



HAL
open science

Développement de méthodes pour le diagnostic, le contrôle, la surveillance de la tuberculose à bacilles ultra-résistants et des souches épidémiques Beijing

Jésutondin Bernice Méline Klotoe

► To cite this version:

Jésutondin Bernice Méline Klotoe. Développement de méthodes pour le diagnostic, le contrôle, la surveillance de la tuberculose à bacilles ultra-résistants et des souches épidémiques Beijing. Santé publique et épidémiologie. Université Paris Saclay (COMUE), 2018. Français. NNT : 2018SACLS379 . tel-01940509

HAL Id: tel-01940509

<https://theses.hal.science/tel-01940509>

Submitted on 30 Nov 2018

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.

DEVELOPPEMENT DE METHODES POUR LE DIAGNOSTIC, LE CONTRÔLE, LA SURVEILLANCE DE LA TUBERCULOSE A BACILLES ULTRA-RESISTANTES ET DES SOUCHES EPIDEMIQUES BEIJING

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°577: Structure et Dynamique des Systèmes
Vivants (SDSV)
Spécialité de doctorat: Sciences de la vie et de la Santé

Thèse présentée et soutenue à Orsay, le 18/10/2018 par

Jésutondin Bernice Méline KLOTUE

Composition du Jury :

Pr Stéphanie Bury-Moné

Professeur de l' Université Paris-Sud, I2BC

Président

Dr Florence Brossier

Bactériologiste, Universitaire Pitié-Salpêtrière /

Rapporteur

Faculté de médecine Pitié-Salpêtrière-UPMC Paris6

Dr Roland Brosch

Chercheur, Institut Pasteur

Rapporteur

(Pathogénomique Mycobactérienne Intégrée)

Dr Mona Dumitrescu

Maître de Conférences des Universités - Praticien Hospitalier

Examineur

UFR de Médecine Lyon Sud

Pr Christophe Sola

Professeur de l' Université Paris – Sud, I2BC, UMR9198

Directeur de Thèse

Dr Guislaine Refrégier

Maitre de conférence / Université Paris – Sud, I2BC

Co-Directrice de Thèse

UMR9198

Remerciements

Du fond de mon coeur, je remercie mon Doux **Jésus-Christ** pour Sa présence à mes cotés. A mon PAPA d'AMOUR MON **DIEU TOUT PUISSANT**. A TOI, l'HONNEUR ET TOUTE LA GLOIRE.

C'est la fin de mon cycle universitaire, je tiens à exprimer toute ma gratitude et ma reconnaissance à tous ceux qui ont participé d'une manière ou d'une autre, de près ou de loin à ma formation tout au long de mon cursus universitaire et dans le déroulement du projet de thèse.

Je remercie particulièrement ma co-Directrice de thèse Madame **Guislain Refregier** pour l'encadrement Scientifique, son dévouement et sa rigueur dans le travail, sa patience, sa disponibilité et sa promptitude à répondre aux questions. Merci pour tout ce que j'ai appris de toi aussi bien scientifiquement qu'humainement, je n'oublie pas bien sur « Excel ». Je remercie mon directeur de thèse, le professeur **Christophe Sola**, de m'avoir accepté dans le laboratoire IGEPE et confier des projets de responsabilités dont la thèse. Merci également pour tes recherches de financements. Merci pour tout ce que j'ai appris durant ces années sur les mycobactéries. Merci à **Egor Shitikov** pour ses explications sur l'analyse des données génomiques des souches Beijing.

Je tiens à remercier l'ensemble des membres du jury d'avoir accepté de prendre le temps, malgré leurs occupations scientifiques et professionnelles de participer à l'examen de mon travail de thèse. Madame **Florence Brossier** et le chercheur **Roland Brosch** d'avoir accepté d'être rapporteurs, Madame **Stéphanie Bury-Moné**, la présidente de jury et Madame **Mona Dumitrescu**, examinatrice.

Merci à mes ex-encadrants Russes, précisément dans le laboratoire de Bio-Ingénierie dirigée par le Professeur **Margarita Scharipova**, où j'ai fait mes premiers pas en Biologie moléculaire sous la responsabilité du Dr **Andrei Cheremin**.

Bien sur le cursus universitaire n'est pas réalisable sans ressources financières. Je remercie mon papa **Hounyevou KLOTUE AGOSSOU** pour ses investissements. J'adresse mes remerciements à l'endroit des Institutions, Entreprises, Organisations et leurs personnels qui par leurs financements ont permis la réalisation du projet de thèse. Je cite:

- Le gourvenement **Béinois** pour l'octroi de la Bourse **Bénin-Russie (2007-2013)**, en vue de l'obtention du diplôme d'Ingénieur-spécialiste en Microbiologie.
- La bourse IDEX-Paris Saclay, France (2013 – 2014) pour l'obtention du Diplôme de Master Professionnel 2, option « Microbiologie appliquée et Génie Biologique ».
- Le Centre National de Recherche Scientifique (CNRS), pour sa collaboration dans les négociations entreprise-institution.
- La start-up Beamedex, **Clément Rippol** et **Gomgnimbou Michel** pour leur implication dans les négociations, et l'Agence Nationale de Recherche et de la Technologique qui a co-financé le projet de thèse.

Je remercie très chaleureusement mon frère **KLOTUE Cédric** pour son soutien moral, ses prières et ses encouragements. Merci à **papa**, mes frères **Donald, Cédric, Brice, Tessorel, Marc-Seth**, mes cousins **Waidou, Abbas** et à ma chère grand-mère **Dossou Josephine**.

Table des matières

Listes des abréviations	5
Liste des Figures et Tables	7
Avant –propos	9
Chapitre 1 : Introduction générale	12
1.1 La tuberculose	13
1.1.1 Historique	13
1.1.2 La maladie de la tuberculose	14
1.1.2.1 Agent responsable de la tuberculose : <i>Mycobacterium tuberculosis</i>	14
1.1.2.2 Pathogénèse et évolution de la tuberculose	17
1.1.2.2.1 Primo-infection	18
1.1.2.2.2 Tuberculose latente	19
1.1.2.2.3 Tuberculose active	19
1.1.3 Prévention de la tuberculose par la vaccination	20
1.1.4 Diagnostic de la tuberculose	21
1.1.4.1 L'intradermo-réaction (IDR) à la tuberculine - le test Interferon Gamma Release Assay (IGRA)	21
1.1.4.2 L'examen Radiologique	22
1.1.4.3 L'examen bactériologique	23
1.1.4.3.1 L'examen direct : la bacilloscopie	23
1.1.4.3.2 La culture	23
1.1.4.3.3 L'identification	24
1.1.4.3.4 L'antibiogramme	24
1.1.5 Traitement de la tuberculose	25
1.1.5.1 Traitement général	26
1.1.5.2 Enjeux du traitement de la tuberculose sensible aux antibiotiques	30
1.1.5.3 Traitement de la tuberculose mono/poly/multi/ultra résistante aux antibiotiques	30
1.1.5.3.1 Résistance de la tuberculose aux antibiotiques	31
1.1.5.3.2 L'Ultra-Résistance (UR) de la tuberculose aux antibiotiques	32
1.1.5.3.3 Résistances aux autres antibiotiques de seconde ligne	34
1.1.5.3.4 Résistances aux nouveaux antituberculeux	35
1.1.6 Méthodes phénotypiques de détection des résistances aux antibiotiques	35
1.1.6.1 Méthodes phénotypiques de détection de sensibilité/résistance sur milieux solides	36
1.1.6.2 Méthodes phénotypiques de détection de sensibilité/résistance sur milieu liquide	36
1.1.7 Méthodes de biologie moléculaire pour la détection des résistances aux antibiotiques	37
1.1.8 La tuberculose dans le monde actuel	40
1.1.8.1 Epidémiologie de la tuberculose dans le monde	40
1.1.9 Epidémiologie et évolution de la résistance aux antibiotiques	43
1.1.10 Tuberculose et coinfection VIH	43
1.2 Le complexe <i>Mycobacterium tuberculosis</i> (MTBc)	43
1.2.1 Génome de <i>Mycobacterium tuberculosis</i>	43
1.2.2 Macroévolution de <i>Mycobacterium tuberculosis</i> et transmission	45
1.3 Génotypage de <i>Mycobacterium tuberculosis</i> : Méthodes de génotypage	46
1.3.1 IS6110-RFLP	47
1.3.2 Typage des espaceurs dans le locus CRISPR – Spoligotypage	48
1.3.3 Typage MIRU-VNTR	49
1.3.4 Analyse des régions de délétions	51
1.3.5 Analyse des génomes entiers et Typage à base des polymorphismes de nucléotides SNP	51

1.4 Classification de Mycobacterium tuberculosis	51
1.5 La lignée 2 « Beijing » de Mycobacterium tuberculosis	53
1.5.1 Diversité et importance de la lignée 2 « Beijing »	53
1.5.2 Pathogénicité et virulence des souches L2/Beijing	55
1.5.3 Diversité des souches de la lignée L2/« Beijing » en fonction des diverses méthodes de génotypage	56
1.5.3.1 Diversité des profils CRISPR (spoligotypes) des souches de la lignée L2/Beijing	56
1.5.3.2 Typage MIRU-VNTR	57
1.5.3.3 Les régions de délétions	58
1.5.3.4 Polymorphisme de nucléotides simples SNP	60
1.5.4 La séquence d'insertion IS6110	62
1.5.4.1 Localisation des copies de IS6110 dans les diverses régions du génome	62
1.5.4.2 IS6110 dans le locus NTF	63
1.5.4.3 Souches épidémiques identifiées par IS6110 et leur classification en sous-groupes de la lignée Beijing	64
Chapitre 2 : La technologie Luminex®, Principes et Application	67
Chapitre 3 : But et problématique de la thèse	70
Chapitre 4 : Résultats et Articles	73
Etude 1 : Développement de méthode haut-débit sur microbilles pour la détection des résistances aux antibiotiques de second ligne sur les souches de Mycobacterium tuberculosis	74
Article 1 :	75
Etude 2 : Etude des sites d'insertion de la séquence d'insertion IS6110 au sein du locus NTF dans la famille « Beijing » de Mycobacterium tuberculosis par hybridation multiplexée sur Luminex®	99
Article 2 :	102
Etude 3 : Analyse de données WGS (whole-genome sequencing) et identification de marqueurs potentiels en vue du développement d'une méthode multiplexée et combinée basée sur le polymorphisme SNPs/RD/IS6110 et utilisant la technologie Luminex®, pour la sous-classification et le suivi mondial des souches de bacilles tuberculeux L2/Beijing.	115
Article 3	118
Etude 4 : Etude de la tuberculose Multi (MDR) et extrêmement (XDR) résistante dans un pays à forte prévalence de souches Beijing, le Kazakhstan	137
Article 4 :	140
Etude 5 : Analyse de la diversité des souches MDR de la lignée 2 circulant au Pérou	164
Article 5 :	167
Etude 6 : Analyse de la diversité génétique des souches de MTBc à Beira, Sofala, Mozambique par spoligotyping ; résultats préliminaires sur l'évolution de la tuberculose au Mozambique	177
Article 6 :	180

<i>Etude 7 : Analyse de la diversité des souches de Mycobacterium tuberculosis dans la région de Balimo en Papouasie Nouvelle Guinée par la technique du Spoligotypage</i>	210
Article 7 :	213
<i>Etude 8 : Epidémiologie moléculaire de Mycobacterium tuberculosis en Italie en 2015: état actuel de la tuberculose</i>	216
Article 8 :	218
<i>Chapitre 5 : Discussion Générale et Perspectives</i>	222
5.1 La détection des résistances aux antibiotiques de 2 ^{de} ligne	224
5.2 Avantage et inconvénients du test IS-NTF/RINT et MLPA-Beijing	225
5.3 Marqueurs des souches épidémiques	228
5.4 Intéret des tests développés pour la santé publique	229
<i>Chapitre 6 : Conclusion Générale</i>	231
<i>Chapitre 7 : Matériel et méthodes</i>	233
<i>Bibliographie</i>	243

Listes des abréviations

ADN: acide désoxyribonucléique

ARN: acide ribonucléique

BAAR: Bacille Acido-Alcool-Resistant

BCG : Bacille de Calmette Guérin

BK : Bacille de Koch

BL : Basal Lineage

Bp: paire de bases

CC : Complexe Clonal

CMI: concentration minimale inhibitrice

CFP-10 : Culture Filtrate Protein 10

CRISPR : Clustered Regularly Interspaced Short Palindromic Repeats

DPO: Dual Priming Oligonucleotide

DR : Direct Repeat

DST : Drug Sensibility Test

EMB: éthambutol

ESAT6 : Early Secretory Antigenic Target Protein 6

ETH: éthionamide

INH: Isoniazide

IDR: Intra Dermal Reaction

IFN- γ : Interféron gamma

IGRA : Interféron Gamma Release Assay

IL12 : Interleukine 12

IS : Insertion Sequence (Séquence d'Insertion)

LJ : Lowenstein – Jensen

LSP : Large Sequence Polymorphisms

MDR: Multi Drug Resistance

MR : Multi – Résistance

MGIT : Mycobacteria Growth Indicator Tube

MIRU : Mycobacterium Interspersed Repetitive Units

MLPA : Multiple Ligation Probe Assay

MTBc : Complexe *Mycobacterium tuberculosis*

NAAT : Test basé sur l'Amplification des Acides Nucléiques

NGS : Nest Generation Sequencing
OADC : Oleic Albumin Dextrose Catalase
OMS/WHO: Organisation Mondiale de la Santé/ World Health Organisation
ORF : Open Reading Frame
PAS : Acide para-Amino Salicylique
PCR: réaction de polymérisation en chaine
PE : Proline - Acide Glutamique
PPE : Proline - Proline - Acide Glutamique
PGRS : Polymorphic GC-Rich Sequences
RD : Region de délétion
RFLP : Restriction Fragment Length Polymorphism
RIF: rifampicine
SNP: Polymorphisme de nucléotide simple
STR: streptomycine
TB: Tuberculose
TNF α : Tumor Necrosis Factor alpha
TST : Tuberculin Skin Test
UR : Ultra - Résistance
VIH : Virus de l'Immunodéficience Humaine
VNTR: Variable Number of Tandem Repeats
XDR: Ultra-résistance/extensively Drug Resistance
ZN : Ziehl - Neelsen
 μ L: microlitre

Liste des Figures et Tables

Figure 1 : Phénotype de <i>Mycobacterium tuberculosis</i>	16
Figure 2: Représentation schématique de la paroi cellulaire de <i>Mycobacterium tuberculosis</i>	17
Figure 3 : Formation et progression histologique du granulome lors d'une infection à <i>M. tuberculosis</i>	19
Figure 4 : Coupe TomoDensitoMétrie (TDM), A- Adénopathie médiastinale, B- Complexe gangliopulmonaire calcifié	20
Figure 5 : Stades d'évolution de la tuberculose après infection par <i>Mycobacterium tuberculosis</i>	21
Figure 6 : Image radiographique du thorax d'un tuberculeux (A), image cavitaire apicale droite et des micronodules en coupe tomодensitométrie (TDM) (B), scanner de la bronche de drainage (C)	24
Figure 7 : Représentation schématique du principe du test Hain	39
Figure 8 : Nombre de cas de TB au plan mondial en 2016 sur tous les continents	43
Figure 9 : Carte circulaire du chromosome de <i>M. tuberculosis</i> souche H37Rv	45
Figure 10 : Illustration de l'évolution du bacille <i>Mycobacterium tuberculosis</i>	47
Figure 11 : Marqueurs de génotypage des souches de <i>Mycobacterium tuberculosis</i>	48
Figure 12 : Exemple de résultat de IS6110-RFLP. a, b, c, d, e, f indiquent les fragments de la souche H37Rv porteurs de IS6110 révélés par Southern	49
Figure 13 : Le spoligotypage	50
Figure 14 : Méthode de génotypage MIRU-VNTR	51
Figure 15: Arbre phylogénétique de 89 séquences de gènes provenant de 108 souches MTBc montrant la structure de la population du complexe <i>M. tuberculosis</i> avec les différentes lignées phylogénétiques	53
Figure 16 : Distribution mondiale des souches Beijing en complexe clonaux (CC) et Lignée Basiques (BL)	55
Figure 17 : Classification phylogénétique déduite des SNP de 358 souches de MTBc	60
Figure 18 : Correspondance entre les classifications phylogénétiques de la lignée 2 de MTB	62
Figure 19 : Séquence d'Insertion IS6110	63
Figure 20 : Distribution de la séquence d'insertion IS6110 dans les génomes des souches Beijing	64
Figure 21 : Locus NTF	65

<u>Figure 22</u> : Identification de la souche B0/W148	67
<u>Figure 23</u> : Principe de la technologie Luminex	70
<u>Figure 24</u> : Schéma du principe des amorces DPO	220
<u>Figure 25</u> : Amplicons obtenues par amplification d'IS6110 dans la région NTF	221
<u>Figure 26</u> : Schema de la representation des amorces et sondes pour l'amplification et l'hybridation/detection des sequences d'insertions IS6110 dans le locus NTF utilisant le cytometer en flux du system Luminex.....	222
<u>Figure 27</u> : Schéma du principe de dessin des sondes destinée à la MLPA	223
<u>Tableau 1</u> : Classification des antituberculeux (WHO 2016), modes d'action, gènes et mécanismes moléculaires impliqué dans la résistance	28
<u>Tableau 2</u> : Type de profils de spoligotypes reportés dans la base de donnée SITVIT	58
<u>Tableau 3</u> : SNP caractéristique des souches épidémiques	215
<u>Tableau 4</u> : Séquences des sondes 1 et 2 dessiner, pour la méthode MLPA	225

Avant –propos

La tuberculose est une maladie ancienne, qui est actuellement en réémergence et constitue un problème de santé publique dans le monde. Selon l’OMS, cent vingt trois pays ont reporté des cas d’ultra-résistance aussi dits XDR (eXtensively Drug Resistant, résistants à la plupart des antibiotiques disponibles, voir introduction pour plus de précisions), et la proportion de patients atteints d’une MDR-TB (Multi-Drug Resistant TuBerculosis, résistants aux deux anti-tuberculeux les plus efficaces, voir introduction pour plus de précisions) qui ont développé une XDR-TB est de 6,2%. Les pays qui ont reporté un nombre important de cas de XDR-TB étaient la Chine, la Biélarussie (exURSS), l’Afrique du Sud, l’Ukraine, l’Inde. Cette réémergence est principalement due à la sélection des souches de *Mycobacterium tuberculosis* multi et ultra-résistantes aux antituberculeux, à cause de l’inefficacité, l’utilisation inappropriée, l’interruption prématurée due à la toxicité des médicaments, la durée, le régime du traitement (long avec l’utilisation des antibiotiques en fonction du pays) ainsi que la qualité des médicaments. Pour le contrôle et/ou l’élimination de la tuberculose, il est primordial de mettre au point des mesures de santé publique. Parmi ces mesures, on note la disponibilité et l’accessibilité à des tests de diagnostic rapides, des traitements rapides, appropriés, efficaces, courts et moins toxiques. Il est aussi très important de développer des outils et/ou méthodes de surveillance afin d’éviter les contaminations, surtout dans les milieux pénitentiaires et hospitaliers.

Les souches de la bactérie *Mycobacterium tuberculosis*, responsable de cette infection sont différentes sur le plan génomique, et en fonction de leurs géolocalisations. Mais, de nombreux cas d’épidémies de tuberculose reportées sont dues à la souche de *Mycobacterium tuberculosis* de génotype « Beijing ». En effet, ce type de souche est la cause d’épidémies de tuberculose en Asie Centrale, en Russie, aux Canaries, en Corée du Sud, et dans plusieurs Etats des USA. Les souches du génotype Beijing originaires du sud de la Chine (Guangxi) se sont répandues dans le monde au travers les âges en fonction des mouvements migratoires et récemment en lien avec la révolution industrielle, les échanges commerciaux, les guerres, les flux migratoires. La particularité de ce génotype, selon plusieurs études, est sa capacité à être hypervirulent, son adaptabilité à l’hôte, son fort pouvoir de transmission et avec un taux de mutation légèrement supérieur aux autres génotypes. L’évolution de la souche de génotype Beijing de *Mycobacterium tuberculosis* montre une grande diversité génomique qui a été sous-estimée. L’étude de la diversité génomique des souches Beijing pourrait améliorer l’identification, le contrôle et la surveillance de ces souches afin de maîtriser leur transmission et limiter les cas d’épidémies de tuberculose. Dans ce contexte, ma thèse a porté sur le développement de nouveaux outils de caractérisation des souches Beijing et des résistances aux antibiotiques, et l’application de ces méthodes et d’autres méthodes de référence pour caractériser l’épidémiologie moléculaire de la tuberculose dans quelques pays.

Dans le double objectif de prévenir la multi/ultra résistance et de maîtriser les épidémies de tuberculose à souches Beijing, nos travaux se sont centrés sur la recherche, le développement et

l'application de méthodes moléculaires rapides et multiplexées. Ces méthodes sont conçues pour détecter la tuberculose XDR, la TB-MDR dans des pays à forte prévalence de souches Beijing, la surveillance des souches Beijing MDR épidémiques. Enfin, la sélection des marqueurs moléculaires pertinents en vue de la discrimination des souches Beijing est aussi un moyen efficace pour leur surveillance. Ainsi nous avons premièrement mis au point deux méthodes prêtes à utilisation : le TB-EFI et le TB IS-NTF/RINT.

Le TB-EFI (Article 1) est un test basé sur l'identification des mutations (polymorphismes de nucléotides simples) dans les gènes associés à la résistance des souches de *Mycobacterium tuberculosis* aux antituberculeux de deuxième ligne : les Fluoroquinolones, les Injectables, l'Éthambutol. A l'aide de la technologie Luminex²⁰⁰, le TB-EFI permet de détecter les mutations en position 90, 94 du gène *gyrA*, responsable de la résistance des souches de *Mycobacterium tuberculosis* aux fluoroquinolones, les mutations en positions 1401, 1402, 1484 du gène *rrs* et les mutations -10, -12, -14 du promoteur *eis* responsable des résistances aux antibiotiques Injectables, et les mutations dans le codon 306 du gène *embB* impliqué dans la résistance à l'éthambutol, un antibiotique de première ligne qui est souvent associé aux traitements de deuxième ligne. Le TB-EFI est un test à haut débit et peut être appliqué en santé publique pour la surveillance et l'évolution de la TB-XDR dans des régions critiques.

Le test IS-NTF/RINT (Article 2) est un test développé pour d'une part suivre l'évolution des cas de tuberculose MDR à souches Beijing et d'autre part déterminer le type génomique de souches Beijing prévalent dans une région spécifique. Le test est développé pour typer la séquence d'insertion IS6110 au sein du locus NTF dont la présence est spécifique aux souches Beijing et détecter les mutations responsables de la résistance de ces souches à la Rifampicine et l'Isoniazide (les deux antibiotiques principaux de première ligne). En effet en fonction du nombre de copies de la séquence d'insertion IS6110, les souches sont subdivisées en Beijing anciennes et modernes. Ce test est d'une importance capitale surtout dans l'identification et le contrôle des souches épidémiques, et permet l'identification des souches d'une épidémie historique, la souche W. Mais au sein de ces deux groupes Beijing ancien et moderne, le test IS-NTF/RINT nous permet de statuer sur l'évolution du phénomène de résistance.

Pour une discrimination complète et identification précise des différents sous-groupes de souches Beijing, nous avons procédé à l'étude des SNP extraites de données WGS, en vue d'identifier des SNPs et régions de délétions spécifiques en collaboration avec Shitikov *et al* (2017) qui ont proposé une classification unifiée. Cette dernière classification identifie 11 sous-groupes dont quatre qui représentent des souches Beijing épidémiques prévalentes en Europe/Russie, et en Asie centrale. Sur différents marqueurs retenus, une méthode sera développée en vue d'une prévention mondiale par l'identification rapide et précise des souches Beijing, qui seront dans les années à venir les souches les plus prévalentes dans plusieurs pays du monde (Article 3).

Par ailleurs, les méthodes existantes dans le laboratoire telles que le spoligotypage sur microbille, nous ont permis d'effectuer des études d'épidémiologie moléculaire de la tuberculose au Kazakhstan, en Nouvelle Guinée Papouasie, Italie, au Mozambique, au Péru (Articles 4 à 8).

Les techniques développées dans cette thèse, sont des méthodes haut débit et pourraient contribuer de manière significative au contrôle de la tuberculose XDR dans les zones « hot-spot », et à la surveillance mondiale sur l'évolution des souches Beijing spécialement ces souches MDR épidémiques.

Chapitre 1 : Introduction générale

Introduction générale

1.1 *La tuberculose*

1.1.1 *Historique*

La tuberculose (TB) est une maladie connue de l'humanité depuis l'antiquité. Elle existait à l'époque néolithique comme en atteste la découverte de lésions osseuses typiques de la tuberculose datant de 7000 ans avant Jésus-Christ sur des squelettes humains du site néolithique de Atlit Yam dans l'Est de la Méditerranée (Hershkovitz Israel , Donoghue Helen D. et al. 2008). Elle a été aussi détectée à différentes périodes sur des squelettes et momies d'Allemagne (1400-1800 AD), de la Hongrie (600-1700 AD) (Fletcher Helen A. , Donoghue Helen D. et al. 2003), au Pérou (Salo Wilmar L. , Aufderheide Arthur C. et al. 1994) et sur des momies égyptiennes de l'Égypte pharaonique (Zink, Haars et al. 2001). Hippocrate, un médecin grec attribue à la tuberculose le nom grec « phtisie » qui signifie « consommation » comme étant la maladie mortelle la plus répandue toujours à cette époque. En 1865, Jean Antoine Villemin, un médecin français prouve la transmission inter-animale de la tuberculose et suppose une transmission interhumaine de la tuberculose dans les logements surpeuplés (Daniel 2015). Sur la base de ses recherches, Robert Koch identifie *Mycobacterium tuberculosis* en 1882 comme étant l'agent responsable de la tuberculose d'où le nom « bacille de Koch » (Daniel 2006). Dès 1890, il parvient à isoler du bacille tuberculeux la « tuberculine », une substance envisagée pour le traitement contre la tuberculose. Malheureusement l'utilisation de la tuberculine comme traitement fut un échec. Au 19^{ème} siècle, cette maladie était un véritable fléau, et par manque de traitements médicamenteux la seule chance de guérison pour les tuberculeux était un repos dans des centres spécialisés suivi d'une cure « hygiéno-diététique ». En 1907, Clemens Von Pirquet invente le test de sensibilité à la tuberculine capable d'identifier des individus qui ont des antécédents d'infection tuberculeuse. Ainsi une réaction positive à la tuberculine reflète une « tuberculose latente ». En 1908 Charles Mantoux introduit l'utilisation d'une aiguille et seringue pour injecter la tuberculine par voie intradermique (Daniel 2006). En 1921, Albert Calmette et Camille Guérin ont développé le vaccin BCG ou Bacille de Calmette Guérin contre la tuberculose dont le principe reposait sur l'injection de bacilles tuberculeux vivants mais de virulence atténuée. Ce vaccin a été obtenu par culture du Bacille sur des milieux contenant de la bile de bœuf (Tan Siang Y. and Erika 2012). En 1943-1944, le premier antibiotique actif contre le bacille tuberculeux, la streptomycine, fut découvert par l'équipe de Selman Abraham Waksman, un microbiologiste américain. Il avait démontré que la streptomycine modifiait

l'évolution de la tuberculose (Daniel 2005). Dans les années qui ont suivi d'autres antituberculeux ont été découverts dont l'isoniazide (synthétisé en 1912, mais dont l'efficacité contre la TB a été montrée en 1945), l'acide para-aminosalicylique, le Pyrazinamide (1952), l'Ethambutol (1961) et la Rifampicine (1969).

Avec l'apparition des antituberculeux et l'amélioration des conditions de vie de la population, la tuberculose a connu une forte régression des taux de mortalité et d'incidence. Mais en 1993 l'OMS déclare la tuberculose une urgence mondiale, suite à l'explosion de nouveaux cas dans les pays industrialisés comme dans les pays en voie de développement en lien avec l'épidémie du Virus d'Immunodéficience Humaine (VIH).

1.1.2 *La maladie de la tuberculose*

1.1.2.1 *Agent responsable de la tuberculose : **Mycobacterium tuberculosis***

M. tuberculosis est une bactérie acido-résistante, microaérobie, asporulée et immobile, sensible à la chaleur et aux UV, détruite en 5 min par l'alcool à 70% (Pfyffer, Murray et al. 2007). Les bacilles sont en forme de bâtonnets légèrement incurvés de 0,2 à 0,6 µm de diamètre sur 1,0 à 10 µm de longueur (**Fig. 1B**) ou droits qui peuvent se ramifier. Ces mycobactéries sont des pathogènes intracellulaires cultivables, à aspect rugueux, avec une croissance très lente de 20 – 21 jours (1^{ère} colonie au 7^{ème} jour) par incubation à 37°C sur les milieux solides (**Fig. 1A**) et de 10-15 jours (1^{ère} trouble bactérien au 3^{ème} jours) pour les milieux liquides.

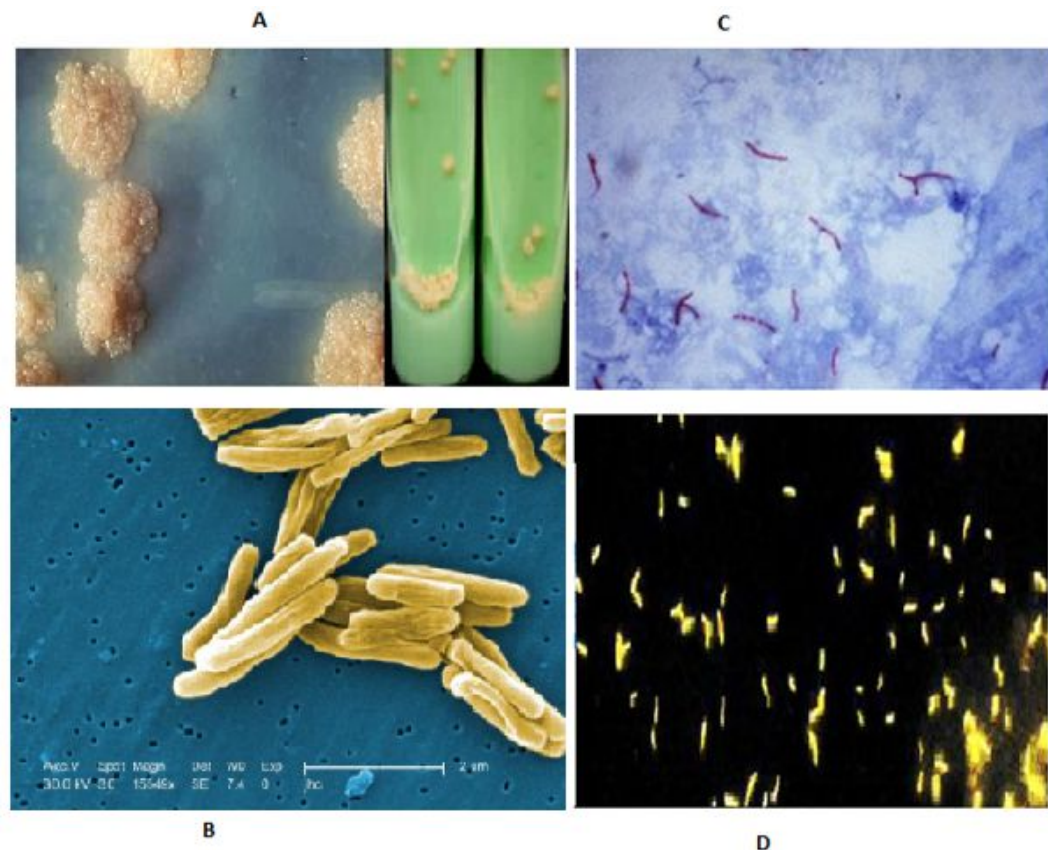


Figure 1 : Phénotype de *Mycobacterium tuberculosis*

A- colonies de *Mycobacterium tuberculosis* sur milieu solide Löwenstein-Jensen (boîtes de Pétri x5 et pente en flacon x1) <http://bacteriologynotes.com/cultural-characteristics-of-mycobacterium-tuberculosis/>, B- Vue au microscope électronique à balayage de *Mycobacterium tuberculosis* <http://textbookofbacteriology.net/tuberculosis.html>, C- Coloration Ziehl-Neelsen de *Mycobacterium tuberculosis* (microscopie optique x1000), D- Coloration à l'auramine de *Mycobacterium tuberculosis* (microscopie à fluorescence x1000) <http://microbia.free.fr/TS2ABM/Mycobacteries/Cours-Mycobacteries.pdf>.

La culture de ces bactéries exige des milieux de culture spéciaux. Sur les milieux solides, seuls les milieux contenant du sérum, de la glycérine, de la pomme de terre glycinée, de l'œuf ou de l'albumine bovine permettent une culture abondante. Ainsi, les milieux de culture les plus utilisés pour la détection de la bactérie sont les milieux solides de Löwenstein-Jensen riches en œuf, les milieux gélosés ou liquides de Middlebrook (7H9, 7H10, 7H11) contenant de l'acide oléique, de l'albumine, du dextrose et de la catalase (OADC). Elles forment des colonies rugueuses de couleur crème (**Fig 1A**).

M. tuberculosis est peu colorée par les colorants ordinaires et à la coloration de Gram à cause de la forte teneur en lipides de sa paroi cellulaire. Elle est donc colorée par deux méthodes : la coloration de Ziehl-Neelsen (**Fig 1C**) et la coloration à l'Auramine (**Fig 1D**) mais la coloration de Ziehl-Neelsen est généralement utilisée. La méthode de Ziehl Neelsen permet de mettre en évidence des bacilles acido-alcool-résistants (B.A.A.R.) après coloration (fuchsine phéniquée à chaud, décoloration par acide-alcool, recoloration par le bleu de méthylène). Les bacilles apparaissent roses sur fond bleu (Pfyffer, Murray et al. 2007). Lors de la coloration à l'Auramine, l'Auramine se fixe sur le bacille et le rend fluorescent. Après traitement à l'acide-l'alcool, les bacilles sont observés au microscope à fluorescence utilisant une diode électroluminescente (LED). Les bacilles apparaissent fluorescents, brillants sur le fond noir de la préparation.

La paroi cellulaire de *Mycobacterium tuberculosis* est très riche en lipides, ce qui la rend particulièrement imperméable à de nombreux agents et lui conférant une résistance aux antibiotiques et agents thérapeutiques. La paroi cellulaire mycobactérienne est constituée d'une couche de peptidoglycane liée de manière covalente à une couche d'arabinogalactane (un polysaccharide) relié par des liaisons esters aux acides gras à longues chaînes ou acides mycoliques (**Fig. 2**).

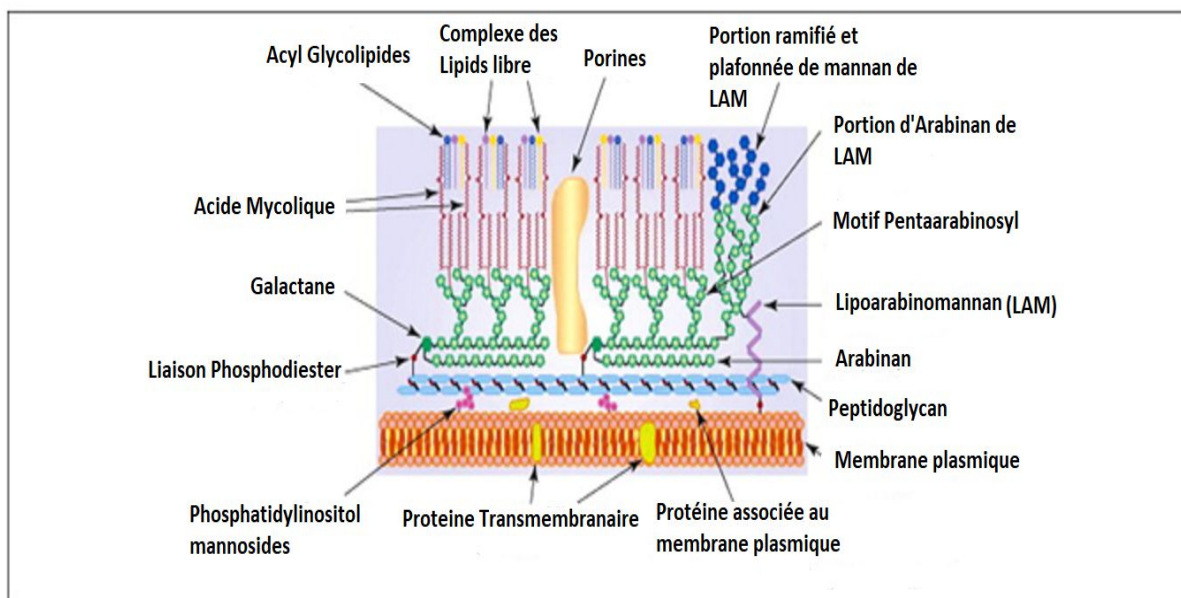


Figure 2 : Représentation schématique de la paroi cellulaire de *Mycobacterium tuberculosis* (Medjahed et al 2009).

1.1.2.2 Pathogénèse et évolution de la tuberculose

La transmission de la tuberculose a lieu d'un malade contagieux à une personne saine par voie aérienne. La bactérie pénètre dans l'organisme par les poumons où elle se développe éventuellement plus tard pour donner forme à la tuberculose pulmonaire, ce qui dépend des conditions de vie de l'hôte, de son âge, de son état de santé (infection à VIH, thérapie immunosuppressive). La contamination se fait par inhalation de sécrétions bronchiques de 1-5 μL projetées dans l'air par un malade contagieux (toux, éternuement ou crachat). Chaque gouttelette peut contenir de 1 à 10 cellules bactériennes (Lee 2016). Le risque de contamination dépend de la concentration des mycobactéries, de la virulence des bacilles, de la durée d'exposition et de l'état du système immunitaire de la personne en contact du bacille. La dose infectante minimale chez l'homme est de 10 bacilles. Une personne infectée (surtout les enfants) a 5-10% de risque en deux ans de développer la maladie contre 10% annuel dans le cas des personnes infectées par le VIH (S., Ramachandran R. et al. 2000; Hiroshi Nakaoka, Lovett Lawson et al. 2006). Le développement de la maladie évolue en trois étapes : la primo-infection, la tuberculose latente et la tuberculose active.

Après inhalation par une personne saine, les bacilles atteignent les alvéoles pulmonaires et peuvent se multiplier. Le système immunitaire active ses cellules de défense pour contenir l'infection. Les bacilles sont phagocytés par les macrophages alvéolaires, suivant une réaction immunitaire innée, mais certains bacilles ne sont pas détruits et restent à l'état dormant. La multiplication des cellules bactériennes engendre la lyse du macrophage et la libération des bacilles qui seront à leur tour phagocytés par d'autres macrophages avec formation d'un foyer infectieux primaire. Les macrophages infectés produisent immédiatement des cytokines comme le $\text{TNF-}\alpha$ (Tumor Necrosis Factor alpha), l'interféron gamma ($\text{IFN-}\gamma$) et l'interleukine 12 (IL12) qui attirent les lymphocytes T avec production d'une réponse immunitaire lente (Cooper 2009). Les lymphocytes T activés prolifèrent, exercent leur activité bactériostatique ou bactéricide et migrent vers le poumon pour s'accumuler sur le site d'infection et former un granulome. Le granulome est ainsi constitué de macrophages alvéolaires détruits ou regroupés, de lymphocytes et d'une couronne fibroblastique (**Fig. 3**).

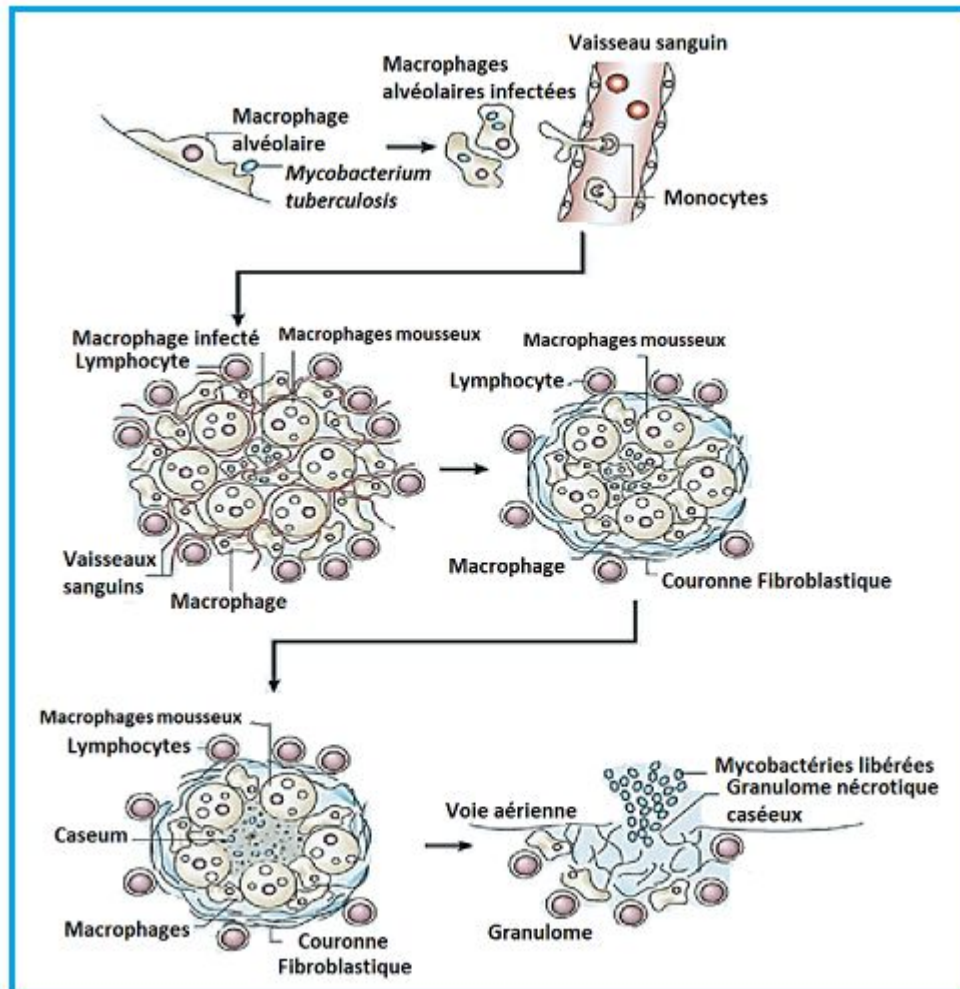


Figure 3 : Formation et progression histologique du granulome lors d'une infection à *M. tuberculosis* (Russell, Cardona et al. 2009).

1.1.2.2.1 Primo-infection

La primo infection tuberculeuse ou tuberculose primaire correspond aux manifestations cliniques (parfois asymptomatiques), biologiques et radiologiques observées à la suite du premier contact entre l'homme et le bacille. Les manifestations cliniques d'ordre symptomatique sont caractérisées par une altération générale associée à une toux. La tuberculose primaire peut s'accompagner d'un érythème noueux sous forme de nodule rouge situées dans le derme ou l'épiderme. Que la primo infection soit symptomatique ou asymptomatique, le virage de la réaction tuberculique (test IntraDermique Réaction,-IDR positif) et la positivité du test IGRA (Interféron-Gamma Release Assay) est la seule preuve de primo infection tuberculeuse (Barben Jürg , Berger Christoph et al. 2012). La primo infection est aussi caractérisée par des anomalies radiologiques telles qu'une adénopathie médiastinale

(Fig 4A) suivie d'une opacité nodulaire du chancre d'inoculation formant un complexe gangliopulmonaire qui sera ultérieurement calcifié (Fig 4B).



Figure 4 : Coupe TomoDensitoMétrie (TDM), A- Adénopathie médiastinale, B- Complexe gangliopulmonaire calcifié. (Carette 2002).

1.1.2.2.2 Tuberculose latente

La tuberculose latente est la forme dormante de la maladie sans manifestation de signes cliniques accompagnée d'une absence d'anomalies sur la radiographie thoracique et une réponse positive aux tests immunologiques IDR et IGRA (Fig.5, paragraphe 1.1.3). La tuberculose « latente » ou « dormante » est très répandue: 1/3 de la population mondiale serait infectée par le bacille (Trajman, Steffen et al. 2013). A ce stade le système immunitaire, première défense de l'organisme, d'une part détruit les bactéries et d'autre part empêche leur prolifération dans les alvéoles pulmonaires. Les personnes ayant une infection latente n'ont aucun symptôme et ne sont pas contagieuses. Lorsque le système immunitaire s'affaiblit, le risque de multiplication des bactéries s'accroît entraînant l'évolution de la maladie vers une tuberculose active suivie de ses symptômes typiques.

1.1.2.2.3 Tuberculose active

La tuberculose active se définit comme l'activation, la multiplication et la propagation des bacilles tuberculeux dans l'organisme lorsque le système immunitaire du sujet infecté est défaillant. Seules 10% des personnes infectées par *Mycobacterium tuberculosis* et qui n'ont pas été traitées pour une tuberculose latente développeront une tuberculose maladie au cours de leur vie (Vynnycky Emilia and M. 2000). Le sujet infecté présente des signes cliniques tels qu'une toux persistante (>3 semaines), une douleur thoracique ou difficulté respiratoire, des expectorations parfois teintées de sang, la fatigue, la fièvre, l'amaigrissement et des transpirations nocturnes (Fig. 5, phase 1). La transition d'une tuberculose latente vers une tuberculose active passe par un stade subclinique où la culture est occasionnellement positive avec une bascilloscopie négative (Fig. 5).

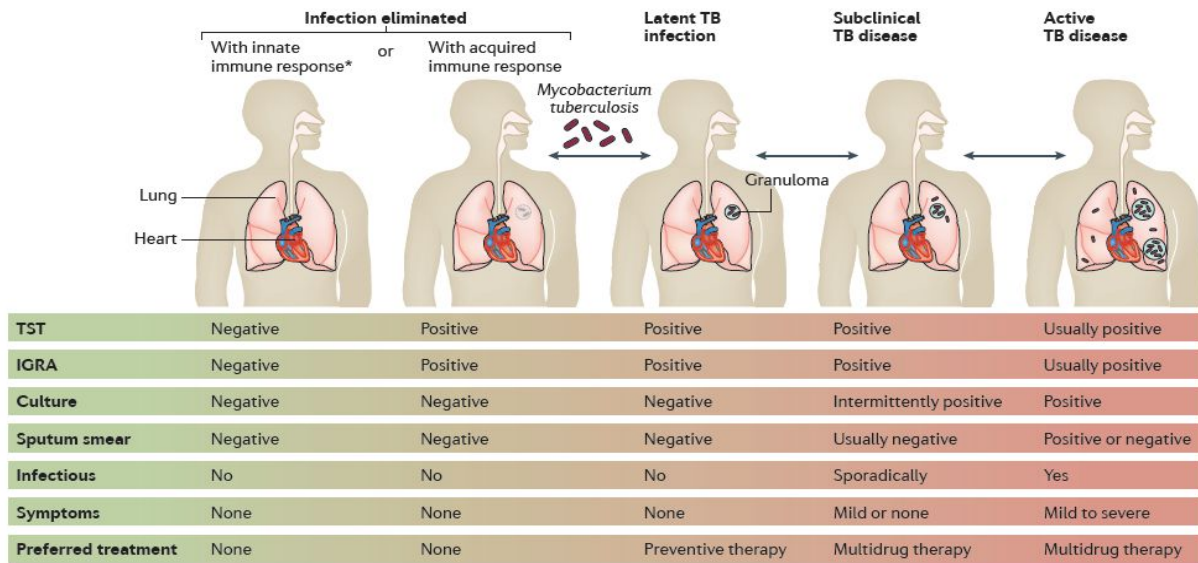


Figure 5 : Stades d'évolution de la tuberculose après infection par *Mycobacterium tuberculosis* (Pai M, Behr MA et al. 2016). TST - *Tuberculin Skin Test*, IGRA – *Interferon Gamma Release Assay*.

1.1.3 Prévention de la tuberculose par la vaccination

Le BCG, actuellement le seul vaccin, vivant atténué, est utilisé pour se protéger contre la tuberculose, (début du 20ème siècle à l'Institut Pasteur par Albert Calmette et Camille Guérin), confère une protection contre les cas graves de tuberculose du jeune enfant.

Mais avec 90% d'efficacité chez l'enfant, le BCG ne protège les adultes que dans environ un cas sur deux. Pour cela, des chercheurs cherchent à mettre au point un vaccin plus efficace pour éradiquer les formes pulmonaires de tuberculose chez l'adulte, qui sont les plus contagieuses.

Ainsi le BCG a été modifié génétiquement afin d'améliorer son immunogénicité en remplaçant le gène codant pour l'uréase C par le gène LLO codant pour la listériolysine de *Listeria Monocytogenes*. La suppression de l'uréase C qui est impliqué dans la neutralisation du phagosome hébergeant le BCG, permet une acidification rapide des phagosomes et favorise la fusion des phagolysosomes. Ainsi le BCG Δ ureC :: hly appelé « VPM1002 » favorise l'apoptose et l'autophagie, et facilite la libération d'antigènes mycobactériens dans le cytosol. Le vaccin a passé les essais cliniques de phase I en Allemagne et en Afrique du Sud, démontrant son innocuité et son immunogénicité chez les jeunes adultes. Il a également été testé avec succès dans un essai clinique randomisé de phase IIa et IIb chez des nouveau-nés en bonne santé, et nouveau-nés exposés et non exposés au VIH respectivement. Un essai

clinique de phase II / III est en cours en Inde, afin d'évaluer l'efficacité contre la récurrence de la tuberculose. Le vaccin VPM 1002 est un vaccin sûr, immunogénique, bien toléré et pourrait être une alternative au BCG dans le futur ((Nieuwenhuizen, Kulkarni et al. 2017), ((Loxton, Knaul et al. 2017))).

1.1.4 *Diagnostic de la tuberculose*

La suspicion d'une tuberculose maladie chez un individu présentant des signes évocateurs de la maladie ou lors d'une enquête épidémiologique conduit à mettre en place une stratégie de diagnostic. L'altération générale de l'état de santé d'une personne avec des symptômes, tels que l'asthénie, l'amaigrissement, une fièvre persistante, d'une toux chronique suivie d'une hémoptysie dans 10% des cas ou d'une dyspnée durant plus de trois semaines oriente le diagnostic du médecin vers la tuberculose. Le diagnostic est aussi proposé à des personnes en contact avec une personne présentant les symptômes de la maladie tuberculeuse, à des personnes provenant des zones à haute prévalence de la tuberculose et à des personnes infectées par le VIH, sous traitement immunodépresseur ou anti-TNF.

1.1.4.1 *L'intradermo-réaction (IDR) à la tuberculine - le test Interferon Gamma Release Assay (IGRA)*

L'IntraDermoRéaction (IDR) ou test de Mantoux a été le premier test destiné à un diagnostic immunologique de la tuberculose latente ou active. Elle consiste en l'injection intradermique d'un volume de 0,1 ml de tuberculine (Nayak and Acharjya 2012), dérivé protéinique purifié à partir de surnageant de culture de *M. tuberculosis*. La tuberculine qui contient plus de 200 antigènes partagés par *M. tuberculosis*, *M. bovis*, le BCG et certaines mycobactéries environnementales. L'IDR mesure la réponse cutanée *in vivo* due à la réaction inflammatoire, avec hypersensibilité retardée. Dans les heures suivant l'injection, il se produit un afflux de monocytes-macrophages, la formation de cellules géantes et l'afflux d'un grand nombre de lymphocytes T CD4+ qui crée une induration dermique. La lecture du diamètre d'induration se fait entre 48 et 72 heures après injection de la tuberculine (Nayak and Acharjya 2012). Une IDR positive (diamètre supérieur à 5mm) traduit une infection à tuberculose latente ou déclarée chez un sujet présentant des défenses immunitaires normales, mais peut aussi être signe d'un antécédent de vaccination par le BCG, ou d'une infection à mycobactérie non tuberculeuse (Nayak and Acharjya 2012). Une IDR négative traduit l'absence de contamination ou un état d'immunodépression (Pai and O'Brien 2008). La spécificité du test est de 70-90%, plus faible chez les personnes immunisées par le vaccin

BCG (Madhukar Pai, Shriprakash Kalantri et al. 2006). Ce test est encore très utilisé comme test de première intention en cas de suspicion de tuberculose.

Le test IGRA est un test sanguin qui aide au diagnostic de l'infection tuberculeuse, en mesurant la réaction immune de l'hôte au *Mycobacterium tuberculosis*. Les lymphocytes T d'une personne infectée par le bacille libèrent l'interféron gamma (IFN- γ) au contact avec les trois antigènes (ESAT 6, CFP-10, Tuberculine 7,7) spécifiques de *Mycobacterium tuberculosis*. Deux tests sont commercialisés: le QuantiFERON (QFT-GIT) et le TB-SPOT (T-Spot). En pratique, les échantillons de sang sont mélangés avec les antigènes, avec un témoin négatif (tube sans antigène) et un témoin positif (tube avec mitogène stimulant de façon aspécifique la production d'IFN- γ). Après plusieurs heures la concentration d' IFN- γ (QFT-GIT) ou le nombre de cellules productrices d'IFN- γ sont mesurés. Le résultat peut être positif, négatif ou indéterminé. Un test IGRA positif traduit une probable infection à *Mycobacterium tuberculosis*, qui peut être latente, en progression vers une tuberculose active, ou une récente exposition aux bacilles. Dans ce cas, il est nécessaire d'exclure l'hypothèse d'une tuberculose active par des examens cliniques, bactériologiques et radiologiques. La spécificité du test IGRA est de 90-100% même chez les personnes immunisées avec le vaccin BCG (Madhukar Pai, Shriprakash Kalantri et al. 2006); ce test est donc plus efficace pour déterminer une infection à tuberculose (Overton, Varma et al. 2018). Malgré son coût élevé et sa complexité technique, il peut être utilisé en addition à l'IDR. Mais ce test n'est pas recommandé dans le diagnostic de la tuberculose latente chez l'enfant, les personnes infectées par le VIH et les agents de santé (Sahal Al Hajoj, Bright Varghese et al. 2016) dans les pays à revenu faible/intermédiaire et dans les zones à forte prévalence (OMS 2011).

Le test IGRA comme l'IDR ne peuvent confirmer ou infirmer la maladie tuberculeuse en cas de suspicion radiologique ou clinique. Ces tests ne peuvent non plus confirmer ou exclure une tuberculose dans les cas cliniquement ou radiologiquement suspects. Ils ne permettent pas non plus de distinguer une infection tuberculeuse latente d'une tuberculose déclarée.

1.1.4.2 L'examen Radiologique

L'examen de base lors d'une suspicion de tuberculose active est la radiographie thoracique. Lors de cet examen, trois types de lésions de morphologie et d'étendue variables, caractéristiques d'une tuberculose pulmonaire sont recherchées: les nodules, les infiltrats et les cavernes. Ces lésions peuvent être associées. Les cavernes siègent au sein d'infiltrats alvéolaires ou macronodules au niveau des segments supérieurs du poumon où la teneur en

oxygène est élevée et favorable pour le bacille tuberculeux qui est aérobic (Carette 2002). Cependant un examen de radiologie positif à la tuberculose n'est pas suffisant pour confirmer la maladie, il doit être suivie d'un examen de bactériologie.

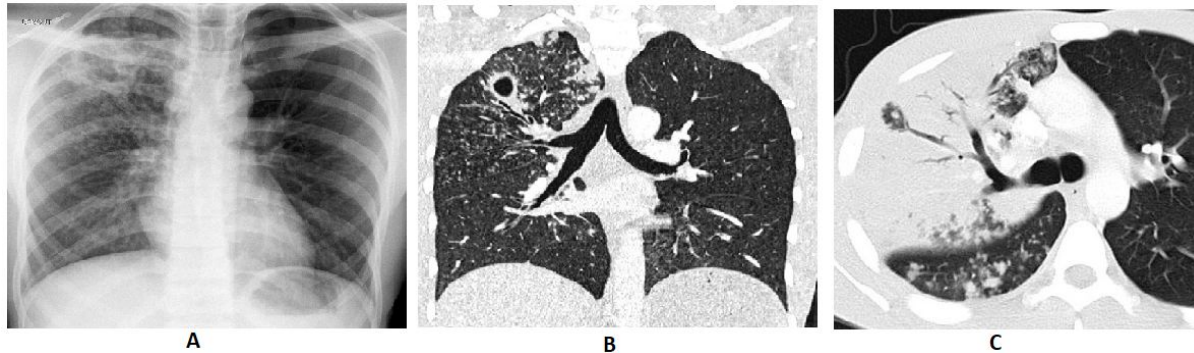


Figure 6 : Image radiographique du thorax d'un tuberculeux (A), image cavitaire apicale droite et des micronodules en coupe tomodensitométrie (TDM) (B), scanner de la bronche de drainage (C) (Carette 2002).

1.1.4.3 L'examen bactériologique

1.1.4.3.1 L'examen direct : la bacilloscopie

La bacilloscopie est une méthode de diagnostic de la tuberculose basée sur un examen microscopique (fluorescent ou optique) direct des expectorations issues d'un malade tuberculeux. Cet examen est souvent le seul pratiqué dans les pays en développement de par sa facilité d'accès, son efficacité et son faible coût. Il permet un diagnostic rapide avec une sensibilité moyenne (charge bacillaire minimale = 10000 organismes/ml de crachat (Desikan 2013) pour la détection des sujets contagieux, ce qui aboutit à la mise en place précoce des mesures pour éviter la contagion. Mais la sensibilité de ce test est plus faible chez les personnes co-infectées par le VIH. L'examen requiert un prélèvement répétitif d'expectorations durant 3 jours sur lesquels on réalise un frottis coloré par la méthode de coloration de Ziehl-Neelsen (**Fig. 1C**) ou fluorescente à l'Auramine (microscope fluorescent) (**Fig 1D**). Ces méthodes s'appuient sur les propriétés acido-alcool-résistantes des bacilles liées à la structure de leurs parois cellulaires. Cependant cet examen ne permet pas de différencier *Mycobacterium tuberculosis* des autres mycobactéries (Akhtar, Bretzel et al. 2000). Pour confirmer définitivement le diagnostic, la bactérie est isolée sur milieu de culture.

1.1.4.3.2 La culture

La mise en culture des prélèvements pulmonaires ou extrapulmonaires est la méthode la plus sensible qui permet d'isoler la souche de *M. tuberculosis* en vue d'établir le diagnostic

bactériologique de la tuberculose et l'infection à d'autres mycobactéries. La culture se fait sur les milieux solide (Löwenstein-Jensen, 20-21 jours d'incubation) ou liquide (Middlebrook, 10-15 jours d'incubation). Le seuil de détection est de $10^1 - 10^2$ bactéries/ml (Parrish and Carroll 2010). A la fin du 20^{ème} siècle, des automates avec incubateurs incorporés, utilisant des milieux liquides plus sensibles ont été développés pour réduire le délai de positivité par rapport au milieu solide. L'automate Bactec MGIT960 (Mycobacterium Growth Indicator Tube) de Becton Dickinson détecte la croissance des bactéries grâce à un indicateur fluorescent sensible à la concentration du milieu en oxygène (la diminution en oxygène génère une fluorescence après excitation à 365 nm). Le BacT/ALERT 3D de BioMérieux identifie la croissance mycobactérienne grâce à un indicateur coloré sensible à l'augmentation du pH, conséquence de la présence de CO₂. Le système VersaTREK/ESP Culture Système II utilise un milieu de croissance sélectif et détecte des modifications de pression dans la partie supérieure du flacon fermé, de production ou consommation de gaz due à la croissance bactérienne. Bien que ces automates soient similaires en performance et caractéristiques opérationnelles, le taux de contamination est plus élevé qu'avec le système semi-automatisé BACTEC 460. Aucune de ces méthodes ne peut détecter les espèces mycobactériennes dans une culture mixte de mycobactéries. Ainsi, la culture en parallèle sur des milieux solides confirme la morphologie unique des colonies et la détection de *M. tuberculosis* ((Gravet, Souillard et al. 2011), (Parrish and Carroll 2010)). Une méthode alternative est d'optimiser le milieu de culture avec de la lécithine d'œuf qui a un effet promoteur de croissance en vue d'une détection précoce de croissance de *M. tuberculosis* (Ghodbane, Raoult et al. 2014).

1.1.4.3.3 L'identification

Après un temps d'incubation les colonies obtenues sont soumises à des tests en vue de l'identification et de la différenciation entre mycobactéries et sous-espèces du complexe *M. tuberculosis*. Les bacilles tuberculeux diffèrent des mycobactéries atypiques par leur durée de croissance (1^{ère} colonie au 7^{ème} jour), leur morphologie, la thermorésistance (pas de croissance à 30°C), la production d'acide nicotinique, la présence de catalase, la sensibilité à l'acide para-amino-salicylique (PAS) et à la D-cyclosérine. Les bacilles *M. tuberculosis*, *M. bovis*, *M. africanum* du complexe *M. tuberculosis* en culture sont différents de par leurs morphologies, la durée d'incubation, la production de l'acide nicotinique (absent chez *M. bovis*), la présence de nitrate réductase (pas chez *M. bovis*). Seul *M. tuberculosis* est résistant à l'acide thiophène carboxylique.

1.1.4.3.4 L'antibiogramme

L'antibiogramme permet de déterminer la sensibilité du bacille tuberculeux aux différents antibiotiques en vue de proposer un traitement adéquat aux patients. La méthode des proportions est la méthode de référence dont le but est de mesurer la proportion de bacilles résistants qui existe dans une souche donnée sur un milieu de culture imprégné d'antibiotique (Canetti, Froman et al. 1963). On obtient cette proportion par le dénombrement de colonies qui se sont développées sur un milieu de culture contenant une concentration critique d'antibiotique par rapport aux colonies viables sur un milieu de culture sans antibiotique. La comparaison de la proportion obtenue aux proportions critiques conventionnellement définies permet de conclure à la sensibilité ou la résistance de la souche étudiée. Lorsque la proportion de mutants résistants est inférieure à 1 % pour l'isoniazide, la rifampicine, la streptomycine, l'éthambutol, et inférieure à 10 % pour le pyrazinamide, alors la souche est dite sensible dans le cas inverse, elle est résistante. L'antibiogramme lors d'une tuberculose ultra résistante sera expliquée plus bas (cf 1.1.5).

1.1.5 *Traitement de la tuberculose*

La tuberculose se traite par l'association de plusieurs antituberculeux et nécessite un respect des règles thérapeutiques par le médecin et le patient. Le traitement est constitué d'antibiotiques de première [Streptomycine (S)/ Isoniazide (H)/ Rifampicine (R)/ Pyrazinamide (Z)/ Ethambutol (E)]; seconde [les Aminoglycosides : Kanamycine (K)/ Amikacine (A)/ Capréomycine (C), les Thioamides : Ethionamide (Eth)/ Prothionamide (P), les quinolones (Q) : Ofloxacin (O)/ Ciprofloxacine (C)/ Levofloxacine (L)/ Moxifloxacine (M)/ Gatifloxacine (G), l'Acide Para-amino Salicylique (PAS), la Cyclosérine (CS)], et troisième lignes (Bedaquiline/ Delamanide) selon le degré de sensibilité de la bactérie (**Tableau 1**). La combinaison des antibiotiques inhibe le métabolisme bactérien particulièrement au niveau de la paroi cellulaire, et prévient les événements mutationnels en limitant les divisions cellulaires. En fonction de leur mode d'action les antituberculeux de 1^{ère} et 2^{ème} lignes sont regroupés en classes: inhibiteur de paroi cellulaire (H/E/Eth/CS), inhibiteur de synthèse d'acide nucléique (R/Q) et protéines (S/K) et inhibiteur du métabolisme d'énergie membranaire (Z). Ces antituberculeux sont efficaces en activité bactériostatique mais sont des stérilisateurs médiocres car incapables de tuer *M. tuberculosis* qui persiste dans les macrophages après la mort des bactéries actives (Martina Casenghi 2006). Tout traitement est rigoureux, toxique et a un double objectif: individuel et collectif. Sur le plan individuel, il guérit les malades atteints de la tuberculose et sur le plan collectif, il empêche la transmission dans la collectivité et la contamination des sujets sains, en stérilisant les sources d'infection.

1.1.5.1 Traitement général

Le traitement général de la tuberculose, dans les années 1952 après la découverte de l'isoniazide, était basé sur l'adjonction de l'isoniazide (H), la streptomycine (S) et l'acide para-amino salicique (PAS) pour une durée de 24 mois (Veziris 2014). Après la découverte de la rifampicine (Sensi et al 1983) et la pyrazinamide, qui sont des stérilisants partiels ((Zhang, Mitchison et al. 2014), (Tasneen, Li et al. 2011)), le traitement a été réduit à 6 mois par la combinaison de plusieurs antibiotiques en deux phases distinctes : phase d'initiation (2 mois de HRZE) et phase de continuation (4 mois de HR) ((WHO 2016), (Wouter, Robert et al. 2010)). La phase d'initiation peut être prolongée lorsque l'examen direct est positif, cela peut être dû à une malabsorption ou une interaction médicamenteuse. La durée du traitement peut être aussi prolongée (9 à 12 mois) en présence de formes graves de la maladie (méningite, tuberculose osseuse, miliaire).

Tableau 1: Classification des antituberculeux (WHO 2016), modes d'action, gènes et mécanismes moléculaires impliqués dans la résistance (Zhang et al 2015, Zhang et al 2016, Martinez et al 2018).

Medicaments	CMI µg/ml	Gènes de résistance	Fonctions des gènes	Mécanismes d'action de l'Antibiotique	Mutations	Fréquences des mutations %	
Antibiotique de 1ère ligne							
Isoniazide (INH, H)	0.02-0.2	<i>KatG</i>	Catalase-peroxidase	Inhibition de la synthèse de l'acide mycolique	mut. <i>katG</i> 315ACC, 315AAC,	50-95	
		<i>inhA</i>	Enoyl ACP reductase		<i>inhA</i> -15T, -8A	8-43	
Rifampicine (Rif, R)	0.05-1	<i>rpoB</i>	Beta sous-unité de l'ARN Polymérase	Inhibition de la synthèse d'ARN	435GTC, 445GAC; 445TAC, 450TTG, 450TGG	95	
Ethambutol (E)	1-2	<i>embB</i>	Arabinosyl transferase	Inhibition de la synthèse de l'arabinogalactane	306GTG, 306ATA	47-65	
Pyrazinamide (P)	16-100	<i>pncA</i>	Nicotinamidase/pyrazinamidase	Epuisement de l'énergie membranaire	mut. dans la region 516 bp de l'ORF	72-99	
		<i>rpsA</i>	Ribosomal S1 protéine	Inhibition de la trans- translation	-		
		<i>panD</i>	Aspartate decarboxylase	Inhibition de la pantothenate et de la synthèse de CoA	-		
Antibiotique de 2ème ligne							
A. Fluoroquinolones	Levofloxacin (L)/ Moxifloxacin (M)/ Gatifloxacin (G)	0.5 - 2.5	<i>gyrA/ gyrB</i>	Sous-unité A et B de l'ADN gyrase	Inhibition de la synthèse de l'ADN	Codons <i>gyrA</i> 90GTG; 91CCG; 94GCC; 94AAC; 94GGC; 94CAC, <i>gyrB</i> 500;538; 540	13-29; 2-7; 7-17; 5-8; 3-14; 21- 36; 0-7; 4; 1; 4

B. 2ème ligne agents injectables	Kanamycine (K)/ Amikacine (A)	2-4	<i>rrs</i>	16S ARNr, aminoglycoside acetyltransferase	Inhibition de la synthèse de protéines	nt 1401, 1484 et 1402 de <i>rrs</i>	38-80; 1-2;	
			<i>eis</i>	Régulateur transcriptionnel		<i>eis</i> -10A; -37T; -12T; -14T	5-33; 0-19; 3-6; 5-24	
	Capreomycine (C)	2-4	<i>rrs</i>	ARN 16S		nt. 1401 et 1402 de <i>rrs</i>	85	
			<i>tlyA</i>	ARNr 2'-O-méthyltransferase		mut. Sporadiques dans tout <i>tlyA</i>	-	
	Streptomycine (S)	2-8	<i>rpsL</i>	S12 proteine ribosomale		<i>rpsL</i> 43	52-59	
			<i>rrs</i>	16S ARNr		nt. 491, 512, 513 de <i>rrs</i>	8-21	
			<i>gidB</i>	16S ARNr méthyltransferase		-		
C. Autres type d'agents de 2ème ligne	Ethionamide (Eth) / Prothionamide	2.5-10	<i>etaA/ethA</i>	Flavin monooxygenase,	Inhibition de la synthèse de l'acide mycolique	mut in <i>inhA</i>	37	
			<i>ethR</i>	Répresseur de transcription		-	56	
			<i>inhA</i>	Enoyl ACP reductase		<i>inhA</i> -15T, -8A	-	
	Cycloserine (C)/ Terizidone (T)	10-40	<i>alr</i>	Alanine racemase	Inhibition de la synthèse des peptidoglycanes	-	A déterminer	
			<i>ddl</i>	D-alanine-D-alanine ligase,		-		
			<i>cycA</i>	D-serine symport de protons		-		
	Linezolid (L)	0.25 - 0.5	<i>rrn, rplC</i>	23S ARNr; protéine ribosomal L3	Inhibition de la synthèse des protéines	mut 2061T, 2576T; 460C <i>rplC</i>	30 de 11 d'isolats MDR	
Clofazimine (C)	0.1 - 0.25	<i>rv0678</i>	Répresseur de transcription du transporteur <i>MmpL5</i>	Production de radicaux oxydant	G193 deletion/insertion; C466T	43; 11.5		
D. Agent supplémentaires, pas impliquée dans le traitement de la MDR-TB	D1	Pyrazinamide (Z)	16 - 100	<i>pncA</i>	Nicotinamidase/ pyrazinamidase	Production d'énergie membranaire/inhibition de la translation	nt. G319T	85
		Ethambutol	1 - 2	<i>embB</i>	Arabinosyl transferase	Inhibition de la synthèse de l'arabinogalactane	306GTG, 306ATA	47-65

	High-dose Isoniazid	-	-	-	-	-	-	
Antibiotique de 3ème ligne								
D2	Bedaquiline	0.06	-	<i>atpE</i>	ATP synthase	Inhibition de la production d'ATP	D28N, A63V	71-83, 16-50
		0.12	-	<i>rv0678</i>	Répresseur de transcription pour le transporteur MmpL5		Mut. au nt. T146C	-
	Delamanide	0,006	-	<i>ddn</i>	Deazaflavin-dépendent nitroreductase	Inhibition de la synthèse de l'acide mycolique	-	-
				<i>fdg1</i>	F420-dépend glucose-6-phosphate déhydrogenase		Production de Nitrogen réactive	-
D3	P-AminoSalicylic acid (PAS)	1-8	<i>thyA</i> ,	Thimidylate synthase	Inhibition du métabolisme de l'acide folique de la Thymine	ACC604GCC	-	
			<i>dfrA</i>	Dihydrofolate reductase		-	-	
			<i>folC</i>	Dihydrofolate synthase		-	-	
			<i>ribD</i>	Enzyme dans la biosynthèses de la riboflavine		-	-	
	Imipenem-cilastatin	Inconnu				Inhibition de la synthèse de la paroi cellulaire bactérienne après fixation aux protéines de liaison des pénicillines (PLP)	-	-
	Meropenem							
	Amoxicillin-clavulanate							
	Thioacetazone							

Colonne 1: Types de médicaments, colonne 2 : médicaments, colonne 3 : concentration minimale inhibitrice (CMI), colonne 4 : gènes impliqués, colonne 5 : fonction des gènes, colonne 6 : mécanisme d'action de l'antibiotique, colonne 7 : mutations au sein des gènes, colonne 8 : fréquences des mutations, mut = mutations, nt = nucléotides.

1.1.5.2 Enjeux du traitement de la tuberculose sensible aux antibiotiques

Le traitement de la tuberculose sensible aux antibiotiques est le traitement standard de 6 mois. Cependant, l'enjeu du traitement de la tuberculose sensible repose sur: 1) la réduction de la durée de traitement (surtout pour la TB pédiatrique) et des interactions médicamenteuses notamment avec les antirétroviraux, 2) l'amélioration de l'observance du traitement pour réduire l'émergence de la résistance, 3) l'amélioration de la tolérance pour les groupes à risque (Nimmo, Lipman et al. 2015). Ainsi, la réduction du traitement à 4 mois avec les schémas de combinaison d'antituberculeux proposés :

- 2 mois de EM_(400mg)RZ / 2 mois de MR_(900mg) (Jindani, Harrison et al. 2014),
- 4 mois de HRZM / 2mois de placebo contre 2 mois de MRZE / 2 mois MHR (Gillespie, Crook et al. 2014) avec la Moxifloxacin pour remplacer l'Ethambutol ou l'Isoniazide,
- 2 mois de HRZG_(400mg/jr) / 2 mois HR pour remplacer l'Ethambutol (Merle, Fielding et al. 2014) ont des efficacités inférieures au régime standard.

Seul le régime 2 mois EM_(400mg) RZ / 4 mois MR_{p(1200mg)} (Jindani, Harrison et al. 2014) a une efficacité équivalente au régime standard. Or, la moxifloxacin présente moins d'effets secondaires et pourrait donc être utilisée à la place de l'Isoniazide pour éviter l'hépto-toxicité et des effets indésirables graves liés au traitement de base (Saukkonen, Cohn et al. 2006). L'augmentation de la dose de Rifampicine (35mg/jour) est sans risque et bien tolérée par le patient, et elle est aussi envisagée pour la réduction de durée de traitement ((Boeree, Diacon et al. 2015), (Boeree, Heinrich et al. 2017), (Milstein, Lecca et al. 2016)) sans effets indésirables comme l'hépto-toxicité (Jindani, Borgulya et al. 2016).

1.1.5.3 Traitement de la tuberculose mono/poly/multi/ultra résistante aux antibiotiques

La mono-résistance de la tuberculose est la résistance de *M. tuberculosis* à un seul antituberculeux de première ligne, le plus souvent à l'Isoniazide, parfois à l'Ethambutol et rarement à la Rifampicine sauf chez les personnes infectées par le VIH. Dans ce cas, le traitement est dans un premier temps limité à trois antituberculeux de première ligne encore actifs dont PZA (d'une durée de 2 mois) et dans un second temps à deux antituberculeux actifs pendant 6 mois (si MonoRésistant E) et 9 mois (si MonoRésistant H ou R).

La poly-résistance est la résistance à au moins deux antituberculeux de première ligne sauf la Rifampicine et l'isoniazide. Après le résultat d'antibiogramme élargi, sont prescrits

trois antituberculeux actifs dont Z et une quinolone dans un premier temps, suivi dans un second temps de deux antituberculeux dont l'activité a été confirmée pour une durée totale de 9 mois.

La multirésistance (MDR) est la résistance de *M. tuberculosis* au moins à l'Isoniazide et Rifampicine. Le traitement dure 18 à 24 mois et est basé sur l'utilisation d'antituberculeux actifs : une quinolone, de l'Amikacine, la Prothionamide en phase initiale et 3 antituberculeux lors l'Amikacine lors des cas sévères en phase de continuation. En 2016, l'OMS a approuvé l'utilisation d'un protocole thérapeutique bref et standardisé pour les cas de tuberculose-MR qui n'ont pas de résistance aux médicaments de deuxième intention. Ce schéma thérapeutique, d'une durée de 9 à 12 mois, est bien moins coûteux que le traitement classique, qui peut prendre jusqu'à 2 ans (WHO 2017).

L'ultra-résistance (XDR pour eXtreme Drug Resistance) est la résistance au moins à la Rifampicine/Isoniazide, à une quinolone et à un antituberculeux injectable (Kanamycine, Amikacine, Capréomycine). Cette forme de résistance particulièrement grave a moins de chance de guérison (Wouter, Robert et al. 2010). Le schéma de traitement de la tuberculose multirésistante n'est pas adapté aux patients ayant une tuberculose-XDR ou résistante aux antituberculeux de seconde intention ; il faut alors prescrire un protocole thérapeutique long auquel on peut ajouter l'un des nouveaux médicaments (bédaquiline et délamanide) et adapter le traitement en fonction du résultat de l'antibiogramme élargi et de l'évolution clinique du patient.

Ces types de résistance sont dus au non-respect thérapeutique du médecin et/ou du patient (résistance acquise) ou à une contamination directe avec des germes résistants (résistance primaire).

1.1.5.3.1 Résistance de la tuberculose aux antibiotiques

1.1.5.3.1.1 La multi-résistance de la tuberculose aux antibiotiques: la Rifampicine et l'Isoniazide

La Multi Résistance (MDR) de *Mycobacterium tuberculosis* est la résistance aux deux principaux antibiotique de première ligne, la Rifampicine et l'Isoniazide. Ces résistances sont due à des mutations au sein des gène codants.

La Rifampicine, dont la cible est la sous-unité beta de l'ARN polymérase, codée par le gène *rpoB*, se lie de façon covalente à sa cible et inhibe l'ARN polymérase, entraînant ainsi la mort de la bactérie par blocage transcriptionnel. La résistance à la rifampicine est liée

essentiellement à des mutations dans une région restreinte (hot-spot) du gène *rpoB*, qui s'étend du codon 430 au codon 452. Les mutations les plus fréquentes sont les substitutions Ser450Leu et His445Tyr. Ces mutations entraînent une modification de l'enzyme, diminuant ainsi la fixation ou l'accessibilité de la rifampicine.

La résistance à l'Isoniazide implique essentiellement des mutations au sein du gène *katG* et du promoteur *inhA*. Le gène *katG* code pour l'enzyme catalase peroxydase de *M. tuberculosis* qui active l'isoniazide. La mutation la plus reportée est la mutation ACC au sein du codon 315 (Heym et al, 1995). Le promoteur du gène *inhA* qui code pour l'*enoyl acyl carrier protein reductase* (protéine impliquée dans l'élongation des acides gras et dans la biosynthèse des acides mycoliques) qui est aussi impliqué dans la résistance à l'Isoniazide et la mutation la plus relevée est la -15T. Ses modifications correspondent le plus souvent à des mutations au niveau du promoteur conduisant à sa surexpression (Banerjee et al, 1994).

1.1.5.3.2 L'Ultra-Résistance (UR) de la tuberculose aux antibiotiques

1.1.5.3.2.1 Mécanisme de résistance aux antibiotiques de deuxième ligne: ultra-résistance

Les antibiotiques de deuxième ligne impliqués dans le traitement classique de la TB-Multidrug Résistante (MDR) sont les fluoroquinolones, les aminoglycosides (Kanamycine, Amikacine), le polypeptide Capréomycine et Streptomycine. D'autres antituberculeux de seconde ligne peuvent être utilisés: Ethionamide/Prothionamide, Cycloserine/Terizidone, Linezolid, Clofazimine (**Tableau 1**).

1.1.5.3.2.1.1 Fluoroquinolones

Les fluoroquinolones sont des antibiotiques bactéricides rapides, concentration-dépendants, et à spectre antibactérien large. Les quinolones agissent par formation d'un complexe ternaire entre ADN/ADN gyrase/topoisomérase IV. Ces enzymes sont directement impliquées dans les mécanismes de déenroulement et de superenroulement de l'ADN au cours de la réplication afin de faciliter l'action de l'ADN polymérase. L'activité anti-bactérienne de la fluoroquinolone passe surtout par l'inhibition des activités de l'ADN gyrase et le blocage de la topoisomérase IV (Kohanski, Dwyer et al. 2010). L'ADN gyrase est une protéine tétramérique ayant deux sous unités A et B. Dans *M. tuberculosis*, *gyrA* et *gyrB* encodent les sous-unités A et B (Howard E. Takiff, Leiria Salazar et al. 1994). Une région de *gyrA* (320 bp) et *gyrB* (375 bp) détermine la résistance à la quinolone, elle a été nommée QRDR. Les mutations identifiées au sein de *gyrA* dans les isolats cliniques de *M. tuberculosis* sont

regroupées au niveau des codons 90, 91, 94 (Lee, Lynn L.H. Tang et al. 2002; Lee 2016), et les plus fréquentes affectent les codons 90 et 94 (Howard E. Takiff, Leiria Salazar et al. 1994; Sun, Zhang et al. 2008). Le polymorphisme au niveau du codon 95 est considéré comme non impliqué dans la résistance aux quinolones (Sreevatsan, Pan et al. 1997).

1.1.5.3.2.1.2 Les antituberculeux Injectables : Kanamycine, Amikacine, Capreomycine

Les aminoglycosides agissent sur la synthèse protéique où la cible principale est le ribosome en particulier la sous-unité 30S. La Kanamycine et son dérivé l'Amikacine sont des inhibiteurs de la synthèse protéique qui modifient la structure de l'ARNr 16S. Dans certaines bactéries, le mécanisme de résistance à la Kanamycine est caractérisé par l'inactivation du médicament par modification enzymatique, et la mutation ou la méthylation de l'ARNr, ce qui perturbe la liaison du médicament au ribosome. Chez toutes les espèces et en particulier chez *M. tuberculosis*, les mutations dans le gène *rrs* qui encode l'ARNr 16S entraînent une résistance de haut niveau à la Kanamycine [Concentration Minimale Inhibitrice (CMI) \geq 80 $\mu\text{g/ml}$], et certaines mutations peuvent conférer une résistance croisée à d'autres médicaments de deuxième ligne, y compris l'amikacine et la capréomycine (Maus et al, 2005). Le gène appelé *tlyA* codant pour la méthyltransférase de l'ARNr est aussi impliqué dans la résistance des souches de *M. tuberculosis* à la capréomycine (antibiotique de type polypeptide).

Les mutations les plus fréquemment rapportées et qui causent la résistance aux aminoglycosides sont A1401G, C1402T et G1484T dans le gène *rrs* (Suzuki et al, 1998). La mutation A1401G cause une forte résistance à la Kanamycine (60%) et 70-80% à l'Amikacine et à la Capreomycine. De même, les mutations identifiées en positions (G-37T) / (C-14T)/(C-12T)/(G-10A) dans la région promotrice du gène *eis* (qui code pour l'aminoglycoside acetyltransférase) confèrent dans 6-22% des cas à la résistance à la Kanamycine et l'Amikacin ((Zaubrecher, Sikes et al. 2009), (Via, Georghiou et al. 2012), (Zhang and Yew 2015)). Mais cliniquement, le plus souvent ces mutations confèrent seulement une résistance à la Kanamycine. Les mutations C1402T et G1484T sont rares; C1402T cause une résistance de haut niveau à la Capreomycine et à l'Amikacine tandis que G1484T confère une résistance aux trois injectables : Kanamycine/l'Amikacine/la Capreomycine. Les mutations 1401, 1402, et 1484 présente simultanément au sein du gène *rrs* favorisent une résistance croisée des souches aux antibiotiques respectifs (Zhang and Yew 2015).

1.1.5.3.2.1.3 Streptomycine

La streptomycine est un antibiotique appartenant à la classe des aminosides. Sa cible est la sous-unité 30S du ribosome qui comporte l'ARNr16S codé par le gène *rrs* et la protéine ribosomale S12 codée par le gène *rpsL*. Après interaction avec la protéine S12, la streptomycine se lie à l'ARNr16S et inhibe la synthèse protéique. La résistance à la streptomycine de *M. tuberculosis* est liée à des mutations du gène *rpsL* (55 % des souches étudiées) et du gène *rrs* (15 % des souches étudiées) (Nadine Honoré and Cole 1994; Zhang and Yew 2015). Pour 30 % des souches, le mécanisme de résistance est inconnu.

1.1.5.3.3 Résistances aux autres antibiotiques de seconde ligne

L'éthionamide et le prothionamide sont des thioamides dérivés de l'acide thioisonicotinique et ont une activité bactéricide contre *M. tuberculosis*. L'éthionamide (ETO) est un promédicament, activé par *EtaA/EthA* (une mono-oxygénase), qui inhibe la même cible que INH, c'est-à-dire le gène *inhA* qui participe à la synthèse de l'acide mycolique, composant de la paroi cellulaire. La structure et l'activité de la prothionamide est identique à celle de l'éthionamide ((Andrea E. DeBarber, Khisimuzi Mdluli et al. 2000), (Baulard, Betts et al. 2000), (Zhang and Yew 2015)).

La cyclosérine inhibe la synthèse du peptidoglycane dans la paroi cellulaire en bloquant l'activité de la D-alanine racémase, impliquée dans la conversion de L-alanine au D-alanine: D-alanine ligase (Ddl) (**Tab. 1**). Cependant le mécanisme de résistance de *M. tuberculosis* n'est encore pas parfaitement élucidé (Zhang and Yew 2015)).

Le linezolide a une activité significative contre *M. tuberculosis*, il inhibe la synthèse des protéines par liaison à la 23S rRNA peptidyl transférase de la sous-unité ribosomale 50S et formant une interaction secondaire avec la sous-unité 30S. Les mutations G2061T et G2576T dans le gène *rrn* codant pour l'ARNr 23S causent un haut niveau de résistance des souches MTB au linezolide. La mutation T460C dans le gène *rplC*, codant pour la protéine ribosomale L3 confère aussi la résistance de *M. tuberculosis* au linezolide. Il est rapporté que seulement 30% des souches MDR ont des mutations dans les gènes *rrn* et *rplC*. Le linezolide est utilisé dans les cas de MDR et XDR-TB compliqués ((Zhang and Yew 2015),(Zhang, Mitchison et al. 2014)).

La clofazimine a une bonne activité contre les mycobactéries, telles que *M. tuberculosis*. Son mécanisme d'action n'est pas clair, et pourrait inclure la production d'oxygène réactif, l'inhibition de la production d'énergie de la NADH déshydrogénase et l'interruption du

mécanisme respiratoire ayant lieu au niveau de la membrane, conduisant à la réduction de la production d'ATP (Zhang and Yew 2015). Les mutations dans le gène *Rv0678* jouent un rôle majeur dans la résistance à la clofazimine. Des mutations dans les gènes *Rv1979c* et *Rv2535c* ont été détectées dans d'autres souches résistantes à la clofazimine ((Zhang and Yew 2015), (Zhang, Chen et al. 2015)).

1.1.5.3.4 Résistances aux nouveaux antituberculeux

La bédaquiline (TMC207) est une diarylquinoline. Elle inhibe spécifiquement l'ATP synthase mycobactérienne, une enzyme essentielle à la production d'énergie chez *Mycobacterium tuberculosis*. L'inhibition de l'ATP synthase entraîne des effets bactéricides tant sur les bacilles tuberculeux en répliation que sur les bacilles dormants. Sa combinaison avec la PZA a un effet très stérilisant dans le modèle murin ((Zhang and Yew 2015), Koul et al 2008).

Le délamanide (OPC-67683) est un promédicament activé par la deazaflavin-dependent nitro-réductase (*Ddn*). Elle inhibe la synthèse de l'acide mycolique ((Koul, Vranckx et al. 2008; Zhang and Yew 2015)). Elle est utilisée dans le traitement de la TB-MDR ou XDR lorsque le régime thérapeutique recommandé n'est pas toléré par le patient.

1.1.6 Méthodes phénotypiques de détection des résistances aux antibiotiques

Les tests phénotypiques relatifs à la sensibilité/résistance du bacille tuberculeux aux antibiotiques ont pour but de donner une indication sur le premier traitement, de confirmer/infirmier la résistance bacillaire lorsqu'un malade n'a pas réagi positivement au premier traitement, d'évaluer l'existence d'une résistance naturelle ou acquise des bacilles tuberculeux présents dans la communauté. Ces tests peuvent être classés en test direct ou indirect. La détection directe (4-6 semaines) de la sensibilité/résistance de *M. tuberculosis* est réalisée sur des expectorations ou sur d'autres échantillons pathologiques cultivés directement sur des milieux contenant l'antibiotique. La détection indirecte (8-10 semaines) est faite sur des isolats issus de la première culture de diagnostic. La différence entre ces deux types de détection est la réduction du temps de diagnostic. La méthode de détection indirecte la plus utilisées dans plusieurs laboratoires est la méthode des proportions réalisée en milieu solide (Löwenstein-Jensen, 7H11 ou 7H12) ou liquide (système Bactec[®] ou MGIT[®]), en vue de déterminer la proportion de bactéries résistantes aux antituberculeux dans une population donnée (Canetti, Rist et al. 1963).

1.1.6.1 Méthodes phénotypiques de détection de sensibilité/résistance sur milieux solides

L'antibiogramme sur milieu solide se fait sur les milieux Löwenstein-Jensen, 7H11 ou 7H12 et il existe trois méthodes de détection de la sensibilité/résistance de *M. tuberculosis* aux antituberculeux : la méthode des proportions, la méthode des concentrations absolues, et la méthode des rapports de résistance (Canetti, Froman et al. 1963).

La méthode des proportions est la méthode de référence et est utilisée dans la plupart des laboratoires. Elle ne nécessite pas forcément de témoins contrairement aux autres méthodes. Comme pour les autres méthodes, elle consiste à ensemencer et incuber à 37°C plusieurs dilutions de la suspension bacillaire à étudier afin de mesurer la proportion de bactéries résistantes à la concentration critique d'antibiotique (0.2µg/mL Isoniazide; 1µg/MI Rifampicine; 5µg/ml Streptomycine; 0.5µg/ml PAS; 1µg/ml Thioacetazone; 20µg/ml Ethionamide). Une souche est considérée résistante si la proportion des résistants est supérieure à 1%.

Pour la méthode des concentrations absolues, une souche est déclarée « résistante » si on observe plus de 20 colonies sur un milieu contenant des antibiotiques à la concentration critique. Dans la méthode de rapport de résistance, le rapport de résistance est le rapport de la concentration minimale inhibitrice (CMI) de croissance des souches testées sur la concentration minimale d'inhibition de croissance des bactéries de référence. La souche est considérée résistante si le rapport est supérieur ou égal à 8. L'antibiogramme de *M. tuberculosis* est aussi réalisé par la méthode du E-test sur gélose Middlebrook 7H11. Cette méthode est essentiellement utilisée pour la mesure des CMI aux antibiotiques des mycobactéries non tuberculeuses avec des résultats en 5-7 jours. Ses résultats sont corrélés à ceux de la méthode des proportions (Wanger and Mill 1996). Les techniques utilisées pour l'antibiogramme ont longtemps varié d'un laboratoire à un autre, d'un pays à un autre, dans les méthodes utilisées ainsi que dans l'interprétation((Canetti, Froman et al. 1963). L'OMS travaille depuis lors afin de standardiser les méthodes en vue d'une bonne détermination du phénotype (sauvage/mutant) des souches.

1.1.6.2 Méthodes phénotypiques de détection de sensibilité/résistance sur milieu liquide

Le test phénotypique sur milieu liquide est plus rapide (3 à 5 jours) et est automatisé. Les plus utilisés sont l'automate Bactec MGIT960 (*Mycobacterium Growth Indicator Tube*,

Becton Dickinson Diagnostic Systemes, USA), le BACTEC 460TB (*Becton Dickinson Diagnostic Systemes, USA*), le BacT/ALERT 3D (*BioMérieux, USA*), le test de sensibilité Myco VersaTREK (*Trek Diagnostics, USA*). L'antibiogramme sur le BACTEC 460TB est une méthode radiométrique mesurant la quantité de CO₂ marqué au ¹⁴C en présence ou en l'absence d'antibiotiques. La sensibilité se manifeste par une inhibition de croissance dans les flacons contenant des antibiotiques comparée à celle observée dans les flacons témoins sans antibiotique (témoins 100 %) et les flacons témoins avec la dilution 1/100 (témoins 1 %) inoculés dans les mêmes conditions ((N. Rastogi, K.S. Goh et al. 1989)). Plus tard le Bactec 460TB qui mettait en jeu des radioéléments a été substitué par le MGIT 960 (*Mycobacterium Growth Indicator Tube, Becton Dickinson Diagnostic Systemes, USA*) qui est basé sur l'ensemencement de tubes contenant des concentrations critiques des différents antituberculeux et de tubes témoins, suivi d'une lecture quotidienne en fluorescence à 365 nm. Une souche est dite résistante si la fluorescence est positive dans un délai de 2 jours par rapport au tube témoin, la fluorescence apparaissant lors de la diminution de la concentration en O₂ du milieu en cas de croissance bactérienne ((Shaun B. Walters and Bruce A. Hanna 1996), (Palaci Moises, Suely Yoko Mizuka Ueki et al. 1996)). Le test de sensibilité Myco VersaTREK (*Trek Diagnostics, USA*) est basé sur la détection du changement de pression dû à la dégradation de l'oxygène. Les résultats sont corrélés à ceux de la méthode des proportions. A cause de la lenteur de croissance des *M. tuberculosis*, la détection de la résistance en milieu liquide requiert au minimum 3 jours, et 10 jours pour les techniques utilisées en milieu hospitalier.

1.1.7 Méthodes de biologie moléculaire pour la détection des résistances aux antibiotiques

Les méthodes par amplification génique sont des techniques de biologie moléculaire rapides qui détectent les mutations corrélées à la résistance par amplification des gènes les plus souvent impliqués suivie d'hybridation à des sondes porteuses des mutations. Elles permettent une initiation du traitement plus rapide par rapport aux méthodes phénotypiques. En 2008, l'OMS a approuvé l'utilisation du test Line Probe Assay (LPA) (Hain life Science technology) pour la détection de la résistance à la Rifampicine/Isoniazide et plus tard aux antibiotiques de seconde ligne (MTBDR*plus V1 et V2*, MTBDR*sl*). Le test de GeneXpert MTB/RIF (Cepheid Sunnyvale, USA), un test automatisé à base de cartouches permettant une analyse entièrement automatisée d'échantillons cliniques, a été reconnu par l'OMS également. Hormis ces trois tests, il existe d'autres techniques de détection : le Séquençage par Sanger ou

le pyroséquençage des gènes de résistance, le séquençage du génome entier, associé à l'analyse bioinformatique, le High Resolution Melting Temperature, la MLPA, l'Allele Specific Primer Extension, SNP-typing sur microbilles en cytomètre en flux Luminex200 (qui permettent le typage des SNP responsables des différentes résistances aux antituberculeux au sein des gènes correspondants). Tous ces tests ont le même but, le diagnostic prédictif rapide de résistance. Dans la suite, nous présenterons le principe, les cibles précises, la sensibilité et la spécificité de ces techniques.

❖ Test d'hybridation inverse sur bandelettes (MTBDR plus, Hain LifeScience)

Le principe de ce test est basé sur l'amplification multiplexée de fractions de gènes codants pour la cible des antituberculeux suivie d'une hybridation avec des sondes correspondant aux gènes sauvages ou mutés présents sur la bandelette (**Fig.7**). Il est utilisé sur des cultures ou des échantillons positifs à l'examen microscopique et les résultats sont fournis en quelques heures. Ce test est subdivisé en deux sous-tests, un pour identifier la résistance à la Rifampicine/Isoniazide et l'autre pour la détection de résistance à la Fluoroquinolone, Amikacine, Kanamycine, Capréomycine et Ethambutol. La deuxième version du test, Geno Type MTDR*plus* 2.0 évaluée en zone endémique, montre une sensibilité de 90% et 93.5% pour la Rifampicine et l'Isoniazide. La deuxième génération du deuxième test de Hain, Geno Type MTDR*sl* destiné à diagnostiquer la tuberculose XDR, détecte les résistances aux différents antituberculeux avec une sensibilité allant de 68 à 95% (FLQ-95%, AMK-83%, KAN-91%, CAP-82%, EMB-68%) ((Yan Feng, Sijun Liu et al. 2013), (Brossier, Guindo et al. 2016)).

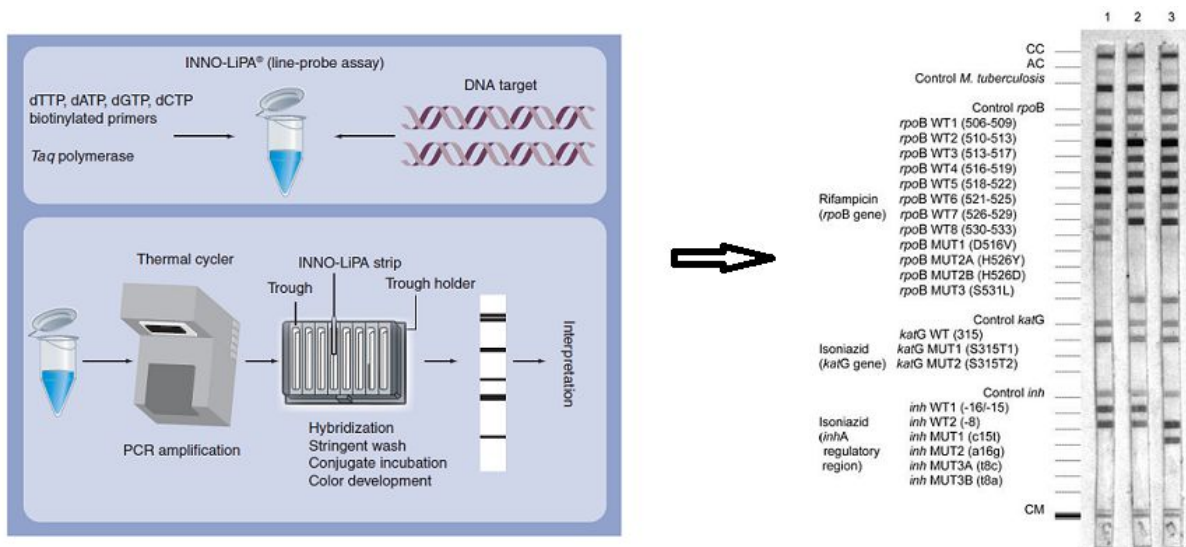


Figure 7 : Représentation schématique du principe du test Hain (Veziris 2014)

❖ Les test GeneXpert MTB/Rif, GeneXpert MTB/Rif Ultra et Xpert XDR

Les tests GeneXpert sont des tests semi-quantitatifs de PCR en temps réel qui permettent de réaliser dans une seule cartouche les différentes étapes d'extraction, purification, amplification d'ADN, hybridation des sondes et une détection multiplexée. Le test de GeneXpert MTB/Rif (Cepheid, Inc, Sunnyvale, CA, USA) est un test automatisé qui permet de détecter en 90 min *M. tuberculosis* ainsi que la résistance à la Rifampicine avec une grande précision dans les échantillons cliniques. Ce test a été recommandé par l'OMS à la place de la microscopie, la culture, la culture avec antibiotique (DST) comme test de diagnostic initial chez l'adulte ou l'enfant suspect d'une tuberculose multirésistante ou associée au VIH. Le test GeneXpert MTB/Rif a une sensibilité de 94% et une spécificité de 98% (Steingart KR, Schiller I et al. 2014). Mais ce test a une faible sensibilité pour détecter la tuberculose difficile à diagnostiquer dans les populations à risques (enfants et personnes vivant avec le VIH). Afin d'améliorer la détection de la résistance à la rifampicine, l'OMS recommande le nouveau test « Next – Generation Xpert/MTB Rif Ultra » pour remplacer le Xpert MTB/Rif ((WHO 2017); (McKenna 2017)). La sensibilité du test Ultra (Cepheid Sunnyvale, USA) à détecter la résistance à la rifampicine est meilleure et a montré des performances significativement meilleures par rapport à la cartouche Xpert® MTB / RIF actuelle pour la détection de *M. tuberculosis* dans des échantillons avec un faible nombre de bacilles, en particulier des échantillons à frottis négatif et à culture positive comme ceux provenant des personnes co-infectées par le VIH (Dorman, Schumacher et al. 2018), des échantillons pédiatriques et des échantillons extra-pulmonaires (Bahr, Nuwagira et al. 2018). Un autre test « Xpert XDR » de Cepheid pas encore approuvé par l'OMS, permet de détecter la résistance à l'Isoniazide, à la Fluoroquinolone, et aux antituberculeux injectable de 2^{ème} ligne avec une sensibilité de INH-98.1%, FQ- 95.8%, KAN-92.7%, AMK- 96.8% ; et une spécificité de INH-98.1%, FQ- 95.8%, KAN-92.7%, AMK- 96.8% par rapport au séquençage ((Boehme 2017); (Lessem 2017)).

❖ Séquençage / le séquençage de génome entier (WGS)

Le séquençage permet la détermination de la séquence nucléotidique, donc la mise en évidence des modifications génétiques responsables d'une résistance par comparaison avec la séquence de bacilles sensibles. Les résultats sont rapides mais coûteux en terme de matériel et d'analyse, cette technique est aujourd'hui applicable en routine grâce à l'utilisation d'automates de séquençage. Le séquençage automatique, contrairement aux techniques

précédentes, permet d'analyser des petites régions d'ADN (150 bp à 1 000 bp selon la technologie). Par exemple avec le séquenceur de nouvelle génération (NGS) de Illumina HiSeq X, en trois jours on peut séquencer un génome humain (<http://dridk.me/ngs.html>). Le séquençage de génome entier, aujourd'hui plus accessible permet non seulement la détection et l'identification de MTBc dans les échantillons cliniques mais aussi favorise le diagnostic des souches MDR/XDR (Louise J Pankhurst, Carlos del Ojo Elias et al. 2016). L'analyse des données de séquençage requiert des outils bioinformatique performant. Le *TBProfiler*, un outils bioinformatique permettant de prédire la résistance, est concordant avec le DST pour les antituberculeux et démontre d'une haute performance par rapport à l'outil *Mykrobe* ((Phelan, O'Sullivan et al. 2016)).

Ainsi, le séquençage et/ou séquençage de génome permettent de prédire la résistance ou la sensibilité des souches MTBc aux antibiotiques, et servent ainsi de guide pour le traitement des patients avec un scénario de résistance bacillaire difficile.

1.1.8 *La tuberculose dans le monde actuel*

1.1.8.1 *Epidémiologie de la tuberculose dans le monde*

L'Organisation Mondiale de la Santé (OMS) a déclaré la tuberculose urgence de santé publique dans le monde en 1993 et publie un rapport annuel depuis 1997 sur la maladie (épidémiologie, diagnostic, traitements, contrôle, R&D). La tuberculose est actuellement la 10^{ème} cause de mortalité dans le monde. En 2016, 10,4 millions de personnes (90% - adultes, 65% - homme, 10% VIH+) ont contracté la tuberculose (toutes formes). La probabilité de développer la tuberculose est plus élevée parmi les personnes infectées par le VIH (1,4 million de nouveaux cas en 2016 dont 374 000 morts), ceux dont le système immunitaire est défaillant, ceux qui sont affectés par des facteurs de risques tels que la malnutrition (3,1-3,3%), le diabète (2,3-4,3%), le tabagisme (1,6-2,5) et l'alcool (1,9 – 4,6) ((Knut Lönnroth, Kenneth G Castro et al. 2010); (WHO 2017)).

Entre 2000 et 2016, on observe une régression mondiale de la tuberculose mais encore lente avec 3% de baisse par an de taux de mortalité et 2% de baisse d'incidence. Régionalement, l'incidence de la tuberculose a chuté de 4,6% (2015-2016) dans les pays d'Europe et de 4% depuis 2010 dans les pays à haute prévalence (Ethiopie, Kenya, Lesotho, Namibie, Russie, Tanzanie, Zambie et Zimbabwe). La mortalité de la tuberculose a aussi chuté depuis 2010 dans les pays d'Europe et du Pacifique de l'Ouest avec 6 et 4,6% respectivement; et avec un taux au-delà de 6% dans les pays à forte prévalence. Cependant en

2016, globalement 16% des personnes qui ont développé la maladie en sont mortes (WHO 2017)

L'incidence globale est plus élevée dans les pays du Sud-Est d'Asie (45%), d'Afrique (25%), du Pacifique de l'ouest (17%) et en faible proportion dans les pays de la Méditerranée de l'est (7%), de l'Europe (3%) et d'Amérique (3%) (Fig. 8A) (WHO 2017).

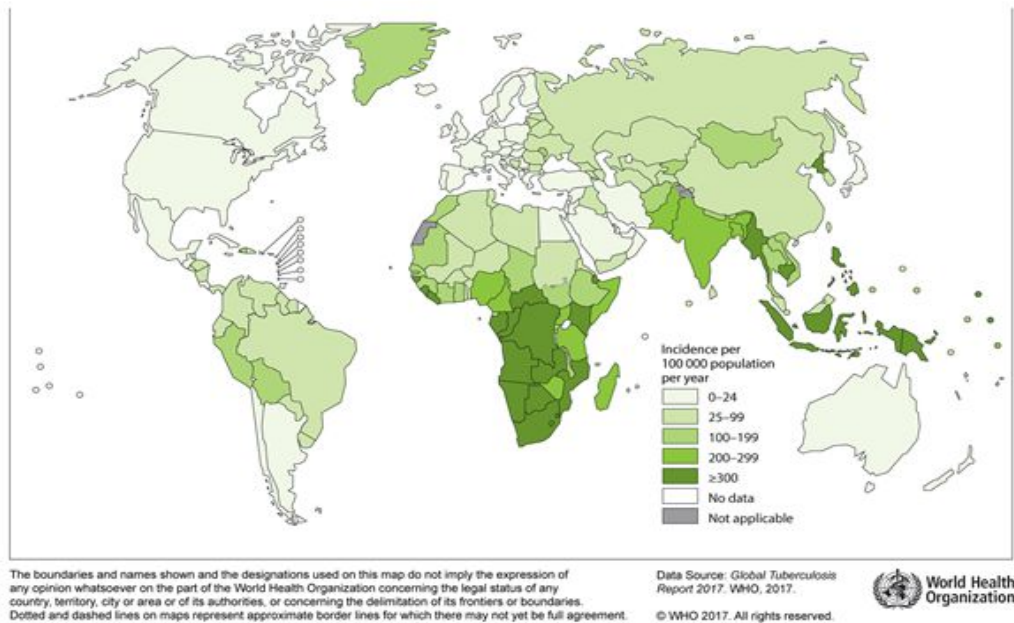


Fig. A

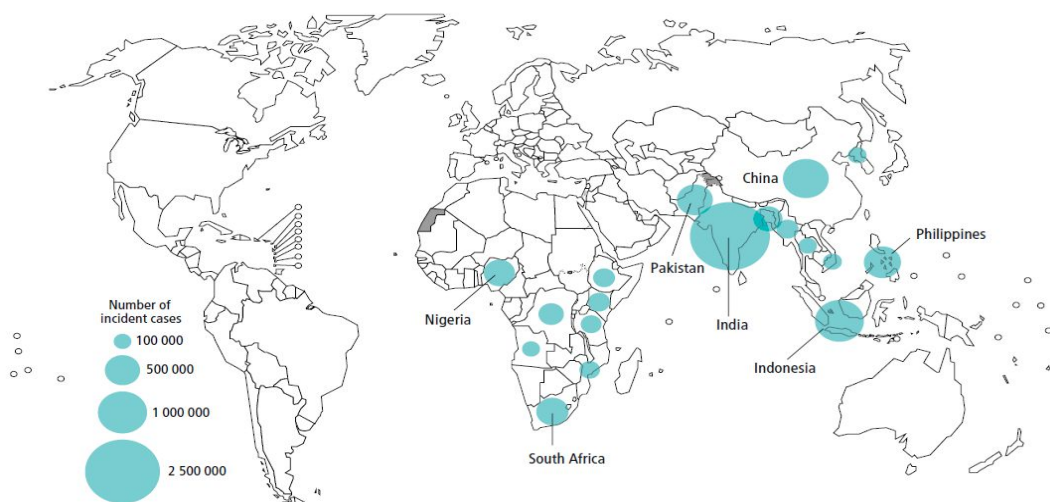


Fig. B

(Voir page suivante)



Fig. C

Figure 8 : Nombre de cas de TB au plan mondial en 2016 sur tous les continents

A- Estimation du taux d'incidence de la TB (par 100.000), 2016; B- Estimation de l'incidence de la TB (nombre de cas) en 2016 pour les pays avec au moins 100000 nouveaux de case; C- Estimation de l'incidence des MDR/RR TB en 2016 pour les pays avec au moins 1000 cas. (FIN)

Les cinq pays aux taux de nouveaux cas, les plus élevés sont l'Inde, l'Indonésie, la Chine, les Philippines, le Pakistan. Ils partagent 56% de nouveaux cas, parmi lesquels 45% de cas sont confirmés (Inde, Indonésie et Chine). L'incidence de la TB au Nigéria et en Afrique du Sud est de 4% de l'incidence mondiale (**Fig. 8B**, (WHO 2017)).

L'incidence de cas de tuberculose est en baisse dans les populations vulnérables qui n'ont pas accès à un service de soins abordable en prix, adéquat et de qualité.

L'OMS a développé une stratégie pour mettre fin à l'épidémie de la tuberculose d'ici 2016-2035, en se basant sur le développement durable. Parmi ses principaux buts, pour éradiquer la tuberculose, il faut mettre fin à la pauvreté sous toutes ses formes; améliorer la nutrition, assurer une vie saine et promouvoir le bien-être de tous.

1.1.9 *Epidémiologie et évolution de la résistance aux antibiotiques*

La résistance de *M. tuberculosis* aux antibiotiques est en pleine croissance par transmission de bacilles résistants, ou résultant d'une mauvaise application, d'une mauvaise qualité des agents thérapeutiques, ou d'une interruption prématurée du traitement par le patient (mauvaise observance). En 2016, l'OMS estime que 600 000 nouveaux cas présentent une résistance à la rifampicine dont 490 000 sont des cas de tuberculose multi-résistante. La tuberculose multi-résistante est en croissance dans trois pays, la Chine, la Russie et l'Inde, ce qui représente près de la moitié des cas dans le monde (**Fig. 8C**). En 2016, près de 6,2% des cas de tuberculose-MDR avaient en fait une tuberculose-XDR et actuellement dans le monde, seulement 54% des cas de tuberculose-MDR et 34% des cas de tuberculose-XDR sont traités avec succès.

1.1.10 *Tuberculose et coinfection VIH*

La coinfection tuberculose VIH survient quand un patient VIH+ a une tuberculose soit latente ou active. La chute du taux des CD4 suivie de l'augmentation de la charge virale est une cause pour le déclenchement d'une tuberculose active. Selon l'OMS, la prévalence des personnes infectées par le VIH dans les nouveaux cas de tuberculose est de 71% en Afrique, le continent le plus touché par la coinfection VIH/TB. Le traitement de la tuberculose est débuté dans les plus brefs délais chez les personnes infectées par le VIH et est similaire à celui d'un sujet séronégatif mais l'OMS recommande une prise quotidienne durant la phase d'initiation. L'initiation des antirétroviraux deux semaines après le début du traitement de la tuberculose améliore significativement la survie des adultes infectés par le VIH avec un taux de CD4+ de 200/mm³ ou plus bas ((Blanc, Sok et al. 2011)).

En se basant, sur les différences phénotypiques observée en culture, la résistance aux antituberculeux, l'épidémiologie des souche MTBc, il est important d'étudier la diversité génomique de ces souches.

1.2 *Le complexe Mycobacterium tuberculosis (MTBc)*

1.2.1 *Génome de Mycobacterium tuberculosis*

Le séquençage du génome complet de la souche de référence H37Rv, première souche de *Mycobacterium tuberculosis* séquençée (Cole 1999) a permis d'approfondir les connaissances d'une part sur l'organisation génétique du bacille mais aussi sur la compréhension des mécanismes biochimiques, physiologiques, génétiques et immunologiques entre la bactérie et

l'hôte. Le génome de H37Rv comprend 4 411 529 bp avec environ 4 000 gènes (50 gènes encodent des ARNr/ARNt et 3924 des protéines), 65% de GC et un grand nombre de séquences répétées (IS6110/MIRU-VNTR, etc...) (Cole 1999). Les génomes des mycobactéries tuberculeux hormis celui de *M. « canettii »* sont à 99,9% identiques. Et on parle de complexe d'espèces de *Mycobacterium tuberculosis* (MTBc) car la subdivision en espèces n'est pas pertinente du point de vue génomique mais bien fondée au niveau écologique (puisque très peu de souches *M. tuberculosis* sont isolées sur le bétail et très peu de *M. bovis* sont isolés sur l'Homme).

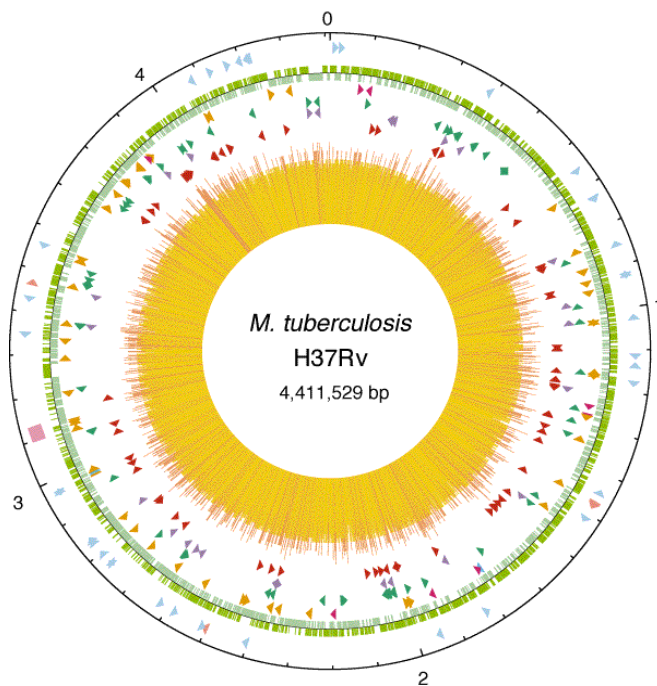


Figure 9 : Carte circulaire du chromosome de *M. tuberculosis* souche H37Rv. 0_ début de la réplication, 1^{er} anneau de l'extérieur: positions des gènes des ARN stables (bleu-ARNt, rose - autres), cube rose- région de répétition directe (CRISPR, Clustered Regularly Interspaced Short Palindromic Repeats); le 2^{ème} anneau montre les positions des séquences codantes sur chaque brin (sens horaire, vert foncé; dans le sens antihoraire, vert clair); le 3^{ème} anneau représente l'ADN répétitif (séquences d'insertion, orange; famille 13E12 REP, rose foncé; prophage, bleu); le 4^{ème} anneau montre les positions des membres de la famille de Proline-Proline-Acide Glutamique « PPE » (vert); le 5^{ème} anneau montre les membres de la famille Proline-Acide Glutamique « PE » (violet, à l'exclusion PGRS); et la 6^{ème} bague montre les positions des PGRS séquences (rouge foncé). L'histogramme (centre) représente le contenu en G + C, avec <65% G + C en jaune, et > 65% G + C en rouge. Elle a été produite avec le logiciel de DNASTAR (* S. Cole et al. 1998).

1.2.2 Macroévolution de *Mycobacterium tuberculosis* et transmission

La comparaison de données génomiques de *M. tuberculosis* isolés de fossiles humains aux données phylogénétiques des isolats contemporains permet de retracer l'origine, l'âge, la distribution géographique, la microévolution, la transmission, et l'interaction hôte-pathogène (Donoghue 2011). Ainsi, la tuberculose existait déjà en Turquie Byzantine, à Bornéo avant la colonisation Européen (Donoghue 2011), en pré-Colombie (Wilmar L. Salo, Arthur C. Aufderheide et al. 1994) et des ADN de MTBc ont été ensuite identifiés en Egypte ancienne (Zink, Haars et al. 2001), en Chine, en Amérique et en Europe (AR Zink, U Reischl et al. 2002; Jordi Gómez i Prat and Souza 2003; Donoghue HD, Spigelman M et al. 2004; Drancourt and Raoult 2005; Donoghue 2008; Donoghue 2009).

L'âge du plus récent ancêtre commun du complexe MTBc a été estimé à 70 000 ans ; il a été évalué par des méthodes génétiques en prenant comme support un ancêtre commun de *M. tuberculosis* originaire de l'Afrique qui se serait propagé durant les migrations des populations humaines (Comas, Coscolla et al. 2013). Cette hypothèse spéculative sur la datation et l'origine n'est pas totalement convaincante et requiert plus d'analyse. Une autre datation de l'ancêtre commun le plus récent de MTBc est estimée à 5 000-6 000 ans ((Bos, Harkins et al. 2014), (Kay, Sergeant et al. 2015), (Achtman 2016)).

La transmission de la tuberculose s'est faite au cours du temps avec les migrations de personnes d'un pays à un autre ou d'un continent à un autre. Avant on pensait que la tuberculose était une maladie bovine et que l'homme avait été infecté par l'animal. Mais la génomique démontre que *M. Bovis* dérive de la branche de *M. tuberculosis* du complexe MTBc (**Fig 10**) ((Brosch, Gordon et al. 2002), (Baker, Lee et al. 2015)).

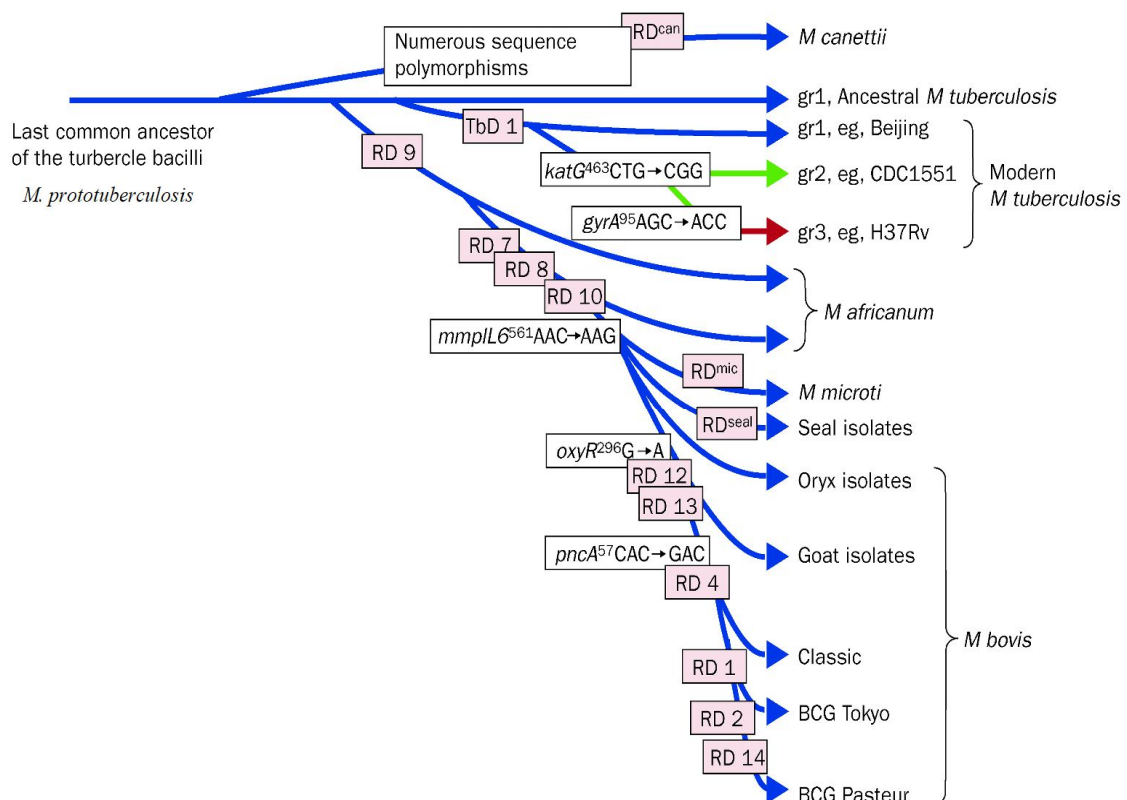


Figure 10 : Illustration de l'évolution du bacille *Mycobacterium tuberculosis*

Les flèches bleues indiquent que les souches sont caractérisées par *katG*⁴⁶³ CTG (Leu), *gyrA*⁹⁵ ACC (Thr), typique pour les organismes du groupe 1. Les flèches vertes indiquent que les souches appartiennent au groupe 2 caractérisé par *katG*⁴⁶³ CGG (Arg), *gyrA*⁹⁵ ACC (Thr). La flèche rouge indique que les souches appartiennent au groupe 3, caractérisé par *katG*⁴⁶³ CGG (Arg), *gyrA*⁹⁵ AGC (Ser), tel que défini par Sreevatsan et al. ((Brosch, Gordon et al. 2002), (Sreevatsan, Pan et al. 1997), (Donoghue 2011)).

Cependant, *M. bovis*, un agent dérivé de *Mycobacterium tuberculosis* responsable de la tuberculose bovine est aussi responsable de la tuberculose humaine : 2,8% et 1,4% de cas de tuberculose zoonotique en Afrique et hors de l'Afrique ((Müller, Dürr et al. 2013)).

Pour distinguer chaque espèce du complexe MTBc, des méthodes de génotypage ont été développées.

1.3 Génotypage de *Mycobacterium tuberculosis* : Méthodes de génotypage

Le génotypage est une technique qui analyse des locus variables et hypervariables, et/ou des polymorphismes de nucléotides simple (SNP). Ce type de technique peut aider à étudier la transmission des souches au sein d'une population. Hormis le séquençage partiel ou du génome complet, les méthodes de génotypage de *M. tuberculosis* utilisées sont le IS6110-

RFLP, le spoligotypage et le typage MIRU-VNTR, l'analyse des régions de différences et du polymorphisme de nucléotides simples.

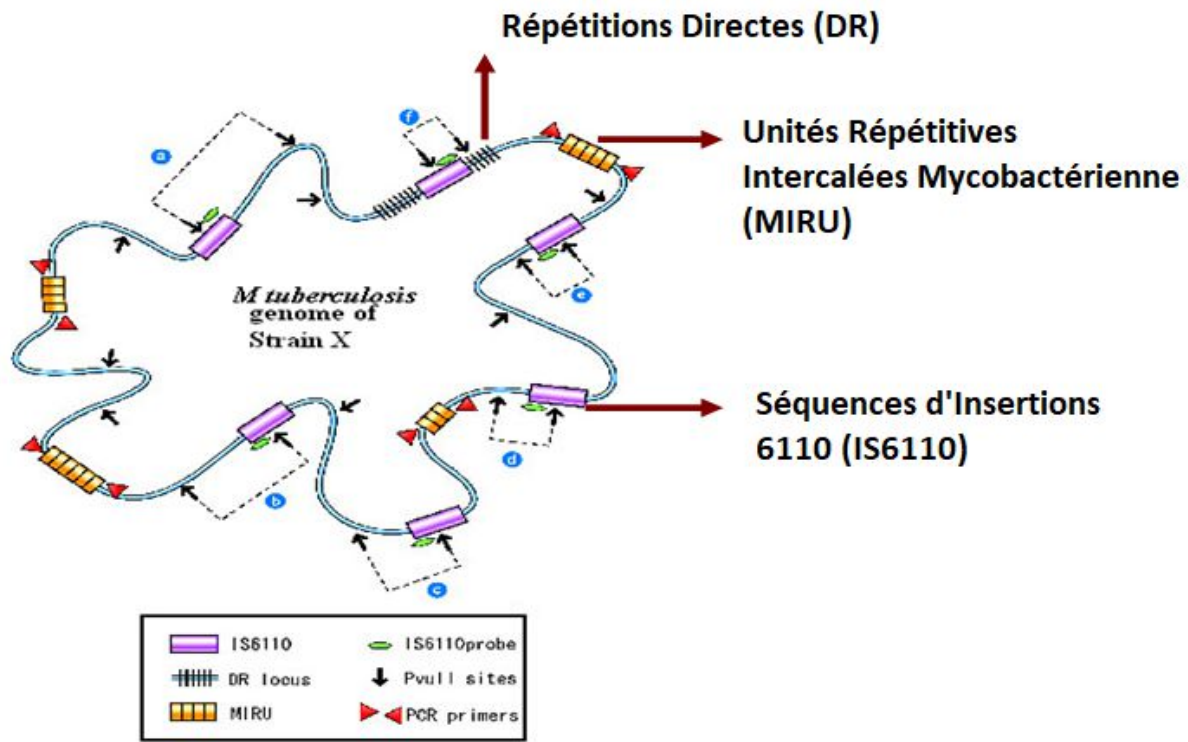


Figure 11 : Marqueurs de génotypage des souches de *Mycobacterium tuberculosis* (Barnes and Cave 2003).

1.3.1 IS6110-RFLP

IS6110 est un élément génétique mobile spécifique dans le génome de MTBc chez d'autres mycobactéries testées (Dominique Thierry, Anne Brisson-Noel et al. 1990). Très peu de souche de MTBc sans insertion IS6110 ont été reportées. Le nombre de copies et les sites d'emplacement de l'insertion IS6110 varient d'une souche à une autre. La méthode de RFLP (Restriction-Fragment-Lenght-Polymorphism) sur la séquence d'insertion IS6110 (IS6110-RFLP) est une méthode de génotypage basée sur le polymorphisme (polymorphisme causé par des délétions ou duplications) du nombre de copies de IS6110 et de leurs sites d'insertion dans le génome du complexe *M. tuberculosis*. Le principe est de séparer sur gel d'agarose l'ADN digéré par une enzyme de restriction (PvuII) puis de révéler les fragments porteurs de IS6110 par hybridation selon la méthode de *Southern* (Fig 12).

Les souches typiques de *M. tuberculosis* ont dans leur génome une ou plusieurs copies d'IS6110 (jusqu'à 25 copies). La souche bovine *M. bovis* a en général un nombre faible de

copies (1 à 2 copies d'IS6110). Dans les souches de *M. tuberculosis* et *M. bovis*, une copie de IS6110 est toujours présente dans la zone du locus CRISPR.

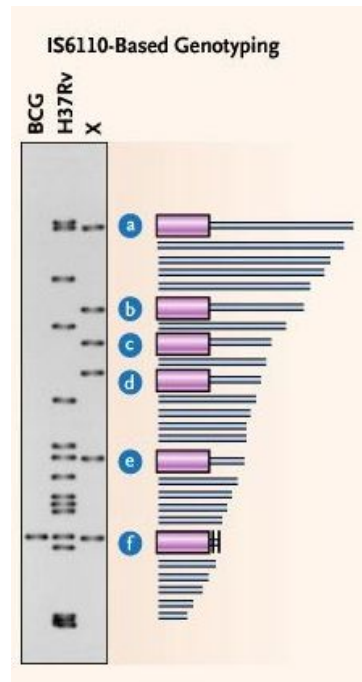


Figure 12 : Exemple de résultat de IS6110-RFLP. a, b, c, d, e, f indiquent les fragments de la souche H37Rv porteurs de IS6110 révélés par Southern (Barnes and Cave 2003).

Cette méthode a longtemps été la méthode de génotypage de référence de MTBc pour l'épidémiologie moléculaire, mais elle est très laborieuse, son pouvoir discriminant est médiocre pour les souches qui comportent un faible nombre de copies de IS6110 ; elle exige par ailleurs de grandes quantités d'ADN. Le fait que l'Europe ait choisi comme souche de référence Mt14323 tandis que les USA conservaient H37Rv, que la technique soit laborieuse à mettre en œuvre, n'a pas facilité l'interchangeabilité des résultats IS6110-RFLP entre laboratoires. Ainsi la méthode IS6110-RFLP a été remplacée progressivement par la combinaison spoligotypage et MIRU-VNTR ((Dick Van Soolingen, Lishi Qian et al. 1995), (Barnes and Cave 2003), (Supply, Allix et al. 2006), (Aze, Sola et al. 2015)).

1.3.2 Typage des espaceurs dans le locus CRISPR – Spoligotypage

Le spoligotypage est une méthode de génotypage par PCR et hybridation inverse basée sur le polymorphisme du locus CRISPR ((Kamerbeek, Schouls et al. 1997)). Le locus CRISPR est une suite de séquences répétées identiques d'environ 21 à 48 bp espacées par des espaceurs (spacers en anglais) de séquences uniques (variant en 26 et 72 bp). Ce type de séquence est présent dans environ 40% des eubactéries et dans plus de 90% d'archéobactéries

((Grissa, Vergnaud et al. 2007; Grissa, Bouchon et al. 2008)). Dans le complexe *Mycobacterium tuberculosis*, il existe un seul locus CRISPR dont la séquence répétée fait 36 bp et la taille des espaceurs varie entre 35 et 41 bp. Il n'acquiert plus de nouveaux espaceurs et son évolution se fait par délétion (délétion de blocs distincts d'espaceurs) soit par recombinaison homologue entre les DR adjacents ou distants soit par transposition d'éléments IS6110 ((Hancock 1995), (Fang, Doig et al. 1999)). Ce marqueur est polymorphe (présence ou absence d'espaceurs) chez MTBc avec une certaine stabilité et sert de marqueur de génotypage. Sur cette base, au cours du spoligotypage, des sondes sont hybridées à des amplicons chevauchant des espaceurs (produits par PCR) (**Fig. 13**). Cette méthode est utile pour les études évolutives et d'épidémiologie portant sur *M. tuberculosis*, qu'il s'agisse de déterminer si la transmission est fortuite (« casual ») ou de ménage (« household ») dans les populations (Barnes and Cave 2003). Mais, elle n'est pas suffisamment discriminante pour être utilisée seule comme méthode d'investigation épidémiologique. Elle est donc toujours associée à d'autres méthodes, le plus souvent le typage MIRU-VNTR.

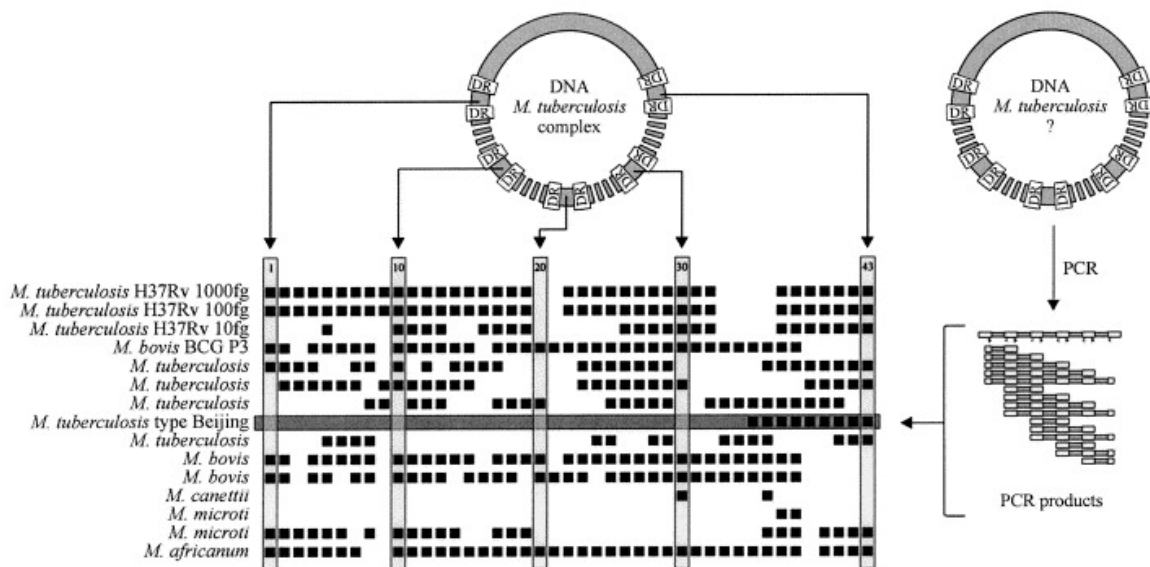


Figure 13 : Le spoligotypage (Bernes et al 2003)

1.3.3 Typage MIRU-VNTR

Les MIRU (Mycobacterial Interspersed Repetitive Units) sont des éléments aussi dits VNTR ou minisatellites (séquences Répétées en Tandem de 40-100bp en Nombre Variable, en anglais Variable Number Tandem Repeats) dispersés dans les régions géniques ou inter-géniques du génomes MTBc. Le typage MIRU-VNTR est basé sur le polymorphisme des loci VNTR. Le principe consiste en une amplification par PCR des VNTR avec des amorces

flanquantes spécifiques à chaque locus. Après amplification, les fragments sont analysés par électrophorèse pour déterminer la taille de chaque amplicon (**Fig 14**). En fonction de la taille du fragment décelé, un nombre de répétitions est attribué à chaque locus. Un ensemble de 12, 15 ou 24 loci VNTR permettent le génotypage de *Mycobacterium tuberculosis* (Supply, Allix et al. 2006). L'utilisation de 24 loci augmente le nombre de types de profils de 40% et de 23% en combinaison avec le spoligotyping comparé à ceux obtenus par 12 loci. La valeur prédictive pour l'évaluation de la transmission de *M. tuberculosis* à base de 15 locus discriminants est égale à celle du typage par IS6110-RFLP. Ainsi le schéma à 15 loci a été proposé comme standard pour la discrimination épidémiologique en routine des isolats de *M. tuberculosis* et les 24 loci comme outils de typage à haute résolution pour les études phylogénétiques (Supply, Allix et al. 2006). Le typage de 24 locus VNTR reste assez laborieux à mettre en pratique et l'analyse des résultats longue et sujette à erreurs.

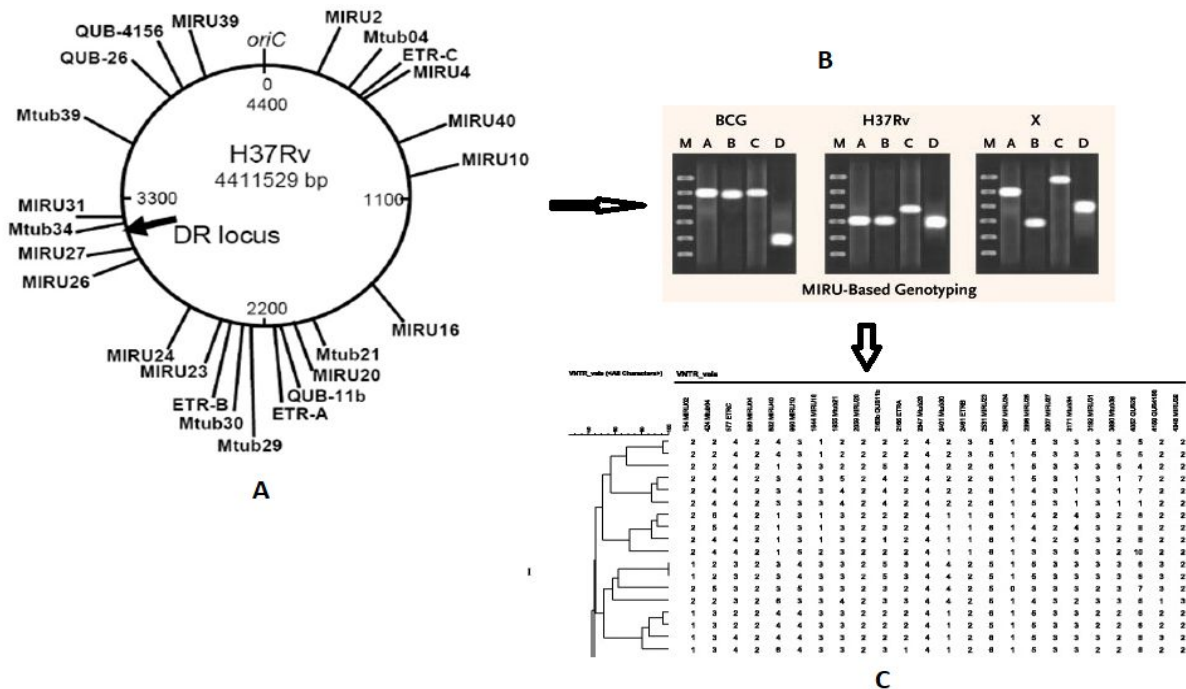


Figure 14 : Méthode de génotypage MIRU-VNTR, A-loci MIRU-VNTR contenus dans le génome de H37Rv, B- électrophorèse sur gel d'agarose d'amplicons de loci VNTR pour 3 souches (BCG, H37Rv, X), C- interprétation phylogénétique de résultats de MIRU-VNTR (Valcheva, Mokrousov et al. 2008).

1.3.4 Analyse des régions de délétions

La présence ou absence de larges séquences (souvent dites « séquences ou régions de délétion ») constitue des marqueurs spécifiques pour la caractérisation, la classification des souches de *M. tuberculosis* dans les lignées (Tsolaki A. et al 2005). Ces délétions dans les mycobactéries sont des événements irréversibles, car il n'y a pas de transfert horizontal de gènes chez *M. tuberculosis*. Une première étude a identifié 68 régions de délétion dans les isolats du complexe *M. tuberculosis*. En se basant sur les régions de délétion, six lignées de *M. tuberculosis* ont été définies (**Fig. 10**) (Gagneux, DeRiemer et al. 2006). Ces pertes successives d'ADN ont pu contribuer à des expansions clonales et des adaptations successives à de nouveaux hôtes (Brosch, Gordon et al. 2002).

1.3.5 Analyse des génomes entiers et Typage à base des polymorphismes de nucléotides SNP

Le polymorphisme de nucléotide simple (SNP) est un des marqueurs les plus performants pour l'étude de la phylogénétique des souches de MTBc. Avec les nouvelles technologies de génomique telles que le séquençage de nouvelle génération (NGS) un panel de SNP fut découvert récemment dans les génomes des souches de MTBc. L'étude des SNP au sein de MTBc permet d'améliorer notre compréhension de la diversité génétique, et de l'épidémiologie de ces souches ((Comas, Homolka et al. 2009), (Coll, McNerney et al. 2014)).

1.4 Classification de *Mycobacterium tuberculosis*

Comme précédemment expliqué, les méthodes de génotypage permettent de surveiller les épidémies. Sur le plan mondial, en se basant sur ces méthodes, sept lignées des souches de *M. tuberculosis* ont été identifiées avec des outils d'analyse et ont été reportés dans les bases de données internationales en ligne (TBminer, MIRU-VNTRplus, SITVITWEB) (Hershberg, Lipatov et al. 2008). Les lignées L1-L7 présentent des distributions géographiques spécifiques (**Fig. 15**) ((Comas, Homolka et al. 2009), (Aze, Sola et al. 2015), (Weniger, Krawczyk et al. 2010), (Demay, Liens et al. 2012)).

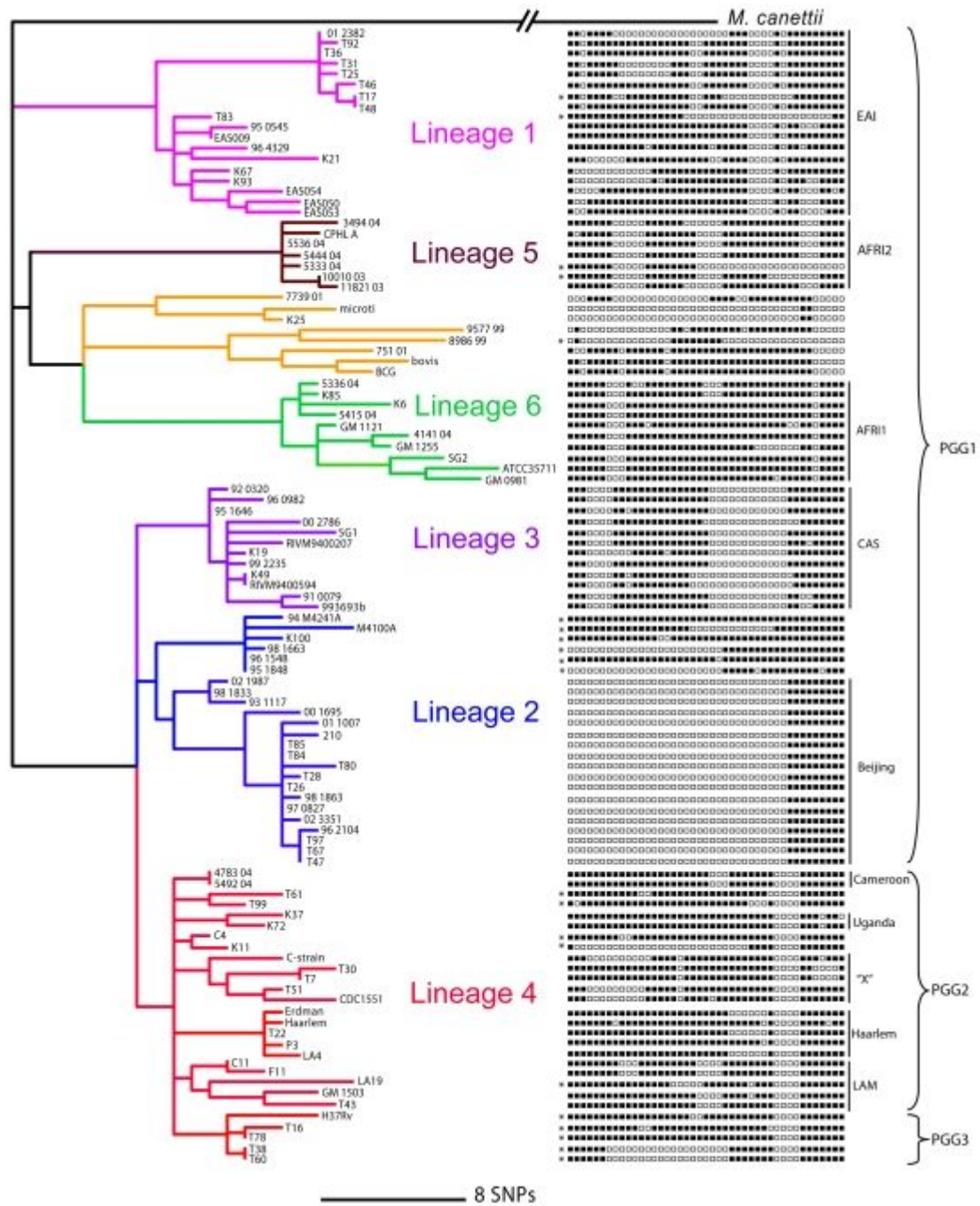


Figure 15: Arbre phylogénétique de 89 séquences de gènes provenant de 108 souches MTBc montrant la structure de la population du complexe *M. tuberculosis* avec les différentes lignées phylogénétiques. Lineage 1 – EAI (East-African-Indian), Lineage 2 – Beijing, Lineage 3 – CAS (Central Asia), Lineage 4 – Euro-American, Lineage 5 – Africanum 1, Lineage 6 – Africanum 2. La barre d'échelle indique le nombre de SNPs. ((Comas, Homolka et al. 2009).

1.5 La lignée 2 « Beijing » de *Mycobacterium tuberculosis*

1.5.1 Diversité et importance de la lignée 2 « Beijing »

Les souches de la lignée 2 tiennent leur origine de la région de Guangxi (55% de souches Beijing par rapport aux souches totales obtenues dans toutes les régions), où 16.5% de ces souches Beijing ont la délétion RD181, marqueurs des souches les plus anciennes (Wan, Liu et al. 2011). La lignée 2 est subdivisée en Beijing et proto-Beijing en fonction du nombre de gènes délétés dans la région RD105. La lignée L2/Beijing regroupe un nombre important de souches autres que la L2/Proto-Beijing.

Les souches de *M. tuberculosis* du génotype Beijing incluant les protos-Beijing sont responsables de plus d'un ¼ de l'épidémie globale de la tuberculose. Ces souches sont très prédominantes dans l'Est de l'Asie, elles se sont répandues à travers le monde et sont associées à des épidémies de souches résistantes (Judith R. Glynn, Kristin Kremer et al. 2006). En particulier, ces souches sont responsables d'au moins sept épidémies tuberculeuses en Union Soviétique, en Corée du Sud (Souche K), aux Etats-Unis (Souche 210 et HN878), aux Canaries (souche aux « Canaries » GC1237). L'analyse d'une collection de souches Beijing par le typage MIRU-VNTR a permis leur classification en complexes clonaux. La distribution de ces souches en complexes clonaux et lignée basique (CC1 à CC6, BL7) révèle une diversité importante en Asie de l'Est et dans les régions éloignées d'Asie de l'Est. Les souches du complexe clonal CC5 montre la prévalence de complex clonal vers l'Océan Pacifique, spécifiquement en Micronésie et Polynésie. Les complexes clonaux CC1 et CC2 sont hautement dominants et en pleine croissance en Asie Centrale, et dans les pays bordant la mer noire (CC1-Russie, CC2-Europe de l'Est). Les souches épidémiques identifiées dans les Etats de l'Union Soviétique se regroupent dans ces deux complexe clonaux (Merker et al 2015). Les souches Beijing des CC3-CC4, CC6-BL7 sont minoritaires parmi les génomes entiers séquencés, et présentes en Asie et en Afrique (**Fig. 16**). La particularité des souches Beijing par rapport aux autres génotypes se situe dans l'augmentation de leur capacité à acquérir la résistance aux antibiotiques, liée à l'hypermutableté ou à la présence des mutations compensatrices, dans leur transmissibilité, et leur hypervirulence, et une progression plus rapide de la maladie après l'infection.

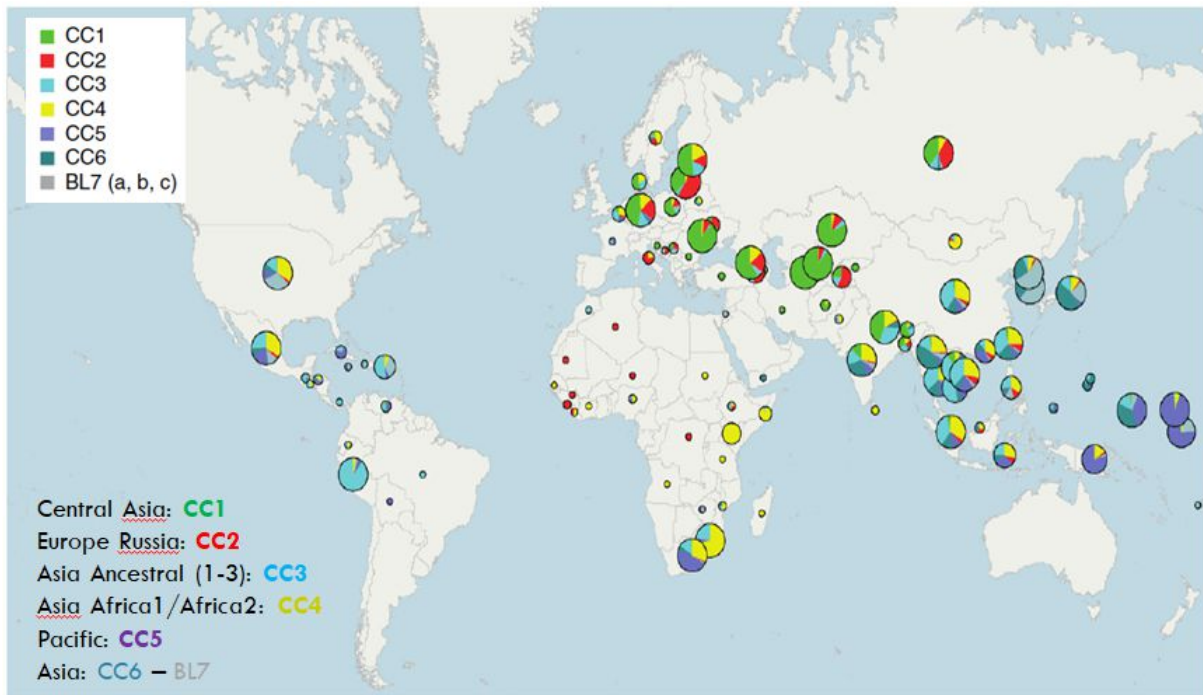


Figure 16 : Distribution mondiale des souches Beijing en complexe clonaux (CC) et Lignée Basiques (BL) (Merker et al 2015).

1.5.2 Pathogénicité et virulence des souches L2/Beijing

Les souches de génotypes Beijing sont plus virulentes que les souches des autres lignées de *M. tuberculosis* sur modèle animal ((Ida Parwati, Reinout van Crevel et al. 2010)) ; cette virulence supérieure a été confortée chez l'homme seulement par une partie des études. Cette virulence pourrait être liée à des propriétés biochimiques intrinsèques de ces souches et leur interaction avec le système de défense de l'hôte humain, ou bien à des différences dans la structure de leur parois cellulaires associées aux lipides, ou encore de l'induction de réponses immunitaires moins efficaces.

Les souches Beijing augmentent l'expression de la protéine homologue α -crystalline (16 kDa), un facteur de virulence de *M. tuberculosis*, et diminuent l'expression des protéines de choc thermique Hsp65, de transport de phosphate et de 47 kDa par rapport à d'autres isolats cliniques et à la souche de référence H37Rv (Pheiffer 2005). Deux autres protéines, la glycogène phosphorylase putative GlgP et l'haloalcane déhalogénase LinB ont une expression élevée après phagocytose de la souche K de *M. tuberculosis*. La forte expression de la protéine GlgP induirait l'activation du facteur inhibiteur de migration du macrophage, ce qui pourrait être avantageux pour les bactéries ((Ryoo, Young Kil Park et al. 2007)). Ces protéines pourraient jouer un rôle dans la pathogénèse de la tuberculose.

In vitro, les souches Beijing produisent une polykétide synthase (synthétise des lipides complexes biologiquement actifs –dérivés du PGL- glycolipide phénolique), ce qui inhibe la libération des médiateurs pro-inflammatoires et conduit à une infection mortelle d'animaux (Reed, Michael B. Reed et al. 2004). Ces souches accumulent aussi une large quantité de triacylglycérides en culture aérobie, coïncidant avec la régulation positive de *Rv3130c*, un membre du régulon de dormance DosR qui est impliqué dans la croissance en milieu microaérophilique et anaérobique durant l'infection ; ces avantages métaboliques pourraient être reliés aux succès des souches Beijing ((Reed, Gagneux et al. 2007), (Smith, Bochkareva et al. 2017), (Domenech, Zou et al. 2017)). Les mutations du PPE38 bloquent complètement la sécrétion de deux grands sous-ensembles de substrats ESX-5 (PPE-MPTR et PE_PGSR). La restauration de la sécrétion dépendant du PPE38 rétablit partiellement le phénotype normal chez une souche Beijing et la délétion de *ppe38* chez *M. tuberculosis* modérément virulent augmente sa virulence. Cela indique que ces substrats ESX-5 jouent un rôle important dans l'atténuation de la virulence (Ates et al 2018).

Dans la cellule monocyttaire THP-1 de leucémie aiguë humaine, l'infection avec des souches de génotype de Beijing (telles que la souche K) induit une faible production de plusieurs cytokines pro-inflammatoires : facteur de nécrose tumorale TNF, l'interleukine 6 et l'interleukine 12b ((Sohn 2009)Sohn et al 2002), ce qui favorise la croissance de *M. tuberculosis* (van Crevel et al 2002). Les cellules THP-1 infectées par la souche hyper virulente Beijing K, présentent un niveau significativement moins élevé d'apoptose que lorsqu'elles sont infectées par la souche de référence H37Rv ((Sohn 2009)). Ces résultats suggèrent que les souches Beijing inhibent l'apoptose cellulaire, un mécanisme de défense de l'hôte qui limite la croissance des mycobactéries dans les macrophages, tout en induisant la nécrose, qui favorise la prolifération des mycobactéries.

1.5.3 Diversité des souches de la lignée L2/« Beijing » en fonction des diverses méthodes de génotypage

1.5.3.1 Diversité des profils CRISPR (spoligotypes) des souches de la lignée L2/Beijing

Le spoligotypage classique contient 43 espaceurs. Le génotype des souches Beijing est caractérisé par l'absence des 34 premier espaceurs et la présence des 9 restants. Les souches de Beijing ont été identifiées partout dans le monde bien que dans des proportions différentes et représentent 13% au plan mondial et 50% en Asie de l'Est. Cependant des études ont identifié des souches Beijing avec d'autres délétions au sein des 9 espaceurs restants qui ont été décrites dans la base de donnée international SITVIT (**Tableau 2**) ((Judith R. Glynn, Jennifer Whiteley et al. 2002), (Flores-Treviño, Morfín-Otero et al. 2015) 2015).

Les souches Beijing ont été classées en six complexes clonaux majeurs (CC1-CC6) et une sous-lignée basale (BL7) à l'aide des 24 loci MIRU-VNTR. Il a été démontré que les complexes clonaux CC1 à CC5 incluent les souches Beijing modernes ou typiques tandis que le complexe clonal CC6 et la sous lignée basique BL7 incluent les souches Beijing ancestrales et atypiques (Merker, Blin et al. 2015). Le MIRU 24 a permis le sous typage et la discrimination des souches du génotype Beijing, et aussi l'identification des clones les plus dominants dont les complexes clonaux 100-32 et 94-32 (Vyazovaya, Mokrousov et al. 2015). Ainsi le système de 15 locus est recommandé mais pour des études plus précises, le MIRU 24 est indispensable.

1.5.3.3 Les régions de délétions

Les souches Beijing sont définies et sous-classées par des régions dites « régions de délétions » ou RD, qui sont absentes dans certaines sous-lignées (Kremer, Glynn et al. 2004; Tsolaki, Gagneux et al. 2005). La délétion RD105 a été trouvée dans toutes les souches Beijing et sert de marqueur pour l'identification de cette lignée. La classification phylogénétique subdivise la lignée 2 en deux clades soeurs: les sous lignées 2.1 et 2.2.

Les souches de la sous lignée 2.1 présentent la région RD207 intacte qui recouvre les premiers espaceurs du locus CRISPR, elles n'appartiennent pas à la famille Beijing *sensu stricto*, elles sont appelées « *proto-Beijing* ». Toutes les souches proto-Beijing ont une très large délétion qui recouvre la RD105, délétion différente de celle du reste des souches de la lignée 2. Cette large délétion est appelée « extended RD105 » (**Fig. 17**).

La sous lignée 2.2 déléetée pour RD207 se subdivise en fonction de l'absence (sous lignée 2.2.1) ou de la présence (sous lignée 2.2.2) de la région RD 181 (Wan, Liu et al. 2011). Parmi 55% de Beijing reportées dans la région Guangxi (Chine) et contrairement aux autres régions, 16.5% de ces souches Beijing montrent une absence de la région RD181. Au Japon, la proportion des souches Beijing ancienne (RD181+) est élevée avec un pourcentage de 5-10% ((Maeda, Hang et al. 2014), (Wada, Iwamoto et al. 2009), (Millet, Miyagi-Shiohira et al. 2012), (Yokoyama, Hachisu et al. 2010)). En Corée il est rapporté 45.3% de souches ayant la RD181+ parmi 64 souches de génotype Beijing (Kang, Wada et al. 2010). En conclusion, on pourrait affirmer que les souches Beijing ont une longue histoire dans le Sud de la Chine et qu'elles ont pu émerger de là (Wan, Liu et al. 2011).

La sous lignée 2.2 se divise en différentes branches selon l'existence des délétions en RD150, RD142, RD131 ((Tsolaki, Gagneux et al. 2005), (Schürch, Kremer et al. 2011; Luo, Tao Luo et al. 2015)).

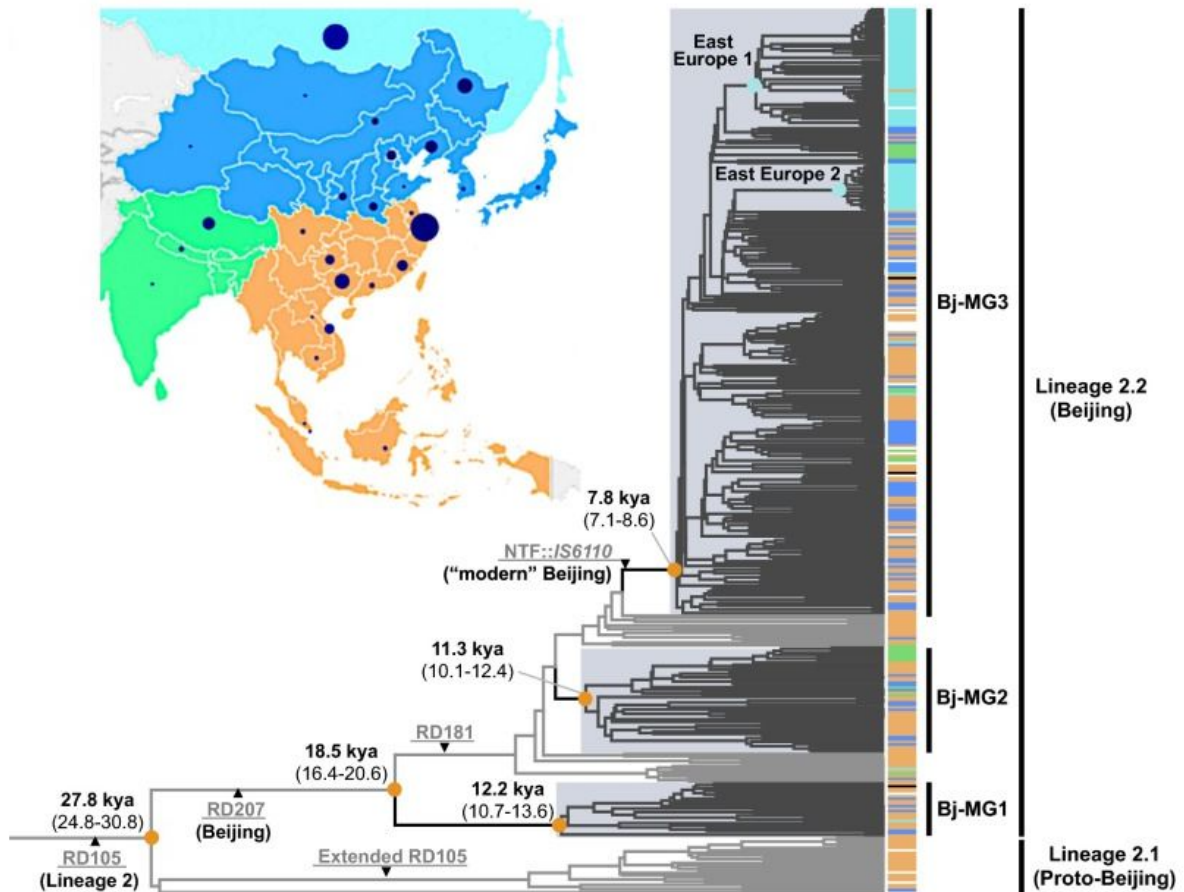
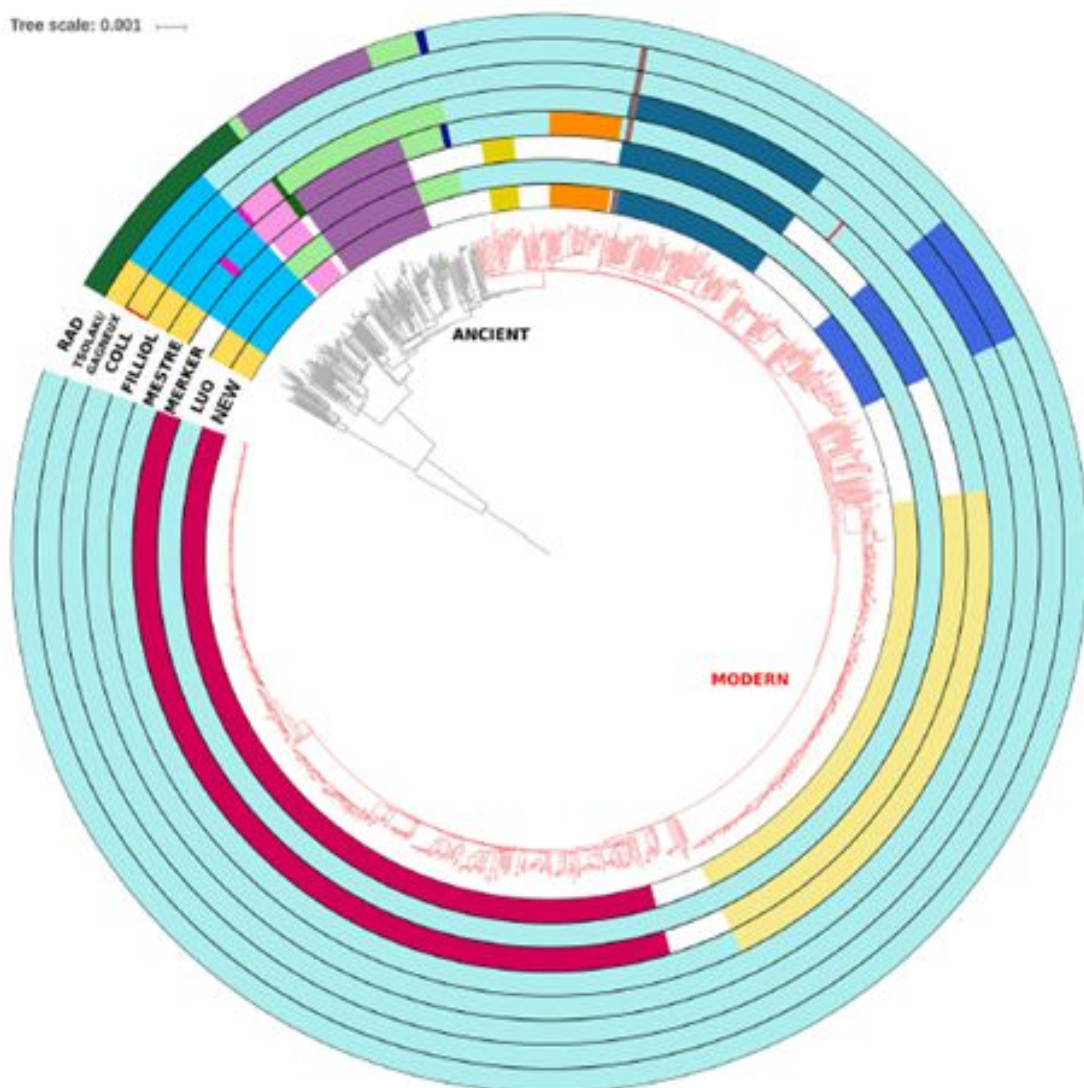


Figure 17 : Classification phylogénétique déduite des SNP de 358 souches de MTBc Proto-Beijing – délétion RD105+ étendue, Beijing groupe monophylétique 1 – RD207+, Beijing groupe monophylétique 2 – RD181+, Beijing groupe monophylétique 3 – RD181+ et 2 IS6110 dans le locus NTF (Luo, Tao Luo et al. 2015).

Ces différentes régions de délétion sont donc des marqueurs pour l'identification des souches Beijing et pour leur sous-classification. Les régions de différences sont aussi faciles à détecter en laboratoire que la détection des SNP, mais ne permettent pas de déterminer la distance génétique entre deux groupes de souches. Le pouvoir discriminant des RD sur la détermination des sous-lignées est limité et requiert une association aux SNP pour la détection des principaux clones épidémiques.

1.5.3.4 Polymorphisme de nucléotides simples (SNP)

Plusieurs études indépendantes ont cherché des SNPs capables de distinguer des clones au sein de la famille Beijing (Rad, Bifani et al. 2003; Filliol, Motiwala et al. 2006; Wan 2007; Mestre, Luo et al. 2011; Schurch, Kremer et al. 2011; Coll, Preston et al. 2014; Merker, Blin et al. 2015). Les isolats se caractérisent par des mutations parmi lesquelles, certaines dans les gènes *mutT2*, *MutT4* et *ogt*. Cependant, ces mutations ne sont pas suffisamment discriminantes pour une sous-classification fine. Plusieurs SNP ont été proposées pour comprendre la taxonomie fine des souches Beijing. L'analyse des différentes approches proposées a permis de comprendre le parcours évolutif et ainsi d'unifier la classification de la lignée 2 de *M. tuberculosis*. A partir du séquençage de génomes entiers (WGS) portant sur 1398 isolats de souches Beijing provenant de 32 pays, 39786 SNPs ont été identifiées et utilisées pour reconstruire la phylogénie de la lignée 2 de *M. tuberculosis* (Shitikov, Kolchenko et al. 2017) (Fig. 18).



New classification	Luo	Merker	Mestre	Filliol	Coll
proto-Beijing	proto-Beijing	Asia Ancestral 1	proto-Bmyc1	ST21 SCG3a (CAS)	2
Asia Ancestral 1	Bj-MG1	Asia Ancestral 2	Bmyc2	ST11	2.1
Asia Ancestral 2	Beijing	Asia Ancestral 3	Bmyc3	STK	2.2.2
Asia Ancestral 3	Bj-MG2	Asian African 1	Bmyc4	ST3	2.2.1
Asian African 1	Bj-MG3	Asian African 2	Bmyc6	proto-ST25	2.2.1.1
Asian African 2		Pacific RD150	Bmyc26	ST25	2.2.1.2
Asian African 2/RD142		Europe/Russia W148 outbreak	Bmyc25	ST10	
Asian African 3		Central Asia	Bmyc26/10	ST22	
Pacific RD150		Tsolaki/Gagneux	Bmyc10	ST8	
Europe/Russia W148 outbreak		gr1	Bmyc11	Rad	
Central Asia		gr2	Bmyc12	WT	
		gr3	Bmyc13	mutT4	
		gr4	Bmyc18	mutT4 ogt(37)	
		gr5		mutT2 mutT4	
				mutT2 mutT4 ogt(12)	

Figure 18 : Correspondance entre les classifications phylogénétiques de la lignée 2 de MTB (Shitikov, Kolchenko et al. 2017).

Cela confirme deux clades phylogénétiques majeures: la lignée 2.1 ou clade proto-Beijing caractérisée par une délétion étendue en RD105 (Luo, Tao Luo et al. 2015), qui appartient au ST21 SCG3a ((Filliol, Motiwala et al. 2006), (Shitikov, Kolchenko et al. 2017)). Le second clade phylogénétique appelé lignée 2.2 caractérisé par la délétion RD207 donne ensuite naissance à la lignée 2.2.2 avec les gènes mutT2, MutT4 et ogt intacts ((Coll, McNERNEY et al. 2014), (Luo, Tao Luo et al. 2015)). Un autre sous groupe de cette sous lignée est caractérisé par les mutations dans le gène MutT4 au niveau du codon 48 et dans le gène ogt au niveau du codon 37, correspondant au groupe Bmyc25, Bj-MG2 et Asia ancestral 3 selon les différents auteurs ((Mestre, Luo et al. 2011), (Luo, Tao Luo et al. 2015), (Mestre, Luo et al. 2011)). Un autre sous-groupe de souches ayant la mutation ogt37 correspond au sous groupe ST25 caractérisé par un autre panel de SNPs (Filliol, Motiwala et al. 2006). Les isolats Beijing se distinguent en souches « atypique » et « typique » Beijing grâce à une SNP spécifique ogt12C-T (Mina Ebrahimi Rad, Pablo Bifani et al. 2003). Avec l'arbre phylogénétique reconstitué de la lignée Beijing, il est possible de sélectionner des SNP spécifiques à chaque sous-groupe pour pouvoir effectuer simultanément une identification rapide, la détection des clones épidémiques, la surveillance, et l'étude de la transmission récente des isolats Beijing dans les régions à forte prévalence et/ou en cas de suspicion d'une épidémie émergente liée à ces souches particulières.

1.5.4 La séquence d'insertion IS6110

La séquence d'insertion IS6110 est composée de nombreuses copies réparties dans tout le génome des souches du génotype Beijing en particulier dans la région intergénique Rv2664-Rv2665 (Fig. 19). La forte variabilité du nombre de copies, la localisation d'IS6110 ainsi que la stabilité d'IS6110 au cours du temps ont fait de la séquence d'insertion IS6110 un marqueur de référence pour le diagnostic, l'épidémiologie et la différenciation entre les souches.

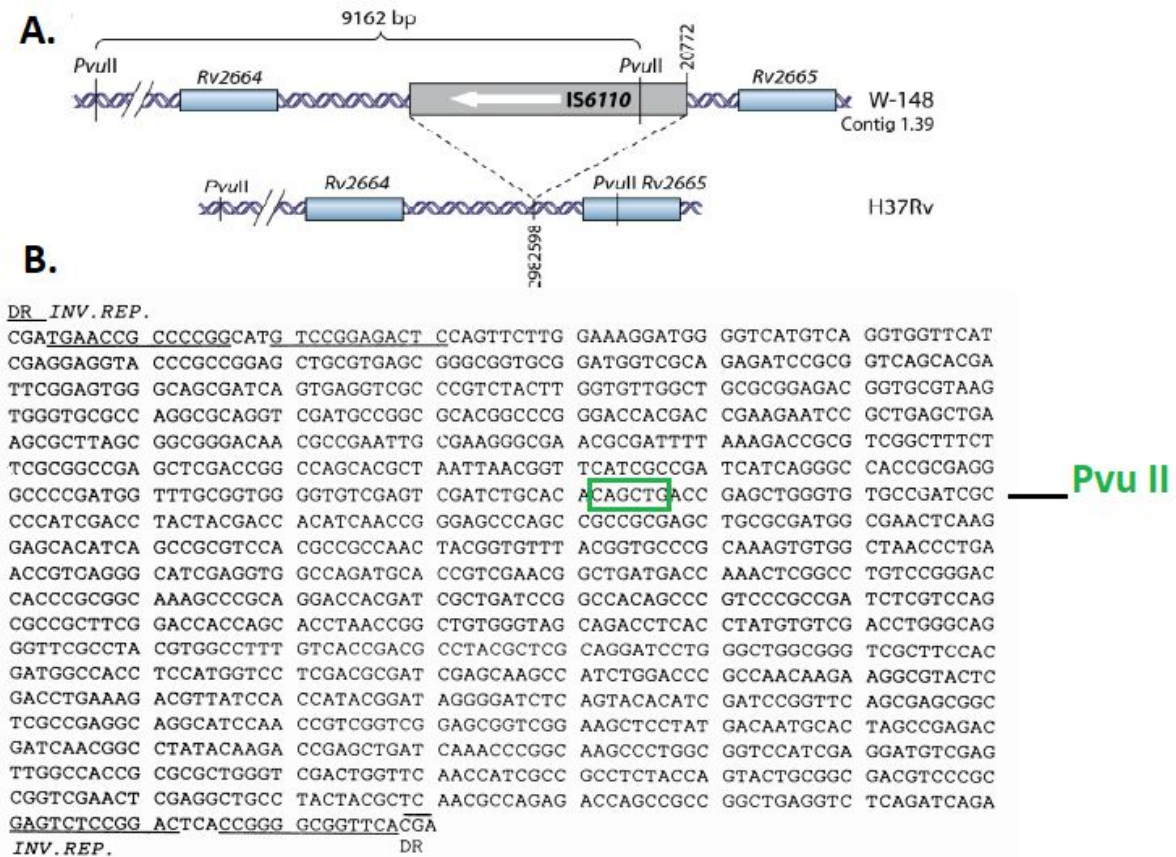


Figure 19 : Séquence d'Insertion IS6110, A- région intergénique Rv2664-Rv2665 contenant IS6110, B- séquence nucléotidique complète d'IS6110 (D. Thierry, M.D.Cave et al. 1990)

1.5.4.1 Localisation des copies de IS6110 dans les diverses régions du génome

Les souches de *M. tuberculosis* contiennent souvent dans leur génome plusieurs copies d'IS6110. Celles de la famille Beijing sont caractérisées par un nombre important de copies d'IS6110 (15-26 copies). Une séquence inversé d'IS6110 (position 19418 à 20772 bp) a été trouvée dans un fragment de digestion PvuII de 9162 bp, et cette séquence est située dans une région intergénique entre les gènes Rv2664 et Rv2665 (Fig. 19). Toutes les souches de la famille Beijing ont par ailleurs en commun une copie d'IS6110 au point d'insertion oriC

connu comme l'insertion A1, entre les gènes Rv1754c-Rv1765c, correspondant à la délétion RD152. Six sites d'insertions sont présents avec une haute fréquence dans les gènes Rv1371, *ctpD* (Rv1469c), Rv2016 et *idsB* (Rv3383c), entre deux IS1532, et au sein de la région *esxR-esxS* (Fig. 20) (Henar Alonso, Sofia Samper et al. 2013). On note également la présence d'une ou deux copies dans la région intergénique *dnaA:dnaN*, mais cette présence d'IS6110 n'est pas exclusive aux souches Beijing ((Henar Alonso, Sofia Samper et al. 2013), (N. E. Kurepina, S. Sreevatsan et al. 1998)). Aussi, dans le génome des souches de génotype Beijing, il est fréquent de trouver une ou deux copie d'IS6110 au sein de la région NTF (Noise TransFer Region, (Liu, Luo et al. 2016)).

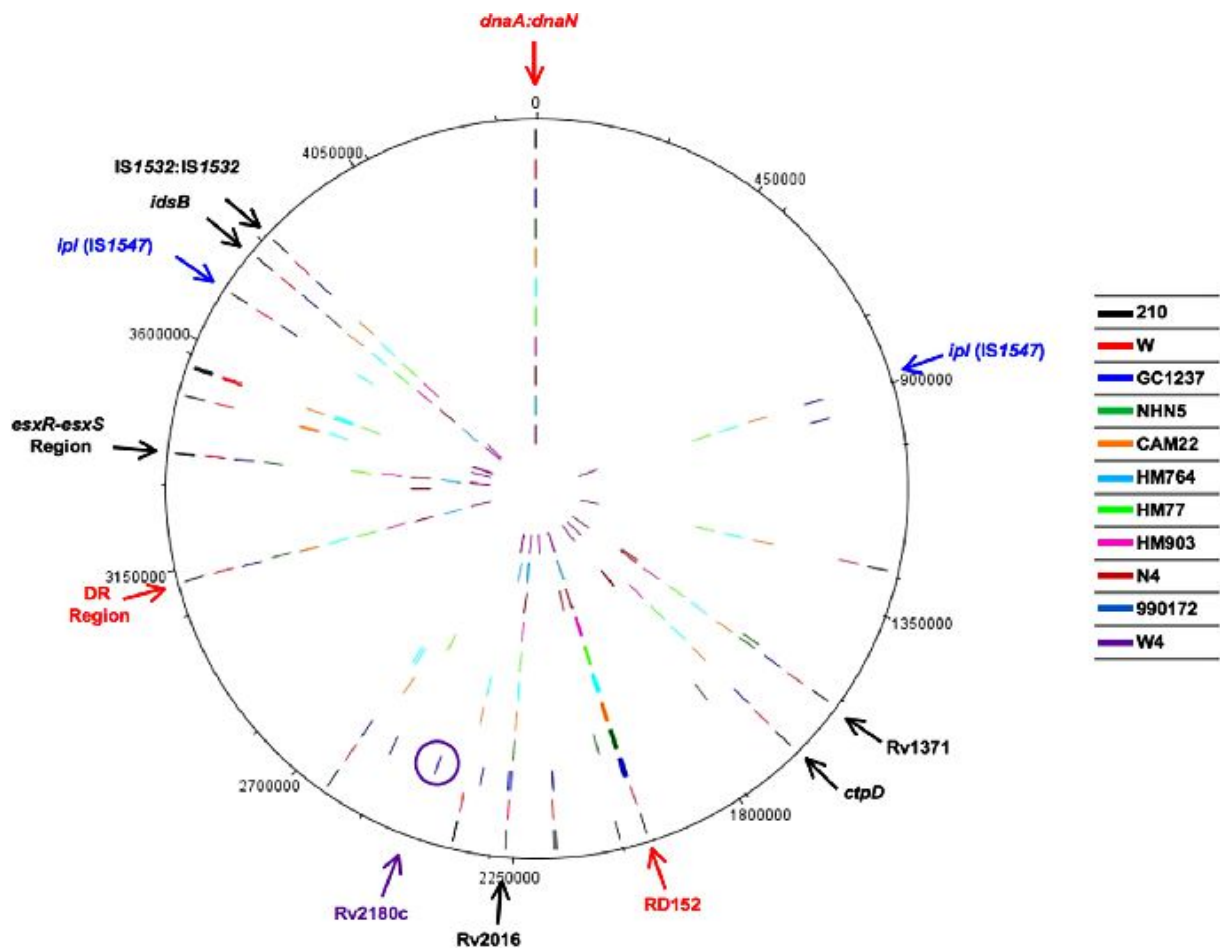


Figure 20 : Distribution de la séquence d'insertion IS6110 dans les génomes des souches Beijing (Henar Alonso, Sofia Samper et al. 2013).

1.5.4.2 IS6110 dans le locus NTF

Dans la région NTF, les souches Beijing sont caractérisées par la présence d'une copie d'IS6110. Dans certaines souches de la même famille Beijing, il a été observé 2 copies d'IS6110 (produit PCR 223 bp et 175 bp) au sein du locus NTF. Ces souches, nommées

« W », étaient résistantes à sept antituberculeux (l’Isoniazide, la Rifampine, l’Ethambutol, Streptomycine, la Kanamycine, l’Ethionamide, la Rifabutine). Elles constituent des représentants de la souche historique « W » multirésistante de l’épidémie de New York des années 90 ((Bonnie B. Plikaytis, Jennifer L. Marden et al. 1994), (Kurepina, Likhoshvay et al. 2005), (Bifani, Mathema et al. 2001)) (Fig. 21).

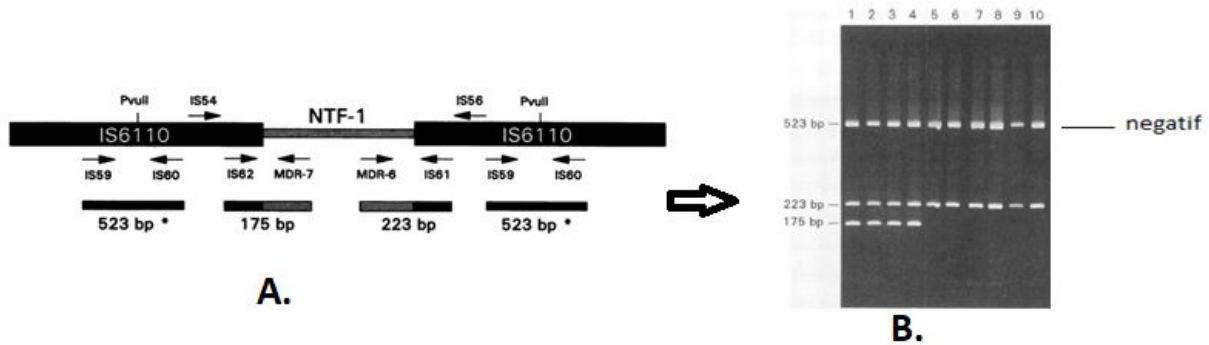


Figure 21 : Locus NTF, A- représentation schématique d’IS6110 au sein du locus NTF de la souche W et les amorces pour l’amplification multiplexée de détection de la souche W (Bonnie B. Plikaytis, Jennifer L. Marden et al. 1994), B- Produits PCR obtenus sur gel agarose typiques de la souche W (puits 1 à 4), de souches dites modernes (puits 5 à 10) (extrait de Plikaytis et al. 1994).

Sur la base du nombre de copies d’IS6110 dans le locus NTF et pour une première discrimination, les souches de géotypes Beijing ont par la suite été subdivisées en lignée phylogénétique ancienne ou atypique (pas d’insertion IS6110, produit PCR =523bp) et modernes ou typiques (NTF : IS6110, présence d’un produit PCR de 223 bp). Cette séquence d’insertion au sein du locus NTF est donc un marqueur d’intérêt d’une part pour l’identification, la discrimination, la surveillance et l’épidémiologie des souches Beijing notamment de la souche « W » (Mokrousov, Ly et al. 2005).

1.5.4.3 Souches épidémiques identifiées par IS6110 et leur classification en sous-groupes de la lignée Beijing

Les souches du géotype Beijing sont présentes au sein des populations mondiales mais en plus grande proportion en Asie et ex-URSS. Sur la base des différents marqueurs, et du mécanisme d’adaptation aux antituberculeux (pression sélective), les souches Beijing sont donc subdivisées en deux sous-groupes: groupe *ancien/atypique* ou « *ancestral* » et groupe *moderne/typique*. Les souches *modernes* sont plus fréquemment associées à la résistante que les souches *anciennes* en Chine (Yang, Luo et al. 2012). La discrimination entre les souches

Beijing *modernes* et *anciennes* assure une subdivision phylogénétique significative, une plus haute résolution de typage permettant ensuite de définir des sous-groupes qui présentent un intérêt épidémiologique.

Par exemple, la souche historique « W » d'identité « 212072 par RFLP » a causé l'épidémie de la tuberculose multirésistante à New-York dans les années 1990, et une forte transmission de souches Beijing en Irlande, dont la circulation était toujours en cours dans les années 2000 (Bifani, Mathema et al. 2001). Cependant, ce clone n'est actuellement plus très actif. De même, les clones 210 et HN878 de la même famille que les souches W sont à la base d'épidémies dans les Etats du Nord des Etats-Unis (Sreevatsan et al 1997, Yang et al 1998).

Un autre clone nommé « B0/W148 Beijing », appartenant à la lignée Beijing a été identifié et est très répandu en Russie. La distribution géographique des souches de génotype Beijing et Beijing B0/W148 dans les différentes populations varie, et représente jusqu'à 10% de B0/W148 dans les différentes régions de la Russie: 16% dans le nord Europe-Russie, 19-22% en Sibérie (21.8% Novosibirsk et 19.4% Tuva), 7.9% en Pskov, pourcentage plus élevé à Moscou) et 7% en Estonie, Abkhazie/Georgie et enfin 3.6% au Kazakhstan). Le clone Beijing B0/W148 serait originaire de la Sibérie (Mokrousov 2013). Sa première dispersion aurait débuté dans les années soixante par l'immigration de populations de la Sibérie vers l'Europe et par sa capacité accrue à acquérir la résistance aux antituberculeux largement utilisés (Mokrousov 2013).

Le clone Beijing B0/W148, d'après le typage IS6110 RFLP est défini par deux bandes caractéristiques dans la partie supérieure du profil RFLP (**Fig. 20**). Hormis la méthode IS6110-RFLP (Mokrousov 2008), une autre méthode alternative est de réaliser une PCR multiplexée de la région Rv2664-Rv2665:: IS6110, utilisée comme cible pour la détection et la surveillance des souches B0/W148 dans les zones épidémiques (Mokrousov, Narvskaya et al. 2012). Le profil 24 MIRU-VNTR des souches B0/W148 au Nord-Ouest de la Russie (St Petersburg) est « 244233352644425173353723 » et dans la base international MIRU-VNTRplus ils est désigné clone 100-32 (Weniger et al. 2010). Avec le MIRU 12, ils sont de type M11 et ont le profil digital « 223325173533 » ((Mokrousov 2005) ; (Mokrousov, Narvskaya et al. 2012)) ; cette signature est aussi appelée MIT17 dans la base de données SITVIT (Demay, Liens et al. 2012). Il est le type le plus largement retrouvé en Asie de l'Est et vient au second rang en Russie. Les souches B0/W148 sont hypervirulentes et ont démontré une augmentation de valeur sélective et de leur croissance dans les cellules THP-1 et macrophages (Lasunskaja, C.M.Ribeiro et al. 2010).

Les souches Beijing venant de la Russie, incluant la souche B0/W148 sont significativement plus cytotoxiques que d'autres souches Beijing, et induisent un faible niveau de TNF- α et un haut niveau de l'interleukine inhibitrice IL-10 produite par les macrophages infectés. Le taux de souches en grappes (clusterisation) sert à estimer la transmissibilité de certaines souches spécialement pour les isolats MDR. L'augmentation de la transmissibilité de ces souches est détectée par la formation de grappes lors des analyses d'épidémiologie moléculaire, c'est particulièrement le cas dans des conditions spécifiques comme les prisons.

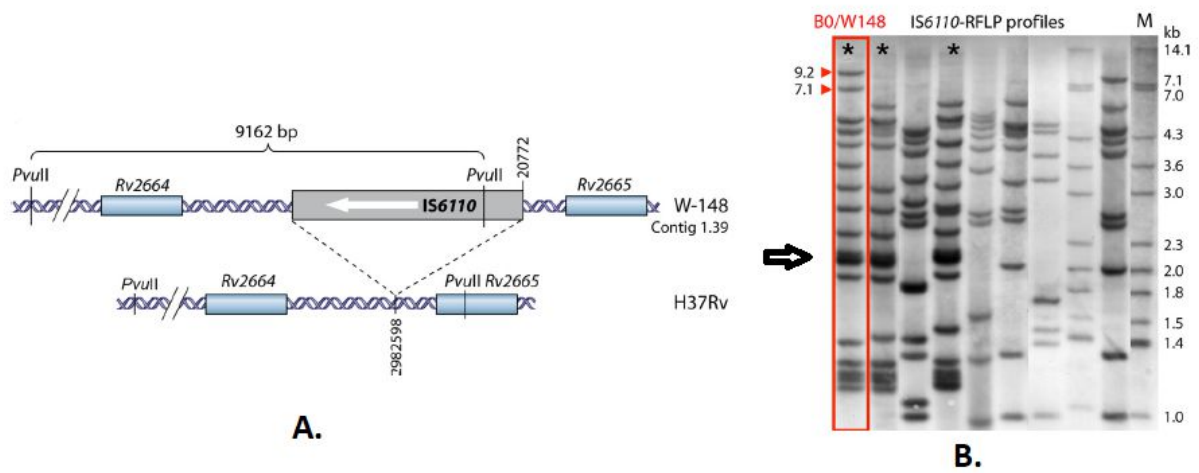


Figure 22 : Identification de la souche B0/W148, A- région intergénique Rv2664-Rv2665 contenant IS6110, B- photo de gel d'électrophorèse des profils IS6110-RFLP montrant le profil de la souche B0/W148 avec deux bandes supplémentaires, comparées aux autres profils IS6110 RFLP d'autres souches (Mokrousov 2013).

Les souches K et GC1237 responsable d'épidémie en Corée et au Canarie respectivement sont des anciennes Beijing avec aussi un nombre important d'IS6110 dans leur génomes (10 à 19 copies) ((Alonso H 2011), (Han, Song et al. 2015)).

Les travaux menés dans la thèse ont été réalisés sur la technologie Luminex que nous développerons dans la seconde partie.

Chapitre 2 : La technologie Luminex®, Principes et Application

La technologie Luminex® : Principes et application

La technologie Luminex (Luminex, Austin, Texas, USA) est basée sur l'utilisation de microbilles (microsphères) en polystyrène (5,6 µm de diamètre) ou magnétiques (6,5 µm de diamètre). Les microbilles contiennent deux colorants internes ou fluorochromes (Dunbar S.A., 2006) dans des proportions différentes d'une région optique à une autre et permettant ainsi de distinguer jusqu'à 100 régions optiques différentes par l'analyseur Luminex 100/200. Parmi les analyseurs Luminex, il y a le Luminex 100/200 dont le principe est fondé sur la cytométrie en flux, mais aussi le MagPix, appareil plus récent d'imagerie de fluorescence.

Au cours d'une analyse, un amplicon obtenu par PCR (1^{ère} étape, **Fig 23**) à partir d'un échantillon biologique ou encore par marquage d'ADN par un marqueur fluorescent ou détectable indirectement facilement (par exemple à l'aide du système avidine-biotine), sera hybridé à une sonde (2^{ème} étape, **Fig 23**) de capture qui aura été préalablement couplée à un type de microbille donné, par une liaison avec un linker en C12, le couplage chimique étant réalisé en présence d'EDC (1-Ethyl-3-[3-diméthylaminopropyl] carbodiimide hydrochloride). Par réaction chimique en présence du catalyseur et dans un milieu tampon MES (acide 4-morpholio-ethane-sulfonate) à pH 4,5 on réalise une liaison covalente amide entre les groupes carboxyliques des microbilles et les groupes amines des oligonucléotides de capture spécifiques des séquences cibles. L'hybridation peut doit être suivie, si les amplicons ne portent pas déjà une extrémité détectable en 5' (par exemple par utilisation d'une amorce marquée par Cy5) d'une étape de détection (3^{ème} étape, **Fig 23**) par ajout de la Streptavidine-phycoérythrine où la phycoérythrine se fixe à une ou plusieurs molécules de biotine présentes sur le brin d'ADN biotinylé de l'amplicon (une des deux amorces de PCR seulement étant biotinylée, celle complémentaire à la sonde de capture). Les microbilles contenues dans l'échantillon à analyser sont aspirées et injectées dans une veine par le liquide de gaine et passent devant le trajet optique permettant un alignement et un passage individuel des microbilles devant les deux lasers. Chaque microbille hybridée est donc individuellement excitée à son passage par les deux lasers (longueur d'onde d'excitation : 633 nm et 532 nm) pour son identification (longueur d'onde d'identification: 633 nm) et la détection de la fluorescence à sa surface (longueur d'onde : 532 nm) (**Fig. 23**). L'intensité de la fluorescence à la surface des microbilles reflète l'interaction oligonucleotide de capture – amplicon (dans le cas où on analyse des produits de PCR) et elle est traduite en valeurs numériques brutes qui

seront ensuite traitées statistiquement pour déterminer les positifs et les négatifs.

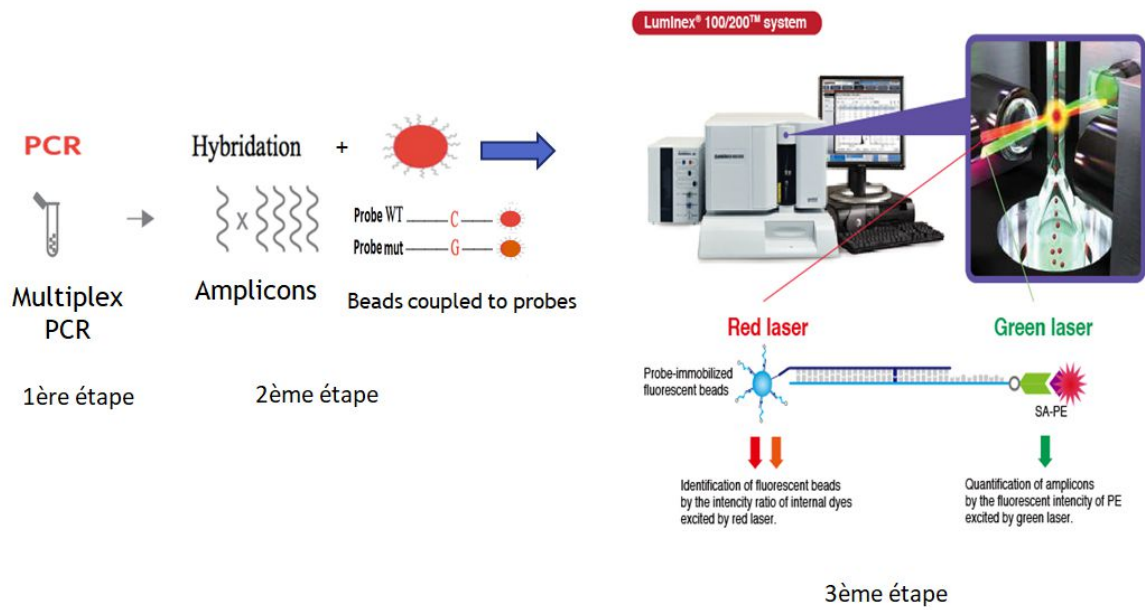


Figure 23 : Principe de la technologie Luminex

1^{ère} étape- PCR multiplexé, 2^{ème} étape – Hybridation de l’amplicon aux sondes couplé aux billes, 3^{ème} étape- Révélation par ajout de la SAPE et détection sur Luminex²⁰⁰.

Chapitre 3 : But et problématique de la thèse

But et problématique de la thèse

La tuberculose causée par la bactérie *Mycobacterium tuberculosis* est l'une des infections les plus répandues dans le monde et demeure un problème majeur de santé publique notamment dans les pays de l'ex-URSS et en Asie. L'expansion des clones épidémiques de la lignée Beijing (lignée 2) de *M. tuberculosis* particulièrement virulente, homogène, et adaptée à la résistance aux antituberculeux, ainsi qu'une forte transmission complique la prise en charge.

L'objectif académique principal est de mieux comprendre la taxonomie et la résistance des souches de la lignée 2 (Beijing). En se basant sur les diverses classifications phylogénétiques proposées par d'autres auteurs, leur corrélation et la proposition récente d'unification (Schikitov et al 2017) il s'agira d'extraire des SNPs spécifiques à chaque clade épidémique et non épidémique, de les associer à d'autres marqueurs moléculaires spécifiques, SNP de résistances, RD et NTF::IS6110, en vue de mettre au point une méthode appliquée pour le diagnostic direct et la surveillance des clones de la lignée Beijing par hybridation sur microbilles et analyse de fluorescence sur des systèmes d'analyse biologique Luminex 200®.

Ainsi pour atteindre cet objectif, quatre différentes études ont été définies. La première étudie les SNP spécifiques à l'ultra-résistance des souches MTBc aux antibiotiques de deuxième génération par des méthodes haut débit (1), le deuxième projet est basé sur l'analyse de la séquence d'insertion IS6110 au sein du locus NTF combinée à la multirésistance des souches MTBc à la Rifampicine/Isoniazide (2), le troisième projet repose sur l'identification des SNPs spécifiques pour chaque sous-groupe/groupe de souches Beijing en vue d'une classification fine, un typage adéquat de ces souches.

Le but de la thèse était donc:

(1) de finaliser un test de détection des mutations de résistance aux antibiotiques de seconde ligne complétant le panel des mutations ciblées par la technologie Luminex qui présente l'avantage de permettre un haut niveau de multiplexage.

(2) de développer une méthode d'identification de souches Beijing sur la base du nombre de copies d'IS6110 au sein du site NTF combinée à la détection de la multirésistance à la Rifampicine/Isoniazide. Dans ce domaine, il s'agira plus spécifiquement de caractériser les

souches « Beijing », un groupe génétique de bacilles tuberculeux, et aussi de suivre leur évolution en fonction de leur profil de résistance.

(3) de comprendre la taxonomie des souches Beijing et identifier des SNP spécifiques dans les génomes de souches Beijing en vue du développement d'une méthode à base des marqueurs « SNP, RD, IS*6110* » pour la caractérisation fine et la surveillance des souches Beijing.

(4) de mettre en pratique les différentes méthodes développées et préexistantes dans l'équipe, afin d'étudier la transmission de la tuberculose dans différents territoires où la présence de souches Beijing est problématique et de suivre la transmission de clones épidémiques, à l'intérieur d'un Etat ou/et entre Etats.

Chapitre 4 : Résultats et Articles

**Etude 1 : Développement de méthode haut-débit
sur microbilles pour la détection des résistances
aux antibiotiques de second ligne sur les souches
de *Mycobacterium tuberculosis***

Article 1 :

Résumé

La tuberculose ultra résistante est reportée au moins dans un pays sur chacun des cinq continents, et la réémergence de la multirésistance et/ou ultra résistance constitue un véritable problème de santé publique dans le monde. Cette émergence est due à l'apparition des souches de *M. tuberculosis* multi-résistantes et ultra-résistantes aux antituberculeux, en plus de la pandémie du VIH/SIDA. Pour le contrôle et/ou l'élimination de la tuberculose, il est primordial de mettre au point des mesures de santé publique telles que; la disponibilité et l'accessibilité à des tests de diagnostic rapides pour la détermination de la sensibilité aux médicaments en particulier de deuxième ligne dans les régions à forte prévalence de souches MDR; des traitements efficaces et des outils de surveillance et de contrôle aux personnes atteintes afin d'éviter les contaminations interhumaines.

Nos travaux concernent la recherche, le développement et l'application de méthodes moléculaires rapides et multiplexées, pour prédire la tuberculose ultrarésistante en identifiant les mutations (polymorphismes de nucléotides simples) dans les gènes associés à la résistance des souches de *Mycobacterium tuberculosis* aux antituberculeux de seconde ligne.

Ainsi à l'aide de la technologie Luminex²⁰⁰, et des amorces Dual-Priming-Oligonucleotide (DPO) pour faciliter la PCR multiplexée, nous avons pu mettre au point le TB-EFI qui nous permet d'identifier les mutations impliquées dans la résistance à l'Ethambutol, les Fluoroquinolones et les trois antituberculeux Injectables (Kanamycine, Amikacine, Capreomycine) et aussi de prédire la sensibilité/la résistance de la souche. La concordance entre la résistance génotypique et phénotypique est de 72-98% avec une faible efficacité pour la fluoroquinolone et l'Ethambutol à cause de mutations non ciblées. Sur des ADNs dont le profil sensible/résistant était non déterminé, le TB-EFI a prédit une souche XDR dans 91 échantillons d'ADNs (Pakistan et Brésil) et 9 souches XDR dans 43 isolats MDR du Kazakhstan.

En se basant sur la détection rapide de la résistance aux antituberculeux de deuxième ligne et celle de l'Ethambutol, TB-EFI est un test efficace et pas coûteux pour une surveillance mondiale à grande échelle de la résistance et le contrôle de la tuberculose XDR.

TB-EFI, a novel 18-Plex microbead-based method for prediction of second-line drugs and ethambutol resistance in *Mycobacterium tuberculosis* complex

Running title: TB-EFI: microbead-based detection of second-line drug-resistant TB

Authors :

Bernice. J. Klotoe¹, Barbara Molina-Moya², Harrison Magdinier Gomes^{1,3}, Michel. K. Gomgnimbou^{1,4,5}, Lorena Oliveira Suzarte¹, Maria H. Féres Saad⁶, Sajid Ali⁷, José Dominguez², Edita Pimkina⁸, Elena Zholdybayeva⁹, Christophe Sola^{1#} and Guislaine Refrégier^{1#}.

Affiliations :

¹Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France. ²Servei de Microbiologia, Hospital Universitari Germans Trias i Pujol. Institut d'Investigació Germans Trias i Pujol, Badalona, Spain. Universitat Autònoma de Barcelona. CIBER Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, Spain. ³Laboratório de Biologia Molecular Aplicada à Micobactérias, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil. ⁴Centre Muraz, Bobo-Dioulasso, Burkina-Faso, ⁵Univ. Polytech, Bobo-Dioulasso, Burkina Faso, ⁶Laboratório de Microbiologia Celular, Instituto Oswaldo Cruz, Av. Brasil, 4365 - 20245, Rio de Janeiro, Brasil, ⁷Microbiology Department, Quaid-i-Azam University, Islamabad Pakistan, ⁸Infectious Diseases and Tuberculosis Hospital, Affiliate of Vilnius University Hospital Santariskiu klinikos, Vilnius, Lithuania. ⁹National Center for Biotechnology, Astana, Kazakhstan.

Co-corresponding authors: guislaine.refregier@u-psud.fr , christophe.sola@i2bc.paris-saclay.fr

Keywords: NAAT, XDR-TB, second-line drug resistance, TB, automatized analysis, allele call.

Abstract

Several diagnostic tests are being developed to detect drug resistance in tuberculosis. In line with previous developments detecting rifampicin and isoniazid resistance using microbead-based systems (spoligotyping and TB-SPRINT), we present here an assay called TB-EFI detecting mutations involved in resistance to ethambutol, fluoroquinolones and the three classical injectable drugs (kanamycin, amikacin and capreomycin) in *Mycobacterium tuberculosis*. The proposed test includes both wild-type and mutant probes for each targeted locus. Basic analysis can be performed manually. An upgraded interpretation is made available in Excel 2016®.

Using a reference set of 61 DNA extracts, we show that TB-EFI provides perfect concordance with pyrosequencing. Concordance between genotypic resistance and phenotypic DST was relatively good (72 to 98% concordance), with lower efficiency for fluoroquinolones and ethambutol due to some untargeted mutations. When compared to phenotypical resistance, performances were similar to those obtained with Hain MTBDRs/ assay, possibly thanks to the use of automatized processing of data although some mutations involved in fluoroquinolone resistance could not be included.

When applied on three uncharacterized sets, phenotype could be predicted for 51% to 98% depending on the setting and the drug investigated, detecting one extensively drug-resistant isolate in each of a Pakistan and a Brazilian set of 91 samples, and 9 XDR among 43 multi-resistant Kazakhstan samples.

By allowing high-throughput detection of second-line drugs resistance and of resistance to ethambutol that is often combined to second-line treatments, TB-EFI is a cost-effective assay for large-scale worldwide surveillance of resistant tuberculosis and XDR-TB control.

Highlights

- TB-EFI detects 2nd-line drug and EMB resistance mutations using microfluidics
- TB-EFI is highly reproducible, has very high specificity and high sensitivity
- Combination of maximum signal and allele call analyses is key to TB-EFI performance
- 11 XDR-TB were detected in this study.

Introduction

The mortality rate of tuberculosis dropped by 47% between 1990 and 2015, still the World Health Organization estimated 10.4 million of new cases, 1.3 million deaths caused by TB only, an additional 0.37 million deaths resulting from TB among HIV-positive people in 2016, and cases of antibiotic resistance are growing with increasing primary resistance (Zager and McNerney 2008; Brites and Gagneux 2012; Eldholm, Monteserin et al. 2015; WHO 2017).

Multidrug-resistant (MDR) TB defines TB cases resistant to the two main first-line drugs, Rifampin (RIF) and Isoniazid (INH). Extensively drug-resistant (XDR) TB refers to resistance to RIF and INH, plus a fluoroquinolone (FLQ, such as levofloxacin LVX, Ofloxacin OFX, and Moxifloxacin MXF) and at least one second-line injectable drug (SLID) (kanamycin, KAN; amikacin, AMK; or capreomycin, CAP). One hundred and twenty-three (123) countries reported XDR-TB at the end of 2016 (WHO 2017). The average proportion of MDR-TB cases with XDR-TB was around 6.2% in 2016, a rate lower than in preceding years (9 to 9.7%) but absolute number of XDR-TB still increases. Highest levels have been observed in Belarus (29% in 2014) and Lithuania (25% in 2013). The five countries that reported the largest numbers of cases in 2016 were China (n=525), Belarus (n=572), South Africa (n=967), Ukraine (n=1195) and India (n=2464) (WHO 2015). These increasing levels of resistance to second-line drugs appeal to resistance detection improvement to provide adequate treatment.

Recommended treatment of MDR-TB is composed of at least five likely effective drugs as per recent WHO updates: one FLQ, one second-line injectable drug (SLID), at least two drugs from group C (Ethionamide/Protionamide, cycloserine, linezolid, clofazimine), and pyrazinamide (PZA). It is also recommended to strengthen the regimen with high-dose INH and/or ethambutol (EMB) (WHO 2016).

Resistance to second-line drugs and EMB is conferred by mutations in drug targets or genes involved in drug metabolism (Takiff, Salazar et al. 1994; Starks, Gumusboga et al. 2009; Zaunbrecher, Sikes et al. 2009; Georghiou, Magana et al. 2012; Malik, Willby et al. 2012; Zhu, Zhang et al. 2012; Chernyaeva, Fedorova et al. 2013; Avalos, Catanzaro et al. 2015; Kambli, Ajbani et al. 2015; Zhao, Sun et al. 2015). Relative frequency of these mutations is similar in different studies suggesting that some mutations confer higher level of resistance and are therefore most frequently selected under antibiotic pressure (**Table 1**).

Table 1 – Mutations conferring resistance to EMB and second-line drugs and Excel 2016® comparison of various NAAT SNP panels

Drug	Resistance loci / prediction tool	Prevalence among resistant isolates	Beamedex TB-EFI	HAIN MTBDRsl V 1.0	HAIN MTBDRsl V 2.0	Beamedex MLPA (TB-SNPID)*
FLQ	gyrA_90-91 wt		wt90	WT1,2,3	WT1,2,3	x
	gyrA90_mut_GTG(A90V)	13-29%	mut90	MUT1	MUT1	
	gyrA91_mut_CCG(S91P)	2-7%		MUT2	MUT2	
	gyrA_94_wt_GAC		wt94			x
	gyrA_94_mut_GCC(D94A)	7-17%		MUT3A	MUT3A	
	gyrA94_mut_AAC(D94N)	5-8%	mut2_94	MUT3B	MUT3B	
	gyrA94_mut_TAC(D94Y)	3-14%		MUT3B	MUT3B	
	gyrA94_mut_GGC(D94G)	21-36%	mut1_94	MUT3C	MUT3C	
	gyrA94_mut_CAC(D94H)	0-7%		MUT3D	MUT3D	
	gyrB_WT				WT	
	gyrB_500	4%				
gyrB_538	1%			MUT1		
gyrB_540	4%			MUT2		
Injectable drugs KAN / AMK / CAP	rrs 1401_wt_A		wt1401	WT1	WT1	x
	rrs1401_mut_G	38-80% / 78-80% / 50-80%	mut1401	MUT1	MUT1	
	rrs1402_wt_C		wt1402			x
	rrs1402_mut_T	1% / 1% / 2%	mut1402			
	rrs1484_wt_G		wt1484	WT2	WT2	
	rrs1484_mut_T	1-5% / 1% / 1%	mut1484	MUT2	MUT2	
	eis WT (G-37)				WT1	
	eis G-37T	0-19% / 0% / 0%				
	eis_wt(-10to-14)_CACAG		wt-10-14		WT2	
	eis WT -2				WT3	
	eis_(-10A)_CACAA	5-33% / 0% / 0%	mut1			x
eis_(-12T)_CATAG	3% / 3-6% / 0%	mut2				
eis_(-14T)_TACAG	5-24% / 2-9% / 0%	mut3		MUT1	x	
EMB	embB306_wt_ATG		wt	WT1		
	embB306_mut-ATA	0-23%	mut2	MUT1A		
	embB306_mut-GTG	34-42%	mut1	MUT1B		
Diagnostic tool references			<i>This work</i>	<i>Hilleman et al. 2009</i>	<i>Brossier et al. 2016</i>	<i>Sengstake et al. 2014</i>

NAAT= Nucleic Acid Amplification Test

DNA sequencing is the standard and most effective method to detect mutations, and tools to ease Next Generation Sequencing raw data analyses are being developed (Bradley, Gordon et al. 2015; Feuerriegel, Schleusener et al. 2015; Phelan, Coll et al. 2016; WHO 2017). However, processes are still relatively tedious, expensive, necessitating at least basic bioinformatic skills and relatively poorly explicit regarding resistance (Feuerriegel, Schleusener et al. 2015). For the time-being, such solutions are rather suited to reference laboratories with large financing capacities and not for low-income countries. For routine detection of first-line drug-resistance the most commonly used methods are Genotype MTBDR*plus* (Hain Lifescience, Germany), and the GeneXpert MTB/RIF (Cepheid, USA) assays. The recent TB-SPRINT (based on Dual-Priming Oligonucleotide), TB-SNPID (based on Multiple-Ligation Probe dependent Amplification), and TRIOL (based on Nested-PCR) are examples of alternative multiplexed assays, all being based on multiplexed PCR and on microbead-based arrays, a well suited format for efficient and cost-effective nucleic acid amplification tests (NAATs) (Hillemann, Rusch-Gerdes et al. 2009; Bergval, Sengstake et al. 2012; Gomgnimbou, Hernandez-Neuta et al. 2013; Sengstake, Bablishvili et al. 2014; Cabibbe, Miotto et al. 2015; Brossier, Guindo et al. 2016; WHO 2016; Yin, Chan et al. 2016). Many recent systems use chips or Line-Probes to investigate first-line drug resistance or second-line-drug resistance (**Table 1**): VerePLEX (Cabibbe, Miotto et al. 2015) Genotype MTBDR*s*/ V2.0 (Brossier, Guindo et al. 2016), microbead-based ASPE assay (Lee, Choi et al. 2015; Kim, Yang et al. 2018), Anyplex II MTB/MDR/XDR (Molina-Moya, Lacoma et al. 2015; Perez-Garcia, Ruiz-Serrano et al. 2017).

The goal of this study was to develop a new add-on multiplexed method to TB-SPRINT for detecting the main mutations involved in resistance to second-line drugs (FLQ and injectable drugs) also including EMB, one of the first line drug. We named this system TB-EFI (Ethambutol, Fluoroquinolones, Injectables). It complements our previous TB-SPRINT and shows the advantage of allowing large sample screening and rapid analysis; it has a reasonable cost (in combination with TB-SPRINT <15 euros), reasonable technical expertise demand, has been validated by several independent studies (Yasmin, Gomgnimbou et al. 2014; Dantas, Suffys et al. 2017; Molina-Moya, Gomgnimbou et al. 2017; Molina-Moya, Agonafir et al. 2018) and has been successfully implemented onto more robust device MagPix® (Gomgnimbou, Hernandez-Neuta et al. 2013). Developing TB-EFI is in the line of proposing more assays to select the most efficient ones and achieve higher rapidity in drug resistance prediction (Somoskovi 2015). We present our assay fitted for the classical LX200

system. TB-EFI targets most common mutations involved in FLQ, KAN, AMK, CAP and EMB resistance (**Table 1**). By applying this test on unknown samples, we also contribute to the description of main mutations frequencies.

Materials and Methods

Reagents and devices

We used the high-throughput flow cytometry-based LX200 Luminex® system (Luminex Corp., Austin, TX) to develop this new assay. Primers and probes were purchased from IDT (IDT Corp. Newark, NJ). Oligonucleotide-coupling onto MicroPlex® microspheres (polystyrene) was performed by the Beamedex® company by EDC coupling (Beamedex®, SAS, Orsay, France).

Strains and phenotypic drug susceptibility testing

Reference set used to develop and validate the test:

Sixty-one (n=61) DNA samples (efi1 to efi61) corresponding to as many *Mycobacterium tuberculosis* complex isolates described in Lacoma *et al.* study were used as a set of reference (Lacoma, Molina-Moya *et al.* 2015). More than half of them are resistant to at least one second-line drug (**Table 2**). Detailed profiles are available in **Appendix Table A.1**. A set of five of these 61 samples were randomly selected to develop the test (five different samples for each probe).

Table 2: Phenotypic resistance to second-line and ethambutol drugs in the reference set of 61 isolates.

		Number of clinical isolates (percentages)	Total (percentage)
FLQ	OFX ^R	21 (35)	61 (100)
	OFX ^S	23 (39)	
	MXF ^R	4 (5)	
	MXF ^S	5 (7)	
	LVX ^R	6 (11)	
	LVX ^S	2 (4)	
Second-line injectable drugs	KAN ^R	29 (48)	61 (100)
	KAN ^S	32 (52)	
	AMK ^R	14 (23)	61 (100)
	AMK ^S	38 (62)	
	Not Determined	9 (15)	61 (100)
	CAP ^R	9 (15)	
	CAP ^S	51 (84)	
	Not Determined	1 (2)	
EMB	EMB ^R	41 (67)	61 (100)
	EMB ^S	20 (33)	

FLQ: Fluoroquinolones, OFX: Ofloxacin, MXF: Moxifloxacin; LVX: Levofloxacin, KAN: Kanamycin, AMK: Amikacin, CAP: Capreomycin, EMB: Ethambutol, R: Resistant, S: Susceptible. Critical concentrations used for isolates efi1 to 52 for OFX, LVX, KAN, CAP, AMK, EMB were: 2 mg/L, 1.5 mg/L, 2.5 mg/L, 3 mg/L, 1.5 mg/L, 5 mg/L in BACTEC MGIT960. Critical concentrations used for isolates efi53 to 61 in BACTEC 460TB for MOX, KAN, CAP and EMB were: 0.5 mg/L, 5 mg/L, 1.25 mg/L, 7.5 mg/L.

Test sets: 43 MDR samples from Kazakhstan, 91 uncharacterized random samples from Brazil and 91 samples from Pakistan. The Pakistan sample was enriched in MDR isolates.

FLQ resistance levels were expected to be above 15% in Pakistan and Kazakhstan, and around 1% for Brazil as it is a country with low-resistance levels (WHO 2016). These uncharacterized sets were used for first-step clinical evaluation of the upgraded analysis to report conclusive results frequency (Mattocks, Watkins et al. 2010).

DNA extraction and PCR amplification

Reference DNA extracts were obtained by thermolysis as described previously (Lacoma, Molina-Moya et al. 2015). They were sent at 4°C to Paris for TB-EFI analysis. DNA from Pakistan, Brazil and Kazakhstan were extracted by the cetyl-trimethyl-ammonium bromide (CTAB) or by a simple thermolysis method (Kazakhstan samples) and shipped at room temperature to Paris for TB-EFI analysis.

Four-plex PCR amplification was performed in a final volume of 25µL containing 1 µL of DNA extract, 1X Flexi buffer, 1.25U of GoTaq Flexi DNA polymerase (Promega Corp., Fitchburg, WI), 1.5 mM of MgCl₂, 0.2 µM of each primer, 200 µM of each dNTP. After initial denaturation at 95°C for 3 min, the reaction mixture was run through 35 cycles of a Touch-down PCR protocol: denaturation at 95°C for 30s, annealing at 67 to 59°C for 30s for 10 cycles and an additional 25 cycles at 57°C, and extension at 72°C for 30s. Final extension at 72°C lasted 5 min.

Primer and probe oligonucleotide sequences and hybridization protocol

Primers were designed using the dual-priming oligonucleotide (DPO) principle (Chun, Kim et al. 2007). They amplify 100 to 350 bp-long fragments covering the targeted mutations. Probes were designed under the following rationale: their first version was 19-bp long and centered on the targeted single nucleotide polymorphism (SNP), their length was adjusted to reach a T_m close to 60°C. Final sequences of primers and probes are shown in **Table 3**. Hybridization involved 1,800 MicroPlex® coupled-microspheres per target in a final volume of 50 µL of Tetramethylammonium chloride buffer (TMAC 1x), including 2 µL of the PCR product as described previously (Gomgnimbou et al, 2013). Detection was done using Streptavidin-Phycoerythrin Lumigrade (Roche Biochemicals, Penzberg, Germany) and was used at 5 µg/mL final concentration. Reading was done in the LX200 after 5 min incubation at 52°C following user's manual, and using a minimum of 50 microspheres count for each target

(Dunbar 2006). A raw data output file under csv format provides Median Fluorescent Intensities (MFI) values for each target for each sample.

Table 3 - Primers and probes sequences.

	Name	Sequence
Primer 1	gyrA_DPO_Fw	CCGGGTGCTCTATGCAATGTTTCGIIIIIGGCTTC
Primer 2	gyrA_DPO_Rv	GCTCTTGCACCCGGCCGTCGTIIIIIGGGATG
Primer 3	rrs_DPO_Fw	AATGGCCGGTACAAAGGGCTCATCIIIIICGAGGT
Primer 4	rrs_DPO_Rv	CCGCACCTTCCGGTACGGCTACIIIIITACGAC
Primer 5	embB_T_fw	GCGCACCTTACCCTGACCGACGIIIIIGTGATA
Primer 6	embB_T_Rv	GCTGCCGAACCAGCGGAAATAIIIIACATGT
Primer 7	eis_DPO_fw	AAATTCGTGCTGATTCTCGCAGTIIIIICACGCT
Primer 8	eis_DPO_Rv	CCGATGAAATCGGTGAAACTGGIIIIIGCCAGT
Probe 1	gyrA 90-91 wt	ACGGCGACG C GTCGATCTAC
Probe 2	gyrA90_mut_GTG	ACGGCGACGTGTCGATCTAC
Probe 3	gyrA 94_wt	TCGATCTAC <u>GAC</u> ASCCTGGTG
Probe 4	gyrA94_mut_GGC	TCGATCTAC <u>GGC</u> ASCCTGGTG
Probe 5	gyrA94_mut_AAC	TCGATCTAC <u>AAC</u> ASCCTGGT
Probe 6	rrs 1401_wt	ACCGCCCGTCA <u>C</u> CGTCATGAAA
Probe 7	rrs1401_mutG	ACCGCCCGTCA G CGTCATGAAA
Probe 8	rrs1402_wt	ACCGCCCGTCA <u>C</u> CGTCATGAAA
Probe 9	rrs1402_mutT	ACCGCCCGTCA <u>T</u> CGTCATGAAA
Probe 10	rrs1484_wt	GATTGGGAC <u>G</u> AAGTCGTA
Probe 11	rrs1484_mutT	GATTGGGAC <u>T</u> AAGTCGTA
Probe 12	eis_wt	GGCATATGCC <u>CACAG</u> TCGGATTCT
Probe 13	eis_(-10A)	GGCATATGCCACA <u>A</u> TCGGATTCT
Probe 14	eis_(-12T)	GGCATATGCCCA <u>C</u> AGTCGGATTCT
Probe 15	eis_(14T)	GGCATATGCT <u>T</u> ACAGTCGGATTCT
Probe 16	embB-wt306	ATCCTGGGC <u>AT</u> GGCCCGAGT
Probe 17	embB-mt306GTG	ATCCTGGGC <u>G</u> TGGCCCGAGT
Probe 18	embB-mt306ATA	CCTGGGC <u>A</u> TAGCCCGAGTC

Codons or nucleotide of 16S RNA affected by the mutation are underlined, mutated nucleotide is shown in bold.

Genotype prediction using basic or upgraded analysis of median fluorescent intensities

Analysis of MFI values to predict genotypes were performed in Excel 2016® (Microsoft Office, Cupertino, CA). Two types of analyses were performed. The first one called “basic analysis” infers genotypes based on the maximal MFI value obtained for the targeted locus: if MFI value of wt_probe is higher than that of the single/all other mutant probes, genotype is

inferred to be wild type; if MFI_value of mut1_probe is higher than that of wt_probe and of other mut_probes, genotype is declared “mut1”, etc.

The second method, called “upgraded TB-EFI analysis” combines maximal MFI detection with allele-call method as described for MOL-PCR method (Wuyts, Mattheus et al. 2015). Briefly, allele-call method compares all signals to noise as detected in negative controls (BLANK) and confronts signals from all alternative probes targeting the same locus. Additional information on the technique is provided in **Appendix file A.2**. Upgraded TB-EFI analysis was automatized in Excel 2016®. Note that allele-call method requires several negative controls (target number of neg. controls n=5). As an output, the automatized analysis provides a list of genotypes at each targeted locus (by columns) for each sample (by lines). Summary of total sample characteristics is provided as a separate table.

Phenotype prediction by TB-EFI

Table 4- Phenotype prediction using TB-EFI genotype predictions

Targeted loci predicting...	COMPLETE resistance	PARTIAL resistance	susceptibility
FLQ	<i>gyrA</i> 90mut GTG or <i>gyrA</i> 94mut1 GGC or <i>gyrA</i> 94mut2 AAC		<i>gyrA</i> 90wt GCG + <i>gyrA</i> 94wt GAC
KAN	<i>rrs</i> 1401mut G or <i>rrs</i> 1484mut T or eis -14mut3 T	eis -12 T or eis -10 A	<i>rrs</i> 1401wt A + <i>rrs</i> 1484wt G +eis -14wt C +eis -12wt C +eis -10wt G
AMK	<i>rrs</i> 1401mut G or <i>rrs</i> 1484mut T	eis -12 T or eis -14 T	<i>rrs</i> 1401wt A + <i>rrs</i> 1484wt G +eis -14wt C +eis -12wt C
CAP	<i>rrs</i> 1402mut T or <i>rrs</i> 1484mut G	<i>rrs</i> 1401 G	<i>rrs</i> 1402wt C + <i>rrs</i> 1401wt A + <i>rrs</i> 1484wt G
EMB	<i>embB</i> 306 GTG or <i>embB</i> 306 ATA		<i>embB</i> 306wt ATG

wt=wild-type.

Automatized TB-EFI analysis also provides phenotype prediction for FLQ, KAN, AMK, CAP and EMB using data extracted from the literature. **Table 4** lists the rules to infer resistance based on most recent publications. Known MDR Isolates predicted to be also resistant to FLQ and at least one SLID are coined as “XDR”. A Summary of total sample characteristics is also provided as a separate table.

Statistical data analyses

Computation of concordance (Co) and Kappa (κ) were performed in Excel 2016®. Agreement between the two methods is traditionally acknowledged as satisfying if κ lies above 0.6, and as excellent if κ lies above 0.8.

Confidence interval of concordance were computed approximating the binomial distribution by the normal distribution (Hess, Shardell et al. 2012): $\text{error} = 1.96 * \sqrt{[\text{Co} * (1 - \text{Co}) / n]}$.

Results

We aimed at designing a test targeting the most frequent mutations involved in second-line drugs and EMB resistance at a reasonable cost. We used one reference set of 61 samples enriched in samples resistant to second-line drugs (Lacoma, Molina-Moya et al. 2015) and 3 uncharacterized sets coming from countries with renowned differences in the level of second-line drug-resistant TB prevalence (Mattocks, Watkins et al. 2010).

Reproducibility of TB-EFI using basic analysis

Table 5- Reproducibility of basic genotype analysis by TB-EFI for the reference samples

		<i>gyrA_90</i>	<i>gyrA94</i>	<i>rrs1401</i>	<i>rrs1402</i>	<i>rrs1484</i>	<i>eis-prom</i>	<i>embB306</i>
Genotypes inferred by pyrosequencing	wt	57 (GCG)	46 (GAC)	54 (A)	61(C)	60 (G)	30 (CACAG)	26 (ATG)
	mut 1	4 (GTG)	11 (GGC)	7 (G)	0 (T)	1 (T)	15 (CACAA)	25 (GTG)
	mut 2		2 (AAC)				5 (CATAG)	6 (ATA)
	mut 3		1				11 (TACAG)	3 (ATC)
	mut 4		1					1 (ATT)
	<i>total</i>	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>
# Genotypes inferred by basic TB- EFI analysis (Max. signal interpretation)	wt	59 (GCG)	50 (GAC)	55 (A)	61(C)	60 (G)	32 (CACAG)	28 (ATG)
	mut 1	2 (GTG)	9 (GGC)	6 (G)	0 (T)	1 (T)	15 (CACAA)	25 (GTG)
	mut 2		2 (AAC)				5 (CATAG)	7 (ATA)
	mut 3						9 (TACAG)	
	ND	0	0	0	0	0	0	0
<i>total</i>	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>
Basic TB-EFI genotypes reproducibility		97%	97%	95%	100%	100%	83%	91%

ND= Not Determined. wt=wild-type. mut=mutant.

Reference sample set amplifications and hybridizations were performed thrice independently. The two first repetitions were used to compute reproducibility, the third repetition was used to confirm results that were discordant with pyrosequencing (**Appendix file A.3**). First analyses were performed using “basic analysis” *i.e.* by identifying maximal MFI values for each locus. Reproducibility was very high for *gyrA* and *rrs* loci (95 to 100%) and good for *eis* promoter and *embB* gene (83 and 91%; **Table 5**). Very similar profiles were obtained in all repetitions; however, in some instances, the probe exhibiting the maximal signals differed from one repetition to the other: when signal/noise ratios were relatively high, experimental variations lead to different maximal signals (data not shown).

Reproducibility and performance of upgraded statistical TB-EFI prediction of genotypes

Table 6- Reproducibility and performance of upgraded genotype analysis by TB-EFI for the reference sample

		<i>gyrA_90</i>	<i>gyrA94</i>	<i>rrs1401</i>	<i>rrs1402</i>	<i>rrs1484</i>	<i>eis-prom</i>	<i>embB306</i>
# Genotypes inferred by Pyro-sequencing	wt	57 (GCG)	46 (GAC)	54 (A)	61(C)	60 (G)	30 (CACAG)	26 (ATG)
	mut1	4 (GTG)	11 (GGC)	7 (G)	0 (T)	1 (T)	15 (CACAA)	25 (GTG)
	mut2		2 (AAC)				5 (CATAG)	6 (ATA)
	mut3		1				11 (TACAG)	3 (ATC)
	mut4		1					1 (ATT)
	<i>total</i>		<i>61</i>	<i>61</i>	<i>61</i>	<i>0</i>	<i>61</i>	<i>61</i>
# Confirmed genotypes inferred by TB-EFI (Genotypes concordant between Max. signal and automatized interpretations)	wt	54 (GCG)	45 (GAC)	54 (A)	59(C)	60 (G)	30 (CACAG)	27* (ATG)
	mut1	2 (GTG)	9 (GGC)	6 (G)	0 (T)	1 (T)	14 (CACAA)	21 (GTG)
	mut2		2 (AAC)				5 (CATAG)	4 (ATA)
	mut3						10 (TACAG)	
	ND	5	5	1	2	0	2	9
	<i>total</i> %	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>	<i>61</i>
ND	8%	8%	2%	3%	0%	3%	15%	
Upgraded TB-EFI reproducibility		100%	98%	100%	100%	100%	100%	98%
Concordance upgraded TB-EFI and Pyrosequencing		100%	100%	100%	100%	100%	100%	98%*

ND= Not Determined. wt=wild-type. mut=mutant.

*The only discordant data for *embB* genotyping between upgraded TB-EFI and Pyrosequencing concerns a sample with clearcut data on TB-EFI on 3 repetitions suggesting an error in pyrosequencing or during sample or result transfer.

We improved data interpretation by taking into account the noise as detected in control samples using allele-call principle. When noise was relatively high, we lowered allele-call threshold to retrieve more genotypes. To benefit from this higher sensitivity without risking false assignments, only genotypes that are conform to basic analysis are validated by this “upgraded” analysis.

We obtained relatively low proportion of inconclusive results (0-8% of undetermined data - ND in **Table 6-** except for embB306 locus that lead to 15% indetermination). We obtained 100% concordance with pyrosequencing for all determined data (the single possible discordance was repeatedly observed with clear signals in 3 TB-EFI repetitions suggesting error either during control analysis or during sample or results transfer).

Performance of TB-EFI for phenotype prediction

Using genotypes at loci involved in resistance, we predicted phenotypic resistance based on the rules indicated in **Table 4**. We compared the predictions of the 61 reference samples to their phenotypic profiles obtained in liquid culture. As expected due to good concordance between pyrosequencing and TB-EFI genotyping, we observed high levels of concordance for all antibiotics, except for ethambutol and fluoroquinolones (72% and 77%, **Table 7**).

Most samples that were predicted Susceptible by TB-EFI and were resistant by phenotypic DST were mutated at loci other than those covered by TB-EFI. For instance, for FLQ resistance, one sample was mutated at codon 91, two at untargeted mutations at *gyrA94* codon. The other eight isolates had no detected mutation by pyrosequencing. We also observed that for most samples for which no phenotype could be predicted (Not Determined samples), one of the targeted locus involved in resistance carried a mutation other than the targeted one. For instance, for FLQ resistance, among the 6 samples with unpredicted phenotype, five were carrying a resistant mutation in *gyrA 94* or *gyrA 91* codon (**Appendix table A.1**).

For SLID resistance, performance was much higher (concordance of 90 to 98%; $\kappa > 0.6$). Of note, all samples mutated at *rrs 1401* were associated with CAP resistance. Mutations in *eis* promoter were partly associated with resistance (**Table 7**; **Appendix table A.1**).

Concordantly, on the subset of samples for which we had HAIN MTBDRs/ v1.0 test results, concordance between the two tests was 92%, and TB-EFI retrieved four more resistance information (7%) as compared to HAIN (**Appendix file A.1**).

Table 7- Agreement between TB-EFI genotypic resistance and phenotypic DST and partial resistance prediction inside the reference sample set.

Genotypic resistance (TB-EFI)		Phenotypic resistance (experimental)				Concordance	Kappa
		R	S	ND	Total		
FQ	R	13			13	77 [67-89]	0.54
	S	12	29		41		
	ND	6	1		7		
	Total	31	30	0	61		
KAN	R	14	2		16	90 [80-99]	0.78
	IR	10	6		16		
	S	2	20		22		
	ND	3	4		7		
	Total	29	32		61		
AMK	R	4		3	7	96 [91-100]	0.87
	IR	7	7		14		
	S	1	26	6	33		
	ND	2	5		7		
	Total	14	38	9	61		
CAP	R	1			1	98 [94-100]	0.66
	IR	6			6		
	S	1	51	1	53		
	ND	1			1		
	Total	9	51	1	61		
EMB	R	25	4		29	72 [60-84]	0.44
	S	12	16		28		
	ND	4			4		
	Total	41	20		61		

Ability of TB-EFI to predict resistance for uncharacterized samples

We used TB-EFI to analyze a set of 43 isolates from Kazakhstan. Despite a relatively high level of indetermination for each locus ($\approx 30\%$), we detected a high level of mutations: 12% of the samples being mutated at *gyrA90*, 21% at *gyrA94* in one of the two targeted mutations, 35% at *rrs* 1401 locus, 63% at *embB* 306 codon. We predicted a total of 9 XDR samples.

We could also implement partial version of TB-EFI onto 91 samples from Brazil and Pakistan. We observed much higher levels of susceptibility for Brazil than for Pakistan sample (79 to 85% vs 39 to 42% susceptibility to all second-line drugs, data not shown). The

level of indetermination was high in the Pakistan set (39 to 49% indetermination) but very low for the Brazilian sample (2%). Among the determined samples, we could detect one case of XDR-TB in Brazil and another one in Pakistan.

Regarding diversity of mutations conferring FLQ resistance, on the restricted sample which was analyzed by the full TB-EFI test, we could detect no statistical difference between the prevalence between GCG90GTG (n=5; 36%), GAC94GGC (n=3; 21%) and GAC94AAC (n=6; 43%) mutations. When including the samples analyzed with a restricted version of TB-EFI excluding *gyrA*90 locus, GGC mutation was significantly more frequent as reported in other studies (n=12; 60% of the 94 locus mutations; **Appendix file A.4**). We detected a high level of mutations in *rrs* 1401 (n=31) and very low levels of mutations in *rrs* 1402 (n=1) and in *eis* promoter (n=3). *embB* 306 GTG mutation was frequent (n=75, 33% of all samples), ATA mutation being detected in only 1 sample out of the 43 samples from Kazakhstan (**Appendix file A.4**).

Discussion

Why does it remain interesting to develop new kits on TB resistance detection at the time of WGS?

The rapid detection of drug-resistant TB cases is a clinical demand that by now the WGE cannot supply due to the cost and sophistication of this technology (Somoskovi 2015; Dominguez, Boettger et al. 2016). TB-EFI development was undertaken using the LX200 and xMAP® technology, aiming to rapidly apply it to the more recent generation of MagPix® LED-based fluorescence imager analyzers (Ocheretina, Merveille et al. 2013). Given the results obtained with previous techniques on the two machines, we are confident in achieving our next step, the transfer on MagPix® analyzers of the TB-EFI.

TB-EFI method detects nucleotide polymorphisms (SNP) in target genes of drugs that are known to confer resistance. It covers mutations on *gyrA* (GCG90GTG, GAC94GGC and AAC), *rrs* gene (A1401G, C1402CT, G1484T), *eis* promoter (G-10A, C-12T, C-14T) and *embB* (ATG306GTG and ATA). It is associated to an automatized analysis of the Luminex raw data inspired from MOL-PCR (Wuyts, Mattheus et al. 2015). The biostatistical analytic tool we developed provides not only detailed information for every single sample (genotype and predicted phenotype), but also summarizes the number of any genotypes and phenotypes in the whole sample.

Other tests target similar sets of loci as ours, some of them having been validated on clinical samples (Bergval, Sengstake et al. 2012; Singh, Bhan et al. 2013; Chakravorty, Simmons et al. 2017) (**Table 1**). The peculiarities of our test, associated with the use of Luminex systems, are its high throughput (2 sets of 91 samples in one day by a single operator starting from DNA sample with a single PCR machine and single Luminex platform including analysis, longest step being the PCR), its possible combination with other analyses such as spoligotyping or TB-SPRINT, and its relatively low cost. As NAATs are meant to rapidly identify resistant strains, and phenotypic DST is still recommended on any susceptible isolate to confirm susceptibility, the absence of some rare mutations in TB-EFI test may not be unacceptable (Gilpin, Korobitsyn et al. 2016).

TB-EFI efficiency

TB-EFI efficiently detected all targeted mutations in our reference sample (100% concordance with pyrosequencing on classified isolates, ~5% ND) albeit with lower efficiency for EMB resistance (98% concordance on classified isolates, 15% ND, **Table 6**).

Regarding resistance prediction, few true discordances between TB-EFI genotypic resistance profiles and phenotypic DST were detected. A first category was due to isolates having a mutation that is not included in TB-EFI. However, in most cases, as these untargeted mutations lie very close to the targeted one, they impair hybridization signals of both wild-type and mutant probes. Accordingly, among the 12 samples for which TB-EFI results were indeterminate for *gyrA* genotype, around 50% (n=5; an additional one having no pyrosequencing data) harbored a near-by mutation. Previously undescribed mutations could also contribute to this first type of discordances.

A second category of discordances between TB-EFI genotypic resistance and phenotypic DST was due to leakiness of some mutations: these mutations provide intermediate resistance so that around 30% of the isolates are classified as susceptible based on *in vitro* testing, although they carry a mutation at position eis -14 or -10 as recently characterized by Kambli et al. (Kambli, Ajbani et al. 2015). Existence of leaky mutations should in fact be taken into account, as proposed in TB-EFI outputs: resistance to SLIDs due to *rrs* mutations is leading to higher MIC than eis promoter mutations. As a consequence, inclusions of SLIDs in the treatment of a patient carrying a strain with eis promoter mutations may be considered, using higher doses than basic treatment. We also detected that eis promoter -14 mutation was linked to CAP resistance in one isolate. This promoter mutation could be responsible for a leaky

phenotype or another mutation in the genome of this isolate could have conferred CAP resistance.

A third category could be due to the use of lower critical concentrations during the Phenotypic DST. Indeed, nine of our reference samples were passed onto media with lower antibiotic concentrations (**Table 2**, see footnote). However, this subset did not harbor a higher level of discordance than the rest of the samples.

Other rare discrepancies existed between phenotypic DST and both TB-EFI and pyrosequencing or between phenotypic DST and TB-EFI only. These rare discrepancies could be due to contamination, sample exchange during the transmission of sample between collaborating laboratories, or heteroresistance (Eilertson, Maruri et al. 2014; Operario, Koepfel et al. 2017).

The relatively good performances of TB-EFI test for predicting resistance are further supported by resistance prevalence in the uncharacterized samples we analyzed: we expected around 1% of FLQ resistance in the Brazilian set and actually detected 2%; we expected around 15% of FLQ resistant isolates in the set from Pakistan and detected 13% samples of the genotyped samples carrying one of the two targeted mutations in *gyrA* gene. In addition, as underlined above, samples that could not be assigned a genotype (ND) at this locus may have ~50% chance of being resistant due to mutations at the same or at neighboring loci (codons 90 or 91). Altogether, mutant prevalence at the QRDR may be as high as 40% in this set of samples.

We used reference *M. tuberculosis* DNAs extracted from solid cultures by either classical CTAB method or by thermolysis and could not detect differences between these two extraction methods. However, sample quality remains a bottleneck to apply diagnostic tests such as TB-EFI. To apply it onto sputum, an important step will be the improvement in biological samples DNA extraction methods. New levers are being discovered such as methods improving handling of Thermodynamically Ultra-Fastened (TUF) DNA (McGinn, Bauer et al. 2016). In this regard, with a 65-66% genome GC-content, *Mycobacterium tuberculosis* complex still constitutes a challenging organism to be highly efficiently sequenced or analyzed whatever the quality of biological sample is.

Regarding potential XDR-TB cases, assuming that isolates carrying second-line resistances were also resistant to rifampicin and isoniazid (which in practice should be confirmed), we detected 9 such cases in our reference sample from Spain and Lithuania, one case in the set from Brazil, one in the Pakistan set. TB-EFI together with its analytic add-on is therefore an efficient tool for rapid identification of potential XDR-TB cases. By providing both mutation

and likely phenotype, it meets recommendations of TBNET/RESIST-TB networks regarding the indication of exact mutations present in the samples for better adjustment of patients' drug regimen (Dominguez, Boettger et al. 2016).

Altogether, although not targeting all loci involved in resistance (**Table 1**), TB-EFI was able to efficiently predict most cases of resistance to FLQ, second-line injectable drugs (KAN, AMK, CAP) and EMB in our hands, with products prepared on-site ensuring optimal test conditions. Further studies showing good implementation in other laboratory will be needed to confirm the robustness of this test (Mattocks, Watkins et al. 2010).

Conclusion

We propose here a new technique to predict resistance to second-line drugs and EMB using a flow-cytometry microbead-based assay. Our test is analyzed with an easy-to-handle analytic spreadsheet implemented in Excel 2016® that provides rapid and accurate prediction of genotypes, detailed resistances and possible XDR prediction. Performance proved not significantly different from MTBDRs/ v1.0 with slightly higher performance for SLID resistance prediction. Our test detected unambiguously 50% of the XDR samples present in our set of reference and allowed to suspect XDR cases in the uncharacterized sets from Pakistan, Brazil and Kazakhstan. As it is meant for high throughput, TB-EFI test might be particularly useful for retrospective follow-up of XDR-TB emergence dynamics for which the slightly lower sensitivity to FLQ resistance as compared to other tests may not be unacceptable. Test on new DNA extracted from clinical samples by improved methods will be undertaken to test whether TB-EFI can serve as a rapid first-line tool to identify best anti-TB regimen for MDR-TB patients.

Acknowledgments

We are grateful to Luminex Corp, Austin, TX and Luminex BV for technical support and to Dr. Pasquale Cutolo and Mr. Clément Ripoll, Beamedex SAS (Orsay, France) for technical and financial support. BJK acknowledges a CIFRE grant co-financed by ANRT and Beamedex SAS.

Author Contributions

BJK performed most experiments and wrote the first draft of the manuscript. BMM contributed to experiments, first analyses, and improvement of the manuscript. HMG performed the analyses on the isolates from Brazil. MKG contributed by coaching first BJK training and contributions to the manuscript. LOSV contributed to tests on the ethambutol probes. MHS, EP provided DNA samples, and MHS contributed to the improvement of the manuscript. JD organized DNA transfers from samples prepared in his laboratory and contributed to the improvement of the manuscript. CS conceived the study, designed first reagents with BJK, coordinated LOSV and HMG participation to the project, and contributed to the manuscript. GR participated to the study design, performed the automation of raw data analysis, analyzed the results and wrote the manuscript.

Competing interests statement

BJ Klotoe PhD fellowship at the University of Paris-Saclay was financed by the Beamedex® company that produces and distributes research and diagnostics kits as RUO (research use only). All other authors declare no conflict of interest.

List of Tables:

Table 1 - Mutations conferring resistance to EMB and second-line drugs and comparison of various NAAT SNP panels.

Table 2: Phenotypic resistance to second-line and ethambutol drugs in the reference set of 61 isolates.

Table 3: Primers and probes sequences.

Table 4: Phenotype prediction using TB-EFI genotype predictions

Table 5: Reproducibility of basic genotype analysis by TB-EFI for the reference sample

Table 6: Reproducibility and performance of upgraded genotype analysis by TB-EFI for the reference sample.

Table 7: Agreement between TB-EFI phenotype prediction and phenotypic DST on the reference sample set.

List of appendixes:

Appendix table A.1: Reference set profiles by phenotypic DST, pyrosequencing, and TB-EFI.

Appendix file A.2: Additional information regarding allele-call technique.

Appendix file A.3: Study design.

Appendix file A.4: Mutations detected in 3 uncharacterized samples (KAZ, PAK, BRA)

References

- Avalos, E., Catanzaro, D., Catanzaro, A., Ganiats, T., Brodine, S., Alcaraz, J. and Rodwell, T.** (2015). Frequency and geographic distribution of *gyrA* and *gyrB* mutations associated with fluoroquinolone resistance in clinical *Mycobacterium tuberculosis* isolates: a systematic review. *PLoS One* **10**, e0120470.
- Bergval, I., Sengstake, S., Brankova, N., Levterova, V., Abadia, E., Tadumaze, N., Bablishvili, N., Akhalaia, M., Tuin, K., Schuitema, A. et al.** (2012). Combined Species Identification, Genotyping, and Drug Resistance Detection of *Mycobacterium tuberculosis* Cultures by MLPA on a Bead-Based Array. *PLoS One* **7**, e43240.
- Bradley, P., Gordon, N. C., Walker, T. M., Dunn, L., Heys, S., Huang, B., Earle, S., Pankhurst, L. J., Anson, L., de Cesare, M. et al.** (2015). Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. *Nat Commun* **6**, 10063.
- Brites, D. and Gagneux, S.** (2012). Old and new selective pressures on *Mycobacterium tuberculosis*. *Infect Genet Evol* **12**, 678-85.
- Brossier, F., Guindo, D., Pham, A., Reibel, F., Sougakoff, W., Veziris, N. and Aubry, A.** (2016). Performance of the New Version (v2.0) of the GenoType MTBDRsl Test for Detection of Resistance to Second-Line Drugs in Multidrug-Resistant *Mycobacterium tuberculosis* Complex Strains. *J Clin Microbiol* **54**, 1573-80.
- Cabibbe, A. M., Miotto, P., Moure, R., Alcaide, F., Feuerriegel, S., Pozzi, G., Nikolayevskyy, V., Drobniowski, F., Niemann, S., Reither, K. et al.** (2015). Lab-on-Chip-Based Platform for Fast Molecular Diagnosis of Multidrug-Resistant Tuberculosis. *J Clin Microbiol* **53**, 3876-80.
- Chakravorty, S., Simmons, A. M., Rowneki, M., Parmar, H., Cao, Y., Ryan, J., Banada, P. P., Deshpande, S., Shenai, S., Gall, A. et al.** (2017). The New Xpert MTB/RIF Ultra: Improving Detection of *Mycobacterium tuberculosis* and Resistance to Rifampin in an Assay Suitable for Point-of-Care Testing. *MBio* **8**.
- Chernyaeva, E., Fedorova, E., Zhemkova, G., Korneev, Y. and Kozlov, A.** (2013). Characterization of multiple and extensively drug resistant *Mycobacterium tuberculosis* isolates with different ofloxacin-resistance levels. *Tuberculosis (Edinb)* **93**, 291-5.
- Chun, J. Y., Kim, K. J., Hwang, I. T., Kim, Y. J., Lee, D. H., Lee, I. K. and Kim, J. K.** (2007). Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of *CYP2C19* gene. *Nucleic Acids Res* **35**, e40.
- Dantas, N. G. T., Suffys, P. N., Carvalho, W. D. S., Gomes, H. M., Almeida, I. N., Figueiredo, L. J. A., Goncalves, A. D., Gomgnimbou, M. K., Refregier, G., Sola, C. et al.** (2017). Correlation between the BACTEC MGIT 960 culture system with Genotype MTBDRplus and TB-SPRINT in multidrug resistant *Mycobacterium tuberculosis* clinical isolates from Brazil. *Mem Inst Oswaldo Cruz* **112**, 769-774.
- Dominguez, J., Boettger, E. C., Cirillo, D., Cobelens, F., Eisenach, K. D., Gagneux, S., Hillemann, D., Horsburgh, R., Molina-Moya, B., Niemann, S. et al.** (2016). Clinical implications of molecular drug resistance testing for *Mycobacterium tuberculosis*: a TBNET/RESIST-TB consensus statement. *Int J Tuberc Lung Dis* **20**, 24-42.
- Dunbar, S. A.** (2006). Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection. *Clin Chim Acta* **363**, 71-82.
- Eilertson, B., Maruri, F., Blackman, A., Herrera, M., Samuels, D. C. and Sterling, T. R.** (2014). High proportion of heteroresistance in *gyrA* and *gyrB* in fluoroquinolone-resistant *Mycobacterium tuberculosis* clinical isolates. *Antimicrob Agents Chemother* **58**, 3270-5.
- Eldholm, V., Monteserin, J., Rieux, A., Lopez, B., Sobkowiak, B., Ritacco, V. and Balloux, F.** (2015). Four decades of transmission of a multidrug-resistant *Mycobacterium tuberculosis* outbreak strain. *Nat Commun* **6**, 7119.
- Feuerriegel, S., Schleusener, V., Beckert, P., Kohl, T. A., Miotto, P., Cirillo, D. M., Cabibbe, A. M., Niemann, S. and Fellenberg, K.** (2015). PhyResSE: a Web Tool Delineating *Mycobacterium tuberculosis* Antibiotic Resistance and Lineage from Whole-Genome Sequencing Data. *J Clin Microbiol* **53**, 1908-14.

- Georghiou, S. B., Magana, M., Garfein, R. S., Catanzaro, D. G., Catanzaro, A. and Rodwell, T. C.** (2012). Evaluation of genetic mutations associated with Mycobacterium tuberculosis resistance to amikacin, kanamycin and capreomycin: a systematic review. *PLoS One* **7**, e33275.
- Gilpin, C., Korobitsyn, A. and Weyer, K.** (2016). Current tools available for the diagnosis of drug-resistant tuberculosis. *Ther Adv Infect Dis* **3**, 145-151.
- Gomgnimbou, M. K., Hernandez-Neuta, I., Panaiotov, S., Bachiyiska, E., Palomino, J. C., Martin, A., Del Portillo, P., Refrégier, G. and Sola, C.** (2013). "TB-SPRINT: TUBERCULOSIS-SPOLIGO-RIFAMPIN-ISONIAZID TYPING"; an All-in-One assay technique for surveillance and control of multi-drug resistant tuberculosis on Luminex® devices. *J Clin Microbiol* **51**, 3527-3534.
- Hess, A. S., Shardell, M., Johnson, J. K., Thom, K. A., Strassle, P., Netzer, G. and Harris, A. D.** (2012). Methods and recommendations for evaluating and reporting a new diagnostic test. *Eur J Clin Microbiol Infect Dis* **31**, 2111-6.
- Hillemann, D., Rusch-Gerdes, S. and Richter, E.** (2009). Feasibility of the GenoType MTBDRsl assay for fluoroquinolone, amikacin-capreomycin, and ethambutol resistance testing of Mycobacterium tuberculosis strains and clinical specimens. *J Clin Microbiol* **47**, 1767-72.
- Kambli, P., Ajbani, K., Sadani, M., Nikam, C., Shetty, A., Udadia, Z., Georghiou, S. B., Rodwell, T. C., Catanzaro, A. and Rodrigues, C.** (2015). Defining multidrug-resistant tuberculosis: correlating GenoType MTBDRplus assay results with minimum inhibitory concentrations. *Diagn Microbiol Infect Dis* **82**, 49-53.
- Kim, K., Yang, J. S., Choi, H. B. and Lee, S. H.** (2018). Detection of resistance to fluoroquinolones and injectable drugs among antituberculosis drugs by allele-specific primer extension on a microsphere-based platform. *J Microbiol Methods* **144**, 111-116.
- Lacoma, A., Molina-Moya, B., Prat, C., Pimkina, E., Diaz, J., Dudnyk, A., Garcia-Sierra, N., Haba, L., Maldonado, J., Samper, S. et al.** (2015). Pyrosequencing for rapid detection of Mycobacterium tuberculosis second-line drugs and ethambutol resistance. *Diagn Microbiol Infect Dis* **83**, 263-9.
- Lee, S. H., Choi, H. B., Yu, S. Y., Chang, U. J., Kim, C. K. and Kim, H. J.** (2015). Detection of first-line anti-tuberculosis drug resistance mutations by allele-specific primer extension on a microsphere-based platform. *Ann Lab Med* **35**, 487-93.
- Malik, S., Willby, M., Sikes, D., Tsodikov, O. V. and Posey, J. E.** (2012). New insights into fluoroquinolone resistance in Mycobacterium tuberculosis: functional genetic analysis of gyrA and gyrB mutations. *PLoS One* **7**, e39754.
- Mattocks, C. J., Watkins, G., Ward, D., Janssens, T., Bosgoed, E. A., van der Donk, K., Ligtenberg, M. J., Pot, B., Theelen, J., Cross, N. C. et al.** (2010). Interlaboratory diagnostic validation of conformation-sensitive capillary electrophoresis for mutation scanning. *Clin Chem* **56**, 593-602.
- McGinn, S., Bauer, D., Brefort, T., Dong, L., El-Sagheer, A., Elsharawy, A., Evans, G., Falk-Sorqvist, E., Forster, M., Fredriksson, S. et al.** (2016). New technologies for DNA analysis--a review of the READNA Project. *N Biotechnol* **33**, 311-30.
- Molina-Moya, B., Agonafir, M., Blanco, S., Dacombe, R., Gomgnimbou, M. K., Spinasse, L., Gomes-Fernandes, M., Datiko, D. G., Edwards, T., Cuevas, L. E. et al.** (2018). Microbead-based spoligotyping of Mycobacterium tuberculosis from Ziehl-Neelsen-stained microscopy preparations in Ethiopia. *Sci Rep* **8**, 3987.
- Molina-Moya, B., Gomgnimbou, M. K., Lafoz, C., Lacoma, A., Prat, C., Refregier, G., Samper, S., Dominguez, J. and Sola, C.** (2017). Molecular Characterization of Mycobacterium tuberculosis Strains with TB-SPRINT. *Am J Trop Med Hyg* **97**, 806-809.
- Molina-Moya, B., Lacoma, A., Prat, C., Pimkina, E., Diaz, J., Garcia-Sierra, N., Haba, L., Maldonado, J., Samper, S., Ruiz-Manzano, J. et al.** (2015). Diagnostic accuracy study of multiplex PCR for detecting tuberculosis drug resistance. *J Infect* **71**, 220-30.
- Ocheretina, O., Merveille, Y. M., Mabou, M. M., Escuyer, V. E., Dunbar, S. A., Johnson, W. D., Pape, J. W. and Fitzgerald, D. W.** (2013). Use of Luminex MagPlex(R) Magnetic Microspheres for High-Throughput Spoligotyping of M. tuberculosis Isolates in Port-au-Prince, Haiti. *J Clin Microbiol*.

Operario, D. J., Koepfel, A. F., Turner, S. D., Bao, Y., Pholwat, S., Banu, S., Foongladda, S., Mpagama, S., Gratz, J., Ogarkov, O. et al. (2017). Prevalence and extent of heteroresistance by next generation sequencing of multidrug-resistant tuberculosis. *PLoS One* **12**, e0176522.

Perez-Garcia, F., Ruiz-Serrano, M. J., Lopez Roa, P., Acosta, F., Perez-Lago, L., Garcia-De-Viedma, D. and Bouza, E. (2017). Diagnostic performance of Anyplex II MTB/MDR/XDR for detection of resistance to first and second line drugs in *Mycobacterium tuberculosis*. *J Microbiol Methods* **139**, 74-78.

Phelan, J., Coll, F., McNerney, R., Ascher, D. B., Pires, D. E., Furnham, N., Coeck, N., Hill-Cawthorne, G. A., Nair, M. B., Mallard, K. et al. (2016). *Mycobacterium tuberculosis* whole genome sequencing and protein structure modelling provides insights into anti-tuberculosis drug resistance. *BMC Med* **14**, 31.

Sengstake, S., Bablishvili, N., Schuitema, A., Bzekalava, N., Abadia, E., de Beer, J., Tadumadze, N., Akhalaia, M., Tuin, K., Tukvadze, N. et al. (2014). Optimizing multiplex SNP-based data analysis for genotyping of *Mycobacterium tuberculosis* isolates. *BMC Genomics* **15**, 572.

Singh, J. A., Bhan, A. and Upshur, R. (2013). Diagnosis of drug-resistant TB and provision of second-line TB treatment in India: some ethical considerations. *Indian J Med Ethics* **10**, 110-4.

Somoskovi, A. (2015). Novel laboratory diagnostic tests for tuberculosis and their potential role in an integrated and tiered laboratory network. *Tuberculosis (Edinb)* **95 Suppl 1**, S197-9.

Starks, A. M., Gumusboga, A., Plikaytis, B. B., Shinnick, T. M. and Posey, J. E. (2009). Mutations at embB codon 306 are an important molecular indicator of ethambutol resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* **53**, 1061-6.

Takiff, H. E., Salazar, L., Guerrero, C., Philipp, W., Huang, W. M., Kreiswirth, B., Cole, S. T., Jacobs, W. R., Jr. and Telenti, A. (1994). Cloning and nucleotide sequence of *Mycobacterium tuberculosis* gyrA and gyrB genes and detection of quinolone resistance mutations. *Antimicrob Agents Chemother* **38**, 773-80.

WHO. (2015). Global Tuberculosis Report 2015, (ed.: WHO, Geneva, Switzerland).

WHO. (2016). Global Tuberculosis Report 2016, (ed.: WHO, Geneva, Switzerland).

WHO. (2017). Global Tuberculosis Report 2017, (ed.: WHO, Geneva, Switzerland).

Wuyts, V., Mattheus, W., Roosens, N. H., Marchal, K., Bertrand, S. and De Keersmaecker, S. C. (2015). A multiplex oligonucleotide ligation-PCR as a complementary tool for subtyping of *Salmonella* Typhimurium. *Appl Microbiol Biotechnol* **99**, 8137-49.

Yasmin, M., Gomgnimbou, M. K., Siddiqui, R. T., Refregier, G. and Sola, C. (2014). Multi-drug resistant *Mycobacterium tuberculosis* complex genetic diversity and clues on recent transmission in Punjab, Pakistan. *Infect Genet Evol* **27C**, 6-14.

Yin, F., Chan, J. F., Zhu, Q., Fu, R., Chen, J. H., Choi, G. K., Tee, K. M., Li, L., Qian, S., Yam, W. C. et al. (2016). Development and in-use evaluation of a novel Luminex MicroPlex microsphere-based (TRIOL) assay for simultaneous identification of *Mycobacterium tuberculosis* and detection of first-line and second-line anti-tuberculous drug resistance in China. *J Clin Pathol*.

Zager, E. M. and McNerney, R. (2008). Multidrug-resistant tuberculosis. *BMC Infect Dis* **8**, 10.

Zaunbrecher, M. A., Sikes, R. D., Jr., Metchock, B., Shinnick, T. M. and Posey, J. E. (2009). Overexpression of the chromosomally encoded aminoglycoside acetyltransferase eis confers kanamycin resistance in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **106**, 20004-9.

Zhao, L. L., Sun, Q., Liu, H. C., Wu, X. C., Xiao, T. Y., Zhao, X. Q., Li, G. L., Jiang, Y., Zeng, C. Y. and Wan, K. L. (2015). Analysis of embCAB mutations associated with ethambutol resistance in multidrug-resistant *Mycobacterium tuberculosis* isolates from China. *Antimicrob Agents Chemother* **59**, 2045-50.

Zhu, C., Zhang, Y., Shen, Y., Siu, G. K., Wu, W., Qian, X., Deng, G., Xu, Y., Lau, R., Fan, X. et al. (2012). Molecular characterization of fluoroquinolone-resistant *Mycobacterium tuberculosis* clinical isolates from Shanghai, China. *Diagn Microbiol Infect Dis* **73**, 260-3.

Etude 2 : Etude des sites d'insertion de la séquence d'insertion IS6110 au sein du locus NTF dans la famille « Beijing » de *Mycobacterium tuberculosis* par hybridation multiplexée sur Luminex®

Résumé

La transmission des souches Beijing de *Mycobacterium tuberculosis* constitue un problème de santé publique mondial. L'expansion de ces clones, particulièrement virulents, homogènes et adaptés à la résistance et l'ultra-résistance aux antituberculeux complique la prise en charge des patients. Ces clones sont responsables d'épidémies de tuberculose dans diverses régions.

Parmi les clones les premiers à être reportés figurent les souches « W » identifiées lors d'une épidémie à New-York dans les années 90, et qui sont résistantes aux antituberculeux (Rifampicine, Isoniazide, Streptomycine, Ethambutol, Kanamycine, Ethionamide), mais aussi le clone « W148 » identifié en Sibérie/Russie et responsable d'épidémie en Europe/Russie, ainsi que le clone Beijing responsable d'épidémie en Asie centrale. La différence entre ces différents clones résistants se situe au niveau de leur structure génomique, notamment au niveau du site d'insertion des séquences d'insertion *IS6110* au locus NTF. Le clone « W » épidémique originel dispose de deux copies de *IS6110* en amont et en aval du locus NTF, tandis que le W148, ne possède qu'une copie de *IS6110* en aval. Ce groupe de clones est dénommé « *modern Beijing* », et en absence d'*IS6110* dans NTF, on parle d'« *ancient Beijing* ».

L'objectif de ce projet est de développer le test « IS-NTF/RINT » une technique haut-débit détectant simultanément la séquence d'insertion *IS6110* dans le locus NTF ainsi que les mutations responsables de la multirésistance aux antituberculeux (R = Rifampicin, I = Isoniazide) dans les gènes associés en vue d'un diagnostic de génomique personnalisé, de prédiction génétique à la résistance aux antituberculeux. La logique de développement de ce test est aussi, de détecter des souches en grappes, qui constituent un risque d'extension des épidémies. Pour cela nous avons utilisé des souches de référence provenant de New-York qui étaient impliquées ou/non dans l'épidémie historique de New York (n = 47) et pour validation du test, nous utiliserons une collection de souche MDR (n = 169) provenant du Kazakhstan.

Avec le test IS-NTF/RINT nous avons détecté 18 souches épidémiques « W », 18 souches avec une seule copie de *IS6110* en NTF, soit 36 souches « modernes Beijing » et 11 souches dites « *ancient Beijing* ». Sur le plan de résistance, le test IS-NTF/RINT a détecté 23 souches MDR évoluant en trois grappes : grappe1 (3 souches MDR « *ancient Beijing* »), grappe2 (17 souches « W » MDR et « modernes Beijing »), grappe3 (3 souches MDR « *modern Beijing* » avec une copie de *IS6110* dans le NTF). La sensibilité et la spécificité du test IS-NTF/RINT sont de 100% respectivement. Sur les souches cliniques du Kazakhstan, le test IS-NTF/RINT a détecté 141 « *modern Beijing* » (spécificité, 99%) avec 101 MDR-

« modern Beijing » (spécificité, 94%) qui évoluent dans la grappe 3 et 20 non-MDR « modern Beijing ». Ainsi nous avons développé une méthode qui a permis à posteriori d'identifier les grappes de souches prédominant à New York et d'autres grappes prédominant aux Kazakhstan. Le test IS-NTF/RINT pourrait être une technique permettant de suivre la transmission des souches Beijing et aussi qui permettrait de contribuer au contrôle et la maîtrise des épidémies à souches Beijing dans différentes régions du Monde.

L'article ci-après sera prochainement soumis à la revue « *Tuberculosis* » comme *Short communication*.

Article 2 :

IS-NTF/RINT, a new method for the epidemiological surveillance of MDR L2/Beijing strains *Mycobacterium tuberculosis*

¹Bernice Klotoe, ²Natalia Kurepina, ³Elena Zholdibayeva, S. Panaiotov^{1,4} ²Barry Kreiswirth, , Sarah Sengstake, Richard Anthony, ¹Christophe Sola*, ¹Guislaine Refrégier*.

¹Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France

²Public Health Research Institute TB Center, International Center for Public, Newark, New Jersey, USA

³National Center for Biotechnology, Astana, Kazakhstan

*co-corresponding authors

Short Communications: These should be no more than 2,500 words, with up to 15 references and a maximum of 3 figures or tables.

Abstract

L2/Beijing lineage of *Mycobacterium tuberculosis* has caused the most virulent tuberculosis and Multi Drug Resistant (MDR) tuberculosis epidemics in several region of world. Among the most reported epidemics due to L2/Beijing strain are the “W” and “W148” from the New York City and Russian outbreaks. Among the difference between both strains was the genomic profile according to IS6110 in NTF locus. We present a high-throughput, microbead-based method, called IS-NTF/RINT for detection of IS in NTF and Rifampicin and Isoniazid Typing. This method provides tuberculosis diagnostic confirmation, detects IS6110 insertions in NTF locus revealing affiliation or not to modern L2/Beijing sublineage and detects mutations involved in resistance to Rifampicin and Isoniazid. This assay could be used for surveillance of transmission and expansion of MDR-L2/Beijing clinical isolates.

I. Introduction

Mycobacterium tuberculosis L2/Beijing is responsible of around one third of tuberculosis cases around the world with highest prevalence in Eastern Europe (27.9%), Asia (50%) and is poorly present in Africa (with the exception of coastal areas and South Africa) and in South America (Cattoir 2004; Demay, Liens et al. 2012); Hoffner S and al. 2017, (Ramazanzadeh and Sayhemiri 2014). L2/Beijing lineage originated in South East of China (Asia), likely in the province of Guangxi ((Wan, Liu et al. 2011), (Luo, Tao Luo et al. 2015) and was first described in 1995 based on similar IS6110-RFLP patterns and specific spoligotyping profiles characterized by the absence of spacers 1-34 and presence of most spacers 35-43 in the CRISPR region (van Soolingen, Qian et al. 1995), (Kamerbeek, Schouls et al. 1997); (Filliol, Ferdinand et al. 2002). The L2/Beijing comprises at least two major sublineages based on the IS6110 copy numbers in the noise transfer function (NTF) chromosomal region: “typical” and “atypical” Beijing strains. Most actively spreading isolates belong to the so-called “modern sublineage” characterized by, one or two insertion sequence IS6110 in the 5’ and 3’ part of Noise Transfer Region (NTF) region, and two mutations in 3R genes (mutT2 and ogt respectively at positions 58 and 12) (Shitikov, Kolchenko et al. 2017). Expansion of L2/Beijing strains, especially that of the modern sublineage into other countries may have resulted from massive population outflow, antibiotic treatment selection as suggested by the ability of this strain to acquire drug resistance (Ford et al, 2013), ability to escape from BCG vaccine-induced protection (Kremer, van-der-Werf et al. 2009) , Lopez and al 2003, (Tsenova, Ellison et al. 2005)), different immunogenic properties linked to higher virulence

(Dormans, Burger et al. 2004); (Manca, Reed et al. 2004), Rios-Barrera and al 2006; Lopez and al 2003; (Pheiffer 2005); (Ida Parwati, Reinout van Crevel et al. 2010); (Caws, Thwaites et al. 2008); (van Crevel, Nelwan et al. 2001); (Ribeiro, Gomes et al. 2014) and faster progression from latent infection to active tuberculosis (de Jong, Hill et al. 2008). One of the first multi-drug resistant (MDR) epidemics was caused by a modern Beijing strain termed “W” strain, characterized by an additional IS6110 insertion in the NTF locus ((Plikaytis, Marden et al. 1994). Since then, many MDR outbreaks involved others modern L2/Beijing isolates, like the outbreak identified in Europe/Russia (W148, VNTR) 100-32 and central Asia (CAO, VNTR 94-32) (Casali, Nikolayevskyy et al. 2014); (Merker, Blin et al. 2015),(Mokrousov, Narvskaya et al. 2008). In this study, we firstly aimed to develop a new high throughput technique using the Luminex technology, to diagnose tuberculosis detecting IS6110 insertion, to classify different Beijing genotype strains by typing IS6110 insertion in the NTF locus, and to identify the drug resistance markers. We previously developed the TB-SPRINT technique to target main resistance mutations associated with Rifampicin (RIF) and Isoniazid (INH) resistance simultaneously with spoligotyping (TB-SPRINT=TB-SPOL+TB-RINT, i.e. 43-Plex +16-Plex). We propose here “IS-NTF RINT” a technique that combines RINT to the detection of IS6110 in the NTF region. With this new developed technique we will be able to characterize strains in different L2/Beijing clusters according to genotype and MDR profiles for monitoring transmission, evolution and expansion of L2/Beijing strains. This technique may help in the improvement of the epidemiological surveillance of different L2/Beijing outbreaks.

II. Material and Methods

II.1. Material

This study relies on a total of 189 strains to set up and validate the technique. Forty-seven (n=47) clinical strains from New-York City (NYC), USA, were used to set up the method. Eighteen of them belonged to the first large historical MDR outbreak in NYC; double checked reference genotypes and partial phenotype characterizations were available for all these strains. All these samples were extracted by Cetyl-Trimethyl-Ammonium Bromide (CTAB) method as described previously (van Soolingen, 1995). A second set of 142 clinical strains from Kazakhstan with L2/Beijing spoligotype pattern and full genotype prediction (characterized by two methods : TB-SNPID, TB-SPRINT, part of another study, results to be described elsewhere) was used as a validation set. DNA was extracted by

thermolysis on loops of solid cultures. Genotype information concerning the IS6110 insertions in the NTF locus was first obtained on agarose gel (using Plikaytis procedure except for I60 probe changed to IS60B = 5' TGATGTGCTCCTTGAGTTCG 3' to amplify a 355 bp fragment; Plikaytis et al, 1993) and was independently confirmed by mutT2_58 mutation included in TB-SNPID (Bergval et al, 2012). Genotype information concerning resistance loci was obtained by TB-SPRINT (a Rifampicin (RIF) and Isoniazid (INH) resistance assay in combination with spoligotyping, Gomgnimbou et al, 2013) and TB-SNPID. The selected isolates had fully concordant profiles with the two reference methods used for each type of locus. BCG Paris strain and a sample from *Salmonella enterica* called Sara were used as negative controls.

II.2. IS-NTF/RINT principle, primers and probes

The RINT part of IS-NTF RINT targets the same SNP as those described by Gomgnimbou et al, (2013): 1) the *rpoB* mutations 435_GTC, 445_GAC, 445_TAC, 450_TTG, 450_TTG; 2) the *katG* mutations 315_C and 315_A; 3) the *inhA* promoter mutations -15_T and -8_A. Please note that for *rpoB* mutations, Gomgnimbou et al. (2013) referred to positions as in *E. coli* i.e. respectively 516, 526 and 531 in place of 435, 445 and 450. The two spanning probes which were included in the TB-SPRINT assay development were not included here, due to the inherent difficulties to interpret signal-off probes in high-throughput assays.

For the IS6110 typing in NTF, we amplified very short fragments (around 150 bp) of the IS6110 sequence using primers proposed by Eisenach et al. (1990), and designed new primers to amplify small fragments overlapping NTF insertion sites: Primer NTF5'_Fw : GGCTGCCTACTACGCTCAAC; NTF5'_Rv1: TCGACAACCCACTTCCCATG; NTF3'_Fw_new : TCACCACTGCTGCGTAATTCG; NTF3'_Rv_new: CCTGACATGACCCCATCCTTTC. We used a published probe to detect the IS6110 sequence (Brocolo et al, 2003). To detect insertions in NTF, we used sequences straddling NTF and IS6110 sequences: IS-NTF5'-probe-final (GGCGGTTCAACATGGTGAC); IS-NTF3'-probe2-final (ATGAGGTCGTGAACCGCC). Probes were coupled to microbeads as recommended by the manufacturer (www.luminex.corp).

The RINT part of IS-NTF/RINT targets the same SNP as those described by Gomgnimbou et al, (2013): 1) the *rpoB* mutations 435_GTC, 445_GAC, 445_TAC, 450_TTG, 450_TTG; 2) the *katG* mutations 315_C and 315_A; 3) the *inhA* promoter mutations -15_T and -8_A. Please note that for *rpoB* mutations, Gomgnimbou et al. (2013)

referred to positions as in *E. coli* i.e. respectively 516, 526 and 531 in place of 435, 445 and 450. All primers and probes were obtained from integrated DNA Technologies (IDT, Leuven, Belgium).

II.2. Amplification and detection on Luminex

The amplification reaction mixture contained 2 μ L of DNA and 23 μ L of reaction mixture (0.2 mM dNTPs, 1 μ M mix of primers, 0.25U GoTaq Flexi, 1.5mM MgCl₂, 1X GoTaq buffer) as recommended by the manufacturer (Promega Corp., Fitchburg, WI). The mixture was heated for 3 min at 96°C and subjected to 35 cycles of 30s at 96°C, 30s at 60°C, and 30s at 72°C. The rest of the procedure was performed as for TB-SPRINT (Gomgnimbou et al, 2013) and hybridized microbeads samples were analyzed on a Luminex 200® device (Luminex, Austin, TX).

II.3. Data interpretation

Median Fluorescent Intensities (MFI) reported by the Luminex® device were analyzed as follows: for the presence/absence of IS6110 and its insertions in NTF region, the standard cut-off for presence (MFI=150) was used with rare adjustment in case the DNA concentration was very low as evidenced by signals on other markers. For SNP genotyping in *rpoB*, *katG* and *inhA* genes, the genotypes were inferred using allele-call method as described in Klotoe et al (2018). Very similar profiles could be obtained using the maximal MFI (Klotoe et al 2018) but with caution when signals are inferior to MFI=40, or for *inhA* promoter profiles due to the high background associated with -8_A mutant probe.

III. Results

III.1. IS-NTF/RINT test design

The test we designed aims at identifying likely epidemic MDR L2/Beijing strains. To do so, our test targets IS6110 insertions in the NTF locus in genome of *M. tuberculosis* to diagnosis tuberculosis and to identify 1) modern strains and their possible affiliation to the “W”-strain L2/sublineage and 2) most frequent mutations implicated in RIF and INH resistance (Fig. 1). Both identifications are necessary to determine the different scalable cluster in each country. Each MFI value is converted to a positive or negative value (black/white squares in Fig. 1B) however the final interpretation method directly confronts

raw data of probes targeting the same site to identify the one reaching the highest levels as compared to the background.

III.2. Sensitivity and specificity on a reference laboratory sample

When tested on the 47 USA reference isolates, IS-NTF/RINT produced bimodal MFI values easily allowing to discriminate between positive and negative signals. Interpretation to predict genotypes could be performed for all targets except for three samples and all outbreak samples on katG_315 locus, due to a double mutation in “W” in katG : AGC>ACA). (n=0 out of 17 outbreak INHR strains; Supp. Table 1). In contrast, for non-outbreak strains, genotype at 315 katG codon was detected in all isolates (100%, n=40; Supp. Table 1; Table 1). The katG 315 genotype of outbreak strains (as identified by the double IS6110 insertion in NTF) proved thus to be mutated (ACC) unless there was a clear signal for the wild-type probe. When including this information, 231 out of 235 genotypes (98.3%) were predicted successfully, the four remaining staying undetermined. Altogether, IS-NTF/RINT exhibited 100% sensitivity and 100% specificity on successfully typed loci. The method detected the 18 outbreak strains, 18 modern L2/Beijing and 11 strains belonging to ancient Beijing clade or to other lineages, and a total of 23 MDR (Table 1). The developed technique compared with previous agarose gel results revealed 100% concordance.

III.3. Evaluation on clinical strains from Kazakhstan

A set of 216 L2/Beijing clinical strains from Kazakhstan with reference genotype information were subjected to IS-NTF/RINT. All were detected modern L2/Beijing by IS-NTF/RINT as compared to 216 as detected by the reference methods (fbpB_238, mutT2_58, acs_1551 mutations and deletions RD105, RD131 detecting by TB-SNPID, and the presence of IS insertion at the 3' end of NTF as detected using Plikaytis method on agarose gel). Sensibility to detect modern L2/Beijing was 100% and specificity could not be evaluated due to the composition of the samples (We do not have in this collection of sample ancient L2/Beijing).

Regarding the detection of resistance mutations, 435 codon of rpoB was often predicted to be mutated, reaching irrelevant mutation frequency. We thereafter excluded this locus from interpretation. As this mutation is rare, we predicted genotypes considering that no mutation had occurred as this position. A total of 190 strains retrieved a genotype in full accordance with the reference genotype. Altogether, concordance concerned 93% of the sites investigated (n=953 out of 1033 data points). We compared modern L2/Beijing MDR prediction by IS-

NTF/RINT to MDR prediction using reference genotypes. Only 9 samples could not be classified as MDR or non-MDR (6%). A total of 146 isolates were correctly predicted as modern MDR-Beijing, corresponding to a sensitivity of 94% (146/155) (Table 2). Twenty out of a total of 26 strains predicted to be non-MDR were detected as such by IS-NTF/RINT (sensitivity=77%). Seven strains predicted to be MDR were detected mono-resistant with katG-315C mutation by IS-NTF/RINT. 32 mono resistant modern Beijing were identified by the test with katG-315C mutation. Global concordance was very good (Cohen's Kappa =0.75). Among non-MDR, IS-NTF/RINT erroneously predicted sensitivity in a single isolate out of 14 INH-resistant samples.

IV. Discussion and conclusion

We set up IS-NTF RINT method to follow-up resistant Beijing strains by identifying most common RIF and INH resistance mutations (445 and 450 positions in rpoB, 315 position in katG, and positions -15 and -8 inhA promoter) and by identifying IS6110 insertions in the NTF locus. We obtained perfect concordance on the reference sample and very good performance and a large set of clinical strains from Kazakhstan. Altogether in this study, IS-NTF/RINT identified 163 modern L2/Beijing MDR strain, among which 17 belonged to the “W”-strain branch (from the outbreak of NYC), and three MDR ancient Beijing. With the assay, we identified 6 clusters according to genetic organization of IS6110 in NTF locus and drug resistance profiles (Fig1B). These 6 clusters are; W-strain modern Beijing (n = 17), most frequent MDR modern Beijing (n = 148), others MDR modern Beijing (n = 6), susceptible modern Beijing (n = 23), most frequent ancient Beijing (susceptible) (n = 8), rare MDR ancient Beijing (n = 3). Moreover, we tested IS-NTF/RINT on some L2/Beijing isolates from Peru, Mozambique, Bulgaria, Pakistan, and all of them were modern MDR L2/Beijing (result not shown here). The modern L2/Beijing with one IS6110 insertion were prevalent in several countries however in USA, the “W”-strain was historically more prevalent. We suggest that the “W”-strain, (outbreak strain in USA) were already endemic and in the course of evolution got one extra insertion IS6110 copy in the 5' of NTF locus. W148, the clinical isolate identified by IS6110-RFLP at the origin of the outbreak in Europe/Russia was also a modern MDR L2Beijing isolate characterized early in the 90s (Fig1A) ((Mokrousov, Narvskaya et al. 2012).

A first type of limitation of IS-NTF/RINT to detect modern MDR Beijing could be due to failure in identifying affiliation to the modern subbranch. Such a failure could be due to the absence of 3' NTF insertion in a modern strain as described by Merker et al (2015). The larger study by Shitikov et al however apparently confirmed a perfect link between the insertion and the affiliation to modern subbranch of Beijing (Shitikov et al). In this study, the authors identified that several ancient Beijing carried other IS6110 insertions in the NTF locus. As IS-NTF/RINT specifically targets the sequences straddling over IS6110 at the exact position described by Plikaytis et al, these other insertions should not produce false positive signals.

Another type of failure to identify modern MDR Beijing could be due to the failure to identify antibiotic resistance. This may happen if resistance is due to mutations outside the targeted sites. RIF Resistance due to 445 and 450 rpoB mutations is expected to constitute around 80% of all RIF Resistance, i.e. up to 20% of MDR-TB samples should be sequenced (Unissa et al, Tuberculosis 2017; (Gomgnimbou, Hernandez-Neuta et al. 2013), and around 70% for INH (Gomgnimbou, Hernandez-Neuta et al. 2013).

As compared to other genotyping methods to predict resistance and/or classification, IS-NTF/RINT cost is very effective (~4 euros) due to the use of a restricted number of targets (17-Plex) and the absence of need for the costly xTAG® Luminex microbeads as used by TB-SNPID, (Bergval, Sengstake et al. 2012). Regarding resistance phenotype prediction, sensitivity is above 75% which may be sufficient for first-line screening of resistance. Another advantage of this technique would be its high throughput since in one working day, as many as >90 samples can be processed.

The quality of the results is highly dependent on DNA concentration and quality, however the method is not more sensitive to impurities than any method based on PCR. If extracted by appropriate kits, this method could be able to be applied on clinical samples which remains to be tested in different setting to ensure reproducibility.

Acknowledgements

François-Xavier Weill is thanked for giving a Salmonella DNA sample to serve as negative control for IS6110 detection.

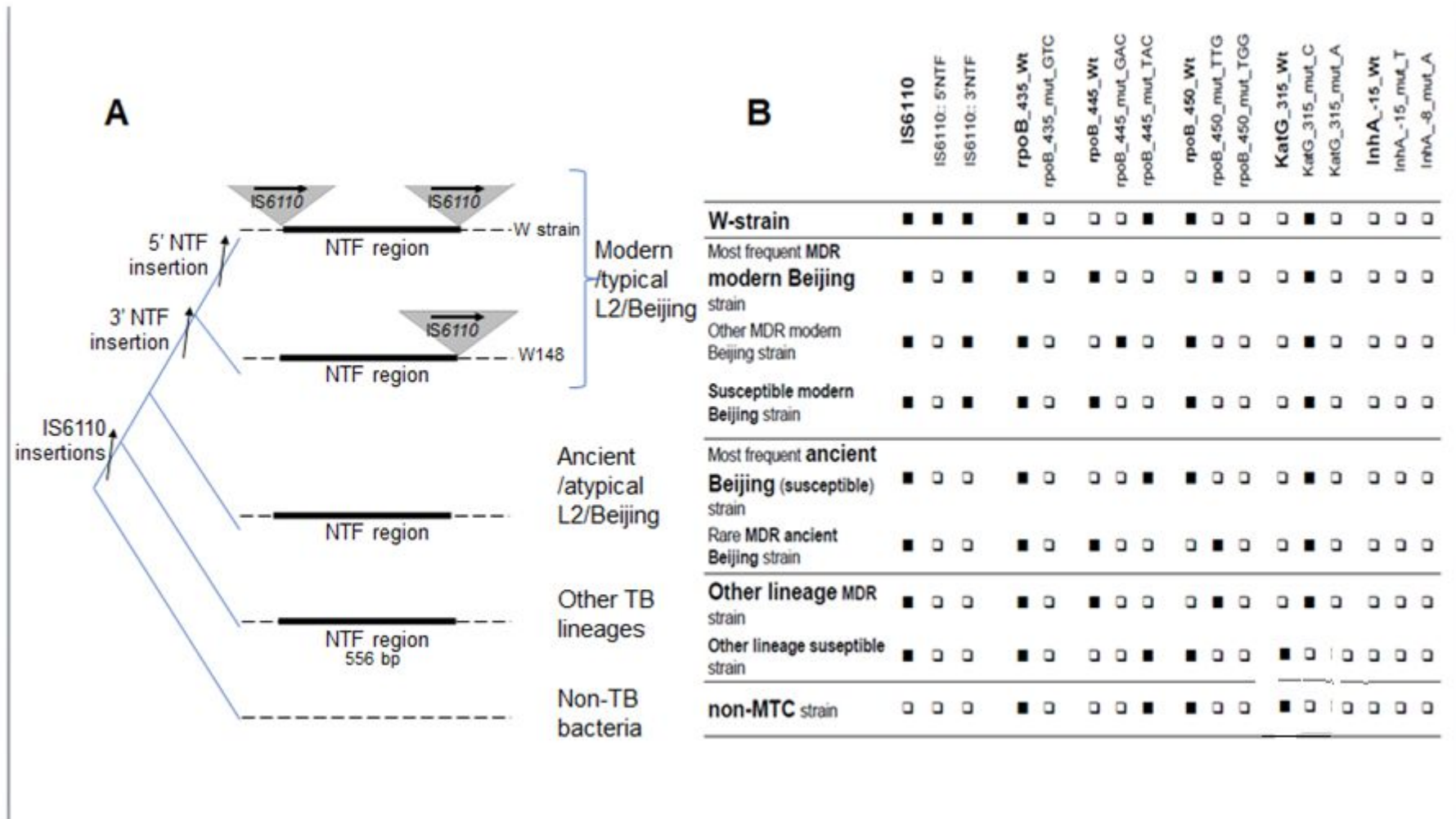


Figure 1 : Evolutionary scenario of IS6110 in NTF the locus of Beijing Strain (A) and corresponding identified cluster according to drug resistant (B).

Table 1- Concordance between IS-NTF/RINT and controls on the reference samples set.

N°	#	NY	IS-NTF RINT profiles			Control profiles		
			IS6110 profile	R. loc. profile	GDST	IS6110 on gel	TB_RINT	PhDST
1	2015	YES	2 IS6110::NTF	435?/445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
2	2706	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445?/315C	MDR
3	3452	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
4	3267	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
5	6138	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
6	8121	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
7	16852	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
8	1582	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
9	2470	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
10	15183	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
11	3542	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
12	4953	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
13	19556	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445?/315C	MDR
14	13223	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
15	19996	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
16	1858	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	MDR
17	19962	YES	2 IS6110::NTF	445TAC/315?	MDR	2 IS6110::NTF	445TAC/315C	ND
18	13250	YES	2 IS6110::NTF	445TAC	R ^R	2 IS6110::NTF	445TAC	ND
19	11455	NO	1 IS6110::NTF	WT	SUSC	1 IS6110::NTF	WT	SUSC
20	14031	NO	1 IS6110::NTF	WT	SUSC	1 IS6110::NTF	WT	SUSC
21	14907	NO	1 IS6110::NTF	WT	SUSC	1 IS6110::NTF	WT	SUSC
22	15803	NO	1 IS6110::NTF	WT	SUSC	1 IS6110::NTF	WT	SUSC
23	16377	NO	1 IS6110::NTF	WT	SUSC	1 IS6110::NTF	WT	SUSC
24	7716	NO	1 IS6110::NTF	WT	SUSC	1 IS6110::NTF	WT	ND
25	7718	NO	1 IS6110::NTF	WT	SUSC	1 IS6110::NTF	WT	ND
26	20935	NO	1 IS6110::NTF	445?/WT	ND	1 IS6110::NTF	WT	ND
27	12783	NO	1 IS6110::NTF	WT	SUSC	1 IS6110::NTF	WT	ND
28	6595	NO	1 IS6110::NTF	WT	SUSC	1 IS6110::NTF	WT	ND
29	20251	NO	1 IS6110::NTF	WT	SUSC	1 IS6110::NTF	WT	ND
30	20217	NO	1 IS6110::NTF	WT	SUSC	1 IS6110::NTF	WT	ND
31	17316	NO	1 IS6110::NTF	WT	SUSC	1 IS6110::NTF	WT	S ^R
32	15550	NO	1 IS6110::NTF	450TTG/-15T	MDR	1 IS6110::NTF	450TTG/-15T	ND
33	16870	NO	1 IS6110::NTF	450TTG/315C	MDR	1 IS6110::NTF	450TTG/315C	ND
34	3219	NO	1 IS6110::NTF	315C	I ^R	1 IS6110::NTF	315C	SI ^R
35	10598	NO	1 IS6110::NTF	450TTG/315C	MDR	1 IS6110::NTF	450TTG/315C	SI ^R
36	11877	NO	1 IS6110::NTF	435?/315C	ND/ I ^R	1 IS6110::NTF	315C	ND
37	12766	NO	IS6110	WT	SUSC	IS6110	WT	SUSC
38	12360	NO	IS6110	WT	SUSC	IS6110	WT	SUSC
39	20082	NO	IS6110	WT	SUSC	IS6110	WT	ND
40	22661	NO	IS6110	WT	SUSC	IS6110	WT	ND
41	33452	NO	IS6110	WT	SUSC	IS6110	WT	ND
42	18842	NO	IS6110	435?/WT	ND	IS6110	WT	ND
43	12551	NO	IS6110	WT	SUSC	IS6110	WT	ND
44	10545	NO	IS6110	450TTG/315C	MDR	IS6110	450TTG/315C	MDR
45	11265	NO	IS6110	450TTG/315C	MDR	IS6110	450TTG/315C	MDR
46	1595	NO	IS6110	450TGG/315C	MDR	IS6110	315C	ND
47	4948	NO	IS6110	315C	I ^R	IS6110	315C	ND

Outb=outbreak; GDST=Genotypic Drug Susceptibility Testing (DST); PhDST=Phenotypic DST; ND=Not Determined. I=Isoniazid; R=Rifampicin; S=Streptomycin; E=Ethionamide; SUSC=Susceptible; ^R/R.=Resistance; loc=locus.

Table 2- Prediction of modern MDR Beijing on a Kazakhstan sample by IS-NTF RINT versus TB-SPRINT/TB-SNPID

	Reference (TB-SNPID, TB-SPRINT, Plikaytis)		
	MDR-modern Beijing	non-MDR modern Beijing	Total
IS-NTF RINT			
MDR-modern Beijing	146	9	155
non-MDR modern Beijing	6	20	26
Total	152	29	181

References

- Bergval, I., S. Sengstake, et al. (2012). "Combined Species Identification, Genotyping, and Drug Resistance Detection of *Mycobacterium tuberculosis* Cultures by MLPA on a Bead-Based Array." PLoS One **7**(8): e43240.
- Casali, N., V. Nikolayevskyy, et al. (2014). "Evolution and transmission of drug-resistant tuberculosis in a Russian population." Nat Genet **46**(3): 279-286.
- Caws, M., G. Thwaites, et al. (2008). "The influence of host and bacterial genotype on the development of disseminated disease with *Mycobacterium tuberculosis*." PLoS Pathog **4**(3): e1000034.
- D. Thierry, M.D.Cave, et al. (1990). "IS6110, an IS-like element of *Mycobacterium tuberculosis* complex.pdf>." Nucleic Acids Research, **Vol. 18, No. 1**.
- de Jong, B. C., P. C. Hill, et al. (2008). "Progression to active tuberculosis, but not transmission, varies by *Mycobacterium tuberculosis* lineage in The Gambia." J Infect Dis **198**(7): 1037-1043.
- Demay, C., B. Liens, et al. (2012). "SITVITWEB – A publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology." Infect Genet Evol **12**(4): 755-766.
- Deshpande, A., J. Gans, et al. (2009). "A rapid multiplex assay for nucleic acid-based diagnostics." J Microbiol Methods **80**(2): 155-163.
- Deshpande, P. S., R. S. Kashyap, et al. (2007). "Evaluation of the IS6110 PCR assay for the rapid diagnosis of tuberculous meningitis." Cerebrospinal Fluid Research **4**(1): 10.
- Dormans, J., M. Burger, et al. (2004). "Correlation of virulence, lung pathology, bacterial load and delayed type hypersensitivity responses after infection with different *Mycobacterium tuberculosis* genotypes in a BALB/c mouse model." Clin Exp Immunol **137**: 460-468.
- Filliol, I., S. Ferdinand, et al. (2002). "Spoligotyping and IS6110-RFLP typing of *Mycobacterium tuberculosis* from French Guiana : a comparison of results with international databases underlines interregional transmission from neighboring countries." Res. Microbiol. **153**: 81-88.
- Gomgnimbou, M. K., I. Hernandez-Neuta, et al. (2013). "'TB-SPRINT: TUBERCULOSIS-SPOLIGO-RIFAMPIN-ISONIAZID TYPING"; an All-in-One assay technique for surveillance and control of multi-drug resistant tuberculosis on Luminex® devices." J Clin Microbiol **51**(11): 3527-3534.
- Ida Parwati, Reinout van Crevel, et al. (2010). "<Possible underlying mechanism for successful emergence of the MTB Beijing genotype strains.pdf>." Lancet Infect Dis **10**:103-111.

Ioannidis, P., D. van Soolingen, et al. (2017). "Multidrug-resistant/extensively drug-resistant tuberculosis in Greece: predominance of *Mycobacterium tuberculosis* genotypes endemic in the Former Soviet Union countries." Clin Microbiol Infect.

Kamerbeek, J., L. Schouls, et al. (1997). "Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology." J Clin Microbiol **35**(4): 907-914.

Kremer, K., M. J. van-der-Werf, et al. (2009). "Vaccine-induced immunity circumvented by typical *Mycobacterium tuberculosis* Beijing strains." Emerg Infect Dis **15**(2): 335-339.

Lia Lima Gomes, Michel Abanto Marin, et al. (2015). "<Genome Comparison of an Ancestral Isolate and a Modern Isolate of *Mycobacterium tuberculosis* of the Beijing Lineage from São Paulo, Brazil.pdf>." Genome Announc **3**(5).

Luo, Tao Luo, et al. (2015). "Southern East Asian origin and coexpansion of *Mycobacterium tuberculosis* beijing family with han Chinese.pdf>." PNAS.

Manca, C., M. B. Reed, et al. (2004). "Differential monocyte activation underlies strain-specific *Mycobacterium tuberculosis* pathogenesis." Infect Immun **72**(9): 5511-5514.

Merker, M. (2017). "Emergence and spread of multidrug resistant tuberculosis in Eurasia ("Combat an old enemy with new weapons")." Oral presentation at ESM congress 2017 **38**: 13.

Merker, M., C. Blin, et al. (2015). "Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage." Nat Genet **47**(3): 242-249.

Mokrousov, I., O. Narvskaya, et al. (2008). "*Mycobacterium tuberculosis* Beijing Genotype in Russia: in Search of Informative VNTR Loci." J Clin Microbiol.

Mokrousov, I., O. Narvskaya, et al. (2012). "Russian "Successful" Clone B0/W148 of *Mycobacterium tuberculosis* Beijing Genotype: a Multiplex PCR Assay for Rapid Detection and Global Screening." Journal of Clinical Microbiology **50**(11): 3757-3759.

Perez-Lago, L., S. Izco, et al. (2017). "A novel strategy based on genomics and specific PCR reveals how a multidrug-resistant *Mycobacterium tuberculosis* strain became prevalent in Equatorial Guinea 15 years after its emergence." Clin Microbiol Infect **23**(2): 92-97.

Pheiffer, C. (2005). "Protein expression by a Beijing strain differs from that of another clinical isolate and *Mycobacterium tuberculosis* H37Rv." Proc. Nati. Acad. Sci. USA **151**(4): 1139-1150.

Plikaytis, B. B., J. L. Marden, et al. (1994). "Multiplex PCR assay for the Multi-drug-resistant strain W of *Mycobacterium tuberculosis*." J. Clin. Microbiol. **32**: 1542-1546.

Ramazanzadeh, R. and K. Sayhemiri (2014). "Prevalence of Beijing family in *Mycobacterium tuberculosis* in world population: Systematic Review and Meta-Analysis." International Journal of Mycobacteriology **3**(1): 41.

Ribeiro, S. C. M., L. L. Gomes, et al. (2014). "*Mycobacterium tuberculosis* Strains of the Modern Sublineage of the Beijing Family Are More Likely To Display Increased Virulence than Strains of the Ancient Sublineage." Journal of Clinical Microbiology **52**(7): 2615-2624.

Shitikov, E., S. Kolchenko, et al. (2017). "Evolutionary pathway analysis and unified classification of East Asian lineage of *Mycobacterium tuberculosis*." Scientific Reports **7**(1).

Tsenova, L., E. Ellison, et al. (2005). "Virulence of selected *Mycobacterium tuberculosis* clinical isolates in the rabbit model of meningitis is dependent on phenolic glycolipid produced by the bacilli." J Infect Dis **192**(1): 98-106.

van Crevel, R., R. H. Nelwan, et al. (2001). "*Mycobacterium tuberculosis* Beijing genotype strains associated with febrile response to treatment." Emerg Infect Dis **7**(5): 880-883.

van Soolingen, D., L. Qian, et al. (1995). "Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of east Asia." J Clin Microbiol **33**(12): 3234-3238.

Vyazovaya, A., I. Mokrousov, et al. (2015). "Tuberculous Spondylitis in Russia and Prominent Role of Multidrug-Resistant Clone *Mycobacterium tuberculosis* Beijing B0/W148." Antimicrobial Agents and Chemotherapy **59**(4): 2349-2357.

Wada, T., T. Iwamoto, et al. (2009). "Genetic diversity of the *Mycobacterium tuberculosis* Beijing family in East Asia revealed through refined population structure analysis." FEMS Microbiology Letters **291**(1): 35-43.

Wan, K., J. Liu, et al. (2011). "Investigation on *Mycobacterium tuberculosis* Diversity in China and the Origin of the Beijing Clade." PLoS One **6**(12): e29190.

Etude 3 : Analyse de données WGS (whole-genome sequencing) et identification de marqueurs potentiels en vue du développement d'une méthode multiplexée et combinée basée sur le polymorphisme SNPs/RD/IS6110 et utilisant la technologie Luminex®, pour la sous-classification et le suivi mondial des souches de bacilles tuberculeux L2/Beijing.

Résumé

L'expansion des souches ultrarésistantes, multirésistantes ou sensibles de la lignée L2/Beijing associée avec différents épisodes pandémiques ou épidémiques dans diverses régions du monde constitue désormais un problème global et une menace sur la santé internationale. La taxonomie et la compréhension de la diversité et de l'évolution de cette famille de souches (Lignée 2) est complexe. Plusieurs études, sur la base de divers marqueurs moléculaires proposant diverses classifications et des arbres phylogénétiques difficiles à interpréter, ont révélé une structure génomique monophylétique de la lignée L2/Beijing. Les marqueurs moléculaires les plus étudiés, soit individuellement soit en combinaison sont les variations nucléotidiques (SNP), les régions de délétions (RD) et les séquences d'insertion *IS6110*. L'objectif de cette étude est d'une part de comprendre la structure génomique de la lignée L2/Beijing en étudiant la correspondance entre les différents marqueurs polymorphes à partir des études précédentes afin de proposer un arbre phylogénétique unique et d'autre part d'identifier des SNP spécifiques et informatives des principaux sous-groupes épidémiques de Beijing, de les corrélérer avec les régions de délétions trouvées, ainsi qu'avec l'étude des sites d'insertion *IS6110* dans le NTF en vue de développer ultimement une méthode de diagnostic discriminante et à haut débit de la lignée L2.

Au cours de ce projet, en se basant sur l'analyse des données génomiques WGS de 1398 souches Beijing issues de 32 pays, ainsi qu'en s'appuyant sur la classification synthétique proposée récemment par Shitikov et al. (2017) nous avons pu améliorer notre compréhension des diverses classifications proposées jusqu'à aujourd'hui et développer une approche intégrative de cette classification. A partir des données génomiques de ces souches, 48275 SNP ont été identifiées avec en moyenne 1601 SNP/souches (Shitikov et al 2017). Pour la mise au point d'une technique utilisant le principe de *Multiplex Ligation depended Probe Amplification Beijing* méthode dite « MLPA_Beijing », nous avons identifié 17 SNP sur 11 gènes, qui permettent de définir et d'identifier 17 sous-groupes de Beijing. Ces SNP ont été choisies sur la base du fait qu'elles sont présentes chez 95 à -99% des souches représentant le sous-groupe. Pour réaliser un suivi encore plus complet des souches Beijing, nous avons les principales régions de délétions identifiées dans la lignée Beijing (RD105, exRD105, RD207, RD181, RD150, RD131, RD142), et nous avons également inclus la détection de la séquence d'insertion dans le locus NTF (voir projet précédent), ainsi que 4 SNPs supplémentaires déjà décrites servant de contrôle. Par la suite en vue d'avoir une collection représentative de la diversité mondiale des souches Beijing, nous avons collecté des ADNs de lignée 2 d'origines géographiques diverses : Nouvelle Guinée Papouasie, Sicile, Bulgarie, Turquie, Mozambique,

Pérou, Pakistan, Kazakhstan, Chine, Cambodge, Djibouti, Ghana, Afghanistan, Arménie, Sénégal, Thaïlande, Soudan, Azerbaïdjan, USA. Certains ADNs ont été séquencés sur les différentes gènes par ailleurs, en vue d'avoir une collection de référence qui soit positive pour les SNPs sélectionnées.

La méthode en cours de développement servira de technique de référence pour l'étude de la structure génomique et de la diversité mondiale des souches Beijing en vue du contrôle de la transmission interhumaine et interpopulations, et qui permettra à faible coût une amélioration de la surveillance épidémiologique des clones Beijing épidémiques en Eurasie, en Asie centrale et ailleurs.

L'article qui suit présente les développements informatifs qui ont abouti et leurs liens avec la classification de Beijing. Ces résultats devront être prolongés par la mise au point de la méthode « MLPA-Beijing ».

Article 3

Improving tuberculosis surveillance by using the new unified classification of East Asian lineage (L2/Beijing) of *Mycobacterium tuberculosis*

¹Bernice Klotoe, ²Egor Shitikov, ⁴Natalia Kurepina, ⁵Elena Zholdibayeva, S. Panaiotov^{1,6}

²Barry Kreiswirth, ¹Christophe Sola*, ¹Guislaine Refrégier*.

¹Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France.

²Federal Research and Clinical Center of Physical-Chemical Medicine, Moscow, Russian Federation.

⁴Public Health Research Institute TB Center, International Center for Public Health, Newark, New Jersey, USA

⁵National Center for Biotechnology, Astana, Kazakhstan

Abstract

Context: East Asian/L2/Beijing lineage of *Mycobacterium tuberculosis* has caused many tuberculous epidemics. The classification of this lineage has been recently unified by SNPs characterizing monophyletic groups among more than one thousand publicly available genome sequencing data. To contribute to Beijing epidemics surveillance, we aimed here at digging deeper in these data to propose, first a SNP combination, and second a kit targeting these SNPs and other markers.

Methods: Using the public data, we searched for SNPs specific of monophyletic subgroups and that could be amplified in pairs. We evaluated their concordance with SNPs proposed in the unified classification and characterized their diversity on a test database including 33 strains from diverse genotypic profiles and countries.

Results: We identified 17 SNPs distributed in 11 amplifiable fragments that gave fully concordance with the unified classification for 14/17 of the subgroups. These SNPs were identified in the test database, which confirmed their position in Beijing evolutionary history. Moreover, we could identify the presence of main sublineages in previously uncharacterized countries. We show for instance that Asia ancestral 3 subgroup is present in Italy and that Pacific RD150 subgroup is present in Western Africa. We also document that epidemic of clade B0/-N90 has already spread from central Russia to Afghanistan.

Conclusion: This work first documents the reliability of the newly proposed unified classification. Second, by going deeper into the precision of the classification with the inclusion of 3 SNPs characteristic of main epidemics, the 17-SNPs combination we propose already identified interesting patterns of Beijing diversity and threats worldwide. This information will be further enriched and confirmed by the addition of some additional markers to propose an easy-to-use kit for Beijing tuberculosis surveillance.

1. Introduction

Effective virulence of many parasites in humans is difficult to precisely assess due to the large impact of environmental factors. Tuberculosis has long been known as a disease of the poor people and/or associated with industrial development. East Asian/L2/Beijing lineage of *Mycobacterium tuberculosis* is the most prevalent lineage causing human tuberculosis. It has been and still is responsible of numerous small outbreaks and causes approximately 50% of tuberculosis deaths in Asia. In Eastern Europe, America and South Africa, most L2/Beijing outbreaks are associated with multidrug resistance ((R. Johnson, R. Warren et al. 2006), (Sonal S. Munsiff, Beth Nivin et al. 2003), (Mokrousov 2006)). Several studies have linked it to higher virulence and mutation rate, but others contradicted these views (Parwati et al 2010, Ates et al 2018). One possible reason for the lack of consensus on the exact virulence of this lineage could be that this virulence depends on the genetic of the host or that only few Beijing sublineages may be hypervirulent. To test for the second hypothesis, precise knowledge on the Beijing subclassification is required.

The *M. tuberculosis* Beijing genotypes have been identified and subclassified for more than 20 years by spoligotyping, mini-satellites (MIRU and/or VNTR) or RFLP patterns, or by the analysis of described regions of difference (RD105, RD207, RD181...) ((Tsolaki, Gagneux et al. 2005), (Gagneux, DeRiemer et al. 2006)). The advent of Whole Genome Sequencing and of methods targeting Single Nucleotide Polymorphisms (SNP) gave rise to new, heterogeneous subclassifications. Despite bringing new knowledge (*e.g.* presence of mutations in Replication, Repair and Recombination, Rad, 2003), these techniques also introduced confusion as correlation with other classifications was rarely investigated.

Recently, Shitikov et al (2017) confronted seven of the existing SNP classifications for 1,439 WGS sequences corresponding to as many strains of lineage 2. Each classification was also confronted to the phylogeny derived from the 48,275 identified SNP. This work led to the proposal of a new, consensual classification of L2/Beijing lineage including eleven subgroups (Shitikov, Kolchenko et al. 2017) (**Fig. 1**). In accordance with previous studies, lineage L2 is first divided into two major phylogenetic clades: 1) “Proto-Beijing” (Luo, Tao Luo et al. 2015) that is a synonym for lineage 2.1 (Coll, McNerney et al. 2014) and corresponding to ST21 SCG3a (Filliol et al, 2003); this group has been characterized by an extended RD105 deletion (Tsolaki et al); and 2) “Beijing” *sensu stricto* that corresponds to lineage 2.2 and that harbours the RD207 deletion. Lineage 2.2 is splitted into ten sub-groups:

Asia Ancestral 1 to 3 (Ancient Beijing); Asian African 1 to 3, Asian African2/RD142, Pacific RD150, Europe/Russia B0/W148 outbreak and Central Asia. They are mainly the same as those proposed by Merker et al, with the addition of Proto-Beijing, Asian African 3 corresponding to Bmyc13 from Mestre et al (2009) and the subgroup Asian African 2/RD142 corresponding to Bmyc18 from Mestre et al (2009), ST8 from Filliol et al (2006), Group 5 from Tsolaki et al and Gagneux et al (2005, 2006) and L2.2.1.2 from Coll et al (2014) (**Table 1**). Altogether, the new classification reinvests all main subgroups from all previous classifications, some of which are complementary. It logically reaches a higher discriminatory power (HGDI 0.79) compared to previous classifications taken separately (Shitikov et al, 2015).

In this study, we propose to further clarify and expand the new L2/Beijing unified classification proposed by Shitikov et al (2017) and implement it in the form of a cheap diagnostic tool targeting “Unified-Outbreak” SNPs dedicated to Beijing subclassification. To do so, we propose a selection of SNPs present as pairs in the same genes which eases multiplex amplification and reduce reagent costs. The questions we handled in this study are: 1) Is the new diagnostic tool providing a classification faithful to the unified classification from Shitikov et al? 2) Is the diagnostic tool able to classify new clinical isolates as compared to those handled by Shitikov et al? 3) Is the exploration of new isolates providing a new picture of L2/ Beijing diversity and evolution characteristics?

2. Material and Methods

2.1 Public reference database and test set

We received from Egor Shitikov (Federal Research and Clinical Center of Physical-Chemical Medicine, Moscow, Russian Federation) the SNP extraction of 1,439 whole genome sequencing data of as many strains from lineage 2 of *M. tuberculosis* collected from 32 countries. The raw data are available on National Center for Biotechnology Information (NCBI) and European Nucleotide Archive (ENA). The corresponding DNA samples could not be obtained, so that these data were used only for in silico analyses. To ease manipulation, we built a reduced version of Shitikov SNP matrix including 79 samples from 27 countries and uncovering the 11 subgroups defined by the new classification.

To further assess Shitikov et al. classification and set up the new technique, we selected 33 DNA from 19 different countries present in our laboratory collection. They were selected based on the known diversity in the country of origin and/or on the diversity of their CRISPR and/or MIRU-VNTR profiles. This selection served as a test set.

2.2 SNP selection of Unified-Outbreak-Beijing set

Our aim was to select SNPs for each meaningful monophyletic group inside L2/ Beijing phylogeny. We used the reduced version of Shitikov matrix to perform this selection.

A first criterion for SNP selection was the presence of the SNP in all strains belonging to the corresponding monophyletic group. The second criterion was the presence of another phylogenetic SNP in close vicinity to reduce the complexity of amplification. Vicinity is determined so that the size of the amplified sequence does not exceed 200 bp because of the technical constraints of the method. The 17 selected SNPs are shown in Table 2. Twelve of them are pairs found on the same gene at a maximal distance of 192 bp for *Rv0177* and a minimal distance of 30 bp for *BioF2*. The five others were SNPs identified by Shitikov et al. were included to cover all subgroups from the unified classification.

2.3 Sequencing

To characterize our test set of 33 samples, we designed primers for amplifying all genes corresponding to selected SNP. The DNA sequencing was done by Genewiz (Takeley, United Kingdom) directly on the amplified fragments.

3. Result

3.1 SNPs selected for Unified-Outbreak-Beijing classification tool

Based on the newly proposed classification by Shikitov et al (2017) and their WGS-derived SNP database, we selected 17 SNP distributed in 10 genes which defined 17 subgroups. These SNPs allow to derive all main Beijing subgroups of the unified classification. Three SNPs were added to identify epidemiological strains responsible for recent tuberculosis outbreaks: Europe-Russia B0/N90, Central Asia Outbreak, and Central Asia-Clade A (**Table 2**). The selected SNPs comprised non-synonymous (8/17) and synonymous (9/17). Our SNP selection included the mutation in codon 12 of *ogt* gene “*methylated DNA-protein-cysteine methyltransferase*” [genome position 1477596 C>T] to allow identification of modern subgroup (subgroups 5 to 11).

3.2 Sensitivity and specificity of Unified-Outbreak-Beijing SNP set according to the public reference database

We compared the 17 selected SNPs to the set of markers proposed by Shitikov et al. for the 1,439 samples included in public reference database. The sensitivity and specificity were 100% for all subgroups: Asia Ancestral 1 (51/51), Asia Ancestral 2 (20/20), Asia Ancestral 3 (63/63), Asian Africa 1 (18/18), Europe/Russia B0/W148 (294/294), Central Asia/CAO (81/81), Asia Africa 2/RD142 (3/3), Pacific RD150 (61/61). The designated SNP for proto-Beijing, Asian Africa 3, Central Asia/Clade A gave respectively 95%, 97,5% 99.5% sensitivity and specificity. As seen on a representative set of 79 strains from the reference collection, all 11 subgroups from Shitikov et al. could be identified by our scheme and rare discordances between classifications concerned only Proto-Beijing, Asia African 3 and some Central-Asia isolates (**Fig. 2**).

3.4 Ability of Unified-Outbreak-Beijing SNP set to classify new clinical isolates

We used the selected SNPs to characterize 33 new samples from a test worldwide database and identify their phylogenetic links (**Table 3, Fig. 3**). One strain was found potentially not belonging to L2 despite its spoligotype pattern. Five strains belonged to ancient Beijing and 27 strains to modern Beijing (**Fig. 3**). Ancient Beijing strains belonged either to Asia Ancestral 1 (n=1) or to Asia Ancestral 3 subgroup (n=3), the fifth isolate potentially belonging to Asia ancestral 2. Among modern Beijing strains, detected subgroups were Asia African 3/Bmyc13 (n=1), B0/Europe-Russia subgroup (n=2, one of which belonging to B0/N90 outbreak), Central Asia Outbreak (n=2), and Pacific RD150 (n=5).

However, 16 strains were only roughly classified as “modern”. The main reason for this rough classification is data missing for the three SNPs characterizing respectively Asia Ancestral 2, Asia African 1 and Central Asia subgroups. We detected only two cases of irrelevant SNP combination: one strain that had the SNPs specific to Asia African 2 and to Pacific RD150, and one strain that carried L2.2.1a mutation but not L2 mutation (**Table 3**). This could indicate either rare convergence events where one of the two mutation was acquired independently from the monophyletic subgroup, or errors during data or sample transfer.

4. Discussion and Perspectives

Using the public genome-based SNP collected by Shitikov et al (2017), we selected specific SNP to discriminate important subgroups of L2/Beijing lineage. The proposed SNP set called “Unified-Outbreak-Beijing” includes 17 SNP distributed in 10 genes and includes both synonymous and non-synonymous SNPs. The corresponding classification was fully concordant to that proposed by Shitikov et al (2017) (>96% sensitivity and specificity for all subgroups) and was able to classify 31/33 samples (94%) from a test set representative of worldwide diversity. We here discuss how this SNP selection may shed light on Beijing evolution and help surveillance of major outbreaks.

4.1 Potential links between Unified-Outbreak-Beijing SNP set and virulence

The inferred biochemical activities of targeted genes include hydrolase, methyltransferase, transferase, nuclease and catalytic activity. Corresponding biological processes are DNA modification (*Rv3204 DNA-methyltransferase*), repair (*ogt*, bifunctional regulatory protein/DNA repair enzyme *AlkA*), transcription (ribonuclease *VapC12*), metal-binding (*Rv3437* gene “*transmembrane protein*”), metabolism (carboxylesterase *LipT*, oxidative stress response regulatory protein *OxyS*, BioF2 gene “*8-amino-7-oxononanoate synthase*) and virulence (*mpa* “*proteasome-associated ATPase*” and *mce3F*). The non-synonymous mutation in *mce3F* gene characterizes B0/ N-90 outbreak and could be associated with its virulence. However, the mutation in *mpa* is synonymous rendering its link with virulence unlikely.

Proto-Beijing is defined by mutation 208782 A>C in *Rv0177*, codon 122 (**Table 2**) while CAER subgroup including Europe/Russia B0/W148 and Central Asia is characterized by 208590 A>C in codon 58 of the same gene. This gene has been described as “*Mce associated protein*” and could also be linked to virulence.

4.2 Unified-Outbreak-Beijing SNP set classification efficiency

We obtained very good sensitivities and specificities for all targeted groups according to the public reference database and we could classify 31/33 strains of our collection taken as a test set with 14 of the 17 SNPs from Unified-Outbreak-Beijing SNP set. Sixteen strains (n=16, 48%) from our test set could only be roughly classified (as “modern Beijing”) and not in the subgroups 5 to 11 from the unified classification. This may largely be due to the 3 missing SNP information but could also be due to lower ability of the chosen SNPs to identify monophyletic subgroups. Further data acquisition on the missing SNPs and on a larger sample

set will be needed to better evaluate the ability of Unified-Outbreak-Beijing SNP set to classify Beijing isolates. Still, the inclusion of outbreak SNPs in Unified-Outbreak-Beijing SNP classification tool allowed us to detect three outbreak strains totaling 9% of our test database. The newly proposed set could thus not only help classify Beijing strains according to the unified classification by Shitikov et al, but also follow up outbreaks.

In our test set, we identified no strain with a synonymous mutation in *mpa* gene (codon 52, pos. 2376135 A>G) characterizing Asian African 2/RD142 subgroup (ST8, Bmyc18, Asia Africa 2) ((Mestre, Luo et al. 2011), (Merker, Blin et al. 2015), (Filliol, Motiwala et al. 2006), (Shitikov, Kolchenko et al. 2017)). This may indicate either that our collection does not represent worldwide diversity or that this subgroup is relatively rare.

The non-synonymous SNP at pos. 3857226 G>T (gene Rv3437, codon 106) that defines defined the subgroup Central Asia Outbreak was found in 2 strains from Kazakhstan known to belong to the highly prevalent VNTR signature 94-32. This SNP could help to track outbreaks at a finer level as compared to *sigE* SNP recently proposed by Mokrousov et al (2018).

4.3 New insights on Beijing diversity and expanding clones

The identification of a selection of SNP to classify isolates is an important step to be able to handle a large number of samples from countries with restricted economical resources. By sequencing 10 genes, we could classify 33 Beijing strains from our worldwide database with reasonable workload and costs. This enabled us to subclassify Beijing strains from countries that were absent from the public Beijing database (Shitikov et al, 2017). **Fig. 4** presents the distribution of Beijing diversity in the public database (using a representative restricted sample, **Fig. 4A**) and in our database (**Fig. 4B**). We document for the first time the presence of Pacific RD150 lineage in Senegal. As previously suspected by Mokrousov (2005), Beijing isolates may have swarmed out of East Asia using slave trade to South Africa and to Europe going through Western Africa. Alternatively, recent worldwide migrations could account for the presence of these strains that are not found in Central Africa. Asia ancestral 3 subgroup was also found in Italy. This subgroup had been already described in Germany (Merker et al, 2015). This confirms that Russian immigration (mainly harbouring Europe-Russia W148 clones) is not the sole entry route for Beijing tuberculosis in Europe.

The B0/W148 outbreak/N-90 outbreak (also referred to as sublineage N-90) has been first described in Samara, Russia. This sublineage can be identified by the non-synonymous

mutation in the virulence gene *mce3F* included in Unified-Outbreak-Beijing SNP set. In our DNA collection, we detected this mutation in one sample from Afghanistan. This indicates that this outbreak is still active and swarming in Central Asia. The broader subgroup Europe/Russia W148 was identified in one strain from Azerbaidjan. This suggests that the whole W148 subgroup is still expanding, although still mainly in Central Asia.

Strain 10545 isolated in the USA and included in our study, initially known as ancestral Beijing (Klotoe et al 2018, to be submitted), and revealed to be part to the Asia Ancestral 1 subgroup. Interestingly, South Korean outbreak strain K also belongs to this subgroup. Likewise, an important outbreak in Gran Canaria belongs to Asia Ancestral 2 subgroup. These results show that Beijing recent transmission is clearly not only linked to modern subgroups (Han, Song et al. 2015).

To conclude, we have selected 17 specific SNPs distributed on only 10 genes to identify and discriminate Beijing strains into different subgroups including outbreaks subgroups. This selection could be useful for Beijing Strains surveillance in hot spot regions. We intend to develop a *Multiplex Ligation depended Probe Amplification Beijing* «MLPA_Beijing » technique to classify and provide surveillance of lineage 2/Beijing strains.

References

1. Bonnie B. Plilkaytis, Jennifer L. Marden, et al. (1994). "< Multiplex PCR Assay Specific for the Multidrug-Resistant Strain W of *Mycobacterium tuberculosis*.pdf>." JOURNAL OF CLINICAL MICROBIOLOGY: 1542-1546.
2. Coll, F., R. McNerney, et al. (2014). "A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains." Nature Communications 5(1).
3. Coll, F., M. Preston, et al. (2014). "PolyTB: A genomic variation map for *Mycobacterium tuberculosis*." Tuberculosis (Edinb) 94(3):346-54(3): 346-354.
4. Filliol, I., A. S. Motiwala, et al. (2006). "Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set." J Bacteriol 188(2): 759-772.
5. Filliol, I., A. S. Motiwala, et al. (2006). "Global Phylogeny of *Mycobacterium tuberculosis* Based on Single Nucleotide Polymorphism (SNP) Analysis: Insights into Tuberculosis Evolution, Phylogenetic Accuracy of Other DNA Fingerprinting Systems, and Recommendations for a Minimal Standard SNP Set." Journal of Bacteriology 188(2): 759-772.
6. Gagneux, S., K. DeRiemer, et al. (2006). "Variable host-pathogen compatibility in *Mycobacterium tuberculosis*." Proceedings of the National Academy of Sciences 103(8): 2869-2873.
7. Han, S. J., T. Song, et al. (2015). "Complete genome sequence of *Mycobacterium tuberculosis* K from a Korean high school outbreak, belonging to the Beijing family." Standards in Genomic Sciences 10(1).
8. Luo, Tao Luo, et al. (2015). "Southern East Asian origin and coexpansion of *Mycobacterium tuberculosis* beijing family with han Chinese.pdf>." PNAS.
9. Merker, M., C. Blin, et al. (2015). "Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage." Nat Genet 47(3): 242-249.
10. Merker, M., C. Blin, et al. (2015). "Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage." Nature Genetics 47(3): 242-249.
11. Mestre, O., T. Luo, et al. (2011). "Phylogeny of *Mycobacterium tuberculosis* Beijing strains constructed from polymorphisms in genes involved in DNA replication, recombination and repair." PLoS One 6(1): e16020.
12. Mina Ebrahimi Rad, Pablo Bifani, et al. "< Mutations in putative mutator genes of *Mycobacterium tuberculosis* strains of the w-beijing family.pdf>." Emerging Infectious Diseases • Vol. 9, No. 7.
13. Mina Ebrahimi Rad, Pablo Bifani, et al. (2003). "Mutations in putative mutator genes of *Mycobacterium tuberculosis* strains of the w-beijing family.pdf>."
14. Mokrousov, I. (2005). "Origin and primary dispersal of the *Mycobacterium tuberculosis* Beijing genotype: Clues from human phylogeography." Genome Research 15(10): 1357-1364.
15. Mokrousov, I. (2006). "*Mycobacterium tuberculosis* Beijing genotype and mycobacterial interspersed repetitive unit typing." J Clin Microbiol 44(4): 1614; author reply 1614-1615.
16. R. Johnson, R. Warren, et al. (2006). "<An outbreak of drug-resistant tuberculosis caused by a Beijing strain in the Western Cape, South Africa .pdf>." INT J TUBERC LUNG DIS 10(12):: 1412-1414.
17. Rad, M. E., P. Bifani, et al. (2003). "Mutations in putative mutator genes of *Mycobacterium tuberculosis* strains of the W-Beijing family." Emerg Infect Dis 9(7): 838-845.
18. Schürch, A. C., K. Kremer, et al. (2011). "SNP/RD Typing of *Mycobacterium tuberculosis* Beijing Strains Reveals Local and Worldwide Disseminated Clonal Complexes." PLoS One 6(12): e28365.
19. Schurch, A. C., K. Kremer, et al. (2011). "Mutations in the regulatory network underlie the recent clonal expansion of a dominant subclone of the *Mycobacterium tuberculosis* Beijing genotype." Infect Genet Evol 11(3): 587-597.

20. Shitikov, E., S. Kolchenko, et al. (2017). "Evolutionary pathway analysis and unified classification of East Asian lineage of *Mycobacterium tuberculosis*." Scientific Reports 7(1).
21. Sonal S. Munsiff, Beth Nivin, et al. (2003). "<Persistence of a highly resistant strain of Tuberculosis in New York city uring 1990-1999.pdf>." The Journal of Infectious Diseases ; 188: 356-363.
22. Tsolaki, A. G., S. Gagneux, et al. (2005). "Genomic Deletions Classify the Beijing/W Strains as a Distinct Genetic Lineage of *Mycobacterium tuberculosis*." Journal of Clinical Microbiology 43(7): 3185-3191.
23. Wan, K. (2007). Study on Genetic Diversity and evolution of *Mycobacterium tuberculosis* the agent of tuberculosis in China. Paris, University Paris-Sud.
24. Zhenhua Yang, Peter F. Barnes, et al. (1998). "<Diversity od DNA Fingerprints of *Mycobacterium tuberculosis* Isolates in the United States.pdf>." Journal of Clinical Microbiology: 1003-1007.

Table 1 : Correspondance with all differentes classifications in comparison with new proposed by Shitikov et al. (2017)

N°	New classification (Shitikov et al 2017)	Merker et al 2015	Luo et al 2015	Tsolaki et al 2005/Gagneux et al 2006)	Coll et al 2014	Rad et al 2003	Filliol et al 2006	Mestre et al 2011	
1	Proto Beijing	-	Proto Beijing	Groupe 1	L2.1	WT mutT2 mutT4/ogt37	ST21	-	
2	Asia Ancestral 1	Asia ancestral 1	Bj-MG1	Groupe 2	L2.2.2		ST11	Bmyc2/3	
3	Asia ancestral 2	Asia ancestral 2	Beijing	Groupe 3	L2.2.1		STK/ST3	Bmyc4	
4	Asia ancestral 3	Asia ancestral 3	Bj-MG2			mutT4/ogt37	ST25	Bmyc25	
5	Asian African 1	Asian African 1	Bj-MG3			Groupe 3	mutT2 (58)/ mutT2 (48)/ ogt12	ST10	Bmyc 10
6	Asian African 3	-							Bmyc13
7	Europe/Russia B0/W148 outbreak	Europe/Russia B0/W148 outbreak							Bmyc12
8	Central Asia	Central Asia							Bmyc10
9	Asian African 2	Asian African 2							ST22
10	Asian African 2/RD142	Asian African 2	Groupe 5			L2.2.1.2	ST8	Bmyc18	
11	Pacific RD150	Pacific RD150	Groupe 4			L2.2.1.1	ST10	Bmyc10	

Table 2 : Selected SNP for Unified-Outbreak-Beijing diagnostic kit and corresponding classification.

Selected SNP according to new proposed classification by Shitikov et al. (2017).

Gene name	Gene #	Codon #	ref	mut	ref codon	mut codon	Type mut.	Corresponding monophyletic group (number in Shitikov et al classif.)
BioF2	Rv0032	572	G	C	Asp	His	NS	L2
	Rv0177	122	A	C	Thr	Thr	S	Proto-Beijing (1)
	Rv3204	54	C	T	Pro	Pro	S	Asia ancestral 1 (2)
LipT	Rv2045c	163	T/A	C/G	Thr	Thr	S	Asia ancestral 2 (3)
OxyS	Rv0117	142	G	A	Asp	Asn	NS	lin2.2.1a (4-11)
ogt	Rv1316c	37	C/G	A/T	Arg	Leu	NS	Asia ancestral 3 (4)
ogt	Rv1316c	12	C/G	T/A	Gly	Gly	S	Modern (5-11)
OxyS	Rv0117	86	G	C	Arg	Arg	S	Asia Afr1 (5)
LipT	Rv2045c	218	G/C	T/A	Ala	Ala	S	Asia Afr3 (6)
	Rv0177	58	A	C	Arg	Arg	S	CentralAsia (8)
VapC12	Rv1720c	32	T/A	C/G	His	Arg	NS	CAER (7-8)
mce3F	Rv1971	410	A	C	Asp	Ala	NS	Eur.Russ.B0/N90 *
BioF2	Rv0032	562	G	A	Ala	Thr	NS	Central Asia-clade A *
	Rv3437	106	G	T	Asp	Tyr	NS	Central Asia-CAO *
mpa	Rv2115c	52	A/T	G/C	Ser	Ser	S	Asia Afr2 (9)
alkA	Rv1317c	11	C/G	T/A	Ala	Thr	NS	Asia Afr2-RD142 (10)
VapC12	Rv1720c	46	A/T	G/C	Arg	Arg	S	Pacific-RD150 (11)

*subgroups corresponding to recent outbreaks

Table 3: Sequences of the 33 test-database clinical isolates

Mutation	Position	Place of origin of clinical isolates and ID																																			
Rv0032-BioF2_572	36008	Netherlands	Italy, Palermo	Italy, Palermo	Bulgaria	Bulgaria	United States, NYC	United States, NYC	Peru	Peru	Djibouti	Senegal	South Africa	South Africa	Mozambique	Mozambique	Pakistan	Turkey	Turkey	Pakistan	Kazakhstan	Kazakhstan	Afghanistan	Azerbaijan	Armenia	Cambodia	Cambodia	Thailand	Thailand	China, Jiangsu	China, Anhui	China, Beijing	Papouasia NG	Papouasia NG			
Rv0177-_122	208782	8000578	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C		
Rv3204-_54	3581476	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C		
Rv2045c-LipT_163	2290732	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)	(A)		
Rv0117-OxyS_142	141623	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
Rv1316c-ogt_37	1477522	(C)	A	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
Rv1316c-ogt_12	1477596	(C)	C	C	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C	T	T	T	T	T		
Rv0117-OxyS_86	141457	(G)	(G)	(G)				(G)	(G)		(G)	(G)	(G)	(G)							(G)	(G)	(G)	(G)			(G)	(G)				(G)	(G)				
Rv2045c-LipT_218	2290567	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	C	
Rv0177-_58	208590	(A)	(A)	(A)				(A)	(A)		(A)	(A)	(A)	(A)							(C)	(C)	(C)	(C)			(A)	(A)	(A)				(A)	(A)			
Rv1720c-VapC12_32	1947325	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
Rv1971-mce3F_410	2216485	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
Rv0032-BioF2_562	35978	G	G	G	G	G	G	G	G	G	G	G	G	G	G	(G)	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
Rv3437-_106	3857226	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	T	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
Rv2115c-mpa_52	2376135	(T)	T	T	T	T	T	T	C		(T)	T	T	C		T	T	T	T	T	T	T	T	T	T			C	(T)	(T)		T	T	T	T	T	
Rv1317c-alkA_11	1479088	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
Rv1720c-VapC12_46	1947282	T	T	T	T	T	T	T	T	T	T	C	T	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C	C	C
		NonL2	AsAnc3	AsAnc3	mod	mod	mod	AsAnc 1	AsAfr2	mod	mod	Pacific	mod	Pacific	Pac/AsAI	?	mod	mod	mod	mod	mod	mod	mod	mod	mod	mod	AsAf3	ASAF2	AsAnc2	AsAnc3	mod	mod	Pacific	Pacific	Pacific		

Figure 1: Shitikov et al phylogeny presenting the new unified classification. Shitikov et al SNP scheme allows to identify the 11 terminal subgroups of the classification (black) and 6 more internal phylogenetic groups (green).

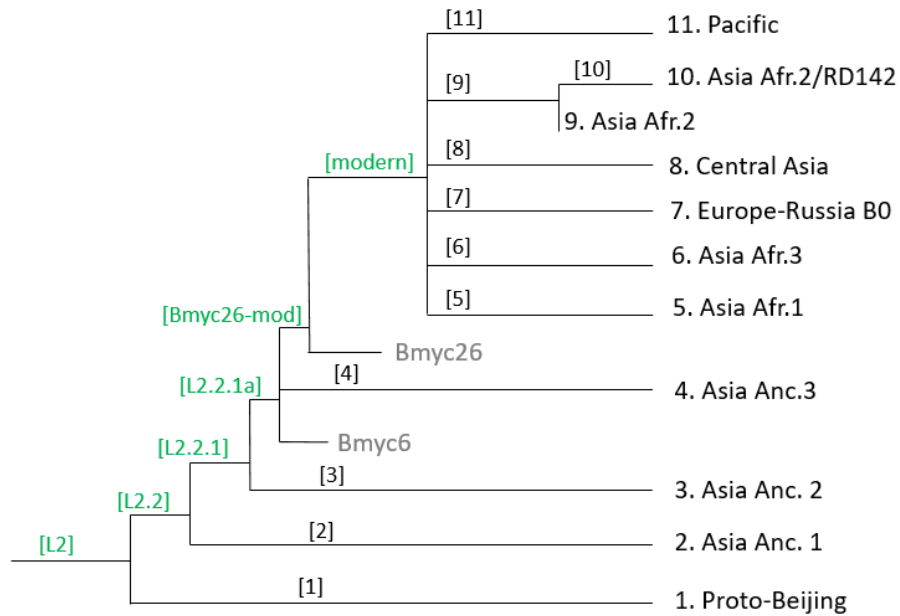


Figure 2: Maximum Likelihood phylogeny of a representative sample of the public reference database based on the 17 selected SNPs for Unified-Outbreak-Beijing classification. Subgroups are indicated under square brackets. Subgroups appearing in black or green are identical to those detected by Shitikov et al scheme (see Fig. 1). Subgroups providing additional information are indicated in red.

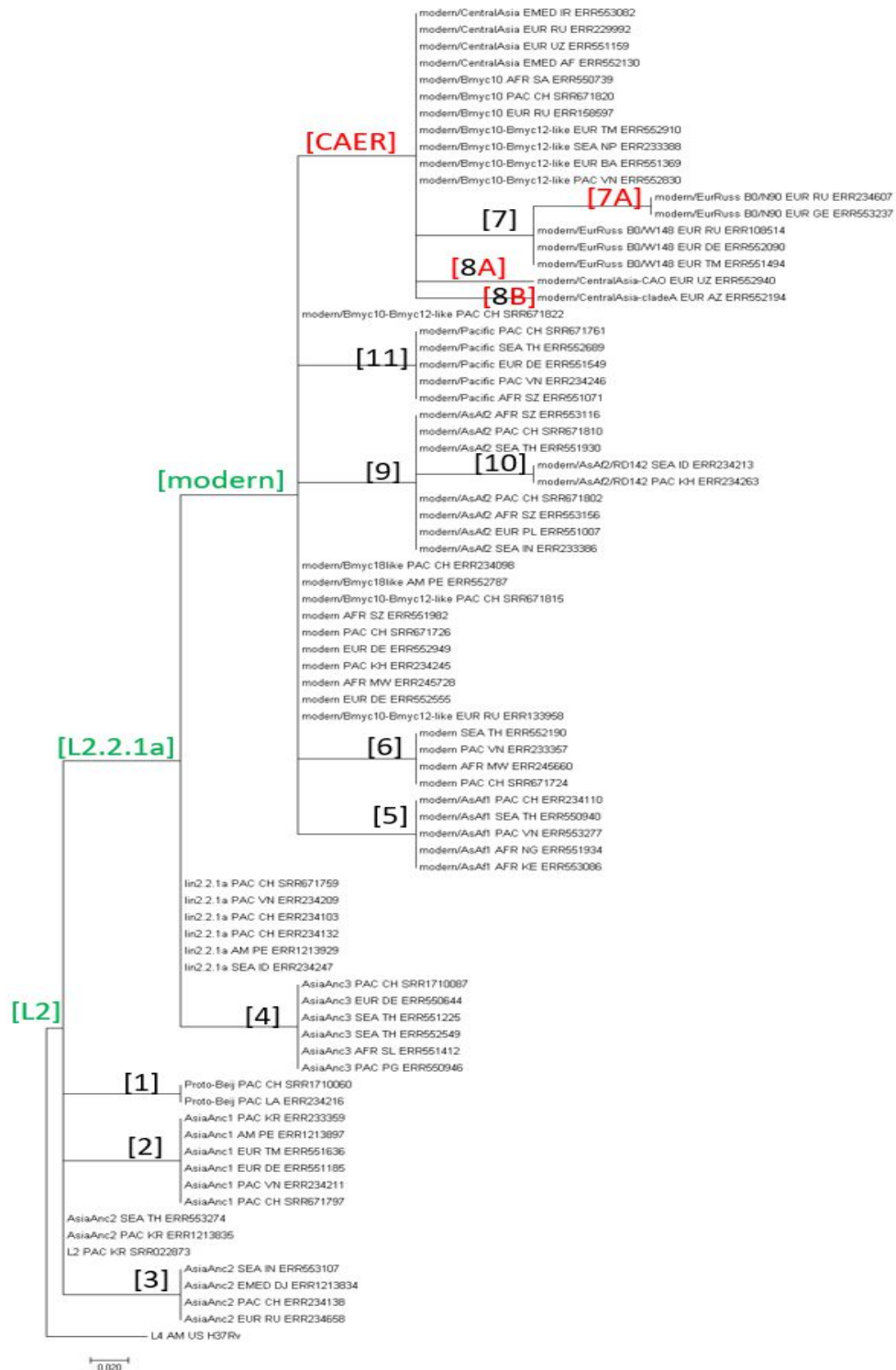


Figure 3: Maximum Likelihood phylogeny of 33 samples test-database based on 14 of the 17 selected SNPs for Unified-Outbreak-Beijing classification. Subgroups appearing in black or green are identical to those detected by Shitikov et al scheme (see Fig. 1). Subgroups providing additional information are indicated in red.

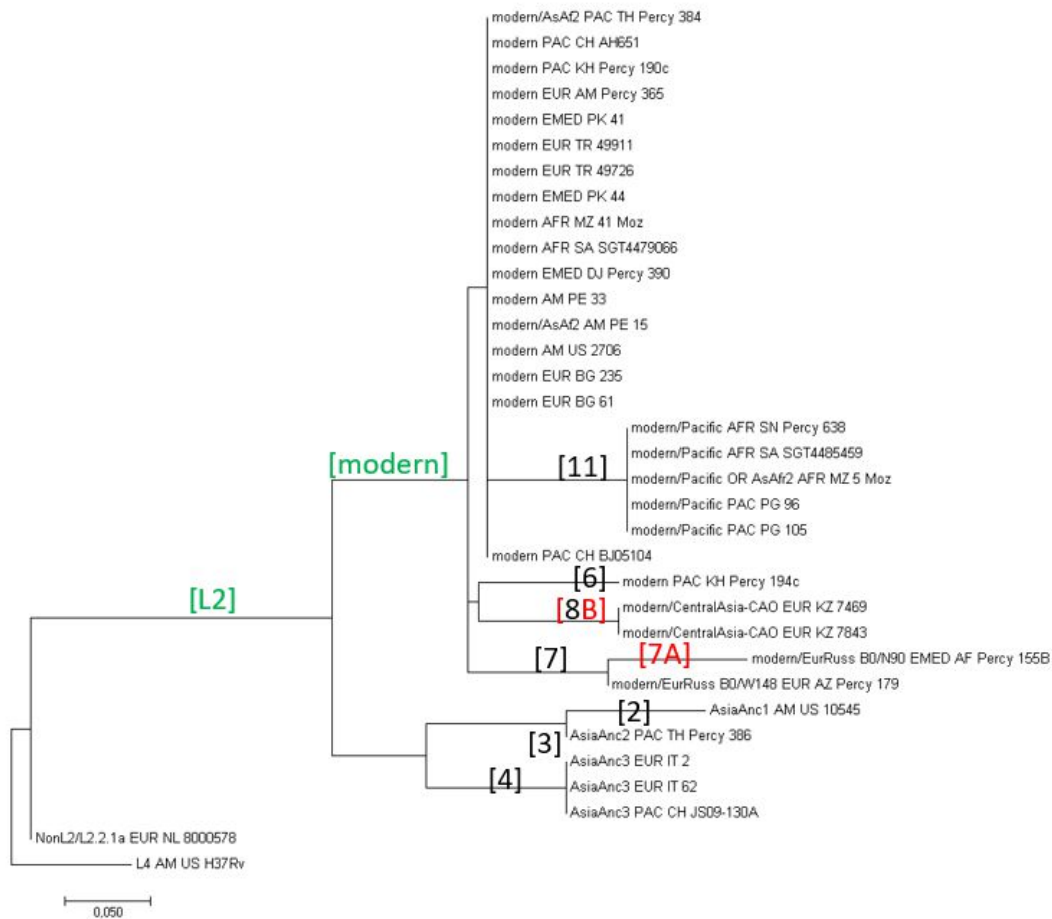
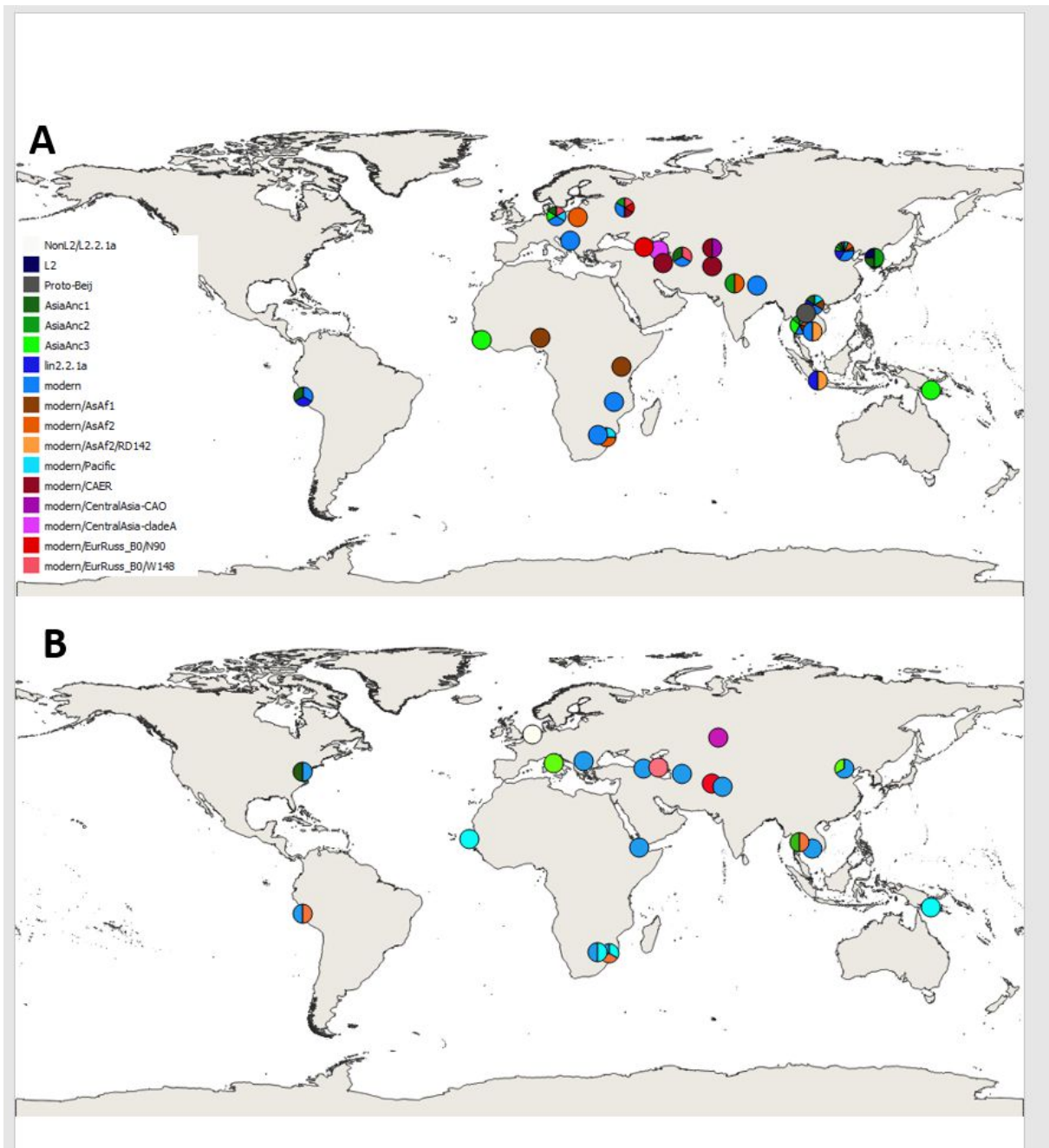


Figure 4 :



Etude 4 : Etude de la tuberculose Multi (MDR) et eXtrêmement (XDR) résistante dans un pays à forte prévalence de souches Beijing, le Kazakhstan

Résumé

Après l'effondrement de l'Union Soviétique et du système de santé, le Kazakhstan a connu une forte augmentation des cas de tuberculose. Ainsi, l'OMS a estimé la prévalence de la tuberculose au Kazakhstan à 127 cas par 100000 habitants en 2014 (WHO 2014), avec un important accroissement du nombre de cas de MDR-TB. L'épidémie de la tuberculose se répand dans tout le pays spécialement dans les régions où le système socio-économique et environnemental est défaillant notamment dans les prisons.

Cependant, il n'y a pas de données récentes sur la distribution et la diversité génomique des diverses souches de *M. tuberculosis* au Kazakhstan, en dehors de l'étude de Y. Skiba *et al.* parue en 2015 et qui portait sur des souches de 2008, et d'une étude allemande encore plus ancienne de T. Kubica *et al.* parue en 2005.

L'objectif de notre étude était donc triple : premièrement étudier la diversité génétique des souches de *Mycobacterium tuberculosis* dans le pays grâce à une sélection exhaustive de souches représentatives de toutes les régions (ou presque), en la comparant aux études de diversité génomique et spatiale antérieures, deuxièmement améliorer la caractérisation des SNPs de multi-résistance et de l'ultra-résistante par la combinaison de deux méthodes à haut débit (TB-SPRINT et TB-SNPID), troisièmement de tenter de caractériser la ou les principaux clone épidémiques qui relèvent d'une transmission active du bacille dans le pays.

L'étude a été réalisée sur 700 souches provenant de 700 patients tuberculeux habitant au Kazakhstan dans 12 des 14 régions et recrutées entre 2010-2015. Ces souches ont été analysées par le TB-SPRINT et le TB-SNPID, deux techniques qui permettent d'identifier à la fois la lignée et de prédire le profil sensibilité/résistance aux principaux antibiotiques sur chaque souche. Deux lignées principalement ont été identifiées: au cours de cette étude : d'une part la lignée L2/Beijing et d'autre part la lignée L4/Euro American. La prévalence de la lignée L2 est d'environ 80% avec 50% de MDR-TB et 20 cas de XDR-TB sur 450.

En parallèle, pour sous-classer les souches Beijing et tenter d'identifier la grappe épidémique prédominante, nous avons typé les sites d'insertions de la séquence IS6110 dans le locus NTF par la technique IS-NTF/RINT (décrite dans un projet précédent), et le polymorphisme de la SNP *SigE* par la PCR-RFLP. Toutes les souches de la lignée L2/Beijing sont des « modern » Beijing, portant la SNP *SigE* caractéristique du clone dit 94-32 selon la nomenclature par 24 MIRU-VNTR.

Cette étude a permis d'identifier le génotype et le groupe de cluster circulant et prédominant au Kazakhstan. L'article qui suit sera prochainement soumis à *BMC Microbiology*.

Article 4 :

The genomic characterization of MDR and XDR-TB in Kazakhstan, analyzed by a combination of high throughput methods, shows a predominant transmission of the 94-32 L2/Beijing cluster.

B. Klotoe¹, S. Kacimi¹, H. Gomes^{1,2}, R. Barcellos^{1,3}, S. Panaiotov^{1,4}, D. Haj Slimene^{1,5}, N. Sikhayeva⁶, S. Sengstake⁷, A. Schuitema⁷, M. Akhalaia⁸, A. Alenova⁹, E. Zholdybayeva⁶, R. Anthony⁷, G. Refrégier¹ and C. Sola^{1*}

¹Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud,

²Laboratory of Molecular Biology Applied to Mycobacteria, FIOCRUZ, Rio de Janeiro, Brazil.

³Fundação Estadual de Produção e Pesquisa em Saúde, Centro de Desenvolvimento Científico e Tecnológico, Porto Alegre, RS, Brasil

⁴National Center of Parasitic and Infectious Diseases, Sofia, Bulgaria.

⁵Institut Pasteur de Tunisie, Tunis, Tunisie.

⁶National Centre for Biotechnology, Astana, Kazakhstan.

⁷Royal Tropical Institute (KIT), Amsterdam, The Netherlands.

⁸Foundation for Innovative New Diagnostics (FIND), Geneva.

⁹National Centre for Tuberculosis Problems, Almaty, Kazakhstan.

*address for correspondance: Christophe Sola, Institute of Integrative Cell Biology (I2BC), UMR9198, CEA-CNRS-UPS, Orsay, France. christophe.sola@i2bc.paris-saclay.fr

Ph : +33 1 69 82 62 00

Summary

A convenience sample of 700 DNA extracted from as many *Mycobacterium tuberculosis* complex (MTC) cultures isolates from 630 tuberculosis patients living in Kazakhstan, recruited within 12 out of 14 regions between 2010-2015, was independently studied by two high-throughput hybridization-based methods, TB-SPRINT (59-Plex, n=700), TB-SNPID (50-Plex, n=543). Almost four hundreds (n=391) clinical isolates DNA were typed by the two methods. Two lineages only, L2/Beijing and L4-Euro American, were detected. L2/Beijing prevalence totaled almost 80% and MDR-TB was present in 50% of all samples by either TB-SPRINT or TB-SNPID. In-Depth molecular identification and Drug-resistance gene mutation characterization is described. Performances of the two assays were compared. Two complementary assays were run on L2/Beijing isolates: an IS6610-NTF insertion site typing assay and *SigE* polymorphism analysis, respectively demonstrating that all L2/Beijing were “modern” Beijing and belonging to the 94-32 Cluster. A dynamic landscape of *M. tuberculosis* genotypes circulating in Kazakhstan is drawn, from 2005 still today that shows the increased spreading of a limited number of specific L2/Beijing clonal complexes in Kazakhstan.

Introduction

Kazakhstan is a 17 million inhabitants's central Asian country that does not belong to the WHO 22 high-TB prevalence list of countries. However it is facing a serious multidrug resistant tuberculosis (MDR-TB) challenge (WHO 2017). Apart from the resistance to Rifampin (RIF) and Isoniazid (INH) that defines MDR, resistance to fluoroquinolones (FLQ) and other second-line injectable drugs (SLID) is becoming more and more common in Former Soviet Union republics and elsewhere (WHO 2017). Such a phenomenon creates a threat to the global effective TB infection control. According to a recent World Health Organization (WHO) report, extensively drug-resistant TB (XDR-TB) had by the end of 2016 been reported by 123 countries (WHO 2017), i.e. in 6 countries more compared to the situation one year before (WHO 2016).

Kazakhstan has established an efficient surveillance system to monitor drug resistance in the past three years, with an index of 34 laboratories, i.e. 6.3 laboratories per 1 million populations, which can perform phenotypic Drug Susceptibility Testing (DST). DST is performed by 22 laboratories and 12 laboratories are reported to be able to run Line Probe Assays (Hain Diagnostics, Germany). However with an increasing trend of 7315 total MDR-TB cases enrolled on MDR-TB treatment, it is clear that high-throughput predictive genotyping of drug susceptibility testing could achieve better results than phenotypic DST, if done in a limited number of certified reference laboratories, that would better support monitoring as well as target DST testing for the bacteriologically confirmed cases (9597 in 2014) (WHO 2015).

TB-SPRINT and TB-SNPID are recent molecular diagnostic methods that provides simultaneous MTC genotyping subspecies identification, as well as first-line and second line (for TB-SNPID) predictive drug resistance genotyping (Bergval, Sengstake et al. 2012; Gomgnimbou, Refrégier et al. 2013; Sengstake, Bablishvili et al. 2014; Dantas, Suffys et al. 2017; Molina-Moya, Gomgnimbou et al. 2017; Santos, Costa et al. 2017). These assays uses the versatility of various types of coupled microspheres that are further analyzed by flow cytometry, in microfluidics or fluorescence imaging-based high-throughput analytical devices (Dunbar 2006; Ocheretina, Merveille et al. 2013). TB-SNPID is a ligation-mediated based assay whereas TB-SPRINT relies on direct hybridization and on a dual-oligonucleotide priming (DPO) principle (Schouten, McElgunn et al. 2002; Chun, Kim et al. 2007). We report in this paper the genotyping results obtained between July-November 2015 on a set of 700 MTC DNAs representative of 630 patients (70 duplicates DNA) that were provided by the National Center for Biotechnology of the Republic of Kazakhstan in Astana, working in collaboration with the National Center for Tuberculosis Problems in Almaty.

Material and methods

Patients

A total of 630 patients living in Kazakhstan (435 men, 195 female) were included in the study. Since the human starting biological samples were obtained for classical TB diagnostics, no informed consent, whether written or oral, nor ethical approval committee was obtained prior to performing this study. 35,6% out of 630 (n=224) were isolated from new cases of tuberculosis and 64,4 % (n=406) were from previously treated patients. Among 630 patients with pulmonary tuberculosis 69 % (n=435) were men and 23% (n=145) were women, 50 (8%) were unknown. The age of TB patients varied from 16 to 80 years, with an average age of 37 years.

Clinical isolates, TB Identification and Phenotypic DST

Clinical isolates of MTB were isolated between 2010-2015 and represented ten regional TB dispensaries (Atyrau, Mangistau, Aktobe, Kostanay, Pavlodar, Semey, Kyzyl Orda, Zhambyl, Taldygorgan and Almaty) and two prisons run by the State-Established Committee for the Management of the Penitentiary System of the Karaganda and Akmola regions. Clinical isolates of *M. tuberculosis* were deposited in the Reference Laboratory of the National Center for Tuberculosis Problems for surveillance studies of anti-TB drug resistance in Kazakhstan. After primary isolation, mycobacteria were subcultured on solid Löwenstein-Jensen (LJ) medium in regional laboratories. Species identification as *M.tuberculosis* and drug susceptibility testing were performed at the Reference Laboratory by standard procedures (David, Levy-Frebault et al. 1989). Rifampicin and isoniazid susceptibility tests were carried out on LJ medium containing 40µg/L rifampicin, 0.2 mg/L isoniazid, or 1 mg/L isoniazid using the absolute concentration method according to the World Health Organization (WHO) recommendations.

DNA was extracted by thermolysis using the protocol described below, from solid or liquid cultures: mycobacterial colonies grown on solid media (Lowenstein-Jensen) medium or mycobacterial pellets obtained from solid or liquid cultures (MGIT, mycobacterial growth indicator tube, Becton-Dickinson, Piscataway, NJ) were resuspended into 200µl 10 mM Tris-HCl, 1 mM EDTA (pH 7.0), in a screwed cap tube or a safe-lock tube. For colonies recovered from solid media, subcultures were made from concentrated stocks solution kept at -20°C. Liquid cultures were centrifuged at 15,000 g x 5 min, the supernatant was discarded and the bacterial pellet resuspended into 200 µl 10 mM Tris-HCl, 1 mM EDTA (pH 7.0). Samples were heated to 95°C for 45 min in an oven or a PCR cycler with a hot lid, if available in the microbiological security facility. The suspension was centrifuged at 15,000 g for 1 min to

pellet the cell debris. The supernatant containing the DNA was harvested and transfer into a new tube. The stocks were stored at $-20\text{ }^{\circ}\text{C}$ until further use, or diluted at 1:50 into sterile water in a new tube. Diluted solutions were stored at $-20\text{ }^{\circ}\text{C}$. 700 DNAs were shipped to France and to the Netherlands for further studies in two shipments (June 2015, n=93, August 2015, n=605).

Multiplexed Genotyping methods

❖ *TB-SPRINT*

TB-SPRINT typing (59-Plex) is a Nucleic-Acid PCR-based Amplification Test (NAAT) that was developed previously (Gomgnimbou, Hernandez-Neuta et al. 2013). In order to minimize failure risks and since the DNA used was neither pure nor concentrated (not Cetyl-Trimethyl-Ammonium-Bromide-extracted), TB-SPRINT was run sequentially, i.e. 43-Plex spoligotyping (TB-SPOL) in a first PCR/hybridization/detection experiment, followed by a second multiplexed PCR (*rpoB*, *katG*, *inhA*) and a 16-Plex drug-resistance genes SNP typing PCR/hybridization/detection experiment (TB-RINT). A total of seven SBS (Society for Biomolecular Screening) 96 plates were run. Repetitions were performed only in case of need. Most experiments were performed by a trained technician however part of the work was also run using an automated pipetting station (Hamilton, Villebon sur Yvette, France) and no differences were observed (results not shown). All readings were performed on a Luminex 200® analyzer (Luminex Corp. Austin, USA) using xMAP® coupled reagents (Beamedex® SAS, Orsay, France). Second-lane drug resistance genes were also tested by TB-EFI, a 16-Plex method and results will be reported elsewhere. Full results are shown in Supplementary Table (S1 and S3 sheets : 677 spoligotyping results with lineage and sublineage).

❖ *TB-SNPID*

TB-SNPID was done as previously described (Bergval, Sengstake et al. 2012) using semi-automated data analysis based on a threshold value as published in (Sengstake, Bablshvili et al. 2014). All readings were performed on a MagPix® analyzer using MagPlex xTAG® beads and other reagents provided by MRC Holland (Amsterdam, The Netherlands) and the Beamedex® SAS company (Orsay, France). DNA samples were first diluted 1:3 in TE containing RNase A (Roche Biochemicals, Pensberg, Germany) in a SBS plate format. Each sample was run in a single well and no samples were repeated. A total of 543 samples were run, however 73 did not pass through quality check hence. Hence the final results analysis was performed on 470 samples and full results are available in Supplementary Table (S2 sheet).

NTF-analysis insertion typing of IS6110 Insertion in L2 isolates

A method described originally published in 1991 was used to analyze the presence/absence of 0-2 copies of IS6110 in the NTF locus of the L2/Beijing isolates DNAs to distinguish between modern and ancient L2/Beijing (Plikaytis, Marden et al. 1994). A microbead-based adapted method of this original protocol was also developed and confirmed the gel-based results (results not shown, method to be described elsewhere).

Taxonomic assignments (nomenclature) of molecular results; analysis of clone 94-32 prevalence

We used the latest whole-genome sequencing-based (WGS) taxonomical assignment in 7 lineages complemented by traditional spoligotyping-based nomenclature for the sublineages when robust spoligotyping signatures is demonstrated (Brudey, Driscoll et al. 2006; Demay, Liens et al. 2012; Coll, Preston et al. 2014; Stucki, Brites et al. 2016). L2/Beijing sublineages designations was done according to (Shitikov, Kolchenko et al. 2017). The prevalence of the “Central Asian” L2/Beijing 94-32 cluster (based on 24 MIRU-VNTR typing and the 15-9 nomenclature) within the L2/Beijing lineage was studied using a specific PCR-RFLP method using agarose gel electrophoresis (Allix-Beguec, Harmsen et al. 2008; Merker, Blin et al. 2015; Mokrousov, Chernyaeva et al. 2018).

GIS systems and maps building

QGIS v2.18 (www.qgis.org, Las Palmas version) was installed on Mac OSX (Apple, Cupertino, CA); Kazakhstan shapefiles were downloaded from <http://www.diva-gis.org> (www.gadm.org, version 2.5, July 2015). Previous published results as well as our results were analyzed and retrieved as .csv files (Kubica, Agzamova et al. 2005; Skiba, Mokrousov et al. 2015). Geographical coordinates of main cities in Kazakhstan were downloaded by Google search and Latitude and Longitude were added to the .csv files. Apple pies charts were built using QGIS following user’s manual.

Results

Species and subspecies Identification results

TB-SPRINT and TB-SNPID were performed in blind in France and in the Netherlands on 700 samples and 543 samples respectively and results were later cross-checked for samples run by the two methods (n=391). 100% concordance between TB-SNPID and TB-SPRINT was achieved on the first chronologically studied set of 93 samples and 99,3% of concordant subspecies identification was found at end of the study. Two unexplained discrepancies remained (cf. Supplementary Table, S4 Sheet). Typeability was better by TB-SPOL (96%; 23/700 failed) than by TB-SNPID (86,6%;73/543 failed). TB-SNPID allowed a sub-classification of L2/Beijing isolates in six subgroups based on SNPs which was not achieved by TB-SPOL (**Fig. 1**). Conversely spoligotyping allowed a fine classification of L4 isolates whereas TB-SNPID assigned only four L4 sublineages (L4.1.1/X, L4.3/LAM, L4.1.2/Haarlem, L4/Euro-American others). The L2/Beijing (SIT1 n=529, SIT190 n=5, SIT260 n=1, SIT1674, n=1) represented 79.6% (536/673) of isolates by TB-SPRINT, and 79.5% using TB-SNPID on a different set (n=374/470). The total number of isolates identified as non-L2 by the two methods was 96/470 by TB-SNPID and 164/700 by TB-SPRINT. By TB-SNPID 20.4% (96/470) belonged to L4/Euro-American. In a country such as Kazakhstan, where L2 is predominant (~80%) spoligotyping was poorly informative. TB-SNPID identifies correctly a previously spoligotyping-misclassified T sublineage, now designated as “LAM-RUS” (Mokrousov 2015; Mokrousov 2018). The most prevalent L4 sublineage was L4.3/LAM (n=51), that includes this LAM-RUS sublineage (previously T1-T5 spoligotyping-mislabelled isolates) (Mokrousov 2015; Mokrousov, Vyazovaya et al. 2016; Mokrousov 2018); the second non-L2 prevalent lineage was the L4.2.1/URAL genotype (Kovalev, Kamaev et al. 2005) (n=39, labeled as “Euro-American Other” by TB-SNPID). By TB-SNPID, five isolates were labeled as “Haarlem” (L4.1.2). By spoligotyping, twelve spoligotypes not reported in the SITVIT database were found, among which 4 were also identified as *bona fide* L4 isolates by TB-SNPID. Three spoligotyping patterns at least, likely linked to mixed infection, were detected (Supplementary Table, S4 Sheet). All L2/Beijing identified isolates belong to the L2/Modern subgroup, i.e. harboring one or two IS6110 copies in the NTF, as shown by the gel-based IS6110-NTF method, confirmed by an NTF-IS6110 microbead based assay (results not shown). L2 was also split by TB-SNPID into: Beijing SA+ (n=163 or 34%), Beijing V+/CHIN+/K2 (n=16 or 3.4%), Beijing (V-) (n=73 or 15.5%), and Beijing K1 (n=22 or 4.6%). The correspondence between this designation and the current L2 sublineage nomenclature is shown in Figure 1 (Shitikov, Kolchenko et al. 2017).

Drug-resistance results

Typability on drug-resistance conferring mutations was 76% for TB-SPRINT, and 86.3% by TB-SNPID. The DPO-based multiplexing strategy is more sensitive to DNA quality than the MLPA-based multiplexing strategy on the drug-resistance gene part of these assays. Among the set of samples run by the two methods (n=391), we detected 257 MDR by TB-SPRINT and 247 by TB-SNPID and 75 fully susceptible isolates by TB-SPRINT and 67 by TB-SNPID.

Full TB-SNPID results

Out of 470 total interpretable results, 93 (19.9%) were found to be sensitive (S) to at least RIF and INH, whereas 86 (18.3%) were Mono-resistant to INH, and 18 (3.8%) were Mono-resistant to RIF and 273 samples (58.1%) were found to be Multi-Drug-Resistant. Among all targets assessed within the *rpoB* gene, *rpoB* mutations were found essentially on codon position 531 (n=273), 522 (n=13), 526 (n=11). No mutations were detected in *rpoB* at codon position 176 whereas the codon position 516 was not targeted in this assay, contrarily to TB-RINT. The *katG* 315 S>T mutation was found in 348 samples, and *inh* -15 mutation (INH resistance) in 34 samples. An *embB* 306 mutation was found in a total of 228 samples (Ethambutol EMB resistance). Second-line Resistance on second-line drugs markers (*gyrA*, *rrs*, *rpsL*, *eis*, *rplC*) were also obtained by TB-SNPID (supplementary Table, S2 Sheet). The most frequent mutation found was the *eisG*-10 A mutation found in 54 samples.

Full TB-RINT results

The total number of samples for which both spoligotyping (TB-SPOL, 43-Plex) and genetic drug-resistance genotyping (TB-RINT results, 16-Plex) were available for comparison was 515 (73.5%). TB-RINT results on this set of 515 isolates provided an overall MDR rate of 66.8% (n=344/515), with 19.4% (n=100/515) of sensitive strains, 13% (n=67/515) INH mono-resistant, and 0.8% (n=4/515) RIF mono-resistant isolates. Of 421 L2/Beijing isolates 320 (62.3%) were MDR whereas 24 only (4.6%) were MDR out of 94 non-L2/Beijing isolates. The statistical significance of this result is very high (odds ratio=9.19 IC₉₅ [5.39, 16.13], $p < 2.2 \cdot 10^{-16}$), showing that Beijing and MDR-TB are strongly linked in Kazakhstan even if the clinical isolates of this study were not randomly selected. In this sample, the main RIF resistance mutations observed were again single point mutations in the *rpoB* gene: 531-TTG (n=316/515; 61.4%), 526-TAC (n=11; 2.1%), 531-TGG (n=4; 0.8%), 516-GTC (n=4; 0.8%), 526-GAC (n=3; 0.6%); uncharacterized (n=4) and double mutations were also observed: 526-TAC/531-TGG (n=6; 1.2%), 516-GTC/526-TAC (n=2; 0.4%), 526-TAC/531-

TTG (n=1; 0.2%), untargeted mutations in codons 509-513 region (n=2; 0.4%), 531-TTG/526-TAC (n=1; 0.2%) and untargeted mutations in codons 509-513 region and codon 526 (n=1; 0.2%). One hundred and sixty samples (31.1%) were WT for all *rpoB* gene positions screened. Some of these were phenotypically MDR. The main INH resistance observed mutations were single point mutations on *katG* gene: 315-C (n=385; 74.8%), 315-A (n=5; 1%); as well *inhA* gene: -15-T (n=5; 1.2%). Double mutations were also observed: *katG*_315-C/*inhA*_ -15-T (n=10; 1.9%) and *inhA*_ -15-T, *katG*_315-C (n=5; 1%). One hundred four (20.2%) were WT for those *katG* and *inhA* gene positions.

Discrepancies analysis between the two NAATs and comparison with Phenotypic Drug Susceptibility Testing

The discrepancies observed between the two NAATs were investigated. Some of these discrepancies are intrinsically linked due to differences in assay design as shown in **Table 1**. A thorough discrepancy analysis was further undertaken firstly on the first set of 93 isolates (Supplementary Table, S4 Sheet) and secondly on all 391 results (Supplementary Table, S5-S6 Sheets). The first set of samples showed that only one true and unexplained discordance (red colour in Supplementary Table, S4 sheet) was found between the two NAATs.

On the global set of results (n=391), taking phenotypic DST as a reference, we detected a total of 45 discordances. In eleven cases, the discrepancies were linked to assay design differences. For example, whereas TB-SPRINT does target the 531_TGG and the 516_GTC mutation, TB-SNPID does not. Conversely *rpoB*_522_TTG or *rpoB*_176_TTC are not directly targeted by TB-SPRINT whereas they are by TB-SNPID. When these differences in assay designed were discarded, the consistency between the two sets of results increases to 91,3%. In nine other cases, a phenotypic DST “Multi-drug resistant” status was not correlated with the NAAT results. For six of these nine cases, there is a partial agreement between the NAATs assays on INH resistance, and the most likely interpretation is that mutations are present elsewhere on *rpoB* and uncovered by our assays, hence different interpretations. For the three other cases, there is a partial overlap between phenotypic DST and NAATs on RIF resistance in at least one of the two assays, suggesting that other mutations linked to isoniazid resistance are not targeted. Less likely possibilities are false negative NAATs or error in Phenotypic DSTs. Hence assay concordance between the two NAATs was shown to be excellent.

Evolutionary Dynamics of L2/Beijing and main clones transmitted in Kazakhstan

To compare our results with previous molecular epidemiological studies done previously in Kazakhstan, we used a geographic information system (GIS) to see the evolution of the prevalence of L2/Beijing sublineages during a 10 year period (2005 to 2015). Results are shown in **Figure 2**. In the 2005 study published by T. Kubica *et al.* on 150 patients recruited in 2001 in 9 provinces of Kazakhstan, 91 isolates could be georeferenced (**Fig 2A**). L4 still remained quite prevalent at that time compared to L2/Beijing (**Fig. 2**, left, top). In the 2015 study published by Y. Skiba *et al.* on a total of 152 patients recruited in 2008 in 9 provinces of Kazakhstan, an increase of the L2/Beijing prevalence could already be noticed (**Fig. 2**, left, central). However in our study, performed on a sample of 632 patients in 13 regions, the prevalence of the blue lineage (L2/Beijing) clearly shows to have outnumbered the L4 everywhere (**Fig. 2**, left, bottom). We also analyzed the evolution of the prevalence of the L2/Beijing epidemic clones in Kazakhstan, that had been characterized by different methods: by IS6110-RFLP in 2005 and by 24 VNTR typing in 2015. Figure 2 shows that a limited number of clones (IS6110-RFLP clusters 4 and 6 in the 2005 article and Beijing 94-32 cluster in the 2015 article) were circulating in Kazakhstan (**Fig. 2** right, top and bottom). Since VNTR typing would have been too laborious to perform on 630 cases we did not perform systematic VNTR-based epidemiological cluster analysis during this study; a preliminary analysis on 8 VNTR and on a limited set of L2/Beijing DNA showed that our results were similar to those obtained by Skyba *et al.* 2015 (results not shown). Hence, we finally decided to analyze a total of n=356 DNAs (seven L4, three “failed”, and 346 L2/Beijing isolates) by a new PCR-RFLP analysis specifically targeting a *SigE* polymorphism that characterizes a mutation specific of the 94-32 L2/Beijing cluster (Allix-Beguec, Harmsen *et al.* 2008; Merker, Blin *et al.* 2015; Mokrousov, Chernyaeva *et al.* 2018). This mutation creates a double band of 38 and 37bp by creation of a new *AluI* site in case of codon 38 change through the CTG ->CTA mutation. We obtained 314 mutant out of 346 L2-isolates characterized, i.e. 90% of L2/Beijing 94-32 clustered isolates which definitively demonstrates an increase of the “Central Asian Outbreak” cluster compared to the previous 2005 and 2015 studies. In a specific case, this technique allowed us to demonstrate that two different L2/Beijing epidemiologically linked clusters are circulating in the prison in Stepnogorsk, with respectively 65 (94-32 cluster) and 8 cases (L2/Beijing other cluster). The study thus also provides a clear picture of L2/Beijing MDR and XDR transmission and demonstrate the increase of the transmission of the 94-32 cluster, part of the very large “Central-Asia” sublineage of L2/Beijing in Kazakhstan.

Discussion

In this study, we performed the largest molecular diagnostics study performed so far in Kazakhstan using two high-throughput multiplexed nucleic acids amplification-based tests (NAATs), and compared in blind their performances with the aim of improving simultaneously MDR-TB surveillance as well as precision and personalized medicine in Kazakhstan; we also demonstrated that the L2/Beijing outbreak in Kazakhstan is almost exclusively linked to the increased prevalence of the 94-32 L2/Beijing cluster.

On one hand we used TB-SNPID, a 50-Plex SNP-based and deletion region typing assay, that uses a Multiple Ligation-mediated Probe Amplification (MLPA) principle using identification and drug-resistance gene markers, and on the other hand, we used TB-SPRINT (TB-SPOL+TB-RINT), a 59-Plex assay that uses the DPO principle and achieves sequentially or simultaneously spoligotyping and first lane SNP-based drug-resistance gene typing (Bergval, Sengstake et al. 2012; Gomgnimbou, Abadia et al. 2012; Gomgnimbou, Hernandez-Neuta et al. 2013; Sengstake, Bablishvili et al. 2014; Yasmin, Gomgnimbou et al. 2014; Dantas, Suffys et al. 2015; Chaidir, Sengstake et al. 2016). Independently acquired TB-SPRINT and TB-SNPID showed highly congruent results, on respectively two large sets of 700 and 553 samples respectively, and 391 clinical samples that were typed by both methods.

TB-SPRINT samples were run sequentially in two runs (TB-SPOL and TB-RINT) by a dedicated trained technician (trained in 2 months) on half of the samples and on a automatized liquid handling device for the second half of the samples (Hamilton France, Villebon-sur-Yvette, France). TB-SNPID samples were run by a trained technician, using published protocols. Experienced users may perform TB-SPRINT in a single 4-Plex PCR (Spoligo and Rif-Inh typing) as demonstrated previously (Gomgnimbou, Hernandez-Neuta et al. 2013). Reagent costs were set to 20 Euros/sample for TB-SNPID and 12 euros/sample for TB-SPRINT. A raw economical analysis suggest a doubling cost to produce final results using european standard labor costs, however such costs are likely to be lower in emerging economies.

The congruence between the techniques was excellent on the molecular identification or drug-resistance subset of markers. The absence of 100% correlation between the two assays and between phenotypic and genotypic DST assays is multifactorial; both phenotypic DST and Luminex-based methods requires expert and dedicated staff and it was impossible to rerun suboptimal results in many cases due to economical constraints.

Our study has some limitations : TB-SPRINT and TB-SNPID are surveillance molecular diagnostic assays that will only detect what they are designed for, *i.e.* a remaining number of samples will always need to be sequenced on drug-resistance genes. Indeed, some

clinical isolates DNA may harbor other yet rare mutations that are not targeted in our assays (Coll, Phelan et al. 2018). However both assays achieve an excellent cost-effectiveness. An independent in-depth multicentric evaluation study of TB-SPRINT in Brazil has shown that untargeted mutations (e.g. 526 CAC->CGC in *rpoB*) may create hardly interpretable TB-SPRINT results (H. Gomes, personal communication); drug-resistance mutation panels could be adapted geographically depending on settings. Another perspective is to increase the number of drug-resistance gene mutations in future versions of assays; TB-SNPID already allows detection of some second-line drug resistance genes mutation analysis and another second-line 16-Plex “TB-EFI” -standing for Ethambutol-Fluoroquinolones-Injectables- that targets mutations in *eis*, *gyrA*, *rrs*, *embB* allows to do the same (Klotoe, Molina-Moya et al. 2018).

A new investigation project could use WGS to investigate the percentage of missed drug-resistance mutations, and adaptation of assays for specific MDR XDR-TB surveillance in Kazakhstan could be done. Among markers to be investigated in the future are the adaptative compensatory mutations such as those acting on *rpoA* and *rpoC* genes (Casali, Nikolayevskyy et al. 2014). Looking for epistasis between SNPs, could also be of added value to understand why the 94-32 cluster is so successful (Borrell, Teo et al. 2013).

The two high throughput techniques have pros and cons. On one hand, one advantage of TB-SPRINT is that it detects both the wild-type and the mutated allele, thus always providing internal controls that facilitate result analysis together with user-friendly biostatistics-guided decision tools that process raw results replacing expert interpretation. TB-SPRINT workload is also less than a one full-day assay for the 59-Plex, when run sequentially and half a day when run in a single experiment. However TB-SPRINT will not perform well in Asia (and for L2) as long as no L2-sublineages markers are used, since spoligotype of L2/Beijing is poorly informative and may be prone to convergence. Another limitation is that resistance gene typing is more prone to failure by TB-SPRINT when DNA quality and quantity are suboptimal.

TB-SNPID targets either the wild-type or the mutated allele with no internal control. Hence, bioinformatical interpretation tools are critical for this assay and if the initial design is not optimal for a given marker, the genetic analysis fails. Two markers: *rpsL* K88R and *rpoB* 516 markers gave no results and need to be redesigned. TB-SNPID requires two-days instead of one for TB-SPRINT; interestingly though, it is a one tube/one run assay on the MagPix®; conversely TB-SPRINT (a 59-Plex) requires two reads in the MagPix® and is consequently more adapted to classical Luminex 200 or FlexMap 3D users. In the current version, TB-

SNPID provides more information than TB-SPRINT since it targets second-line drug resistance markers and provides a finer, yet suboptimal L2/Beijing infra-lineage taxonomical identification to be compared to the current WGS L2/Beijing unified definition (Merker, Blin et al. 2015; Shitikov, Kolchenko et al. 2017). Conversely TB-SPRINT provides a better taxonomical identification for all other lineages than L2, based on the almost universal spoligotyping discriminatory power and on the ability to compare results with worldwide sustained databases (SITVITWEB, http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/) that will soon be updated (SITVIT-EXTEND, Rastogi N., personal communication). An improved 32-Plex version of the TB-SPRINT assay, (designed TB-SPRINT *plus*), could deliver results on first and second line drug-resistance markers simultaneously (*rpoB*, *inhA*, *katG*, *eis*, *gyrA*, *rrs*, *embB*, *pncA*) and for a reasonable price.

Even if phenotypic DST remains a gold standard, the inherent difficulties linked to this method and requires quality insurance procedures to track potential errors (Ahmad, Mokaddas et al. 2016). Hence, predictive WGS-based prediction with bioinformatical pipelines represent an important step forward for TB drug-resistance containment and these methods will progressively spread even if WGS analysis remains quite expensive (200 US\$)(Aung, Tun et al. 2018) and considering that bioinformatical pipelines are in progress though not standardized (Phelan, O'Sullivan et al. 2016). Discrepancies observed in this study between phenotypic and genotypic DST were limited and may be either due to errors in phenotypic DST or in genotypic DST. Infections with mixed resistance genotypes in the patient population are also likely and may account for a proportion of such discrepancies (Streicher, Bergval et al. 2012).

Achieving improved global tuberculosis control and patient treatment simultaneously is now feasible thanks to WGS and high-throughput SNPs assays, and improved training create favorable conditions for shared and sustained laboratory capacity strengthening, that could be applied to other infectious threats (Anonymous 2015; Anonymous 2015). Genomic assays may also allow clinical microbiology to change from targeted to nontargeted investigations (Lecuit and Eloit 2015). However the trend towards closed diagnostics systems development does not create sustainable conditions for education and democratization in health systems but generates inequalities due to unaffordable and unsustainable costs in international health (Sidibe 2016) and may create new frontiers in global health (Kehr 2018). Given the strongly regionalized phylogeographical structure of the TB outbreaks, a “*one-size fits all*” strategy for global MDR-and XDR-TB control is definitively not a solution; as demonstrated in Swaziland

by the prevalence of rare mutations such as the *rpoB* I491F mutation, that is not included in the GenXpert MTB/RIF system (Cepheid, Sunnyvale, CA), underdiagnosis and inadequate TB treatment and infection control in some countries could prosper under monopolistic markets (Sanchez-Padilla, Merker et al. 2015). Global economic issues remain with respect to sustainability and long-term routine work on two complex WGS-based assays in emerging economies, even if lot of resources are spent to train bioinformaticians (McNerney, Cunningham et al. 2015). That is the reason why, at a post-genomic era, other less sophisticated assays, that may be locally developed using new NGS results may also become important within resource-limited countries (Perez-Lago, Izco et al. 2017; Mokrousov, Chernyaeva et al. 2018; Yasmin, Refregier et al. 2018). Our study, with limited economical means is paradigmatic of such attempts.

In sum, this study in Kazakhstan demonstrates that for a limited cost, two non WGS though information-rich genomic assays could satisfactorily be run independently and successfully, with minimal costs and minimal logistics to produce highly valuable public health datasets, both for patient benefit (total results were delivered to the Kazakhstan public health system within 6-8 weeks after DNA delivery) and for public health improvement through improved molecular surveillance with minimal add-on (simple PCR-RFLP). Because of economic issues, TB-SPRINT and TB-SNPID were run only once, and typability (respectively 76% and 86%) would have undoubtedly increased with optimized sampling and/or analysis repetition. Kazakhstan has also benefitted from training and technology transfer for implementation of such methods with support of the NCB to the national reference TB program towards control of MDR XDR-TB in Kazakhstan. Other countries, such as Brazil are also investing in multiplexed bead-based methods for surveillance of MDR and XDR-TB (Dantas, Suffys et al. 2017; Santos, Costa et al. 2017).

Our study had some limitations in terms of epidemiological output. Our results if they demonstrate the spread of the L2/Beijing 94-32 cluster do not provide exactly the same level of information than a WGS-based study (Casali, Nikolayevskyy et al. 2014). However, an in-depth WGS-based analysis of L2/Beijing 94-32 cluster was done recently by I Mokrousov *et al.* that found 24 specific SNPs in L2/Beijing 94-32 cluster genomes A wide (Central-Asia) detailed polymorphism analysis of these 24 SNPs may allow to retrace the fine history of the Central L2-Asian sublineage and to date its introduction into Kazakhstan (Mokrousov, Chernyaeva et al. 2018).

Acknowledgement

This work was supported by the Ministry of Research in Kazakhstan through the project : “High-throughput and cost-effective techniques for control and surveillance of infectious diseases in Kazakhstan: the multi-drug and extremely-drug resistance tuberculosis problem” (**Priority area:** «Life Science» «Development of biomedicine and genetic engineering») **Time frame of the project:** January 2015 – December 2017, three years. We are grateful to Luminex BV (Hertogenbosch, Belgium), M. J. Van Gils, M. E. Bennink, and Luminex Corp. (Austin, TX) for constant technical support. This work was also supported by the Institut for Integrative Cell Biology (I2BC, UMR9198) and The Royal Tropical Institute (KIT). We are also grateful to M. J. Schouten for support, MRC Holland (Amsterdam; the Netherlands). This work was financed by a special grant of the MoH of Kazakhstan attributed to the NCB, Astana.

Author contributions

EZ, RA, SS and CS conceived the study. SK, HMG, SP, RCB ran and analyzed TB-SPRINT assays in Paris. AS, SS, ran and interpreted TB-SNPID assays in the Netherlands. BK ran the IS6110-NTF assays and trained NS and EZ in Kazakhstan. AA facilitated the access to MDR-TB samples in relation to EZ. DHS performed the SNPs-RFLP assay to analyze 94-32 prevalence. GR run statistical analysis and provided coordination of the study in France together with EZ in Kazakhstan and SS in the Netherlands. CS together with EZ were at the origins of scientific grant submission to the MoH of Kazakhstan; CS coordinated samples logistics reception and shipment in Paris, analyzed compiled results, built maps, wrote the manuscript and performed publishing process helped by GR. MA put in contact CS with EZ thus allowing the study to be performed.

Acknowledgements

We are grateful to the MoH of Kazakhstan, to Dr. Y. Ramankulov, Kazakhstan NCB head, for allowing this project to be run, and to Beamedex® (Pasquale Cutolo, Michel K Gomgnimbou, Clément Ripoll) and Luminex BV staff in the Netherlands (J van Gils, E. Benning) for technical support. BK is a PhD from Beamedex® supported by a Beamedex® CIFRE (convention industrielle de formation par la recherche en entreprise) PhD working contract, supported by ANRT (Agence Nationale de la Recherche Technologique) through a CNRS-Beamedex research and development contract. CS is grateful to the CNRS and Institut for Integrative Cell Biology, UMR9198, for financial support including the purchase and maintenance of the Hamilton StarLet automatized liquid handling system. Hamilton Corp

(Hamilton France, Les Ulis, L. Lorida and J. Vicenzi) is also acknowledged for technical support.

Conflict of Interest

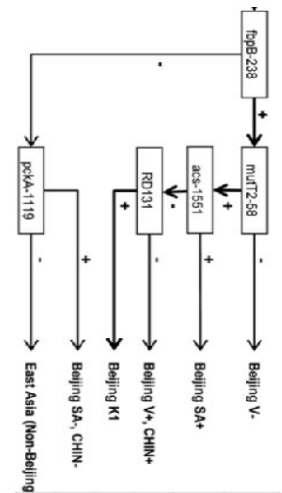
CS and GR were among the founders and shareholders of Beamedex SAS®, at start of the study. They declare not to hold any share in Luminex Corp. (Austin, TX) nor to receive any salary from Beamedex® and do not hold Beamedex® shares anymore. BK is a PhD student from Beamedex® supported by a French ANRT-CIFRE subsidized contract, and CNRS-Beamedex contract.

Table 1 : Drug-resistance markers analyzed in the two high-throughput microbead-based assays

DR Markers assessed	TB-SPRINT	MLPA (TB-SNPID)
rpoB_176_mut		x
rpoB_516_GAC_wt	x	
rpoB_516_GTC_mut1	x	x
rpoB_516_TAC_mut2	x	
rpoB_522		x
rpoB_526_CAC_wt	x	
rpoB_526_TAC_mut1	x	x
rpoB_526_GAC_mut2	x	x
rpoB_531_TCG_wt	x	x
rpoB_531_TTG_mut1	x	
rpoB_531_TGG_mut2	x	
katG_315_AGC_wt	x	x
katG_315_ACC_mut1	x	
katG_315_AAC_mut2	x	
inhA_-15_C_wt	x	
inhA_-15_T_mut1	x	
inhA_-16_A_wt		
inhA_-16_G_mut1		
inhA_-16_G_wt		
inhA_-8_T_wt		
inhA_-8_A_mut1	x	
gyrA_90-91_GCGTCG_wt		x
gyrA_94_wt_GAC		x
rrs_1401_wt_A		x
rrs1402_wt_C		x
eis_(-10A)_CACAA		x
eis_(-14T)_TACAG		x
rpsl-43		x
rpsl-88		x
rplC-460		x
Reference	<i>Gomgnimbou et al. 2012;</i> <i>Gomgnimbou et al. 2013</i>	<i>Bergval et al. 2012;</i> <i>Sengstake et al. 2014</i>

Bergval et al. 2012

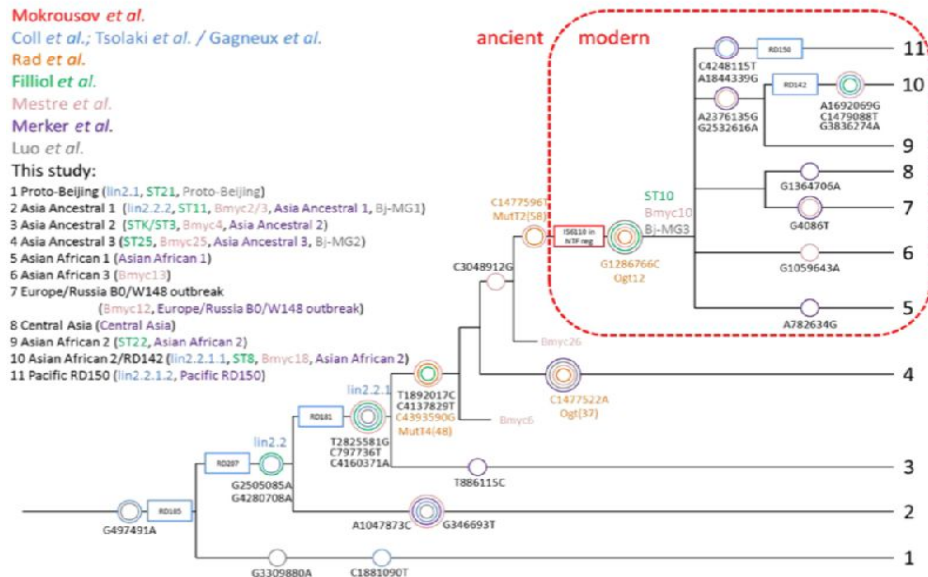
Shitikov et al. 2017



Mokrousov et al.
 Coll et al.; Tsolaki et al. / Gagneux et al.
 Rad et al.
 Filliol et al.
 Mestre et al.
 Merker et al.
 Luo et al.

This study:

- 1 Proto-Beijing (lin2.1, ST21, Proto-Beijing)
- 2 Asia Ancestral 1 (lin2.2.2, ST11, Bmyc2/3, Asia Ancestral 1, Bj-MG1)
- 3 Asia Ancestral 2 (STK/ST3, Bmyc4, Asia Ancestral 2)
- 4 Asia Ancestral 3 (ST25, Bmyc25, Asia Ancestral 3, Bj-MG2)
- 5 Asian African 1 (Asian African 1)
- 6 Asian African 3 (Bmyc13)
- 7 Europe/Russia B0/W148 outbreak (Bmyc12, Europe/Russia B0/W148 outbreak)
- 8 Central Asia (Central Asia)
- 9 Asian African 2 (ST22, Asian African 2)
- 10 Asian African 2/RD142 (lin2.2.1.1, ST8, Bmyc18, Asian African 2)
- 11 Pacific RD150 (lin2.2.1.2, Pacific RD150)



-	-	-	+	+	-
-	-	-	+	-	-
-	-	-	+	-	-
-	-	+	-	-	-
-	-	-	+	-	-
-	-	-	+	-	-
-	-	-	-	-	+
-	+	-	-	-	-
-	-	-	-	-	-

Central Asian Outbreak 94-32



Figure 1: Algorithm of current updated taxonomy of *Mycobacterium tuberculosis* L2/Beijing isolates by TB-SNPID adapted from Shitikov et al. 2017 (11 sublineages based on WGS). The 4 genes used in TB-SNPID are mentioned on top left with SNP variation description : *fbpB*, *pckA*,

acs, *mutT2* and on RD131 (Niemann et al. 2009, Bergval et al. 2012); TB-SNP-ID naming and classification (top right corner) refers to the geographically named groups based on Niemann et al. 2009 (K1) and on Schürch et al. 2011 (CHIN-, CHIN+, SA-, SA+, V-, V+). The matrix depicts the correspondance between Shitikov's and previously published taxonomy. Orange boxes are all fbp+. On the tree, all current validated SNP markers are reported with previous other author's nomenclatures. The *ancient* versus *modern* L2/Beijing nomenclature refers to previous works from I. Mokrousov et al. 2005 based on the presence/absence of IS*6110* copies in the NTF.

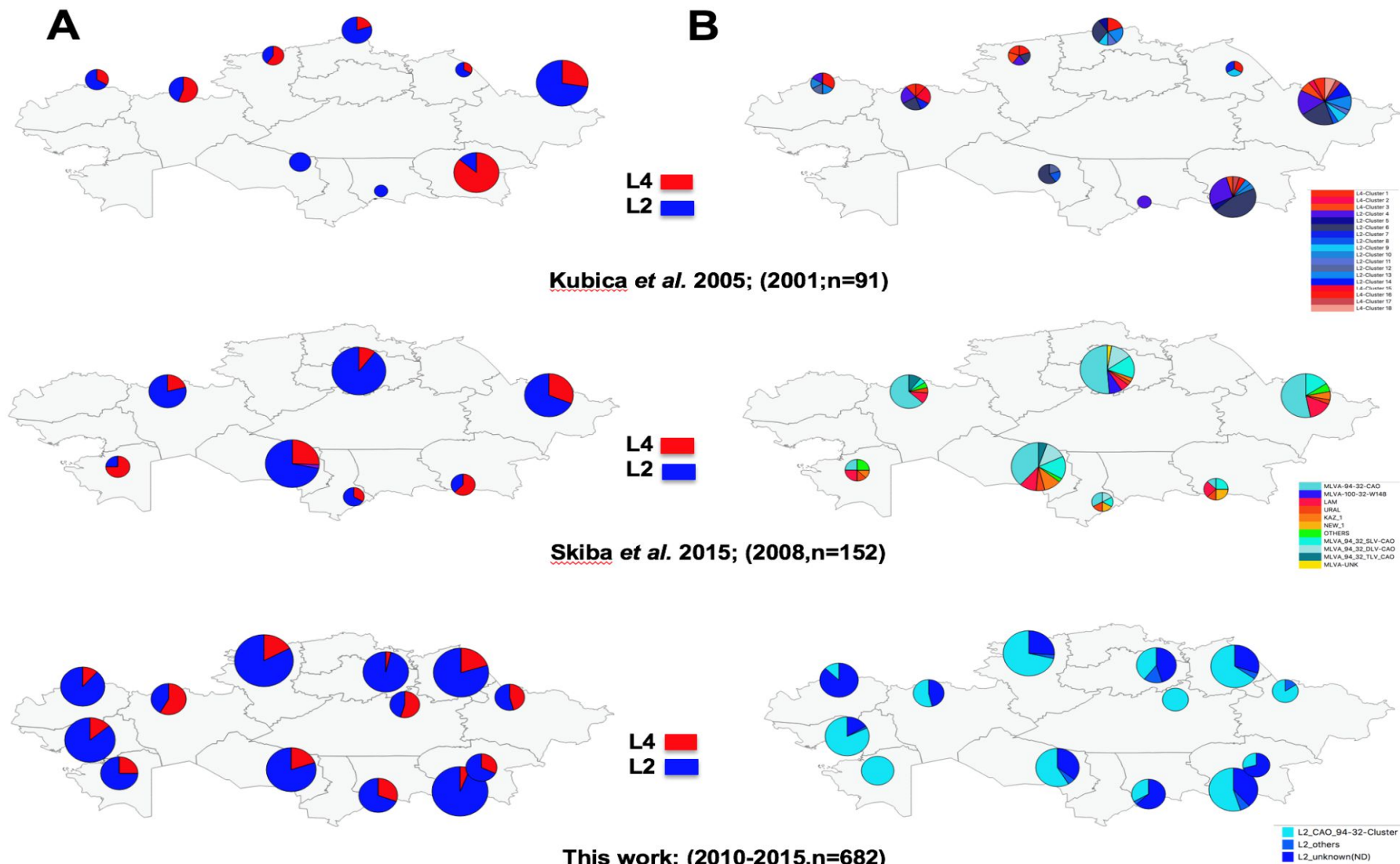


Figure 2 : Maps of Kazakhstan showing *Mycobacterium tuberculosis* diversity L2/Beijing (Blue) and L4-Euro-American (Red) relative

prevalence based on three different studies published in 2005, 2015, and in our study, based on samples recruited respectively in 2001 (n=91), 2008 (n=152) and 2014-2015 (n=632).

A – at the lineage (L2, L4), B – at the sublineage level and based on different markers depending on the study. - -

Supplementary Table

S1 Sheet: detailed TB-SPRINT results on 700 MTC DNA samples from Kazakhstan (n=677 spoligotypes including 3 potential "mixed infection", 23 failed).

S2 Sheet: detailed TB-SNPID results on 470 MTC DNA samples from Kazakhstan (n=543 analysed, 73 failed)

S3 Sheet: Lineage and Sublineage assignation of Spoligotyping results (n=677).

S4 Sheet: All isolates (n=93) of "set1" (i.e. chronologically first set) typed by two methods: TB-SPRINT and TB-SNPID

S5 Sheet: Detailed results for n=391 samples run both by TB-SPRINT and TB-SNPID.

S6 Sheet: Analysis of discrepant drug-resistance results between TB-SPRINT and TB-SNPID on the 391 set of isolates

References

1. WHO (2017) Global Tuberculosis Report 2017. (WHO, Geneva, Switzerland).
2. WHO (2016) Global Tuberculosis Report 2016. (WHO, Geneva, Switzerland).
3. WHO (2015) World Health Statistics 2015. (WHO Geneva), pp 1-164.
4. Bergval I, *et al.* (2012) Combined Species Identification, Genotyping, and Drug Resistance Detection of *Mycobacterium tuberculosis* Cultures by MLPA on a Bead-Based Array. *PLoS One* 7(8):e43240.
5. Gomgnimbou MK, Refrégier G, & Sola C (2013) High Throughput combin methods for resistanc emergence analysis, diagnostic and Epidemiology: the multi-drug resistant tuberculosis paradigm. *10th International Meeting on Microbial Epidemiological Markers*.
6. Sengstake S, *et al.* (2014) Optimizing multiplex SNP-based data analysis for genotyping of *Mycobacterium tuberculosis* isolates. *BMC genomics* 15(1):572.
7. Molina-Moya B, *et al.* (2017) Molecular Characterization of *Mycobacterium tuberculosis* Strains with TB-SPRINT. *The American journal of tropical medicine and hygiene* 97(3):806-809.
8. Dantas NGT, *et al.* (2017) Correlation between the BACTEC MGIT 960 culture system with Genotype MTBDRplus and TB-SPRINT in multidrug resistant *Mycobacterium tuberculosis* clinical isolates from Brazil. *Memorias do Instituto Oswaldo Cruz* 112(11):769-774.
9. Santos P, *et al.* (2017) Detection of tuberculosis drug resistance: a comparison by *Mycobacterium tuberculosis* MLPA assay versus Genotype(R)MTBDRplus. *Memorias do Instituto Oswaldo Cruz* 112(6):396-403.
10. Dunbar SA (2006) Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection. *Clinica chimica acta; international journal of clinical chemistry* 363(1-2):71-82.
11. Ocheretina O, *et al.* (2013) Use of Luminex MagPlex(R) Magnetic Microspheres for High-Throughput Spoligotyping of *M. tuberculosis* Isolates in Port-au-Prince, Haiti. *J Clin Microbiol*.
12. Schouten J, *et al.* (2002) Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucl. Acid. Res.* 30(12):e57.
13. Chun JY, *et al.* (2007) Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of CYP2C19 gene. *Nucleic Acids Res* 35(6):e40.
14. David H, Levy-Frebault V, & Thorel MF (1989) Méthodes de laboratoire pour Mycobactériologie clinique. (Institut Pasteur, Paris), pp 1-87.
15. Gomgnimbou MK, *et al.* (2013) “TB-SPRINT: TUBERCULOSIS-SPOLIGO-RIFAMPIN-ISONIAZID TYPING”; an All-in-One assay technique for surveillance and control of multi-drug resistant tuberculosis on Luminex® devices. *J Clin Microbiol* 51(11):3527-3534.
16. Plikaytis BB, *et al.* (1994) Multiplex PCR assay specific for the multidrug-resistant strain W of *Mycobacterium tuberculosis*. *J Clin Microbiol* 32(6):1542-1546.
17. Coll F, *et al.* (2014) PolyTB: A genomic variation map for *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* 94(3):346-54(3):346-354.
18. Stucki D, *et al.* (2016) *Mycobacterium tuberculosis* lineage 4 comprises globally distributed and geographically restricted sublineages. *Nature genetics*.
19. Brudey K, *et al.* (2006) *Mycobacterium tuberculosis* complex genetic diversity : mining the fourth international spoligotyping database (SpolDB4) for classification, Population Genetics, and Epidemiology. *BMC Microbiol.* 6(6):23.
20. Demay C, *et al.* (2012) SITVITWEB – A publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. *Infect Genet Evol* 12(4):755-766.
21. Allix-Beguec C, Harmsen D, Weniger T, Supply P, & Niemann S (2008) Evaluation and user-strategy of MIRU-VNTRplus, a multifunctional database for on-line analysis of genotyping data and phylogenetic identification of *Mycobacterium tuberculosis* complex isolates. *J Clin Microbiol* 46(8):2692-2699.
22. Merker M, *et al.* (2015) Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nature genetics* 47(3):242-249.
23. Mokrousov I, *et al.* (2018) Rapid Assay for Detection of the Epidemiologically Important Central Asian/Russian Strain of the *Mycobacterium tuberculosis* Beijing Genotype. *J Clin Microbiol* 56(2).
24. Kubica T, *et al.* (2005) The Beijing genotype is a major cause of drug-resistant tuberculosis in Kazakhstan. *Int J Tuberc Lung Dis* 9(6):646-653.
25. Skiba Y, *et al.* (2015) Molecular snapshot of *Mycobacterium tuberculosis* population in Kazakhstan: a country-wide study. *Tuberculosis (Edinb)* 95(5):538-546.
26. Mokrousov I (2015) Molecular structure of *Mycobacterium tuberculosis* population in Russia and its interaction with neighboring countries. *Int J Mycobacteriol* 4 Suppl 1:56-57.
27. Mokrousov I (2018) On sunspots, click science and molecular iconography. *Tuberculosis (Edinb)* 110:91-95.

28. Mokrousov I, *et al.* (2016) Latin-American-Mediterranean lineage of *Mycobacterium tuberculosis*: Human traces across pathogen's phylogeography. *Molecular phylogenetics and evolution* 99:133-143.
29. Kovalev SY, Kamaev EY, Kravchenko MA, Kurepina NE, & Skorniakov SN (2005) Genetic analysis of *Mycobacterium tuberculosis* strains isolated in Ural region, Russian Federation, by MIRU-VNTR genotyping. *Int J Tuberc Lung Dis* 9(7):746-752.
30. Shitikov E, *et al.* (2017) Evolutionary pathway analysis and unified classification of East Asian lineage of *Mycobacterium tuberculosis*. *Scientific reports* 7(1):9227.
31. Klotoe BJ, *et al.* (2018) TB-EFI, a novel 18-Plex microbead-based method for prediction of second-line drugs and ethambutol resistance in *Mycobacterium tuberculosis* complex. *J Microbiol Methods*.
32. Gomgnimbou MK, *et al.* (2012) << Spoligoriotyping >> a DPO-based direct-hybridization assay for TB control on a multianalyte microbead-based hybridization system. *J Clin Microbiol* 50(10):3172-3179.
33. Yasmin M, Gomgnimbou MK, Siddiqui RT, Refregier G, & Sola C (2014) Multi-drug resistant *Mycobacterium tuberculosis* complex genetic diversity and clues on recent transmission in Punjab, Pakistan. *Infect Genet Evol* 27C:6-14.
34. Dantas N, *et al.* (2015) Multidrug-resistant *Mycobacterium tuberculosis* Genetic Diversity in Minas Gerais state, Brazil. *BMC Infectious Diseases* 15:306.
35. Chaidir L, *et al.* (2016) Predominance of modern *Mycobacterium tuberculosis* strains and active transmission of Beijing sublineage in Jayapura, Indonesia Papua. *Infect Genet Evol* 39:187-193.
36. Coll F, *et al.* (2018) Genome-wide analysis of multi- and extensively drug-resistant *Mycobacterium tuberculosis*. *Nature genetics*.
37. Casali N, *et al.* (2014) Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nature genetics* 46(3):279-286.
38. Borrell S, *et al.* (2013) Epistasis between antibiotic resistance mutations drives the evolution of extensively drug-resistant tuberculosis. *Evolution, medicine, and public health* 2013(1):65-74.
39. Ahmad S, Mokaddas E, Al-Mutairi N, Eldeen HS, & Mohammadi S (2016) Discordance across Phenotypic and Molecular Methods for Drug Susceptibility Testing of Drug-Resistant *Mycobacterium tuberculosis* Isolates in a Low TB Incidence Country. *PLoS One* 11(4):e0153563.
40. Aung HL, *et al.* (2018) Association between anti-tuberculosis drug resistance-conferring mutations and treatment outcomes in Myanmar. *Infect Dis (Lond)* 50(5):388-390.
41. Phelan J, *et al.* (2016) The variability and reproducibility of whole genome sequencing technology for detecting resistance to anti-tuberculous drugs. *Genome medicine* 8(1):132.
42. Streicher EM, *et al.* (2012) *Mycobacterium tuberculosis* population structure determines the outcome of genetics-based second-line drug resistance testing. *Antimicrob Agents Chemother* 56(5):2420-2427.
43. Anonymous (2015) Tuberculosis Diagnostics Technology and Market Landscape. (WHO, Geneva, UNITAID Secretariat Chemin de Blandonnet 10 – BIBC III – 8th Floor 1214 Vernier Switzerland), pp 1-96.
44. Anonymous (2015) TB diagnostics market in select high-bruden countries : current market and future opportunities for novel diagnostics. (WHO, Geneva, UNITAID Secretariat Chemin de Blandonnet 10 – BIBC III – 8th Floor 1214 Vernier Switzerland), pp 1-68.
45. Lecuit M & Eloit M (2015) The potential of whole genome NGS for infectious disease diagnosis. *Expert review of molecular diagnostics* 15(12):1517-1519.
46. Sidibe M (2016) Democratising the global health agenda: why we need France. *Lancet* 6736(16):30060-30065.
47. Kehr J (2018) 'Exotic no more': Tuberculosis, public debt and global health in Berlin. *Glob Public Health* 13(3):369-382.
48. Sanchez-Padilla E, *et al.* (2015) Detection of drug-resistant tuberculosis by Xpert MTB/RIF in Swaziland. *N Engl J Med* 372(12):1181-1182.
49. McNerney R, Cunningham J, Hepple P, & Zumla A (2015) New tuberculosis diagnostics and rollout. *Int J Infect Dis* 32:81-86.
50. Perez-Lago L, *et al.* (2017) A novel strategy based on genomics and specific PCR reveals how a multidrug-resistant *Mycobacterium tuberculosis* strain became prevalent in Equatorial Guinea 15 years after its emergence. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* 23(2):92-97.
51. Yasmin M, *et al.* (2018) Reverse line probe assay for cheap detection of Single Nucleotide Polymorphisms in *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* 110:52-55.

**Etude 5 : Analyse de la diversité des souches
MDR de la lignée 2 circulant au Pérou**

Résumé

Cette étude s'inscrit dans le cadre du projet international ERANET-LAC « TRANS-TB TRANS » coordonné par l'Espagne (Hôpital G. Maranon, Madrid, Investigateur Principal, Dr. Dario Garcia de Viedma) et qui vise à développer des méthodes rapides de surveillance de clones multi-résistants. Ainsi, en partenariat avec différents pays d'Amérique latine (Argentine, Panama, Pérou) et d'Europe (Italie, Espagne) 60 souches Multirésistantes (MDR) diagnostiquées entre 2014-2015 au centre de santé de San Juan de Lurigancho, un quartier au Nord-Est de Lima, la capitale, du Pérou, ont été analysées par TB-SPRINT (spoligotypage-et résistance à la rifampicine et l'isoniazide), MIRU-VNTR et WGS. Ces souches MDR représentent 20% de tous les cas de MDR-TB dans toute la capitale.

L'analyse par MIRU-VNTR et TB-SPRINT montre que les souches étudiées appartiennent à la lignée 4 et la lignée 2, et parmi elles, aux sous-lignées L4.3/LAM, L4.1.2/Haarlem et L2/Beijing. La transmission récente d'isolats MDR est élevée : ils forment neuf grappes épidémiques. En comparaison d'une étude antérieure (étude de 2010-2011), cette étude révèle un plus faible nombre de grappes MDR-TB. Nous avons voulu tester l'hypothèse de la persistance des cas de MDR-TB de 2010-2011. Les données de génotypage de 2010-2011 ont été comparées à celles de 2014-2015. Ces études prouvent une transmission active persistante des mêmes souches MDR dans le quartier de San Juan de Lurigancho à Lima. Nous avons ensuite exploré si ces souches ont été exportées dans d'autres pays par des migrations humaines transcontinentales. Ces transmissions intercontinentales ont été démontrées par la suite grâce aux données de séquençage de génome complet : la même souche MDR a causé des cas de MDR-TB à Lima (2 cas), Florence (11 cas), à Milan (2 cas) (Italie), à Madrid (1 cas) (Espagne).

En se basant sur cette étude préliminaire de caractérisation des clones MDR circulant à Lima, on déduit donc une transmission active des souches MDR de Lima, transmission qui s'est étendue à d'autres pays d'Europe, ce qui pourrait constituer un sérieux problème de Santé Publique. D'une part cette étude préliminaire nous a permis de prouver la persistance et l'augmentation de la transmission des souches MDR-TB au Pérou, et d'autre part de servir de base à l'identification de cas de transmission au sein d'autres populations. La transmission des souches MDR est difficile à maîtriser dans le cadre de l'augmentation des phénomènes d'immigration en Europe en provenance de pays, comme le Pérou. Constituer des réseaux de surveillance est essentiel pour détecter l'importation et la dissémination des cas de MDR-TB en Europe, et ainsi identifier des facteurs de risque de dissémination de tuberculose multirésistante. Les résultats de cette première étude ont été rendus possibles tout d'abord par

l'utilisation de techniques performantes et à faible coût comme TB-SPRINT et de résultats obtenus par MIRU-VNTR 24. Ils ont été poursuivis par le travail de WGS sur des lots restreints de souches. Ces derniers ont conduit à la mise au point de techniques simples de type *Allele-Specific Primer Extension* (ASPE) sur des SNPs caractérisant ces clones épidémiques (d'après l'analyse WGS). Cette dernière technique à très bas coût permettra une surveillance soutenable des futures épidémies.

Le manuscrit qui suit sera prochainement soumis pour publication.

Article 5 :

Exportation of multidrug-resistant tuberculosis to Europe from a setting with actively transmitted persistent strains in Peru

Fermín Acosta^{1,2}, Juan Agapito³, Andrea María Cabibbe⁴, Tatiana Cáceres³, Christophe Sola⁵, Laura Pérez-Lago^{1,2}, Estefanía Abascal^{1,2}, Marta Herranz^{1,2,6}, Erika Meza³, **Bernice Klotoe**⁵, Patricia Muñoz^{1,2,6}, Gian María Rossolini⁷, Alessandro Bartoloni⁸, Enrico Tortoli⁴, Daniela María Cirillo⁴, Eduardo Gotuzzo³, Darío García de Viedma^{1,2,6+}

¹Instituto de Investigación Sanitaria Gregorio Marañón, ²Hospital General Universitario Gregorio Marañón, Servicio de Microbiología, Madrid, Spain; ³TB Research Unit, Instituto de Medicina Tropical Alexander von Humboldt, Universidad Peruana Cayetano Heredia, Lima, Peru; ⁴Emerging Bacterial Pathogens Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy; ⁵Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, France; ⁶CIBERES, CIBER Enfermedades Respiratorias, Spain; ⁷Clinical Microbiology and Virology Unit, Florence Careggi University Hospital, Florence, Italy; ⁸Infectious and Tropical Diseases Unit, Careggi Hospital, Florence, Italy

+corresponding author: dgviedma2@gmail.com

; Phone: (+34) 914265104. Postal address: C/Dr. Esquerdo 46 (28007), Madrid, Spain.

Summary

Multidrug-resistant tuberculosis (MDR-TB) has become a major health problem worldwide. Although the highest rates of MDR-TB are concentrated in a few countries, migratory movements from these areas could export high-risk strains to Europe. An integrated cross-border molecular epidemiology analysis allowed us to illustrate the magnitude and persistence of MDR-TB transmission in Peru and how this led to the exportation of an MDR-TB strain that replicated similar uncontrolled transmission in Europe.

Keywords: Tuberculosis, MDR, migration, exportation, transmission, molecular epidemiology

International migratory movements have created the need for cross-border surveillance of tuberculosis (TB). Monitoring of the transmission dynamics of multidrug-resistant TB (MDR-TB) deserves special attention and broader analysis (Zignol, Dean et al. 2016). In 2016, 89.7% of incident cases of MDR-TB worldwide were concentrated in 30 countries (2016) (WHO 2017). Local pools of MDR strains of *Mycobacterium tuberculosis* can therefore be exported through migratory movements to lower-prevalence settings. In the host country, these events could replicate, albeit on a lower scale, the problems related to MDR-TB that are common in the countries of origin.

We performed a cross-border molecular epidemiology analysis of MDR-TB covering a setting with one of the highest rates of resistance in Latin America, namely, Lima, Peru (WHO 2017) and 2 European settings hosting Peruvian immigrant communities (Florence, Italy and Madrid, Spain). Our aims were, first, to obtain a snapshot of the MDR-TB transmission dynamics in the source area, Peru, and second, to try to establish relationships with the strains infecting the Peruvian immigrant populations hosted in Europe.

Actively transmitted MDR-TB in Lima

For the Peruvian sample, we selected 60 consecutive MDR-TB cases diagnosed at the 32 health centers of San Juan de Lurigancho district (SJL, North-East Lima) in 2014-15. Lima is an interesting geographic target, because it is home to 82% of all MDR cases in Peru, and SJL is one of the poorest districts, with the highest incidence of TB in the city (193 per 100,000 inhabitants) (Barletta, Otero et al. 2015). The MDR isolates compiled (around 20% of the total MDR cases in Lima) were assigned to lineages based on the analysis of lineage marker SNPs and/or on spoligotyping (Stucki, Malla et al. 2012; Molina-Moya, Gomgnimbou et al. 2017). Of the 60 cases, 56 corresponded to lineage 4 and the remaining 4 to lineage 2. MIRU-VNTR analysis (Supply, Allix et al. 2006; Aze, Sola et al. 2015) revealed that the most frequent sublineages were LAM (L4.3, 65%), Haarlem (L4.1.2, 15%), and Beijing (L2, 6.7%) (Table 1).

MIRU-VNTR-based fingerprinting revealed a high percentage of recent transmission, with 36 isolates (60%) included in 9 clusters (2-10 isolates). One of the smallest clusters (N=2, C9-LPMDR) involved a Beijing strain (Table 2). None of the clusters corresponded to cases exclusively from one health center but were distributed among various health centers. The percentage of clustered cases in our analysis was higher than the average values obtained elsewhere, even for susceptible strains, which are often considered more efficiently

transmitted, as their fitness is not affected by the cost of acquiring resistance. Our data support the accepted observations that MDR strains harbouring mutations with no fitness-cost can be transmitted at similar rates to susceptible strains (Gagneux 2009). Moreover, adaptive mutations could restore fitness in other MDR strains (Casali, Nikolayevskyy et al. 2014).

It is important to remember that our sample included only 20% of the total cases and that the recruitment period was shorter than recommended (Borgdorff, van den Hof et al. 2011) to efficiently capture transmission clusters. Therefore, the true proportion of transmission could be even higher. A previous study performed in the same district in Lima in 2010-11 identified only 14/51 clustered MDR cases (27.4%) (Barletta, Otero et al. 2015), suggesting that MDR transmission rates seem to have been increasing alarmingly in recent years. It is true that the previous study enrolled exclusively first TB episodes, thus disregarding the role of previously treated cases in transmission dynamics.

Actively transmitted persistent MDR strains in Lima

Once highly active transmission of MDR-TB in Lima was identified, we moved on to evaluate whether the MDR strains were subject to constant turnover in this population or, alternatively, whether the MDR strains that are currently being transmitted constitute a persistent problem in this setting. To clarify this issue, we took advantage of genotyping data available for 228 strains (including MDR, monoresistant, polyresistant, and susceptible strains) from new TB cases, which had been genotyped in the same district of Lima 4 years earlier (March 2010-December 2011) (Barletta, Otero et al. 2015). The integration of the 2014/15 and 2010/11 datasets revealed that 6 out of the 9 actively transmitted strains in 2014/15 were already found in the 2011 survey (Table 2). Four were already MDR strains (corresponding to the 2 largest clusters [10 and 8 strains] and to 2 smaller clusters [C9-LPMDR and C4-LPMDR], including 2 and 3 cases) (Table 2). Three of these 4 strains were already part of MDR transmission clusters in 2011 (including 2, 4 and 6 cases).

Exported MDR strains

The description of a group of persistent MDR strains involved in currently active transmission events in Lima could indicate that some of these strains were exported through migratory movements from Peru to Europe. To address this issue, we selected 2 countries that traditionally host immigrants from Peru, namely, Spain (Madrid) and Italy (Florence) (Zammarchi, Tortoli et al. 2014).

For comparison with the Italian strains, we used a dataset of MIRU-VNTR genotypes obtained from 87 isolates from Peruvian cases in Florence for the period 2001-10 (Zammarchi, Tortoli et al. 2014), together with more than 300 MDR isolates genotyped retrospectively as part of a national MDR surveillance program since 2007 (REPORT 2017).

We found 2 matched genotypes between the Lima and Italian datasets. One (C2-LPMDR) corresponded to a non-MDR strain in Italy. The second one corresponded to a MDR strain infecting 9 cases in Florence and 2 in Milan (mostly Peruvians, among those with a known nationality) during 2007-2014 (Table 3), currently transmitted in Lima (C8-LPMDR), and also identified in Lima in 2010-11, when it was still susceptible to RIF and INH. More recently, 2 additional isolates of this strain were identified by WGS from 2 Peruvian patients in Florence in 2016 and 2017.

Exportation of an MDR strain detected in Lima to a population receiving Peruvian immigrants in Italy was also described for the second European host country, Spain. We used MIRU-VNTR to analyze 14 MDR isolates obtained from Peruvian immigrants during the period 2003-9. We found 3 MDR strains, each matching 3 of the 9 MDR clusters previously found in Lima. One corresponded to the genotype shared by the active Italian MDR cluster.

We performed WGS analysis, as described in Herranz *et al.* (Herranz, Pole et al. 2017), for the 11 isolates available involved in the cross-border MDR cluster C8-LPMDR (7 from Florence, 1 from Milan, the two isolates from Peru and the one from Madrid). The isolates were distributed along two branches in the network (Figure). One of the branches included the Florence isolates, with short distances between them (0-2 SNPs) thus confirming to be result of recent transmission in Florence. The recent transmission is consistent with the fact that patients from countries other than Peru (Italy, Albania) were also part of the Florence cluster, thus ruling out a founder effect. One of the isolates from Lima, Peru, (6068) was positioned only at 3 SNPs from the Florence isolates, emphasizing the close relationship between the isolates transmitted in Florence and one of the two isolates in Lima, which, again supports that exportation was the starting point for this branch.

The second branch in the network included two cases, from Madrid and Milan, together with the second isolate from Lima, Peru. The fact that the most recent common ancestor positioned as a hub between this branch and the Florence one and the finding that each of the two isolates from Lima are part of one different branch, suggests that these two

branches represents two independent exportations of two variants of a prevalent strain at source that diversified after a prolonged period.

The data from Italy and Spain revealed exportation of high-risk strains from Lima to 2 Mediterranean countries within Europe. Not only were the strains exported, but a prevalent strain at origin caused a prolonged, uncontrolled, and still active transmission event in one of the host countries, Italy, whose socio-epidemiological context differs from that of Peru, thus indicating that it can become established in various settings. In summary, the strain was responsible for at least 2 cases in Lima, 11 in Florence, 2 in Milan, and 1 in Madrid.

A recent joint cross-border study investigated the suspected international distribution of an MDR strain involving 10 cases from Romania, Austria, and Germany (Fiebig, Kohl et al. 2017). Our event is geographically wider, covering the continents of America and Europe. In addition, it is even greater in magnitude, since it includes 16 cases, with recent transmission in one of the involved countries, Italy, leading to a cluster of 12 cases that is more extended in time, with secondary cases spanning 11 years.

Only integrative transnational efforts can provide a clearer picture of the transmission of MDR-TB, which has become much more complex owing to increasing international migration. Based on a cooperative analysis involving Peru, Italy, and Spain, we detected a serious problem of active transmission of MDR-TB in Lima. This transmission led to a pool of persistent and efficiently transmitted strains that were also responsible for recreating similar uncontrolled transmission events after exportation to Europe via immigration.

Table 1. Distribution of lineages for the MDR isolates from Lima, Peru, 2014–2015

Lineage	No. of isolates (%)
LAM	39 (65)
Haarlem	9 (15)
Cameroon	3 (5)
X	3 (5)
S	2 (3.3)
Beijing	4 (6.7)
Total	60

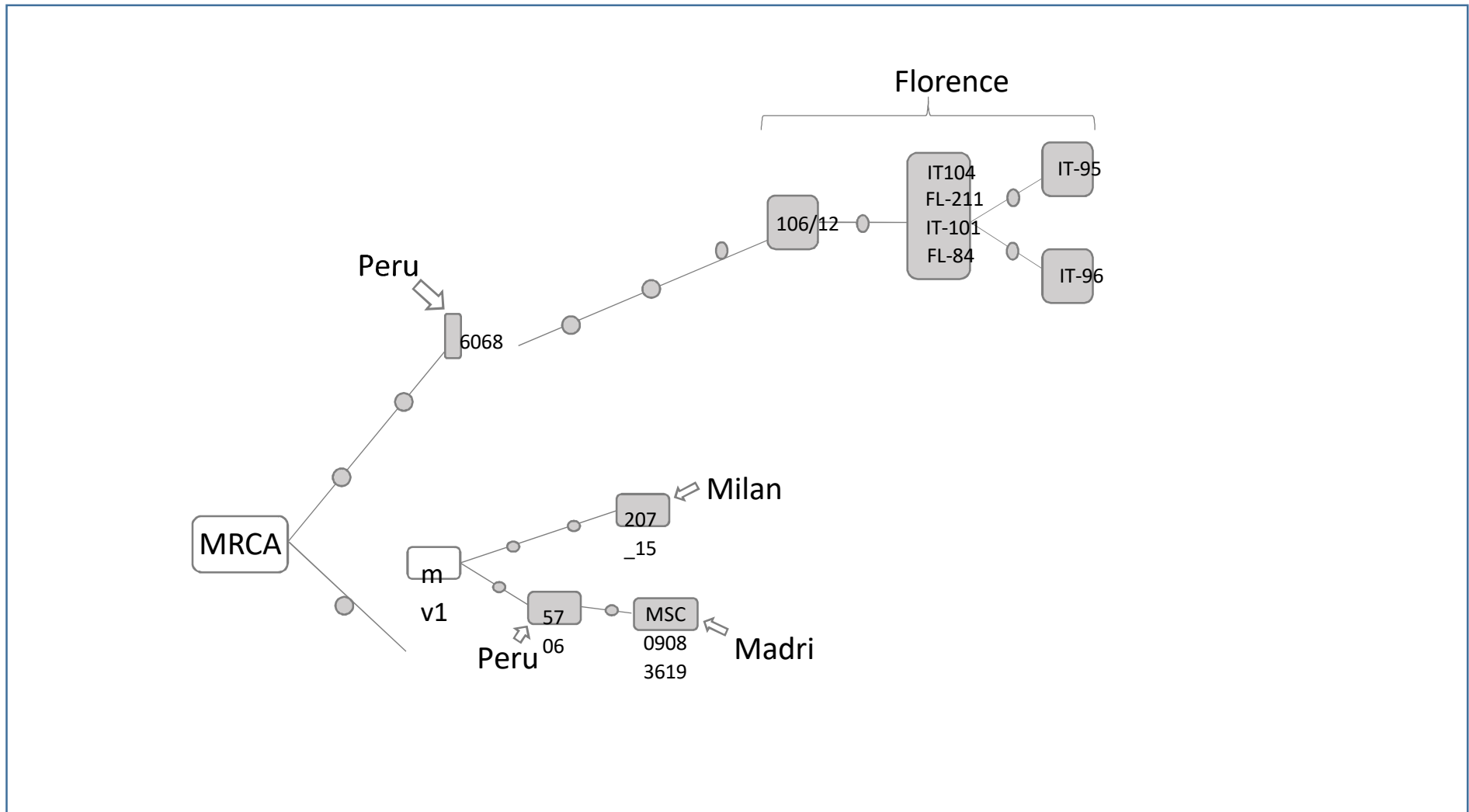


Figure: Median-joining network for the isolates, with WGS data available, involved in cluster C8-LPMDR (Network 4.6.1.6 program). Each dot along the lines linking the isolates corresponds to a single-nucleotide polymorphism (SNP). Isolates within the same box share identical

sequences (0 SNPs between them). The mv1 corresponds to an unsampled case which is inferred from the network topology. MRCA: Most recent ancestor. Sequences were deposited in <http://www.ebi.ac.uk> (accession number PRJEB25765).

Table 3. Features of the cases included in the MDR cluster in Florence, Italy matching with Lima, Peru.

Cluster name	City of isolation	Country of origin	Year of isolation
C8-LPMDR	Florence	Latin America	2007
	Milan	Peru	2009
	Florence	Peru	2010
	Milan	Latin America	2011
	Florence	Latin America	2012
	Florence	Peru	2013
	Florence	Italia	2013
	Florence	Albania	2013
	Florence	Peru	2013
	Florence	Latin America	2014
	Florence	Latin America	2014
	Florence	Peru	2016
	Florence	Peru	2017

Acknowledgments

We thank Thomas O'Boyle for proofreading the manuscript. The project was funded by ERANET-LAC (ELAC2015/T08-0664, E035-ERAnet-LAC/J110-2016/FONDECYT, PER-2012-ELAC2015/T08-0664) and ISCIII (AC16/00057, FIS15/01554, 13/01207). LPL holds a Miguel Servet grant (CP15/00075). FA holds a grant from IFARHU-SENACYT (270-2016-293).

References

1. Zignol M, Dean AS, Falzon D, van Gemert W, Wright A, van Deun A, et al. Twenty Years of Global Surveillance of Antituberculosis-Drug Resistance. *N Engl J Med.* 2016 Sep 15;375(11):1081-9.
2. WHO. Global tuberculosis report; 2017.
3. Barletta F, Otero L, de Jong BC, Iwamoto T, Arikawa K, Van der Stuyft P, et al. Predominant *Mycobacterium tuberculosis* Families and High Rates of Recent Transmission among New Cases Are Not Associated with Primary Multidrug Resistance in Lima, Peru. *J Clin Microbiol.* 2015 Jun;53(6):1854-63.
4. Stucki D, Malla B, Hostettler S, Huna T, Feldmann J, Yeboah-Manu D, et al. Two new rapid SNP-typing methods for classifying *Mycobacterium tuberculosis* complex into the main phylogenetic lineages. *PLoS One.* 2012;7(7):e41253.
5. Molina-Moya B, Gomgnimbou MK, Lafoz C, Lacoma A, Prat C, Refregier G, et al. Molecular Characterization of *Mycobacterium tuberculosis* Strains with TB-SPRINT. *Am J Trop Med Hyg.* 2017 Sep;97(3):806-9.
6. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rusch-Gerdes S, Willery E, et al. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. *J Clin Microbiol.* 2006 Dec;44(12):4498-510.
7. Aze J, Sola C, Zhang J, Lafosse-Marin F, Yasmin M, Siddiqui R, et al. Genomics and Machine Learning for Taxonomy Consensus: The *Mycobacterium tuberculosis* Complex Paradigm. *PLoS One.* 2015;10(7):e0130912.
8. Gagneux S. Fitness cost of drug resistance in *Mycobacterium tuberculosis*. *Clin Microbiol Infect.* 2009 Jan;15 Suppl 1:66-8.
9. Casali N, Nikolayevskyy V, Balabanova Y, Harris SR, Ignatyeva O, Kontsevaya I, et al. Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nat Genet.* 2014 Mar;46(3):279-86.
10. Borgdorff MW, van den Hof S, Kalisvaart N, Kremer K, van Soolingen D. Influence of sampling on clustering and associations with risk factors in the molecular epidemiology of tuberculosis. *Am J Epidemiol.* 2011 Jul 15;174(2):243-51.
11. Zammarchi L, Tortoli E, Borroni E, Bartalesi F, Strohmeyer M, Baretta S, et al. High prevalence of clustered tuberculosis cases in peruvian migrants in florence, Italy. *Infect Dis Rep.* 2014 Nov 19;6(4):5646.
12. REPORT. Molecular typing for surveillance of multidrug-resistant tuberculosis in the EU/EEA. SURVEILLANCE REPORT. 2017.
13. Herranz M, Pole I, Ozere I, Chiner-Oms A, Martinez-Lirola M, Perez-Garcia F, et al. *Mycobacterium tuberculosis* Acquires Limited Genetic Diversity in Prolonged Infections, Reactivations and Transmissions Involving Multiple Hosts. *Front Microbiol.* 2017;8:2661.
14. Fiebig L, Kohl TA, Popovici O, Muhlenfeld M, Indra A, Homorodean D, et al. A joint cross-border investigation of a cluster of multidrug-resistant tuberculosis in Austria, Romania and Germany in 2014 using classic, genotyping and whole genome sequencing methods: lessons learnt. *Euro Surveill.* 2017 Jan 12;22(2).

Etude 6 : Analyse de la diversité génétique des souches de MTBc à Beira, Sofala, Mozambique par spoligotyping ; résultats préliminaires sur l'évolution de la tuberculose au Mozambique

Résumé

Au Mozambique, le nombre de nouveaux cas de tuberculose reste toujours élevé avec 552 cas/100 000 habitants (WHO 2017). Peu d'études antérieures avaient décrit la diversité génétique de la tuberculose au Mozambique, il n'y avait par ailleurs pas de données sur la diversité, l'origine, l'évolution de la lignée prédominante dans la région de Sofala, qui constitue un territoire spécifique à l'histoire coloniale riche et une zone clé d'échanges commerciaux historiques en Afrique de l'Est. Beira est une ville portuaire du Mozambique et la capitale de la province de Sofala avec une démographie importante et un riche passé (culture swahili). Cette ville est d'une grande importance pour le Mozambique ainsi que pour certains pays voisins tels que le Malawi, la Zambie et le Zimbabwe qui n'ont pas accès à l'Océan Indien.

L'objectif de ce projet était d'une part d'étudier la diversité génétique des souches de MTBc à Beira, d'autre part d'étudier la distribution génétique spatiale de la tuberculose au Mozambique et dans les pays voisins. Le Sud-Est de l'Afrique est caractérisé par la prédominance de la lignée 1/SIT129 (étude par séquençage du génome entier au Malawi).

Pour cela, nous avons utilisé en première intention la méthode de spoligotypage qui permet d'obtenir rapidement des résultats de génotypage, et nous avons comparé nos résultats avec des données obtenues dans les pays voisins. Les lignées représentées à Beira sont la lignée 4 (46/89 isolats, 51,7%) et la lignée 1 (36/89 isolats, 40,5%) avec la prédominance des sous lignées L1.1.3/EAI9_MOZ et L4.3.4.2.1/LAM11_ZWE. Le spoligotype le plus prévalent est le SIT702 (23/36, 53.5%), sans doute apparenté au SIT129 qui est aussi retrouvé au Malawi et que nous assimilons à la sous lignée L1.1.3 avec EAI9_MOZ (Malawi-Mozambique). Le SIT129 ne diffère en effet du SIT702 que par la perte d'un espaceur. De même, tous les isolats L4.3.4.2.1/LAM11_ZWE trouvés dans cette étude dérivent du SIT59 (32/46, 70%) décrit au Zimbabwe, eux-mêmes descendants de la lignée 4.3/LAM (SIT42 ancestral).

Par ailleurs, pour déterminer l'origine du SIT129, et avoir une visibilité plus claire de la phylogéographie et la circulation génotypique des souches MTBc observées au Mozambique, nous avons construit à l'aide du logiciel QGIS une carte montrant la diversité génétique des lignées L1 à L4 dans 8 des 11 états du Mozambique, et dans les pays voisins : Malawi, Zimbabwe, Zambie, Tanzanie, et au Nord-Est de l'Afrique du Sud. Au Mozambique le SIT129 de la lignée L1 a plus été retrouvé dans les régions côtières et son origine pourrait être liée à des importations historiques de l'Asie du Sud-Est, de Madagascar à travers le commerce Indo-Océanique et la traite des esclaves dans l'Océan Indien. Le Zimbabwe est au cœur de la

propagation de la sous-lignée L4.3.4.2.1/SIT59 (Lignée L4). Quant à la lignée L2, elle progresse sans doute par radiation en provenance de l'Afrique du Sud, tandis que la lignée L3 est prédominante en Tanzanie.

Par cette étude, qui reste à publier, nous avons commencé à décrire la diversité génétique des souches MTBc à Sofala, Mozambique et l'origine et l'évolution possible des lignées/ sous lignées de tuberculose prédominant historiquement au Mozambique, en particulier l'histoire de la lignée 1.

Article 6 :

On the origin of L1/SIT129 *Mycobacterium tuberculosis* complex in Beira, Sofala, and on the phylogeography and molecular evolution of tuberculosis in Mozambique.

Christophe Sola^{1,2*}, Livia Anselmo², Bernice Klotoe¹, Stefan Panaiotov^{1,3}, Cinara Feliciano⁴, Evangelina Inacio Namburete⁵, Emilyn Costa⁶, Josefo Joao Ferro⁵ and Valdes Bollela²

¹Institut de Biologie Intégrative de la Cellule (I2BC), CEA, CNRS, Univ. Paris Sud, Université Paris-Saclay, F-91198 Gif-sur-Yvette, France

²Internal Medicine Department, School of Medicine of Ribeirão Preto, University of São Paulo (FMRP-USP), Ribeirão Preto, Brazil

³National Center of Infectious and Parasitic Diseases, Sofia, Bulgaria

⁴Clinics Hospital, School of Medicine of Ribeirão Preto, University of São Paulo (FMRP-USP), Ribeirão Preto, Brazil

⁵ Universidade Católica de Mozambique, Faculdade de Ciências de Saúde, Medicina Interna, Mozambique

⁶University Federal de Rio de Janeiro, Brazil

* corresponding author : Institut de Biologie Intégrative de la Cellule (I2BC), CEA, CNRS, Univ. Paris Sud, Université Paris-Saclay, F-91198 Gif-sur-Yvette, France ; christophe.sola@i2bc.paris-saclay.fr

Summary:

There is scarce genotyping information on *Mycobacterium tuberculosis* complex (MTBC) strains circulating in the Mozambique. We therefore characterized MTBC clinical isolates that had been collected during a former study in the port of Beira, Sofala, Mozambique, using spoligotyping. A total of 89 different isolates were successfully genotyped. 46 isolates belonged to L4 (51,7%), 36 isolates belonged to L1 (40.5%), 6 isolates belonged to L2 (6,7%), one isolate belonged to L3 (1,1%). Two main genotypes dominated: the sublineage L.1.1.3/EAI9_MOZ/SIT129 (formerly designated as EAI6_BGD) within L1/EAI and the sublineage L.4.3.4.2.1/LAM11_ZWE (also called the South-African 1 family) within L4. Further Comparison of VNTR between L1/SIT129 and L1/SIT702 isolates from various settings, including Mozambique showed minimal VNTR variations in 6 of 24 loci, supporting the close relationships of SIT129/SIT702 isolates and their recent clonal expansion. Spoligotyping results were further compared with available results from neighbouring countries including North-East of South Africa, Zimbabwe, Malawi, Zambia, Tanzania and Madagascar. This comparison indicated that the prevalence of the L1/SIT129 decreases with increasing distance from the Indian Ocean. We discuss human genetics, historical and anthropological facts that could explain the origin of SIT129.

Introduction

Tuberculosis (TB) remains a major health problem in Mozambique, a 28 million people country located in South East-Africa (WHO 2016). Last year a study performed by Namburete *et al.* underlined the problem of MDR-TB in Beira, Mozambique and they characterized the main genetic mutations responsible of first and seconde line resistance drugs (Namburete, Tivane *et al.* 2016). Little genomic data have been produced in South-East Africa, with the exception of South Africa and of the northern Malawi (Karonga district) studies (Guerra-Assuncao, Crampin *et al.* 2015; Manson, Cohen *et al.* 2017, Dheda, 2017 #10861). Three previous studies described a first estimate of *Mycobacterium tuberculosis* complex (MTC) genetic diversity in Mozambique, however none encompassed the Sofala region, a region with a strong historical, economical and anthropological interest (Viegas, Machado *et al.* 2010; Viegas, Machado *et al.* 2013; Viegas, Ghebremichael *et al.* 2015). A recent study characterized the MTC isolates from Beira, though only by characterization of region of deletions (RDs) (Saifodine, Fyfe *et al.* 2014). Results suggested a majority of L1 (61%), 37% of L4, a single isolate of L3, and no L2/Beijing (Saifodine, Fyfe *et al.* 2014). A fine geographical characterization of the L2/Beijing isolates that had been described in 2010 in Mozambique, showed that they were mainly present in the South of the country (Maputo capital and Maputo region) and that there was a statistically significant association between HIV coinfection and L2/Beijing MTC isolates (Viegas, Machado *et al.* 2013). However, this association was not investigated in the other study (Saifodine, Fyfe *et al.* 2014). Hence, there remains discrepant results and a poor characterization of the genetic diversity of MTC in Sofala and other regions of Mozambique. From a clinical perspective, even if drug-susceptibility is the main information for personalized treatment, differential molecular diagnostics between L1-L4 sublineages may become increasingly useful, as more and more evidence are found in clinical expression variation, pathogeny, and drug-resistance according to lineages (Click, Winston *et al.* 2013; Sharma, Bloss *et al.* 2016; Manson, Cohen *et al.* 2017).

Beira is is both an historical and recently (XIXth century) redeveloped port city, 3rd of importance at the demographic level, and of great importance for Mozambique inland and for Malawi, Zambia and Zimbabwe that have no access to the Indian Ocean (Newitt 1995). Beira was historically involved in international trading since the Xth century (Newitt 1995). Demographically, Mozambique is a large but poorly populated country, approximatively 5 times less than Malawi. There is also a geographical and historical continuity between these two young nations, as shown by their pre-colonial history (Newitt 1995).

In this study, our first goal was to characterize MTC genotypes in Beira and compare it to its neighborhood. Our second goal was to start to draw a larger historical and geographical landscape of TB spatial and temporal dynamics in Mozambique and its neighbouring countries in South-East Africa, in particular at the light of major WGS results obtained in Malawi and elsewhere on Lineage 1/SIT129; in Malawi the spread of the Lineage 1/ SIT129 was demonstrated and WGS results were obtained (Guerra-Assuncao, Crampin et al. 2015; Duarte, Nery et al. 2017). In Brazil, similar clinical isolates were also sequenced (Duarte et al., 2017, Costa Conceição, manuscript in preparation). We report here the results on genotyping by conventional genotyping (spoligotyping) on Beira's MTC clinical isolates, and compare it to other MTC populations in bordering countries. We suggest how Lineage 1 may have evolved, and propose some introduction hypothesis based on combined genetic, anthropological and historical sources.

Material and Methods

Patients and Clinical isolates

All patients were adults with cough of more than two weeks of duration, attending health facility in the Sofala Hospital and were recruited in 2015 during a 12 months period. Out of a total of 54 patients for which demographical data were recovered, (men n=28, women n=26) the mean age of the patients was 35 years (men slightly older 37 versus 33). For 44 patients for which the HIV status was known (84%) there were 38 HIV positive patients and 6 HIV negative. 21 HIV positive patients were women and 17 were men. All clinical isolates were isolated from tuberculosis patients and diagnosed in the mycobacteriology reference laboratory of Beira, Sofala Region, Mozambique. MTC clinical isolates were sent after approval of regulative authorities using express carrier means to the Clinical Hospital of Ribeirão Preto, SP, Brazil (Faculty of Medicine of Ribeirão Preto, University of São Paulo) where subcultures were grown and DNA was extracted by the classical Cetyl-trimethyl-ammonium based (CTAB) method and quantified using 260/280 nm spectrophotometry (vanSoolingen, Hermans et al. 1991). Since all these isolates were recruited for clinical diagnostics and DST purpose and since the project had been approved by ethical boards (see below), no ethical consent approval was fulfilled by the patients.

Spoligotyping

CTAB extracted desiccated DNAs were sent by express carrier to the Infection Genetics Emerging Pathogen Evolution (IGEPE) team at the Institute of Integrative Cell Biology, I2BC, UMR9198, CEA-CNRS-Université Paris-Sud, for further molecular characterization by high-throughput microbead-based spoligotyping which was performed as previously described (Zhang, Abadia et al. 2010), using TB-SPOL xMAP®-based Beamedex® kits purchased and run at Beamedex, www.beamedex.com, (Orsay, France).

Nomenclature, Data comparison to databases

All clinical isolates designation was done following the latest whole-genome-based taxonomical designation as found in (Coll, Preston et al. 2014; Stucki, Brites et al. 2016). When discrepant naming or possible confusions may arise, it is mentioned. Because spoligotyping remains a very useful standard within the MTC and sometimes allows a very fine taxonomic and phylogeographical designation, all previous names described in the literature were also kept, in particular those based on the SpolDB4 and SITVITWEB designation (Brudey et al., 2006; Demay et al., 2012). Sometimes this naming may have been

misleading (e.g. SIT129 was named in SITVITWEB as EAI6_BGD based on the absence of spacer 23). In other cases, the name was providing clear origin : e.g. LAM11_ZWE, or SAF1 (South-African 1) and in these cases, the naming was kept (Kamerbeek, Schouls et al. 1997; Brudey, Driscoll et al. 2006; Demay, Liens et al. 2012). To provide a continuity between ancient and new designation we tried to adopt when possible a hierarchical naming of clinical isolates with the following system : Lineage/Sublineage/Spoligotyping-based Lineage naming/SIT designation according to the SITVITWB database numbering (Demay, Liens et al. 2012).

Geographic Information Systems, Map Building

A Mac OSX (OS 10.9.5) Version of QGIS was downloaded and installed from :

<http://www.kyngchaos.com/software/qgis>. African maps with administrative level 1 were downloaded as shapefiles (ESRI format) from : <http://maplibrary.org/library/stacks/Africa/index.htm>

Latitude and Longitude of main cities within the studied states or nationas were downloaded from Google® search using their names and were translated into decimal latitudes and longitudes when needed. The maps shown in Figure 2 to 4 were produced using QGIS after importation of CSV files prepared by adding the latitude and longitude to Excel files, that compiled L1 to L6 genotype frequencies in each studied country and/or region and /or state (cf. Supplementary file 1 for data source details). All files were imported into QGIS and a new project was built ; PNG files were exported as images into Powerpoint for MacOSX (Version 14.7.3). Legends were added by capture of QGIS defined legends. Images were recorded as PDF files and converted using Adobe Photoshop (version 12.0) as High quality TIFF pictures. The colors chosen for L1 to L4 are those initially reproted by Comas et al. 2009. L1 : pink, L2 : blue, L3 : purple, 4 : red.

Ethics

This project was approved by the Ethics and Research Committee of Hospital das Clínicas, School of Medicine of Ribeirão Preto (HC FMRP-USP) -Report: 944117- 01/02/2015. As it involved *M. tuberculosis* isolates from Mozambique the project was also submitted and approved by the National Bioethics Committee for Health of Mozambique (Opinion: Ref: 259 / CNBS / 14 on November 3, 2014). The isolates from Mozambique had been brought to Brazil with funds from the Federal Grant project: 5D43TW001038, in partnership with the University of Pittsburgh-USA.

Results

Population structure, Taxonomic belonging of the isolates, Cluster description

Out of a total of 93 samples analyzed within the Sofala region, we obtained 89 spoligotyping results as four samples gave either no results or hard to interpret results that were discarded. All results are available in **Table 1**. In this table, we inferred the taxonomic status of each isolate, based on congruence between previously published WGS results on a world-wide genome collection, and known spoligotyping signatures (Coll, Preston et al. 2014; Stucki, Brites et al. 2016). The most precise taxonomic characteristic status of each isolate is suggested in column A of the **Table 1**. When no correspondance between WGS and spoligotype signature was available, lineage was inferred from the nearest spoligotype signature prototype as described in SpoIDB4 (Brudey, Driscoll et al. 2006). For one single isolate, we report a potential case of homoplasia in the spoligotyping data (isolate belonging to SIT4), a case that is also found in the literature in Maputo and reported earlier in South Africa (Warren, Streicher et al. 2002; Viegas, Ghebremichael et al. 2015). The taxonomic status of this isolate could be either to L4.4 or to L4.3 and will be solved in the future by WGS or further SNP typing. 46/89 isolates belong to L4 (51,7%), 6 isolates belong to L2 (6,7%), one single isolate belong to L3 (1,1%) and 36 isolates belong to L1 (40,5%). Our results are slightly different from thus by published Saifodine *et al.* that found more L1 (61%) and no Beijing, which emphasizes the importance of sampling duration for any representative study on MTC (one month in Saifodine *et al.*, versus one year in our study). Within the two most predominant sublineages, L1/L1.1.3/EAI9_MOZ and L4/L.4.3.4.2.1/LAM11_ZWE prototypic spoligotyping patterns were easily recognized, respectively SIT129 and SIT59.

A more thorough analysis of the distribution of Beira isolates within the L1, provides the following clustering results: 23 isolates belonged to SIT702, the most prevalent spoligotype (L1.1.3/EAI9_MOZ, 53,5%), a likely offspring of SIT129. Two isolates belonged to a new Cluster (NEW_A), an offspring of SIT702 that was also found in Malawi (see Supplementary Table 2). Five isolates belonged to SIT48 (L1.2.2/EAI1_SOM, 12%). One isolate belonged to SIT1374 and was already described in Malaysia. One isolate belonged to SIT806 that was already described in Mozambique in 1995, previously found in Norway from Vietnamese or Zambian immigrants (n=2), in USA, from Immigrants from Botswana or Cambodia, and in Zimbabwe (n=2). One isolate belong to SIT1467 that was previously described in Malawi, South Africa and Zambia. One isolate belong to SIT1365 that was described previously in Malaysia, Philippines, Singapore and Tanzania. One orphan isolate matched with an isolate

from Malawi (designated as NEW_B), and one was definitively orphan, i.e. unidentified in the SITVITWEB database. All these results points to the strong phylogeographical specificity of spoligotypes found in Mozambique, with strong link to Malawi and the Indian Ocean. All results are shown in **Table 1** and the full comprative results are shown in the **Supplementary Table 2**.

Given the strong phylogeographical specifity of SIT129 (Malawi-Mozambique) (Glynn, Alghamdi et al. 2010), and given previously unpublished results (Costa-Conceicao, personal communication), we thus decided to rename SIT129 and derived, as a new independent sublineage within L1, which we designate as EAI9_MOZ, that is for the time-being likely to be similar to L1.1.3 described in Malawi (Coll, Preston et al. 2014). Prevalence results on SIT129 obtained by Viegas *et al.* in other regions of Mozambique were similar (Viegas, Machado et al. 2010). A more ancient study in Madagascar also described a few SIT129 samples (Ferdinand, Sola et al. 2005).

All L4.3.4.2.1 isolates (South African 1 or SAF1) found in this study are also likely to be derived from the prototypic SIT59 which was first described in Zimbabwe (SIT59, SIT811 and SIT815 signatures) (Heyderman, Goyal et al. 1998; Easterbrook, Gibson et al. 2004; Chihota, Apers et al. 2007). This widespread family of clinical isolates was already described by a South-african research team that pinpoints its major role in the TB Outbreak in Zimbabwe, and neighboring countries (Chihota, Apers et al. 2007).

The six clinical isolates of the L2/Beijing harbored the classical SIT1 signature. Four of these 6 isolates were further investigated to analyze the SNPs the copy number of IS6110 in the NTF region (Plikaytis, Marden et al. 1994). Four of the six L2/beijing isolates were shown to be modern strains, *i.e.* with a single copy of IS6110 in the 3' NTF region using a new proprietary suspension-array based IS6110-NTF typing method (Klotoe *et al* ; in preparation, results not shown). Lastly, the single Lineage 3 Central-Asian (CAS) isolate was found to harbor the SIT794 signature.

Within the Lineage 4 (n=46), the detailed cluster analysis was as follows : L4.3.4.2.1/LAM11_ZWE/SIT59/811/815/1471/2487, n=32 (70%) ; L4.4/S/SIT71, n=3 ; L4.3/LAM9/SIT42, n=3 ; L4_Others, n=7 (this included L4/T1/SIT53, n=2, L4/T2/SIT52, n=1, L4/SIT51, n=1, L4/SIT122var n=1, L4/SIT205 n=1, L4.1.1.3/X/SIT1273 n=1). One isolate could not be assigned between L4.4 and L.4.3 because of the potential homoplasic signature of SIT4.

Molecular Evolution of SIT129 and SIT59 as suggested by spoligotyping

The molecular evolution of the two main families of genotypes that are found in Mozambique, the L4.3.4.2.1/LAM11_ZWE/SIT59/811/815/ and the L1.1.3/EAI6_BGD1/SIT129/702, as suggested by Spoligotyping is shown in **Figure 1**. **Figure 1** suggests that SIT236 is the most likely spoligotyping ancestor that allows to reach, in one or more steps, the SIT129 signature. Only SIT43 is found as the likely step 1 intermediate, in the SITVITWEB database. The phylogeographical specificity of SIT43 is poorly established ; it is found in Saudi Arabia, Malaysia, Pakistan and Somalia and recently in India. Its first description goes back to 1994, to a Swedish immigrant of unknown origin, or to Mexico importation (See Supplementary Table 2). 23 isolates of this genotype only are reported in SITVITWEB. Microevolution of SIT129 allows to obtain SIT702 by single spacer loss, and the new pattern observed in this study (new_A), previously found in Malawi, can be explained by a new spacer loss (See Supplementary table 2 for more details). However another evolutionary pathway remains possible through SIT733, also found in Malaysia (**Figure 1**).

Figure 1 also shows that SIT59, the classical spoligotype signature of the South-African 1 family (L4.3.1.4.2 by Coll *et al.*), is likely to be an offspring of a SIT42 (ancestor spoligotype). It is suggested to have evolved recently in Zimbabwe where it spread within the country (Chihota, Apers et al. 2007) ; in Mozambique too, some specific variants of this type emerged (SIT815 and 811 from one hand, SIT2487 from another hand). The numbers reported on **Figure 1** alongside to each spoligotype structure are those of the distribution of these genotypes, whether in our study (labelled A) or in the Viegas 2010 study (labelled V) or with geographic designation as found in the SITVITWEB database (Demay, Liens et al. 2012). SIT702 differs from SIT129 by the absence of spacer 17. Absence of this spacer could be due either to a deletion of this spacer or to an asymmetric IS6110 Insertion (Filliol, Sola et al. 2000). *In silico* obtention of spoligotypes using WGS will allow to test this hypothesis (Coll, Mallard et al. 2012; Xia, Teo et al. 2016).

Molecular Evolution of SIT129 as suggested by VNTR typing

Further results on SIT129 and SIT702, using Variable Number of Tandem Repeat (VNTR) polymorphism results were extracted from SITVITWEB (Demay, Liens et al. 2012). They were compared to two recently described isolates from Mozambique (EBOV 13-12 and 13-41) (Viegas, Ghebremichael et al. 2015) and to available VNTR results from other SIT129

isolates. Results are presented in **Table 2**. VNTR results show that all these isolates, even if not epidemiologically linked, are evolutionary linked and present minimal VNTR variation, with only 5 VNTR (often one copy variation) polymorphic loci. MIRU39, VNTR1955, VNTR2163b, ETRA, VNTR3690 and VNTR4052 could be used for further VNTR characterization of SIT129 isolates in Mozambique, since they present slight polymorphism whereas all other loci (n=18) seems to be identical within this spoligotyping signature. SIT702 are characterized by a change of VNTR2165/ETRA with a loss of one copy of the repeat (n=8) as shown in two isolates in SITVITWEB, with VNTR2461/ETRB remaining unchanged to 5 copies. Full VNTR data remain however scarce on this genotype and these results should be confirmed at a larger scale. Further studies on SIT129 should try to characterize the 6 above mentioned loci in priority, whereas SIT702 should better be characterized with an extensive 24 loci set.

Phylogeographical specificity of main genotypes found in Beira and comparison within Mozambique and with neighboring countries.

All results from previously published studies in Mozambique were first used to build a region-based genotype frequency Excel file (**Table 3**). Then, a genotypic map showing MTC genotypes distribution in Mozambique was built using the QGIS software. These results are shown in **Figure 2 and 3**. Figure 2 shows the percentage of L1 to L4 in 8 of the 11 administrative states of Mozambique. Figure 3 shows the sublineage distribution for L4. There remains three black spots, i.e. without any data, in three administrative regions of Mozambique : Tete, Zambezia and Manica states. Further studies in these regions should be done to provide basic genotype frequency information. The L4.3.(most of the time the 4.3.4.2.1/LAM11_ZWE) dominates throughout the country and it represents the major L4 sublineage, with anecdotal presence of others L4 sublineages. When looking at the spatial distribution of the second most prevalent type, L1.1.3, the Niassa (hinterland) and Nampula or Sofala (coastal) varies in prevalence, suggesting that L1.1.3. could be more prevalent on the shores. We statistically compared by a X^2 homogeneity test, the percentage of L1 in coastal states and in Niassa (hinterland) and found a significant difference of prevalence ($p=0,01$), suggesting that SIT129 is more prevalent on coastal areas and hence, that the origin of SIT129 may be linked to an historical Indian-Ocean importation from Asia, Madagascar or elsewhere. To further evaluate this hypothesis and study in detail the phylogeography of the genotypes observed in Mozambique, we then reconstructed spatial maps showing the genetic diversity of L1-L4 (L5 to L7 being absent from this part of Africa) in Mozambique neighboring

countries : Malawi, Zimbabwe, Zambia, Tanzania and North-East of South Africa, using previously published results (Easterbrook, Gibson et al. 2004; Chihota, Apers et al. 2007; Glynn, Alghamdi et al. 2010; Said, Kock et al. 2012). The results are shown in **Figure 4**. As shown in this Figure, the geographical gradients of L1-L4 MTC prevalence are more obvious : Zimbabwe is the heart of a monophyletic spread of a Lineage 4 sublineage, Lineage 4.3.1.4.2 (red). L1 spread (pink lineage) is a coastal (shores) phenomenon that decreases when going deeper in the inland, and appears to be at maximum on Mozambique shores. L2 (blue lineage) is emerging with South Africa as a spreading force toward North-East neighboring South-African countries, mainly in large cities (e.g. Beira, Dar es Salaam). L3 (purple lineage) is a predominant lineage in Tanzania and also in the north (Kenya, results not shown).

Discussion

In this paper, we described the genetic diversity of MTC in Sofala, Mozambique. We also increased our knowledge on the MTC genetic diversity and phylogeography in South-East Africa, with Beira, Mozambique, as center of study ; Sofala, an important city today as a harbor, was already a center of early islamic trade more than 1000 years ago (Vernet 2009). Buzurg Ibn Shahriyar of Ramhormuz, was a Muslim traveler, that completed a collection of narratives from Muslim sailors that traveled to India, Malaysia, Indonesia, East Africa and China. He mentioned around 950 that Zanj slaves were caught in the areas between Sofala and Zanzibar to be sold in Oman (Vernet 2009). Zanj refer to the name given by muslim geographers to the East African coast and to the areas inhabited by Bantu's.

We will discuss below two major hypotheses of how Lineage 1/EAI was introduced in Mozambique. From one hand, Indian Ocean trade historical migration from the South or South-East Asian continent seems very likely. We will see that the presence of ancestral genotypes on the maritime route to Mozambique was found at various locations. From another hand, the southward Bantu expansion and migration could also explain, through Malawi, the southward migration of Lineage 1 to Mozambique.

The very designation of EAI when initially described in 2001 was under this paradox: L1 was named « *East-African-Indian* », to create a continental continuity between India and Africa, thus suggesting the potential dual origin of these isolates (Asia or Africa). As soon as naming the L1/EAI, the question of its origin was raised (Sola, Filliol et al. 2001). Gagneux *et al.* named it « *Indian-Ocean* », suggesting a maritim continuity, i.e. an evolutionary link between both continent (Brudey, Driscoll et al. 2006; Gagneux, DeRiemer et al. 2006; Gagneux and Small 2007).

L1 clinical isolates were initially *IS6110* low-copy strains however L1/SIT129 have become high *IS6110* copy number isolates in Malawi, with up to 13 bands for SIT129 (Glynn, Crampin et al. 2008). Such characteristics can be found in up to 15% of L1 (Roychowdhury, Mandal et al. 2015). The existence of an ancient homogeneous *IS6110* low-copy MTC Lineage found both in South-East Asia (Malaysia) in Mascate and in Oman was described as early as 1994 (Fomukong, Tang et al. 1994). The SIT129 molecular evolution that is suggested in this paper (**Figure 1**) may have taken place in Asia or in Africa. This issue could be clarified by WGS that may identify specific SNPs inside L1/SIT129. SNPs number could be used to compute likely divergence time with SIT236 making assumption on clock rates, to

infer an evolutionary time frame to rebuild the SIT129 evolutionary history and to determine the date of its divergence from SIT236, similarly as what we did recently between two genotypes found in Japan and Turkey (Refregier, Abadia et al. 2016).

Even though SIT129 was first mentioned in SpolDB2, and SIT702 in SpolDB3 (Sola, Filliol et al. 2001a ; Filliol, Driscoll et al. 2003) the potential phylogeographical specificity of this type emerged only recently (Glynn, Alghamdi et al. 2010; Viegas, Machado et al. 2010; Viegas, Machado et al. 2013; Coll, Preston et al. 2014; Viegas, Ghebremichael et al. 2015; Costa Conceição, Lopes et al. 2017; Duarte, Nery et al. 2017). 22 isolates of SIT129 and 21 of SIT702 only are found in the SITVITWEB database (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/ query date : Aug 12th 2017). SIT129 was also found in Mozambique, and described in Zimbabwe in 2004 (Easterbrook, Gibson et al. 2004). However it was not present in the original study published in Zimbabwe by Heydermann in 1998, a study that was done on sputum extracts and showed exclusively L4/SIT59 and derived types (Heyderman, Goyal et al. 1998). The fact that sputum conversion is shorter for L1 lineage could explain that, on sputa extracts, L1 was not found in Zimbabwe, even if it might have been present (Click, Winston et al. 2013).

L1.1.3 /SIT129 is not found in the recent paper by Singh et al. describing spoligotypes from all over India (Singh, Sankar et al. 2015). However the likely ancestor type, SIT43, mentioned above was found in Kerala (Joseph, Soman et al. 2013). The main geographical origins of the patients harboring L1.1.3 /SIT129 in SITVITWEB are Malawi, Zimbabwe, Zambia, South Africa, Congo, and French Guiana or Brazil in the Americas (Demay, Liens et al. 2012). Hence, today L1.1.3 /SIT129 distribution looks as being african rather than asian.

One knows more about the recent L1.1.3 /SIT129 and L.4.3.2.1/SIT59 dynamics thanks to sustained studies done in northern Malawi (Glynn, Crampin et al. 2008; Glynn, Alghamdi et al. 2010; Guerra-Assuncao, Crampin et al. 2015) and in Zimbabwe (Easterbrook, Gibson et al. 2004; Chihota, Apers et al. 2007) respectively. In Malawi, a dynamic observation was performed since the end of the 80s. When following the distribution of various MTC clonal complex, during more than 20 years, a steady increase in L1.1.3 /SIT129 could be observed : from one single isolate during the 1986-1991 period, two isolates during the 1992-1996 period, three isolates during the 2001-2005 period, there was a sudden increase (n=38 cases) during the 2006-2008 period (Glynn, Alghamdi et al. 2010). This shows that L1.1.3 /SIT129 was already present in Malawi before its recent spread. The geographical distribution of L1.1.3 /SIT702 as seen in SITVITWEB also points to Malawi and Zambia in Africa, and to Brazil, Cuba and French Guiana in Americas (see Supplementary Table 2). The L1.1.3

/SIT702 isolate that was recently sequenced, originating in the Bahia state in Brazil, shows that the L1.1.3 sublineage is the most frequent L1 sublineage in Brazil (Duarte, Nery et al. 2017). Even if L1 is rare in Brazil, L1.1.3/SIT129 is well established in the Para state, North-East of Brazil (Belem in particular) where a recent extensive spoligotyping-based study found a lot of isolates sharing L1.1.3/SIT129 and/or L1.1.3/SIT702 spoligotypes (Costa Conceição, Lopes et al. 2017). Extensive SNPs analysis and WGS of these isolates has been done and will be reported elsewhere (Costa Conceição in preparation). Extensive ongoing WGS work from all over the world will also allow to infer specific SNPs and the long-term history of the L1 lineage family.

Using a combination of genetic data and geographic information systems, we showed that the L1.1.3/SIT129 and L1.1.3/SIT702 prevalence is likely to be linked to the coastal states in Mozambique since their prevalence is between 40-60%, almost equal to L4 or superior in certain states, whereas it decreases to less than 30% when going further deep into the land (**Figure 1**). Such a geographical association is statistically supported ($p=0,01$ Chi² test). Decreasing gradients of L1 on the North (Tanzania) and on the South (South Africa) of Mozambique are observed, Mozambique shores around Beira being the center of these gradients (**Figure 2**).

Another striking observation is the detection of L1.1.3/SIT129 isolates in Madagascar, however only on the north western shores facing Mozambique (Rasolofo-Razanamparany V, personal communication). Three isolates of the L1.1.3/SIT129 type were found in 2005 and one in 2009 in Boeny, and Sofia. Slave trade between North West Madagascar (Boeny) and East Africa (Sofala) and Arabia is well documented during various epochs, and various kingdoms (Vernet 2009).

A co-evolution hypothesis of *M. tuberculosis-Homo sapiens sapiens*, in relation to demography, migration, genetics drift, could explain the long-lasting presence of L1.1.3 on the Eastern shores of India, Arabia, Madagascar and Africa (Comas, Coscolla et al. 2013). Such a co-evolution process between human beings and MTC is also observed in Western Africa (Asante-Poku, Yeboah-Manu et al. 2015). The MTC population structure could be as diverse, conserved and fine as the *Homo sapiens sapiens* phylogeographical structure and could go back as far as 70,000 years B.P (Comas, Coscolla et al. 2013). Alternatively, a younger age of MTC, in relation to the neolithic transition, could also be possible (Achtman 2016). Two major hypothesis may explain the MTC population structure that we observed in Mozambique : an ancient route that may be linked to Indian Ocean trading history from Asia,

a more recent continental route linked to Bantu's southward expansion. Both dating though remains quite imprecise for the time-being (60,000-5,000 Years b.p).

The *Periplus maris Erythraei* is a greek novel of maritime exploratory routes, that described navigation and trade opportunities from roman-egyptian harbours to red sea shores, East Africa and India (https://en.wikipedia.org/wiki/Periplus_of_the_Erythraean_Sea). *An even more ancient maritime route in the Indian Ocean, that could go back up to 8,000 years b.p is suggested by other authors* (Fernandes, Triska et al. 2015). *The sustained Indian Ocean trade stretching from Java, to Zanzibar and Mombassa can be split into nine periods (Early, Hellenic, Roman, Hindu Buddhist, Muslim, Chinese, Portuguese, Dutch and English, and Japanese)* https://en.wikipedia.org/wiki/Indian_Ocean_trade. India's links with Mozambique goes back over half a millennium. Indian Muslim traders from the South India's Malabar region plied the trade routes of the Indian Ocean bringing them up and down the Eastern coast of Africa (Newitt 1995). The SIT43 profile, which could be ancestral to SIT129 is still found in countries such as Malaysia, Pakistan, Saudi Arabia and importantly was also found at low frequency too in Kerala, (i.e. the Malabar Region), thus conforiting the historical slave trade sources (Joseph, Soman et al. 2013). More recently, during the XVIIIth and XIXth century, Indians from Gujarat migrated to Mozambique and were involved, with the Portuguese, in the transatlantic slave trade https://en.wikipedia.org/wiki/Indians_in_Mozambique . One may thus not discard the hypothesis of a XVIIIth-XIXth recent Gujarati importation of L1.1.3/SIT43 or SIT129. French were also trading slaves between East-Africa and their Indian Ocean islands (Madagascar, Reunion, Mauritius, Comores) during these times (Newitt 1995). Altogether, and until we get the full genome these isolates sequenced (similarly as what was done for similar isolates from Malawi) we have some evidence that maritime routes hypothesis could explain how and when SIT129 (or its ancestor SIT43) reached Africa from Indian sources. The Swahili culture (also part of the Niger-Congo Bantu language family and thus of the Bantu's expansion), before the arrival of Portuguese traders, as early as during the XIIth and XIIIth century was also known to trade slaves (Vernet 2009). The footprints of their trades on the East-African shore is also of great importance (Vernet 2009).

Many archeological sites in Mozambique predates the portuguese arrival, and an ancient continental local co-evolution of L1, parallel as what is observed for mitochondrial DNA haplotype frequencies could be suggested to explain the population structure of MTC in Mozambique and the major prevalence of L1.1.3. If we consider ethnic groups diversity and

demography, some local tribes may indeed have been living isolated on these shores, without being disturbed and without contact with other tribes, before trading ivory, gold and slaves. Keeping in mind the anthropologist perspective that isolated communities of between 180 to 440 people may have pass the tuberculosis disease from one generation to another, i.e. creating endemicity and conditions of coexistence and coevolution ((McGrath 1988), and even if L1 is apparently more diverse in Asia, one should also distinguish the highly different demography between the two continents. Whereas the two active centers of population development in human history have been Europe and Asia, Africa until its recent expansion was not a dense continent, hence, some genotypes may have bene transmitted unchanged from one generation to another during millenias. Sofala was a center with around 10,000 inhabitants during the early XVIth century (Newitt 1995), thus providing TB endemicity likelihood to be maintained at low prevalence.

A continental, southward spread from a Bantu's core thus, cannot be eliminated. Bantu's important demographic expansion to the south can approximately be dated to 500 A.D, however it may have been much more ancient (Silva, Alshamali et al. 2015). Considering the demic diffusion model of the neolithic revolution and recent results obtained in Mozambique on Mitochondrial DNA L2 frequency gradients in African population (in particular considering the Frequency distribution maps for mtDNA haplogroup L2a and L2e, with gradients centered on Mozambique) (Silva, Alshamali et al. 2015), an african origin of L1 should not be excluded. In favour of this hypothesis, some ancestral genotypes of the L1 lineage (SIT236) were found both in Uganda and in Burkina Faso (Coll, Preston et al. 2014). However intermediates types (such as SIT43) are not reported in Africa but only in Asia for the time-being which decreases the likelihood of an ancient southward continental spreading route hypothesis.

1 Mozambique's major ethnic groups are comprising three main Bantu's groups : Makua, north of the Zambezi, Karanga (Kalanga) and Tonga on the south (Newitt 1995). Makua populations were established during the first millenium A.D. and were trading with Swahilis and Gujaratis. They suffered a lot from the growth of Portuguese interests in Brazil and an increasing slave demand https://en.wikipedia.org/wiki/Makua_people. Kalanga populations were suggested to be linked to both L3/SIT21 and L1/SIT129 in a previous study (Glynn, Alghamdi et al. 2010). These people are known to have inhabited south-eastern Zimbabwe, and Mozambique and were trading on the shores since 1000 AD too (https://en.wikipedia.org/wiki/Kalanga_people).

2 Finally another less likely hypothesis of introduction of L1 to Mozambique could also be linked to the portuguese sailors that may have introduced this genotype by progression of their maritime routes during the XIVth century from West Africa till East Africa and India ; indeed, it is also known that L1 strains are found in Guinea-Bissau, another portuguese-speaking country, where L1 represented up to 6% of the isolates in this country, and where one spoligotype of SIT129 was also detected (Groenheit, Ghebremichael et al. 2011).

The purpose of this article was not to investigate the origin of L1.1.3/SIT129 presence in Brazil, however we should also lay some hypothesis on L1.1.3/SIT129/SIT702 introduction to Brazil, since it is likely to be linked to transatlantic slave trade (Costa, in preparation). Portuguese started to trade slaves to Brazil as early as 1508, i.e. eight years only after Brazil discovery (Rout 1976). After the portuguese, the Dutch were already involved in the Atlantic slave trade at the early XVIth (van Welie 2008). They had experience in buying and controlling slaves. They thus imported slaves as the cheapest labour option. Slaves were imported from a variety of places, including the east coast of Africa (Mozambique and Madagascar), but the majority came from East Africa and Asia, especially the Indonesian Islands, which were controlled by the Dutch at the time (van Welie 2008). This explains, for instance, why there is a relatively large number of people of Malaysian descent in the Cape (the so-called Cape Malays). <http://www.sahistory.org.za/topic/slavery-south-africa> accessed on August 11th 2017). Hence, importation of L1.1.3/SIT129 from Java or Indonesia, i.e. the Indonesian peninsula, including Malaysia by Dutch is also possible. L1.1.3/SIT702, for example, is still found in Malaysia (SITVITWEB search on August 10th). However, against this hypothesis, except for the presence of a single L1/SIT733 isolate in Malaysia, no L1.1.3/SIT129 or its likely ancestor L1.1.3/SIT43 were found in Indonesia (Sasmono et al. 2012, Ismail et al. 2014, (Lisdawati, Puspadari et al. 2015).

In the most extensive study performed on the genomic epidemiology of L1.1.3/SIT129 in Malawi, striking differences between the lineages in cluster size, in proportion of disease due to recent transmission and transmissibility in Lineage 1 were shown (Guerra-Assuncao, Crampin et al. 2015). Lineage-1 formed the smallest clusters, with the largest SNP differences. Patients with L1 strains were the least likely to have disease due to recent transmission and were less likely to transmit and cause new cases than those with L2 or L3 (Guerra-Assuncao, Crampin et al. 2015). Taken together, these observations suggest a lower propensity of L1 to cause disease, which may also explain L1's association with HIV

infection. This study provided strong evidence for differences in transmission patterns and virulence between *M. tuberculosis* lineages, corroborating ancient or more recent results in animal models (Bhatia, Csillag et al. 1961 ; Bhatia, Jacob et al. 1963; Narayanan, 2008 #4468; Reiling, Homolka et al. 2013). L1 clinical isolates were also characterized in host-pathogen interaction by low uptake, low cytokine induction, and a low replicative potential in human primary macrophages and in mice infected by the aerosol route (Reiling, Homolka et al. 2013).

To conclude, we provided in this article a precise description of MTC genotypes circulating in Beira, with a strong prevalence of L1.1.3/SIT702 isolates. Most are historically linked to the same specific subtype, L1.1.3/SIT129, also found in Malawi and define Lineage 1.1.3 with EAI9_MOZ (SIT129 as prototype) being part of it. Whether linked to a continental southward advancing force from a Bantu's center, or to an Indian-Ocean based maritime expansion, a more likely explanation, in relation to Madagascar-East Africa-Arabian and Indian shores, L1 in Malawi and Mozambique may also have been co-evolving during centuries if not millenias in this region of Africa. Trying to disentangle of these two hypothesis remains difficult for the time-being. We hope that comparative genomics and improved genetic diversity studies at a finer geographical scale in Mozambique will provide more results to better explain which of the maritime route or the continental spread may explain the preferential presence of L1.1.3/SIT129 on Mozambican shores.

Acknowledgements

Liliana Rutaihwa and Sebastien Gagneux' from the Swiss Tropical Institute are acknowledged for helpful comments on the draft and for taking in charge WGS of some of the isolates described in this paper whose results are to be reported elsewhere. Guislaine Refrégier is acknowledged for biostatistics help in Chi2 analysis. This work was financed by the Federal Grant: 5D43TW001038-15 (REVISED AIDS International Training and Research Program) and by an University of São Paulo-Institut de France « Chaire franco-brésilienne » 6 weeks inviting professor grant attributed to Pr. Christophe Sola. FAEPA: Fundação de Apoio Ao Ensino, Pesquisa e Assistência do HCFMRP-USP is also warmly acknowledged. Christophe Sola is thankful to Campus France, and to Mrs Arianne Megdoud, Institut de France in São Paulo, who allowed this research to be done.

Legends of Figures

Table 2 : L1.1.3/EA19_MOZ/SIT129 Variable Number Tandem Repeat polymorphism found in the literature

	MIRU 02	MIRU 04/ETRD	MIRU 10	MIRU 16	MIRU 20	MIRU 23	MIRU 24	MIRU 26	MIRU 27	MIRU 31/ETRE	MIRU 39	MIRU 40	VNTR 04/24	VNTR 057/ETRC	VNTR 19/55	VNTR 21/33b	VNTR 21/56/ETRA	VNTR 23/47	VNTR 24/01	VNTR 24/61/ETRB	VNTR 31/71	VNTR 36/90	VNTR 40/52	VNTR 41/56	MIRU12	VNTR5
SIT129_K21 (Comas et al. 2009)	2	5	4	3	1	6	2	2	3	6 ^a	4	3	2	4	8	10	9	3	2	5	3	8	3	1		
EBOV 13-12 (Viegas et al. 2015)	2	5	4	3	1	6	2	2	3	5 ^a	3	3	2	4	6	7	7	3	2	5	3	6	6	1		
EBOV 13-41 (Viegas et al. 2015)	2	5	4	3	1	6	2	2	3	5 ^a	3	3	2	4	7	10	2 ^a	3	2	5	3	6	7	1		
GUF0X1996076 (SITVITWEB)	2	5	4	3	1	6	2	2	3	5 ^a	4	3		4			9								254316223543	95465 ^d
ESP0620040MRC-04015937 (ibid)	2	5	4	3	1	6	2	2	3	5 ^a	3	3													254316223533	
GUF0X20030102 (ibid)		5								5 ^a				4			9			5						95465 ^d
NLD000016700 (ibid)	2	5	4	3	1	6	2	2	3	6 ^a	4	3		4		9				5					254316223643	95466 ^d
ZAF0920020RXH329 (ibid)	2	5	4	3	2 ^a	6	1 ^b	2	3	6 ^a	3	3													254326123633	
Variant					+/-		+/-				x			x	x		x					x	x			
Invariant	x	x	x	x	x	x	x	x	x	x	x	x	x	x				x	x	x	x			x		x

^a: doubtful result

^b: wrong result

^c: difference may be due to bioinformatic

treatment see also note ^d

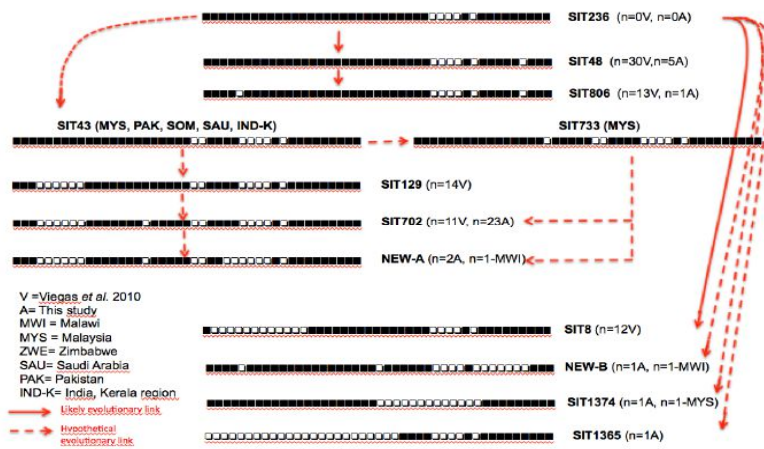
^d: see Sola et al. 2003

Table 3: Genetic studies describing MTC diversity in Mozambique

Number of Samples	Year of recruitment	Methods	Region of Sampling	Reference	Results
445	2007-2008	spoligotyping RT_-PCR on Region of Deletions, VNTR typing (7 or 9 VNTR respectively for L1 and L4) plus 22 isolates by 12 VNTR	all but Sofala, Tete, Zambezia, Manica	Viegas et al. 2010	see also Figure 1
67	2009		Beira (Sofala)	Saifodine et al. 2014	61% L1; 1,7% L3; 37,3% L4;
43	2014	spoligotyping-VNTR	Maputo	Viegas et al. 2015	44%L1; 11,6%L2; 2,3%L3; 41,8%L4
89	2016	spoligotyping	Beira (Sofala)	This study	40% L1; 51,6% L4; 6,7% L2; 1,1% L3

601

Lineage 1.1.3



Lineage 4.3.4.2

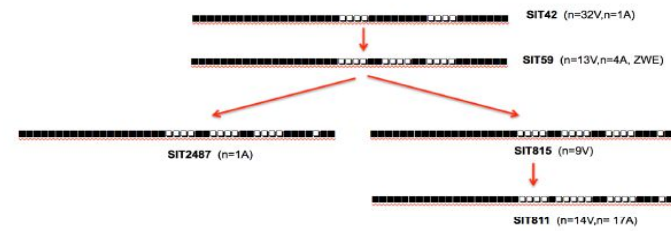


Figure 1 : Spoligotyping-based likely evolution of the two main MTC sublineages found in this study : L1.1.3 (left) and L4.3.1.2 (right)

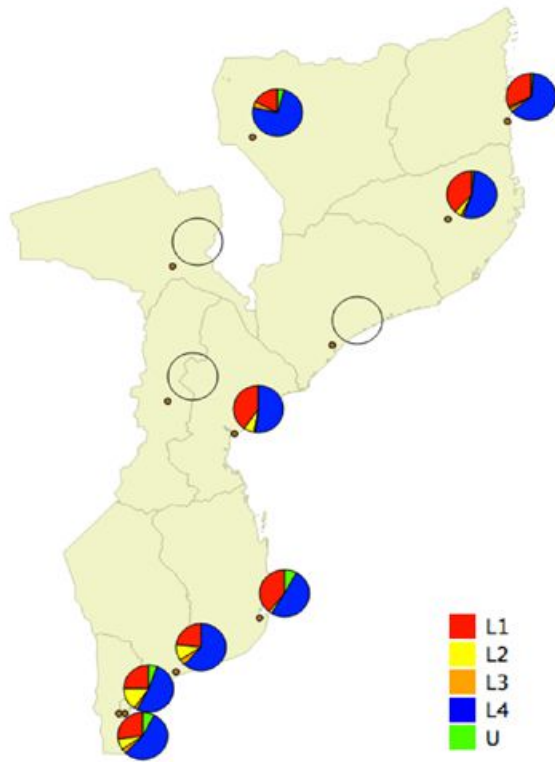


Figure 2 : QGIS-built maps showing Lineage 1-4 (plus Unknown) distribution in Mozambique.

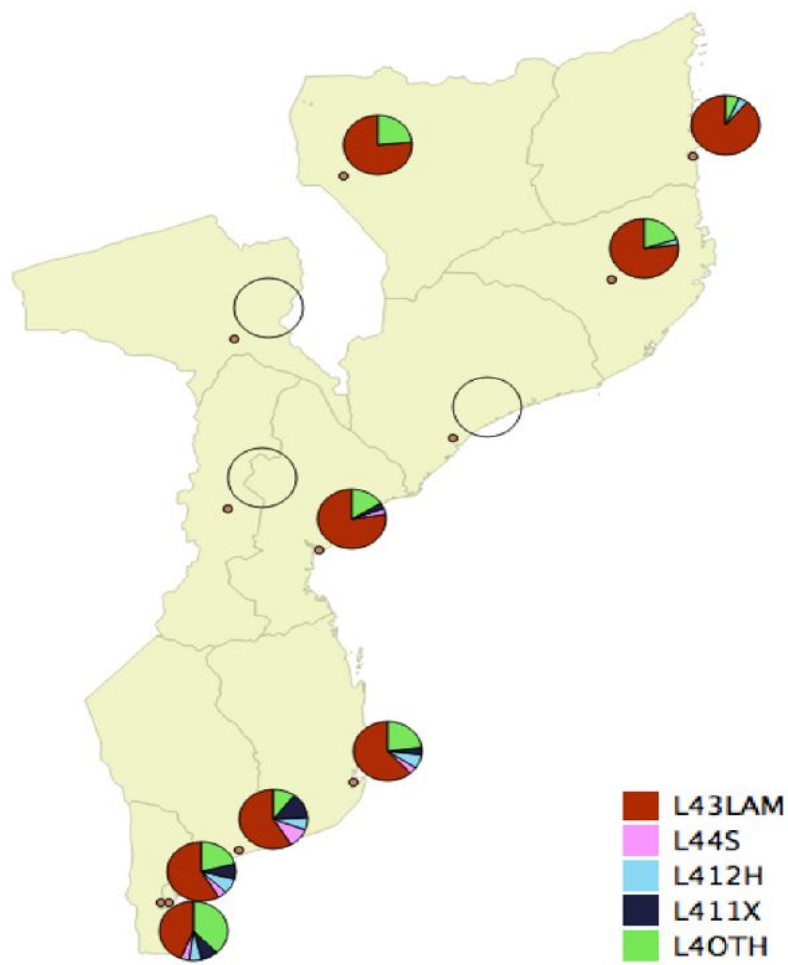


Figure 3 : QGIS-built maps showing within L4 sublineages distribution (right).

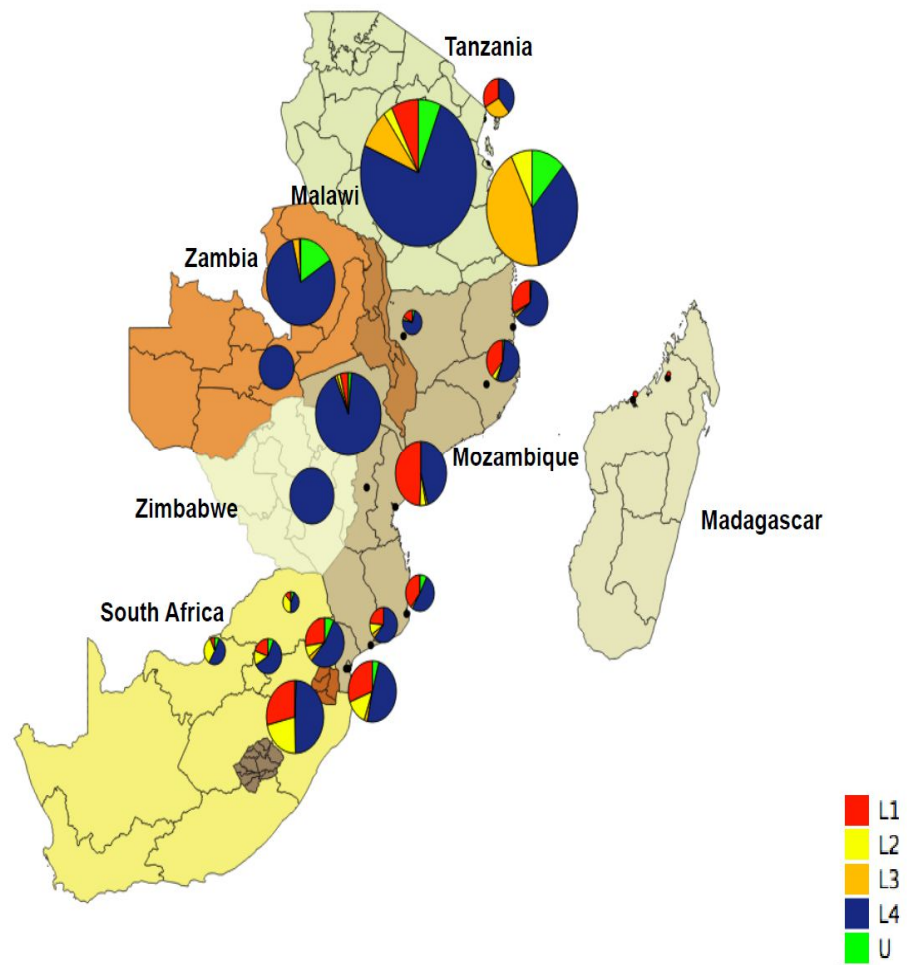


Figure 4: QGIS-built maps showing Lineage 1-4 (plus Unknown) distribution in Mozambique and its borders. Diameter is correlated to sampling size. Madagascar data are missing except for two L.1.3. Isolates on northwestern shore as discussed in text.

References

1. Achtman, M., 2016. How old are bacterial pathogens? *Proceedings. Biological sciences / The Royal Society* 283.
2. Asante-Poku, A., Yeboah-Manu, D., Otchere, I.D., Aboagye, S.Y., Stucki, D., Hattendorf, J., Borrell, S., Feldmann, J., Danso, E., Gagneux, S., 2015. *Mycobacterium africanum* is associated with patient ethnicity in Ghana. *PLoS neglected tropical diseases* 9, e3370.
3. Bhatia, A.L., Csillag, A., Mitchison, D.A., Selkon, J.B., Somasundaram, P.R., Subbaiah, T.V., 1961. The virulence in the guinea-pig of tubercle bacilli isolated before treatment from South Indian patients with pulmonary tuberculosis. 2. Comparison with virulence of tubercle bacilli from British patients. *Bulletin of the World Health Organization* 25, 313-322.
4. Bhatia, A.L., Jacob, C.V., Hitze, K.L., Ramachandran, K., Selkon, J.B., 1963. A Comparison of the Virulence in the Guinea-Pig of Tubercle Bacilli from Thai and British Patients. *Bulletin of the World Health Organization* 29, 483-490.
5. Brudey, K., Driscoll, J., Rigouts, L., Prodinger, W.M., Gori, A., Al-Hajj, S.A.M., Allix, C., Aristimuno, L., Arora, J., Baumanis, V., Binder, L., Cafrune, P., Cataldi, A., Cheong, S., Diel, R., Ellermeier, C., Evans, J.T., Fauville-Dufaux, M., Ferdinand, S., Garcia de Viedma, D., Garzelli, C., Gazzola, L., Gomes, H.M., Gutierrez, M.C., Hawkey, P.M., van Helden, P.D., Kadival, G.V., Kreiswirth, B.N., Kremer, K., Kubin, M., Kulkarni, S.P., Liens, B., Lillebaek, T., Ly, H.M., Martin, C., Martin, C., Mokrousov, I., Narvskaja, O., Ngeow, Y.F., Naumann, L., Niemann, S., Parwati, I., Rahim, M.Z., Rasolofonandrianana, V., Rasolonavalona, T., Rossetti, M.L., Rüsck-Gerdes, S., Sajduda, A., Samper, S., Shemyakin, I.G., Singh, U.B., Somoskovi, A., Skuce, R., Van Soolingen, D., Streicher, E.M., Suffys, P.N., Tortoli, E., Tracevska, T., Vincent, V., Victor, T.C., Warren, R., Yap, S.F., Zaman, K., Portaels, F., Rastogi, N., Sola, C., 2006. *Mycobacterium tuberculosis* complex genetic diversity : mining the fourth international spoligotyping database (SpolDB4) for classification, Population Genetics, and Epidemiology. *BMC Microbiol.* 6, 23.
6. Chihota, V., Apers, L., Mungofa, S., Kasongo, W., Nyoni, I.M., Tembwe, R., Mbulo, G., Tembo, M., Streicher, E.M., van der Spuy, G.D., Victor, T.C., van Helden, P., Warren, R.M., 2007. Predominance of a single genotype of *Mycobacterium tuberculosis* in regions of Southern Africa. *Int J Tuberc Lung Dis* 11, 311-318.
7. Click, E.S., Winston, C.A., Oeltmann, J.E., Moonan, P.K., Mac Kenzie, W.R., 2013. Association between *Mycobacterium tuberculosis* lineage and time to sputum culture conversion. *Int J Tuberc Lung Dis* 17, 878-884.
8. Coll, F., Mallard, K., Preston, M.D., Bentley, S., Parkhill, J., McNERNEY, R., Martin, N., Clark, T.G., 2012. SpolPred: rapid and accurate prediction of *Mycobacterium tuberculosis* spoligotypes from short genomic sequences. *Bioinformatics* 28, 2991-2993.
9. Coll, F., Preston, M., Guerra-Assuncao, J.A., Hill-Cawthorn, G., Harris, D., Perdigo, J., Viveiros, M., Portugal, I., Drobniowski, F., Gagneux, S., Glynn, J.R., Pain, A., Parkhill, J., McNERNEY, R., Martin, N., Clark, T.G., 2014. PolyTB: A genomic variation map for *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* 94(3):346-54, 346-354.
10. Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K.E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., Thwaites, G., Yeboah-Manu, D., Bothamley, G., Mei, J., Wei, L., Bentley, S., Harris, S.R., Niemann, S., Diel, R., Aseffa, A., Gao, Q., Young, D., Gagneux, S., 2013. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nature genetics*. 45(10):1176-82.
11. Costa Conceição, E., Lopes, M.L., Rastogi, N., Couvin, D., Furlaneto, I.P., Gomes, H.M., Vasconcelos, S.E.G., Suffys, P.N., Schneider, M.P.C., Sousa, M.S.d., Sola, C., de Paula

- Souza e Guimarães R.J., Duarte, R.S., Lima, K.V.B., 2017. Genetic diversity of *Mycobacterium tuberculosis* from Pará, Brazil, reveals a higher frequency of the lineage East African India in South America. *Infection Genetics Evolution* submitted.
12. Demay, C., Liens, B., Burguière, T., Hill, V., Couvin, D., Millet, J., Mokrousov, I., Sola, C., Zozio, T., Rastogi, N., 2012. SITVITWEB – A publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. *Infect Genet Evol* 12, 755-766.
 13. Duarte, T.A., Nery, J.S., Boechat, N., Pereira, S.M., Simonsen, V., Oliveira, M., Gomes, M.G., Penha-Goncalves, C., Barreto, M.L., Barbosa, T., 2017. A systematic review of East African-Indian family of *Mycobacterium tuberculosis* in Brazil. *Braz J Infect Dis*.
 14. Easterbrook, P.J., Gibson, A., Murad, S., Lamprecht, D., Ives, N., Ferguson, A., Lowe, O., Mason, P., Ndudzo, A., Taziwa, A., Makombe, R., Mbengeranwa, L., Sola, C., Rastogi, N., Drobniowski, F., 2004. High rates of clustering of strains causing tuberculosis in Harare, Zimbabwe: a molecular epidemiological study. *J Clin Microbiol* 42, 4536-4544.
 15. Ferdinand, S., Sola, C., Chanteau, S., Ramarokoto, H., Rasolonavalona, T., Rasolof-Razanamparany, V., Rastogi, N., 2005. A study of spoligotyping-defined *Mycobacterium tuberculosis* clades in relation to the origin of peopling and the demographic history in Madagascar. *Infect Genet Evol* 5, 340-348.
 16. Fernandes, V., Triska, P., Pereira, J.B., Alshamali, F., Rito, T., Machado, A., Fajkosova, Z., Cavadas, B., Cerny, V., Soares, P., Richards, M.B., Pereira, L., 2015. Genetic stratigraphy of key demographic events in Arabia. *PLoS One* 10, e0118625.
 17. Filliol, I., Driscoll, J.R., van Soolingen, D., Kreiswirth, B.N., Kremer, K., Valetudie, G., Dang, D.A., Barlow, R., Banerjee, D., Bifani, P.J., Brudey, K., Cataldi, A., Cooksey, R.C., Cousins, D.V., Dale, J.W., Dellagostin, O.A., Drobniowski, F., Engelmann, G., Ferdinand, S., Gascoyne-Binzi, D., Gordon, M., Gutierrez, M.C., Haas, W.H., Heersma, H., Kassa-Kelembho, E., Ho, M.L., Makristathis, A., Mammina, C., Martin, G., Mostrom, P., Mokrousov, I., Narbonne, V., Narvskaya, O., Nastasi, A., Niobe-Eyangoh, S.N., Pape, J.W., Rasolof-Razanamparany, V., Ridell, M., Rossetti, M.L., Stauffer, F., Suffys, P.N., Takiff, H., Texier-Maugein, J., Vincent, V., de Waard, J.H., Sola, C., Rastogi, N., 2003. Snapshot of moving and expanding clones of *Mycobacterium tuberculosis* and their global distribution assessed by spoligotyping in an international study. *J Clin Microbiol* 41, 1963-1970.
 18. Filliol, I., Sola, C., Rastogi, N., 2000. Detection of a previously unamplified spacer within the DR locus of *Mycobacterium tuberculosis*: Epidemiological implications. *J. Clin. Microbiol.* 38, 1231-1234.
 19. Fomukong, N.G., Tang, T.H., Al-Maamary, S., Ibrahim, W.A., Ramayah, S., Yates, M., Zainuddin, Z.F., Dale, J.W., 1994. Insertion sequence typing of *Mycobacterium tuberculosis*: characterization of a widespread subtype with a single copy of IS6110. *Tubercle and Lung Disease* 75, 435-440.
 20. Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B.C., Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M.C., Hilty, M., Hopewell, P.C., Small, P.M., 2006. Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America* 103, 2869-2873.
 21. Gagneux, S., Small, P.M., 2007. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *The Lancet. Infectious diseases* 7, 328-337.
 22. Glynn, J.R., Alghamdi, S., Mallard, K., McNERney, R., Ndlovu, R., Munthali, L., Houben, R.M., Fine, P.E., French, N., Crampin, A.C., 2010. Changes in *Mycobacterium*

- tuberculosis* genotype families over 20 years in a population-based study in Northern Malawi. PLoS One 5, e12259.
23. Glynn, J.R., Crampin, A.C., Traore, H., Chaguluka, S., Mwafulirwa, D.T., Alghamdi, S., Ngwira, B.M., Yates, M.D., Drobniewski, F.D., Fine, P.E., 2008. Determinants of cluster size in large, population-based molecular epidemiology study of tuberculosis, northern Malawi. *Emerg Infect Dis* 14, 1060-1066.
 24. Groenheit, R., Ghebremichael, S., Svensson, J., Rabna, P., Colombatti, R., Riccardi, F., Couvin, D., Hill, V., Rastogi, N., Koivula, T., Kallenius, G., 2011. The Guinea-Bissau Family of *Mycobacterium tuberculosis* Complex Revisited. PLoS One 6, e18601.
 25. Guerra-Assuncao, J.A., Crampin, A.C., Houben, R.M., Mzembe, T., Mallard, K., Coll, F., Khan, P., Banda, L., Chiwaya, A., Pereira, R.P., McNERney, R., Fine, P.E., Parkhill, J., Clark, T.G., Glynn, J.R., 2015. Large-scale whole genome sequencing of *M. tuberculosis* provides insights into transmission in a high prevalence area. *eLife* 4.
 26. Heyderman, R.S., Goyal, M., Roberts, P., Ushewokunze, S., Zizhou, S., Marshall, B.G., Makombe, R., Van Embden, J.D., Mason, P.R., Shaw, R.J., 1998. Pulmonary tuberculosis in Harare, Zimbabwe: analysis by spoligotyping. *Thorax* 53, 346-350.
 27. Joseph, B.V., Soman, S., Radhakrishnan, I., Hill, V., Dhanasooraj, D., Kumar, R.A., Rastogi, N., Mundayoor, S., 2013. Molecular epidemiology of *Mycobacterium tuberculosis* isolates from Kerala, India using IS6110-RFLP, spoligotyping and MIRU-VNTRs. *Infect Genet Evol* 16, 157-164.
 28. Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., Bunschoten, A., Molhuizen, H., Shaw, R., Goyal, M., van Embden, J., 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* 35, 907-914.
 29. Lisdawati, V., Puspadari, N., Rifati, L., Soekarno, T., M, M., K, S., Ratnasari, L., Izzatun, N., Parwati, I., 2015. Molecular epidemiology study of *Mycobacterium tuberculosis* and its susceptibility to anti-tuberculosis drugs in Indonesia. *BMC Infect Dis* 15, 366.
 30. Manson, A.L., Cohen, K.A., Abeel, T., Desjardins, C.A., Armstrong, D.T., Barry, C.E., 3rd, Brand, J., Consortium, T.B.G.G., Chapman, S.B., Cho, S.N., Gabrielian, A., Gomez, J., Jodals, A.M., Joloba, M., Jureen, P., Lee, J.S., Malinga, L., Maiga, M., Nordenberg, D., Noroc, E., Romancenco, E., Salazar, A., Ssengooba, W., Velayati, A.A., Winglee, K., Zalutskaya, A., Via, L.E., Cassell, G.H., Dorman, S.E., Ellner, J., Farnia, P., Galagan, J.E., Rosenthal, A., Crudu, V., Homorodean, D., Hsueh, P.R., Narayanan, S., Pym, A.S., Skrahina, A., Swaminathan, S., Van der Walt, M., Alland, D., Bishai, W.R., Cohen, T., Hoffner, S., Birren, B.W., Earl, A.M., 2017. Genomic analysis of globally diverse *Mycobacterium tuberculosis* strains provides insights into the emergence and spread of multidrug resistance. *Nature genetics* 49, 395-402.
 31. McGrath, J.W., 1988. Social networks of disease spread in the lower Illinois valley: a simulation approach. *American journal of physical anthropology* 77, 483-496.
 32. Namburete, E.I., Tivane, I., Lisboa, M., Passeri, M., Pocente, R., Ferro, J.J., Harrison, L.H., Bollela, V.R., 2016. Drug-resistant tuberculosis in Central Mozambique: the role of a rapid genotypic susceptibility testing. *BMC Infect Dis* 16, 423.
 33. Narayanan, S., Gagneux, S., Hari, L., Tsolaki, A.G., Rajasekhar, S., Narayanan, P.R., Small, P.M., Holmes, S., Deriemer, K., 2008. Genomic interrogation of ancestral *Mycobacterium tuberculosis* from south India. *Infect Genet Evol* 8, 474-483.
 34. Newitt, M., 1995. A history of Mozambique. Indiana University Press.
 35. Plikaytis, B.B., Marden, J.L., Crawford, J.T., Woodley, C.L., Butler, W.R., Shinnick, T.M., 1994. Multiplex PCR assay specific for the multidrug-resistant strain W of *Mycobacterium tuberculosis*. *J Clin Microbiol* 32, 1542-1546.

36. Refregier, G., Abadia, E., Matsumoto, T., Ano, H., Takashima, T., Tsuyuguchi, I., Aktas, E., Comert, F., Gomgnimbou, M.K., Panaiotov, S., Phelan, J., Coll, F., McNerney, R., Pain, A., Clark, T.G., Sola, C., 2016. Turkish and Japanese *Mycobacterium tuberculosis* sublineages share a remote common ancestor. *Infect Genet Evol.* 45:461-473.
37. Reiling, N., Homolka, S., Walter, K., Brandenburg, J., Niwinski, L., Ernst, M., Herzmann, C., Lange, C., Diel, R., Ehlers, S., Niemann, S., 2013. Clade-specific virulence patterns of *Mycobacterium tuberculosis* complex strains in human primary macrophages and aerogenically infected mice. *mBio* 4, 4.
38. Rout, L.B.J., 1976. Race and Slavery in Brazil. *The Wilson Quarterly* 1, 73-89.
39. Roychowdhury, T., Mandal, S., Bhattacharya, A., 2015. Analysis of IS6110 insertion sites provide a glimpse into genome evolution of *Mycobacterium tuberculosis*. *Scientific reports* 5, 12567.
40. Said, H.M., Kock, M.M., Ismail, N.A., Mphahlele, M., Baba, K., Omar, S.V., Osman, A.G., Hoosen, A.A., Ehlers, M.M., 2012. Molecular characterization and second-line antituberculosis drug resistance patterns of multidrug-resistant *Mycobacterium tuberculosis* isolates from the northern region of South Africa. *J Clin Microbiol* 50, 2857-2862.
41. Saifodine, A., Fyfe, J., Sievers, A., Coelho, E., Azam, K., Black, J., 2014. Genetic diversity of *Mycobacterium tuberculosis* isolates obtained from patients with pulmonary tuberculosis in Beira city, Mozambique. *Int J Mycobacteriol* 3, 94-100.
42. Sharma, A., Bloss, E., Heilig, C.M., Click, E.S., 2016. Tuberculosis Caused by *Mycobacterium africanum*, United States, 2004-2013. *Emerg Infect Dis* 22, 396-403.
43. Silva, M., Alshamali, F., Silva, P., Carrilho, C., Mandlate, F., Jesus Trovoada, M., Cerny, V., Pereira, L., Soares, P., 2015. 60,000 years of interactions between Central and Eastern Africa documented by major African mitochondrial haplogroup L2. *Scientific reports* 5, 12526.
44. Singh, J., Sankar, M.M., Kumar, P., Couvin, D., Rastogi, N., Singh, S., Indian, T.B.D.N., 2015. Genetic diversity and drug susceptibility profile of *Mycobacterium tuberculosis* isolated from different regions of India. *The Journal of infection* 71, 207-219.
45. Sola, C., Filliol, I., Gutierrez, M.C., Mokrousov, I., Vincent, V., Rastogi, N., 2001a. Spoligotype database of *Mycobacterium tuberculosis*: biogeographic distribution of shared types and epidemiologic and phylogenetic perspectives. *Emerg Infect Dis* 7, 390-396.
46. Sola, C., Filliol, I., Legrand, E., Mokrousov, I., Rastogi, N., 2001. *Mycobacterium tuberculosis* phylogeny reconstruction based on combined numerical analysis with IS1081, IS6110, VNTR, and DR-based spoligotyping suggests the existence of two new phylogeographical clades. *Journal of molecular evolution* 53, 680-689.
47. Stucki, D., Brites, D., Jeljeli, L., Coscolla, M., Liu, Q., Trauner, A., Fenner, L., Rutaihwa, L., Borrell, S., Luo, T., Gao, Q., Kato-Maeda, M., Ballif, M., Egger, M., Macedo, R., Mardassi, H., Moreno, M., Vilanova, G.T., Fyfe, J., Globan, M., Thomas, J., Jamieson, F., Guthrie, J.L., Asante-Poku, A., Yeboah-Manu, D., Wampande, E., Ssengooba, W., Joloba, M., Boom, W.H., Basu, I., Bower, J., Saraiva, M., Vasconcellos, S.E., Suffys, P., Koch, A., Wilkinson, R., Gail-Bekker, L., Malla, B., Ley, S.D., Beck, H.P., de Jong, B.C., Toit, K., Sanchez-Padilla, E., Bonnet, M., Gil-Brusola, A., Frank, M., Penlap Beng, V.N., Eisenach, K., Alani, I., Ndung'u, P.W., Revathi, G., Gehre, F., Akter, S., Ntoumi, F., Stewart-Isherwood, L., Ntinginya, N.E., Rachow, A., Hoelscher, M., Cirillo, D.M., Skenders, G., Hoffner, S., Bakonyte, D., Stakenas, P., Diel, R., Crudu, V., Moldovan, O., Al-Hajoj, S., Otero, L., Barletta, F., Carter, E.J., Diero, L., Supply, P., Comas, I., Niemann, S., Gagneux, S., 2016. *Mycobacterium tuberculosis* lineage 4 comprises

- globally distributed and geographically restricted sublineages. *Nature genetics*. 48, 1535-1543.
48. van Welie, R., 2008. Slave trading and slavery in the Dutch colonial empire : a global comparison. *New West Indian Guide* 82, 47-96.
 49. vanSoolingen, D., Hermans, P.W.M., Haas, P.E.W.d., Sool, D.R., Embden, J.D.A.v., 1991. The occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. *J. Clin. Microbiol.* 29, 2578-2586.
 50. Vernet, T., 2009. Slave trade and slavery on the Swahili coast (1500-1750), in: B.A. Mirzai, Montana, I.M., Lovejoy, P. (Eds.), *Slavery, Islam and Diaspora*. Africa World Press, pp. pp.37-76.
 51. Viegas, S.O., Ghebremichael, S., Massawo, L., Alberto, M., Fernandes, F.C., Monteiro, E., Couvin, D., Matavele, J.M., Rastogi, N., Correia-Neves, M., Machado, A., Carrilho, C., Groenheit, R., Kallenius, G., Koivula, T., 2015. *Mycobacterium tuberculosis* causing tuberculous lymphadenitis in Maputo, Mozambique. *BMC microbiology* 15, 268.
 52. Viegas, S.O., Machado, A., Groenheit, R., Ghebremichael, S., Pennhag, A., Gudo, P.S., Cuna, Z., Langa, E., Miotto, P., Cirillo, D.M., Rastogi, N., Warren, R.M., van Helden, P.D., Koivula, T., Kallenius, G., 2013. *Mycobacterium tuberculosis* Beijing genotype is associated with HIV infection in Mozambique. *PLoS One* 8, e71999.
 53. Viegas, S.O., Machado, A., Groenheit, R., Ghebremichael, S., Pennhag, A., Gudo, P.S., Cuna, Z., Miotto, P., Hill, V., Marrufo, T., Cirillo, D.M., Rastogi, N., Kallenius, G., Koivula, T., 2010. Molecular diversity of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in Mozambique. *BMC microbiology* 10, 195.
 54. Warren, R.M., Streicher, E.M., Sampson, S.L., van der Spuy, G.D., Richardson, M., Nguyen, D., Behr, M.A., Victor, T.C., van Helden, P.D., 2002. Microevolution of the direct repeat region of *Mycobacterium tuberculosis*: implications for interpretation of spoligotyping data. *J Clin Microbiol* 40, 4457-4465.
 55. WHO, 2016. *Global Tuberculosis Report 2016*. WHO, Geneva, Switzerland.
 56. Xia, E., Teo, Y.Y., Ong, R.T., 2016. SpoTyping: fast and accurate in silico *Mycobacterium* spoligotyping from sequence reads. *Genome medicine* 8, 19.
 57. Zhang, J., Abadia, E., Refregier, G., Tafaj, S., Boschiroli, M.L., Guillard, B., Andremont, A., Ruimy, R., Sola, C., 2010. *Mycobacterium tuberculosis* complex CRISPR genotyping: improving efficiency, throughput and discriminative power of 'spoligotyping' with new spacers and a microbead-based hybridization assay. *J Med Microbiol* 59, 285-294.

**Etude 7 : Analyse de la diversité des souches de
Mycobacterium tuberculosis dans la région de
Balimo en Papouasie Nouvelle Guinée par la
technique du Spoligotypage**

(Article en préparation)

Résumé

La tuberculose est un sérieux problème de santé publique en Nouvelle Guinée Papouasie avec une estimation de 30000 nouveaux cas et 3800 morts/an soit 432 cas/100 000 habitants (WHO 2015). Et la proportion de la tuberculose multi-résistante est estimée à 3% dans les nouveaux cas (WHO 2017).

Peu d'études ont caractérisé la diversité génétique des souches de *Mycobacterium tuberculosis* dans ce pays, et certaines provinces n'ont pas été du tout échantillonnées.

L'objectif de cette étude était dans un premier temps d'analyser la diversité génétique des souches de *M. tuberculosis*, précisément la lignée dominante dans la région ouest Balimo en Nouvelle Guinée Papouasie. Ce travail a été réalisé par spoligotypage sur Luminex. Dans un second temps, nous souhaitons identifier les clusters spoligotypes prédominants et leur répartition dans l'ensemble du pays, en vue de déterminer les zones de fortes transmissions. Pour cela, nous avons complété nos données de spoligotypage avec des données épidémiologiques (VNTR, SNP, résistance aux anti-tuberculeux) issues d'autres études.

L'étude a été faite sur une collection de 231 isolats de *M. tuberculosis* diagnostiqués à l'hôpital de la ville de Balimo chez des patients provenant de diverses villes de la région de Balimo. Ces ADNs ont été extraits de lames 3+/, de culture ou de sang. Après analyse, selon les profils non ambigus, les lignées prédominantes dans cette région étaient la lignée L2 (Beijing) et la lignée L4 (European-American), avec des pourcentages respectifs de 56% et 43%. Une seule souche de lignée 1 a été identifiée (1%). Il est à noter que le pourcentage de profils ambigus était très élevé dans cette étude : 23 à 70% selon la rigueur des seuils de significativité appliqués. Ces profils peuvent être dus soit à des infections mixtes (80% de ces profils présentent une absence claire des espaceurs 33 et 34 comme dans le cas d'une superposition des profils de L2 et L4) ou à des contaminations post-prélèvements, ou à une qualité d'ADN insuffisante.

Dans une étude antérieure, la prévalence des lignées 2 et 4 variait significativement entre les régions. Par exemple, dans Madang (capitale de la province Madang) et ses alentours et Wewak (capitale de la province du Sépik oriental), la lignée L4 était la plus prévalente (76,6% - 80%, Ballif et al, 2012) tandis que vers Goroka (capitale de la province des Hautes terres orientales) et Alotau (capitale de la province de baie de Milne), la lignée L2 était dominante avec un taux de 60,5% et 84,4% respectivement (Ley et al 2014). Par le taux de 56% de L2 relevé, notre étude montre que la diversité des souches de tuberculosis de Balimo est comparable à celle de Goroka (**Fig 1**).

La surveillance de la maladie de *M. tuberculosis* peut être améliorée par la connaissance des géotypes et leurs dynamiques de transmission.

L'article proposé est une ébauche qui reste à documenter.

Article 7 :

Structure de la population du complexe *Mycobacterium tuberculosis* dans la region de Balimo, Nouvelle Guinée Papuasie par la technique de spoligotypage sur le luminex²⁰⁰: TB-Spol

Vanina Guernier ^{a,*}, Tanya Diefenbach-Elstob^{a,b}, Bernice Klotoe^c, Guislaine Refregier^c,
Christophe Sola^c, Daniel Pelowa^c, Emma S. McBryde^{a,d}, Jeffrey Warner^{a,b}

^aAustralian Institute of Tropical Health and Medicine, James Cook University,
Townsville, Queensland, Australia

^bCollege of Public Health, Medical and Veterinary Sciences, James Cook University,
Townsville, Queensland, Australia

^cBalimo District Hospital, Balimo, Western Province, Papua New Guinea

^dDepartment of Medicine, Royal Melbourne Hospital, University of Melbourne,
Melbourne, Victoria, Australia

^eInstitute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud,
Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France

Abstract

Background

Tuberculosis (TB) is a serious health problem in Papua New Guinea (PNG) with an estimated 30000 new cases and 3800 deaths each year. The epidemiological studies on tuberculosis are rare in all regions of country and the recent study by *Ballif M.* and al 2012 in Madang region revealed that, the most prevalent lineage are the Euro-American lineage (Lineage 4) and the East-Asian lineage (Lineage 2). The study aimed to characterize the genetic diversity of *Mycobacterium tuberculosis* in Balimo region, the Western Province of Papua New Guinea (PNG) and confront our results to the spoligotype profiles reported in database SITVITWEB in neighbouring regions.

Material and Methods

M. tuberculosis was isolated from patients in Balimo hospital. 231 isolates were extracted by 10% chelex from sputum smear slides (lam+/3+, ZN stained slides) collected from the Balimo region (2012-2014), from cultures and from blood samples. On these samples, we performed spoligotyping on Luminex technology.

Results

The 231 *M. tuberculosis* isolates collected belonged predominantly to the European-American (L4) and to East-Asian lineage (L2) at the percentages of 43% and 56% respectively with 70% of mixed infection and undetermined (Fig 1). The reported and collected patterns of spoligotype from SITVITWEB, revealed that although the lineage L2 and L4 are predominant, these percentage differ in the regions of Papua New Guinea.

Conclusion

The obtained data from this study provide the molecular characteristic of *M. tuberculosis* in the Balimo city (Balimo District) of Papua New Guinea, and suggest substantial mixed infection.

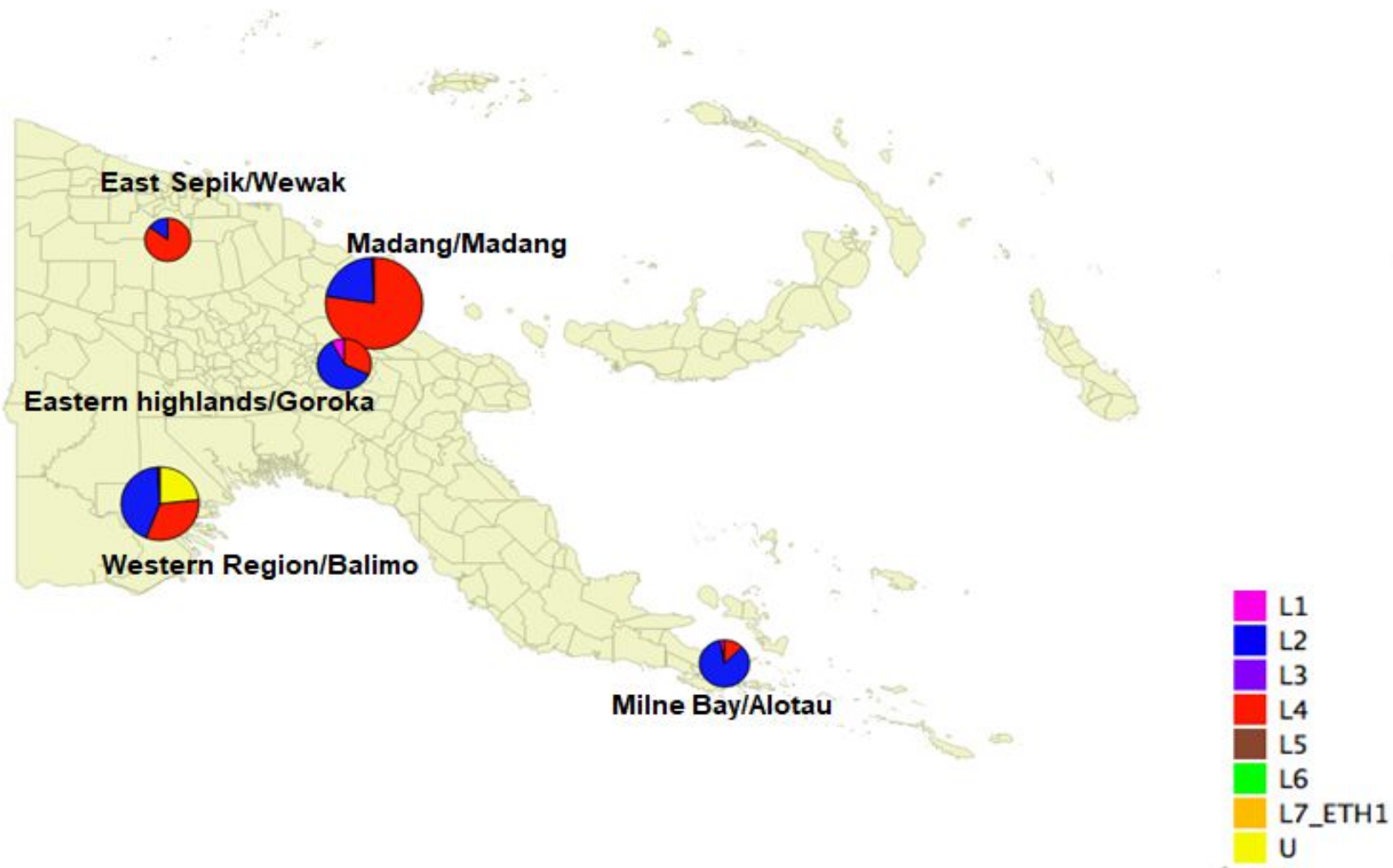


Figure 1: Carte de la Nouvelle Guinée Papouasie représentant la prevalence des lignées de *M. tuberculosis*, L1- EAI, L2 – Beijing, L3 – CAS, L4 – Euro American, L5 – Africanum 1, L6 – Africanum 2, L7 – Ethiopie, U- Infection Mixte/indéterminé.

**Etude 8 : Epidémiologie moléculaire de
Mycobacterium tuberculosis en Italie en 2015:
état actuel de la tuberculose**

(Article en préparation)

Résumé

L'épidémiologie moléculaire de *M. tuberculosis* est complexe et a connu un rapide et profond changement à Palermo (Italie) entre 1994 et 2000 avec la prédominance des sous-types Haarlem, T, LAM, S et l'apparition des sous-lignées CAM, U, EAI, CAS, Tur, X, Bovis, Ural (C. Bonura et al 2014).

L'objectif de cette étude était de faire une enquête d'épidémiologie moléculaire de *M. tuberculosis* sur la diversité génétique des souches de bacilles tuberculeux isolés dans les diverses localités de Sicile (Palermo, Ragusa, Catania, Messina), afin d'identifier des éventuelles grappes de bacilles identiques, suggestives de cas de transmission récente. Le second but du projet est de faire une étude retrospective en comparant les données épidémiologiques actuelles à celles obtenues en 2014. Cette enquête a été faite sur une collection de 200 souches de *Mycobacterium tuberculosis* au cours de l'année 2015, prélevées chez des patients Italiens (n = 57) et non Italiens (n = 113). Ces souches sont extraites par le kit Qiagen et analysées par la méthode à haut débit de typage moléculaire sur microbilles (TB-SPOL) permettant d'envisager une approche globale du contrôle de la tuberculose. Nous avons identifié 7 différentes lignées et 11 clades correspondants dont les plus prévalents étaient T (29,5%), Haarlem (14,7%), LAM (10,7%) et autres clades moins dominants dont Beijing (L2), CAS1 et CAS-Dehli (L3), indo-oceanic (EAI, L1), Africanum 1 (L5) et 2 (L6). En comparaison avec les données épidémiologiques obtenues en 2014, on note une inversion des proportions de souches entre les clades T et Harleem de la lignée Euro-American à Palerme. Il convient donc une mise en place de méthodes épidémiologiques dans les laboratoires des villes d'Italie en vue d'une surveillance constante et rapprochée des souches MTBc.

L'article qui suit est une ébauche qui reste à documenter

Article 8 :

Molecular diversity of *Mycobacterium tuberculosis* in Italy in 2015

Celestino Bonura¹, Bernice J. KLOTUE², Xavier Odessa², Guislaine Refrégier²,
Christophe Sola², and Caterina Mammina^{1,3}

¹ Department of Sciences for Health Promotion and Mother-Child Care "G. D'Alessandro",
University of Palermo, Palermo, Italy. celestino.bonura@unipa.it.

² CNRS-Université Paris-Sud, Institut de Génétique et Microbiologie, UMR8621, Infection
Genetics Emerging Pathogen Evolution Team, Orsay, France.

³ Department of Sciences for Health Promotion and Mother-Child Care "G. D'Alessandro",
University of Palermo, Palermo, Italy. caterina.mammina@unipa.it.

Abstract

Background

The present study aimed to investigate the molecular epidemiology of *M. tuberculosis* complex (MTBc) isolates in the province of Palermo, Ragusa, Catania, Messina, (Sicily, Italy), by characterizing 200 isolates identified in 2015. A comparison with 183 MTBC strains identified in 2014 in the geographic area of Palermo was also carried out.

Material and Methods

Two hundred MTBC isolates identified in the four main cities of Sicily in the 2015 (Palermo n = 80; Ragusa, n = 29; Catania, n = 47; Messina, n = 17) were analyzed by spoligotyping on Luminex technology platform. Furthermore, the spoligotyping dataset obtained from 183 MTBC isolates identified in 2014 in Palermo City was reanalyzed. Distributions into lineages and clustering of isolates in the two periods were compared.

Results

One hundred seventy-three out of the 200 isolates of MTBC submitted to molecular typing were fully characterized. Of these isolates, 57 were from Italian-born and 113 from foreign-born individuals. Seven different lineages and 11 subclades were identified with the most represented being T (29.5%), Haarlem (14.7%), LAM (10.7%) from Lineage 4 (**Fig 1**). Other clades were Beijing (L2), CAS1 and CAS-Dehli (L3), indo-oceanic (EAI, L1), Africanum 1 (L5) et 2 (L6). T clade was overrepresented in foreign-born patients. In Messina, the LAM clade was not found.

When the results of this study were compared to MTB_C isolates identified in 2014 in Palermo City, the proportion of Haarlem clade decreased (26.5% in 2014 and 11.25% in 2015) (**Fig 2**).

Conclusion: We observed a large geographical heterogeneity and genetical diversity among the MTBc strains isolated in 2015. Comparison between distributions of lineages in the two consecutive periods depicts rapid and deep changes in the TB epidemiology in Palermo, Italy. A universal and continued laboratory-based surveillance of TB in Sicily is required.

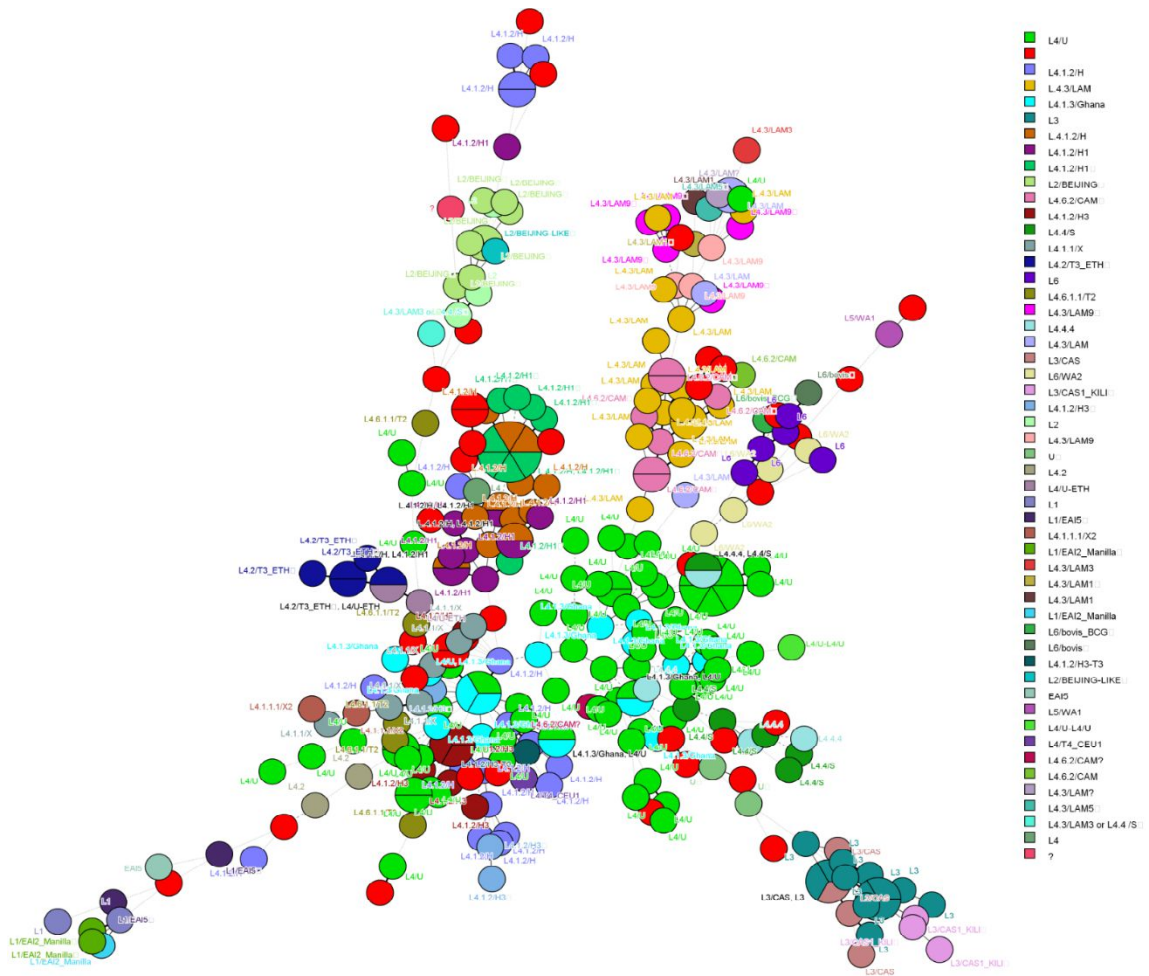


Figure 1: Diversité des souches de MTBc obtenue en 2015

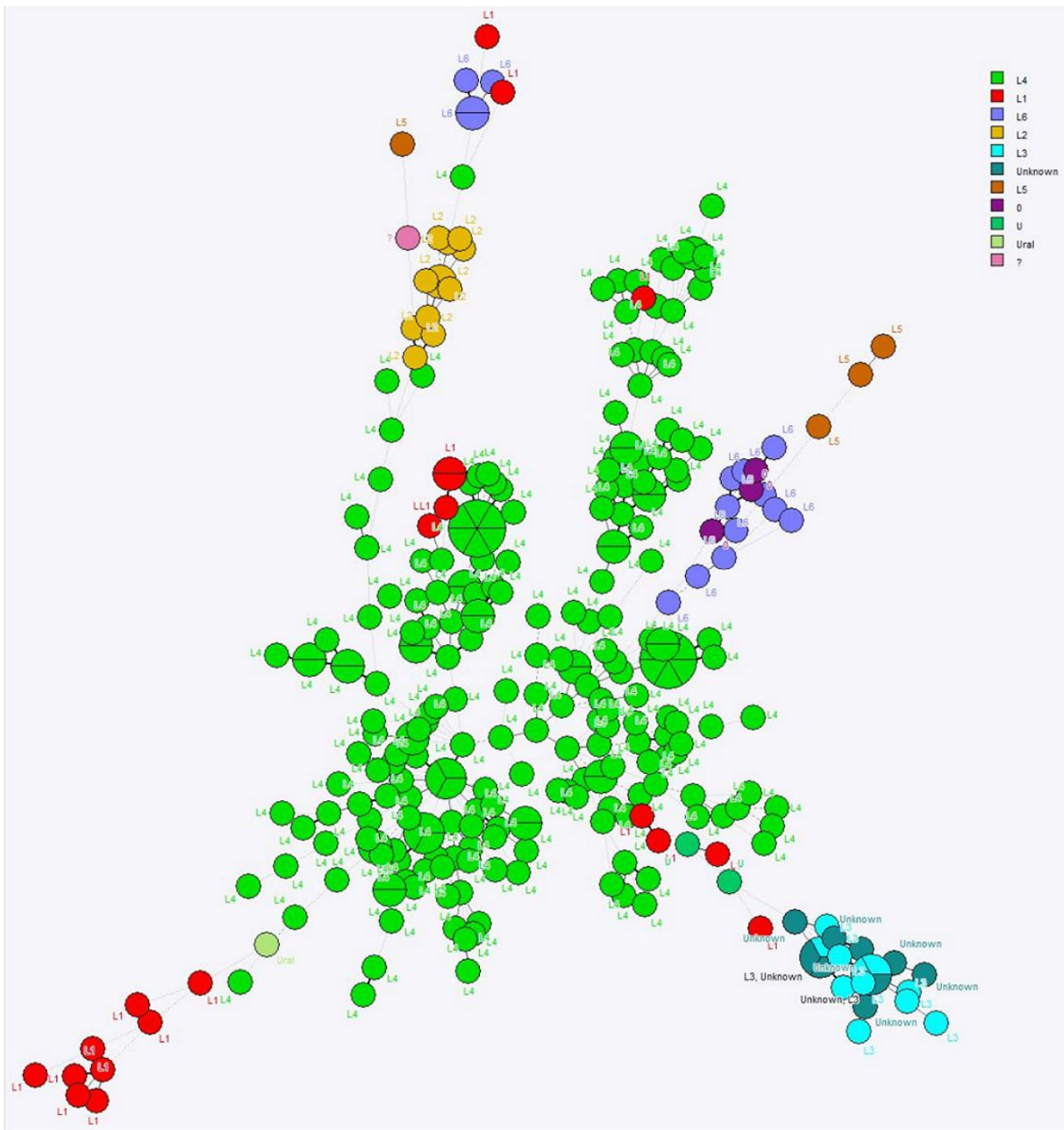


Figure 2 : Comparaison des profils spoligotypes des isolats de MTBc de 2015 et 2014.

Chapitre 5 : Discussion Générale et Perspectives

Discussion Générale et Perspectives

Faire face à la tuberculose sensible, multi-résistante (MDR-TB), ultra-résistante (XDR-TB) et aux épidémies à souche de génotype Beijing est un défi mondial. Dans ce contexte, nous avons développé deux tests de diagnostic avec deux types d'objectifs : 1) identification de la résistance génotypique aux antibiotiques, 2) identification de profils génétiques assez précis pour contribuer à la surveillance et la maîtrise des souches Beijing résistantes et épidémiques. Pour la détection rapide de l'ultra-résistance des souches de *Mycobacterium tuberculosis* nous avons développé le test TB-EFI, et pour l'identification et le suivi des souches Beijing MDR et épidémique, nous avons mis au point le test IS-NTF/RIT.

Le TB-EFI est basé sur la détection de mutations liées à des résistances ; il prédit la résistance phénotypique des souches de tuberculose (résistance génotypique) aux antibiotiques de deuxième ligne (les fluoroquinolones, les aminoglycosides dont la Kanamycine, l'Amikacine, un polypeptide injectable, la Capréomycine) et un antibiotique de première ligne, l'Ethambutol.

Le IS-NTF/RIT étudie la diversité des souches Beijing, une famille de souches de *Mycobacterium tuberculosis*, la plus virulente et répandue, responsable d'épidémie dans diverses régions. Le test est basé sur le typage génotypique de la séquence d'insertion IS6110 dans le locus NTF, et la détection de résistance de ces souches à la Rifampicine et à l'Isoniazide. Tous les tests ont été développés sur la technologie Luminex200®, après hybridation de l'amplicon aux sondes spécifiques couplées à des microbilles.

Parallèlement, nous avons prolongé l'analyse d'un panel de génomes Beijing des bases de données publiques et plus particulièrement les SNP qui les distinguent. Nous proposons une nouvelle combinaison de SNPs de caractérisation de Beijing qui : 1) est plus facilement analysable que les combinaisons proposées précédemment du fait de leur proximité par paires sur le génome et 2) permet une discrimination plus fine des épidémies les plus fréquentes. Cette combinaison comprend 17 SNP spécifiques de 17 sous-groupes importants. Ces SNPs pourraient à terme être combinés aux régions de délétions et à la séquence d'insertion dans le locus NTF pour confirmer et/ou affiner cette combinaison. L'objectif final est de proposer un test fiable et peu coûteux en vue du diagnostic et de la transmission des souches L2/Beijing au sein des populations.

Nous avons également caractérisé en détail des souches MDR et XDR au Kazakhstan par des méthodes haut-débit (MLPA, TB-SPRINT et PCR-RFLP) et ainsi à partir de leurs marqueurs moléculaires spécifiques, déterminé la transmission prédominante du clone 94-32

de la lignée L2/Beijing. Par la technique du Spoligotypage, nous avons aussi des études sur l'épidémiologie moléculaire de la tuberculose en Nouvelle Guinée Papouasie, en Italie, au Mozambique, la transmission active de la tuberculose multi-résistante au Pérou et son impact sur les pays de l'Europe.

Cette partie résume les résultats obtenus et discute de l'applicabilité, l'importance, les perspectives futures des différentes méthodes développées dans l'épidémiologie moléculaire, la surveillance et le contrôle de la tuberculose, un problème de santé publique.

5.1 La détection des résistances aux antibiotiques de 2^{de} ligne

La connaissance et l'identification des mutations au sein du gène cible des antituberculeux contre *Mycobacterium tuberculosis*, favorisent une rapide prise en charge adaptée au patient, ce qui est important pour une guérison complète et évite la sélection d'autres souches mutants. Les tests et méthodes de génomiques appliquées sont des solutions qui fournissent des informations génétiques pour chaque isolat clinique et ainsi facilitent sa détermination dans un nombre d'échantillons issus de patients. L'aspect important dans le développement des méthodes moléculaires est leur capacité à être concordant/discordant avec les méthodes existantes et classiques. En se basant sur des analyses statistiques, la concordance du TB-EFI par rapport au pyroséquençage est 100% (n = 61 souches, 1098/1098 données) pour les mutations ciblées pour la résistance à la Fluoroquinolone, aux Injectables et de 98% pour l'Ethambutol. TB-EFI a été développé sur la technologie Luminex et sur des extraits d'ADN de type thermolysats. Le pyroséquençage est une technique de séquençage ciblé rapide, haut débit, et à moindre coût que le séquençage classique de Sanger. Le TB-EFI et le pyroséquençage sont deux techniques, utilisant la PCR, en multiplex pour le TB-EFI et en simplex pour le pyroséquençage. Dans les deux techniques, plusieurs échantillons peuvent être analysés en série sur des plaques de 96 puits, mais par l'analyse simultanée de plusieurs loci, le TB-EFI est de plus haut débit que le pyroséquençage. L'interprétation des résultats est aussi différente, mais aucune des deux ne requiert des analyses bio-informatiques. Le TB-EFI a démontré sa capacité à donner des résultats de même qualité que le pyroséquençage. A ce jour, le séquençage du génome entier (Whole Genome Sequencing, WGS) reste la méthode ultime pour la détection précise et simultanée des profils de résistance aux antituberculeux. Mais elle reste plus de 10 fois plus coûteuse, avec une complexité supérieure dans l'analyse des résultats. Le WGS n'est pas accessible à tous les laboratoires surtout dans les pays à faibles revenus, d'une part à cause de son coût économique et d'autre part à cause de l'analyse des résultats qui nécessite une expertise en Bio-informatique.

Le taux de concordance de la résistance déduite du génotype par TB-EFI et du phénotype obtenu par DST est de 72 à 98% pour tous les antituberculeux ciblés, avec un taux inférieur de 72% et 77% respectivement pour l’Ethambutol et la Fluoroquinolone. Cela se justifie par le fait qu’on n’a pas ciblé toutes les mutations présentes dans la collection de souches analysées.

Les outils d’analyse de TB-EFI permettent la prédiction du phénotype des souches pour chaque antibiotique. Pour les injectables, ces outils aident à prévoir la force de la résistance (intermédiaire ou forte). En compilant les prédictions, TB-EFI a aussi la capacité de diagnostiquer la tuberculose XDR. Il peut éventuellement révéler les profils phénotypiques des isolats qui ont été ratés par le test de sensibilité phénotypique (DST). Mais en aucun cas, le TB-EFI ne peut remplacer la méthode conventionnelle DST. Le DST est nécessaire pour l’accès précis au degré (fort/faible) de résistance d’une souche par rapport à un antibiotique afin d’initier, d’augmenter la dose ou exclure l’antibiotique du traitement.

Le TB-EFI a été testé avec succès sur des souches MDR du Kazakhstan (n = 43), du Brésil et du Pakistan, deux pays où le taux de tuberculose multirésistante n’est pas négligeable et moindre au Brésil, avec une prédiction phénotypique de 51 à 98%. Et dans cette collection de souches du Kazakhstan, du Brésil et du Pakistan, le TB-EFI a détecté 9 XDR, 1 XDR, 1XDR respectivement.

Ainsi, par la simplicité d’interprétation et son haut-débit, le TB-EFI est un test adapté à l’étude régulière de l’état de résistance aux antituberculeux de seconde ligne. Son utilisation dans les pays à forte prévalence de MDR/pré-XDR/XDR permettrait d’y avoir de temps en temps une photo sur l’évolution de la situation. Il peut aussi contribuer à l’épidémiologie moléculaire en identifiant si les souches potentiellement transmises entre individus traités portent bien les mêmes mutations de résistance.

La validation du TB-EFI sera effective et totale lorsqu’on fera des tests sur un nombre plus important de souches et dans divers contextes géographiques, précisément des souches MDR venant de la Chine, la Russie et l’Inde où le taux de TB XDR est important. Par ailleurs, comme le développement et la validation de méthode ont été faits sur des ADN thermolysats, il est important de le tester sur des ADN extrait d’échantillons cliniques.

5.2 [Avantage et inconvénients du test IS-NTF/RINT et MLPA-Beijing](#)

Le test IS*6110*/RINT étudie la séquence d’insertion IS*6110* au sein du locus NTF en vue de génotyper donc avoir des informations pour suivre l’évolution, l’expansion, la transmission

des souches Beijing. Le groupe Beijing constitue le groupe de souches de *Mycobacterium tuberculosis*, le plus répandu et un groupe épidémiologique important. Initialement, les souches Beijing étaient structurées en Anciennes/Atypiques et Modernes/Typiques Beijing en fonction du nombre de copies de la séquence d'insertion au sein du locus NTF. Les souches anciennes n'ont aucune copie d'IS6110 au sein du locus tandis que les souches modernes Beijing ont 1 IS6110 en 3' du locus NTF. Ainsi, le test IS6110-NTF/RINT permet non seulement d'identifier l'espèce *Mycobacterium tuberculosis* par la présence de cette séquence, mais aussi le type de Beijing et le profil de résistance à la Rifampicine et à l'Isoniazide. Les souches incluses dans le développement et la validation du test sont issues des USA (n = 47) et du Kazakhstan (n = 216). D'après cette collection de souche, le sous-groupe « *moderne Beijing* » est surreprésenté par rapport aux « *ancien Beijing* » (89 ; 4%). Le test NTF/RINT a pu détecter les souches Beijing modernes dont la souche « W » responsable de l'épidémie de tuberculose aux USA dans les années 90. Nous avons identifié les trois types de génotype Beijing regroupés en *modern et ancien Beijing*. Par ailleurs le génotype ancien Beijing est plus prévalent au Japon (Iwamoto et al 2009), ce qui s'explique par le fait de la migration des populations du Sud de la Chine. Par l'analyse génotypique des souches Beijing du Pérou, du Mozambique, de la Bulgarie, du Pakistan et de la Turquie avec le test IS-NTF/RINT, nous avons documenté l'expansion mondiale du génotype moderne Beijing. Certes les génotypes anciens/modernes Beijing sont présents partout, mais la prédominance de l'un par rapport à l'autre varie d'une région ou d'un pays donné, suggérant une récente évolution ou une coévolution hôte-pathogène permettant une rapide émergence. Le typage des souches Beijing en se basant sur la séquence d'insertion IS6110 dans le locus NTF permet ainsi de différencier les sous-groupes de génotype Beijing et produit un aperçu de la structure globale de cette population de souches.

En fonction de la région géographique et du régime thérapeutique, les souches de la lignée L2/Beijing sont souvent associées à la multi-résistance aux antituberculeux de première ligne dont la Rifampicine et l'Isoniazide. La combinaison de la détection de résistance des souches à la Rifampicine/l'Isoniazide (mutation en position 445 et 450 dans le gène *rpoB*, position 315 en *katG*, et -15 / -8 *inhA* dans le promoteur *eis*), au typage génomique en fonction de la séquence d'insertion IS6110 en NTF, dans le test IS-NTF/RINT est très importante pour leur différenciation en sous-lignées/clones et ainsi permet de comprendre l'évolution et la transmission de ce génotype. Grâce au test, nous avons pu identifier 6 clones où le plus fréquent et dominant est le clade MDR moderne Beijing avec 62% (163/263) de souches

définis par 1 IS**6110** :: NTF/ mut rpoB-450TTG/ mut katG-315C, ou 2 IS**6110** :: NTF/ mut rpoB-450TTG/ mut katG-315C.

La sensibilité du test IS-NTF/RINT est de 100% et sa spécificité de 100% sur tous les locus typés. La concordance du test par rapport à la détection sur gel d'électrophorèse est de 100%. La différence entre les deux techniques est que la technique haut-débit fournit des résultats numérisés. Cependant, le test permet juste d'identifier les grands sous-groupes de la famille L2/Beijing (Ancien/Moderne) et le profil de résistance. Il ne contient pas de marqueurs permettant une classification génétique fine, il n'utilise donc pas toute la finesse de la taxonomie (la phylogénie évolutive et la répartition de ces souches en sous-groupe selon des marqueurs moléculaires spécifiques).

Lorsqu'il s'agit de comprendre l'évolution et la phylogénie des souches Beijing de *Mycobacterium tuberculosis*, les marqueurs moléculaires les plus pertinents et performants sont les SNPs et les régions de délétion RD. Sur cet aspect, plusieurs classifications et arbres phylogénétiques reconstituant l'histoire évolutive des souches Beijing ont été proposés par différents auteurs. Nous avons initié un travail de recherche de corrélation entre ces classifications moléculaires. Parallèlement à nos travaux, un groupe de recherche russe a proposé une classification cohérente en analysant les génotypes de 1 398 génomes Beijing issus de séquençages de génomes entiers. Cette nouvelle classification subdivise la lignée L2 en proto-Beijing et Beijing. Le groupe L2/Beijing au sens strict est divisé en groupes *ancient Beijing et modern Beijing*. Ces groupes correspondent respectivement à trois (Asia Ancestral 1, Asia Ancestral 2, Asia Ancestral 3) et sept groupes (Asian African 1, Asian Africa 2, Asian Africa 2/RD142, Asian Africa 3, Pacific RD150, Europe/Russia B0/W148 outbreak et Central Asia) (Shitikov et al). En se basant sur cette nouvelle classification et avec les données de génomes et de SNP, nous avons identifié des SNPs rapprochées sur le génome, spécifiques pour chaque groupe identifié, ainsi que des SNPs définissant la lignée L2, ainsi que les groupes de toutes les souches épidémiques permettant et la différenciation des « ancient Beijing » des « modern ». Nous disposons donc de 17 SNPs (Tab. 2, projet 3) spécifiques avec une sensibilité de 100%. Nous avons identifié des marqueurs qui pourraient être ajoutés à cette combinaison : 3 SNPs spécifiques de Coll *et al* 2014. Servant de contrôle pour 3 SNPs de notre combinaison, les marqueurs identifiant IS**6110** dans le locus NTF et ceux correspondant aux régions de délétion les plus utilisées pour décrire la diversité des souches : RD105, RD207, RD181, RD142, RD150. Au final, nous retenons 31 marqueurs moléculaires sur lesquels nous prévoyons de développer un test de discrimination des souches Beijing en utilisant le principe de la « Multiple Ligation-mediated Probe Amplification ». Nous

prévoyons de nommer ce test, le « MLPA-Beijing ». Pour cela nous avons dessiné les sondes pour identifier les marqueurs, séquencé des ADNs d'isolats de diverses régions en vue de disposer de contrôles positifs. Ce travail est encore en cours.

Notons que le nouveau test en développement, aidera à mieux classer les souches Beijing, à détecter la transmission des clones épidémiques dans les régions à forte prévalence, et à suivre l'évolution de cette famille dans le temps.

Donc pour une classification globale des souche Beijing en *Modern/Ancient*, et une identification rapide des souches MDR, nous recommandons le test IS-NTF/RIT, mais pour une taxonomie fine de ces souches, la MLPA-Beijing sera une meilleure option. Le séquençage WGS reste la référence.

5.3 Marqueurs des souches épidémiques

Les souches Beijing sont responsables au moins huit épidémies tuberculeuses majeures reportées depuis les années 1990. Ces épidémies de tuberculose multi résistantes ont été causées par la souche « W » (à New York dans les années 90), les souches épidémiques B0/W148 et A0/94-32 circulant dans les pays de l'ex URSS et d'Asie (réparties en clade B, A et CAO, clades différents mais issus d'un même ancêtre), la souche GC1237 (épidémie à Gran Canaria), la souche K (épidémie en Corée du Sud), la souche 210 et HN878 (à New York) (Merker et al 2015, Casali et al 2014, Han et al 2015, Sreevatsan et al 1997, Yang et al 1998) (**Tab 3**). Ces souches épidémiques appartiennent aux sous-groupes Asia Ancestral 1 (Strain K), Asia ancestral 3 (Strain GC1237), Asian Africa2/RD142 (Strain 210 et HN878), Europe/Russia B0/W148 (Souche W148), CAO (souche spécifique en Asie Central), Clade A (souche du clone 94-32). Les SNPs spécifiques de ces clades ont été identifiées, elles pourront aider à traquer les différentes épidémies. Par exemple au sein du sous-groupe, nous avons identifié une SNP spécifique dans le gène de virulence *mce3F* au sein du groupe B0/W148, un sous-groupe fils nommé N-90 qui est spécifique pour des souches Beijing de la région de Samara en Russie (**Tab. 3**). Il est possible que cette SNP ait favorisé la virulence de la souche. L'avantage de prendre cette SNP c'est qu'elle cible bien ce qui a permis à la souche d'être plus virulente, l'inconvénient est que si cette SNP est vraiment avantageuse, elle pourrait avoir déjà été sélectionnée dans une autre branche évolutive de la tuberculose, et donc ne pas/plus être spécifique du clone recherché.

Tableau 3: SNP caractéristique des souches Beijing épidémiques, Syn – Synonyme, NonSyn – Non Synonyme

Pays	Souche Beijing	Groupe	Position	Gène	Codon	Ref	Mut	SNP
Corée du Sud	Strain K	Asia Ancestral 1	3581476	Rv3204	54	C	T	Syn
Gan Canaria	Strain GC1237	Asia Ancestral 3	22290732	LipT	163	T/A	C/G	Syn
USA	Strain 210 et HN878	Asian Africa 2/RD142	2376135	mpa	52	A	G	Syn
Europe/Russia	Strain B0/W148	Europe Russia B0/W148	1947325	Vap C12	32	T	C	NonSyn
Russia	Strain de N/90	B0/ W148/ N-90	2216485	mce	410	A	C	NonSyn
Asia	Central Asian strain	Central Asia Outbreak	3857226	Rv3437	106	G	T	NonSyn
Asia/Russia	Central Asian/ Russian strain	Clade A de Central Asia	35978	BioF2	562	G	A	NonSyn

5.4 Intérêt des tests développés pour la santé publique

Identifier la cause d'une maladie infectieuse humaine permet de développer des mesures de préventions pour stopper l'expansion de la maladie, maîtriser les chaînes de transmissions, ce qui sauve des vies humaines donc contribue à la santé publique. Mais tracer la cause exacte d'une maladie lors des transmissions interhumaines est parfois très difficile et requiert des méthodes de génotypage adaptées. La connaissance rapide du profil génomique et phénotypique d'une souche par rapport à un antituberculeux permet au médecin d'initier rapidement un traitement adapté aux patients. Un traitement adapté est important pour une guérison et évite la sélection d'autres mutations. Les tests de génomique appliquée sont des solutions qui fournissent des informations génétiques pour chaque souche dans un échantillon clinique et ainsi facilitent sa détermination.

Le TB-EFI conçu sur la base des SNPs fréquentes dans les gènes cibles des antituberculeux dont la Fluoroquinolone, les injectables et l'Ethambutol, est un test génomique pour prédire la TB-pré XDR et XDR, surveiller l'évolution de la tuberculose MDR et XDR dans les pays à forte prévalence.

La combinaison du IS-NTF/RINT aux 17 SNPs de MLPA-Beijing permet directement la détection d'un clone Beijing dans un pays, elle pourrait remplacer les techniques standard de Spoligotypage/MIRU-VNTR dans les régions à forte prévalence de Beijing. Par exemple avec le test IS-NTF/RINT on a pu déterminer que la collection des souches du Kazakhstan

sont des « *modern Beijing* » (Article 2), et avec la SNP de « MLPA-Beijing » identifier leur appartenance au clade CAO (Article 3), qui est identique au clone dit 94-32 d'après le MIRU (Article 4).

Néanmoins l'utilisation du spoligotypage et/ou du MIRU-VNTR reste utile pour l'identification des géotypes de clones circulant dans une région ou pays (Etudes 4-8). C'est le cas des études d'épidémiologie moléculaire effectuées en Papouasie Nouvelle Guinée, au Mozambique, au Pérou, en Sicile qui nous ont permis de révéler la lignée, le géotype prévalent ainsi que les clones prédominants. Ainsi la confrontation du géotypage avec les données de localisation permet d'identifier l'étendue géographique des épidémies et son évolution dans les pays voisins en Europe comme en Afrique.

Aujourd'hui les outils d'épidémiologie classique et moléculaire sont d'une grande importance pour la santé publique, car ils permettent l'étude des transmissions de cas et la surveillance des clones et maladies. Les données socio-démographiques, données antérieures d'épidémiologie-moléculaire et de géolocalisation sont aussi importantes à confronter à celles d'épidémiologie classique en vue d'établir des liens de maladie entre patients, villes, pays et continents. Ainsi avec le géotypage des souches Beijing MDR/XDR à base d'IS6110 en NTF, les SNPs, les RD, qui sont les marqueurs les plus informatifs, on est en marche pour une surveillance mondiale des souches Beijing, principalement de ses clones épidémiques, à des coûts plus compatibles avec les moyens financiers des pays les plus touchés.

Le IS-NTF/RINT, le TB-EFI combiné à la MLPA-Beijing (en cours de développement), avec la disponibilité des données sociodémographiques, le report dans une base de données (Ex SITVITWEB) et les études d'épidémiologie sont susceptibles de faciliter le suivi des clones Beijing MDR/XDR.

Chapitre 6 : Conclusion Générale

Conclusion Générale

Nous avons finalisé le développement de deux méthodes :

- le TB-EFI, une méthode moléculaire rapide et multiplexée pour la détection de la résistance génotypique des souches de MTBc à la Fluoroquinolone, la Kanamycine, l'Amikacine, la Capréomycine, et l'Ethambutol,

- le test IS-NTF/RINT, une méthode moléculaire de typage des souches Beijing par rapport au nombre de copies de la séquence d'insertion *IS6110* dans la région NTF, et de détection de la résistance à la Rifampicine et l'Isoniazide pour l'identification des transmissions interhumaines, le contrôle, et la surveillance des clones Beijing épidémiques.

Nous avons initié le développement de la méthode « MLPA-Beijing » pour la classification précise des souches au sein de la lignée 2/Beijing.

Ces méthodes sont développées sur technologie haut débit, moins onéreuse que le séquençage de génome complet. Elles constituent en plus des méthodes internes existant dans notre laboratoire (TB-Sprint/MLPA) avec d'une part un bénéfice pour les pays à forte prévalence de tuberculose MDR et XDR et d'autre part un avantage pour le suivi et la maîtrise de l'expansion des clones Beijing.

Chapitre 7 : Matériel et méthodes

Matériel et méthodes

7.1 Amorces Dual Priming Oligonucléotides (DPO)

Le système des amorces DPO a été développé par Chun et al. Pour la détection du polymorphisme de SNP du gène CYP2C19 chez l'homme et la détection multiplexée de virus respiratoires. Une amorce DPO comprend trois parties qui jouent différents rôles dans le processus d'hybridation et d'extension du fragment cible. Ces parties sont : une longue partie 5' de 18 à 25 nucléotides, qui initie en premier une hybridation stable, une courte partie 3' de 6-12 nucléotides qui détermine la spécificité de l'élongation et enfin un polydéoxyinosine (3-8 bases déoxyinosines) qui relie les deux parties 5' et 3'. Ce système d'amorce DPO permet d'éviter les problèmes de dimer, de « hairpin » et autres structures secondaires avec les amorces classiques. Il facilite le multiplexage en PCR, et une économie de temps en design. Pour développer le TB-EFI, ainsi que dans le IS-NTF/rint, nous avons donc utilisé des amorces DPO pour amplifier simultanément 4 locus (*gyrA*, *rrs*, *embB* genes et eis- promoter) et 3 locus (*rpoB*, *KatG* genes, et *inhA* promoter).

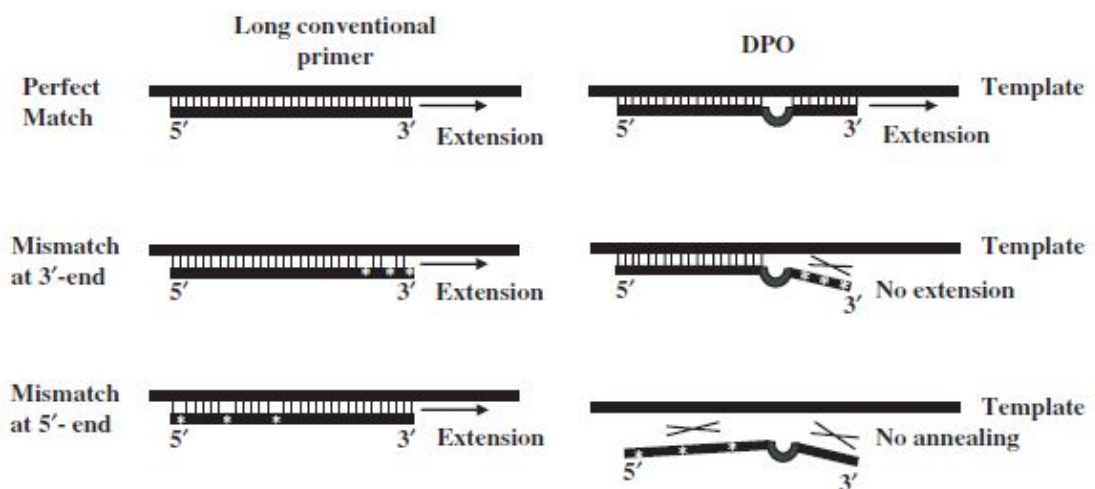


Figure 24 : Schéma du principe des amorces DPO et comparaison à des longues amorces classiques. Une amplification non spécifique avec les amorces classiques est obtenue malgré les mélangements en 5' ou 3'. Un mélangement à l'extrémité 3' de l'amorce DPO empêche l'élongation mais un mélangement à l'extrémité 5' conduit à une absence d'hybridation et dans les deux cas de mélangement il n'y a pas d'amplification (Chun et al).

7.2 Oligonucléotide de capture ou Sonde

Le dessin des oligonucléotides de captures est l'étape la plus importante, car il faut qu'elles arrivent à bien s'hybrider avec les produits de PCR biotinylés. Certaines sondes sont dessinées à l'aide du programme en ligne PrimerPlus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) et d'autres manuellement. Mais toutes les sondes ont été analysées avec le logiciel idt oligo analyzer (<https://eu.idtdna.com/calc/analyzer>) pour vérifier leurs spécificités. Les amorces pour le séquençage ont été aussi dessinées à l'aide des mêmes programmes et logiciels.

7.3 Détection des copies d'IS6110 copies dans la région NTF sur gel d'agarose

Nous avons premièrement établi une procédure proche de celle proposée par Plykaitis et al (1994) pour identifier la présence d'IS6110 dans les isolats, et sa présence en 5' et/ou en 3' extrémité du locus NTF. La seule modification comparée au protocole de Plykaitis est la réduction de la taille de l'amplicon de 523 bp à 355 bp pour une utilisation potentielle sur le système Luminex qui requiert un amplicon moins de 400 bp. Pour cela, nous avons dessiné une nouvelle amorce reverse appelé IS60B combinée à la forward IS59 forward (IS60B sequence = TGATGTGCTCCTTGAGTTCG).

La réaction de d'amplification contient 2 µl d'ADN et 23 µl de réaction mix (0.2 mM dNTPs, 1 µM d'amorces, 0.25U de GoTaq Flexi, 1.5mM MgCl₂), comme recommandé par Promega (Promega Corp., Fitchburg, WI). Le mélange a été préchauffé pendant 3min à 95°C et soumis à 25 cycles de 30s à 95°C, 30s à 55°C, et 30s à 72°C. Le produit amplifié obtenu était analysé par électrophorèse sur gel d'agarose de 2% et était visualisé au BET (Bromure d'éthidium).

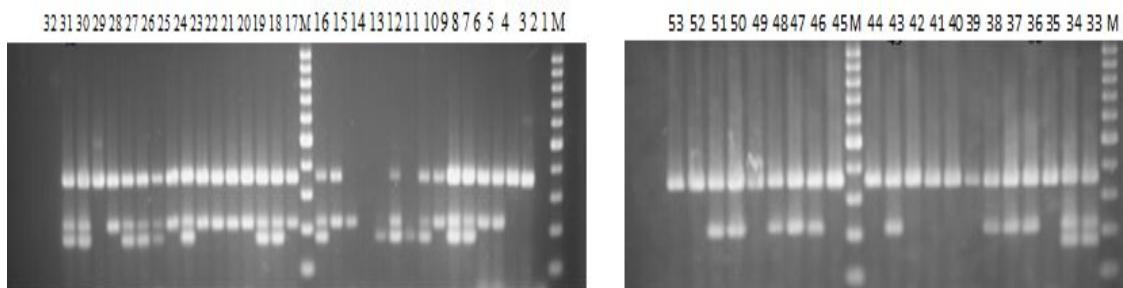


Figure 25 : Amplicons obtenus par amplification d'IS6110 dans la région NTF des ADN, des USA et détection sur gel. M = ladder (100bp), IS6110_NTF5' = 175 bp, IS6110_NTF3' = 223 bp, IS6110 = 355 b.p,

3 bandes sur gel = IS6110_NTF5'/IS6110_NTF3'/IS6110 = Moderne L2/Beijing W-strain,
 2 bandes sur gel = IS6110_NTF3'/IS6110 = Modern L2/Beijing),
 1 bande sur gel = IS6110 = Ancient L2/Beijing).

On a ainsi analysé la collection d'ADN du Kazakhstan par typage d'IS6110 dans la région NTF sur gel d'électrophorèse, et les 216 souches analysées sont des « *Modern Beijing* ».

7.4 Detection of IS6110 insertion in NTF-locus on microsphere-based flow cytometry Luminex200

En se basant sur le génome de la souche W148 et à l'aide du program ApE (A plasmid Editor by M. Wayne Davis), nous avons identifiée la region NTF et la séquence d'insertion IS6110 en 3', ensuite représenté un prototype de la souche « W » (2 IS6110). Avec cette approche, nous avons dessiné des amorces, et utilisé certaines amorces publiées afin d'amplifier un petit fragments d'IS6110 et de cette insertion en 5' et 3' extrémités de la région NTF. Nous avons aussi dessiné des sondes pour détecter les amplicons avec la technologie Luminex.

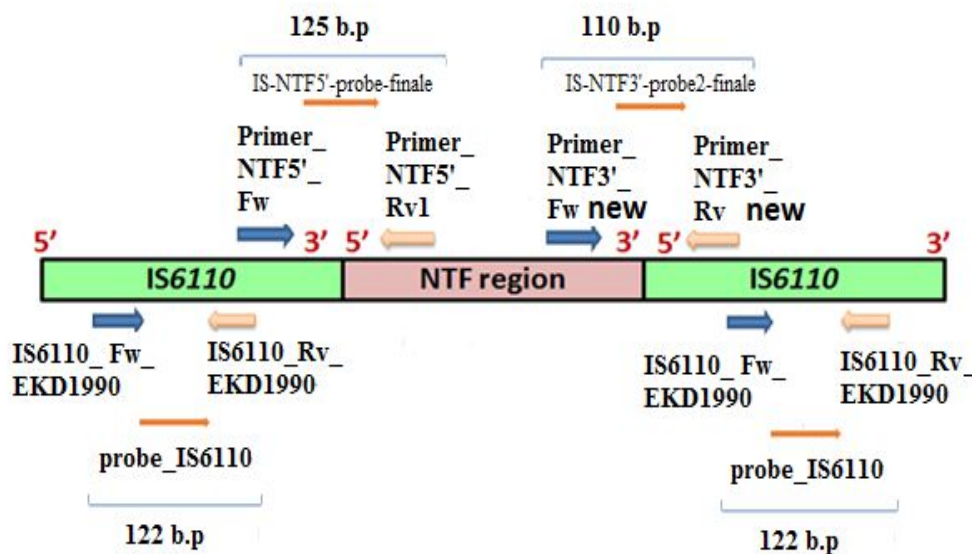


Figure 26 : Schéma de la représentation des amorces et sondes pour l'amplification et l'hybridation/détection des séquences d'insertion IS6110 dans le locus NTF utilisant le cytomètre en flux du système Luminex.

7.5 Dessin des oligonucléotides pour les marqueurs de la MLPA/Beijing

Pour développer la MLPA/Beijing, nous avons sélectionné 31 marqueurs SNP, région de délétion (RD), séquences d'insertion (IS6110 en NTF).

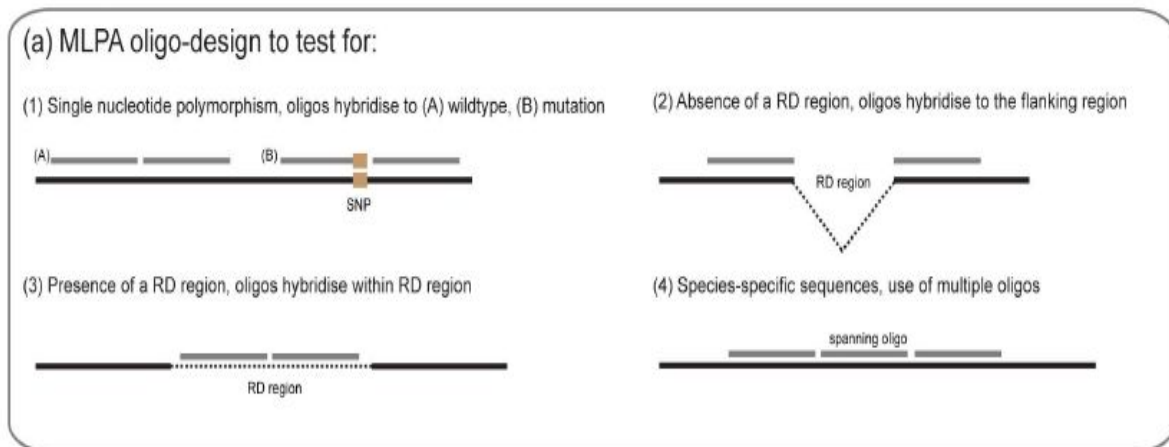


Figure 27 : Schéma du principe de dessin des sondes destinée à la MLPA

Les sondes MLPA sont dessinées pour identifier le polymorphisme de nucléotide simple, l'absence (2) et la présence (3) de région de délétions (RD), (4) type de séquence spécifique

Tableau 4 : Séquences des sondes 1 et 2 dessiner, pour la méthode MLPA, séquence en rouge – amorce Fw, séquence en orange – Séquence xTAG, nucléotide en rouge foncé – SNP mut/wt

Groupe de classification	GENE	SNP (codon)	Assemblage 1, Oligonucleotides Synthétique et Specificque with XTAG
Modern Beijing	ogt	G12A	GGGTTCCCTAAGGGTTGGACATAATCAATTTCAACTTTCTACTCGCACCATCGATAGCCCCATCGGA
Asia ancestral 3	ogt	G37T	GGGTTCCCTAAGGGTTGGACAAATACATAATCTTACATTCCTCTGCGGATGCTCGAGCAGACGTATGAGCCAAGCC
Asia outbreak (AO)	BioF2	G562	GGGTTCCCTAAGGGTTGGAAATTTCTTCTTTCTTTTCCACAATTGTTTCAGCATGGAAGGCACCGTCG
Beijing	BioF2	G572C	GGGTTCCCTAAGGGTTGGATACTTCTTTACTACAATTTACAACCTGG CCA CCA TCG CCG AGC TTG CC
Eu/Ru BO/W148 et CA outbreak	Rv0177	A58C	GGGTTCCCTAAGGGTTGGAACACTTATCTTTCAATTCAATTACTCCCATGAATCGCACCAGCGC
Proto Beijing	Rv0177	A122	GGGTTCCCTAAGGGTTGGAATACTTTACAAACAAATAACACAC TCCTGATTGCATCTCCGGGGTGAACCGACCAC
Asia ancestral 1	Rv3204	C54T	GGGTTCCCTAAGGGTTGGACTTTCTCATACTTTCAACTAATTTGCGGACCGATTCTCGGATCTGCC
Asia ancestral 2	lipT	T163C	GGGTTCCCTAAGGGTTGGATCAAACCTCTCAATTCTTACTTAATGGT GTC TTG ACC TGT CGT CCT TGT CGA CC
Asian African3_Bmyc13	lipT	C218T	GGGTTCCCTAAGGGTTGGACAAACAAACATTCAAATATCAATCATATCACCGCCACACTGTTGGCGGTGCCAGC
Asian african 1	OxyS	G86C	GGGTTCCCTAAGGGTTGGACTTTCTTAATACATTACAACATACAAGGCCGAGGTGGATGCGGTGCGC
ST25(Bmyc6,25,26) + modern Beijing	OxyS	G142A	GGGTTCCCTAAGGGTTGGATACATTCAACACTCTTAAATCAAAGCTGTACCGACGGCTGCGGAATTCGAGCTC
BO/W148 outbreak	VapC12	T32C	GGGTTCCCTAAGGGTTGGATAACTTACACTTAACATATCATCTTACATTTTCGATGTCGAGGTGATCGGGGCGATACG
Pacific RD150	VapC12	A46G	GGGTTCCCTAAGGGTTGGACACTTAATTCATTCTAAATCTATCGCTTGGCGGGCGAAACCGTGCG
CAO	Rv3437	G106T	GGGTTCCCTAAGGGTTGGATACTACTTCTATAACTCACTTAAAGGG TGC GAG CCG CCG CCT
BO/W148 outbreak, N-90	mce3F	A410	GGGTTCCCTAAGGGTTGGACTTAACATTTAACTTCTATAACACAACGACCGCTACAACCTGGAAGGGCGA
Asian African 2	mpa	A52	GGGTTCCCTAAGGGTTGGAACACTTATTCTCAAACCTAATA CGGGCTTCGAGTTGATGTATATCGCGGGCA
Asian African 2/ RD142	IS6110 in NTF	C11	GGGTTCCCTAAGGGTTGGAACCTATTTCTTCACTACTATATCAACCGGGCGTCTTTGGACTGGATCGC
Beijing	IS6110	Pres	GGGTTCCCTAAGGGTTGGACATCTTCATATCAATTTCTTATTACC TAA CCG GCT GTG GGT AGC AGA CCT T
Détection of W-strain	IS-NTF5'	Pres	GGGTTCCCTAAGGGTTGGAATTAACAACCTTTAACTACACAACGGACTCACCGGGGCGGTTCAA
Détection of modern Beijing	IS-NTF3'	Pres	GGGTTCCCTAAGGGTTGGATACAACATCTCATTAAACATATACATTCGCAGGGCCGGCGATGAGGTCG
Présence of deletion RD	RD105	Absc	GGGTTCCCTAAGGGTTGGAATTCATACTATCTAACACTTACTGGCGTGCCGTGCGCCCCGTCCA
Présence of deletion RD	RD207	Absc	GGGTTCCCTAAGGGTTGGACACTACACATTTATCATAACAAATTTCCGTCCCCTCTCGGGTTTTGGGTCTGACGACG

Présence of deletion RD	RD181	Absc	GGGTTCCCTAAGGGTTGGATCATCACTTTCTTTACTTTACATTGCAGACTAAGACCGGCAGTCCCCGCCAGAAAGGG
Présence of deletion RD	RD ex105	Absc	GGGTTCCCTAAGGGTTGGATACACAATATTCATCATAACTAACAGTCTCACGACGAGCAGCTGCAGCGCCGCC
Targeting the absence of the RD	RD131	Pres	GGGTTCCCTAAGGGTTGGAACAAATATCTAACTACTATCACAAGAACGCCATGATTTCTGGTCGGAA
Absence of dele	RD150	Pres	GGGTTCCCTAAGGGTTGGAACCTTTCTCTCTATTCTTATTTACCGATCGGCACAGCGAGACGAACAATGCGTGGA
lineage 2	glnH	wt	GGGTTCCCTAAGGGTTGGAACCTTTCTCTCTATTCTTATTTACGCCGTAGGGCTGGTCGGCCAT
Llineage 2.2	CobC	wt	GGGTTCCCTAAGGGTTGGAACCTTTCTCTCTATTCTTATTTAGCACCAGCACATCGGGCAGCGAGTCATCG
lineage 2.2.1	Rv0697	wt	GGGTTCCCTAAGGGTTGGAACCTTTCTCTCTATTCTTATTTAGCGATTAGTCAGCTCATCTGTTGATGCTC
peu avant mod Beijing	mutT4	wt	GGGTTCCCTAAGGGTTGGAACCTTTCTCTCTATTCTTATTTACCCCGCAACGGCGAAGCGGTCC
Présence of deletion RD	RD142		GGGTTCCCTAAGGGTTGGATTAACAATCTACTATTCAATCACGCTGACCGGGTACTCTCGACGAC G
LumH	LumH		GATGAATATAGTAAGTATTGAGTATCTCTTTAAACACATTCAACAATA
LumQ	LumQ		GGGTTCCCTAAGGGTTGGAATCAACACACAATAACATTATCTAGATTGGATCTTGCTGGCGGTCC
LumD	LumD		GGGTTCCCTAAGGGTTGGACAATTTACATTTCACTTTCTTATCGACATCGTGGTGATCGACTCGGT

Groupe de classification	GENE	SNP (codon)	Assemblage de sonde 2, MLPA
Modern Beijing	ogt	G12A	CCATTAACCCTGGCCGGGCATGGCTTCTAGATTGGATCTTGCTGGCAC
Asia ancestral 3	ogt	G37T	TCACACACTGGACACCCGACCCCTCTAGATTGGATCTTGCTGGCAC
Asia outbreak (AO)	BioF2	G562	CCGACCTGGCCACCATCGCCGAGCTTCTAGATTGGATCTTGCTGGCAC
Beijing	BioF2	G572C	CACCGGCACGGCTGCCGGTCTATCTAGATTGGATCTTGCTGGCAC
Eu/Ru BO/W148 et CA outbreak	Rv0177	A58C	GCAGCGCACAAGGATATCGTGATGCTCAGTGATGTCTAGATTGGATCTTGCTGGCAC
Proto Beijing	Rv0177	A122	AGGAACGGTTCTAGACGCGGGCGTACAGTCTAGATTGGATCTTGCTGGCAC
Asia ancestral 1	Rv3204	C54T	TGGCACCGGTGATCAGAGCCTCCG TCTAGATTGGATCTTGCTGGCAC
Asia ancestral 2	lipT	T163C	CCGCAGATCACCTCGACAGCAACGTGTACCTTCTAGATTGGATCTTGCTGGCAC
Asian African3_Bmyc13	lipT	C218T	TGCCAAAGGCCTATTCGCCAGGGCTCTAGATTGGATCTTGCTGGCAC
Asian african 1	OxyS	G86C	TCCGGGATAACCGGGACGCTTCGGCTAGGCTCTAGATTGGATCTTGCTGGCAC
ST25(Bmyc6,25,26) + modern Beijing	OxyS	G142A	AATGCCGTCATCGTGACCCCGAGTCTAGATTGGATCTTGCTGGCAC

BO/W148 outbreak	VapC12	T32C	CCAAGCGGTCGTGCGTCAGCTGATCATCTAGATTGGATCTTGCTGGCAC
Pacific RD150	VapC12	A46G	CGCCCTGCACATTCGATGTCGAGGTGATCTCTAGATTGGATCTTGCTGGCAC
CAO	Rv3437	G106T	ACGGGGCGTCCGTTCCCCACGCTCTAGATTGGATCTTGCTGGCAC
BO/W148 outbreak, N-90	mce3F	A410	CCAACGCCACGGTGCCGGTTTTCTAGATTGGATCTTGCTGGCAC
Asian African 2	mpa	A52	GAGCGCGTCGGGGCATGTGATCCGATCTAGATTGGATCTTGCTGGCAC
Asian African 2/ RD142	IS6110 in NTF	C11	CCGGTAGCAGCGTTCGAAGTCGTCGTGCACCTCTAGATTGGATCTTGCTGGCAC
Beijing	IS6110	Pres	CACCTATGTGTCGACCTGGGCAGGGTTCGCCTACGCTCTAGATTGGATCTTGCTGGCAC
Détection of W-strain	IS-NTF5'	Pres	CATGGTGACCCTCGTGAGGCATGGGAAGTGGGCTCTAGATTGGATCTTGCTGGCAC
Détection of modern Beijing	IS-NTF3'	Pres	TGAACCGCCCCGGCATGTCCGGAGCTCTAGATTGGATCTTGCTGGCAC
Présence of deletion RD	RD105	Absc	CGGCGACCACGGCGTCCACCGGGTCCACTCTAGATTGGATCTTGCTGGCAC
Présence of deletion RD	RD207	Absc	TTCAGGTCGGCCGTGCCATCGAGGATCTAGATTGGATCTTGCTGGCAC
Présence of deletion RD	RD181	Absc	AGTGATTGCATGGCCAAGGATCTGGTCGCCATCTAGATTGGATCTTGCTGGCAC
Présence of deletion RD	RD ex105	Absc	CCATCATGGCGAATCTTGTTGGACGAGCTGCTTCTAGATTGGATCTTGCTGGCAC
Targeting the absence of the RD	RD131	Pres	AGCCGGTCGCCGTCGAGCTCGGCTGCCAGCAGTGCCTAGATTGGATCTTGCTGGCAC
Absence of dele	RD150	Pres	TGTCGTTGCAGTTGCCACCGAGCAGTCTAGATTGGATCTTGCTGGCAC
lineage 2	glnH	wt	GTCGGGCCCGACGATATGCAGGTAGGGTCTAGATTGGATCTTGCTGGCAC
Llineage 2.2	CobC	wt	GCCAACGATTGCGGCTCGCCGGGAATCTAGATTGGATCTTGCTGGCAC
lineage 2.2.1	Rv0697	wt	TCGGGCGTCGGCGAGGAGGAGTCTAGATTGGATCTTGCTGGCAC
peu avant mod Beijing	mutT4	wt	CGGTCCCGCTCACCTCGTCGCGGGTCTAGATTGGATCTTGCTGGCAC
Présence of deletion RD	RD142		AGGTCCGGCTGGCTTGCTGATCATGCTCGTCTAGATTGGATCTTGCTGGCAC
LumH	LumH		
LumQ	LumQ		
LumD	LumD		GGCGGCGCTGGTGCCGCGCGGGTCTAGATTGGATCTTGCTGGCAC

7.6 Le principe de la MLPA

La MLPA ou « *Multiple Ligation-dependent Probe Amplification* » est une méthode de PCR multipléxée, détectant des polymorphismes de nucléotide simple (SNP) utilisant une seule séquence de sonde pour identifier la SNP. Le principe de cette technique consiste en une étape de design des oligonucléotides, une étape d'hybridation des oligonucléotides à la cible d'ADN, une étape de ligation des oligos hybridés, une étape d'amplification des oligonucléotides liés par PCR. En général, après la PCR, la détection se fait par électrophorèse capillaire mais dans notre cas, nous avons prévue une détection avec la technologies xTAG de Luminex sur la plateforme MAGPIX. Dans ce cas, on utilise un unique xTAG (24 nt) présent sur chaque sonde MLPA amplifiée, qui est complémentaire à l'anti Xtag sur une bille dans un panel donné. L'avantage de cette technique est qu'il facilite la détection de plusieurs marqueurs. On l'a ainsi choisi en vue du développement du test MLPA/Beijing.

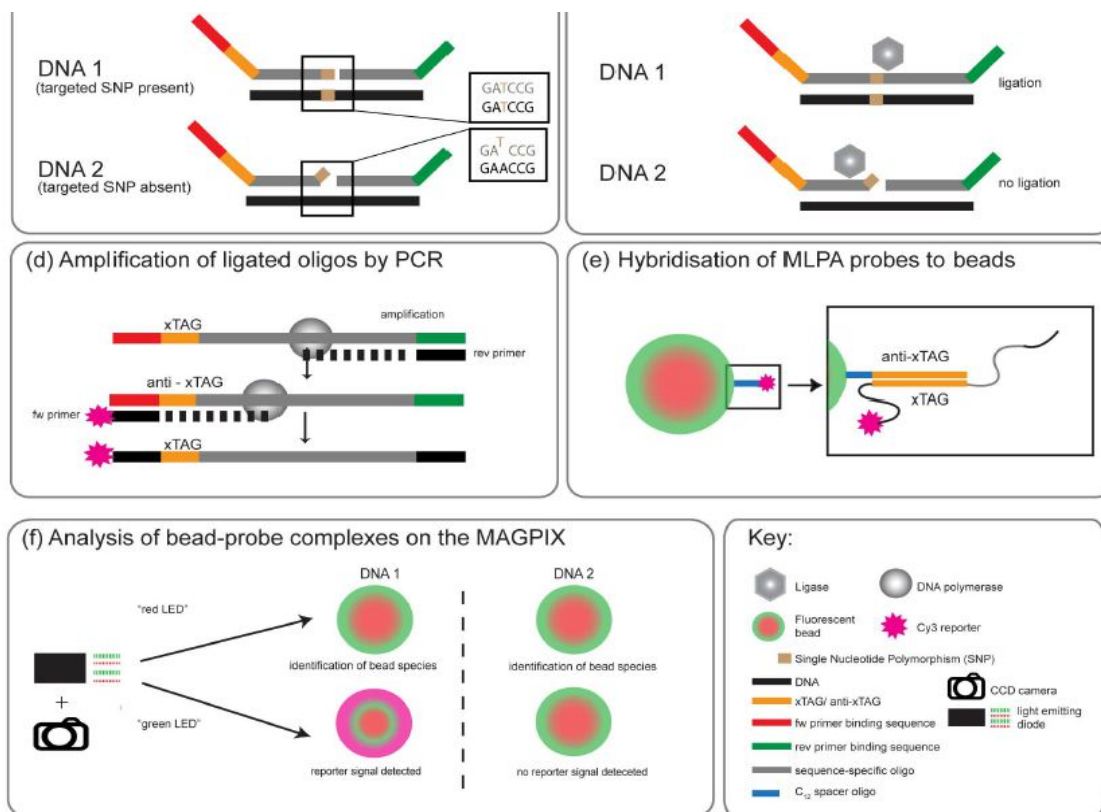


Figure 24 : Principe de la MLPA « *Multiple Ligation-dependent Probe Amplification* » avec des microbilles sur MAGPIX.

(b) Hybridation des sondes MLPA aux ADN cibles : la séquence spécifique de chaque sonde s'hybride à la l'ADN 1 ou 2. Chaque sonde correspond à une séquence cible spécifique (en vert), et un unique xTAG (orange), lié à des séquence forward et reverse (en rouge et vert). Par exemple la sonde MLPA s'hybride parfaitement à la séquence de l'ADN qui contient la SNP 1. (c) Ligation des oligos hybridés : seul les oligos qui sont hybridisé drectement à l'autre partie adjacentes sont ligué. (d) Amplifications des oligos ligués par PCR, tous les oligos ligués sont amplifiés dans une réaction de PCR avec des amorces forward lié à Cy3 et un reverse non lié. (e) Hybridation du produit MLPA obtenue aux billes, les sondes amplifiées s'hybrident à leur anti-xTAG couplé à un seul type de billes. (f) Analyse et détection de la combinaison bille-sonde sur le MAGPIX : le laser rouge (LED) et la camera CCD identifie le type de bille, suivie de l'exitation de la molécule cy3 attaché aux sonde par le laser vert. Le signal est transmis en forme de MFI. Par exemple, pour l'ADN 1, le MFI est détecté sur le type de bille indiquant la présence de la SNP (Bergval et al 2012).

Bibliographie

1. Achtman, M. (2016). "How old are bacterial pathogens?" Proceedings of the Royal Society B: Biological Sciences **283**(1836): 20160990.
2. Achtman, M. (2016). "How old are bacterial pathogens?" Proc Biol Sci **283**(1836).
3. Ahmad, S., E. Mokaddas, et al. (2016). "Discordance across Phenotypic and Molecular Methods for Drug Susceptibility Testing of Drug-Resistant Mycobacterium tuberculosis Isolates in a Low TB Incidence Country." PLoS One **11**(4): e0153563.
4. Akhtar, M., G. Bretzel, et al. (2000). "Guide technique, diagnostic de la tuberculose par examen microscopique direct des expectorations dans les pays à faibles revenus.pdf>." Union Internationale Contre la Tuberculose et les Maladies Respiratoires.
5. Allix-Beguec, C., D. Harmsen, et al. (2008). "Evaluation and user-strategy of MIRU-VNTRplus, a multifunctional database for on-line analysis of genotyping data and phylogenetic identification of Mycobacterium tuberculosis complex isolates." J Clin Microbiol **46**(8): 2692-2699.
6. Alonso H, A. J., Samper S, Caminero JA, Campos-Herrero MI, Gicquel B, Brosch R, Martín C, Otal I. (2011). "Deciphering the role of IS6110 in a highly transmissible Mycobacterium tuberculosis Beijing strain, GC1237." Tuberculosis (Edinb) **91**(2): 117-126.
7. Andrea E. DeBarber, Khisimuzi Mdluli, et al. (2000). "Ethionamide activation and sensitivity in MDR MTB.pdf>." PNAS vol. **97** no. **17**: 9677-9682.
8. Anonymous (2015). TB diagnostics market in select high-bruden countries : current market and future opportunities for novel diagnostics. Geneva, UNITAID Secretariat Chemin de Blandonnet 10 – BIBC III – 8th Floor 1214 Vernier Switzerland, WHO: 1-68.
9. Anonymous (2015). Tuberculosis Diagnostics Technology and Market Landscape. Geneva, UNITAID Secretariat Chemin de Blandonnet 10 – BIBC III – 8th Floor 1214 Vernier Switzerland, WHO: 1-96.
10. AR Zink, U Reischl, et al. (2002). "Molecular analysis of ancient microbial infections." FEMS Microbiol Lett **213**: 141–147.
11. Asante-Poku, A., D. Yeboah-Manu, et al. (2015). "Mycobacterium africanum is associated with patient ethnicity in Ghana." PLoS Negl Trop Dis **9**(1): e3370.
12. Aung, H. L., T. Tun, et al. (2018). "Association between anti-tuberculosis drug resistance-conferring mutations and treatment outcomes in Myanmar." Infect Dis (Lond) **50**(5): 388-390.
13. Avalos, E., D. Catanzaro, et al. (2015). "Frequency and geographic distribution of gyrA and gyrB mutations associated with fluoroquinolone resistance in clinical Mycobacterium tuberculosis isolates: a systematic review." PLoS One **10**(3): e0120470.
14. Aze, J., C. Sola, et al. (2015). "Genomics and machine learning for taxonomy consensus: the Mycobacterium tuberculosis complex paradigm." PLoS one **10**(7): e0130912.
15. Aze, J., C. Sola, et al. (2015). "Genomics and Machine Learning for Taxonomy Consensus: The Mycobacterium tuberculosis Complex Paradigm." PLoS One **10**(7): e0130912.
16. Bahr, N. C., E. Nuwagira, et al. (2018). "Diagnostic accuracy of Xpert MTB/RIF Ultra for tuberculous meningitis in HIV-infected adults: a prospective cohort study." The Lancet Infectious Diseases **18**(1): 68-75.
17. Baker, O., O. Y. C. Lee, et al. (2015). "Human tuberculosis predates domestication in ancient Syria." Tuberculosis **95**: S4-S12.
18. Barben Jürg , Berger Christophe, et al. (2012). "Manuel de la tuberculose, Handbuch_Tuberculose_2012_fr.pdf." Ligue pulmonaire suisse.
19. Barletta, F., L. Otero, et al. (2015). "Predominant Mycobacterium tuberculosis Families and High Rates of Recent Transmission among New Cases Are Not Associated with Primary Multidrug Resistance in Lima, Peru." J Clin Microbiol **53**(6): 1854-1863.

20. Barnes, P. F. and M. D. Cave (2003). "Molecular epidemiology of tuberculosis..pdf>." NEJM **349**: 1149-1156.
21. Baulard, A. R., J. C. Betts, et al. (2000). "Activation of the pro-drug Ethionamide is regulated in mycobacteria..pdf>." JBC Papers in Press.
22. Bergval, I., S. Sengstake, et al. (2012). "Combined Species Identification, Genotyping, and Drug Resistance Detection of Mycobacterium tuberculosis Cultures by MLPA on a Bead-Based Array." PLoS One **7**(8): e43240.
23. Bhatia, A. L., A. Csillag, et al. (1961). "The virulence in the guinea-pig of tubercle bacilli isolated before treatment from South Indian patients with pulmonary tuberculosis. 2. Comparison with virulence of tubercle bacilli from British patients." Bull World Health Organ **25**: 313-322.
24. Bhatia, A. L., C. V. Jacob, et al. (1963). "A Comparison of the Virulence in the Guinea-Pig of Tubercle Bacilli from Thai and British Patients." Bull World Health Organ **29**: 483-490.
25. Bifani, P. J., B. Mathema, et al. (2001). "Global dissemination of the Mycobacterium tuberculosis, W-Beijing family strain.pdf>." Trends in Microbiology **10**.
26. Blanc, F.-X., T. Sok, et al. (2011). "Earlier versus Later Start of Antiretroviral Therapy in HIV-Infected Adults with Tuberculosis." New England Journal of Medicine **365**(16): 1471-1481.
27. Boehme, C. (2017). "New Diagnostic Technologies for MDR-TB (Abstract 52). Oral abstract presented at annual Conference on Retroviruses and Opportunistic Infections (CROI)." <http://www.croiconference.org/sessions/new-diagnostic-technologies-mdr-tb>.
28. Boeree, M. J., A. H. Diacon, et al. (2015). "A Dose-Ranging Trial to Optimize the Dose of Rifampin in the Treatment of Tuberculosis." American Journal of Respiratory and Critical Care Medicine **191**(9): 1058-1065.
29. Boeree, M. J., N. Heinrich, et al. (2017). "High-dose rifampicin, moxifloxacin, and SQ109 for treating tuberculosis: a multi-arm, multi-stage randomised controlled trial." The Lancet Infectious Diseases **17**(1): 39-49.
30. Bonnie B. Plikaytis, Jennifer L. Marden, et al. (1994). "Multiplex PCR Assay Specific for the Multidrug-Resistant Strain W of Mycobacterium tuberculosis.pdf>." JOURNAL OF CLINICAL MICROBIOLOGY **Vol. 32, No. 6**: 1542-1546.
31. Borgdorff, M. W., S. van den Hof, et al. (2011). "Influence of sampling on clustering and associations with risk factors in the molecular epidemiology of tuberculosis." Am J Epidemiol **174**(2): 243-251.
32. Borrell, S., Y. Teo, et al. (2013). "Epistasis between antibiotic resistance mutations drives the evolution of extensively drug-resistant tuberculosis." Evol Med Public Health **2013**(1): 65-74.
33. Bos, K. I., K. M. Harkins, et al. (2014). "Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis." Nature **514**(7523): 494-497.
34. Bradley, P., N. C. Gordon, et al. (2015). "Rapid antibiotic-resistance predictions from genome sequence data for Staphylococcus aureus and Mycobacterium tuberculosis." Nat Commun **6**: 10063.
35. Brites, D. and S. Gagneux (2012). "Old and new selective pressures on Mycobacterium tuberculosis." Infect Genet Evol **12**(4): 678-685.
36. Brosch, R., S. V. Gordon, et al. (2002). "A new evolutionary scenario for the Mycobacterium tuberculosis complex." Proceedings of the National Academy of Sciences **99**(6): 3684-3689.
37. Brossier, F., D. Guindo, et al. (2016). "Performance of the New Version (v2.0) of the GenoType MTBDRsl Test for Detection of Resistance to Second-Line Drugs in Multidrug-Resistant Mycobacterium tuberculosis Complex Strains." J Clin Microbiol **54**(6): 1573-1580.
38. Brossier, F., D. Guindo, et al. (2016). "Performance of the New Version (v2.0) of the GenoType MTBDRsl Test for Detection of Resistance to Second-Line Drugs in Multidrug-

- Resistant Mycobacterium tuberculosis Complex Strains." Journal of Clinical Microbiology **54**(6): 1573-1580.
39. Brudey, K., J. Driscoll, et al. (2006). "Mycobacterium tuberculosis complex genetic diversity : mining the fourth international spoligotyping database (SpolDB4) for classification, Population Genetics, and Epidemiology." BMC Microbiol. **6**(6): 23.
 40. Cabibbe, A. M., P. Miotto, et al. (2015). "Lab-on-Chip-Based Platform for Fast Molecular Diagnosis of Multidrug-Resistant Tuberculosis." J Clin Microbiol **53**(12): 3876-3880.
 41. Canetti, G., S. Froman, et al. (1963). "Mycobacteria: Laboratory Methods for testing drug sensitivity and resistance.pdf>." Bull. Org. mond. Santé **29**: 565-578.
 42. Canetti, G., N. Rist, et al. (1963). "Mesure de la sensibilité du bacille Tuberculeux aux drogues antibacillaires par la methode des proportions " Revue de Tuberculose et de Pneumologie **27**: 217-272.
 43. Carette (2002). "Imagerie typique et atypique de la tuberculose bronchopulmonaire.pdf>."
 44. Casali, N., V. Nikolayevskyy, et al. (2014). "Evolution and transmission of drug-resistant tuberculosis in a Russian population." Nat Genet **46**(3): 279-286.
 45. Cattoir, V. (2004). "<Identification moléculaire des mycobactéries et détection de la résistance aux antibiotiques.pdf>." Ann Biol Clin, vol. 62, n° 4.
 46. Caws, M., G. Thwaites, et al. (2008). "The influence of host and bacterial genotype on the development of disseminated disease with Mycobacterium tuberculosis." PLoS Pathog **4**(3): e1000034.
 47. Chaidir, L., S. Sengstake, et al. (2016). "Predominance of modern Mycobacterium tuberculosis strains and active transmission of Beijing sublineage in Jayapura, Indonesia Papua." Infect Genet Evol **39**: 187-193.
 48. Chakravorty, S., A. M. Simmons, et al. (2017). "The New Xpert MTB/RIF Ultra: Improving Detection of Mycobacterium tuberculosis and Resistance to Rifampin in an Assay Suitable for Point-of-Care Testing." MBio **8**(4).
 49. Chernyaeva, E., E. Fedorova, et al. (2013). "Characterization of multiple and extensively drug resistant Mycobacterium tuberculosis isolates with different ofloxacin-resistance levels." Tuberculosis (Edinb) **93**(3): 291-295.
 50. Chihota, V., L. Apers, et al. (2007). "Predominance of a single genotype of Mycobacterium tuberculosis in regions of Southern Africa." Int J Tuberc Lung Dis **11**(3): 311-318.
 51. Chun, J. Y., K. J. Kim, et al. (2007). "Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of CYP2C19 gene." Nucleic Acids Res **35**(6): e40.
 52. Click, E. S., C. A. Winston, et al. (2013). "Association between Mycobacterium tuberculosis lineage and time to sputum culture conversion." Int J Tuberc Lung Dis **17**(7): 878-884.
 53. Cole, S. T. (1999). "Learning from the genome sequence of Mycobacterium tuberculosis H37Rv.pdf>." FEBS **452**: 7-10.
 54. Coll, F., K. Mallard, et al. (2012). "SpolPred: rapid and accurate prediction of Mycobacterium tuberculosis spoligotypes from short genomic sequences." Bioinformatics **28**(22): 2991-2993.
 55. Coll, F., R. McNerney, et al. (2014). "A robust SNP barcode for typing Mycobacterium tuberculosis complex strains." Nature Communications **5**(1).
 56. Coll, F., J. Phelan, et al. (2018). "Genome-wide analysis of multi- and extensively drug-resistant Mycobacterium tuberculosis." Nat Genet.
 57. Coll, F., M. Preston, et al. (2014). "PolyTB: A genomic variation map for Mycobacterium tuberculosis." Tuberculosis (Edinb) **94**(3):346-54(3): 346-354.
 58. Comas, I., M. Coscolla, et al. (2013). "Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans." Nat Genet **45**(10): 1176-1182.
 59. Comas, I., M. Coscolla, et al. (2013). "Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans." Nature Genetics **45**(10): 1176-1182.

60. Comas, I., S. Homolka, et al. (2009). "Genotyping of genetically monomorphic bacteria: DNA sequencing in mycobacterium tuberculosis highlights the limitations of current methodologies." *PLoS One* **4**(11): e7815.
61. Cooper, A. M. (2009). "Cell-Mediated Immune Responses in Tuberculosis." *Annual Review of Immunology* **27**(1): 393-422.
62. Costa Conceição, E., M. L. Lopes, et al. (2017). "Genetic diversity of Mycobacterium tuberculosis from Pará, Brazil, reveals a higher frequency of the lineage East African India in South America." *Infection Genetics Evolution* **submitted**.
63. D. Thierry, M.D.Cave, et al. (1990). "IS6110, an IS-like element of Mycobacterium tuberculosis complex.pdf>." *Nucleic Acids Research*, **Vol. 18, No. 1**.
64. Daniel, T. M. (2005). "Selman Abraham Waksman and the discovery of Streptomycin.pdf>." *INT J TUBERC LUNG DIS* **9**(2): 120-122.
65. Daniel, T. M. (2006). "The history of tuberculosis." *Respiratory Medicine* **100**(11): 1862-1870.
66. Daniel, T. M. (2015). "Jean-Antoine Villemin and the infectious nature of tuberculosis." *The International Journal of Tuberculosis and Lung Disease* **19**(3): 267-268.
67. Dantas, N., P. Suffys, et al. (2015). "Multidrug-resistant Mycobacterium tuberculosis Genetic Diversity in Minas Gerais state, Brazil." *BMC Infectious Diseases* **15**: 306.
68. Dantas, N. G. T., P. N. Suffys, et al. (2017). "Correlation between the BACTEC MGIT 960 culture system with Genotype MTBDRplus and TB-SPRINT in multidrug resistant Mycobacterium tuberculosis clinical isolates from Brazil." *Mem Inst Oswaldo Cruz* **112**(11): 769-774.
69. David, H., V. Levy-Frebault, et al. (1989). *Méthodes de laboratoire pour Mycobactériologie clinique*. Paris, Institut Pasteur: 1-87.
70. de Jong, B. C., P. C. Hill, et al. (2008). "Progression to active tuberculosis, but not transmission, varies by Mycobacterium tuberculosis lineage in The Gambia." *J Infect Dis* **198**(7): 1037-1043.
71. Demay, C., B. Liens, et al. (2012). "SITVITWEB – A publicly available international multimer database for studying Mycobacterium tuberculosis genetic diversity and molecular epidemiology." *Infect Genet Evol* **12**(4): 755-766.
72. Desikan, P. (2013). "Sputum smear microscopy in tuberculosis: Is it still relevant?." *Indian J Med Res.* **137**(3): 442–444.
73. Dick Van Soolingen, Lishi Qian, et al. (1995). "Predominance of a single Genotype of Mycobacterium tuberculosis in countries of East Asia.pdf> " *JOURNAL OF CLINICAL MICROBIOLOGY*: 3234-3238.
74. Domenech, P., J. Zou, et al. (2017). "Unique Regulation of the DosR Regulon in the Beijing Lineage of Mycobacterium tuberculosis." *Journal of Bacteriology* **199**(2): e00696-00616.
75. Dominguez, J., E. C. Boettger, et al. (2016). "Clinical implications of molecular drug resistance testing for Mycobacterium tuberculosis: a TBNET/RESIST-TB consensus statement." *Int J Tuberc Lung Dis* **20**(1): 24-42.
76. Donoghue, H. D. (2008). "Paleomicrobiology of Tuberculosis.pdf>." *Raoult D, Drancourt M, eds. Paleomicrobiology—past human infections*. Berlin:Springer-Verlag, 2008;: 75–97.
77. Donoghue, H. D. (2009). "Human tuberculosis--an ancient disease, as elucidated by ancient microbial biomolecules." *Microbes Infect* **11**(14-15): 1156-1162.
78. Donoghue, H. D. (2011). "Insights gained from palaeomicrobiology into ancient and modern tuberculosis." *Clinical Microbiology and Infection* **17**(6): 821-829.
79. Donoghue HD, Spigelman M, et al. (2004). "Tuberculosis: from prehistory to Robert Koch, as revealed by ancient DNA." *Lancet Infect Dis* **4**(9): 584-592.
80. Dorman, S. E., S. G. Schumacher, et al. (2018). "Xpert MTB/RIF Ultra for detection of Mycobacterium tuberculosis and rifampicin resistance: a prospective multicentre diagnostic accuracy study." *The Lancet Infectious Diseases* **18**(1): 76-84.

81. Dormans, J., M. Burger, et al. (2004). "Correlation of virulence, lung pathology, bacterial load and delayed type hypersensitivity responses after infection with different *Mycobacterium tuberculosis* genotypes in a BALB/c mouse model." Clin Exp Immunol **137**: 460-468.
82. Drancourt, M. and D. Raoult (2005). "Palaeomicrobiology: current issues and perspectives." Nat Rev Microbiol **3**(1): 23-35.
83. Duarte, T. A., J. S. Nery, et al. (2017). "A systematic review of East African-Indian family of *Mycobacterium tuberculosis* in Brazil." Braz J Infect Dis.
84. Dunbar, S. A. (2006). "Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection." Clin Chim Acta **363**(1-2): 71-82.
85. Easterbrook, P. J., A. Gibson, et al. (2004). "High rates of clustering of strains causing tuberculosis in Harare, Zimbabwe: a molecular epidemiological study." J Clin Microbiol **42**(10): 4536-4544.
86. Eilertson, B., F. Maruri, et al. (2014). "High proportion of heteroresistance in *gyrA* and *gyrB* in fluoroquinolone-resistant *Mycobacterium tuberculosis* clinical isolates." Antimicrob Agents Chemother **58**(6): 3270-3275.
87. Eldholm, V., J. Monteserin, et al. (2015). "Four decades of transmission of a multidrug-resistant *Mycobacterium tuberculosis* outbreak strain." Nat Commun **6**: 7119.
88. Fang, Z., C. Doig, et al. (1999). "IS6110-mediated deletions of wild-type chromosomes of *Mycobacterium tuberculosis*." J Bacteriol **181**(3): 1014-1020.
89. Ferdinand, S., C. Sola, et al. (2005). "A study of spoligotyping-defined *Mycobacterium tuberculosis* clades in relation to the origin of peopling and the demographic history in Madagascar." Infect Genet Evol **5**(4): 340-348.
90. Fernandes, V., P. Triska, et al. (2015). "Genetic stratigraphy of key demographic events in Arabia." PLoS One **10**(3): e0118625.
91. Feuerriegel, S., V. Schleusener, et al. (2015). "PhyResSE: a Web Tool Delineating *Mycobacterium tuberculosis* Antibiotic Resistance and Lineage from Whole-Genome Sequencing Data." J Clin Microbiol **53**(6): 1908-1914.
92. Fiebig, L., T. A. Kohl, et al. (2017). "A joint cross-border investigation of a cluster of multidrug-resistant tuberculosis in Austria, Romania and Germany in 2014 using classic, genotyping and whole genome sequencing methods: lessons learnt." Euro Surveill **22**(2).
93. Filliol, I., J. R. Driscoll, et al. (2003). "Snapshot of moving and expanding clones of *Mycobacterium tuberculosis* and their global distribution assessed by spoligotyping in an international study." J Clin Microbiol **41**(5): 1963-1970.
94. Filliol, I., S. Ferdinand, et al. (2002). "Spoligotyping and IS6110-RFLP typing of *Mycobacterium tuberculosis* from French Guiana : a comparison of results with international databases underlines interregional transmission from neighboring countries." Res. Microbiol. **153**: 81-88.
95. Filliol, I., A. S. Motiwala, et al. (2006). "Global Phylogeny of *Mycobacterium tuberculosis* Based on Single Nucleotide Polymorphism (SNP) Analysis: Insights into Tuberculosis Evolution, Phylogenetic Accuracy of Other DNA Fingerprinting Systems, and Recommendations for a Minimal Standard SNP Set." Journal of Bacteriology **188**(2): 759-772.
96. Filliol, I., A. S. Motiwala, et al. (2006). "Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set." J Bacteriol **188**(2): 759-772.
97. Filliol, I., C. Sola, et al. (2000). "Detection of a previously unamplified spacer within the DR locus of *Mycobacterium tuberculosis*: Epidemiological implications." J. Clin. Microbiol. **38**(3): 1231-1234.

98. Fletcher Helen A. , Donoghue Helen D. , et al. (2003). "Molecular analysis of Mycobacterium tuberculosis DNA from a family of 18th century Hungarians." Microbiology **149**(1): 143-151.
99. Flores-Treviño, S., R. Morfín-Otero, et al. (2015). "Genetic Diversity of Mycobacterium tuberculosis from Guadalajara, Mexico and Identification of a Rare Multidrug Resistant Beijing Genotype." PLoS One **10**(2): e0118095.
100. Fomukong, N. G., T. H. Tang, et al. (1994). "Insertion sequence typing of Mycobacterium tuberculosis: characterization of a widespread subtype with a single copy of IS6110." Tubercle and Lung Disease **75**: 435-440.
101. Gagneux, S. (2009). "Fitness cost of drug resistance in Mycobacterium tuberculosis." Clin Microbiol Infect **15 Suppl 1**: 66-68.
102. Gagneux, S., K. DeRiemer, et al. (2006). "Variable host-pathogen compatibility in Mycobacterium tuberculosis." Proceedings of the National Academy of Sciences **103**(8): 2869-2873.
103. Gagneux, S., K. DeRiemer, et al. (2006). "Variable host-pathogen compatibility in Mycobacterium tuberculosis." Proc Natl Acad Sci U S A **103**(8): 2869-2873.
104. Gagneux, S. and P. M. Small (2007). "Global phylogeography of Mycobacterium tuberculosis and implications for tuberculosis product development." Lancet Infect Dis **7**(5): 328-337.
105. Georghiou, S. B., M. Magana, et al. (2012). "Evaluation of genetic mutations associated with Mycobacterium tuberculosis resistance to amikacin, kanamycin and capreomycin: a systematic review." PLoS One **7**(3): e33275.
106. Ghodbane, R., D. Raoult, et al. (2014). "Dramatic reduction of culture time of Mycobacterium tuberculosis." Scientific Reports **4**(1).
107. Gillespie, S. H., A. M. Crook, et al. (2014). "Four-Month Moxifloxacin-Based Regimens for Drug-Sensitive Tuberculosis." New England Journal of Medicine **371**(17): 1577-1587.
108. Gilpin, C., A. Korobitsyn, et al. (2016). "Current tools available for the diagnosis of drug-resistant tuberculosis." Ther Adv Infect Dis **3**(6): 145-151.
109. Glynn, J. R., S. Alghamdi, et al. (2010). "Changes in Mycobacterium tuberculosis genotype families over 20 years in a population-based study in Northern Malawi." PLoS One **5**(8): e12259.
110. Glynn, J. R., A. C. Crampin, et al. (2008). "Determinants of cluster size in large, population-based molecular epidemiology study of tuberculosis, northern Malawi." Emerg Infect Dis **14**(7): 1060-1066.
111. Gomgnimbou, M. K., E. Abadia, et al. (2012). "<< Spoligotyping >> a DPO-based direct-hybridization assay for TB control on a multianalyte microbead-based hybridization system." J Clin Microbiol **50**(10): 3172-3179.
112. Gomgnimbou, M. K., I. Hernandez-Neuta, et al. (2013). "'TB-SPRINT: TUBERCULOSIS-SPOLIGO-RIFAMPIN-ISONIAZID TYPING"; an All-in-One assay technique for surveillance and control of multi-drug resistant tuberculosis on Luminex® devices." J Clin Microbiol **51**(11): 3527-3534.
113. Gomgnimbou, M. K., G. Refrégier, et al. (2013). High Throughput combin methods for resistanc emergence analysis, diagnostic and Epidemiology: the multi-drug resistant tuberculosis paradigm. 10th International Meeting on Microbial Epidemiological Markers, Institut Pasteur.
114. Gravet, A., N. Souillard, et al. (2011). "La culture et l'antibiogramme de mycobactéries sur automate VersaTREK." Pathologie Biologie **59**(1): 32-38.
115. Grissa, I., P. Bouchon, et al. (2008). "On-line resources for bacterial micro-evolution studies using MLVA or CRISPR typing." Biochimie **90**(4): 660-668.
116. Grissa, I., G. Vergnaud, et al. (2007). "The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats." BMC Bioinformatics **8**: 172.

117. Groenheit, R., S. Ghebremichael, et al. (2011). "The Guinea-Bissau Family of Mycobacterium tuberculosis Complex Revisited." *PLoS One* **6**(4): e18601.
118. Guerra-Assuncao, J. A., A. C. Crampin, et al. (2015). "Large-scale whole genome sequencing of M. tuberculosis provides insights into transmission in a high prevalence area." *Elife* **4**.
119. Han, S. J., T. Song, et al. (2015). "Complete genome sequence of Mycobacterium tuberculosis K from a Korean high school outbreak, belonging to the Beijing family." *Standards in Genomic Sciences* **10**(1).
120. Hancock, J. M. (1995). "The contribution of slippage-like processes to genome evolution." *J. Mol. Evol.* **41**(6): 1038-1047.
121. Henar Alonso, Sofía Samper, et al. (2013). "Mapping IS6110 in high-copy number Mycobacterium tuberculosis strains shows specific insertion points in the Beijing genotype.pdf>." *BMC Genomics*.
122. Herranz, M., I. Pole, et al. (2017). "Mycobacterium tuberculosis Acquires Limited Genetic Diversity in Prolonged Infections, Reactivations and Transmissions Involving Multiple Hosts." *Front Microbiol* **8**: 2661.
123. Hershberg, R., M. Lipatov, et al. (2008). "High functional diversity in Mycobacterium tuberculosis driven by genetic drift and human demography." *PLoS Biol* **6**(12): e311.
124. Hershkovitz Israel , Donoghue Helen D. , et al. (2008). "Detection and Molecular Characterization of 9000-Year-Old Mycobacterium tuberculosis from a Neolithic Settlement in the Eastern Mediterranean." *PLoS One*.
125. Hess, A. S., M. Shardell, et al. (2012). "Methods and recommendations for evaluating and reporting a new diagnostic test." *Eur J Clin Microbiol Infect Dis* **31**(9): 2111-2116.
126. Heyderman, R. S., M. Goyal, et al. (1998). "Pulmonary tuberculosis in Harare, Zimbabwe: analysis by spoligotyping." *Thorax* **53**(5): 346-350.
127. Hillemann, D., S. Rusch-Gerdes, et al. (2009). "Feasibility of the GenoType MTBDRsl assay for fluoroquinolone, amikacin-capreomycin, and ethambutol resistance testing of Mycobacterium tuberculosis strains and clinical specimens." *J Clin Microbiol* **47**(6): 1767-1772.
128. Hiroshi Nakaoka, Lovett Lawson, et al. (2006). "Risk for Tuberculosis among Children.pdf." *Emerging Infectious Diseases* **Vol. 12, No. 9**: 1383-1388.
129. Howard E. Takiff, Leiria Salazar, et al. (1994). "Cloning and Nucleotide Sequence of Mycobacterium tuberculosis gyrA and gyrB Genes and Detection of Quinolone.pdf>" *American Society for Microbiology*.
130. Ida Parwati, Reinout van Crevel, et al. (2010). "<Possible underlying mechanism for successful emergence of the MTB Beijing genotype strains.pdf>." *Lancet Infect Dis* **10**103-111.
131. Iwamoto, T., S. Yoshida, et al. (2007). "Hypervariable loci that enhance the discriminatory ability of newly proposed 15-loci and 24-loci variable-number tandem repeat typing method on Mycobacterium tuberculosis strains predominated by the Beijing family." *FEMS Microbiology Letters* **270**(1): 67-74.
132. Jindani, A., G. Borgulya, et al. (2016). "A randomised Phase II trial to evaluate the toxicity of high-dose rifampicin to treat pulmonary tuberculosis." *The International Journal of Tuberculosis and Lung Disease* **20**(6): 832-838.
133. Jindani, A., T. S. Harrison, et al. (2014). "High-Dose Rifapentine with Moxifloxacin for Pulmonary Tuberculosis." *New England Journal of Medicine* **371**(17): 1599-1608.
134. Jordi Gómez i Prat and S. M. M. d. Souza (2003). "Prehistoric Tuberculosis in America, Adding comments to a Literature Review .pdf>." *Mem Inst Oswaldo Cruz, Rio de Janeiro* **98**: 151-159.
135. Joseph, B. V., S. Soman, et al. (2013). "Molecular epidemiology of Mycobacterium tuberculosis isolates from Kerala, India using IS6110-RFLP, spoligotyping and MIRU-VNTRs." *Infect Genet Evol* **16**: 157-164.

136. Judith R. Glynn, Jennifer Whiteley, et al. (2002). "Worldwide Occurrence of Beijing/W Strains of Mycobacterium tuberculosis: A Systematic Review .pdf>." Emerging Infectious Diseases **Vol. 8, No. 8**.
137. Judith R. Glynn, Kristin Kremer, et al. (2006). "<Beijing-W genotype Mycobacterium tuberculosis and drug resistance.pdf>." Emerging Infectious Diseases **Vol. 12, No. 5**.
138. Kambli, P., K. Ajbani, et al. (2015). "Defining multidrug-resistant tuberculosis: correlating GenoType MTBDRplus assay results with minimum inhibitory concentrations." Diagn Microbiol Infect Dis **82(1)**: 49-53.
139. Kamerbeek, J., L. Schouls, et al. (1997). "Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology." J Clin Microbiol **35(4)**: 907-914.
140. Kang, H. Y., T. Wada, et al. (2010). "Phylogeographical particularity of the Mycobacterium tuberculosis Beijing family in South Korea based on international comparison with surrounding countries." Journal of Medical Microbiology **59(10)**: 1191-1197.
141. Kay, G. L., M. J. Sergeant, et al. (2015). "Eighteenth-century genomes show that mixed infections were common at time of peak tuberculosis in Europe." Nature Communications **6(1)**.
142. Kehr, J. (2018). "'Exotic no more': Tuberculosis, public debt and global health in Berlin." Glob Public Health **13(3)**: 369-382.
143. Kim, K., J. S. Yang, et al. (2018). "Detection of resistance to fluoroquinolones and injectable drugs among antituberculosis drugs by allele-specific primer extension on a microsphere-based platform." J Microbiol Methods **144**: 111-116.
144. Klotoe, B. J., B. Molina-Moya, et al. (2018). "TB-EFI, a novel 18-Plex microbead-based method for prediction of second-line drugs and ethambutol resistance in Mycobacterium tuberculosis complex." J Microbiol Methods.
145. Knut Lönnroth, Kenneth G Castro, et al. (2010). "Tuberculosis control and elimination 2010–50: cure, care, and social development." Lancet **375**: : 1814-1829.
146. Kohanski, M. A., D. J. Dwyer, et al. (2010). "How antibiotics kill bacteria: from targets to networks." Nature Reviews Microbiology **8(6)**: 423-435.
147. Koul, A., L. Vranckx, et al. (2008). "Diarylquinolines Are Bactericidal for Dormant Mycobacteria as a Result of Disturbed ATP Homeostasis." Journal of Biological Chemistry **283(37)**: 25273-25280.
148. Kovalev, S. Y., E. Y. Kamaev, et al. (2005). "Genetic analysis of mycobacterium tuberculosis strains isolated in Ural region, Russian Federation, by MIRU-VNTR genotyping." Int J Tuberc Lung Dis **9(7)**: 746-752.
149. Kremer, K., J. R. Glynn, et al. (2004). "Definition of the Beijing/W lineage of Mycobacterium tuberculosis on the basis of genetic markers." J Clin Microbiol **42(9)**: 4040-4049.
150. Kremer, K., M. J. van-der-Werf, et al. (2009). "Vaccine-induced immunity circumvented by typical Mycobacterium tuberculosis Beijing strains." Emerg Infect Dis **15(2)**: 335-339.
151. Kubica, T., R. Agzamova, et al. (2005). "The Beijing genotype is a major cause of drug-resistant tuberculosis in Kazakhstan." Int J Tuberc Lung Dis **9(6)**: 646-653.
152. Kurepina, N., E. Likhoshvay, et al. (2005). "Targeted Hybridization of IS6110 Fingerprints Identifies the W-Beijing Mycobacterium tuberculosis Strains among Clinical Isolates." Journal of Clinical Microbiology **43(5)**: 2148-2154.
153. Lacoma, A., B. Molina-Moya, et al. (2015). "Pyrosequencing for rapid detection of Mycobacterium tuberculosis second-line drugs and ethambutol resistance." Diagn Microbiol Infect Dis **83(3)**: 263-269.
154. Lasunskaja, E., S. C.M.Ribeiro, et al. (2010). "Emerging multidrug resistant Mycobacterium tuberculosis strains of the Beijing genotype circulating in Russia express a pattern of biological properties associated with enhanced virulence" Microbes and Infection **Volume 12**(Issue 6): 467-475.

155. Lecuit, M. and M. Eloit (2015). "The potential of whole genome NGS for infectious disease diagnosis." *Expert Rev Mol Diagn* **15**(12): 1517-1519.
156. Lee, A. S. G., Lynn L.H. Tang, et al. (2002). "Characterization of pyrazinamide and ofloxacin resistance among drug resistance Mycobacterium tuberculosis isolates from Singapore .pdf>." *Int J Infect Dis* **6**: 48-51
157. Lee, S. H. (2016). "Tuberculosis Infection and Latent Tuberculosis." *Tuberc Respir Dis (Seoul)* **79**(4): 201-206.
158. Lee, S. H., H. B. Choi, et al. (2015). "Detection of first-line anti-tuberculosis drug resistance mutations by allele-specific primer extension on a microsphere-based platform." *Ann Lab Med* **35**(5): 487-493.
159. Lessem, E. (2017). "The Tuberculosis Diagnostics Pipeline.pdf>." *WHO Pipeline-Report-TB-Diagnostics*.
160. Lisdawati, V., N. Puspendari, et al. (2015). "Molecular epidemiology study of Mycobacterium tuberculosis and its susceptibility to anti-tuberculosis drugs in Indonesia." *BMC Infect Dis* **15**: 366.
161. Liu, Q., T. Luo, et al. (2016). "Genetic features of Mycobacterium tuberculosis modern Beijing sublineage." *Emerging Microbes & Infections* **5**(2): e14-e14.
162. Louise J Pankhurst, Carlos del Ojo Elias, et al. (2016). "Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study." *Lancet Respir Med*. 2016 Jan **4**(1): 49–58.
163. Loxton, A. G., J. K. Knaul, et al. (2017). "Safety and Immunogenicity of the Recombinant Mycobacterium bovis BCG Vaccine VPM1002 in HIV-Unexposed Newborn Infants in South Africa." *Clinical and Vaccine Immunology* **24**(2).
164. Luo, Tao Luo, et al. (2015). "Southern East Asian origin and coexpansion of mycobacterium tuberculosis beijing family with han Chinese.pdf>." *PNAS*.
165. Madhukar Pai, Shriprakash Kalantri, et al. (2006). "New tools and emerging technologies for the diagnosis of tuberculosis, latent tuberculosis.pdf>." *Expert Rev. Mol. Diagn.* **6**(3)(1473-7159): 413-422.
166. Maeda, S., N. T. L. Hang, et al. (2014). "Mycobacterium tuberculosis strains spreading in Hanoi, Vietnam: Beijing sublineages, genotypes, drug susceptibility patterns, and host factors." *Tuberculosis* **94**(6): 649-656.
167. Malik, S., M. Willby, et al. (2012). "New insights into fluoroquinolone resistance in Mycobacterium tuberculosis: functional genetic analysis of gyrA and gyrB mutations." *PLoS One* **7**(6): e39754.
168. Manca, C., M. B. Reed, et al. (2004). "Differential monocyte activation underlies strain-specific Mycobacterium tuberculosis pathogenesis." *Infect Immun* **72**(9): 5511-5514.
169. Manson, A. L., K. A. Cohen, et al. (2017). "Genomic analysis of globally diverse Mycobacterium tuberculosis strains provides insights into the emergence and spread of multidrug resistance." *Nat Genet* **49**(3): 395-402.
170. Martina Casenghi (2006). "Development of new drugs for TB Chemotherapy.pdf>." *Book*.
171. Mattocks, C. J., G. Watkins, et al. (2010). "Interlaboratory diagnostic validation of conformation-sensitive capillary electrophoresis for mutation scanning." *Clin Chem* **56**(4): 593-602.
172. McGinn, S., D. Bauer, et al. (2016). "New technologies for DNA analysis--a review of the READNA Project." *N Biotechnol* **33**(3): 311-330.
173. McGrath, J. W. (1988). "Social networks of disease spread in the lower Illinois valley: a simulation approach." *Am J Phys Anthropol* **77**(4): 483-496.
174. McKenna, L. (2017). "The Tuberculosis Diagnostics and Treatment Pipeline for Children.pdf>." *WHO PIPELINE REPORT*.

175. McNerney, R., J. Cunningham, et al. (2015). "New tuberculosis diagnostics and rollout." *Int J Infect Dis* **32**: 81-86.
176. Merker, M., C. Blin, et al. (2015). "Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage." *Nat Genet* **47**(3): 242-249.
177. Merker, M., C. Blin, et al. (2015). "Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage." *Nature Genetics* **47**(3): 242-249.
178. Merle, C. S., K. Fielding, et al. (2014). "A Four-Month Gatifloxacin-Containing Regimen for Treating Tuberculosis." *New England Journal of Medicine* **371**(17): 1588-1598.
179. Mestre, O., T. Luo, et al. (2011). "Phylogeny of Mycobacterium tuberculosis Beijing Strains Constructed from Polymorphisms in Genes Involved in DNA Replication, Recombination and Repair." *PLoS One* **6**(1): e16020.
180. Millet, J., C. Miyagi-Shiohira, et al. (2012). "High-resolution MIRU-VNTRs typing reveals the unique nature of Mycobacterium tuberculosis Beijing genotype in Okinawa, Japan." *Infection, Genetics and Evolution* **12**(4): 637-641.
181. Milstein, M., L. Lecca, et al. (2016). "Evaluation of high-dose rifampin in patients with new, smear-positive tuberculosis (HIRIF): study protocol for a randomized controlled trial." *BMC Infectious Diseases* **16**(1).
182. Mina Ebrahimi Rad, Pablo Bifani, et al. (2003). "Mutations in putative mutator genes of mycobacterium tuberculosis strains of the w-beijing family.pdf>."
183. Mokrousov, I. (2005). "Origin and primary dispersal of the Mycobacterium tuberculosis Beijing genotype: Clues from human phylogeography." *Genome Research* **15**(10): 1357-1364.
184. Mokrousov, I. (2006). "Mycobacterium tuberculosis Beijing genotype and mycobacterial interspersed repetitive unit typing." *J Clin Microbiol* **44**(4): 1614; author reply 1614-1615.
185. Mokrousov, I. (2008). "Genetic geography of Mycobacterium tuberculosis Beijing genotype: a multifacet mirror of human history?" *Infect Genet Evol* **8**(6): 777-785.
186. Mokrousov, I. (2013). "Insights into the origin, Emergence, and Current spread of a Successful Russian clone of Mycobacterium tuberculosis.pdf>." *Clin. Microbiol.Rev* **26**: 342-360.
187. Mokrousov, I. (2015). "Molecular structure of Mycobacterium tuberculosis population in Russia and its interaction with neighboring countries." *Int J Mycobacteriol* **4 Suppl 1**: 56-57.
188. Mokrousov, I. (2018). "On sunspots, click science and molecular iconography." *Tuberculosis (Edinb)* **110**: 91-95.
189. Mokrousov, I., E. Chernyaeva, et al. (2018). "Rapid Assay for Detection of the Epidemiologically Important Central Asian/Russian Strain of the Mycobacterium tuberculosis Beijing Genotype." *J Clin Microbiol* **56**(2).
190. Mokrousov, I., H. M. Ly, et al. (2005). "Origin and primary dispersal of the Mycobacterium tuberculosis Beijing genotype: clues from human phylogeography." *Genome Res* **15**(10): 1357-1364.
191. Mokrousov, I., O. Narvskaya, et al. (2008). "Mycobacterium tuberculosis Beijing Genotype in Russia: in Search of Informative VNTR Loci." *J Clin Microbiol*.
192. Mokrousov, I., O. Narvskaya, et al. (2012). "Russian "Successful" Clone B0/W148 of Mycobacterium tuberculosis Beijing Genotype: a Multiplex PCR Assay for Rapid Detection and Global Screening." *Journal of Clinical Microbiology* **50**(11): 3757-3759.
193. Mokrousov, I., A. Vyazovaya, et al. (2016). "Latin-American-Mediterranean lineage of Mycobacterium tuberculosis: Human traces across pathogen's phylogeography." *Mol Phylogenet Evol* **99**: 133-143.
194. Molina-Moya, B., M. Agonafir, et al. (2018). "Microbead-based spoligotyping of Mycobacterium tuberculosis from Ziehl-Neelsen-stained microscopy preparations in Ethiopia." *Sci Rep* **8**(1): 3987.

195. Molina-Moya, B., M. K. Gomgnimbou, et al. (2017). "Molecular Characterization of Mycobacterium tuberculosis Strains with TB-SPRINT." *Am J Trop Med Hyg* **97**(3): 806-809.
196. Molina-Moya, B., A. Lacoma, et al. (2015). "Diagnostic accuracy study of multiplex PCR for detecting tuberculosis drug resistance." *J Infect* **71**(2): 220-230.
197. Müller, B., S. Dürr, et al. (2013). "Zoonotic Mycobacterium bovis-induced Tuberculosis in Humans." *Emerging Infectious Diseases* **19**(6): 899-908.
198. Murase, Y., S. Mitarai, et al. (2008). "Promising loci of variable numbers of tandem repeats for typing Beijing family Mycobacterium tuberculosis." *Journal of Medical Microbiology* **57**(7): 873-880.
199. N. E. Kurepina, S. Sreevatsan, et al. (1998). "Characterization of the phylogenetic distribution and chromosomal insertion sites of five IS6110 elements in Mycobacterium tuberculosis: non-random integration in the dnaA – dnaN region.pdf>." *Tubercle and Lung Disease* **79**(1): 31-42.
200. N. Rastogi, K.S. Goh, et al. (1989). "DRUG SUSCEPTIBILITY TESTING IN TUBERCULOSIS:A COMPARISON OF THE PROPORTION METHODS USING LOWENSTEIN4ENSEN, MIDDLEBROOK 7H10 AND 71111 AGAR MEDIA AND A RADIOMETRIC METHOD and a radiometric method.pdf>." *Res Microbiol*: 405-417.
201. Nadine Honoré and S. T. Cole (1994). "Streptomycin Resistance in Mycobacteria.pdf>." *ANTIMICROBIAL AGENTS AND CHEMOTHERAPY*: 238-242
202. Namburete, E. I., I. Tivane, et al. (2016). "Drug-resistant tuberculosis in Central Mozambique: the role of a rapid genotypic susceptibility testing." *BMC Infect Dis* **16**: 423.
203. Nayak, S. and B. Acharjya (2012). "Mantoux test and its interpretation." *Indian Dermatology Online Journal* **3**(1): 2.
204. Newitt, M. (1995). *A history of Mozambique*, Indiana University Press.
205. Nieuwenhuizen, N. E., P. S. Kulkarni, et al. (2017). "The Recombinant Bacille Calmette–Guérin Vaccine VPM1002: Ready for Clinical Efficacy Testing." *Frontiers in Immunology* **8**.
206. Nimmo, C., M. Lipman, et al. (2015). "Shortening treatment of tuberculosis: lessons from fluoroquinolone trials." *The Lancet Infectious Diseases* **15**(2): 141-143.
207. Ocheretina, O., Y. M. Merveille, et al. (2013). "Use of Luminex MagPlex(R) Magnetic Microspheres for High-Throughput Spoligotyping of M. tuberculosis Isolates in Port-au-Prince, Haiti." *J Clin Microbiol*.
208. OMS (2011). "Use of tuberculosis IGRA in low-and middle-income countries.pdf>." *WHO Library Cataloguing-in-Publication Data*.
209. Operario, D. J., A. F. Koepfel, et al. (2017). "Prevalence and extent of heteroresistance by next generation sequencing of multidrug-resistant tuberculosis." *PLoS One* **12**(5): e0176522.
210. Overton, K., R. Varma, et al. (2018). "Comparison of Interferon- γ Release Assays and the Tuberculin Skin Test for Diagnosis of Tuberculosis in Human Immunodeficiency Virus: A Systematic Review." *Tuberculosis and Respiratory Diseases* **81**(1): 59.
211. Pai M, Behr MA, et al. (2016). "Tuberculosis." *Nat Rev Dis Primers*.
212. Pai, M. and R. O'Brien (2008). "New Diagnostics for latent and Active Tuberculosis, State of the Art and Future Prospects_ Abstract.pdf>." *Thieme E-Journals - Seminars in Respiratory and Critical Care Medicine*.
213. Palaci Moises, Suely Yoko Mizuka Ueki, et al. (1996). "Evaluation of Mycobacteria growth indicator tube for recovery and drug susceptibility testing of M.tuberculosis isolates from respiratory specimens, JCM.pdf>." *JCM*.
214. Parrish, N. M. and K. C. Carroll (2010). "Role of the Clinical Mycobacteriology Laboratory in Diagnosis and Management of Tuberculosis in Low-Prevalence Settings." *Journal of Clinical Microbiology* **49**(3): 772-776.
215. Perez-Garcia, F., M. J. Ruiz-Serrano, et al. (2017). "Diagnostic performance of Anyplex II MTB/MDR/XDR for detection of resistance to first and second line drugs in Mycobacterium tuberculosis." *J Microbiol Methods* **139**: 74-78.

216. Perez-Lago, L., S. Izco, et al. (2017). "A novel strategy based on genomics and specific PCR reveals how a multidrug-resistant *Mycobacterium tuberculosis* strain became prevalent in Equatorial Guinea 15 years after its emergence." *Clin Microbiol Infect* **23**(2): 92-97.
217. Pfyffer, G. E., P. R. Murray, et al. (2007). "Mycobacterium: General characteristics, laboratory detection, and staining procedures." *Eds Manual of Clinical Microbiology., 9th Edition, American Society for Microbiology* 543-572.
218. Pfeiffer, C. (2005). "Protein expression by a Beijing strain differs from that of another clinical isolate and *Mycobacterium tuberculosis* H37Rv." *Proc. Natl. Acad. Sci. USA* **151**(4): 1139-1150.
219. Phelan, J., F. Coll, et al. (2016). "Mycobacterium tuberculosis whole genome sequencing and protein structure modelling provides insights into anti-tuberculosis drug resistance." *BMC Med* **14**: 31.
220. Phelan, J., D. M. O'Sullivan, et al. (2016). "The variability and reproducibility of whole genome sequencing technology for detecting resistance to anti-tuberculous drugs." *Genome Med* **8**(1): 132.
221. Plikaytis, B. B., J. L. Marden, et al. (1994). "Multiplex PCR assay for the Multi-drug-resistant strain W of *Mycobacterium tuberculosis*." *J. Clin. Microbiol.* **32**: 1542-1546.
222. Plikaytis, B. B., J. L. Marden, et al. (1994). "Multiplex PCR assay specific for the multidrug-resistant strain W of *Mycobacterium tuberculosis*." *J Clin Microbiol* **32**(6): 1542-1546.
223. R. Johnson, R. Warren, et al. (2006). "<An outbreak of drug-resistant tuberculosis caused by a Beijing strain in the Western Cape, South Africa .pdf>." *INT J TUBERC LUNG DIS* **10**(12):: 1412-1414.
224. Rad, M. E., P. Bifani, et al. (2003). "Mutations in putative mutator genes of *Mycobacterium tuberculosis* strains of the W-Beijing family." *Emerg Infect Dis* **9**(7): 838-845.
225. Ramazanzadeh, R. and K. Sayhemiri (2014). "Prevalence of Beijing family in *Mycobacterium tuberculosis* in world population: Systematic Review and Meta-Analysis." *International Journal of Mycobacteriology* **3**(1): 41.
226. Reed, Michael B. Reed, et al. (2004). "A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response.pdf>." *NATURE* **VOL 431**.
227. Reed, M. B., S. Gagneux, et al. (2007). "The W-Beijing Lineage of *Mycobacterium tuberculosis* Overproduces Triglycerides and Has the DosR Dormancy Regulon Constitutively Upregulated." *Journal of Bacteriology* **189**(7): 2583-2589.
228. Refregier, G., E. Abadia, et al. (2016). "Turkish and Japanese *Mycobacterium tuberculosis* sublineages share a remote common ancestor." *Infect Genet Evol* **45**: 461-473.
229. Reiling, N., S. Homolka, et al. (2013). "Clade-specific virulence patterns of *Mycobacterium tuberculosis* complex strains in human primary macrophages and aerogenically infected mice." *MBio* **4**(4).
230. REPORT (2017). "Molecular typing for surveillance of multidrug-resistant tuberculosis in the EU/EEA." *SURVEILLANCE REPORT*.
231. Ribeiro, S. C. M., L. L. Gomes, et al. (2014). "Mycobacterium tuberculosis Strains of the Modern Sublineage of the Beijing Family Are More Likely To Display Increased Virulence than Strains of the Ancient Sublineage." *Journal of Clinical Microbiology* **52**(7): 2615-2624.
232. Rout, L. B. J. (1976). "Race and Slavery in Brazil." *The Wilson Quarterly* **1**(1): 73-89.
233. Roychowdhury, T., S. Mandal, et al. (2015). "Analysis of IS6110 insertion sites provide a glimpse into genome evolution of *Mycobacterium tuberculosis*." *Sci Rep* **5**: 12567.
234. Russell, D. G., P.-J. Cardona, et al. (2009). "Foamy macrophages and the progression of the human tuberculosis granuloma." *Nature Immunology* **10**(9): 943-948.
235. Ryoo, S. W., Young Kil Park, et al. (2007). "Comparative proteomic analysis of virulent Korean M. tuberculosis K-strain with other mycobacteria strain following infection of U-937 macrophage.pdf>." *The Journal of Microbiology*; p. 268-271.

236. S., S., Ramachandran R. , et al. (2000). "Risk of development of tuberculosis in HIV-infected patients.pdf." INT J TUBERC LUNG DIS **4(9)**: 839-844.
237. Sahal Al Hajoj, Bright Varghese, et al. (2016). "Interferon Gamma Release Assay versus Tuberculin Skin Testing among Healthcare Workers of Highly Diverse Origin in a Moderate Tuberculosis Burden Country." PLoS One **11(5)**: e0154803.
238. Said, H. M., M. M. Kock, et al. (2012). "Molecular characterization and second-line antituberculosis drug resistance patterns of multidrug-resistant Mycobacterium tuberculosis isolates from the northern region of South Africa." J Clin Microbiol **50(9)**: 2857-2862.
239. Saifodine, A., J. Fyfe, et al. (2014). "Genetic diversity of Mycobacterium tuberculosis isolates obtained from patients with pulmonary tuberculosis in Beira city, Mozambique." Int J Mycobacteriol **3(2)**: 94-100.
240. Salo Wilmar L. , Aufderheide Arthur C. , et al. (1994). "Identification of Mycobacterium tuberculosis DNA in a pre-Columbian Peruvian mummy.pdf " Proc. Nati. Acad. Sci. USA **Vol. 91**: pp. 2091-2094.
241. Sanchez-Padilla, E., M. Merker, et al. (2015). "Detection of drug-resistant tuberculosis by Xpert MTB/RIF in Swaziland." N Engl J Med **372(12)**: 1181-1182.
242. Santos, P., E. R. D. Costa, et al. (2017). "Detection of tuberculosis drug resistance: a comparison by Mycobacterium tuberculosis MLPA assay versus Genotype(R)MTBDRplus." Mem Inst Oswaldo Cruz **112(6)**: 396-403.
243. Saukkonen, J. J., D. L. Cohn, et al. (2006). "An Official ATS Statement: Hepatotoxicity of Antituberculosis Therapy." American Journal of Respiratory and Critical Care Medicine **174(8)**: 935-952.
244. Schouten, J., C. McElgunn, et al. (2002). "Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification." Nucl. Acid. Res. **30(12)**: e57.
245. Schürch, A. C., K. Kremer, et al. (2011). "SNP/RD Typing of Mycobacterium tuberculosis Beijing Strains Reveals Local and Worldwide Disseminated Clonal Complexes." PLoS One **6(12)**: e28365.
246. Schurch, A. C., K. Kremer, et al. (2011). "Mutations in the regulatory network underlie the recent clonal expansion of a dominant subclone of the Mycobacterium tuberculosis Beijing genotype." Infect Genet Evol **11(3)**: 587-597.
247. Sengstake, S., N. Bablshvili, et al. (2014). "Optimizing multiplex SNP-based data analysis for genotyping of Mycobacterium tuberculosis isolates." BMC Genomics **15(1)**: 572.
248. Sharma, A., E. Bloss, et al. (2016). "Tuberculosis Caused by Mycobacterium africanum, United States, 2004-2013." Emerg Infect Dis **22(3)**: 396-403.
249. Shaun B. Walters and Bruce A. Hanna (1996). "Testing of susceptibility of M. tuberculosis to Isoniazid and Rifampin by Mycobacterium Growth Indicator Tube Method, .pdf>." JCM **Vol. 34, No. 6**: 1565-1567
250. Shitikov, E., S. Kolchenko, et al. (2017). "Evolutionary pathway analysis and unified classification of East Asian lineage of Mycobacterium tuberculosis." Scientific Reports **7(1)**.
251. Shitikov, E., S. Kolchenko, et al. (2017). "Evolutionary pathway analysis and unified classification of East Asian lineage of Mycobacterium tuberculosis." Sci Rep **7(1)**: 9227.
252. Sidibe, M. (2016). "Democratising the global health agenda: why we need France." Lancet **6736(16)**: 30060-30065.
253. Silva, M., F. Alshamali, et al. (2015). "60,000 years of interactions between Central and Eastern Africa documented by major African mitochondrial haplogroup L2." Sci Rep **5**: 12526.
254. Singh, J., M. M. Sankar, et al. (2015). "Genetic diversity and drug susceptibility profile of Mycobacterium tuberculosis isolated from different regions of India." J Infect **71(2)**: 207-219.

255. Singh, J. A., A. Bhan, et al. (2013). "Diagnosis of drug-resistant TB and provision of second-line TB treatment in India: some ethical considerations." Indian J Med Ethics **10**(2): 110-114.
256. Skiba, Y., I. Mokrousov, et al. (2015). "Molecular snapshot of Mycobacterium tuberculosis population in Kazakhstan: a country-wide study." Tuberculosis (Edinb) **95**(5): 538-546.
257. Smith, L. J., A. Bochkareva, et al. (2017). "Cmr is a redox-responsive regulator of DosR that contributes to M. tuberculosis virulence." Nucleic Acids Research **45**(11): 6600-6612.
258. Sohn (2009). "Induction of Cell Death in Human Macrophages by a Highly Virulent Korean Isolate of Mycobacterium tuberculosis and the Virulent Strain H37Rv -pdf>." Scandinavian Journal of Immunology.
259. Sola, C., I. Filliol, et al. (2001a). "Spoligotype database of Mycobacterium tuberculosis: biogeographic distribution of shared types and epidemiologic and phylogenetic perspectives." Emerg Infect Dis **7**(3): 390-396.
260. Sola, C., I. Filliol, et al. (2001). "Mycobacterium tuberculosis phylogeny reconstruction based on combined numerical analysis with IS1081, IS6110, VNTR, and DR-based spoligotyping suggests the existence of two new phylogeographical clades." J Mol Evol **53**(6): 680-689.
261. Somoskovi, A. (2015). "Novel laboratory diagnostic tests for tuberculosis and their potential role in an integrated and tiered laboratory network." Tuberculosis (Edinb) **95 Suppl 1**: S197-199.
262. Sonal S. Munsiff, Beth Nivin, et al. (2003). "<Persistence of a highly resistant strain of Tuberculosis in New York city using 1990-1999.pdf>." The Journal of Infectious Diseases ; **188**: 356-363.
263. Sreevatsan, S., X. Pan, et al. (1997). "Restricted structural gene polymorphism in the Mycobacterium tuberculosis complex indicates evolutionarily recent global dissemination." Proc Natl Acad Sci U S A **94**(18): 9869-9874.
264. Starks, A. M., A. Gumusboga, et al. (2009). "Mutations at embB codon 306 are an important molecular indicator of ethambutol resistance in Mycobacterium tuberculosis." Antimicrob Agents Chemother **53**(3): 1061-1066.
265. Steingart KR, Schiller I, et al. (2014). "Xpert MTB, Rif assay for pulmonary tuberculosis and rifampicin resistance in adults.pdf>."
266. Streicher, E. M., I. Bergval, et al. (2012). "Mycobacterium tuberculosis population structure determines the outcome of genetics-based second-line drug resistance testing." Antimicrob Agents Chemother **56**(5): 2420-2427.
267. Stucki, D., D. Brites, et al. (2016). "Mycobacterium tuberculosis lineage 4 comprises globally distributed and geographically restricted sublineages." Nat Genet **48**(12): 1535-1543.
268. Stucki, D., D. Brites, et al. (2016). "Mycobacterium tuberculosis lineage 4 comprises globally distributed and geographically restricted sublineages." Nat Genet.
269. Stucki, D., B. Malla, et al. (2012). "Two new rapid SNP-typing methods for classifying Mycobacterium tuberculosis complex into the main phylogenetic lineages." PLoS One **7**(7): e41253.
270. Sun, Z., J. Zhang, et al. (2008). "Comparison of gyrA gene mutations between laboratory-selected ofloxacin-resistant Mycobacterium tuberculosis strains and clinical isolates." International Journal of Antimicrobial Agents **31**(2): 115-121.
271. Supply, P., C. Allix, et al. (2006). "Proposal for Standardization of Optimized Mycobacterial Interspersed Repetitive Unit-Variable-Number Tandem Repeat Typing of Mycobacterium tuberculosis." Journal of Clinical Microbiology **44**(12): 4498-4510.
272. Supply, P., C. Allix, et al. (2006). "Proposal for Standardization of Optimized Mycobacterial Interspersed Repetitive Unit-Variable-Number Tandem Repeat Typing of Mycobacterium tuberculosis." J Clin Microbiol **44**(12): 4498-4510.

273. Supply, P., S. Lesjean, et al. (2001). "Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units." *J Clin Microbiol* **39**: 3563-3571.
274. Takiff, H. E., L. Salazar, et al. (1994). "Cloning and nucleotide sequence of *Mycobacterium tuberculosis gyrA* and *gyrB* genes and detection of quinolone resistance mutations." *Antimicrob Agents Chemother* **38**(4): 773-780.
275. Tan Siang Y. and K. Erika (2012). "Albert Calmette (1863–1933): Originator of the BCG vaccine.pdf." *Singapore Med J* **53**(7): 433
276. Tasneen, R., S.-Y. Li, et al. (2011). "Sterilizing Activity of Novel TMC207- and PA-824-Containing Regimens in a Murine Model of Tuberculosis." *Antimicrobial Agents and Chemotherapy* **55**(12): 5485-5492.
277. Trajman, A., R. E. Steffen, et al. (2013). "Interferon-Gamma Release Assays versus Tuberculin Skin Testing for the Diagnosis of Latent Tuberculosis Infection: An Overview of the Evidence." *Pulmonary Medicine* **2013**: 1-11.
278. Tsenova, L., E. Ellison, et al. (2005). "Virulence of selected *Mycobacterium tuberculosis* clinical isolates in the rabbit model of meningitis is dependent on phenolic glycolipid produced by the bacilli." *J Infect Dis* **192**(1): 98-106.
279. Tsolaki, A. G., S. Gagneux, et al. (2005). "Genomic deletions classify the Beijing/W strains as a distinct genetic lineage of *Mycobacterium tuberculosis*." *J Clin Microbiol* **43**(7): 3185-3191.
280. Tsolaki, A. G., S. Gagneux, et al. (2005). "Genomic Deletions Classify the Beijing/W Strains as a Distinct Genetic Lineage of *Mycobacterium tuberculosis*." *Journal of Clinical Microbiology* **43**(7): 3185-3191.
281. Valcheva, V., I. Mokrousov, et al. (2008). "Molecular characterization of *Mycobacterium tuberculosis* isolates from different regions of Bulgaria." *J Clin Microbiol* **46**(3): 1014-1018.
282. van Crevel, R., R. H. Nelwan, et al. (2001). "Mycobacterium tuberculosis Beijing genotype strains associated with febrile response to treatment." *Emerg Infect Dis* **7**(5): 880-883.
283. van Soolingen, D., L. Qian, et al. (1995). "Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of east Asia." *J Clin Microbiol* **33**(12): 3234-3238.
284. van Welie, R. (2008). "Slave trading and slavery in the Dutch colonial empire : a global comparison." *New West Indian Guide* **82**(1 & 2): 47-96.
285. vanSoolingen, D., P. W. M. Hermans, et al. (1991). "The occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis." *J. Clin. Microbiol.* **29**: 2578-2586.
286. Vernet, T. (2009). Slave trade and slavery on the Swahili coast (1500-1750). *Slavery, Islam and Diaspora*. B.A. Mirzai, I. M. Montana and P. Lovejoy, Africa World Press: pp.37-76.
287. Veziris, N. (2014). "Détection rapide du bacille tuberculeux et de sa résistance. pdf.pdf>."
288. Via, L. E., S. B. Georghiou, et al. (2012). "Evaluation of Genetic Mutations Associated with *Mycobacterium tuberculosis* Resistance to Amikacin, Kanamycin and Capreomycin: A Systematic Review." *PLoS One* **7**(3): e33275.
289. Viegas, S. O., S. Ghebremichael, et al. (2015). "Mycobacterium tuberculosis causing tuberculous lymphadenitis in Maputo, Mozambique." *BMC Microbiol* **15**: 268.
290. Viegas, S. O., A. Machado, et al. (2013). "Mycobacterium tuberculosis Beijing genotype is associated with HIV infection in Mozambique." *PLoS One* **8**(8): e71999.
291. Viegas, S. O., A. Machado, et al. (2010). "Molecular diversity of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in Mozambique." *BMC Microbiol* **10**: 195.

292. Vyazovaya, A., I. Mokrousov, et al. (2015). "Tuberculous Spondylitis in Russia and Prominent Role of Multidrug-Resistant Clone Mycobacterium tuberculosis Beijing B0/W148." Antimicrobial Agents and Chemotherapy **59**(4): 2349-2357.
293. Vynnycky Emilia and F. P. E. M. (2000). "Lifetime risks, incubation period, and serial interval of tuberculosis.pdf>." American Journal of Epidemiology **Vol. 152, No. 3**: 247–263.
294. Wada, T., T. Iwamoto, et al. (2009). "Genetic diversity of the Mycobacterium tuberculosis Beijing family in East Asia revealed through refined population structure analysis." FEMS Microbiology Letters **291**(1): 35-43.
295. Wan, K. (2007). Study on Genetic Diversity and evolution of Mycobacterium tuberculosis the agent of tuberculosis in China. Paris, University Paris-Sud.
296. Wan, K., J. Liu, et al. (2011). "Investigation on Mycobacterium tuberculosis Diversity in China and the Origin of the Beijing Clade." PLoS One **6**(12): e29190.
297. Wanger, A. and K. Mill (1996). "Testing of M.tuberculosis susceptibility to Ethambutol, Isoniazid, Rifampin and Streptomycin by using Etest.pdf>." JOURNAL OF CLINICAL MICROBIOLOGY: 1672-1676.
298. Warren, R. M., E. M. Streicher, et al. (2002). "Microevolution of the direct repeat region of Mycobacterium tuberculosis: implications for interpretation of spoligotyping data." J Clin Microbiol **40**(12): 4457-4465.
299. Weniger, T., J. Krawczyk, et al. (2010). "MIRU-VNTRplus: a web tool for polyphasic genotyping of Mycobacterium tuberculosis complex bacteria." Nucleic Acids Res **38**(Web Server issue): W326-331.
300. WHO (2015). Global Tuberculosis Report 2015, WHO, Geneva, Switzerland.
301. WHO (2015). World Health Statistics 2015, WHO Geneva: 1-164.
302. WHO (2016). Global Tuberculosis Report 2016, WHO, Geneva, Switzerland.
303. WHO (2016). "WHO, Treatment guidelines for drug-resistant tuberculosis 2016.pdf>."
304. WHO (2017). Global tuberculosis report.
305. WHO (2017). Global Tuberculosis Report 2017, WHO, Geneva, Switzerland.
306. WHO (2017). "WHO- Xpert MTB_Rif Ultra compared to Xpert MTB_Rif 2017.pdf>." World Health Organization 2017.
307. WHO (2017). "WHO_GlobalTB-report-2017.pdf>."
308. Wilmar L. Salo, Arthur C. Aufderheide, et al. (1994). "Identification of MTB DNA in a pre-Columbian Peruvian mummy.pdf>." Proc. Natl. Acad. Sci. USA **91**: 2091-2094.
309. Wouter, A. d. O., C. Robert, et al. (2010). "Diagnostic et traitement de la tuberculose, Manuel pratique.pdf>." FARES.
310. Wuyts, V., W. Mattheus, et al. (2015). "A multiplex oligonucleotide ligation-PCR as a complementary tool for subtyping of Salmonella Typhimurium." Appl Microbiol Biotechnol **99**(19): 8137-8149.
311. Xia, E., Y. Y. Teo, et al. (2016). "SpoTyping: fast and accurate in silico Mycobacterium spoligotyping from sequence reads." Genome Med **8**(1): 19.
312. Yan Feng, Sijun Liu, et al. (2013). "Rapid Diagnosis of Drug Resistance to Fluoroquinolones, Amikacin, Capreomycin, Kanamycin and Ethambutol Using Genotype MTBDRsl Assay: A Meta-Analysis." PLoS One.
313. Yang, C., T. Luo, et al. (2012). "Mycobacterium tuberculosis Beijing Strains Favor Transmission but Not Drug Resistance in China." Clinical Infectious Diseases **55**(9): 1179-1187.
314. Yasmin, M., M. K. Gomgnimbou, et al. (2014). "Multi-drug resistant Mycobacterium tuberculosis complex genetic diversity and clues on recent transmission in Punjab, Pakistan." Infect Genet Evol **27C**: 6-14.
315. Yasmin, M., G. Refregier, et al. (2018). "Reverse line probe assay for cheap detection of Single Nucleotide Polymorphisms in Mycobacterium tuberculosis." Tuberculosis (Edinb) **110**: 52-55.

316. Yin, F., J. F. Chan, et al. (2016). "Development and in-use evaluation of a novel Luminex MicroPlex microsphere-based (TRIO) assay for simultaneous identification of Mycobacterium tuberculosis and detection of first-line and second-line anti-tuberculous drug resistance in China." J Clin Pathol.
317. Yokoyama, E., Y. Hachisu, et al. (2010). "Concordance of variable-number tandem repeat (VNTR) and large sequence polymorphism (LSP) analyses of Mycobacterium tuberculosis strains." Infection, Genetics and Evolution **10**(7): 913-918.
318. Zager, E. M. and R. McNerney (2008). "Multidrug-resistant tuberculosis." BMC Infect Dis **8**: 10.
319. Zammarchi, L., E. Tortoli, et al. (2014). "High prevalence of clustered tuberculosis cases in peruvian migrants in florence, Italy." Infect Dis Rep **6**(4): 5646.
320. Zaunbrecher, M. A., R. D. Sikes, Jr., et al. (2009). "Overexpression of the chromosomally encoded aminoglycoside acetyltransferase eis confers kanamycin resistance in Mycobacterium tuberculosis." Proc Natl Acad Sci U S A **106**(47): 20004-20009.
321. Zhang, J., E. Abadia, et al. (2010). "Mycobacterium tuberculosis complex CRISPR genotyping: improving efficiency, throughput and discriminative power of 'spoligotyping' with new spacers and a microbead-based hybridization assay." J Med Microbiol **59**(Pt 3): 285-294.
322. Zhang, S., J. Chen, et al. (2015). "Identification of novel mutations associated with clofazimine resistance in Mycobacterium tuberculosis." J Antimicrob Chemother **70**(9): 2507-2510.
323. Zhang, Y., D. Mitchison, et al. (2014). "Mechanisms of Pyrazinamide Action and Resistance." Microbiology Spectrum **2**(4).
324. Zhang, Y. and W. W. Yew (2015). "Mechanisms of drug resistance in Mycobacterium tuberculosis: update 2015." The International Journal of Tuberculosis and Lung Disease **19**(11): 1276-1289.
325. Zhao, L. L., Q. Sun, et al. (2015). "Analysis of embCAB mutations associated with ethambutol resistance in multidrug-resistant mycobacterium tuberculosis isolates from China." Antimicrob Agents Chemother **59**(4): 2045-2050.
326. Zhu, C., Y. Zhang, et al. (2012). "Molecular characterization of fluoroquinolone-resistant Mycobacterium tuberculosis clinical isolates from Shanghai, China." Diagn Microbiol Infect Dis **73**(3): 260-263.
327. Zignol, M., A. S. Dean, et al. (2016). "Twenty Years of Global Surveillance of Antituberculosis-Drug Resistance." N Engl J Med **375**(11): 1081-1089.
328. Zink, A., C. J. Haars, et al. (2001). "Molecular analysis of skeletal tuberculosis in an ancient Egyptian population pdf.pdf." J. Med. Microbiol **V-50**: 355-366.

Titre : Développement de méthodes pour le diagnostic, le contrôle, la surveillance de la tuberculose à Bacilles Ultra-résistantes et des souches épidémiques Beijing.

Mots clés : *Mycobacterium tuberculosis*, Souche Beijing, MDR-XDR TB, Épidémiologie moléculaire, phylogénie.

Résumé : La tuberculose MDR/XDR (multi et ultrarésistante aux antituberculeux) causée par *Mycobacterium tuberculosis* constitue un problème de santé publique mondial. L'étude et l'identification des mutations responsables de la résistance sont des facteurs clés pour le contrôle et la surveillance de la tuberculose MDR/XDR. L'expansion de lignée L2/Beijing, une famille de souches originaire du Sud-Est de la Chine (Guangxi) potentiellement plus virulente, complique la maîtrise de cette maladie. Dans ce contexte, nous avons développé le TB-EFI et le TB IS-NTF/RINT, deux méthodes moléculaires rapides, multiplexées et haut débit (développées sur le système Luminex xMap), prêtes à utilisation. Nous avons initié le développement d'une méthode moléculaire par la sélection de marqueurs moléculaires pertinents en vue de la discrimination des souches Beijing par la technique MLPA-Beijing.

Le TB-EFI est un test qui permet d'identifier les mutations fréquentes (polymorphismes de nucléotides simples) dans les gènes associés à la résistance des souches de *Mycobacterium tuberculosis* aux antituberculeux de deuxième ligne dont la Fluoroquinolone, les Injectables, et à l'antituberculeux de première ligne, l'Éthambutol.

Le TB-EFI pourrait être un test utilisable dans les études rétrospectives en vue du suivi de la résistance d'une population.

Le test IS-NTF/RINT est un test spécifique aux souches Beijing qui type la séquence d'insertion IS6110 au sein du locus NTF (Ancien/moderne) et détecte les mutations responsables de la résistance de ces souches à la Rifampicine et l'Isoniazide (les deux antibiotiques principaux de première ligne). Ce test est d'une importance capitale pour l'identification et le contrôle des souches épidémiques, mais aussi pour une vision sur l'évolution du phénomène de résistance dans le temps et l'espace. Il est peu discriminant pour la différenciation des souches Beijing. En vue d'une discrimination complète et précise des souches Beijing, nous avons proposé un lot de SNP qui serviront pour la technique MLPA-Beijing.

Par ailleurs, ces méthodes ainsi que le spoligotyping sur microbille, nous ont permis d'effectuer des études d'épidémiologie moléculaire de la tuberculose au Kazakhstan, en Nouvelle Guinée Papouasie, en Italie, au Mozambique, au Pérou.

Les techniques développées dans cette thèse pourraient contribuer de manière significative au contrôle de la tuberculose XDR dans les zones «hot-spot», et à la surveillance mondiale de l'évolution des souches Beijing spécialement des souches MDR épidémiques.

Title : Development of methods for the diagnostic, the control and monitoring of tuberculosis with Bacilles extensively resistant and epidemic Beijing strain.

Keywords : *Mycobacterium tuberculosis*, Beijing Strain, MDR-XDR TB, molecular epidemiology, phylogeny.

Abstract: MDR / XDR (multidrug and extensively resistant to tuberculosis) TB caused by *Mycobacterium tuberculosis* is still a global public health problem. The study and identification of mutations responsible for resistance are important factors for the control and surveillance of MDR / XDR TB. The expansion of the L2 / Beijing lineage, a family of strains originating from South-East of China (Guangxi) and potentially more virulent, complicates the control of this disease. In this context, we have developed TB-EFI and TB IS-NTF / RINT, two high-speed, multiplexed and high-throughput molecular methods ready to use (developed on the Luminex xMap system). We initiated the development of a molecular method by the selection of relevant molecular markers for the discrimination of Beijing strains by the MLPA technique. TB-EFI is a test that identifies frequent mutations (single nucleotide polymorphisms) in the genes associated with the resistance of *Mycobacterium tuberculosis* strains to second-line anti-TB drugs including Fluoroquinolone, Injectable and first-line antituberculosis drug, Ethambutol. TB-EFI may be used in retrospective studies to monitor resistance in a population.

The IS-NTF / RINT test is a test specific to Beijing strains that types the IS6110 insertion sequence within the NTF locus (Ancient / Modern) and detects the mutations responsible for the resistance of these strains to Rifampicin and Isoniazid (the two leading primary antibiotics). This test is of paramount importance for the identification and control of epidemic strains, but also for a vision on the evolution of the phenomenon of resistance in time and space. It is not very discriminating among Beijing strains. In view of complete and precise discrimination of the Beijing strains, we have proposed a set of SNPs that will be used for a technique that will be called MLPA-Beijing.

In addition, these methods as well as spoligotyping on microbeads allowed us to carry out molecular epidemiological studies of tuberculosis in Kazakhstan, Papua New Guinea, Italy, Mozambique and Peru.

The techniques developed in this thesis could contribute significantly to the control of XDR tuberculosis in hot-spot areas, and to the global monitoring of the evolution of Beijing strains especially epidemic MDR strains.



