier *et al.*, 2003; Smith *et al.*, 2009) ce qui fait qu'il s'agit stricto sensu d'une seule espèce bactérienne. Cette restriction de la diversité nucléotidique peut être expliquée soit par une fidélité de réplication inhabituelle, soit par un système de réparation des erreurs très efficace, soit par une origine évolutive très récente, soit encore par un taux de mutation très faible compte tenu du style de vie de ce bacille (intracellulaire), aucune de ces raisons n'étant exclusive. Le génome est aussi caractérisé par l'utilisation du codon Guanine-Tyrosine-Guanine (GTG) comme codon d'initiation dans plus de 35% des gènes.

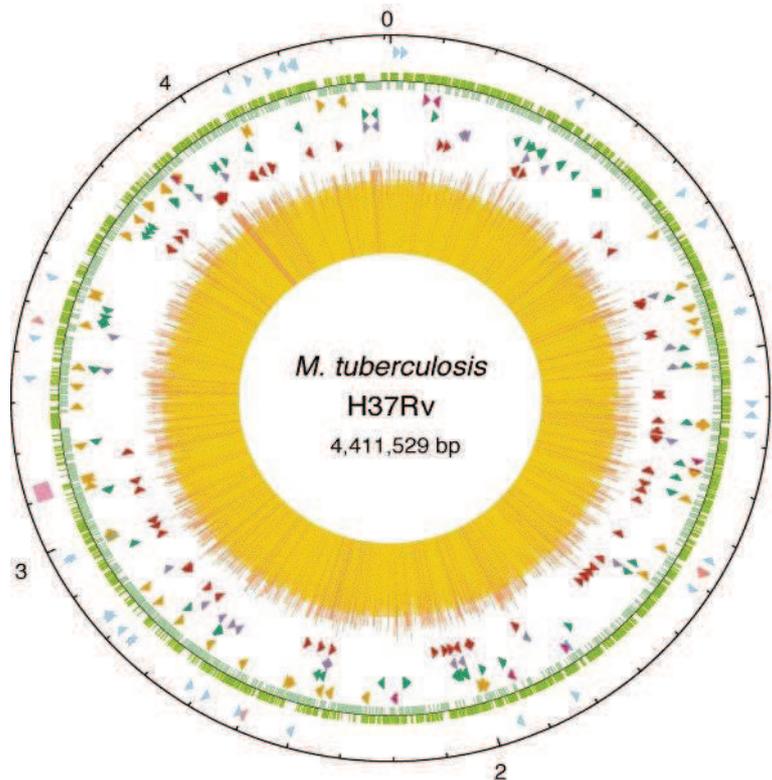


Figure 8: Représentation du génome de *M. tuberculosis* H37Rv

Le cercle extérieur indique l'échelle en megabases, 0 représente l'origine de réplication. Le premier cercle de l'extérieur représente les positions des gènes codant des ARN stables (ARNt en bleu, les autres en rose) et la région « Direct Repeat » (carré rose); le second cercle de l'intérieur montre les séquences codantes par brin (dans le sens horaire, vert foncé; antihoraire, vert clair); le troisième cercle représente les séquences répétées (les séquences d'insertion en orange; la famille REP13E12 en rose foncé; les prophages en bleu); le quatrième cercle montre les positions des membres de la famille des PPE (en vert); le cinquième cercle montre les membres de la famille des PE (en violets, à l'exclusion des PGRS); et le sixième cercle montre les positions des séquences PGRS (en rouge foncé). L'histogramme du centre représente le contenu en G+C, avec <65% G+C en jaune, et > 65% G+C en rouge. La figure a été générée avec le logiciel DNASTAR®. (Cole *et al.*, 1998)

II.3.4 LES MARQUEURS GENETIQUES DE *M. tuberculosis*

Le génome de *M. tuberculosis* est aussi caractérisé par des séquences nucléotidiques particulières polymorphes utilisées comme des marqueurs génétiques d'évolution et/ou d'épidémiologie. Ces marqueurs sont utilisés pour des études d'épidémiologie moléculaire, de classification ou de phylogénie. Parmi ces marqueurs génétiques, on trouve des séquences répétées (les CRISPR, les mini satellites ou les éléments d'insertion), des délétions ou insertions de courts ou de longs fragments d'ADN, des régions PE (Proline – Acide Glutamique) et PPE (Proline – Proline - Acide Glutamique) ou tout simplement des variations nucléotidiques individuelles ou SNPs.

Des techniques de typage spécifiques du MTBC conçues pour la détermination de ces variations au niveau de ces marqueurs ont été développées telles que : la RFLP IS6110 (Van Embden *et al.*, 1993), qui fut longtemps la méthode de référence, le spoligotypage (Kamerbeek *et al.*, 1997; Zhang *et al.*, 2010), les MIRU-VNTR (« Mycobacterial Interspersed Repetitive Units – Variable Number of Tandem Repeat ») (Supply *et al.*, 2000; Supply *et al.*, 2006), et depuis peu la génomique comparative par séquençage de nouvelle génération (« Whole Genome Sequencing » - WGS / « Next Generation Sequencing » - NGS : séquençage du génome de nouvelle génération). Ces techniques vont être décrites dans le paragraphe § 1.3.5. à la page 21.

II.3.4.1 Les éléments d'insertion (IS)

Le génome des mycobactéries présente 29 différentes séquences d'insertion transposables de taille de l'ordre de 1,4-2,5kb pouvant s'insérer dans de multiples sites du génome (Gordon *et al.*, 1999). Des régions « Hotspot » d'intégration de ces IS sont retrouvées au niveau d'autres familles de séquences telles que la région « Direct Repeat » (DR), les PE/PPE (Proline-Acide Glutamique / Proline-Proline-Acide Glutamique) ou des gènes codant pour des antigènes potentiels (Cole *et al.*, 1998; Warren *et al.*, 2000). L'IS le plus étudié chez les MTBC est l'IS6110 de taille égale à 1361 paires de bases (Thierry *et al.*, 1990). Le polymorphisme du nombre et de la position des IS6110 dans le génome des souches du MTBC a été utilisé pour leur génotypage (Van Embden *et al.*, 1993). Les IS6110 sont conservées et très répandues dans les bactéries MTBC. Le génome de H37rv possède 16 copies IS6110 réparties de manière inéquitable sur tout le génome (figure 9) (Cole *et al.*, 1998). Les rôles biologiques des IS ou la conséquence de leur insertion dans le génome de *M. tuberculosis* ne sont pas encore bien connus même si l'insertion de IS6610 peut avoir des conséquences sur l'expression de certains gènes (Soto *et al.*, 2004). Une compensation de la perte de fitness a été observée chez une souche *M. bovis* provenant d'Espagne après insertion d'un IS6110 en amont du locus *phoPR* (Gonzalo-Asensio *et al.*, 2014).

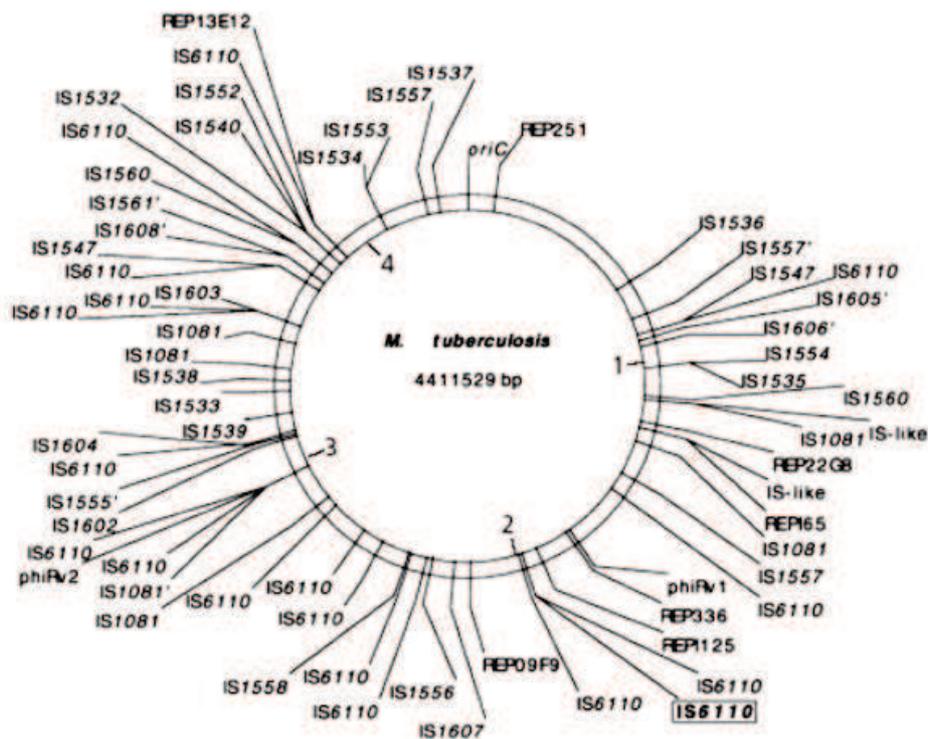


Figure 9: Distribution des 29 IS, des 2 prophages PhiRv1 et PhiRv2 et des 7 éléments REP13E12 sur le génome de la souche H37rv

oriC : origine de réplication. Echelle en mégabase à l'intérieur du cercle (Gordon *et al.*, 1999)

II.3.4.2 Les motifs PE et PPE

Les séquences nucléotidiques PE/PPE sont des régions du génome des mycobactéries contenant des gènes codant pour des protéines acides particulièrement riches en Proline et Acide Glutamique (Cole *et al.*, 1998). Ces protéines sont caractérisées par la présence de motifs « Proline-Acide Glutamique » (PE) ou « Proline-Proline-Acide Glutamique » (PPE) au niveau de leur extrémité N-terminale.

Les régions PE/PPE représentent environ 10% du génome de la souche H37Rv. Les gènes de ces séquences PE/PPE sont regroupés en clusters et répartis en multiples copies polymorphes sur tout le génome. Dans le génome de H37Rv, une centaine de régions PE/PPE ont été identifiées. Celles-ci ont été classées en 3 sous-familles, selon la nature et la présence ou non de motifs répétés associés aux PE/PPE : les PE-PGRS, les PPE-MPTR et les PE ou PPE.

La sous-famille des PGRS ou « polymorphic GC-rich sequences » particulièrement riche en G+C sont organisées en microsatellites dont les motifs varient d'une souche de

mycobactérie à une autre. La classe PE_PGRS comporte 61 membres. Elle est caractérisée par des motifs répétés en tandem, c'est-à-dire en unité monomérique associées aux PE ou aux PPE. Les PE_PGRS sont retrouvées exclusivement dans les bactéries MTBC.

Les fonctions de tous les gènes PE/PPE n'ont pas été précisées, il semble qu'ils soient impliqués dans des facteurs de virulence de la bactérie tels que la persistance des mycobactéries (Camacho *et al.*, 1999; Talarico *et al.*, 2005).

II.3.4.3 Les prophages PhiRv1 et PhiRv2

Deux régions du génome de H37Rv possèdent des organisations similaires avec des gènes phagiques adjacents situés dans un même locus et ont été considérées comme des prophages dénommés PhiRv1 et PhiRv2 respectivement (Cole *et al.*, 1998). La présence et/ou la localisation des gènes de PhiRv1 semble être assez variable d'une souche à une autre. Une étude comparative des régions de différences entre quelques mycobactéries a en effet montré que le prophage PhiRv1 est absent dans 16% des souches cliniques étudiées et aussi absent chez la souche vaccinale BCG (Mahairas *et al.*, 1996). Les gènes du prophage PhiRv2 sont généralement toujours présents et sont relativement plus stables. La position des 2 prophages dans le génome de H37rv est représentée dans la figure 16. Le prophage PhiRv1 possède 7 sites d'attachements potentiels (ATT) dans le génome de H37Rv (Bibb et Hatfull, 2002; Bibb *et al.*, 2005; Hendrix *et al.*, 1999).

Le prophage PhiRv1 est par ailleurs flanqué de séquences REP13E12 (§ 1.3.4.4-a). Il a été suggéré que l'opéron biotine de la souche H37Rv a d'abord été interrompu par des séquences REP13E12 suite à un stress. Cette séquence REP13E12 aurait ensuite servi de site d'insertion du prophage PhiRv1 dans le génome de H37Rv (Gordon *et al.*, 1999).

Le rôle biologique de ces prophages dans les bactéries MTBC n'est pas encore clairement établi. Leur présence semble cependant indiquer leur contribution dans des transferts horizontaux impliqués dans la dynamique et la structure du génome des bactéries MTBC.

II.3.4.4 Les séquences répétées

a. **Les éléments REP13E12**

Il s'agit de séquences répétées, de la famille des séquences entrecoupées ou «Interspersed sequences», de 1,3 kb à 1,5 kb. Leur dénomination vient des répétitions (REP) mises en évidence dans le cosmide MTCY13E12. On retrouve 7 éléments REP13E12 dans le génome de la souche H37Rv. Les protéines REP13E12 n'ont aucune homologie de séquences avec des protéines connues. Ces REP13E12 semblent être uniquement observées dans les bactéries MTBC. Le rôle des REP13E12 n'est pas encore connu. Leur expression serait induite en réponse à une altération de l'ADN. Dans le génome de H37Rv, 4 régions REP13E12 possèdent des sites d'attachement potentiels pour le prophage PhiRv1 (Bibb *et al.*, 2005; Cole *et al.*, 1998; Gordon *et al.*, 1999).

b. Le locus DR

Le locus DR (« Direct Repeat ») fait partie des séquences de la famille des CRISPR (« Clustered Regularly Interspersed Short Palindromic Repeats ») (Marraffini et Sontheimer, 2010a). Le locus DR chez *M. tuberculosis* est formé par des séquences nucléotidiques identiques répétées (les DR) de 36 pb entrecoupées par de courtes séquences uniques, non variables, de 35 à 41 pb appelées espaceurs. La structure de la région DR qui est unique chez *M. tuberculosis* est montrée sur la figure 10.

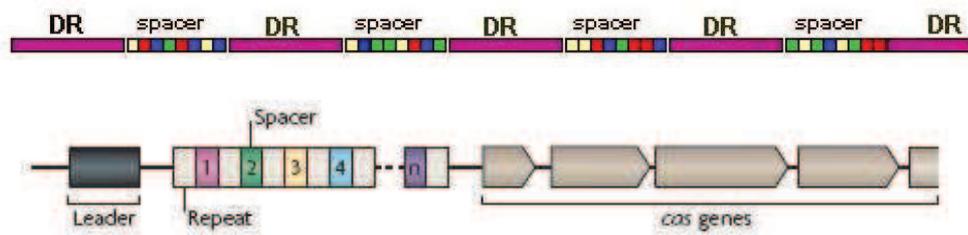


Figure 10: Structure de la région DR (Marraffini et Sontheimer, 2010b)

L'ordre des espaceurs est identique chez toutes les souches mais leur présence/absence varie d'un isolat à l'autre. Quarante-quatre espaceurs ont été retrouvés jusqu'à maintenant dans toutes les souches du MTBC dont certains espaceurs spécifiques de certaines espèces (tableau 1) (Caimi *et al.*, 2001; Van der Zanden *et al.*, 2002).